

**MODELO PREDITIVO DA CRIMINALIDADE
GEOREFERENCIAÇÃO AO CONCELHO DE LISBOA**

Por

PAULO ABEL DE ALMEIDA JOÃO

Dissertação apresentada como requisito parcial para a obtenção do grau de:

MESTRE EM ESTATÍSTICA E GESTÃO DE INFORMAÇÃO

Pelo

INSTITUTO SUPERIOR DE ESTATÍSTICA E GESTÃO DE INFORMAÇÃO

Da

UNIVERSIDADE NOVA DE LISBOA

2009

ORIENTADOR

PROFESSOR DOUTOR VICTOR JOSÉ DE ALMEIDA E SOUSA LOBO

Professor Associado Convidado com Agregação

Doutor em Engenharia Informática (FCT/UNL)

COORIENTADOR

PROFESSOR DOUTOR FERNANDO JOSÉ FERREIRA LUCAS BAÇÃO

Professor Associado

Doutor em Gestão de Informação (ISEGI/UNL)

*Aos meus filhos:
Ana, Nelson e Susana, como exemplo
Ao meu Pai:
Pelo que me ensinou
Á minha Mãe, in memoriam:
Pela falta que me fez.*

Qualquer resultado em investigação é relativo a uma problemática, ao esquema teórico no qual se baseia directa ou indirectamente, e à metodologia através da qual foi obtido.

(Sierpiska, Kilpatrick & Balacheff, 1993)

PREFÁCIO

Quando o Sistema Estratégico de Informação da PSP (SEI) foi finalmente lançado em 2004 surgiram no seio da Instituição dois sentimentos antagónicos mas naturais: um de expectativa na qualidade e modernidade que o Sistema pressupunha e outro de resistência natural à mudança.

Passados quatro anos de “vigência” do SEI continuam a observar-se os mesmos dois sentimentos.

O SEI foi, indubitavelmente, o maior passo, até ao momento, dado na PSP na aposta ao recurso às novas tecnologias de informação.

Sendo, na prática, uma grande Base de Dados onde são registadas determinadas variáveis que caracterizam uma ocorrência de índole policial, o SEI permite o seu cruzamento em tempo real, de forma a obter-se informação e conhecimento para a caracterização de um ou mais fenómenos criminais.

Mas esta análise – importante mas não suficiente – apenas permite adquirir valor acrescentado na referência e definição do perfil dos crimes inseridos na base de dados, não permitindo, contudo, que com base nessa mesma análise se prevejam futuras ocorrências.

Tal limitação condiciona a tomada de decisões de natureza táctica e operacional, dificultando uma correcta gestão de meios e impossibilitando, de forma eficaz, a prevenção criminal, missão nuclear da Polícia de Segurança Pública.

Quando o Dr. Paulo João me abordou sobre o tema do seu trabalho e a metodologia que iria seguir para o desenvolver, apoiei-o de imediato.

Os modelos preditivos da criminalidade são, nos dias de hoje, os ingredientes que distinguem as polícias do século XXI das do século passado.

Aplicá-los à PSP, tendo por base a análise que o SEI não responde, é não só um projecto ambicioso mas, seguramente, um trabalho da mais prestimosa pertinência para o caminho de modernidade, ainda não alcançado, na PSP.

Prever a ocorrência criminal, a sua evolução, estagnação ou redução, com base em critérios científicos, leva-nos a acreditar numa Polícia eficaz, capaz e credível. Numa Polícia orientada pela Intelligence. Numa Polícia que responde com qualidade e eficiência ao sentimento de insegurança, reduzindo-o e criando laços de confiança entre a comunidade civil e os seus profissionais.

Este trabalho é o ponto de partida. Parabéns ao seu autor.

Subintendente *Alexandre Coimbra*
Chefe da Divisão de Análise de Informações
Policiais da Direcção Nacional da PSP
Lisboa, 23 de Abril de 2008

AGRADECIMENTOS

Ao ISEGI e a todos os seus docentes e colaboradores que directa ou indirectamente contribuíram para este trabalho;

À Professora Doutora Ana Maria Ramalho Correia, pelos preciosos esclarecimentos da metodologia científica aplicável;

Aos Orientadores deste trabalho, pela preciosa ajuda e disponibilidade sem o qual não teria sido possível;

Ao gabinete de informática da PSP pela cedência dos dados das ocorrências criminais em Lisboa no ano de 2007, que serviu de base ao modelo preditivo e sem o qual este não seria fidedigno da realidade;

Ao Subintendente Alexandre Coimbra pela colaboração na elaboração do prefácio, disponibilidade e apoio neste projecto;

Ao Dr. Abel João Rosa, pelo constante apoio e revisão final do texto.

A todos, muito obrigado.

RESUMO

Pretende-se elaborar um modelo preditivo ou processo analítico e sistemático de descoberta do conhecimento, orientado segundo os princípios da pertinência e da oportunidade, que detecte os *hot spots* da criminalidade, que faça uma previsão e propensão de ocorrência e ainda, que faça uma previsão da sua evolução, estagnação ou redução, sendo realizado a partir do estabelecimento de correlações entre conjuntos de dados criminais ocorridos no primeiro semestre do ano de 2007 no concelho de Lisboa. Este modelo poderá posteriormente ser aplicado a outras regiões do país.

ABSTRACT

The intends is to prepare a predictive model or analytical and systematic process of knowledge discovery, oriented according to the principles of the relevance and opportunity, which detects the criminality hot spots and does a forecasting of incident and his evolution, stagnation or reduction, being carried out from the establishment of correlations between sets of criminal data occurred in the first semester of year 2007 in the district of Lisbon. This model can be applied to other regions of the country.

PALAVRAS-CHAVE

Análise Cluster

Análise criminal

Análise de risco

Data Mining

Geoestatística

Hot-Spot

Intelligence led policing

Mapeamento criminal

Modelo preditivo

Policiamento orientado pelo tipo de problema

Movimentação baseada no risco

Sistema de Informação Geográfica

KEYWORDS

Cluster analysis

Crime analysis

Crime mapping

Data mining

Geographic Information Systems

Geostatistics

Hot-Spot

Intelligence led policing

Predictive model

Problem Oriented Policing

Risk and thread assessment

Risk based deployment

ABREVIATURAS E ACRÓNIMOS

ACPO	<i>Association of Chief Police Officer</i>
BI	<i>Business Intelligence</i>
BMCP	<i>Bicriterion Median Clustering Problem</i>
C4.5	Algoritmo de treino de árvores de decisão
CAOP	Carta Administrativa Oficial de Portugal
CART	<i>Classification and Regression Trees</i>
CAU	<i>Crime Analysis Unit</i>
CEJ	Centro de Estudos Judiciários
CEMET	<i>Cells Monitora Espacio-Temporal</i>
CEWS	<i>Crime Early Warning System</i>
CGT	<i>Criminal Geographic Targeting</i>
CHAID	<i>Chi Square Automatic Interaction Detection</i>
CLAP	<i>Cluster Analysis Prediction</i>
CLP	<i>Crime Led Policing</i>
CLUSTER	Classificação de indivíduos ou variáveis com características
CMAP	<i>Crime Mapping and Analysis Program</i>
CNIG	Centro Nacional de Informação Geográfica
CNPD	Comissão Nacional de Protecção de Dados
COP	<i>Community Oriented Policing</i>
COPDAT	<i>Crime and Offender Pattern Detection Analysis Technique</i>
COMPSTAT	<i>Computerized Comparison Crime Statistics</i>
CPAS	<i>Crime Pattern Analysis System</i>
CPTED	<i>Crime Prevention Through Environmental Design</i>
CRISP-DM	<i>Cross Industry Standard Process for Data Mining</i>
CRMC	<i>Crime Mapping Research Center</i>
DDC	<i>Distance Decay Curve</i>
DM	<i>Data Mining</i>
ECIM	<i>European Criminal Intelligence Model</i>
EDA	<i>Exploratory Data Analysis</i>
EMPACT	<i>Electronic Monitoring Protection and Crime Tracking</i>
ESDA	<i>Exploratory Spatial Data Analysis</i>
ESRI	<i>Environmental Systems Research Institute</i>
GNR	Guarda Nacional Republicana
GUI	<i>Graphic User Interface</i>
HMIC	<i>Her Majesty's Inspectorate of Constabulary</i>

HOT-SPOT	Ponto-quente, local o área de concentração da criminalidade
IACA	<i>International Association of Crime Analysts</i>
IBDDC	<i>Incident Based Distance Decay Curve</i>
IBR	<i>Incident-Based Reports</i>
ILP	<i>Intelligence Led Policing</i>
INE	Instituto Nacional de Estatística
ISCPSI	Instituto Superior de Ciências Policiais e Segurança Interna
J2C	<i>Journey To Crime</i>
KDD	<i>Knowledge Discovery in Databases</i>
LAC	<i>Local Area Command</i>
LCA	<i>Local Crime Analysis</i>
LEO	<i>Law Enforcement Online</i>
LISA	<i>Local Indicator of Spatial Association</i>
MAUP	<i>Modifiable Areal Unit Problem</i>
NARX	<i>Nonlinear Autoregressive Exogenous Model</i>
NCIS	<i>National Criminal Intelligence Service</i>
NIBRS	<i>National Incident-Based Reporting System</i>
NSCR	<i>Netherlands Institute for the Study of Crime and Law Enforcement</i>
NNI	<i>Nearest Neighbor Index</i>
NORA	<i>Nonobvious Relationship Analysis</i>
NP405	Norma portuguesa para referências bibliográficas e citações
OSCOT	Observatório de Segurança, Criminalidade Organizada e Terrorismo
PESI	Plano Estratégico de Sistemas de Informação
PIPP	Programa Integrado de Policiamento de Proximidade
POP	<i>Problem Oriented Policing</i>
PPAC	<i>Police Patrol Area Covering</i>
PSP	Polícia de Segurança Pública
RASI	Relatório Anual de Segurança Interna
RCMP	<i>Royal Canadian Mounted Police</i>
SARA	<i>Scanning, Analysis, Response and Assessment</i>
SEI	Sistema Estratégico de Informação, Gestão e Controlo Operacional
SEMMA	<i>Sample, Explore, Modify, Model, Assess</i>
SIG	Sistema de Informação Geográfica
SIIG	Sistema Integrado de Informação Criminal
SOM	<i>Self Organized Map</i>
SPSS	<i>Statistical Package for Social Science</i>
SQL	<i>Structured Query Language</i>
STAC	<i>Spatial and Temporal Analysis of Crime</i>
UCR	<i>Uniform Crime Reports</i>
URL	<i>Universal Resource Locator</i>
ViCLAS	<i>Violent Crime Linkage Analysis System</i>
WWW	<i>World Wide Web</i>

GLOSSÁRIO

- ALGORITMOS GENÉTICOS** Técnica de optimização da aprendizagem dos modelos inspirada na teoria Darwinista sobre a evolução das espécies; esta técnica procura a solução do problema através de um processo de evolução da população (e.g. geração aleatória).
- AGENT-BASED MODELLING** Tipo de modelo de simulação, onde os elementos importantes são representados por agentes, um agente pode representar uma pessoa, organização ou vizinhança; Cada agente tem determinadas características e comportamentos próprios.
- ANÁLISE CLUSTER** É um método de classificação por agrupamento de indivíduos com características homogéneas com base num conjunto de variáveis; permite abordagens de análise com métodos hierárquicos (e. g. dendograma) e métodos não hierárquicos (e. g. *K-means*), os métodos hierárquicos, subdividem-se ainda em aglomerativos ou divisivos. Basicamente, fazer *clustering* é agrupar populações heterogéneas em subgrupos homogéneos tendo como únicos critérios a semelhança entre indivíduos.
- ÁRVORES DE DECISÃO** Ferramenta de classificação e previsão organizada em estruturas que representam regras de decisão para classificar os dados de input; dos algoritmos mais conhecidos, destacam-se o CART, C4.5 e o CHAID.
- CEMET** Técnica de monitorização espaço-temporal de células para a localização de padrões criminais, é sensível á qualidade dos dados disponíveis e permite explorar mudanças nos padrões criminais, usa um método probabilístico para as referências geográficas de cada ocorrência de modo a detectar a probabilidade de nova ocorrência dentro da área geográfica coberta.
- DATA MINING** Metodologia de descoberta de conhecimento em grandes bases de dados; baseia-se em métodos não paramétricos para extracção do conhecimento. Procura entre as variáveis encontrar padrões ocultos. As principais tarefas da *data mining* são: classificação, estimação, predição e *clustering*.

HEAT-MAP	Matriz que associa o nível e tipo de criminalidade de acordo com o turno de serviço das forças de segurança, tipicamente, quando mais escuro estiver a quadrícula, maior o número de ocorrências nesse turno de serviço.
HOT-SPOT	Metodologia típica de representação dos índices de criminalidade sob a forma de elipses. Nos Sistemas de Informação Geográfica, representa uma área onde os índices de criminalidade estão acima do que é normal para esse local (pode ser associado a outras variáveis para além da criminalidade), ou seja, é uma área com uma concentração de ocorrências de crimes superior à média.
LIFT	É o rácio que mede a mudança na concentração de uma classe específica quando um modelo é utilizado para seleccionar uma amostra propositadamente enviesada a partir da população.
MATRIZ CONFUSÃO	É uma matriz que avalia a qualidade do modelo preditivo indicando os resultados: falsos-positivos, falsos-negativos, verdadeiros-positivos e os verdadeiros-negativos.
MODELO DETERMINÍSTICO	É baseado em conhecimento, eminentemente preditivo, consistem em formular equações “fechadas” que definem deterministicamente a forma como os <i>outputs</i> são obtidos a partir dos <i>inputs</i> , sendo todas as restantes características constantes.
MODELO NÃO PARAMÉTRICO	É baseado nos dados, não faz uso de hipóteses estatísticas <i>a priori</i> , limita-se a encontrar as relações entre as variáveis de forma generalizada. Recorre normalmente a grandes bases de dados como fonte de informação secundária. Este modelo pode ter ou não pré-processamento dos dados.
MODELO PARAMÉTRICO	É baseado em pressupostos, faz uso de conhecimento <i>a priori</i> de hipóteses e preconceitos estatísticos como as distribuições e propriedades clássicas, recorre normalmente a fontes primárias de dados, recolhidos especialmente para o efeito pretendido, as equações matemáticas definem claramente os inputs e os outputs do modelo (e. g. regressão).
MODELO PREDITIVO	Modelo baseado em regressão ou classificação, capaz de prever o nível de probabilidade do valor de determinada variável com base em dados passados.

NARX Modelo exógeno não linear e autoregressivo, (*nonlinear autoregressive exogenous model*) é utilizado na modelação de séries temporais. O modelo apresenta o valor actual das series temporais, para valores do passado da mesma série de dados; No entanto apresenta sempre um erro residual que nos inibe de prever o valor actual exacto da série temporal. Algebricamente o modelo representa-se da seguinte forma:

$$y_t = F(y_{t-1}, y_{t-2}, y_{t-3}, \dots, u_{t-1}, u_{t-2}, u_{t-3}, \dots) + \varepsilon_t$$

Onde y é a variável de interesse e u é uma variável correlacionada com y , ou seja, o facto de termos informação sobre u ajuda-nos a prever o valor de y . F é uma função não linear polinomial, em algumas aplicações pode ser uma rede neuronal.

NNI Índice do vizinho mais próximo (*nearest neighbor index*), é o rácio das médias das distâncias ao vizinho mais próximo pela média aleatória das distâncias ao vizinho mais próximo.

NORMALIZAÇÃO Normalizar significa pré-processar os dados de *input* de modo a que todos tenham o mesmo peso e medida, evitando que dados com valores muito elevados ou muito baixos produzam resultados com distribuições enviesadas degradando o modelo; Dos diversos processos de normalização dos dados podemos referir: médias, *Z-score*, desvio-padrão, *Min-Max* e sigmoideal. A normalização é um passo no pré-processamento dos dados de elevada importância, porque os modelos não-paramétricos, assumem, implicitamente que as distâncias nas diferentes direcções do espaço de input, possuem o mesmo peso ou importância.

OVERFITTING Tendência dos algoritmos de aprendizagem do conjunto de treino dos dados em captar as particularidades e de não ser capaz de os generalizar, ou seja, o modelo “aprende” tão bem os dados que depois não os pode extrapolar para a população por serem demasiados específicos. (e. g. quando o modelo capta o ruído como se fosse um dado do conjunto de treino).

REDES NEURONAIS Modelos preditivos não lineares inspirados nas redes neuronais biológicas; possuem um processo de aprendizagem supervisionada ou não supervisionada. Estas redes são baseadas em pesos sinápticos atribuídos aos dados de *input* e retropropagação do erro como meio de optimização.

- REGRESSÃO** Modelo estatístico para funções lineares que agrega indivíduos de acordo com a sua tendência ou recta da regressão, pode ser simples ou múltipla conforme o número de variáveis exógenas.
- No modelo de regressão, o comportamento da variável endógena depende linearmente do valor das variáveis exógenas que a explicam.
- SOM** Para Bação (2006) um Mapa Auto-Organizado (*Self Organized Map*), é uma rede neuronal não supervisionada (sem variável *target*); pode ser observado como sendo um conjunto de neurónios de *input* e *output* organizados em forma de grelha (mapa topológico); tem a capacidade para adaptar o conjunto de dados a uma camada de neurónios de *output*; nesta adaptação os neurónios vizinhos são “puxados” ou “empurrados” de acordo com a proporção de distância definida pela taxa de aprendizagem do modelo utilizado; esta taxa de aprendizagem é a magnitude que separa o neurónio do indivíduo.
- TÁBUAS DE PROGNOSE** São tabelas de probabilidade estatística que se baseiam em factores sobre o comportamento futuro, penalmente relevante, de um indivíduo, podem destinar-se a predizer a reincidência, a adaptação às condições do regime de liberdade condicional, os resultados do tratamento institucional ou até a antecipar a potencial delinquência futura de crianças. De acordo com Dias, G. (1984:149), do ponto de vista de política criminal, acentua-se a possibilidade de a tábua de prognose ter um efeito amplificador ou reproduzidor da delinquência, actuando como *self-fulfilling prophecy*, a apressada ou errada classificação de um indivíduo como pertencente a uma categoria com elevado grau de probabilidade de reincidência, pode acabar por determinar a reincidência efectiva.
- TESTE GAMMA** O teste *Gamma*, é uma técnica estatística para estimação do ruído e a variância do erro em dados amostrais, é análogo à soma do erro quadrático em regressões lineares, no entanto aplica-se a qualquer função *smooth* (uma função *smooth*, é a que tem derivadas de todas as ordens de variação contínua). Basicamente, estima o erro quadrado médio que pode ser conseguido quando modelamos os dados usando um modelo contínuo de ajustamento (e.g. regressão por mínimos quadrados ou redes neuronais para uma função desconhecida).

TABELA DE CONTEÚDO

Introdução	1
1. Questão de Investigação, objectivos e motivações	7
1.1 Distinção entre dados, informação e conhecimento.....	8
1.2 Objectivos gerais.....	11
1.3 Interesse pessoal.....	13
1.4 Interesse profissional.....	14
1.4.1 Policiamento comunitário ou de proximidade.....	15
1.4.2 Policiamento orientado para os problemas.....	19
1.4.3 Policiamento orientado pelas informações.....	25
2. Revisão critica da literatura	28
2.1 Metodologias de mapeamento da criminalidade.....	30
2.2 Situação em Portugal.....	33
3. Metodologia de Investigação	40
3.1 Abordagem de investigação.....	41
3.2 Recursos e dados disponíveis.....	42
3.2.1 Interpretação e dificuldades das estatísticas oficiais.....	45
3.2.2 Os perigos da interpretação <i>Post Factum</i>	46
3.3 Fontes de informação e pesquisas.....	46
4. Processamento dos dados disponíveis	48
4.2 Pré-processamento dos dados.....	49
4.3 Descrição das variáveis.....	51
4.4 Estatística descritiva dos dados.....	53

5. Modelo preditivo	58
5.1 Selecção e desenvolvimento do modelo teórico	59
5.2 Modelo <i>enterprise miner</i> do SAS.....	63
5.2.1 Árvore de decisão	66
5.2.2 Descrição da rede neuronal 1	67
5.2.3 Descrição da rede neuronal 2	68
5.2.4 Descrição da regressão.....	69
5.3 Análise dos Resultados	70
6. Análise de dados com SIG	74
6.1 Análise espacial da criminalidade	81
6.2 Análise temporal da criminalidade.....	83
6.3 Modelo SIG para representação dos dados	84
6.4 Mapas Temáticos de Lisboa.....	86
7. Trabalhos futuros	92
7.1 Projecto SIGpol	92
7.2 Esquadra virtual	93
7.3 Cooperação policial ao nível das tecnologias da informação	94
7.4 A GEO-Polícia	96
Conclusões	98
Referências Bibliográficas	104
Apêndices	114
Apêndice 1 - Cronograma de actividades	114
Apêndice 2 – Mapas temáticos	115
Apêndice 3 – Tabela com os dados agregados	125
Apêndice 4 – Histogramas das variáveis	125
Apêndice 5 – <i>Box-plot</i> das variáveis	132
Apêndice 6 – Matriz de correlação de <i>Pearson</i>	139

Anexos	139
Anexo 1 - Modelo de dados CRISP do SPSS	140
Anexo 2 - Abordagens de investigação	140
Anexo 3 - Metodologias de <i>crime mapping</i>	141
Anexo 4 - Comparação entre os modelos de <i>data mining</i>	142
Anexo 5 - Metodologia SEMMA do <i>enterprise miner</i>	142
Anexo 6 – Pequena história das freguesias de Lisboa	143
Anexo 7 – Legislação e documentos de referência.....	148

LISTA DE TABELAS

Tabela 1 - Vertentes estratégicas da actuação policial	15
Tabela 2 - Procedimentos de policiamento.....	18
Tabela 3 - Características do modelo de policiamento.....	24
Tabela 4 - Exemplo dos dados originais.....	49
Tabela 5 - Exemplo dos dados nominais finais	50
Tabela 6 - Descrição das variáveis.....	51
Tabela 7 - Codificação da variável tipo.....	51
Tabela 8 - Codificação da variável subtipo.....	51
Tabela 9 - Codificação da variável local	52
Tabela 10 - Codificação da variável freguesia.....	52
Tabela 11 - Exemplo dos dados numéricos finais	53
Tabela 12 - Descrição do <i>role</i> das variáveis.....	65
Tabela 13 - Estatística descritiva das variáveis.....	65
Tabela 14 - Erro do treino e validação da árvore de decisão.....	66
Tabela 15 - Erro quadrático médio da árvore de decisão.....	66
Tabela 16 - Estatística da rede neuronal 1	67
Tabela 17 - Estatística da rede neuronal 2.....	68
Tabela 18 - Estatística da regressão	69
Tabela 19 - Comparação das previsões dos métodos	70
Tabela 20 - Resultados do <i>assessment</i>	71
Tabela 21 - Comparação do valor real com o previsto no <i>score</i>	72

—

LISTA DE FIGURAS

Figura 1 - Processo de conversão dos dados	10
Figura 2 - Processo de tratamento dos dados.....	12
Figura 3 - Avaliação de desempenho das forças policiais.....	22
Figura 4 - Processo <i>ILP</i>	25
Figura 5 - Programas de proximidade	33
Figura 6 - Balança da criminalidade violenta	43
Figura 7 - Evolução da criminalidade violenta denunciada na última década.....	43
Figura 8 - Geografia da criminalidade.....	44
Figura 9 - Criminalidade denunciada	44
Figura 10 - Registos por entidades policiais	45
Figura 11 - Metodologia de pré-processamento dos dados	48
Figura 12 - Ocorrências criminais mensais.....	53
Figura 13 - Análise global da criminalidade	54
Figura 14 - Tipo de ocorrências	54
Figura 15 - População por freguesias.....	55
Figura 16 - Ocorrências por freguesia	55
Figura 17 - Número de ocorrências por cada 1000 habitantes	56
Figura 18 - Local da ocorrência	56
Figura 19 - Crimes de furto / roubo por cada 1000 habitantes	57
Figura 20 - Modelo preditivo genérico	58
Figura 21 - Metodologia SEMMA.....	60
Figura 22 - Modelo CLAP.....	61
Figura 23 - Modelo preditivo no SAS <i>Enterprise Miner</i>	63
Figura 24 - Características do conjunto de <i>input</i>	63
Figura 25 - Descrição da partição dos dados	64
Figura 26 - Anel da árvore de decisão.....	66
Figura 27 - Gráfico de aprendizagem da árvore de decisão	66
Figura 28 - Organograma da árvore de decisão.....	67
Figura 29 - Gráfico de aprendizagem da rede neuronal 1	67
Figura 30 - Gráfico de aprendizagem da rede neuronal 2.....	68
Figura 31 - Gráfico da previsão da regressão	69
Figura 32 - Erro quantificado do modelo.....	73
Figura 33 - Comparação dos resultados obtidos com valor real	73
Figura 34 - Mapa da cólera em Londres.....	75
Figura 35 - Bairros de risco na grande Lisboa.....	79
Figura 36 - Modelo de processamento de pontos	81

Figura 37 - Diagrama do modelo SIG	85
Figura 38 - Mapa das 53 freguesias de Lisboa.....	86
Figura 39 - Mapa da densidade de ocorrências pela área	87
Figura 40 - Mapa da densidade de ocorrências por cada mil habitantes.....	87
Figura 41 - Mapa da densidade do crime de furto por cada mil habitantes	88
Figura 42 - Mapa da densidade do crime de furto e roubo pela área.....	88
Figura 43 - Mapa da população por freguesia	89
Figura 44 - Mapa da densidade populacional pela área.....	89
Figura 45 - Mapa dos valores REAIS para o mês de Julho.....	90
Figura 46 - Mapa dos valores PREVISTOS para o mês de Julho.....	91
Figura 47 - Logotipo da GEO-Polícia	96

INTRODUÇÃO

O crime é um fenómeno que afecta todas as sociedades humanas e para que um delito tenha alguma probabilidade de concretizar-se, devem existir simultaneamente os elementos que o tornam possível, tais como: o criminoso e o alvo que procura, que poderá ser um bem ou uma pessoa, tudo isto numa envolvente de espaço, tempo e oportunidade.

A relatividade do conceito de crime, aplica-se no tempo como no espaço em virtude de que, alguns actos que hoje são considerados crimes em Portugal, podem não ter sido classificados como tal no passado ou podem não ser considerados crimes noutros países (e. g. pequenas quantidades de estupefacientes têm consumo e comercialização legalizada na Holanda enquanto que em Portugal tal é considerado crime com pena de prisão), deverá haver sempre um enquadramento legislativo e penal na definição da criminalidade.

Os motivos que levam alguém a praticar um crime são diversos, mas o conhecimento exacto onde os crimes são praticados pode contribuir para um plano mais eficaz de prevenção.

Para Falletti (1998), a par desta criminalidade, dita pequena e média criminalidade, temos assistido nos últimos anos ao aumento de uma criminalidade organizada de contornos internacionais como por exemplo o terrorismo, o tráfico de estupefacientes, o tráfico de pessoas, o branqueamento de capitais, *etc.*

Para além da criminalidade objectiva, vários estudos têm dado conta da existência de um sentimento de insegurança, que se desenvolve num plano psicológico, produzindo sentimentos de medo pessoal e de preocupações securitárias.

Lourenço (1998), define o sentimento de insegurança como “Conjunto de manifestações de inquietação, de perturbação e de medo, quer individuais quer colectivas, cristalizadas sobre o crime”.

São invocados vários argumentos para explicar o aumento da criminalidade e o aumento do sentimento de insegurança.

Para Giddens (1996:15), as causas destes fenómenos relacionam-se com o processo de globalização que provocou a descaracterização dos sistemas sociais, levando à ruptura das relações sociais dos contextos locais de interacção e à sua reestruturação através de extensões indefinidas de tempo-espço.

Tal fenómeno provocou a dissolução das solidariedades sociais e a ruptura dos mecanismos de confiança social, contribuindo de forma indirecta para o aumento da criminalidade e para o aumento do sentimento de insegurança.

Já para Roché (1998) a questão relaciona-se com a crise económica e com o processo de urbanização da vida em sociedade ou com as incivilidades.¹

O Procurador-geral da República - Pinto Monteiro, na sua directiva de 15 de Janeiro de 2008, refere:

“Os indicadores existentes permitem identificar alguns fenómenos e tendências criminosas que merecem uma particular atenção por serem susceptíveis de contribuir para o aumento do sentimento de insegurança, pelo que a sua repressão eficaz e atempada é essencial para reforçar a confiança dos cidadãos no sistema de justiça e nos valores do estado de direito, por outro lado, a execução da política criminal, não pode alhear-se da importância e da necessidade de um adequado tratamento da pequena criminalidade, quer na perspectiva da prevenção quer na perspectiva da resocialização dos seus agentes. Para que tais objectivos sejam alcançados, importa promover, neste âmbito, a aplicação de medidas de consenso e de sanções não privativas da liberdade, privilegiando a justiça restaurativa e a celeridade dos procedimentos.”

Estudos de grande importância para o desenvolvimento da criminologia e em especial na criação daquilo que hoje se designa por *Crime Mapping*, foram realizados por investigadores de onde se destaca o francês André Guerry (1833).

¹ Consideram-se incivilidades os pequenos actos de desordem e de acções que nem sempre constituem crime, mas que, pela sua repetição e mal estar que provocam, estão na origem de formas de crime mais graves e do agravamento do sentimento de insegurança junto das comunidades.

Guerry (1833), criou um mapa de localização de crimes baseado numa recolha de dados em termos nacionais, considerou os factores sócio-económicos de cada região, o que permitiu a elaboração da primeira teoria positivista da criminologia.²

A análise criminal do Séc. XIX era trabalhosa, especialmente se tivermos em conta que os mapas eram feitos manualmente e com recursos escassos comparados com a tecnologia que temos hoje. O desenvolvimento dos computadores e de novas tecnologias (e. g. SIG – Sistemas de Informação Geográfica e o GPS – *Global Position System*), permitiu uma localização mais fácil das ocorrências criminais e a sua interligação com o meio geográfico envolvente.

Face a estes recentes desenvolvimentos, o *Crime Mapping* voltou a emergir como uma técnica para muitos projectos, planos de prevenção de segurança e como ferramenta de diagnóstico nos processos de decisão e gestão de recursos materiais e humanos das diversas forças e serviços policiais que têm por missão garantir a segurança pública.

As políticas de segurança pública tradicionalmente desenvolvidas em Portugal pecam por falta de diagnósticos, objectivos, cronogramas, projectos e avaliações.

Não existe um planeamento de projectos de segurança, embora haja uma atitude política de mudança face a anos anteriores, veja-se o Relatório Anual de Segurança Interna (2007:38) quando refere:

² in <http://www.revistapersona.com> “A teoria positivista da criminologia surge no século XIX, segundo a qual os actos das pessoas podem ser previsíveis (determinismo). Ao serem determinados esses actos, excluindo dos indivíduos o livre arbítrio, os teóricos positivistas concluíram que a gênese do crime (que é um facto) pode ser localizada a partir do estudo da pessoa que comete o delito. Desta forma, o delito passa a ser, não um acto consciente mas sim um sintoma cujas causas não podem ser combatidas pelas sanções penais. Trata-se de uma enfermidade que atinge a pessoa, deve, por isso, ser tratada tendo como objectivo final à proteção da sociedade. Três teorias estão envolvidas neste contexto positivista: a) Bioantropológica, que, a partir do determinismo, chega ao entendimento de que existem pessoas inclinadas para a prática de crimes segundo características biológicas; b) Psicodinâmica, para cujos adeptos o homem que comete crimes não o faz exclusivamente segundo factores genéticos mas sim por causa de distúrbios e falhas no seu processo de aprendizagem e socialização (relativização do eu), que interferem no seu crescimento, levando-o à prática de crimes; c) Psico-sociológica, para quem as atitudes são determinadas pela predominância de elementos sociais e situacionais sobre a personalidade.”

“Tendo como objectivo um salto qualitativo e eficaz nas políticas de combate a toda a violência de género, pretende-se dinamizar, com as diferentes estruturas, quer do Estado quer da sociedade civil, um trabalho conjunto na consolidação de uma política de prevenção.”

Históricamente, as abordagens e estratégias policiais estão concentradas basicamente na acção e repressão das forças e serviços de segurança, entendida como instância privilegiada, senão exclusiva, dos dispositivos policiais para a manutenção da ordem pública e a redução dos índices de criminalidade.

Todavia, a acção da polícia é fundamentalmente reactiva, tentando oferecer uma resposta aos últimos acontecimentos perturbadores da ordem pública sobretudo aqueles que tenham grande impacto na opinião pública. São os chamados “Mega-Processos” (e. g. o caso Maddie, caso Joana, Apito Dourado, Casa Pia, caso *Freeport*, Face oculta, etc.).

Normalmente, são estas ocorrências de grande divulgação que obrigam a uma resposta policial imediata, por darem origem aos crimes que afectam áreas nobres da cidade e vítimas ou arguidos de notoriedade social.

A formação das forças e serviços de segurança, não tem contribuído para melhorar este quadro de actuação, continuando a privilegiar os casos mais mediáticos como prioritários na investigação, o que indícia uma polícia com grande influência política, o que não abona nada em transparência da actuação policial em prol da redução da criminalidade e sentimento de insegurança junto da população.

As acções repressivas são focalizadas sobretudo em zonas de baixo rendimento, nomeadamente bairros sociais, estas intervenções policiais são inspiradas no paradigma da “guerra contra o crime” tantas vezes divulgada nos meios de comunicação social.

A ênfase está muito mais na repressão do que na prevenção e a repressão policial está voltada, sobretudo, contra os crimes patrimoniais e crimes pessoais.

A investigação costuma ser feita de forma precária, o que tradicionalmente provoca um elevado número de processos pendentes ou arquivados por falta de provas ou por se ter esgotado o prazo legal de instauração do processo.

Apenas nos últimos anos, as forças policiais começaram a mostrar os frutos de uma investigação mais qualificada, investindo mais na investigação criminal e formação dos agentes com programas específicos de policiamento de proximidade.

Pela primeira vez na nossa ordem jurídica, a Lei n.º 51/2007, de 31 de Agosto, veio definir os objectivos, prioridades e orientações de política criminal para o biénio de 2007-2009, em cumprimento da Lei n.º 17/2006, de 23 de Maio, que aprovou a Lei-quadro da Política Criminal.

Aquela Lei, define como objectivos gerais da política criminal: prevenir, reprimir e reduzir a criminalidade, promovendo a defesa de bens jurídicos, a protecção da vítima e a reintegração do agente do crime na sociedade.

Tradicionalmente, as forças e serviços de segurança, funcionam como instrumentos capazes de gerar uma percepção de ordem e segurança junto da população. Para Monet (1993:99):

*“A noção de segurança pública, encontra, hoje, todo um conjunto de objectivos assinalados por textos jurídicos que regulamentam a actividade policial. E integra também os pedidos múltiplos e heterogéneos que os cidadãos solicitam á polícia.”*³

Rigorosamente, o que está em questão hoje, não é o formato institucional, mas sim as falsas premissas sobre as quais se constrói todo o sentimento de insegurança na população.

A mais problemática delas é a que supõe que a política de segurança está voltada preferencialmente para o combate ao crime, sendo por isso, dado destaque à figura do criminoso.

A base de uma redefinição da abordagem para a questão da segurança está situada no reconhecimento de que este é um campo que diz respeito fundamentalmente às condições necessárias para o incremento da qualidade de vida de toda a população e ao sentimento de segurança ou insegurança percebido pelas pessoas.

Roché (1994), descreve esta noção vaga de insegurança, a partir do sentimento de insegurança, procurando ver de que modo os actores sociais vivem e sentem a realidade.

³ Monet (1993) *“La notion de sécurité publique recouvre, aujourd’hui, tout un ensemble d’objectifs assines par les textes juridiques qui réglementent l’activité policière. Mais elle integre aussi las demandes multiples et hétéroclites que le citoyen adresse à la police.”*

Roché (1994:25), dividiu em dois vectores o conceito de sentimento de insegurança:

- O **medo**, que se vai operacionalizar a partir de variáveis como o medo no domicílio, o medo nas ruas, as medidas de protecção no domicílio;
- A **preocupação** com a ordem, que se vai operacionalizar a partir de variáveis como a autoridade na família, maior repressão dos poderes políticos, estigmatização dos delinquentes, entre outros.

Deve-se por isso alterar a forma de encarar o problema da criminalidade passando de uma acção repressiva para uma atitude preventiva. A investigação aqui descrita é um primeiro passo inovador nesse sentido.

Esta análise da criminalidade. tem por objectivo final, apoiar as decisões estratégicas dos órgãos de polícia criminal. Nomeadamente nas áreas operacional e tática, otimizando assim, os recursos humanos e materiais das forças de segurança e simultaneamente prevenindo e reduzindo a criminalidade. Tem ainda por objectivo, utilizar os sistemas de informação geográfica para visualizar os dados de forma individualizada ou agrupada (*clusters*), associando estes dados com outros de índole demográfica e sócio-económicos, que de algum modo possam estar correlacionados com a criminalidade.

1. QUESTÃO DE INVESTIGAÇÃO, OBJECTIVOS E MOTIVAÇÕES

Segundo, Quivy & Campenhoudt (2003:32), a questão de investigação constitui normalmente um primeiro meio para pôr em prática uma das dimensões essenciais do processo científico: a ruptura com os preconceitos e as noções prévias.

Esta é a melhor forma de começar um trabalho de investigação, ou seja, enunciar o projecto sob a forma de uma questão. Perante os autores supracitados, uma boa questão de investigação deve permitir meios de auscultação e tratamento de informação, para lhe dar resposta.

Na construção desta questão foram tidas em conta as características que Quivy & Campenhoudt (2003:36) enunciam como as que devem estar presentes para que a mesma possa ser tratada num trabalho de investigação:

- **Clareza** – consiste em ter uma questão que seja precisa, concisa e unívoca. Uma pergunta precisa, não encerra imediatamente o trabalho numa perspectiva restritiva e sem possibilidades de generalização. Permite-nos simplesmente saber aonde nos dirigimos e comunicá-lo aos outros. Ao elaborar a questão pretende-se que seja unívoca, para que seja compreensível e permita perceber o objectivo que persegue;
- **Exequibilidade** – tem a ver com o carácter realista do trabalho que a pergunta deixa antever. Tem de permitir delimitar de forma precisa a temática de estudo, para que não sejam abarcadas matérias de tal forma extensas, que seja impossível obter uma resposta.
- **Pertinência** – é necessário que seja uma verdadeira questão no sentido que não deve conter em si, já a resposta, mas sim, levar o investigador a elaborar um enquadramento teórico, acompanhado pelo menos, de uma abordagem metodológica que se conjuguem na resposta aos objectivos propostos.

Assim, de acordo com os objectivos propostos, formulam-se algumas questões de investigação para o presente estudo:

- Qual a freguesia onde se regista maior ocorrência criminal *per capita*?
- Qual a correlação entre nível populacional e criminalidade?
- Qual o tipo de crime mais praticado?

1.1 Distinção entre dados, informação e conhecimento

A probabilidade da tomada de decisão estratégica ter sucesso, depende da qualidade da informação que o decisor tem sobre o meio envolvente.

Esta informação pode ter diversas interpretações e é importante compreender a diferença entre o que se entende por dado, informação e conhecimento para entender todo o processo de tratamento a que a informação é sujeita, já que os termos dado, informação e conhecimento, são muitas vezes usados de modo diferenciado dependendo do contexto em que estão inseridos:

- **Dados** – É a recolha primária e directa de informação não estruturada. Os dados são recolhidos de forma sequencial não respeitando nenhum critério ou objectivo, por vezes originam bases de dados com pouca qualidade, valores omissos ou incorrectos que consomem muito tempo para corrigir, (pré-processamento dos dados).

Varajão (1998:45) define dados como “representações não estruturadas que poderão ou não ser pertinentes ou úteis numa determinada situação. Por outras palavras, são apenas elementos ou valores discretos que isoladamente não tem qualquer utilidade e cuja simples posse não assegura a obtenção de quaisquer benefícios.”

Já para Rascão (2001:17), dados são “factos, eventos, imagens ou sons que podem ser pertinentes ou úteis para o desempenho de uma tarefa, mas que por si só não conduzem a uma compreensão de determinado facto ou situação.”

- **Informação** – São dados agrupados de forma estruturada e com determinado critério, consideram-se dados secundários. A informação é o estado inicial que “alimenta” o modelo preditivo (*input*). Para Varajão & Amaral (2000:8), informação “É aquele conjunto de dados que, quando fornecido de forma e a tempo adequado, melhora o conhecimento da pessoa que o recebe, ficando ela mais habilitada a desenvolver determinada actividade ou a tomar determinada decisão.”

Para King & Kraemer (1988) *apud* Varajão & Amaral (2000:9), A informação deve ser gerida da mesma forma que os restantes recursos da organização e à semelhança de outros bens económicos.

De acordo com Le Moigne (1978) *apud* Rascão (2001:18), A informação é um “objecto formatado, criado artificialmente pelo homem, tendo por finalidade representar um tipo de acontecimento identificável por ele no mundo real, integrando um conjunto de registos ou dados e um conjunto de relações entre eles, que determinam o seu formato”.

- **Conhecimento** – É a descoberta de padrões de informação úteis em bases de dados, o conhecimento que resulta do modelo preditivo depois de devidamente interpretado (*output*), é o objectivo do tratamento da informação útil. Rascão (2001:17) entende que “conhecimento é uma combinação de instintos, ideias, regras e procedimentos que guiam as acções e as decisões”.

Para Steven, A. (1992) *apud* Rascão (2001:19), conhecimento pode ainda ser definido como “capacidade de uma pessoa relacionar estruturas complexas de informação para um novo contexto, novos contextos implicam mudança – acção, dinamismo. O conhecimento não pode ser partilhado, embora a técnica e os componentes da informação possam ser partilhados.”

Alter, S. (1992:82), representou a relação entre dados, informação e conhecimento, da seguinte forma:

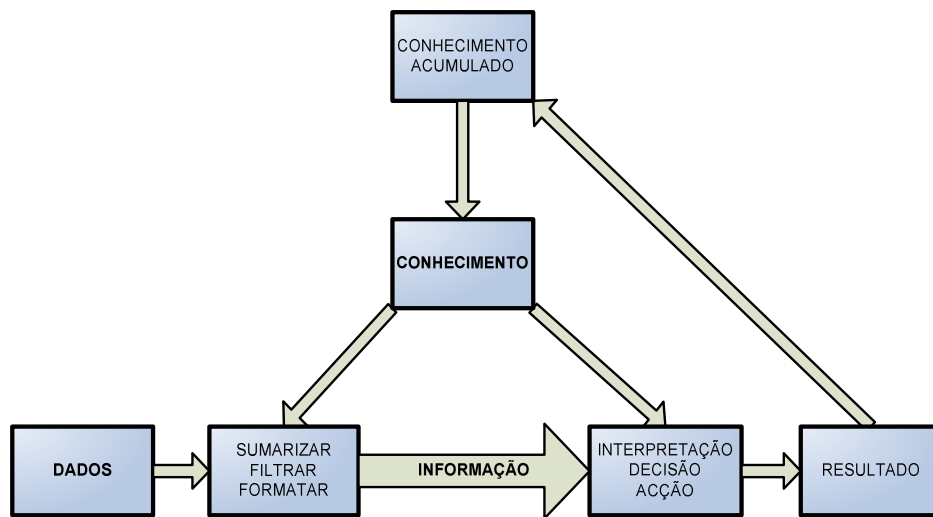


Figura 1 – Processo de conversão dos dados

Sendo os dados fáceis de identificar porque são registos primários e o ponto de partida de todo o processo, o mesmo já não acontece com a relação entre conhecimento e informação. Esta é mais subtil e leva a várias interpretações consoante o autor e envolvente, carece por isso de algum esclarecimento.

Para Kock *et al.* (1997) *apud* Bação (2006:7) refere que “a dicotomia conhecimento / informação, baseia-se no facto de a informação ser descritiva, ou seja, relacionar-se com o passado e o presente; sendo que o conhecimento é iminentemente preditivo, por outras palavras, proporciona as bases para a predição do futuro, com determinado grau de certeza, baseado na informação referente ao passado e presente.”

1.2 Objectivos gerais

O presente trabalho de investigação pretende usar os dados disponibilizados pela PSP referente a criminalidade denunciada e a partir desses mesmos dados, aplicar um modelo preditivo supervisionado, baseado em diversos métodos que definam uma previsão de uma variável *target*, nomeadamente: redes neuronais, árvores de decisão e regressão,

Pretendemos, construir um modelo preditivo que seja inovador na área onde se insere, conseguindo com isso, produzir um método de análise útil e aplicável a casos concretos (e. g. forças e serviços de segurança), que defina geograficamente a criminalidade, no Concelho de Lisboa. Permitindo assim, por antecipação, tomar as devidas medidas de utilização eficiente dos recursos materiais e humanos de modo a maximizar a sua produtividade na prevenção e redução da criminalidade.

Pretende-se quantificar os ilícitos criminais, georeferenciando esses actos ao local de ocorrência ao nível de freguesia. Para isso irá ser utilizada uma base de dados com toda a criminalidade denunciada e registada na PSP no Concelho de Lisboa, no período entre 1 janeiro de 2007 até 31 de Julho de 2007, referente a crimes pessoais e crimes patrimoniais.⁴

Dessa base de dados, irá retirar-se as variáveis necessárias a esta investigação, preservando-se os dados pessoais que não são necessários.⁵

Para alcançar os objectivos, e responder às questões de investigação iniciais, será feito um mapeamento da criminalidade através da visualização dos níveis da criminalidade referentes ao período em análise, nas diversas freguesias de Lisboa.

⁴ Por limitações na qualidade dos dados recolhidos, não nos é possível desagregar os dados para além da freguesia, pois estes não possuem informação de qualidade nem as variáveis necessárias que possam ser úteis no presente trabalho. Salienta-se ainda, que: a criminalidade denunciada não é a criminalidade total, visto que, para além dos muitos crimes que ficam por denunciar, apenas estamos a trabalhar com os dados registados pela PSP e disponibilizados pelo SEI.

⁵ Não são necessários dados pessoais porque este estudo tem por objectivo a previsão criminal e não a investigação criminal das ocorrências ou identificação dos seus autores.

Com a aplicação do modelo preditivo e posterior tratamento dos dados pretendemos ainda detectar quais serão as previsões de criminalidade ao nível da freguesia com base nos dados disponíveis e qual o erro associado face a dados reais.

Este tratamento dos dados têm extrema importância para a implementação de estratégias de policiamento e redução da criminalidade (*risk based deployment*). É com base na actuação policial, que se elaboram estratégias e programas de policiamento.

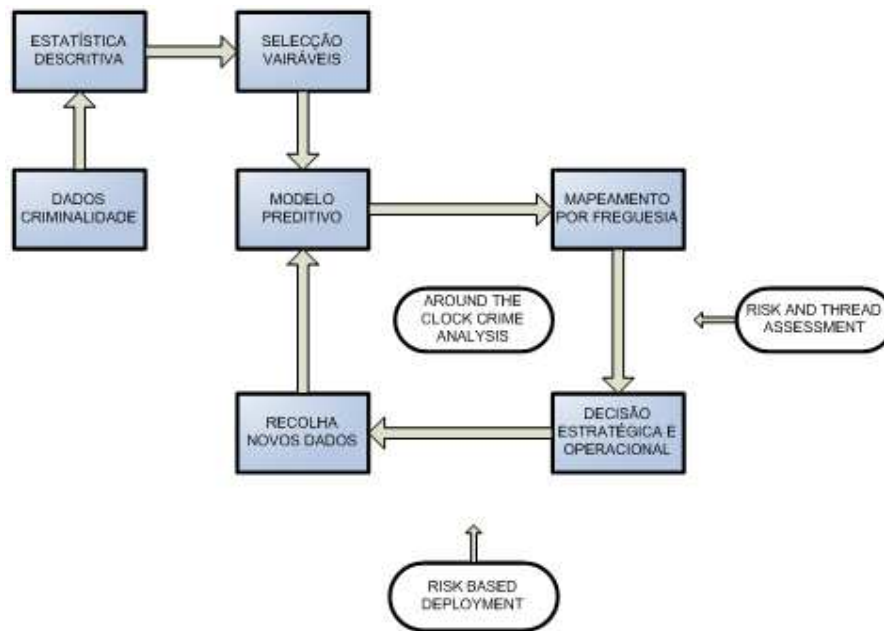


Figura 2 - Processo de tratamento dos dados

A Análise criminal, toma assim, uma importância estratégica no combate á criminalidade que, de acordo com Osborne, D. & Wernicke, S. (2006:4-5), analisar o crime justifica-se nas seguintes razões:

- Para tirar vantagem da abundância de informação existente nas forças e serviços policiais, justiça criminal e domínio público;
- Para maximizar os recursos limitados das forças e serviços de segurança;

- Para informar as polícias sobre as tendências gerais e particulares e agrupamentos criminais de forma atempada;
- Como objectivo para aceder a problemas locais, regionais e nacionais de criminalidade e favorecer a cooperação internacional;
- Como forma proactiva na detecção e prevenção criminal;
- Como objectivo das necessidades das forças de segurança e mudanças na sociedade.

1.3 Interesse pessoal

Os modelos de previsão usados em *Data Mining*, embora tenham por base modelos teóricos suportados em técnicas e pressupostos estatísticos, necessitam de dados reais para permitir testar a sua aplicabilidade e qualidade e assim produzir resultados fidedignos.

Esta investigação, que presumimos ser inovadora em Portugal, é um precioso contributo para os sistemas de suporte à decisão das forças e serviços de segurança nacionais.

Pretendemos com este estudo aliar os conhecimentos teóricos com a experiência profissional do seu autor de modo a produzir um modelo que possa ser extrapolado para outros tipos de crimes ou regiões, no âmbito operacional e estratégico policial. O modelo pode a todo o momento, receber novos dados e produzir novos *output's* que suportem a decisão estratégica.⁶

⁶ É o que se chama de: *around the clock crime analysis*.

1.4 Interesse profissional

O tema desta investigação insere-se na área de trabalho profissional do autor, que ao longo dos anos tem vindo a verificar o tipo de formação policial das forças e serviços de segurança, com o qual lida diariamente no seu local de trabalho.

Tal formação policial, é manifestamente a de repressão da criminalidade, investigação criminal, políticas e programas de proximidade com o cidadão, não se dando a devida atenção a estratégias preventivas.

Em termos operacionais, a actuação policial, desagrega-se em três vertentes estratégicas de actuação para o combate á criminalidade:

- **Policiamento comunitário ou de proximidade** (*NCP - neiborhoud and community policing*) é o tipo mais comum de actuação, onde se enquadra o patrulhamento apeado;
- **Policiamento orientado pelos problemas** (*POP - problem-oriented policing*) define estratégias e programas de redução da criminalidade (e. g. programa de apoio aos idosos, escola segura, operação natal, etc.);
- **Policiamento orientado pelas informações** (*ILP - intelligence led policing*) aqui são os dados do passado que são analisados para elaborar novas estratégias de actuação. É muito exigente em termos de meios tecnológicos e qualificação dos recursos humanos, sendo também o que apresenta os melhores resultados por ter uma abordagem, tanto preventiva como repressiva.

A seguinte tabela, compara os diversos modelos de policiamento em termos de implementação, recursos e resultados obtidos:

	Dificuldade implementação	Recursos humanos	Recursos tecnológicos	Custos públicos	Aplicação geográfica	Resultados Prevenção criminal	Tipo Informação	Tipo actuação
Neighbourhood And Community policing	Baixo	Elevado	Baixo	Elevado	Baixo	Baixo	Retrospectiva	Preventiva
Problem Oriented Policing	Médio	Elevado	Médio	Elevado	Médio	Médio	Retrospectiva	Repressiva
Intelligence Led policing	Elevado	Baixo	Elevado	Médio	Elevado	Elevado	Prospectiva	Preventiva e Repressiva

Tabela 1 – Vertentes estratégicas da actuação policial

O modelo de *intelligence led policing*, é o de mais difícil implementação mas é também o mais versátil na actuação e o que obtém melhores resultados.

1.4.1 Policiamento Comunitário ou de Proximidade

Segundo Carter, D. (1994:373-404), o policiamento comunitário ou de proximidade - *Neighbourhood and Community Policing* - resulta de 10 premissas ou mandamentos que aqui se sintetizam:

1. A polícia orientada para a comunidade é uma filosofia e uma organização estratégica que permite à polícia e à comunidade de residentes, trabalhar estreitamente em conjunto e de novas formas para resolver problemas de crime, medo do crime, desordens físicas e sociais e degradação das vizinhanças;
2. Exige, em primeiro lugar, que todas as pessoas ligadas à polícia, incluindo o pessoal civil, aprofundem formas e processos de traduzir essa filosofia na prática;

3. Os departamentos policiais devem criar e desenvolver um novo tipo de agente policial, o Agente de Polícia de Comunidade, que actua como uma ligação directa entre a polícia e as pessoas da comunidade;
4. O amplo papel do Agente de Polícia de Comunidade exige um contacto contínuo e sustentado com as pessoas cumpridoras da lei na comunidade, para que, em conjunto, possam criar sinergias de mútua confiança e explorar soluções criativas para as preocupações locais, em termos de crime, medo do crime, desordem e degradação, servindo os cidadãos como voluntários não pagos;
5. Implica um novo contrato entre a polícia e os cidadãos por ela servidos, que ofereça a esperança de ultrapassar a crescente e invasora apatia, mas restringindo-se todos os impulsos para a criação de «vigilantes» pois estes poderiam usurpar o papel exclusivo da polícia como órgão de polícia criminal;
6. Junta um elemento de protecção vital ao papel reactivo tradicional da polícia, resultando isto num serviço policial com ampla aplicação;
7. Desenvolve novos processos de protecção e de apoio às vidas dos mais vulneráveis, jovens, idosos, minorias, pobres, deficientes e sem-abrigo;
8. Promove o uso judicioso da tecnologia, mas também se baseia na ideia de que nada substitui o que humanos dedicados, trabalhando e dialogando em conjunto, podem conseguir [comunidades de prática];⁷

⁷ Wenger & Snyder (2000), “as Comunidades de Prática (*COP’s Communities of practice*), são novas formas organizacionais para a criação de conhecimento e têm como características o facto de serem: Informais, movidas pelo desejo de partilhar saber especializado, definirem a sua própria agenda, encontrarem um formato próprio, sustentadas pelo interesse e paixão dos participantes. Estas comunidades ao interagirem num ambiente de aprendizagem criam benefícios para os participantes e para a própria organização.”

9. Deve concretizar um processo totalmente integrado que envolva todos na polícia, com os Agentes de Polícia de Comunidade no papel de especialistas na ligação entre a polícia e as pessoas servidas;
10. Proporciona um serviço à comunidade com características descentralizadas e personalizadas, assume que a polícia não pode impor ordem na comunidade a partir de fora dela, mas que as pessoas devem ser encorajadas a pensarem na polícia como um recurso que podem usar na ajuda à resolução de problemas que preocupam as comunidades actuais.

É, hoje, consensual que, a implementação de um programa de policiamento de proximidade implica, antes de mais, uma doutrina de emprego dos meios, que é conhecida e repousa em 5 modos de actuação destinados a criar confiança entre a polícia e a população:

1. Territorialização da polícia;
2. Contacto permanente com a população;
3. Aplicação de valências policiais sobre determinado território em função do contexto situacional (prevenção, ordem pública, investigação criminal, informações policiais, *etc.*);
4. Responsabilização dos elementos policiais;
5. Especial atenção às vítimas.

Ora, o que se tem constatado é que, se esta doutrina não definir, também, concretas práticas de actuação, será “letra morta”.

Na verdade, se a doutrina de emprego dos meios numa polícia de proximidade está estabelecida, o saber fazer no terreno não está estabilizado e encontramos, em bom rigor, numa fase de experimentação.

Porém, como em qualquer situação policial tradicional, a simples presença policial é insuficiente para o combate eficaz da criminalidade. Para se conseguir os resultados desejados é necessário conseguir o apoio da comunidade.

Quando os níveis de criminalidade aumentam, numa área geográfica, a resposta imediata dos decisores policiais é a de reforçar o policiamento dessa área.

De uma forma geral, segundo Ferret (2003:91), o que se sabe neste domínio pode ser sintetizado na seguinte tabela:

Procedimentos	Espectativas	Resultados
Aumento do efectivo policial	Quanto mais polícias se empregam, mais reduz a criminalidade	Impacto indeterminado
Patrulhas aleatórias	Quanto mais a polícia é percebida como omnipresente, mais impedirá a criminalidade	Ineficaz
Utilização intensiva do poder de detenção	Quanto mais a polícia responde com detenções aos comportamentos ilícitos, mais a criminalidade desce	Eficaz para a violência conjugal; Contraproducente nos menores
Contacto, no sentido lato do termo, com a comunidade	Quanto mais contactos em qualidade e quantidade entre a polícia e os cidadãos, mais a criminalidade desce	Geralmente ineficaz, a não ser quanto ao reforço da legitimidade da polícia
Contactos informais com os jovens	Os contactos informais entre a polícia e os jovens dissuadem os potenciais delinquentes de infringirem a lei	Geralmente ineficaz
Respostas rápidas aos pedidos de auxílio	Quanto mais rápida for a resposta da polícia, mais a criminalidade desce	Resultados duvidosos
Esforço sobre a criminalidade mais grave	Quanto mais detenções houver na criminalidade grave, mais baixa a taxa desta criminalidade	Resultados positivos quanto aos delinquentes reincidentes. Menos eficaz no caso da droga
Afectação de patrulhas a lugares concretos	Quanto mais as patrulhas de polícia se concentram em locais de criminalidade, menos ilícitos aí ocorrerão	Eficaz na resolução de problemas muito localizados
Esforço sobre a vitimação recorrente	Podemos reduzir a criminalidade protegendo as vítimas contra novos ilícitos	Eficaz em certos tipos de criminalidade
Colaboração com diferentes organizações	A polícia pode prevenir a criminalidade através de parcerias com outros organismos	Eficaz e útil para a polícia
Acção orientada para os problemas	Se a polícia poder identificar as causas da criminalidade local, poderá encontrar as soluções para reduzir os crimes	Não há estudos conclusivos

Tabela 2 - Procedimentos de policiamento

Este esforço visa sobretudo a prevenção de novas situações e a detecção de ocorrências criminais. Infelizmente, constata-se que, tem um impacto pouco significativo na redução da criminalidade, já que os criminosos não deixam de cessar a actividade criminosa, apenas deslocam-se geográficamente para outra zona.⁸

De facto, não se sabe em que consistem as boas práticas em matéria de polícia de proximidade e sobretudo, as condições de transmissão do saber fazer.

⁸ Constata-se com frequência que a criminalidade onde a polícia actua, sofre uma redução, mas algures nas imediações sofre um aumento.

Mas para acompanhar o aumento do efectivo policial, também é necessário promover acções específicas de prevenção criminal de forma a obter o apoio da comunidade e conseguir motivar essa comunidade para a importância da tomada de certas medidas individuais e colectivas que reforcem a segurança de todos.⁹

1.4.2 Policiamento orientado para os problemas

O policiamento orientado para os problemas - *Problem Oriented Policing* - foi proposto por Herman Goldstein (1979), como uma resposta às crescentes críticas ao modelo tradicional de policiamento (policiamento comunitário ou de proximidade).

O conceito, teve na sua origem, a resposta a um conjunto de ineficiências, nomeadamente, o facto de a polícia concentrar a maioria dos seus recursos humanos na gestão interna e nas actividades tradicionais de patrulhamento apeado, ignorando os objectivos e os fins estratégicos da polícia e as expectativas da comunidade.

Goldstein, H (1995), verificou que a comunidade assumia que a polícia tinha uma capacidade científica que lhe permitia planear, estabelecer prioridades e utilizar os seus recursos de forma estratégica, no entanto, a realidade era bem diferente, a totalidade da actividade policial, consistia em reagir a solicitações da população, ao que Goldstein, H. (1995) chamou: *Incident-drive*.

Esta foi a primeira contradição identificada, pois a comunidade responsabilizava a polícia pela prevenção da criminalidade, Goldstein, H. (1995:6) acrescenta:

*“Devido ao enorme investimento em responder rapidamente aos pedidos de auxílio, a polícia é primariamente, reactiva. Prevenção é, no seu melhor, uma função periférica na maioria das organizações policiais”.*¹⁰

⁹ Na actuação policial, por vezes é necessário restringir as liberdades, direitos e garantias individuais, para proteger a comunidade.

¹⁰ Goldstein, H. (1995) *“Given the enormous investment in responding quickly to calls for service, police are primarily reactive. Prevention is, at best, a peripheral function in most police agencies”.*

Uma segunda contradição encontrada por Goldstein, H. (1995), relacionava-se com a natureza da organização formal e com a necessidade de actuação e de aplicação das competências por parte dos agentes policiais:

“A polícia está organizada de forma análoga com as forças militarizadas, com características e formalidades semelhantes, rigidez e gestão top-down. Na realidade, a natureza das funções policiais, requerem informalidade, flexibilidade e poder de decisão nas camadas mais baixas da organização.”¹¹

Para Goldstein (1995), é neste contexto que o policiamento orientado para a resolução de problemas teve a sua origem, o objectivo era reduzir e resolver os conflitos organizacionais e operacionais da polícia, na medida em que:

“A peça central que relaciona todos estes elementos, é o compromisso para analisar peças específicas da actividade policial, na expectativa de construir uma base de conhecimento acerca da actividade operacional e criar esforços concertados para resolver problemas relevantes e expectáveis pela polícia”.¹²

Segundo Goldstein (1995), é preciso reconhecer que a polícia não faz apenas cumprir a lei, mas desempenha um conjunto alargado de tarefas dirigidas aos problemas comunitários.

O policiamento não pode ser apenas dirigido para incidentes pontuais, tem também de reconhecer a relação entre incidentes agregando-os de acordo com as variáveis comuns e correlacionando-os.

¹¹ Goldstein, H. (1995) *“The police are still commonly organized along semi-military lines, with attendant characteristics of formality, rigidity and top-down management. In reality, the nature of police function requires informality, flexibility and decision-making at the lowest levels of the organization.”*

¹² *Idem Ibidem “Its centrepiece, to which all of these elements relate, is a commitment to analyzing specific pieces of police business – thereby building a body of knowledge about the business of policing and building into change efforts an appropriate balance of concern for the substantive problems the police are expected to handle”.*

Goldstein, H. (1984), pioneiro no policiamento orientado aos problemas, reconhece que, identificar claramente os problemas é uma tarefa difícil, pelo que propõe o processo SARA para identificar e auxiliar a análise dos problemas, propor uma concepção da resposta adequada, que pode, ou não, ser de natureza policial e a avaliação, de modo a verificar se a resposta é a mais apropriada para um problema em concreto.¹³

Por vezes existe uma sobreposição entre o policiamento comunitário e este modelo, na medida em que ambos projectam as actividades da polícia para lá da tradicional actividade policial. Por outro lado, o policiamento orientado para a resolução de problemas, alarga o conceito de policiamento dirigido, expandindo o mesmo à análise dos problemas da comunidade como o sentimento de insegurança.

Existem diversos factores que contribuem para que o sentimento de insegurança possa ser aumentado ou diminuído, como refere Lourenço (1995:55):

“Sem pretender minimizar a importância do aumento do número de crimes, é preciso buscar noutros factores – em combinação com o crime – a origem deste [agravamento] do sentimento de insegurança.”

Certas características relativas ao indivíduo influenciam directamente o seu sentimento de insegurança, existem estudos que conferem alguma importância à componente genética do sujeito, dadas as diferenças nos níveis de insegurança.

Variáveis como a idade, o sexo, o grau de instrução, a actividade profissional ou mesmo o nível de rendimentos, são determinantes.

À semelhança do medo, que pode ser pessoal ou de ordem social, também o sentimento de insegurança pode ser individual ou social.

O medo pessoal, engloba características intrínsecas do indivíduo como a idade, sexo, grau de instrução, actividade profissional, socialização e extrínsecas ao indivíduo, como as respeitantes ao local de residência ou as incivildades.

¹³ Goldstein, H. (1984) pioneiro neste processo que é conhecido pela abreviatura SARA de: *Scanning problem identification, Analysis in depth evaluation, Response designing & implementing a strategy and Assessment for evaluate the intended effect*. Foi inicialmente implementado na Polícia de *Newport News*, Vancouver, com resultados muito positivos na redução da criminalidade, o objectivo era estimular o pensamento sobre a estratégia de resolução de problemas e não limitar-se a standardizar os processos que são apropriados para cada problema.

Já o medo de ordem social caracteriza-se sobretudo pela interacção social com o crime, as fontes indirectas, experiências de vitimações, isolamento e a actuação policial.

Um dos factores que influenciam o medo em relação ao crime é vitimização repetida, querem tenha sido o próprio, ou pessoas que lhe sejam próximas.¹⁴

Para além deste factor, também a gravidade da experiência de vitimação, o número de vezes que o foram, a facilidade em recuperar de tal acontecimento e, naturalmente, o objecto do crime, provocam o medo generalizado nas comunidades.

Parece contudo que alguma experiência de vitimação podem reduzir o sentimento de insegurança, dado que é mais assustador temer o desconhecido, que é continuamente referenciado e emolado, do que temer algo pelo qual já se passou e do qual não resultaram danos de maior.¹⁵

Em Portugal, houve um aumento do sentimento de insegurança na população de 2000 para 2005, contudo, houve também um aumento na confiança no trabalho das forças policiais.¹⁶

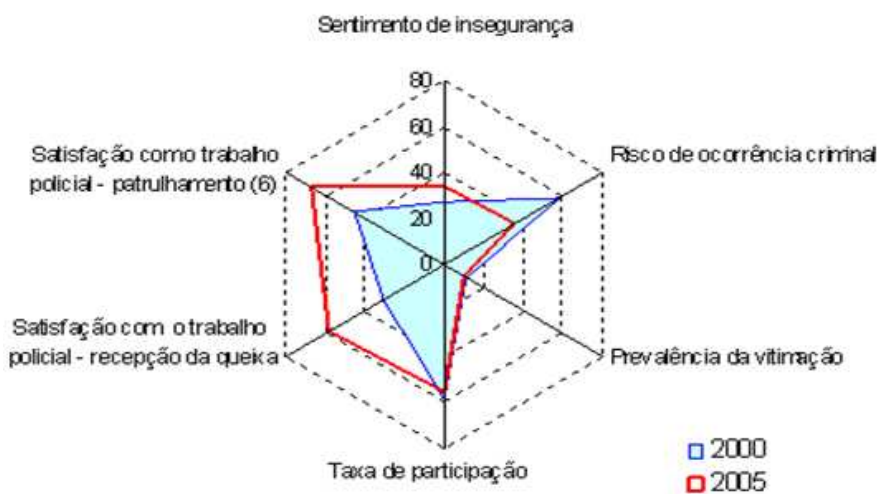


Figura 3 – Avaliação de desempenho das forças policiais

¹⁴ Existem estudos que abordam exaustivamente o fenómeno da vitimização repetida.

¹⁵ Machado & Cabral (2000:53) *apud* Costa, P. (2002:28), detectaram que os delitos contra residências têm efeitos mais nefastos do que os delitos contra pessoas.

¹⁶ Adaptado de <http://www.portalseguranca.gov.pt>

O sentimento de insegurança está também relacionado com a avaliação que o indivíduo faz das forças de segurança, nomeadamente se lhes reconhece competência para resolver o problema que o afecta. A relação estabelece-se pela simpatia que nutre pelas forças de segurança e pela eficácia que lhes reconhece.

Esta tática de policiamento dirigido para a comunidade, mas orientada para a resolução de problemas em zonas problemáticas, acaba por ter um maior impacto na redução da criminalidade e sentimento de insegurança da população.

Contudo, existem factores que deverão ser melhorados: a demora nas investigações policiais leva a que pareça, junto das populações, que as leis favorecem a impunidade dos delinquentes e que o crime compensa.

Para Machado (2004), estas preocupações são legítimas porque são livremente expressas pelos cidadãos, embora se considere que nem sempre são justificadas pelas estatísticas criminais ou pelos acontecimentos reais.

São, por conseguinte, preocupações que se baseiam no discurso do medo que Machado (2004) esclarece como:

“A conceptualização do medo do crime é um processo interpretativo, construído e sustentado através da adopção de um discurso que magnifica o risco criminal e a vulnerabilidade face a ele”.

Segundo Oliveira (2002:82) as características do Modelo de Policiamento orientado pelos problemas (*problem oriented policing*) versus Modelo de Policiamento Comunitário ou de Proximidade (*neighbourhood and community policing*) são as seguintes:

Questão	Policiamento orientado pelos problemas	Policiamento comunitário ou de proximidade
O que é a polícia?	Um organismo do Estado, responsável pela aplicação e cumprimento da lei	Polícia é o público e o público é a polícia, cada agente policial deve envolver-se nos problemas dos cidadãos
Qual é a relação que a polícia estabelece com os outros organismos?	Conflitos frequentes ao nível do estabelecimento das prioridades de acção	A polícia é apenas um dos organismos responsável pela melhoria da qualidade de vida
Qual é o papel da Polícia?	A sua acção concentra-se na detecção e resolução dos crimes	Uma abordagem mais alargada na resolução dos problemas
Quais as prioridades da polícia?	Sobretudo os crimes que envolvam grande alarme público ou violência	Qualquer problema ou desordem que provoque distúrbios na comunidade
De que trata especificamente a polícia?	De incidentes	Problemas e preocupações dos cidadãos
O que é que determina a eficiência da polícia?	Os tempos de resposta às solicitações	A colaboração e cooperação com o público
Qual é a atitude da polícia face às solicitações do público?	Dá-lhes uma resposta se não existir entretanto um trabalho policial para fazer	É uma função vital e uma grande oportunidade
O que se entende por profissionalismo policial?	Combate ao crime de forma eficaz e eficiente	Interagir com as comunidades
Quais as informações mais importantes?	Informação sobre o crime	Informações sobre indivíduos e grupos criminosos
O que é que a polícia deve observar na sua acção?	A centralização, respeito integral dos regulamentos, legitimação pela lei	Ênfase na responsabilização local perante as necessidades de segurança das comunidades
Qual o papel das estruturas de comando?	Desenvolver os regulamentos e emitir directivas	Desenvolver valores organizacionais
Qual o papel das relações públicas?	Manter o enfoque dos média longe da actividade operacional para que estes possam desenvolver o seu trabalho	Coordenar um canal de comunicação essencial com a comunidade
De que modo encara a polícia as detenções?	Como um objectivo importante	Como uma ferramenta entre muitas outras

Tabela 3 – Características do modelo de policiamento

1.4.3 Policiamento Orientado pelas Informações

Ratcliffe (2003:3), defende uma concepção do modelo de policiamento orientado pelas informações - *Intelligence Led Policing* ou *ILP* - como um processo contínuo e permanente de interacção entre as informações, os decisores policiais e o ambiente criminal, da seguinte forma:

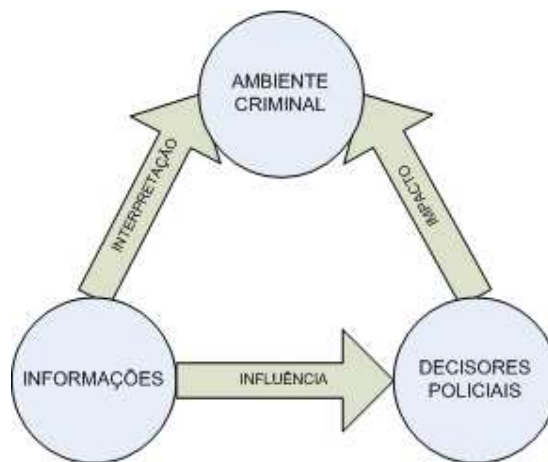


Figura 4 – Processo *ILP*

Esta concepção do modelo ILP, pressupõe uma capacidade das organizações policiais de interpretar o ambiente criminal com recurso a especialistas em análise criminal inseridos em unidades orgânicas e descentralizadas geograficamente e que fazem uso de todas as fontes possíveis de informação.

A inteligência policial, vulgo informações policiais, marca a vida urbana “os *actos mais claros de um homem*», escreveu *Joseph Conrad (s. d.) apud Tofler, A. (1991:355), “têm um lado secreto.”*¹⁷

A inteligência policial favorece a predição de perigos e o afastamento de incivildades socialmente alarmantes e facilita a gestão de incidentes na via pública.

¹⁷ Tofler, A. (1991), “Os novos poderes”.

O saber fazer policial não se resume apenas à resposta reactiva, por isso mesmo, as forças e serviços de segurança devem manter um sistema interno de informações ao nível operacional e estratégico, para o acompanhamento de certos fenómenos criminais que perturbam e aumentam o sentimento de insegurança junto da comunidade.

Por certo, as informações contribuem para a pro-actividade e por isso, para a eficácia da acção policial, tanto na manutenção da ordem pública, como na resposta à prática criminal. Desta forma, a inteligência policial proporciona a reposição célere da legalidade e reforça o dever cívico do cumprimento da lei.

Os resultados da análise do ambiente criminal são indicadores úteis aos decisores policiais que lhes permitem tomar decisões estrategicamente acertadas de acordo com os recursos humanos e materiais disponíveis, na prossecução de medidas de “combate” e redução dos índices de criminalidade.

Ratcliffe (2003), identifica este modelo, três elementos críticos:

1. A interpretação do ambiente criminal;
2. O processo de decisão;
3. O impacto no ambiente criminal.

A interpretação do ambiente criminal, exige avultados investimentos materiais e humanos sendo de difícil implementação no terreno, no entanto espera-se que, caso se concretize, o sistema de informações disponha da necessária capacidade de análise e tratamento da informação do ambiente criminal.

O processo de decisão, é controlado pelos decisores policiais que controlam todos os recursos que podem influenciar o ambiente criminal, no entanto, nem todas as decisões podem ser implementadas por haver recursos escassos, o que invalida a capacidade de resposta em tempo útil prejudicando a eficácia policial.

É necessário hierarquizar e priorizar as medidas a adoptar face aos recursos disponíveis e definir estratégias de acção que potenciem os resultados.

Para Sherman *et al.* (1998), por vezes estas medidas implicam uma profunda mudança da visão organizacional, processos instituídos e cultura policial.

O impacto no ambiente criminal, não depende do sucesso dos outros elementos críticos, é necessário ter a capacidade de influenciar o ambiente criminal de maneira proactiva e orientada para os objectivos definidos estrategicamente.

Sherman *et al.* (1998), acrescenta que por vezes são tomadas medidas pouco eficazes sob o ponto de vista policial, mas que são apelativas e geram apoio popular da comunidade.

O presente estudo, vem ao encontro da necessidade de colmatar uma lacuna existente de *Intelligence Led Policing* nas organizações policiais de modo a possibilitar uma nova abordagem da prevenção da criminalidade baseada em metodologia científica e optimização dos processos existentes com uma melhor gestão dos recursos humanos e materiais.

2. REVISÃO CRÍTICA DA LITERATURA

O tema desta investigação está em constante desenvolvimento e inovação, trata-se de um assunto que lida com os comportamentos das pessoas o que origina muita imprevisibilidade e requer apoio de outras ciências como a econometria, sociologia e mesmo a psicologia.

Todos os dias, novos métodos são implementados e novas técnicas de visualização são testadas, o que nos leva a procurar literatura em vários meios, nomeadamente em artigos científicos e livros de referência sobre o tema.

Foi efectuada uma revisão crítica da literatura com artigos directamente relacionados com o tema em investigação, relegando outros assuntos mais especializados para as outras ciências e especialistas dessas temáticas.

Este tipo de abordagem é tendencialmente não paramétrico e recorre normalmente a grandes bases de dados como fonte secundária de informação.

Devido á elevada densidade populacional e elevado índice criminal, o Brasil e os Estados Unidos da América são pioneiros na implementação de sistemas de *Crime Mapping*. A principal área de actuação destes países, refere-se à aplicação de Sistemas de Informação Geográfica, vulgo SIG, nos sistemas tradicionais de mapeamento da criminalidade.¹⁸

A *Royal Canadian Mounted Police* (RCMP) é a polícia federal do Canadá e tem competências nas áreas de investigação e prevenção criminal. Um dos sistemas mais elogiados desta polícia é o ViCLAS - *Violent Crime Linkage Analysis System* (Sistema de Análise e Relacionamento de Crimes Violentos).¹⁹

O ViCLAS é um sistema informático que liga em rede os departamentos da RCMP e onde são inseridos dados sobre todos os crimes violentos cometidos no Canadá: homicídios e tentativas de homicídio, assaltos, crimes sexuais, pessoas desaparecidas e sequestros.

¹⁸ Estes sistemas tradicionais de *crime mapping* concretizavam-se usando mapas em papel onde se marcavam as ocorrências criminais.

¹⁹ In <http://www.rcmp-grc.gc.ca>

O sistema está programado para analisar e procurar automaticamente possíveis relacionamentos entre os crimes que se encontram na base de dados.

Actualmente o ViCLAS tem aproximadamente 20 mil casos inseridos no sistema e já foi possível estabelecer 3.200 correlações entre eles.

Já o Reino Unido, utiliza uma abordagem mais “paramétrica” do mapeamento da criminalidade, usando sistemas e métodos estatísticos clássicos na elaboração de modelos preditivos que possam ser fiáveis e úteis no combate e prevenção da criminalidade.

Segundo Sykes, G. (1978:7), ao estudar o crime devemos ter consciência de que as descobertas científicas, normalmente consideradas como impessoais e objectivas, trazem invariavelmente consigo a marca do tempo e do lugar.

Para Groff (2001:31), a capacidade de prever as localizações de crimes futuros, pode servir como um valor inestimável como fonte de conhecimento para as forças e serviços de segurança, nomeadamente nas perspectivas táticas e estratégicas.

Paralelamente com as táticas tradicionais de redução da criminalidade, terá de haver uma estratégia para os novos tipos de crimes.

O desenvolvimento das tecnologias de informação tem vantagens inegáveis para a sociedade, no entanto também tem uma vertente obscura de novas formas de actuação, entre os quais o crime informático. Os sistemas informáticos permitem formas sofisticadas para o cometimento de crimes.

A sociedade moderna depende dos sistemas informáticos de modo que qualquer intervenção ilícita pode colocar vidas em perigo.

Apesar de existir unanimidade quanto à existência do fenómeno da criminalidade informática, ainda não foi possível chegar a uma definição consensual do que é o crime informático.²⁰

²⁰ Em Portugal, a Lei 109/91, de 17 de Agosto tipifica os crimes referentes à criminalidade informática e a Lei 109/2009 de 15 Setembro Aprova a Lei do Cibercrime, transpondo para a ordem jurídica interna a Decisão Quadro 2005/222/JAI, de 24 de Fevereiro

2.1 Metodologias de mapeamento da criminalidade

Um tipo de trabalho como este, envolve um conjunto substancial de variáveis e áreas do conhecimento, no entanto não se pretende analisar exaustivamente todas essas vertentes. Pretende-se apenas fazer uma análise quantitativa face a dados concretos da criminalidade denunciada. Recolher os dados do passado, fazer uma análise preditiva usando um modelo preditivo, validar esse mesmo modelo pela medição da taxa de erro e visualizar os resultados obtidos por freguesia.

O modelo preditivo faz uma previsão das ocorrências criminais nas freguesias de Lisboa, conjugando variáveis correlacionadas como: o tipo e subtipo de crime, o local da ocorrência, o índice urbanístico e populacional. Posteriormente, os resultados são convertidos em mapas temáticos.

Para Groff (2001:34) existem diversos métodos de mapeamento criminal sendo os mais comuns os seguintes:

- **Hot-Spot** (ponto-quente), é a metodologia mais utilizada na representação criminal, assume que as localizações de concentração criminal do passado irão manter-se no futuro, no entanto os resultados efectivos deste método dependem do período de tempo em análise, normalmente este método só é robusto e produz bons resultados quando aplicado a séries temporais de curta duração (Spellman, 1995).

Para Admans-Fuller's, (2001), uma característica interessante na detecção de *Hot-Spots* é a sua persistência e coincidência ao longo do tempo como demonstra Anselin *et al.* (2000), os *Hot-Spots* reflectem inicialmente elevados níveis de criminalidade moderada, mas, ao longo do tempo, tendencialmente, essa criminalidade altera-se para tipos de crimes mais violentos (*e. g.* actos de vandalismo que evoluem para crimes de furto).

Por isso, deverão ser contidos e controlados atempadamente a fim de prevenir ocorrências mais gravosas para as pessoas e património na área geográfica coberta pelo *Hot-Spot*.

De acordo Eck, *et al.* (2005), este método, parte do princípio de que se deve mapear os locais e não as ocorrências criminais para assim se compreender porque é que, certos locais, têm mais facilmente ocorrências criminais enquanto outros parecem inibir essas mesmas ocorrências.

Ainsworth (2001:88), define *Hot-Spot* como sendo:

“Um Hot-Spot criminal é normalmente entendido como uma localização ou pequena área com fronteiras bem identificadas onde existe uma concentração de incidentes criminais, que excedem o normal para essa área; o termo pode também ser usado para descrever localidades que demonstram um crescimento da criminalidade num determinado período de tempo...”

- **Repeat Victimization** (vitimização repetida), também chamado de *hot-dots*, refere-se a modelos preditivos de curta duração, o conceito aplica-se a indivíduos ou locais que foram vítimas uma vez e que terão probabilidades elevadas de voltarem a ser vítimas de novo, sendo o tempo curto entre ocorrências, normalmente dois a três meses.

A pesquisa de Anderson *et al.* (1995), neste tipo de método, sugere que locais de *Repeat Victimization*, tendem a ser no futuro próximo, locais de concentração da criminalidade (*Hot-Spots*);

- **Métodos Univariados**, estes métodos usa o valor de uma determinada variável para prever a ocorrência de crimes, é um método atractivo porque requer pouca informação *a priori*; estes métodos vão desde o simples “*random walk*” (utilizado sobretudo quando temos mudanças de padrões de criminalidade frequentes) e “*naive lag*”, (método de análise de séries temporais baseado no momento imediatamente anterior para prever o momento imediatamente posterior) até métodos mais sofisticados que envolvem variáveis sazonais e séries temporais complexas (e. g. utilização de métodos de *Kalman Filter*). Para Gorr & Olligschlaeger (2001), os Métodos univariados, embora sejam de fácil implementação, estão ainda longe de serem os mais fiáveis e correctos por apresentarem um nível elevado de erro, embora a inclusão de mais variáveis, torne o modelo mais robusto mas também mais complicado, entre esses métodos temos: “*exponential smoothing*” e o “*pooling data*”.

- **Time Span** (lapso temporal), é um método relativamente novo associado à criminalidade onde se analisa o tempo decorrido entre ocorrências criminais, são utilizadas redes neuronais associadas a modelos autoregressivos NARX com *time series* de dados criminais onde o *lag* nos *inputs* serve de previsão para o *output*, ou seja, quando as entradas estão desfasadas por meses, na saída obtemos a previsão do mês seguinte para determinado *cluster* de ocorrências criminais, (difere do *naive lag* por basear-se na estrutura do momento anterior e não apenas no seu valor), este método é especialmente adequado para previsões a curto prazo e assume um bom comportamento das séries temporais.

Este conceito relativamente único e associado à criminalidade; por vezes é apenas conhecido pela vítima, por exemplo, que a sua casa foi assaltada depois de regressar das férias, ou seja, não se sabe o momento, nem a duração, apenas um espaço temporal de ocorrência da infracção.

Ratcliffe (2000), foi precursor nesta abordagem no que é conhecido por *Aoristic Analysis*, que apostam em “algoritmos” e probabilidades para tentar vislumbrar um período temporal dentro do *time span* para a ocorrência do crime. Um bom exemplo para entendermos a dificuldade, é a questão dos “três roubos a habitações”. Imaginemos que três roubos ocorreram num fim-de-semana, porém a primeira família saiu sexta-feira ao final do dia, a segunda família sábado de tarde e a terceira família sábado de madrugada, regressando as duas primeiras famílias domingo à noite e a terceira ao final do dia de sábado. De facto, empiricamente, sábado à tarde parece ser a marca mais “lógica” para os três roubos, sendo sobre este tipo de problemas que se debruça a literatura referida, através de análises probabilísticas e cálculos matemáticos.

2.2 Situação em Portugal

Portugal, no caso concreto da PSP, presume-se que não haja tratamento da informação tendo em vista o mapeamento criminal, embora exista a figura do analista de informações, na prática, este, limita-se a elaborar relatórios da criminalidade, não fazendo o devido tratamento de inferência estatística.

As esquadras de polícia utilizam o conhecimento tácito dos agentes como meio de localizar os locais mais susceptíveis de ocorrências criminais, os diversos locais de concentração da criminalidade, exigem um método próprio e uma adequada actuação, e que, por cada tipo de crime deve-se prevenir adequadamente e em tempo útil.²¹

As estratégias estão muito viradas para a repressão policial para conter a criminalidade, ainda não existe nenhum sistema que permita chegar a modelos preditivos que possam prever e actuar em estratégias de prevenção com justificação científica.

Nos últimos anos, nota-se uma viragem para os programas de policiamento de proximidade da polícia para com populações alvo específicas (e. g. Escola Segura, Apoio a Idosos, Operação Natal, Operação Páscoa, Operação Férias, etc.).²²

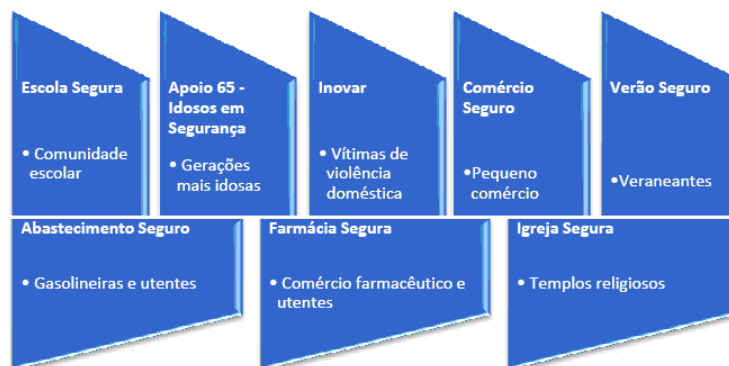


Figura 5 – Programas de proximidade

²¹ Destaca-se em especial os bairros problemáticos onde são constantemente solicitados a intervir; o jornal Diário de Notícias de 11 de Maio de 2009 página 4, refere que: “Cerca de 70% da criminalidade mais grave da área da Grande Lisboa foi praticada por residentes dos bairros problemáticos.”

²² PIPP – Programa Integrado de Policiamento de Proximidade

Nenhum destes programas tem a componente previsão, limitando-se a actuar de acordo com as datas e objectivos estabelecidos *a priori*.²³

Para melhor aferir a realidade portuguesa, no dia 31 de Março de 2008, foi feita uma entrevista semi-estruturada com o Chefe da Divisão de Informações Policiais da PSP (Subintendente Alexandre Coimbra), a fim de detectar as lacunas existentes e oportunidades de melhoria dos processos de recolha e tratamento das informações policiais na PSP, as conclusões a que chegamos juntamente com alguns *reviews* de artigos permite-nos detectar o *state of the art* desta temática, que será desenvolvido posteriormente.

Transcreve-se de seguida o resumo e análise de conteúdo da entrevista efectuada, que, embora semi estruturada através de um guião, deixou liberdade ao entrevistado para expor e desenvolver os temas como bem entender:

- A PSP faz uma análise quantitativa e qualitativa no aspecto policial e sociológico dos dados da criminalidade que são captados e registados pelos diversos comandos no território nacional;
- A análise qualitativa resume-se a estudar os novos fenómenos criminais a fim de estabelecer padrões e perfis da criminalidade, em concreto: o fenómeno *carjacking*, roubo de caixas multibanco, burla com idosos *etc.* Basicamente são aquelas ocorrências que mais influenciam o sentimento de insegurança junto da comunidade;
- Não existe *software* específico no tratamento da informação, sendo os dados registados em folhas de cálculo *Excel* e agrupados numa base de dados *Access*, desenvolvida especialmente para o efeito, refira-se que esta base de dados é externa ao SEI;
- A base de dados é preenchida mensalmente originando um lapso temporal de 30 dias em relação à ocorrência dos factos que os originaram;

²³ Pensamos que estes programas eram uma excelente oportunidade da PSP para recolher dados dos cidadãos para posterior análise e elaboração de estratégias de intervenção.

- Os dados enviados para a base de dados, são agrupados de acordo com o tipo de crime e actividade operacional;
- Posteriormente, os dados são agrupados com outros dados de outras entidades (PJ e GNR), que originam os relatórios mensais de toda a criminalidade denunciada;²⁴
- Estes relatórios mensais são recebidos das outras entidades em papel e posteriormente colocados na base de dados do crime violento;
- Para a PSP, informação, são os dados trabalhados, ou seja, o seu cruzamento com outros dados, originando valor acrescentado;
- Ainda não existe tratamento desta informação de modo a gerar conhecimento útil ao nível estratégico e operacional;
- O controlo de qualidade dos dados é feito localmente e depende das pessoas que estão á frente da esquadra;
- Cabe ao comandante da esquadra a preocupação [ou não] de averiguar a qualidade dos dados inseridos no SEI;
- Não existe nenhum mecanismo centralizado de controlar a qualidade, mesmo o SEI é ineficaz neste ponto, já que não tem filtros na entrada dos dados, originando, redundância e diversos erros sistemáticos;
- Existem diversas atribuições entre departamentos, assim, cabe ao departamento de operações, instruir os comandantes de esquadra sobre os procedimentos a adoptar e ao departamento de informações receber e analisar os dados;

²⁴ Perintrep (*Periodic Intelligence Report*) e Sitrep (*Situation Report*).

- Nos cursos de formação de Agentes já existe uma disciplina de informações policiais que procura sensibilizar os futuros Agentes para que a recolha de dados seja feita correctamente;
- Existe ainda um curso de análise de informações policiais para o pessoal de investigação criminal que contempla a preocupação de estabelecer critérios de qualidade na recolha dos dados;
- O SEI é o único *standard* existente, no entanto possui imensas falhas, nomeadamente: a não configuração dos dados de entrada, originando diferentes formatos para os mesmos dados e registos redundantes ou mesmo dados omissos nos registos, É a informatização do papel e não uma verdadeira mudança de procedimentos;
- O departamento de informações, sente diariamente esta falta de qualidade dos dados, levando bastante tempo para os corrigir [quando é possível] de modo mais ou menos satisfatório, já que não são os elementos primários da recolha dos dados;
- Os elementos do departamento de informações policiais, gastam mais tempo a corrigir os dados do que propriamente na sua análise;
- O SEI foi programado para recolher os dados mais importantes numa ocorrência criminal: o quê?, a onde?, quando?, porquê? e quem?. No entanto a maneira como esses dados são registados, originam muita informação redundante e “suja” ou omissa por falta de filtros na entrada dos dados;
- Outro aspecto importante é a omissão na recolha de variáveis geográficas do local das ocorrências criminais, sendo apenas registado o tipo de local: estabelecimento de ensino, farmácia, banco, ourivesaria, estação correios, *etc.*;
- Com os dados inseridos no SEI apenas se faz: estatística descritiva ou exploratória, relatórios da criminalidade e relatórios semestrais. Não existe

nenhum tratamento ao nível da inferência estatística ou utilização de qualquer método preditivo;

- Em termos de salvaguarda dos dados pessoais, houve a preocupação em aplicar a Lei 5/95;
- O registo das ocorrências criminais, subdivide-se em dois grandes grupos: crime comum e crime violento e grave;
- É o crime violento e grave que gera maior sentimento de insegurança junto das populações;
- O entrevistado, considera que é necessário e fundamental um sistema preditivo da criminalidade num futuro próximo, para se detectar quais são as prioridades no combate ao crime e alocação de meios humanos e materiais;
- Ninguém olha para as informações para tomar decisões, observa-se a tendência da evolução da criminalidade, mas ninguém analisa o porquê dessa situação e muito menos se alteram os processos instituídos tendo em vista melhores práticas de actuação;
- Existe uma preocupação em dividir a actuação interventiva e repressiva da actuação preventiva, de acordo com as valências e formação dos recursos humanos. Para isso é importante canalizar os recursos para essas áreas de intervenção e repressão (mais especializados e equipados) ou prevenção (visibilidade policial);
- O entrevistado considera que um *hot spot* criminal é aquele local ou locais ou zonas onde existe um determinado tipo de criminalidade que preocupa a polícia em termos de segurança do cidadão e que carece de actuação policial quer em termos preventivos quer em termos repressivos ou reactivos;
- Considera ainda, de acordo com a sua experiência profissional, que, mesmo com intervenção policial o *hot spot* vai transferir-se para outro local, mas

regride no local da intervenção policial, o que origina falta de meios para actuar em outro local num determinado período de tempo, originando uma constante reorientação de meios humanos e materiais para fazer face a estas mudanças;

- A análise de *hot-spots* tem de partir do geral para o particular, a visualização deve localizar os locais de maior criminalidade e depois descrevê-los em cada localização (tipo de crime praticado) e ao mesmo tempo irmos reduzindo a escala geográfica até ao nível de rua,²⁵
- Os meios de acção e reacção são diferentes, cada tipo de crime tem uma medida de actuação específica, (pessoal a civil ou fardado?, reactiva ou preventiva?, mais visibilidade policial ou menos?, equipamento necessário?, manutenção ou reposição da ordem publica? etc.);
- A análise de informações policiais não possui nenhum *software* específico para o efeito;
- ILP (*intelligence led policing*) é a utilização dos dados do terreno de forma inteligente e de acordo com os meios existentes, é muito pouco utilizado na PSP, existe pouca sensibilidade e falta formação nesta área específica.

Como conclusões da entrevista efectuada, pode-se afirmar que:

A PSP não faz o devido tratamento dos dados recolhidos no “terreno” de forma útil e atempada para o processo de decisão estratégica e operacional. Os dados são simplesmente objecto de estatísticas descritivas básicas, não se aprofundando a sua análise ou descoberta de padrões que são a base de novo conhecimento explícito.

A colaboração com outras entidades ainda se faz de forma arcaica, não obedecendo a critérios de qualidade, interoperacionalidade e análise global do fenómeno da criminalidade.

A recolha e qualidade dos dados inseridos no SEI, necessita de melhoramentos significativos, o sistema de informação não se encontra adaptado para aprofundar uma inferência estatística dos dados.

²⁵ Esta atitude parece sensata mas na realidade é impraticável porque “no terreno” não são recolhidos dados que permitam uma georeferenciação eficaz das ocorrências.

Não existe um controle de qualidade directo na recolha primária dos dados, deixando ao elemento que insere os dados a opção de como, quando e com que qualidade o faz.

3. METODOLOGIA DE INVESTIGAÇÃO

O tipo de abordagem metodológica será a quantitativa, tendo em conta os dados utilizados e os objectivos a atingir.

A pesquisa quantitativa, também conhecida por análise positivista, é um dos mais importantes métodos de investigação de fenómenos organizacionais.

Para Gefen, Straub & Boudreau (2000:1-78), a pesquisa quantitativa é:

“...Ortogonal ao modelo analítico, um método científico que depende apenas de derivações matemáticas e de suposições para interpretar a realidade...”²⁶

Acrescenta ainda que é um tipo de pesquisa, dedutiva por natureza e:

“...Menos adequada para elaborar teorias porque os dados numéricos não tem a riqueza contextual e social que os dados qualitativos...”²⁷

Este tipo de limitação pode ser ultrapassado através da incorporação de dados quantitativos recolhidos especialmente para este efeito.

O questionário é a forma mais comum de obter os dados quantitativos, que perante Martyn & Lancaster (s. d.) *apud* Heather & Stone (1984), deve ser:

“...Breve, atractivo, as perguntas não são ambíguas, é de fácil interpretação e completo; pode ser analisado sem esforço e interpretado sem dificuldade para prover informação clara e concisa...”²⁸

²⁶ Gefen, Straub & Boudreau, (2000) *“...orthogonal to analytical (math) modeling, a scientific method that typically depends on mathematical derivations and assumptions for interpreting reality....”*

²⁷ *Idem Ibidem* *“...less suited to building theories because numbers are not contextually and socially rich as soft data...”*

A elaboração de um questionário que reúna todas estas qualidades é de difícil execução.

No presente estudo optou-se por solicitar á PSP os dados da criminalidade denunciada, referente a crimes pessoais e patrimoniais no primeiro semestre do ano de 2007 no Concelho de Lisboa.

Actualmente existem programas específicos para trabalhar e visualizar dados criminais dos quais os mais conhecidos são o *CrimeSTAT* e o *Mapinfo*, mas o facto de exigirem uma licença paga, não serem adaptados para língua portuguesa e de necessitarem de dados georeferenciados, impossibilita a sua utilização, neste trabalho.²⁹

Assim, utilizamos na fase de pré-processamento e tratamento dos dados: o *Microsoft Excel* com o *add-in XLstat 2009*, para as estatísticas descritivas iniciais e selecção de variáveis relevantes, para o modelo preditivo foi utilizado o *SAS Enterprise Miner*, quanto á visualização dos dados e mapeamento criminal foi utilizado o *Geomedia Pro*.

3.1 Abordagem de investigação

Será utilizada uma abordagem baseada maioritariamente na análise quantitativa dos dados e com uma finalidade descritiva e positivista. Este tipo de abordagem caracteriza-se por ser maioritariamente quantitativa e interpretativa, onde o autor é alheio ao que está a investigar.

²⁸ Heather, P. & Stone, S., (1984) “...*brief, is attractive, asks unambiguous questions, is interesting and easy to complete, can be analysed with little effort and interpreted without difficulty to provide clear and concise information’s...*”.

²⁹ Para mais informação do *CrimeSTAT* consultar: <http://www.icpsr.umich.edu/CRIMESTAT/>, para o *Mapinfo*: www.mapinfo.co.uk/, uma aplicação prática do *CrimeSTAT* pode ser visualizada em: <http://www.winnipeg.ca/crimestat/d6.stm>

A descrição desta abordagem é dada por Creswell (2003:4):

“Abordagem experimental, empirista ou pós-positivista - emana da escola do pensamento que considera que todo o conhecimento deriva da observação directa e de inferências lógicas baseadas na mesma (observação directa) – utiliza os métodos estatísticos para descobrir relações e padrões e expressar esses padrões em dados numéricos.”

Creswell (2003:7), acrescenta ainda:

“Está associada a uma filosofia determinista segundo a qual as causas determinam os resultados; os problemas estudados pelos positivistas reflectem a necessidade de examinar as causas que determinam os efeitos, portanto as questões são estudadas através de experiências.”

Iremos utilizar esta abordagem por ser a mais adequada para validar os objectivos a que nos propormos e responder às questões de investigação, esta abordagem justifica-se devido ao tratamento dos dados utilizado, quer como estatística descritiva quer como estatística inferencial ou preditiva.

3.2 Recursos e dados disponíveis

Para o presente estudo recorreu-se a fontes secundárias para obter os dados necessários, estas fontes provêm de dados primários que são, de acordo Mann, P. (1983:63) “recolhidos em primeira mão por outras pessoas”.

Os dados recolhidos para o presente estudo, necessitam de ser fidedignos, para isso, recorreu-se a dados reais, da base de dados da PSP, embora se saiba *a priori*, que estes dados são apenas uma amostra da criminalidade denunciada, no entanto, consideram-se ser suficientes para a elaboração e validação do modelo preditivo.

Foram obtidas outras informações recorrendo ao Relatório Anual de Segurança interna 2008, onde refere que, a criminalidade denunciada em 2008 teve um aumento significativo face a 2007, superando mesmo os máximos registados em 2003 e 2004.

Os crimes patrimoniais (54% da criminalidade participada) foram os mais praticados em Portugal no ano de 2007 logo seguidos pelos crimes pessoais (24% da criminalidade participada). Justifica-se assim, a escolha destes grupos de crimes como objecto de estudo do presente trabalho.

Já de acordo com o RASI 2008, os crimes contra as pessoas voltaram a subir 1,5% e os crimes contra o património 13,9% face a 2007.



Figura 6 – Balança da criminalidade violenta

Estes dois tipos de crimes, foram os que mais contribuíram para o aumento da criminalidade violenta na última década, registando o ano de 2007 valores semelhantes ao ano de 2001, no entanto, em 2008, houve uma subida substancial, atingindo praticamente os valores registados em 2006 o que inverteu a tendência de descida que se vinha a verificar em 2007.

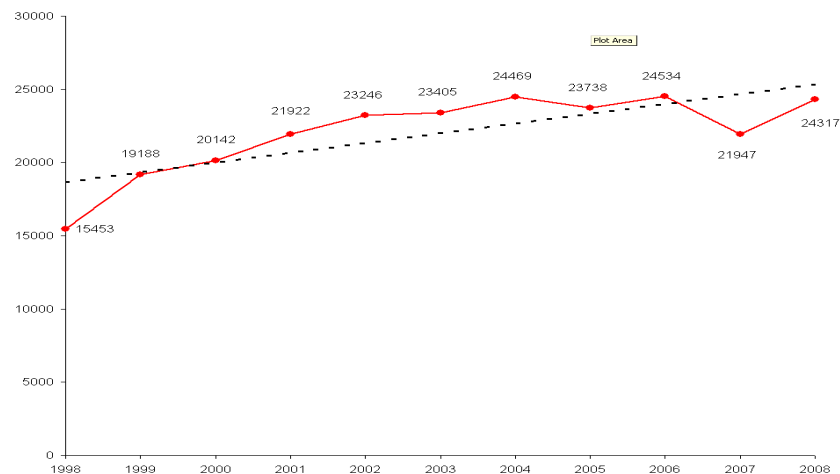


Figura 7 – Evolução da criminalidade violenta denunciada na última década

Os dados disponibilizados pela PSP, são apenas uma pequena amostra das ocorrências criminais. Apenas foram solicitadas os dados referentes a crimes pessoais e patrimoniais, o que pensamos ser suficiente, mesmo assim deve-se alertar para o facto de nem todas as ocorrências criminais serem denunciadas e destas apenas uma parte ser efectivamente registada no sistema estratégico de informações (SEI) de onde se recolheu os dados para este estudo.

Em termos geográficos, de acordo com o RASI 2008, Leiria foi o distrito com maior aumento dos níveis de criminalidade em 2008 com 15% (16.060 ocorrências), mas Lisboa continua a ter o maior número de ocorrências criminais tendo registado um aumento de 10,4% em 2008 face ao ano de 2007 com 110.211 ocorrências criminais registadas.

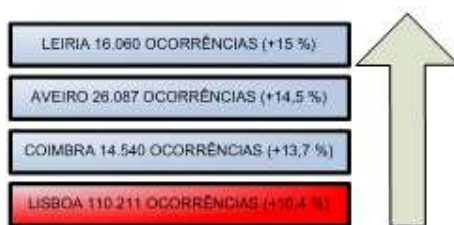


Figura 8 – Geografia da criminalidade

Em termos de recolha de informação e como já foi referido, os dados utilizados representam uma pequena parte da realidade criminal, pois apenas estamos a considerar os crimes denunciados e registados pela PSP o que nos afasta do número real de ocorrências cometidas.



Figura 9 - Criminalidade denunciada

Se tivermos em conta as ocorrências registadas na GNR, PSP e PJ, o total de ocorrências sobe para 421.037 no ano de 2008 o que corresponde a um aumento de 7,5 % face ao ano de 2007.

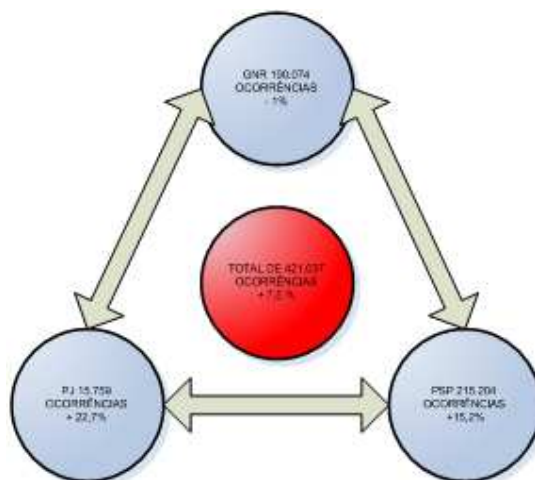


Figura 10 – Registos por entidades policiais

3.2.1 Interpretação e dificuldades das estatísticas oficiais

Para Manheim, H. (1985:170-183), a expressão “crimes conhecidos pela polícia” revela já a nossa consciência de uma limitação comprometedora do valor das estatísticas criminais indicando a existência de crimes que persistem inteiramente desconhecidos pela polícia ou que não são, pelo menos, clarificados, isto é susceptível de várias explicações:

1. O cometimento do crime que pode permanecer inteiramente desconhecido (e. g. o desaparecimento de pessoas que vivem sozinhas sem qualquer relação de parentesco, furtos sem que as pessoas notem que foram atingidas no seu património, a não comunicação de violação por vergonha da vítima, etc.);
2. Falta de um determinado objecto mas explicar-se o evento de forma inócua (e. g. um homicídio ser facilmente disfarçado como acidente ou mesmo suicídio, etc.).

Em certa medida, as estatísticas oficiais da criminalidade, reflectem a eficiência da polícia, há contudo, factores concorrentes: crimes pelos quais a polícia se limita a reprimir e identificar os seus autores, reduzem o número de condenações sem se reflectirem necessariamente nas estatísticas oficiais de eficiência da polícia, pois são ocorrências não registadas.

Há portanto, um trabalho, leia-se, científico, a fazer, para reduzir esta margem de desconhecimento das ocorrências criminais. Há que deduzir da informação acessível, certas conclusões que permitam, pelo menos, ter uma ideia sobre a dimensão efectiva desta lacuna estatística. O trabalho científico é essencial como tentativa para reduzir a margem do desconhecido através duma utilização adequada dos factos já conhecidos.

3.2.2 Os perigos da interpretação Post Factum

Manheim, H. (1985:198-199), considera ainda que, estes perigos centram-se no facto de, os dados serem frequentemente recolhidos e só depois serem submetidos a tentativas de interpretação.

Merton, R. (1957), considera que a interpretação intervém só depois de as observações terem sido feitas, em vez do processo mais correcto que consiste na comprovação empírica duma hipótese pré-elaborada, não é por via de regra muito difícil produzir uma hipótese para suportar as observações.

Há contudo que ter presente o perigo inverso, ou seja, se partirmos já com uma hipótese pré-elaborada, isto pode condicionar e prejudicar as próprias observações.

3.3 Fontes de informação e pesquisas

Para a realização deste estudo, é imprescindível, obter dados de fonte segura, que sejam fidedignos e o mais actualizados possíveis, para isso, recorreu-se a utilização da base de dados de registo de ocorrências policiais da PSP.

Estes dados são o ponto de partida de toda a análise efectuada neste trabalho.

Como fontes primárias, foram encontrados alguns artigos na *internet* sobre o tema da investigação que foram averiguados quanto a proveniência e qualidade, antes de se optar por incluir esses mesmos artigos no presente trabalho,

Como fontes secundárias, foi efectuado um estudo exaustivo dos livros referenciados na bibliografia e posterior síntese e análise das partes relevantes que enriqueçam este estudo.

Foi ainda feita uma pesquisa por palavras-chave às bases de dados *online* disponibilizadas pelo ISEGI, nomeadamente: B-On e *Web of Knowledge*, não se encontrando artigos relevantes sobre o tema, o que nos levou a fazer uma procura mais abrangente pelo motor de busca *Google* e *Wikipedia*.³⁰

³⁰ Através da *internet*, é possível aceder aos sites de várias associações que trabalham neste tema, como por exemplo a ACPO (*Association of Chief Police Officer*) e IACA (*International Association of Crime Analyst's*).

4. PROCESSAMENTO DOS DADOS DISPONÍVEIS

*There are three kinds of lies: lies, damned lies, and statistics.*³¹

Os dados chegaram-nos em formato texto tabelado, de forma não organizada, o que nos levou imenso tempo para compreender a estrutura de cada registo, Para além disso, o ficheiro era demasiado grande para ser aberto no *Excel* e demorava muito tempo (quando não bloqueava o computador) se fosse aberto no *Access* ou mesmo no *SAS*. Após a redução do ficheiro inicial, o processamento foi mais fluido seguindo-se o procedimento de pré-processamento dos dados, da seguinte forma:

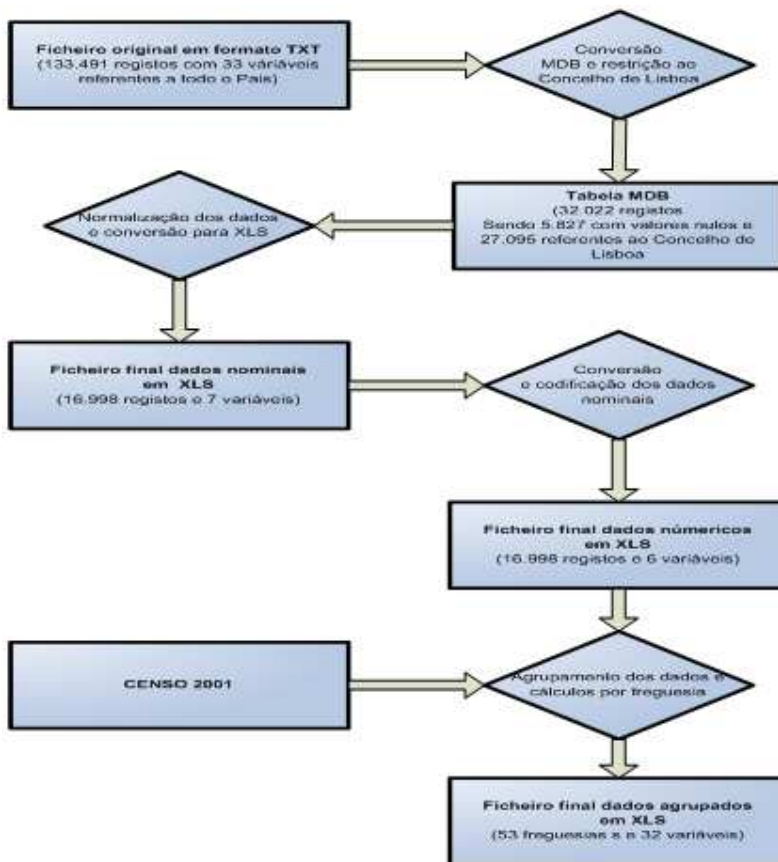


Figura 11 – Metodologia de pré-processamento dos dados

³¹ Disraeli, B.(1804-1881), “Existem três tipo de mentiras: “mentiras, malditas mentiras e estatísticas”.

4.2 Pré-processamento dos dados

Bação (2006:41), considera que: “o objectivo fundamental da fase de pré-processamento dos dados consiste em facilitar e simplificar o problema a tratar sem excluir ou danificar informação importante para a modelação e entendimento do problema.”, e acrescenta “...o pré-processamento permite, em muitas circunstâncias reduzir o ruído e amplificar o sinal.” e ainda “...o pré-processamento pode ajudar a reduzir o espaço de *input* melhorando a performance dos modelos.”

Do ficheiro inicial com de 133.491 registos e 33 variáveis, relativos a todos os distritos de Portugal continental, destes, foram seleccionadas as variáveis que interessam ao estudo, restringindo os registos ao Concelho e Lisboa e aos crimes patrimoniais e pessoais, ficamos com 16.998 registos e 6 variáveis cada: dia, mês, freguesia, tipo, subtipo e local.

processo	ano	dataocorrencia	tipooocorrencia	tipopessoa	subtipoocorrencia	datanascimento	sexo
17	2007	2007-01-01 00:00:00.000	Crimes	contra o património	Danos	1972-10-01 00:00:00.000	Masculino
35	2007	2007-01-01 00:00:00.000	Crimes	contra o património	Furto/Roubo/Receptação	1988-03-17 00:00:00.000	Masculino
36	2007	2007-01-01 00:00:00.000	Crimes	contra o património	Furto/Roubo/Receptação	1988-04-11 00:00:00.000	Masculino
36	2007	2007-01-01 00:00:00.000	Crimes	contra o património	Furto/Roubo/Receptação	1972-09-11 00:00:00.000	Masculino
37	2007	2007-01-01 00:00:00.000	Crimes	contra o património	Furto/Roubo/Receptação	1977-02-28 00:00:00.000	Masculino
40	2007	2007-01-01 00:00:00.000	Crimes	contra o património	Furto/Roubo/Receptação	1974-04-17 00:00:00.000	Masculino
62	2007	2007-01-01 00:00:00.000	Crimes	contra o património	Furto/Roubo/Receptação	1991-04-19 00:00:00.000	Feminino
68	2007	2007-01-01 00:00:00.000	Crimes	contra o património	Furto/Roubo/Receptação	1988-03-28 00:00:00.000	Masculino
98	2007	2007-01-01 00:00:00.000	Crimes	contra o património	Furto/Roubo/Receptação	1974-11-19 00:00:00.000	Feminino
102	2007	2007-01-01 00:00:00.000	Crimes	contra o património	Furto/Roubo/Receptação	1984-10-07 00:00:00.000	Masculino
111	2007	2007-01-01 00:00:00.000	Crimes	contra pessoas	Agressões	NULL	NULL
111	2007	2007-01-01 00:00:00.000	Crimes	contra pessoas	Agressões	NULL	Masculino
120	2007	2007-01-01 00:00:00.000	Crimes	contra pessoas	Agressões	NULL	Masculino
120	2007	2007-01-01 00:00:00.000	Crimes	contra pessoas	Agressões	1988-03-14 00:00:00.000	Feminino

Tabela 4 – Exemplo dos dados originais

Para Bação (2006:42-43), criar variáveis baseadas em rácios é uma das formas mais efectivas de proceder para combinar duas ou mais variáveis num único *input* e assim reduzir o espaço de *input*. Este procedimento foi adoptado na criação de novas variáveis baseadas em rácios que agregam os dados originais com os dados do Censo 2001.

Todas as restantes variáveis que não puderam ser aproveitadas por falta de qualidade, foram ignoradas, assim como as variáveis que apresentavam percentagens elevadas de dados omissos ou incorrectamente inseridos.

Verificou-se que existiam muitas variáveis que não interessavam para este estudo, tais como os dados referentes ao ofendido e infractor, apenas foram aproveitados os dados que de algum modo se relacionam com o tipo de crime e local onde foi cometido, para além da data e hora do mesmo.

Constatou-se ainda uma redundância excessiva dos registos, ou seja, por cada ocorrência, fosse ela qual fosse, havia sempre: um registo para a vítima, um registo para o autor e um registo para cada testemunha, o que tornava o registo das ocorrências redundante com variáveis duplicadas.

Ou seja, uma única ocorrência criminal tinha tantos registos como pessoas intervenientes no processo o que aumentava exponencialmente o tamanho da base de dados e tempo requerido para o seu tratamento.

Todas as ocorrências registadas no posto de turismo da PSP, não têm directamente relação com o local da ocorrência do crime, por isso foi criada uma categoria de “outros” para estes casos.

As ocorrências que aconteceram num determinado local mas são denunciadas e registadas noutra local, considera-se para efeitos estatísticos a freguesia a que pertence o posto policial de registo.

DATA	TIPO	SUBTIPO	FREGUESIA	LOCAL	DESCRICAO	REGISTO
2007-01-05	Património	Furto	Ajuda	Habitação	Rua das Açucenas	Telheiras
2007-01-06	Pessoas	Agressões	Ajuda	Associação desportiva	Rua do Jardim Botânico	Calvario
2007-01-08	Pessoas	Agressões	Ajuda	Habitação	Rua Fonseca Benevides	Praca Alegria
2007-01-09	Património	Furto	Ajuda	Habitação	Rua do Sítio do Casalinho da Ajuda	Telheiras
2007-01-12	Património	Danos	Ajuda	Via pública	Rua da Paz	Campo Ourique
2007-01-16	Pessoas	Ameaça	Ajuda	Via pública	Travessa D. Vasco	Telheiras
2007-01-20	Pessoas	Ameaça	Ajuda	Habitação	Rua 13	Moscavide
2007-01-22	Património	Furto	Ajuda	Habitação	Rua do Meio à Ajuda	Musqueira
2007-01-30	Património	Furto	Ajuda	Habitação	Travessa da Memória	Rego
2007-02-01	Património	Furto	Ajuda	Habitação	Rua Eduardo Bairrada	Telheiras
2007-02-02	Património	Furto	Ajuda	Via pública	Rua João Castilho	Rato
2007-02-03	Património	Furto	Ajuda	Via pública	Calçada da Memória	Camide
2007-02-03	Património	Furto	Ajuda	Via pública	Calçada da Memória	Rego
2007-02-05	Pessoas	Agressões	Ajuda	Posto combustível	Estrada de Caselas	Horta Nova
2007-02-06	Património	Furto	Ajuda	Escola	Rua da Junqueira	Benfica
2007-02-07	Pessoas	Ameaça	Ajuda	Habitação	Rua Raul Proença	Olivais
2007-02-08	Pessoas	Agressões	Ajuda	Via pública		Rossio
2007-02-09	Património	Furto	Ajuda	Escola		Benfica
2007-02-09	Património	Furto	Ajuda	Restauração	Largo da Ajuda	Camide
2007-02-09	Pessoas	Agressões	Ajuda	Habitação	Rua Armando Lucena	Lapa
2007-02-12	Património	Furto	Ajuda	Via pública	Calçada da Ajuda	Benfica
2007-02-13	Património	Danos	Ajuda	Via pública	Bairro Calçada dos Mestres	Aeroporto
2007-02-14	Património	Danos	Ajuda	Via pública	Bairro Calçada dos Mestres	Aeroporto
2007-02-14	Património	Furto	Ajuda	Habitação	Travessa do Giestal	Mourana
2007-02-16	Património	Furto	Ajuda	Habitação	Rua Jorge Álvares	Rossio
2007-02-17	Património	Furto	Ajuda	Restauração	Travessa D. Vasco	Benfica
2007-02-18	Património	Furto	Ajuda	Via pública	Travessa João Alves	Benfica
2007-02-20	Património	Furto	Ajuda	Via pública	Bairro Calçada dos Mestres	Campo Grande
2007-02-21	Património	Furto	Ajuda	Via pública	Calçada da Ajuda	Rego
2007-02-22	Património	Danos	Ajuda	Via pública	Rua Alfredo da Silva	Campo Ourique
2007-02-22	Património	Furto	Ajuda	Habitação	Rua das Açucenas	Telheiras
2007-02-23	Património	Furto	Ajuda	Estabelecimento comercial	Travessa da Boa Hora	Arroios
2007-02-24	Património	Furto	Ajuda	Via pública	Calçada da Ajuda	Benfica
2007-02-24	Património	Furto	Ajuda	Via pública	Rua Nossa Senhora Ajuda	Caminhos Ferro

Tabela 5 – Exemplo dos dados nominais finais

4.3 Descrição das variáveis

Com exceção da data, todas as restantes variáveis encontravam-se em formato nominal, houve por isso necessidade de transformar e codificar as variáveis mais significativas para o estudo.

Nome	Amplitude	Descrição
freguesia	1 a 53	ID da freguesia
dia	1 a 31	Dia da ocorrência
mes	1 a 7	Mês da ocorrência
freguesia	1 a 53	Identifica a freguesia
tipo	1 e 2	Tipo de crime
subtipo	1 a 7	Subtipo agregado de crime
local	1 a 25	Local agregado da ocorrência

Tabela 6 – Descrição das variáveis

Após selecção e validação das variáveis, retemos as que se afiguram como estatisticamente relevantes no presente estudo, de seguida, as variáveis foram codificadas para o formato numérico e agrupadas numa tabela.³²

As tabelas seguintes, representam as codificações utilizadas para codificar e agrupar as variáveis:³³

TIPO	
1	Patrimoniais
2	Pessoais

Tabela 7 - Codificação da variável tipo

SUBTIPO AGREGADO	
1	Burla
2	Danos
3	Furto
4	Agressões
5	Ameaça
6	Injúrias
7	Outro

Tabela 8 – Codificação da variável subtipo

³² A tabela em referência encontra-se no apêndice 3.

³³ Para as variáveis, dia e mês, não houve necessidade de codificar os valores por estes já se encontrarem em formato numérico.

LOCAL AGREGADO	
1	Aeroporto
2	Agencia Bancária
3	Centro Comercial
4	Embaixada
5	Escola
6	Esquadra Policia
7	Estabelecimento Comercial
8	Estação Correios
7	Farmácia
10	Garagem
11	Habitação
12	Hospital
13	Hotel
14	Igreja
15	Museu
16	Organismo Público
17	Ouriversaria
18	Parque Estacionamento
17	Posto Combustivel
20	Recinto Desportivo
21	Recinto Espectaculos
22	Restauração
23	Transportes Publicos
24	Via publica
25	Outro

Tabela 9 – Codificação da variável local

FREGUESIA	
1	Ajuda
2	Alcantara
3	Alto Pina
4	Alvalade
5	Ameixoeira
6	Anjos
7	Beato
8	Benfica
9	Campo Grande
10	Campolide
11	Carnide
12	Castelo
13	Charneca
14	Coracao Jesus
15	Encarnacao
16	Graca
17	Lapa
18	Lumiar
17	Madalena
20	Martires
21	Marvila
22	Merces
23	Nossa Senhora Fatima
24	Pena
25	Penha Franca
26	Prazeres
27	Sacramento
28	Santa Catarina
27	Santa Engracia
30	Santa Isabel
31	Santa Justa
32	Santa Maria Belem
33	Santa Maria Olivais
34	Santiago
35	Santo Condestavel
36	Santo Estevao
37	Santos-o-Velho
38	Sao Cristovao Sao Lourenco
37	Sao Domingos Benfica
40	Sao Francisco Xavier
41	Sao Joao
42	Sao Joao Brito
43	Sao Joao Deus
44	Sao Jorge Arroios
45	Sao Jose
46	Sao Mamede
47	Sao Miguel
48	Sao Nicolau
47	Sao Paulo
50	Sao Sebastiao Pedreira
51	Sao Vicente Fora
52	Se
53	Socorro
54	Outro

Tabela 10 – Codificação da variável freguesia

Em apêndice encontra-se a tabela final dos dados já processados e agrupados, a matriz de correlação de *Pearson*, os histogramas e *box-plot* das variáveis originais e calculadas.

DIA	MÊS	FREGUESIA	TIPO	SUBTIPO AGREGADO	LOCAL AGREGADO
1	1	1	1	3	11
1	1	1	2	4	20
1	1	1	2	4	11
1	1	1	1	3	11
1	1	1	1	2	24
1	1	1	2	5	24
1	1	1	2	5	11
1	1	1	1	3	11
1	1	1	1	3	11
1	2	1	1	3	11
1	2	1	1	3	24
1	2	1	1	3	24
1	2	1	1	3	24
1	2	1	2	4	17
1	2	1	1	3	5
1	2	1	2	5	11
1	2	1	2	4	24
1	2	1	1	3	5
1	2	1	2	4	11
1	2	1	2	4	24
1	2	1	1	3	22
1	2	1	1	3	24
1	2	1	1	2	24
1	2	1	1	3	11
1	2	1	1	2	24
1	2	1	1	3	11
1	2	1	1	3	22
1	2	1	1	3	24
1	2	1	1	3	24
1	2	1	1	3	24
1	2	1	1	3	24
1	2	1	1	3	24
1	2	1	1	3	11
1	2	1	1	2	24
1	2	1	1	3	7
1	2	1	1	3	24
1	2	1	1	3	24
1	2	1	1	3	7
1	3	1	1	3	5
1	3	1	1	2	24
1	3	1	1	3	24
1	3	1	1	3	24
1	3	1	1	3	24

Tabela 11 – Exemplo dos dados numéricos finais

4.4 Estatística descritiva dos dados

Tendo em conta os dados descritos, e depois de agregados, chegamos á seguinte análise descritiva:

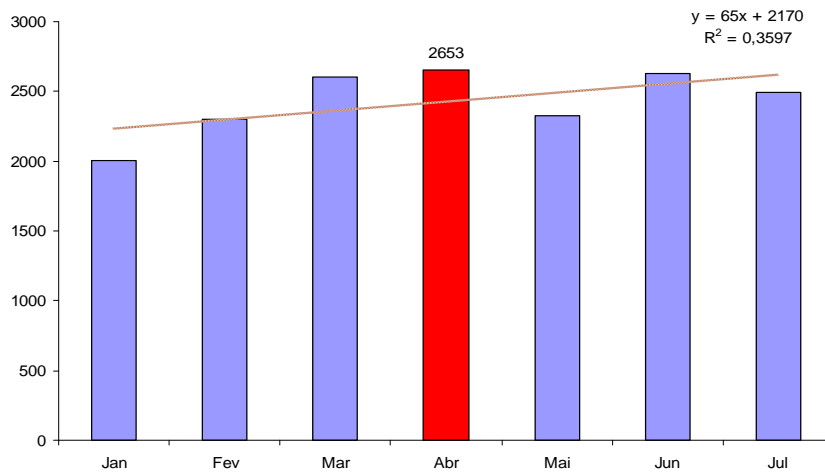


Figura 12 - Ocorrências criminais mensais

É notório que durante o período em análise, houve um aumento gradual da criminalidade, destacando-se os meses de Março, Abril e Junho em que houve um número de ocorrências acima da média.

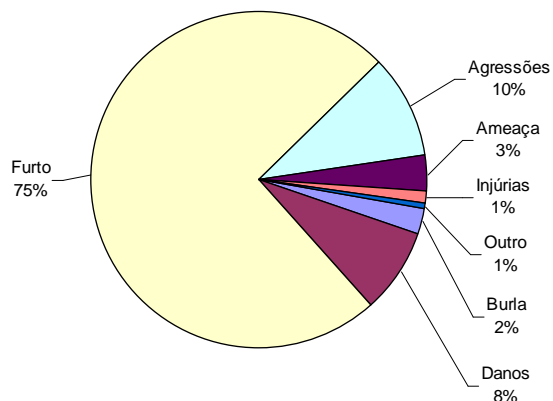


Figura 13 - Análise global da criminalidade

Os crimes de furto, agressões e danos, foram os mais cometidos, representando cerca de 94 % de todas as ocorrências registadas.

Os crimes patrimoniais representam 84% de todos os registos, sendo os restantes, crimes pessoais. De todos os crimes patrimoniais, destaca-se o crime de furto com 75% do total.

Acrescenta-se que, o crime de furto referenciado, já inclui o crime de roubo, este, destaca-se do primeiro por envolver contacto físico com a vítima.

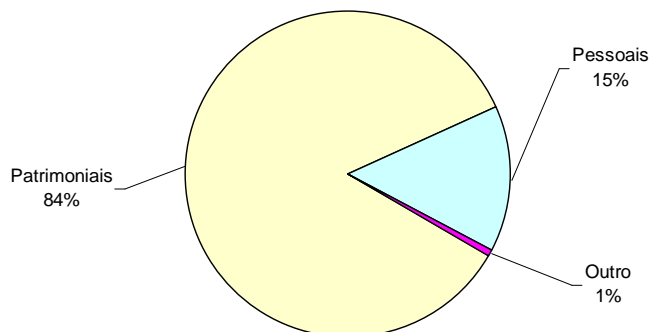


Figura 14 - Tipo de ocorrências

Quanto á totalidade dos crimes pessoais, destaca-se o crime de Agressões com 73,81% seguido do crime de Ameaças com 24,08%

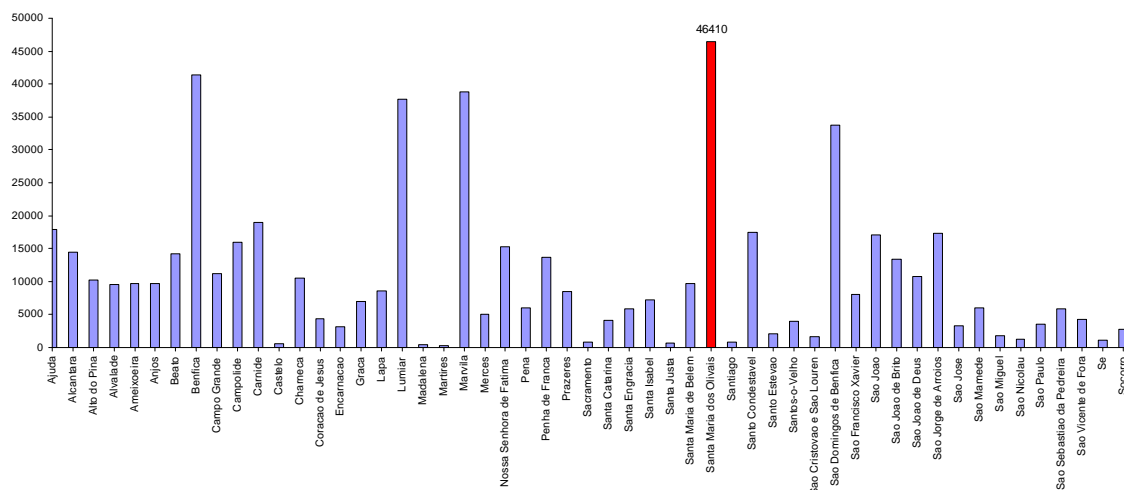


Figura 15 - População por freguesias

Santa Maria dos Olivais é a freguesia mais populosa com 46.410 habitantes, logo seguida por Benfica, Marvila, Lumiar e São Domingos de Benfica, todas as restantes estão abaixo dos 15.000 habitantes.

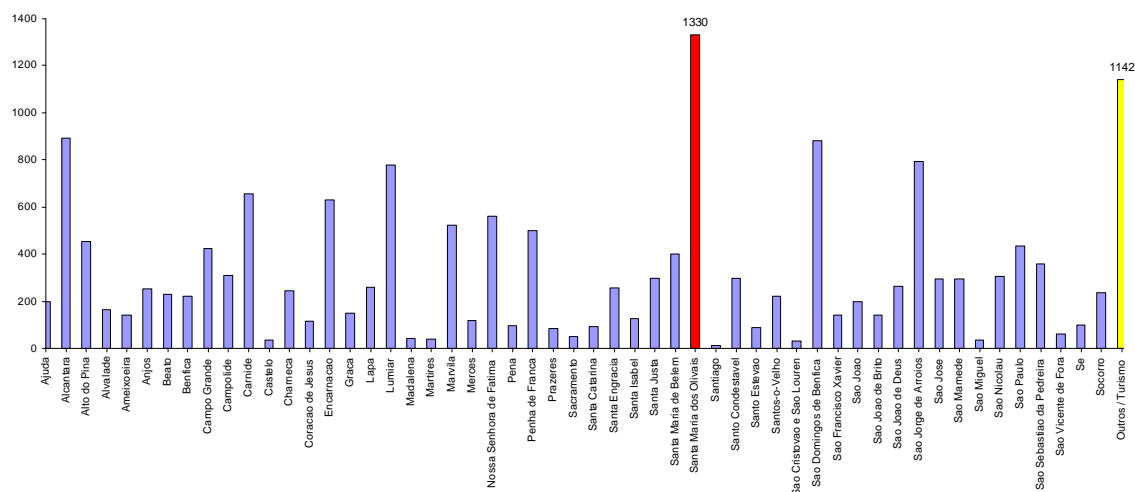


Figura 16 - Ocorrências por freguesia

Quanto ao número de ocorrências, é liderado também pela freguesia de Santa Maria dos Olivais com 1.330 ocorrências criminais no período em análise, seguindo-se Alcântara e São Domingos de Benfica. Note-se que se registaram 1.142 ocorrências no posto de turismo da PSP, o que coloca este posto com o segundo maior nível de ocorrências registadas em Lisboa.

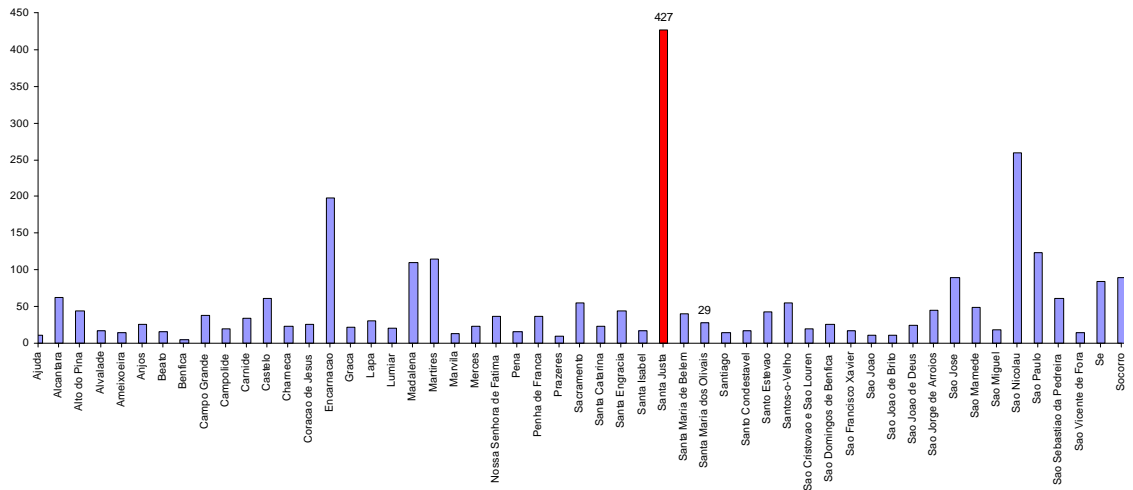


Figura 17 – Número de ocorrências por cada 1000 habitantes

Quanto ao rácio entre a população e o número de ocorrências verifica-se que é Santa Justa que obtém, o maior número de ocorrências por cada 1000 habitante, com 427 ocorrências, seguindo-se São Nicolau e Encarnação.

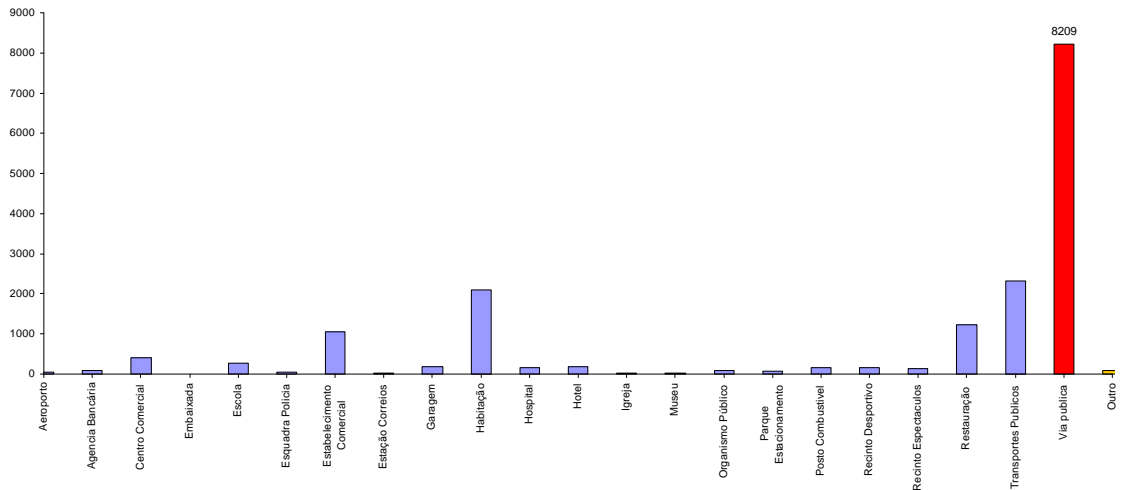


Figura 18 - Local da ocorrência

Quanto ao local onde se deu a ocorrência criminal, destacam-se a via pública com 8.209 ocorrências, seguindo-se os transportes públicos e a habitação (inclui *hall* de entrada e arrecadação), os estabelecimentos de restauração e estabelecimentos comerciais, também são representativos.

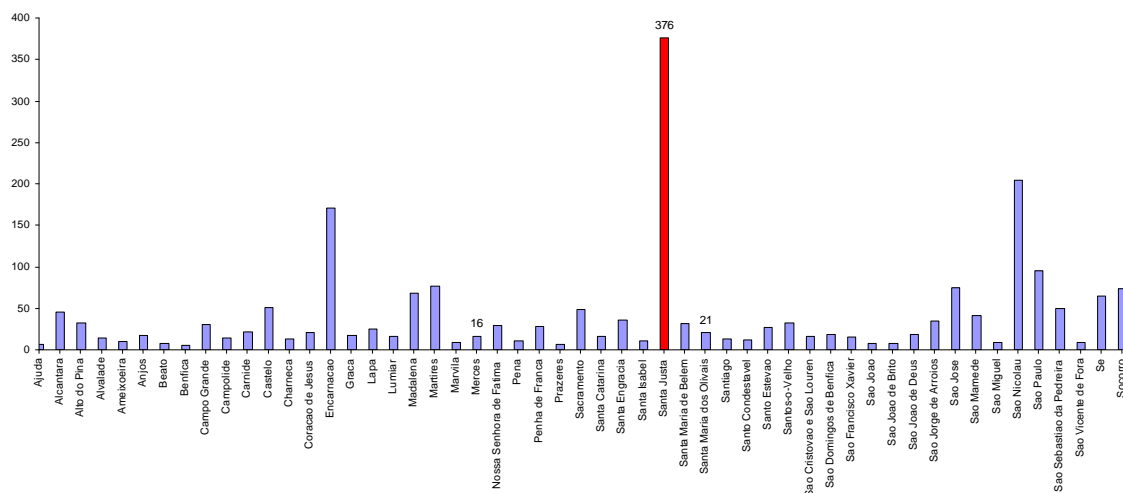


Figura 19 - Crimes de furto / roubo por cada 1000 habitantes

Tendo em conta que os crimes de furto e/ou roubo são os que mais se destacam com 75% de todas as ocorrências criminais, convém analisar estes por freguesia, assim, Santa Justa toma a liderança com 376 ocorrências por cada 1000 habitantes, seguindo-se São Nicolau e Encarnação.

Este número de ocorrências faz com que a freguesia de Santa Justa tenha o maior rácio de crimes por habitante e que, (provavelmente) levou a que fossem tomadas medidas de combate à criminalidade, como a recente autorização Comissão Nacional de Protecção de Dados, para a instalação de 27 câmeras de videovigilância no Bairro Alto, no entanto e de acordo artigo publicado no Jornal Destak de 27 de Novembro de 2009 página 4, este sistema não poderá captar ou gravar som e está limitado ao horário das 22h até 7h, para não ter de “sacrificar os direitos dos cidadãos”.

5. MODELO PREDITIVO

O objectivo fulcral da análise preditiva não é o de encontrar um modelo que se adapte bem aos dados (a preocupação usual), mas sim a construção de um modelo probabilístico adequado aos dados que permita prever com o máximo de probabilidade o comportamento aleatório de observações futuras, com base na informação do passado.

Em Ciência, existem três tipos de modelos: modelos determinísticos, que são baseados no conhecimento; modelos paramétricos, baseados em pressupostos e os modelos não-paramétricos que são baseados nos dados.

Bação (2006:47), descreve o processo de desenvolvimento de modelos genéricos não-paramétricos da seguinte forma:

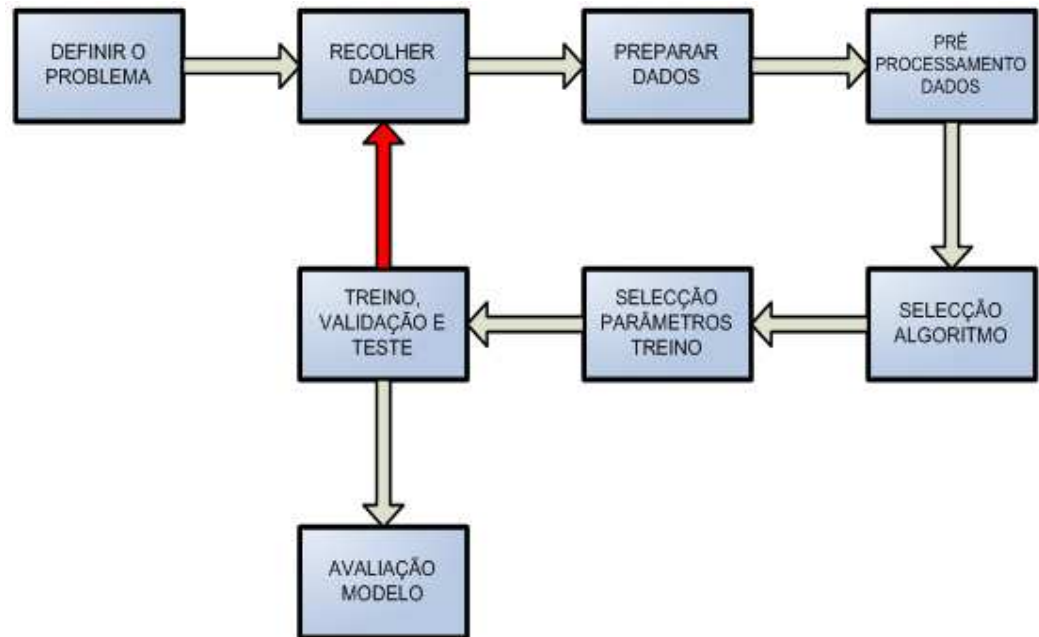


Figura 20 - Modelo preditivo genérico

Foram os modelos não-paramétricos que deram origem as actuais técnicas de *data mining*. Utilizam normalmente, grandes quantidades de dados que têm por premissa, o facto de as relações que ocorrem frequentemente de forma agrupada, tendem a repetir-se no futuro, desde que haja um conhecimento mínimo do fenómeno e possibilidade de ajustar o modelo para o objectivo a atingir.³⁴

5.1 Selecção e desenvolvimento do modelo teórico

A estratégia de investigação baseia-se num modelo preditivo não paramétrico, para extracção de conhecimento dos dados iniciais, o modelo tem um ciclo redundante que possibilita o ajuste fino do modelo aos dados.

Embora existam outros modelos de extracção do conhecimento (e. g. *Clementine* do SPSS), iremos utilizar o modelo proposto pelo SAS que obedece ao Modelo SEMMA como método de extrair conhecimento dos dados esparsos.

O acrónimo SEMMA - *Sample, Explore, Modify, Model, Assess*, refere-se ao processo base de *data mining* utilizado pelo SAS, inicia-se com uma representação estatística da amostra dos dados.

O SEMMA, propicia a aplicação da análise exploratória dos dados e técnicas de visualização, selecção e transformação das variáveis mais significativas, permite modelar as variáveis de modo a prever resultados e confirmar a precisão do modelo utilizado.³⁵

³⁴ Quando temos uma variável *Target*, o modelo designa-se de supervisionado.

³⁵ in <http://www.sas.com>

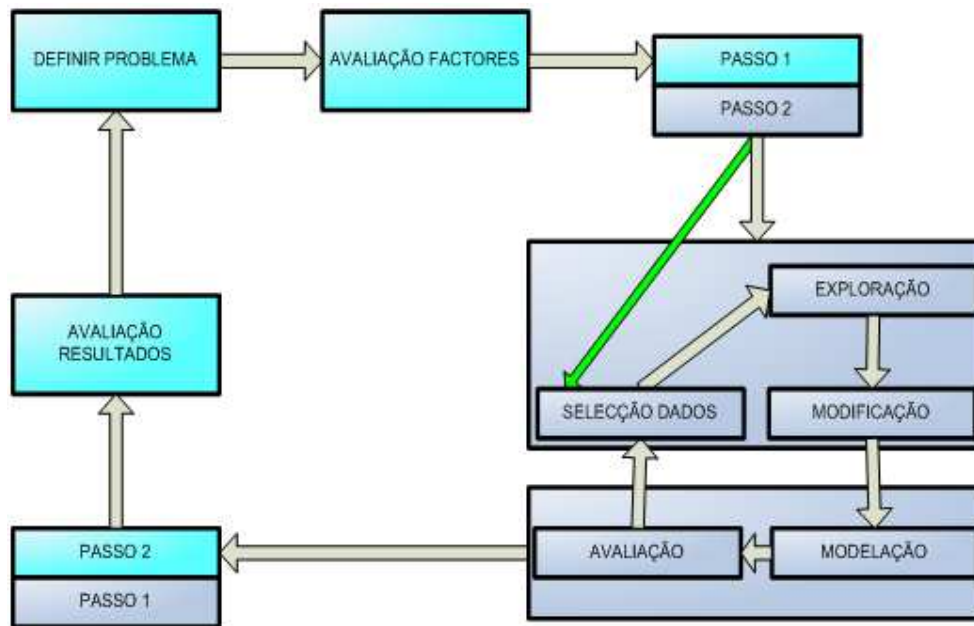


Figura 21 - Metodologia SEMMA

- **Sample** (amostra) - amostra dos dados representativa da população, é normalmente particionada em: treino, validação e teste;
- **Explore** (exploração) - ajuda a redefinir todo o processo de descoberta de conhecimento pela procura de tendências e anomalias nos dados através de técnicas estatísticas é o pré-processamento dos dados;
- **Modify** (modificar) - permite seleccionar e transformar as variáveis tendo em vista o tipo de modelo utilizado, baseia-se na fase exploratória para manipular a informação. (e. g. converter variáveis nominais em numéricas);
- **Model** (modelo) - através de modelos de *data mining*, procura combinações na informação que melhor preveja o resultado esperado com o modelo;
- **Assess** (avaliação) - avalia os resultados obtidos através da medição da *performance* do processo de *data mining* permitindo otimizar os resultados pelo ajuste do modelo.

O modelo CLAP (*Cluster Analysis Prediction*) proposto por Jonathan (2003), tem a vantagem de agrupar as ocorrências em *clusters* e de aplicar um modelo preditivo específico para cada tipo de *cluster*, otimizando o resultado final e minimizando o erro gerado através da aplicação de um teste *Gamma*.

Para Jonathan (2003), este tipo de abordagem faz todo o sentido visto que, cada tipo de crime tem a sua especificidade e tem de ser analisado de forma específica e é especialmente robusto na previsão a curto prazo desde que se tenha uma quantidade razoável de dados, esta “quantidade razoável” é determinada pelo “teste M” que nos indica quais o número mínimo de dados para definir a função com um erro mínimo, ignorando os restantes.

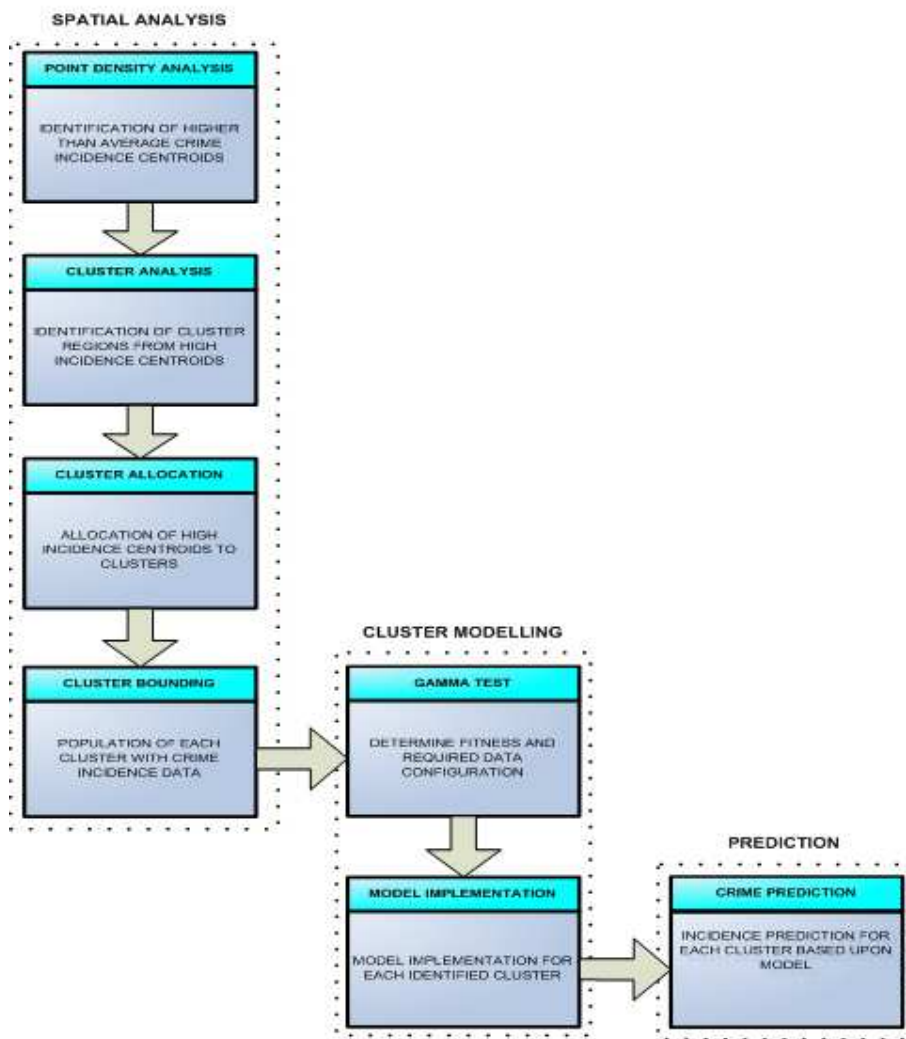


Figura 22 - Modelo CLAP

O teste *Gamma*, estima o melhor erro quadrático médio (MSE) que pode ser conseguido no tratamento da informação usando um modelo contínuo como os regressores de mínimos quadrados ou uma rede neuronal, para uma função desconhecida.³⁶

Se seguida descreve-se cada etapa deste processo:

- ***Spatial Analysis***, análise espacial – identifica os clusters geográficos com base em centroídes;
- ***Cluster Modelling***, modelação do cluster – determina a qualidade da informação de cada cluster associando os dados ao centroíde desse cluster;
- ***Prediction***, previsão – desenvolve o modelo preditivo baseado numa correspondente rede neuronal (uma por cada cluster) com especificações de previsão autoregressiva. (o *lag* de cada rede neuronal é determinado pelo teste *Gamma* de modo a minimizar o erro).

Embora o modelo CLAP seja adequado, o mesmo não é possível de implementar no corrente estudo por falta de qualidade dos dados e ausência de variáveis geográficas, optando-se pelo modelo SEMMA do SAS.

O SAS permite-nos criar um modelo preditivo no *Enterprise Miner* ou seja, um modelo capaz de fazer o ajustamento dos dados de acordo com vários métodos de regressão ou classificação simultaneamente, o modelo com melhor performance em termos de ajuste do erro quadrático médio, é seleccionado pelo *score node* para fazer a previsão dos dados.

De seguida, Iremos descrever o modelo preditivo supervisionado adoptado e quais os parâmetros de cada método.

³⁶ MSE – *Mean Square Error*

5.2 Modelo Enterprise Miner do SAS

A seguinte figura, representa o modelo utilizado no presente estudo, com todos os nós (*nodes*) do SAS *enterprise miner*.

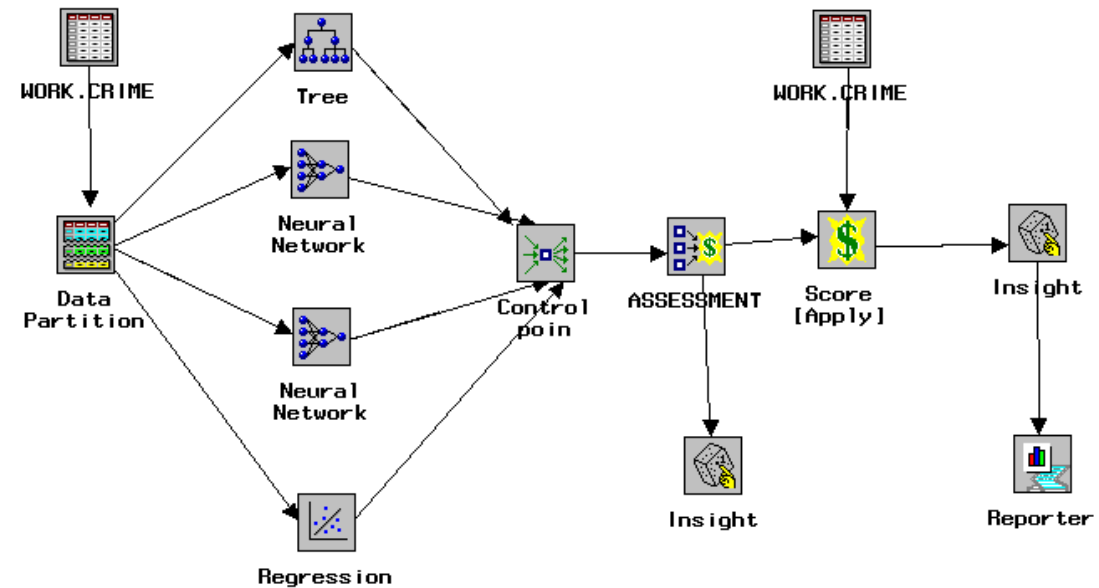


Figura 23 – Modelo preditivo no SAS *Enterprise Miner*

A tabela de *INPUT* do modelo foi alterada para *RAW* para que o modelo possa efectuar uma partição dos dados inseridos.

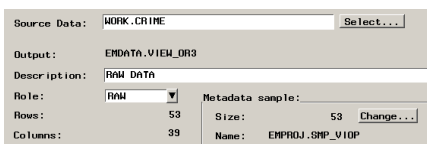


Figura 24 – Características do conjunto de *input*

Para Bação (2006:41-42), “todos os métodos de descoberta de conhecimento sofrem da tendência de se adaptar demasiado ao conjunto de dados que é utilizado para treinar [sobreaprendizagem]. Só com os dados de treino não existe forma de saber se as relações encontradas podem generalizar-se a toda a população em estudo ou se só ocorrem no conjunto de treino e por isso não têm valor para generalizar à população.”

Para evitar a sobreaprendizagem do modelo, as 53 observações referentes às freguesias de Lisboa foram divididas pelo *partition node* da seguinte forma: 70% para o conjunto de treino e 30% para o conjunto de validação. A grande percentagem para os dados de treino, tem por fim evitar a sobreaprendizagem do modelo.

Não se optou por criar um conjunto de teste porque não se pretende aferir a qualidade da performance na generalização do modelo, a semente – *Seed* – foi de 100.

The image shows a software interface for data partitioning. It consists of two main panels. The left panel, titled 'Method:', contains three radio button options: 'Simple Random' (which is selected), 'Stratified', and 'User Defined'. Below this is a 'Random Seed:' section with a 'Generate New Seed' button and a text input field containing the number '100'. The right panel, titled 'Percentages:', shows a distribution: 'Train:' at 70%, 'Validation:' at 30%, 'Test:' at 0%, and 'Total:' at 100%.

Figura 25 – Descrição da partição dos dados

As variáveis de *input* do modelo, deverão estar causalmente relacionadas com o output para que a modelação tenha sucesso e os resultados obtidos sejam válidos.

Após a análise da correlação entre as diversas variáveis, verificaram-se algumas correlações espúrias com a variável *target* que podem degradar o modelo.

Tais correlações espúrias, devem-se, provavelmente, ao facto de termos poucas observações e ao relativamente elevado número de variáveis de *input*.

A fim de minimizar este “ruído” no modelo, algumas variáveis de *input* correlacionadas com a variável *target* (nomeadamente as referentes aos totais por tipo de crime e totais mensais) foram transformadas para percentagem, a fim de não influenciar os resultados obtidos, pois poderiam facilmente introduzir erros aleatórios pela possibilidade matemática de calcular a variável *target*. (e. g. variável Fur - furto - foi convertida para PFur que corresponde à percentagem que o valor do crime de furto representa face aos outros subtipos de crimes).

Não houve necessidade de normalizar as variáveis de *input*, por estas não apresentarem valores muito elevados e encontrarem-se na mesma escala.

Em termos de funções ou papel - *role* - desempenhado das variáveis no modelo a variável *target* corresponde ao total de ocorrências do mês de Julho, usando-se todas as outras variáveis e as novas variáveis (excepto FID e variáveis originais que necessitaram de transformação) como *input*.

A tabela 13 mostra a estatística descritiva ou exploratória das variáveis.

Name	Model	Role	Measurement	Type	Name	Min	Max	Mean	Std Dev.
FREGUESIA	id		nominal	char	ABR	1	229	47.189	46.333
AGEK	input		interval	num	AGEK	2	287	45.528	52.714
AGRMH	input		interval	num	AGR	1	159	32.472	32.149
FURK	input		interval	num	AGRMH	0	31	4.8113	5.6808
FURMH	input		interval	num	AME	0	59	10.302	11.328
HABEDI	input		interval	num	BUR	0	24	4.2075	5.0853
HABK	input		interval	num	DAN	1	127	24.472	25.76
MOCOM	input		interval	num	FEV	3	172	37.434	33.223
OCOK	input		interval	num	FID	1	53	27	15.443
OCOMH	input		interval	num	FUR	11	972	222.51	201.78
PABR	input		interval	num	FURK	25	3627	370.81	554.49
PAGR	input		interval	num	FURMH	5	376	40.057	60.154
PAME	input		interval	num	HABEDI	4	22	9.3396	4.274
PBUR	input		interval	num	HABK	2878	29617	10619	6262.2
PDAN	input		interval	num	INJ	0	17	3.6792	3.8719
PEDI	input		interval	num	JAN	2	213	37.868	37.214
PFEV	input		interval	num	JUL	2	197	43.849	40.846
PFUR	input		interval	num	JUN	2	188	46.547	41.588
PINJ	input		interval	num	MAI	0	178	41.226	38.887
PJAN	input		interval	num	MAR	2	183	45.057	38.988
PJUL	input		interval	num	MOCOM	2	190	42.717	38.246
PJUN	input		interval	num	OCOK	28	4193	476.85	643.92
PMAI	input		interval	num	OCOMH	5	427	51.717	70.155
PMAR	input		interval	num	OUT	0	10	1.5283	2.1177
POCO	input		interval	num	PABR	476	2937	1518.6	421.8
POUT	input		interval	num	PAGR	208	3030	1116.4	629.08
PPOP	input		interval	num	PAME	0	935	330.74	205.02
TARE	input		interval	num	PBUR	0	428	139.89	113.75
TEDI	input		interval	num	PDAN	135	1951	905.92	456
TOCO	input		interval	num	PEDI	0	6	1.8679	1.5571
TPOP	input		interval	num	PFEV	588	3636	1387.4	490.67
ABR	rejected		interval	num	PFUR	4545	9099	7330.1	1047.6
AGR	rejected		interval	num	PINJ	0	606	131.34	122.75
AME	rejected		interval	num	PJAN	457	2340	1273.3	377.23
BUR	rejected		interval	num	PJUL	556	3333	1458.3	483.17
DAN	rejected		interval	num	PJUN	213	3333	1550.7	507.74
FEV	rejected		interval	num	PMAI	0	2431	1263.3	449.27
FID	rejected		interval	num	PMAR	625	2500	1548.4	409.94
FUR	rejected		interval	num	POCO	0	8	1.8113	1.6179
INJ	rejected		interval	num	POUT	0	235	45.679	59.175
JAN	rejected		interval	num	PPOP	0	8	1.9057	1.9539
JUN	rejected		interval	num	TARE	0	11	1.5283	2.1979
MAI	rejected		interval	num	TEDI	60	3090	1024.3	792.31
MAR	rejected		interval	num	TOCO	13	1330	299.17	267.29
OUT	rejected		interval	num	TPOP	341	46410	10654	10919
JUL	target		interval	num					

Tabela 13 – Estatística descritiva das variáveis

Tabela 12 – Descrição do *role* das variáveis

5.2.1 Árvore de decisão

Foi utilizada uma árvore de decisão com as seguintes características: critério de divisão baseado na redução da variância e avaliação do modelo com base no erro quadrático médio.

SOURCE	STATISTIC	VALUE
TRAIN	N	37
TRAIN	AVERAGE	37.891891892
TRAIN	AVE SQ ERR	424.83449163
TRAIN	R SQUARED	0.7337208406
VALID	N	16
VALID	AVERAGE	57.625
VALID	AVE SQ ERR	321.55446854
VALID	R SQUARED	0.8262224621

Tabela 14 – Erro do treino e validação da árvore de decisão

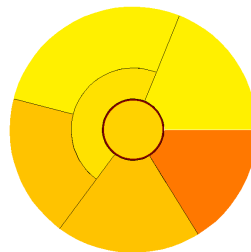


Figura 26 – Anel da árvore de decisão

Leaves	Training	Validation
5	424.8345	321.5545
6	420.7237	323.8701

Tabela 15 – Erro quadrático médio da árvore de decisão

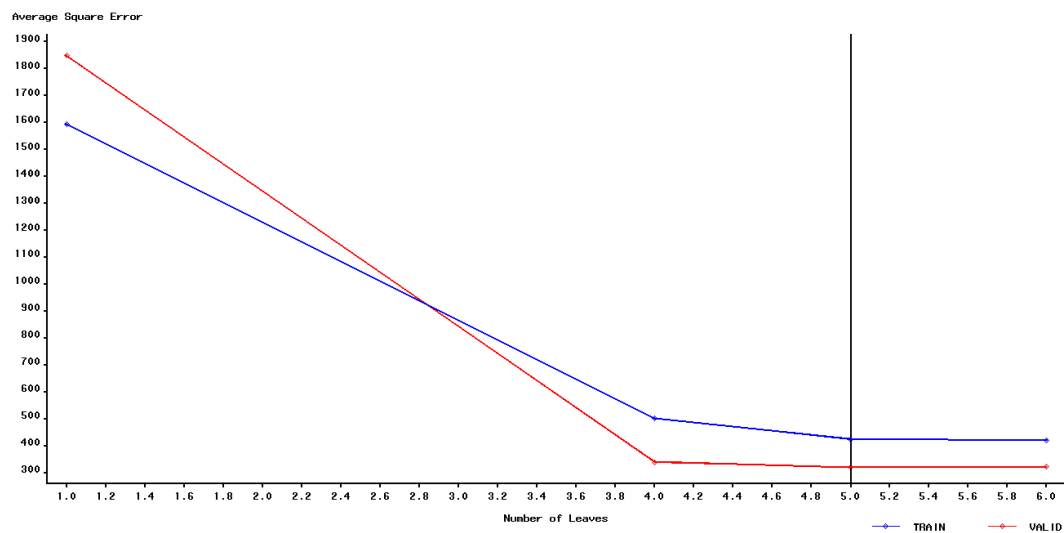


Figura 27 – Gráfico de aprendizagem da árvore de decisão

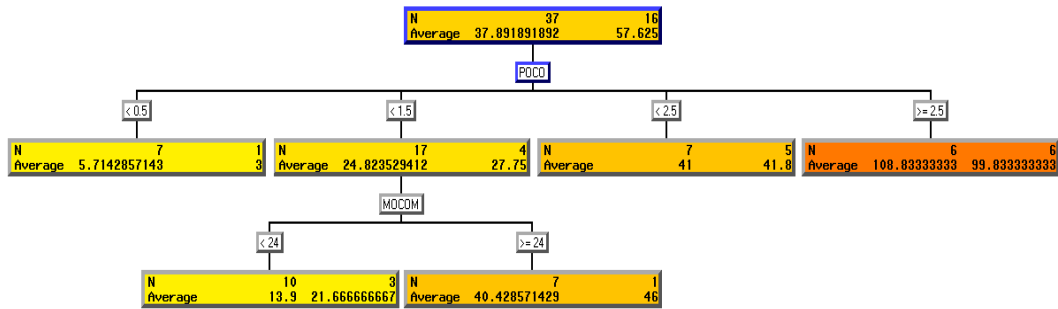


Figura 28 – Organograma da árvore de decisão

5.2.2 Descrição da rede neuronal 1

Para a rede neuronal 1, foi seleccionado um perceptrão multicamada, com uma camada escondida e o critério de selecção é o erro médio.

Fit Statistic	Training	Validation
[TARGET=Jul]	.	.
Average Error	1.5216055966	216.11855479
Average Squared Error	1.5216055966	216.11855479
Sum of Squared Errors	56.299407073	3457.8968767
Root Average Squared Error	1.2335337841	14.700971219
Root Final Prediction Error	5.1602420428	.
Root Mean Squared Error	3.751646541	14.700971219
Error Function	56.299407073	3457.8968767
Mean Squared Error	14.074851768	216.11855479
Maximum Absolute Error	2.8650066151	45.005389984
Final Prediction Error	26.62809794	.
Divisor for ASE	37	16
Model Degrees of Freedom	33	.
Degrees of Freedom for Error	4	.
Total Degrees of Freedom	37	.
Sum of Frequencies	37	16
Sum Case Weights * Frequencies	37	16
Akaike's Information Criterion	81.531345363	.
Schwarz's Bayesian Criterion	134.69163648	.

Tabela 16 – Estatística da rede neuronal 1

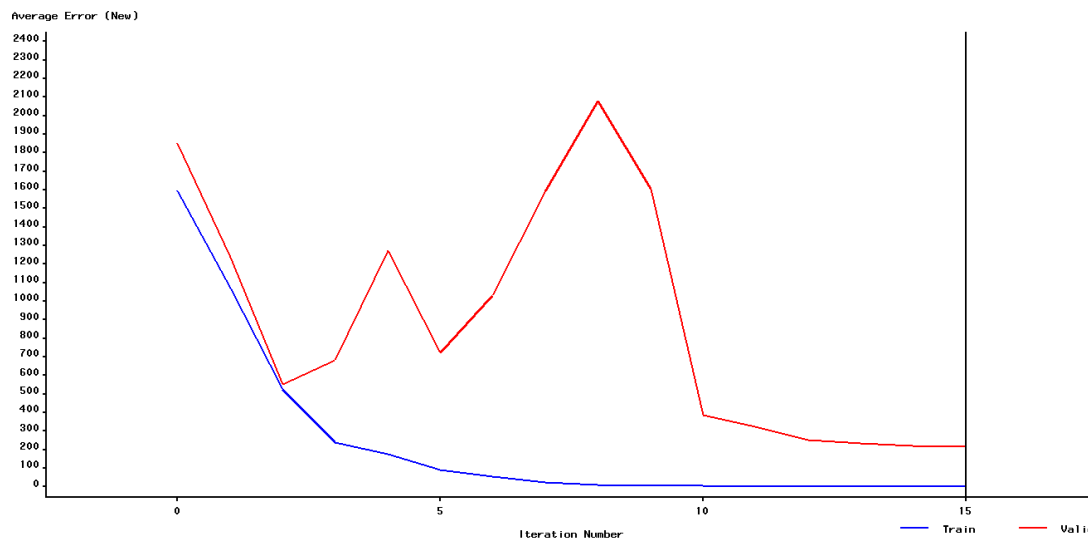


Figura 29 – Gráfico de aprendizagem da rede neuronal 1

5.2.3 Descrição da rede neuronal 2

Para a rede neuronal 2, foi seleccionado um perceptrão multicamada, com duas camadas escondidas e o critério de selecção é o erro médio.

Fit Statistic	Training	Validation
[TARGET=Jul]	.	.
Average Error	341.1654141	445.94356229
Average Squared Error	341.1654141	445.94356229
Sum of Squared Errors	12623.120322	7135.0969967
Root Average Squared Error	18.470663607	21.117375838
Root Mean Squared Error	.	21.117375838
Error Function	12623.120322	7135.0969967
Mean Squared Error	.	445.94356229
Maximum Absolute Error	98.75568261	39.820153358
Divisor for ASE	37	16
Model Degrees of Freedom	65	.
Degrees of Freedom for Error	-28	.
Total Degrees of Freedom	37	.
Sum of Frequencies	37	16
Sum Case Weights * Frequencies	37	16

Tabela 17 – Estatística da rede neuronal 2

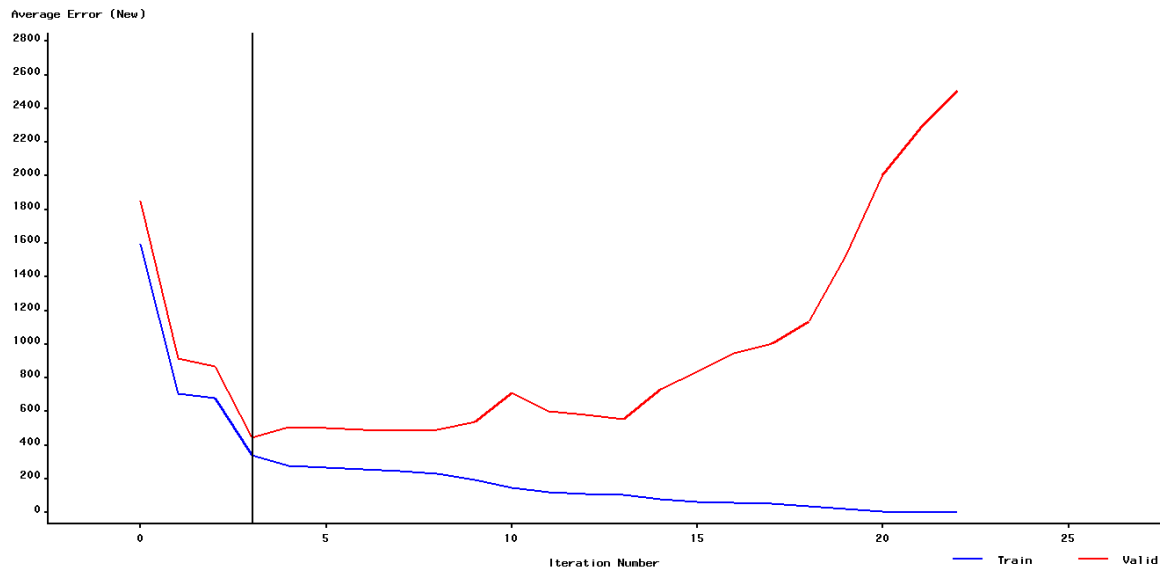


Figura 30 – Gráfico de aprendizagem da rede neuronal 2

5.2.4 Descrição da regressão

A regressão é linear e o critério de selecção é o erro de validação.

Fit Statistic	Label	Training	Validation
AIC	Akaike's Information Criterion	100.70784722	.
ASE	Average Squared Error	3.0048032361	129.1516261
AVERR	Average Error Function	3.0048032361	129.1516261
DFE	Degrees of Freedom for Error	7	.
DFM	Model Degrees of Freedom	30	.
DFT	Total Degrees of Freedom	37	.
DIV	Divisor for ASE	37	16
ERR	Error Function	111.17771974	2066.4260176
FPE	Final Prediction Error	28.760259546	.
MAX	Maximum Absolute Error	3.9737510374	20.21840359
MSE	Mean Square Error	15.882531391	129.1516261
NOBS	Sum of Frequencies	37	16
NW	Number of Estimate Weights	30	.
RASE	Root Average Sum of Squares	1.7334368278	11.364489698
RFPE	Root Final Prediction Error	5.3628592696	.
RMSE	Root Mean Squared Error	3.9852893736	11.364489698
SBC	Schwarz's Bayesian Criterion	149.0353846	.
SSE	Sum of Squared Errors	111.17771974	2066.4260176
SUMW	Sum of Case Weights Times Freq	37	16

Tabela 18 – Estatística da regressão

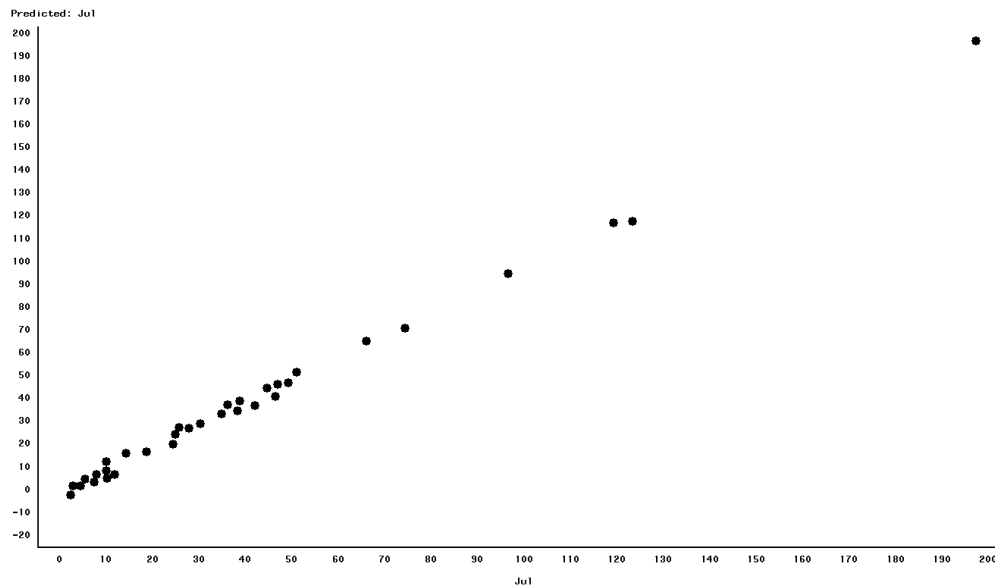


Figura 31 – Gráfico da previsão da regressão

5.3 Análise dos Resultados

É comum ouvir dizer que: “quando as coisas funcionam demasiado bem é porque algo está mal”, a análise dos resultados do modelo preditivo em causa não é excepção. Após algumas execuções, verificou-se que o modelo parecia estar demasiado correcto, pelo que houve uma preocupação em avaliar se tudo estava bem com a metodologia adoptada, com as variáveis que estavam a ser trabalhadas e com os parâmetros dos nós do modelo.

Analisando as diversas previsões do modelo preditivo para o conjunto de treino, que corresponde a 37 freguesias, escolhidas aleatoriamente pelo *partition node* do modelo, obtiveram-se as seguintes previsões e erros associados:

Conjunto Treino	Real Julho	Previsto Decision Tree	erro DT	Previsto Neural Network 1	erro NN1	Previsto Neural Network 2	erro NN2	Previsto Regression	erroREG	total erros
Ajuda	47	40	-7	48	1	41	-6	48	1	-11
Alcantara	119	109	-10	120	1	98	-21	118	-1	-31
Ameixoeira	24	14	-10	23	-1	25	1	22	-2	-13
Beato	35	40	5	35	0	29	-6	36	1	1
Benfica	74	40	-34	74	0	43	-31	73	-1	-65
Campolide	42	41	-1	41	-1	54	12	38	-4	7
Castelo	2	6	4	1	-1	1	-1	0	-2	-1
Charneca	38	40	2	38	0	29	-9	37	-1	-8
Coracao de Jesus	10	14	4	11	1	9	-1	9	-1	3
Encarnacao	97	109	12	97	0	91	-6	96	-1	5
Lapa	39	41	2	38	-1	34	-5	40	1	-3
Madalena	6	6	0	7	1	9	3	6	0	4
Martires	8	6	-2	8	0	0	-8	8	0	-10
Merces	10	14	4	13	3	12	2	14	4	12
Pena	19	14	-5	20	1	18	-1	19	0	-6
Penha de Franca	66	109	43	66	0	56	-10	67	1	34
Prazeres	8	14	6	8	0	10	2	8	0	8
Sacramento	8	6	-2	10	2	9	1	8	0	1
Santa Catarina	7	14	7	8	1	6	-1	6	-1	6
Santa Engracia	30	41	11	30	0	29	-1	31	1	11
Santa Isabel	11	14	3	11	0	17	6	8	-3	6
Santa Maria de Belem	45	41	-4	44	-1	45	0	46	1	-4
Santa Maria dos Olivais	197	109	-88	197	0	98	-99	199	2	-185
Santo Estevao	10	14	4	8	-2	9	-1	8	-2	-1
Santos-o-Velho	25	40	15	23	-2	27	2	25	0	15
Sao Cristovao e Sao Louren	5	6	1	3	-2	0	-5	5	0	-7
Sao Joao	36	40	4	36	0	30	-6	38	2	0
Sao Joao de Brito	14	14	0	15	1	18	4	17	3	8
Sao Jorge de Arroios	123	109	-14	122	-1	97	-26	119	-4	-45
Sao Jose	49	41	-8	46	-3	49	0	50	1	-10
Sao Miguel	3	6	3	3	0	8	5	3	0	8
Sao Nicolau	46	41	-5	46	0	51	5	45	-1	-2
Sao Paulo	51	109	58	54	3	50	-1	54	3	62
Sao Sebastiao da Pedreira	36	41	5	36	0	38	2	37	1	9
Sao Vicente de Fora	8	6	-2	10	2	3	-5	8	0	-5
Se	26	14	-12	28	2	33	7	28	2	-1
Socorro	28	40	12	29	1	24	-4	29	1	10
total	1402	1402	0	1406	4	1201	-201	1402	0	-197

Tabela 19 – Comparação das previsões dos métodos

Pela análise da tabela pode-se verificar que: a um valor elevado dos dados reais, corresponde um erro total previsto, também elevado, no entanto, é aqui que a rede neuronal com uma camada escondida – *neural network 1* - se adapta melhor aos dados, pois foi o método que teve menores erros para valores elevados de *input* do conjunto de treino.

O nó de avaliação – *assessment node*, dá-nos os seguintes erros quadráticos médios de treino e validação do modelo:

Name	Root ASE
decision tree	20.61151357
neural network 2	18.470663607
regression	1.7334368278
neural network 1	1.2335337841

Tabela 20 – Resultados do *assessment*

Pode-se constatar que, a performance da rede neuronal com uma camada escondida – *neural network 1* - é o método com o erro mais baixo, sendo por isso o mais eficaz na previsão por se ajustar melhor aos dados. No entanto, constatou-se que, alterando ligeiramente a *seed* do nó da partição dos dados (*partition node*), ou seja alterando ligeiramente os exemplos escolhidos para treino e validação, alteram-se drasticamente os erros obtidos. O modelo está muito sensível a pequenas variações nos parâmetros, o que se justifica devido ao número reduzido de observações que se está a aplicar aos diversos métodos. Neste caso concreto, uma regressão é igualmente eficaz para prever os dados.

No modelo utilizado, seleccionou-se *apply* no *score node*, para que o conjunto de treino seja avaliado juntamente com o conjunto de dados e não isoladamente.

A previsão do *score* para o mês de Julho para todas as 53 freguesias é produzido pelo modelo com menor erro, neste caso é a rede neuronal 1 – *neural network 1* - com uma camada escondida. Indica-se de seguida os resultados obtidos:

Freguesia	Real Julho	Previsto Julho	Erro
Ajuda	47	47,63	0,63
Alcantara	119	120,43	1,43
Alto do Pina	84	63,00	-21,00
Alvalade	25	18,56	-6,44
Ameixoeira	24	22,87	-1,13
Anjos	46	43,81	-2,19
Beato	35	35,24	0,24
Benfica	74	74,27	0,27
Campo Grande	56	44,56	-11,44
Campolide	42	41,40	-0,60
Carnide	87	88,22	1,22
Castelo	2	0,93	-1,07
Charneca	38	37,98	-0,02
Coracao de Jesus	10	10,88	0,88
Encarnacao	97	96,86	-0,14
Graca	22	20,49	-1,51
Lapa	39	38,35	-0,65
Lumiar	138	141,74	3,74
Madalena	6	6,69	0,69
Martires	8	7,64	-0,36
Marvila	79	100,56	21,56
Merces	10	12,87	2,87
Nossa Senhora de Fatima	76	77,90	1,90
Pena	19	19,66	0,66
Penha de Franca	66	65,78	-0,22
Prazeres	8	8,27	0,27
Sacramento	8	10,14	2,14
Santa Catarina	7	7,81	0,81
Santa Engracia	30	30,04	0,04
Santa Isabel	11	11,36	0,36
Santa Justa	40	85,01	45,01
Santa Maria de Belem	45	44,11	-0,89
Santa Maria dos Olivais	197	196,93	-0,07
Santiago	3	12,30	9,30
Santo Condestavel	43	47,00	4,00
Santo Estevao	10	7,85	-2,15
Santos-o-Velho	25	23,09	-1,91
Sao Cristovao e Sao Louren	5	2,85	-2,15
Sao Domingos de Benfica	135	144,28	9,28
Sao Francisco Xavier	18	7,82	-10,18
Sao Joao	36	35,69	-0,31
Sao Joao de Brito	14	14,84	0,84
Sao Joao de Deus	34	38,50	4,50
Sao Jorge de Arroios	123	122,03	-0,97
Sao Jose	49	46,49	-2,51
Sao Mamede	36	32,05	-3,95
Sao Miguel	3	3,08	0,08
Sao Nicolau	46	45,62	-0,38
Sao Paulo	51	53,52	2,52
Sao Sebastiao da Pedreira	36	36,45	0,45
Sao Vicente de Fora	8	9,85	1,85
Se	26	27,57	1,57
Socorro	28	28,81	0,81

Tabela 21 – Comparação do valor real com o previsto no score

Em termos de resultados obtidos, as previsões da rede neuronal 1, ficaram bastante próximas do valor real do mês de Julho, com uma margem de erro total de apenas 2,05%, o que corresponde a uma previsão superior, em aproximadamente 48 (47,66) ocorrências, do que o valor real para o mês de Julho.

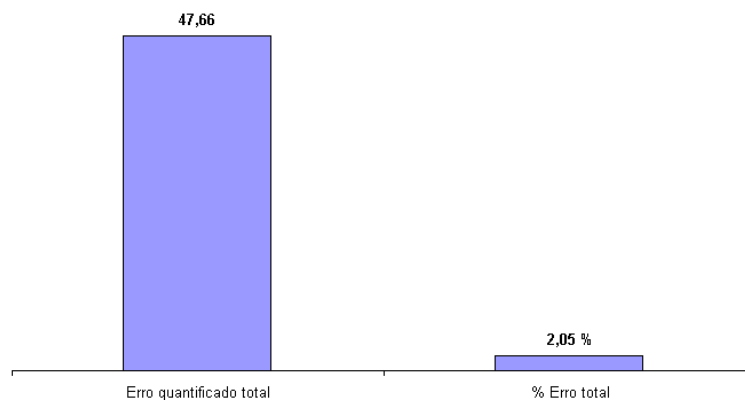


Figura 32 - Erro quantificado do modelo

O facto de, em termos globais, ocorrerem mais ocorrências na previsão (e não menos) é um indicador de uma tendência para um aumento generalizado dos níveis de criminalidade denunciada.³⁷

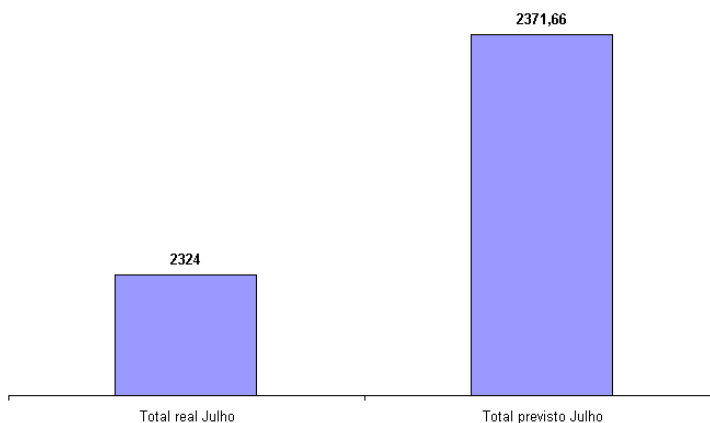


Figura 33 - Comparação dos resultados obtidos com valor real

³⁷ Este facto, veio a constatar-se no Relatório Anual da Procuradoria-Geral da República de 2008, quando refere na página 15 o seguinte: “Em 2008 foram registados 557.884 inquéritos, ou seja, mais 77.462 do que em 2007 o que representa assim, um aumento de 16,1 % na criminalidade participada.”

6. ANÁLISE DE DADOS COM SIG

*We should not be fighting the crime of the twenty-first century with the tools of the nineteenth.*³⁸

Para Matos, J. (2001:39), a construção de um modelo geográfico para descrição da realidade tem como principal problema a diversidade do universo de fenómenos a incorporar. O conceito de fenómeno é aqui utilizado na acepção de Kant como tudo o que é objecto de experiência possível, isto é, tudo o que aparece no tempo ou no espaço.

Já há mais de um século, os epidemiologistas e outros cientistas da área da medicina, começaram a explorar o poder dos mapas na ajuda à assimilação e compreensão da dinâmica espacial da doença, através do uso de mapas, indicando a distribuição geográfica das mortes provocadas pela cólera.

John Snow (1854) foi talvez o primeiro a fazer o mapeamento gráfico de uma série de ocorrências com o intuito de descobrir a sua origem, através do registo de óbitos provocados pela cólera no Verão de 1854.

Snow, fez uma análise de *cluster* em Londres, com recurso ao desenho de um mapa, a sua hipótese era a de que, a doença era transportada pela água contaminada numa das ruas de Londres, por isso, desenhou a localização das bombas de água, identificando a fonte do problema pela proximidade das vítimas de cólera (*pump A- Broad Street*).³⁹

³⁸ Jack Straw (1997) "Não devemos combater o crime do século XXI com as ferramentas do século XIX"

³⁹ Mapa original elaborado por John Snow (1854)

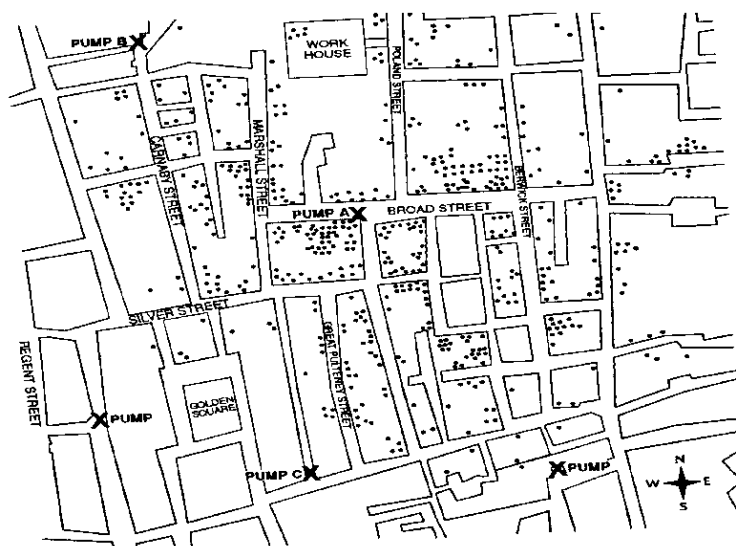


Figura 34 – Mapa da cólera em Londres

Demonstrou assim, que, as associações entre mortes por cólera e abastecimento de água contaminada, resultaram numa distribuição geográfica coincidente.

Para Matos, J. (2001:41), a modelação geográfica resulta de um compromisso entre, sintetizar conhecimento a partir de um conjunto de dados e simultaneamente, providenciar a informação com conteúdo tão completo quanto possível, de forma a poder ainda ser operada com vista a representar conhecimento sob outro enquadramento de análise.

Hoje em dia, é usado a metodologia do *crime mapping* com os mesmos princípios mas com ferramentas diferentes, nomeadamente os Sistemas de Informação Geográfica – SIG.

Para Painho, M. (2006), os SIG definem-se como: “um sistema composto por hardware, software e um ambiente institucional que permite armazenar, processar, visualizar e analisar dados de natureza espacial, referenciados à superfície da terra”.

Maguire (1991) *apud* Matos (2001:55) refere ainda que:

“As especificações funcionais de um SIG, devem apoiar a resolução das questões relativas a identificação de objectos visualizados, interrogação e visualização de objectos que verifiquem uma dada condição, identificação de alterações e de padrões espaciais e modelação de fenómenos com desenvolvimento espacial”.

Um SIG tem muitos campos de aplicação e poderá deste modo sofrer adaptações consoante a área ou domínio de utilização, no entanto, as suas capacidades de utilização, de integração de operações correntes de gestão de bases de dados, como inquirição e análise estatística, com os benefícios de visualização e de análise geográfica, proporcionada pela utilização de mapas, distinguem o SIG de outros sistemas de informação.

De acordo com Matos, J. (2001:265), um SIG, distingue-se ainda em 2 tipos de utilização:

- Primeiro, o SIG de gestão, onde prevalecem as preocupações com: segurança, integridade, concepção e desenvolvimento de aplicações específicas de utilização, distribuição e manutenção dos dados;
- Segundo, o SIG de projecto, o objectivo é a resolução de um dado problema, sem preocupação com a posterior utilização e manutenção, ainda que os resultados possam vir a ser reutilizados.

O SIG está organizado por camadas, da seguinte forma: os objectos espaciais são organizados em níveis de informação ou temas, cada nível de informação é constituído por um conjunto de objectos geográficos ligados topologicamente ao nível de atributos.

Como ferramenta de representação da realidade, um SIG utiliza 2 tipos de modelos geográficos de dados:

1. Raster;
2. Vectorial.

Executa essencialmente 5 tipos de processos ou tarefas:

1. Entrada de dados;
2. Manipulação;
3. Gestão;
4. Inquirição;
5. Análise e visualização.

Quanto ao tipo de dados, estes podem ser: temas tipo polígono, temas tipo linha e temas tipo ponto.

A título de introdução sobre as teorias espaciais da criminalidade descritas por Patrícia & Paul Brantingham (1981), existem 4 dimensões para cada crime:

1. Dimensão da infracção legal;
2. Dimensão vítima do crime;
3. Dimensão do autor do crime;
4. Dimensão espacial ou local do crime.

A última dimensão, espacial, foi descrita como um local e o tempo em que ocorre a infracção. Tal como refere Chainey (2005), estas componentes da equação criminal podem ser mapeadas e avaliados com recurso nos SIG - Sistemas de Informação Geográfica.

Segundo Kate (Kate *et al.* 2004), avaliar os resultados do modelo preditivo, envolve a medição do grau de incerteza associado às previsões. Mapas agradáveis à vista podem não ser adequados para termos um modelo eficiente. Um dos problemas com as técnicas de *crime mapping* foi a falta de avaliação dos modelos preditivos em prever a ocorrência de novos incidentes (Gorr & Ollitschlaeger, 2002).⁴⁰

Segundo Kate *et al.* (2004), numa primeira abordagem, parece fácil avaliar um modelo preditivo: simplesmente inserimos dois anos de informação, usamos o primeiro ano para definir o *Hot-Spot* e o segundo ano para mapear os resultados e comparar com os do modelo preditivo. No entanto encontramos alguns problemas nesta análise:

⁴⁰ Normalmente usa-se a *Confusion Matrix* - Matriz Confusão, para avaliar a qualidade do modelo preditivo, esta matriz, dá-nos a percentagem de falsos positivos e falsos negativos do modelo.

- Este tipo de abordagem é basicamente “visual”. Quando se lida com cobertura de imagem do tipo *grid* ou *raster* e não com o formato de vectores, não é possível trabalhar o SIG para que possa “contar” o número de ocorrências por cada *Hot-Spot* ou *grid*. Isto inviabiliza qualquer técnica de avaliação sistemática;
- Existe relativamente pouca documentação que efectivamente faça comparações entre diferentes técnicas de mapeamento. É necessário um método de comparação para avaliar os resultados obtidos, para além de ser necessário, um conjunto de teste (independente dos restantes dados) para testar a fiabilidade do modelo preditivo;
- Existem características que se devem considerar quando se avalia a fiabilidade dos mapas de criminalidade produzidos como: o número de ocorrências por *Hot-Spot*, a área geográfica coberta, a distância e o número de *Hot-Spots* ideal, para uma prevenção eficaz tendo em conta os meios operacionais disponíveis, *etc.*

Antes de continuar, resta justificar o interesse de criar este mapeamento espacial ou temporal dos crimes. Se estes ocorressem de forma aleatória, não existiriam razões para realizar este estudo, nem existiriam evidências de que crimes futuros seriam realizados no mesmo espaço temporal ou físico de crimes passados e portanto nenhuma previsibilidade daí adviria. Felizmente esse não é o caso.

A polícia no desempenho da sua missão terá de averiguar a perigosidade dos locais da sua área geográfica de responsabilidade independentemente da dificuldade em cumprir esse objectivo, pois isso reflecte-se no sentimento de insegurança da comunidade.

A polícia agir deste modo, está em concordância com o que o Ministro da Administração Interna, Dr. Rui Pereira, quando disse “...não há nenhum sítio onde a polícia tenha dificuldade em entrar.”⁴¹

Existem de facto zonas mais perigosas, bem como horas do dia mais perigosas onde determinado tipo de crimes têm mais probabilidade de ocorrência.

⁴¹ Programa da RTP “Prós e Contras” do dia 30/3/2009

Recentemente foram identificados os bairros de risco na grande Lisboa tendo como principais causas:⁴²

- Zonas com condicionantes arquitectónicas preocupantes;
- Elevada densidade populacional;
- Elevado número de residentes com antecedentes criminais;
- Número elevado e preocupante de casos de desordem pública;
- Número elevado e preocupante de agressões a elementos policiais;
- Elevando número de casos de criminalidade imanente.⁴³

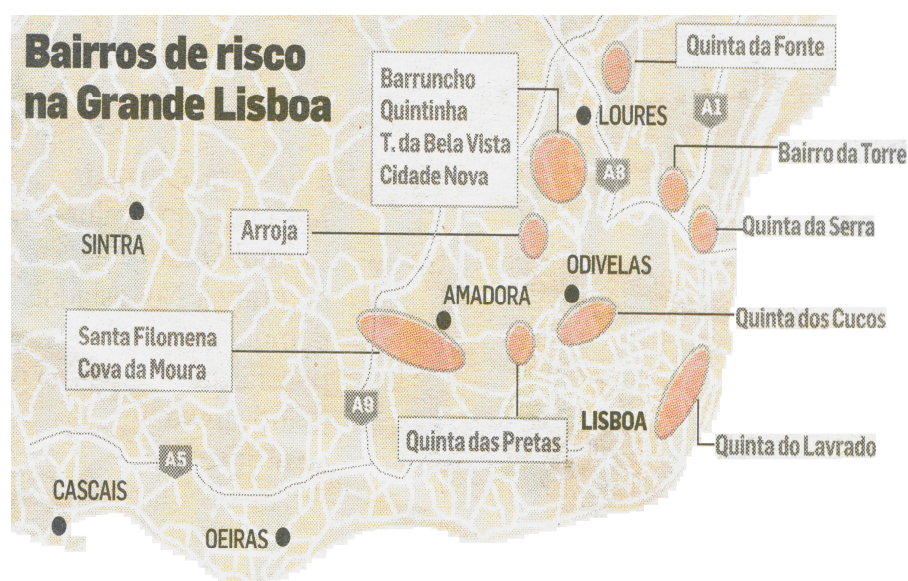


Figura 35 – Bairros de risco na grande Lisboa

Esta opção estratégica de identificação dos bairros problemáticos, vem ao encontro da teoria *broken windows*, segundo a qual o factor desencadeador da criminalidade e da violência nos bairros das cidades são: as incivildades, os desvios comportamentais, a degradação do ambiente e os atentados à qualidade de vida dos residentes.

Para Gomes (2004) este clima de "paz podre" gera um sentimento de tolerância social e institucional e de crescente impunidade dos delinquentes que tende a estimulá-los à prática de actos cada vez mais violentos.

⁴² Artigo publicado no Diário de Notícias de 11 de Maio de 2009:4

⁴³ Projecção da criminalidade para fora dos bairros.

Torna-se, assim, necessário, a partir de determinado momento crítico e como forma de restabelecer uma paz e tranquilidade sustentáveis, adoptar uma estratégia policial de “tolerância-zero” face às condutas aparentemente menos censuráveis - incivildades que atentam contra a segurança rodoviária e a qualidade de vida - mas que são um factor que contribui em grande medida para o sentimento de insegurança dos cidadãos.

Vejam os casos concretos, Ratcliffe (2001), incluiu entrevistas a autores de furtos a habitações e escritórios, com base nos estudos realizados e no mapeamento de zonas da cidade e horas do dia, os furtos a residências são realizados durante o dia, com o seu expoente máximo durante a tarde e os furtos a escritórios, durante a noite, tendo o padrão inverso.

O mesmo acontece com as zonas da cidade, algumas mais residenciais e outras mais de serviços.

Sem este mapeamento, empiricamente, acreditaríamos provavelmente que os furtos a residências ocorreriam, por exemplo, maioritariamente durante a noite em períodos de férias.

As explicações para estes acontecimentos podem ser: as actividades de rotina da população, a atractividade dos bens furtados, entre outras características, que não se encontram no âmbito do presente estudo.

Concluiu Ratcliffe (2001) que, a necessidade de mapeamento possui três espaços distintos:

- Aquele que temos a capacidade de controlar;
- Aquele onde consideramos coexistir um misto de oportunidade e atractividade;
- Aquele onde ocorrem crimes.

Neste mapeamento deveremos também ter em atenção a noção de vizinhança, ou seja, um determinado crime deve afectar uma zona e sua vizinhança e não apenas um ponto no mapa (princípio do mapeamento *Hot-Spot*).

6.1 Análise espacial da criminalidade

Uma das primeiras perguntas realizadas aos SIG, pelas Forças e Serviços de Segurança ou pelo cidadão comum, será saber onde são as zonas de maior índice de criminalidade.

A literatura, responde a esta questão com diversas aproximações aos modelos espaciais e preditivos, donde se destacam os seguintes métodos:⁴⁴

- *Hot spots* (e. g. Jefferis);⁴⁵
- Análise da taxa de criminalidade agregada (e. g. Osgood);⁴⁶
- Séries Temporais (e. g. Gorr);⁴⁷
- Redes Neurais (e. g. Olligschlaeger);⁴⁸
- Modelos de processamento de pontos (e. g. Liu & Brown).⁴⁹

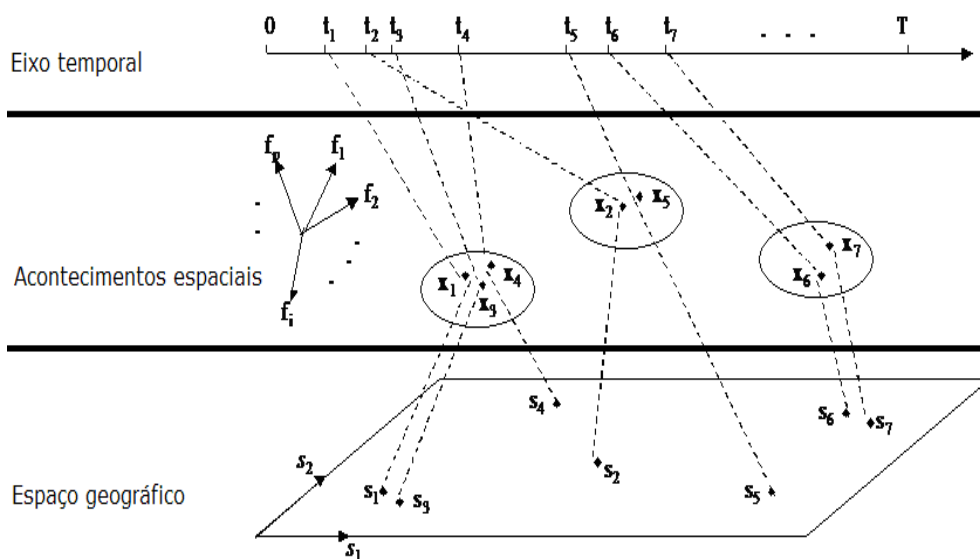


Figura 36 - Modelo de processamento de pontos

⁴⁴ vide Chaaney (2005) ou para exemplos de utilização Canter (1995).

⁴⁵ Jefferis, E. (1999) in "A multi-method exploration of crime hot spots"

⁴⁶ Osgood, W. (2000) "Poisson-based regression analysis of aggregate crime rates"

⁴⁷ Gorr, D. & Tita, G. (2000) in "Spatial analyses of crime"

⁴⁸ Olligschlaeger, A. (1997) in "Artificial neural networks and crime mapping".

⁴⁹ Liu & Brown (1999) in "A new approach to spatial-temporal criminal event prediction"

O desafio relativo à identificação de *Hot-Spots* pode ser perceptível através do exemplo de comparação de criminalidade em zonas rurais e urbanas, um nível de criminalidade numa zona rural pode ser alto localmente, mas baixo quando comparado com uma zona urbana.

Contudo, Canter (1995), esclarece que, um *Hot-Spot* não deixa de ser um *Hot-Spot* por ser apenas local, pelo que é necessário reflectir sobre os métodos e estratégias que assegurem a sua correcta identificação, tais como:

- Frequência de ocorrência do crime;
- Geografia da ocorrência do crime;
- Espaço temporal.

Chainey (2005) diz-nos que, após a identificação dos *Hot-Spots*, os Sistemas de Informação Geográfica podem seguir diversas estratégias de visualização da informação criminal:

- **Pontos** – equivalente a colocar *pins* numa parede de cortiça em cima de uma carta militar. De facto é o método mais popular para um SIG representar a criminalidade (ou para qualquer outra informação);
- **Mapas temáticos** – o número de crimes ou outras métricas são agregados, utilizando paletes de cores (mais escuro para mais crimes, mais claro para menos crimes), em limites administrativos ou outros.

Este método é bastante utilizado porque se percebe, por exemplo por concelho, aquelas onde a criminalidade tem mais incidência. No entanto, e embora este sendo também um método popular, é pouco provável que a criminalidade seja equitativa por Concelho (e. g. Lisboa), pelo que pode levar a interpretações que não correspondem exactamente à “verdade” porque a densidade populacional é diferente em cada local, neste caso, o mais sensato é usar rácios na representação das cores. Este tipo de mapas gera um conflito jurisdicional entre a área geográfica administrativa e policial, quando estas são diferentes (como é o caso de Lisboa).⁵⁰

⁵⁰ Foi esta a estratégia de visualização do presente estudo.

- **Mapa de rede** – o mapa é dividido em células equitativas (princípio dos SOM e mapas de *Varano*), sendo associado a cada uma delas uma cor, de acordo com uma paleta de cores e intervalos pré-definidos, estas células, que podem variar de tamanho, levam a uma análise mais cuidada dado que não se “prendem” com limites administrativos ou policiais.

Chainey (2005), conclui dizendo que: a análise espacial da criminalidade não se trata apenas de uma forma de visualização mas também da visualização ou a percepção e compreensão das variáveis que potencialmente contribuem para o crime.⁵¹

6.2 Análise temporal da criminalidade

Como refere Chainey (2005), é importante “entender o mosaico do comportamento criminal” ao longo do tempo. Um dos aspectos focados por, Georges (1978) e por Chainey (2005), será a observação de um determinado espaço temporal, numa zona piloto, após a implementação de estratégias de combate à criminalidade, servindo-nos da ciência e SIG como validação para a estratégia implementada.

Para Chainey (2005), podem ser consideradas 5 categorias temporais de análise da criminalidade:

- **Momentos** – normalmente indicam a data e hora de ocorrência de um crime;
- **Duração** – em alguns tipos de crime será muito importante ter uma indicação da sua duração (e. g. através das câmeras de vigilância). Esta medida ajuda a estabelecer possíveis níveis de serviço para a resposta das forças policiais;

⁵¹ A raiz dos problemas criminais têm sido alvo de hipóteses e a título de exemplo, podemos referir os trabalhos relacionados com o desemprego de Weatherburn (2001) e sazonalidade de Farrel (1994), entre outros.

- **Tempo estruturado** – é, por exemplo, quando um analista verifica que um infractor esteve activo durante alguns meses, o que nos permite descrever o número de infracções mensalmente ou anualmente, para potenciar, por exemplo, a comunicação entre forças e serviços de segurança;
- **Tempo como distância** – a verificação do tempo como distância tem particular relevância nesta matéria. Quando um Juiz indica uma medida de Termo de Identidade e Residência para um determinado cidadão e este se apresenta numa esquadra de Lisboa todas as noites, este não deverá estar envolvido em confrontos na Baixa do Porto num desses dias.
O tempo como distância indica, para os diversos tipos de crime, a distância por exemplo que o infractor teve de percorrer desde a sua casa até ao roubo, ou de um roubo para outro, por forma às forças policiais melhor combaterem o crime após a sua ocorrência.

6.3 Modelo SIG para representação dos dados

Em termos de densidade urbanística e populacional, Lisboa caracteriza-se da seguinte forma:

- Lisboa tem uma área de 84,6 km quadrados;
- Tem um rácio de edifícios por km quadrado superior a 642;
- Tem um rácio de habitante por km quadrado superior a 6674;
- Em média existem 642 edifícios por km quadrado com 6674 habitantes;
- Em média existem 10 habitantes por cada edifício;

Para representar os dados no Sistema de Informação Geográfica, foi utilizado o *software Geomedia Professional*, com os dados dos Censo de 2001, os quais foram pré-processados no *Excel* juntamente com os dados das ocorrências de 2007, o resultado deste pré-processamento foi posteriormente agrupado (JOINT) com um *Shapefile* obtido na CAOP 2008, obtendo-se os mapas temáticos finais.

Os mapas temáticos obtidos, obedecem a uma classificação em 6 classes e o método utilizado nos intervalos é o *equal range*, para se obter uma visualização mais eficaz das distribuições dos dados e assim facilitar a interpretação dos mapas.⁵²

O modelo SIG utilizado obedeceu às seguintes etapas:⁵³

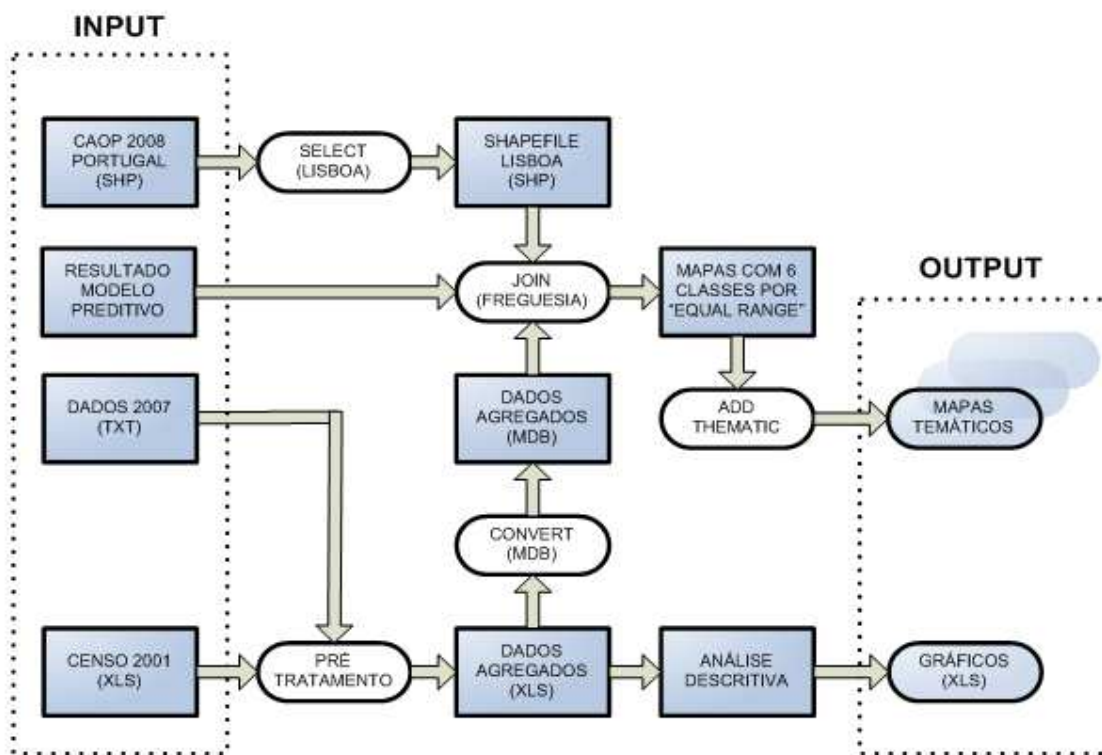


Figura 37 – Diagrama do modelo SIG

Para a representação dos mapas temáticos, foram analisadas as 53 freguesias de Lisboa, que se encontram divididas geograficamente nos seguintes mapas.

⁵² Embora existam outros métodos (e. g. *min-max*, *quantiles*, *natural-breaks*, *nested means*, *rácios*, *desvio-padrão*) para classificar os dados, o método adoptado por intervalos de *equal range* é o que representa (neste caso) visualmente melhor as diferentes densidades entre os atributos, este método consiste subdividir os dados iniciais em subcamadas com exactamente a mesma amplitude, no entanto, certas distribuições particulares - uniformes - dos dados não se reflectem com uma boa visualização nos resultados obtidos. (tradução livre do artigo de Geoffrev Dutton “*Bivariate Construction of Equal-Class Thematic Maps*” s. d.).

⁵³ O ficheiro com a Carta Administrativa Oficial de Portugal de 2008 - CAOP, pode ser descarregado directamente do endereço: <http://www.igeo.pt/produutos/cadastro/caop/inicial.htm>

6.4 Mapas Temáticos de Lisboa

O concelho de Lisboa é constituído por 53 freguesias, caracterizando-se por um crescimento em anel, ou seja, partindo do centro onde encontramos uma grande concentração de freguesias com pequena dimensão, geograficamente desorganizadas, que estiveram na génese da história da cidade, chegamos a um anel exterior onde se encontram as freguesias mais recentes, com grande área geográfica e grande densidade populacional.

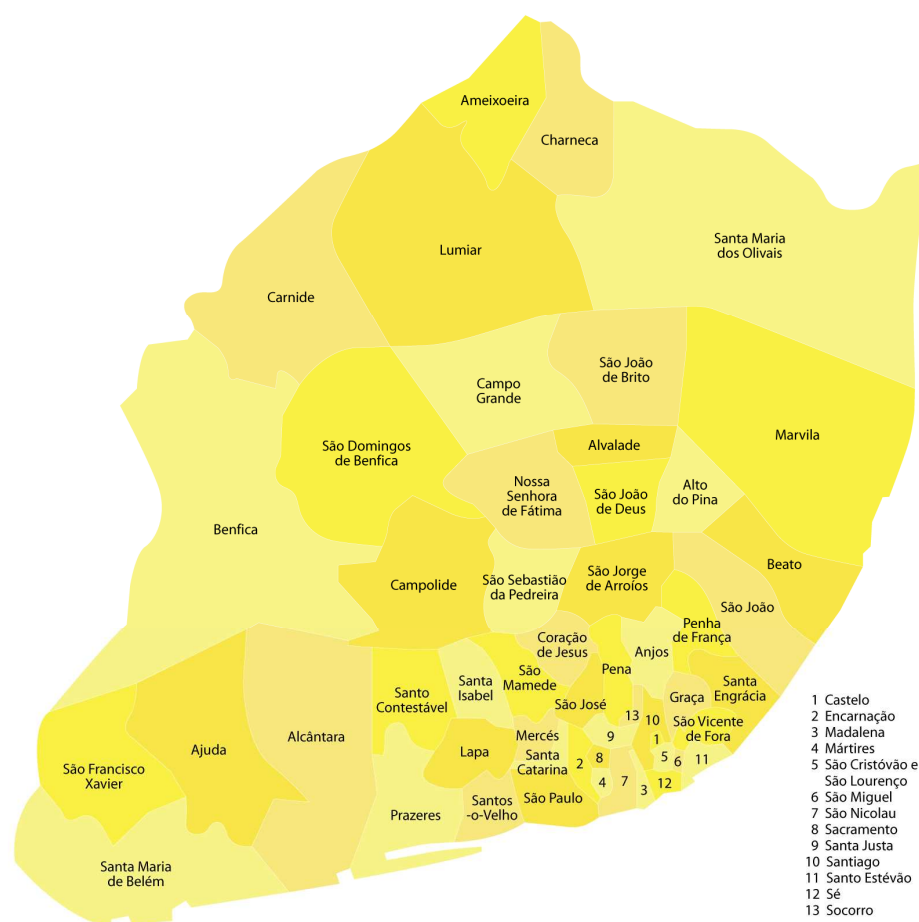


Figura 38 – Mapa das 53 freguesias de Lisboa

Seguem-se os mapas temáticos mais relevantes para este trabalho, obtidos com os dados que nos foram disponibilizados juntamente com os dados do Censo 2001.⁵⁴

⁵⁴ Em anexo encontram-se os restantes mapas temáticos obtidos

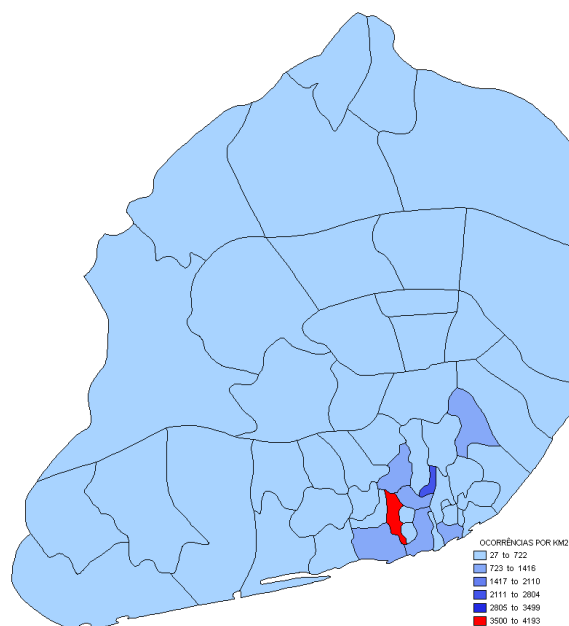


Figura 39 – Mapa da densidade de ocorrências pela área

Freguesia com maior densidade de ocorrências pela área:
Encarnação (4.193).

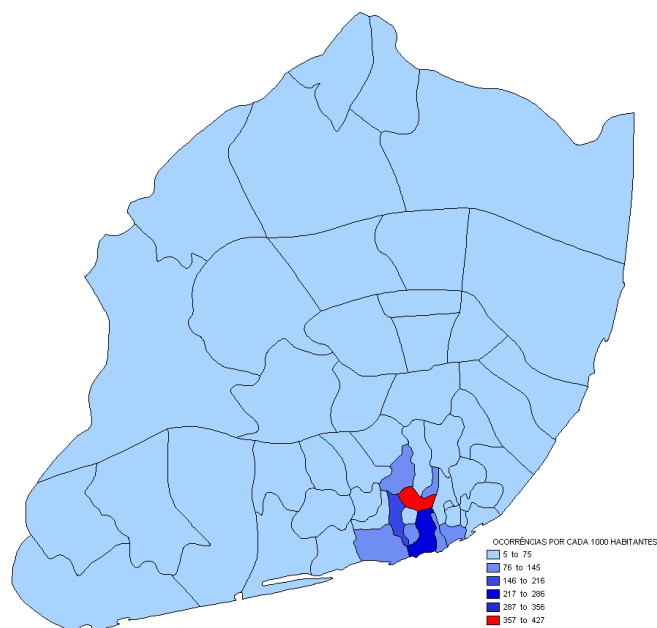


Figura 40 – Mapa da densidade de ocorrências por cada mil habitantes

Freguesia com maior densidade de ocorrências por cada mil habitantes:
Santa Justa (427,14).

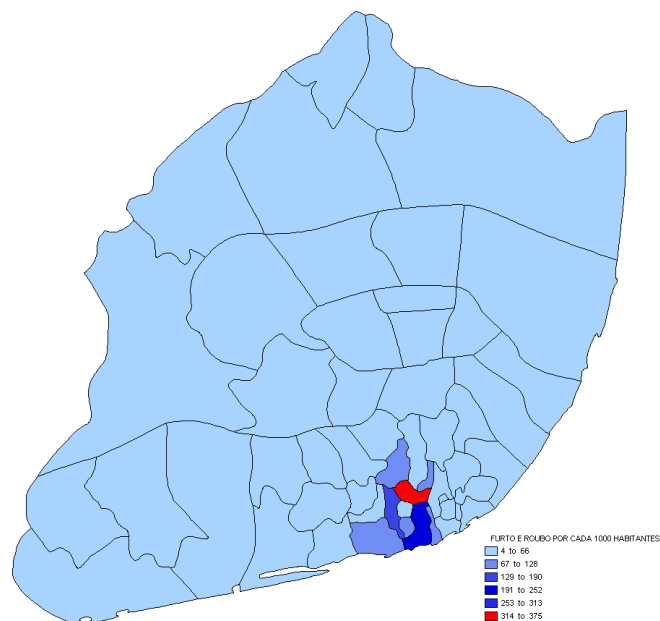


Figura 41 - Mapa da densidade do crime de furto por cada mil habitantes

Freguesia com maior densidade do crime de furto e roubo por cada mil habitantes:

Santa Justa (375,71).

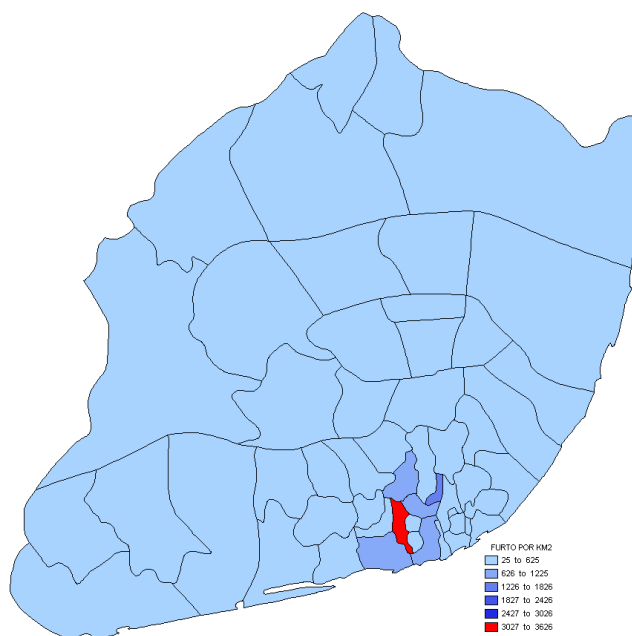


Figura 42 - Mapa da densidade do crime de furto e roubo pela área

Freguesia com maior densidade do crime de furto e roubo pela área:

Encarnação (3.627).

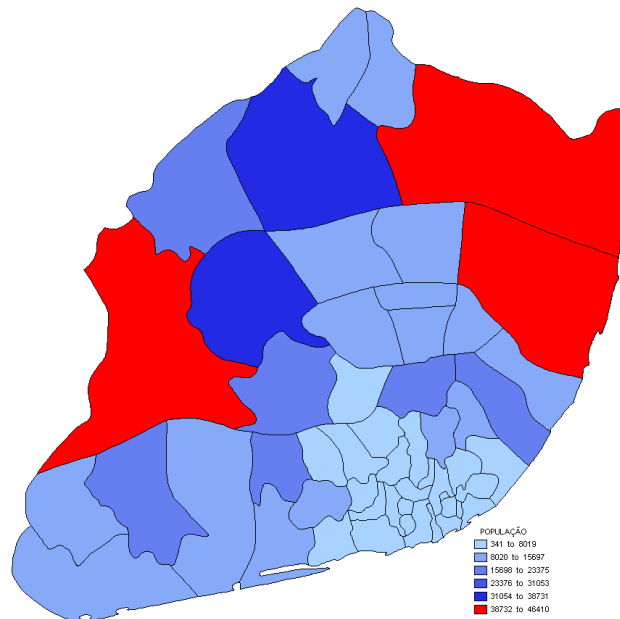


Figura 43 – Mapa da população por freguesia

Freguesias com maior densidade populacional:

Santa Maria dos Olivais (46.410), Marvila (38.767) e Benfica (41.368).

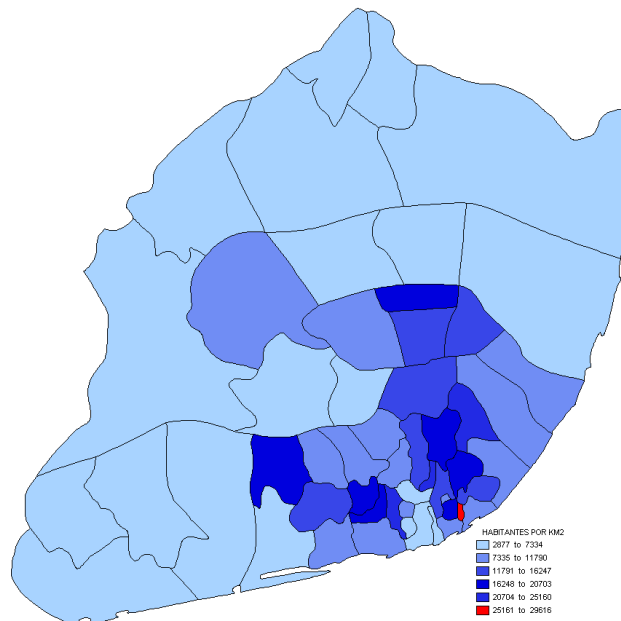


Figura 44 - Mapa da densidade populacional pela área

Freguesia com maior densidade populacional pela área:

São Miguel (29.617).

Para efeitos de comparação dos mapas temáticos do valor real e o mapa temático com o valor previsto pelo modelo, a classificação passou para 12 classes, mantendo-se o método *equal range* para diferenciar os intervalos.

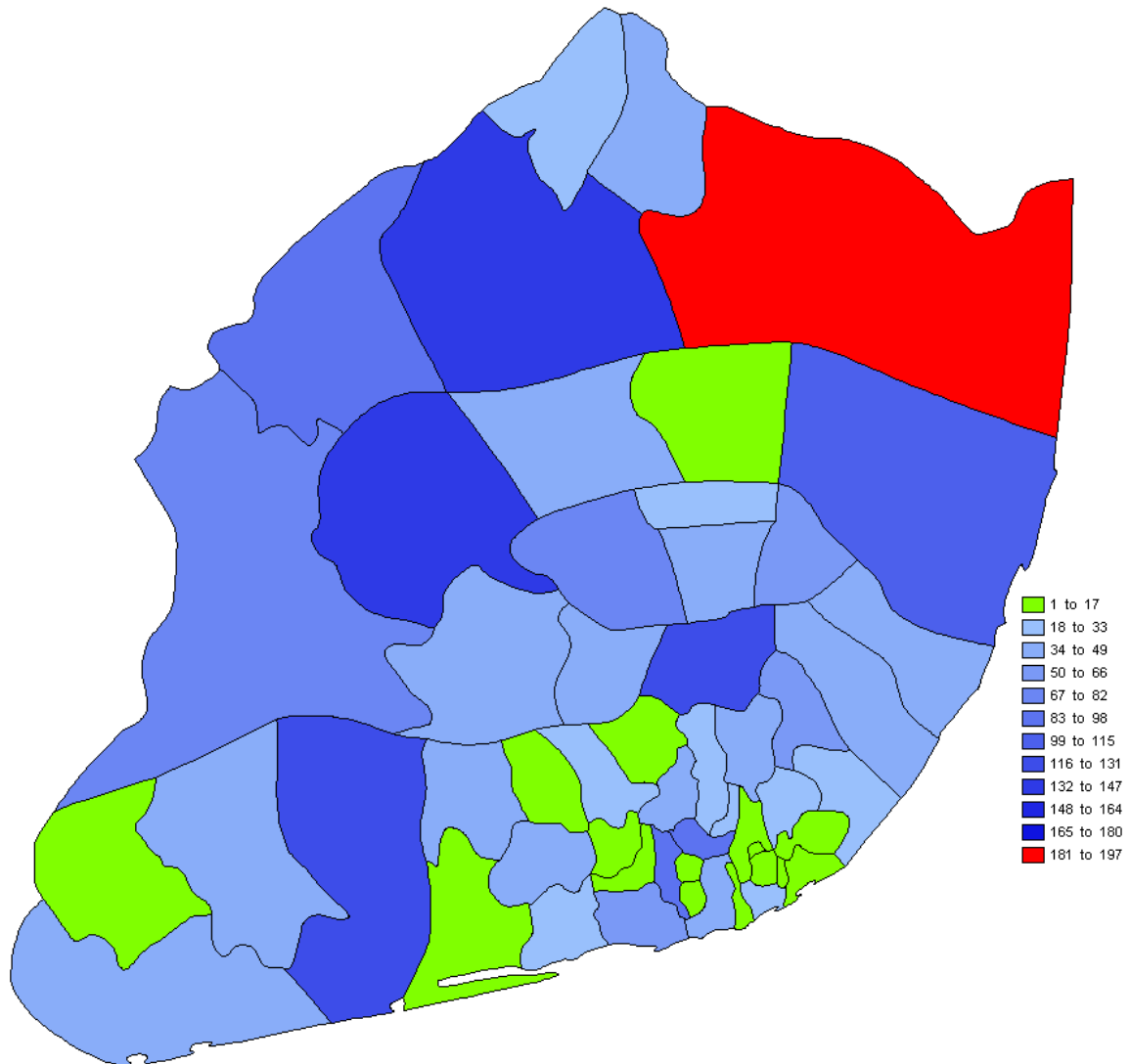


Figura 45 - Mapa dos valores REAIS para o mês de Julho

Freguesia com maior densidade de crimes em Julho:
Santa Maria dos Olivais (197).

Comparando os valores reais com os valores previstos, nota-se uma maior diferença nas freguesias de Marvila (mais 21,56 ocorrências) e Santa Justa (mais 45,01 ocorrências).

Usando agora os novos dados obtidos pelo modelo preditivo para o mês de Julho, obtemos o seguinte mapa:

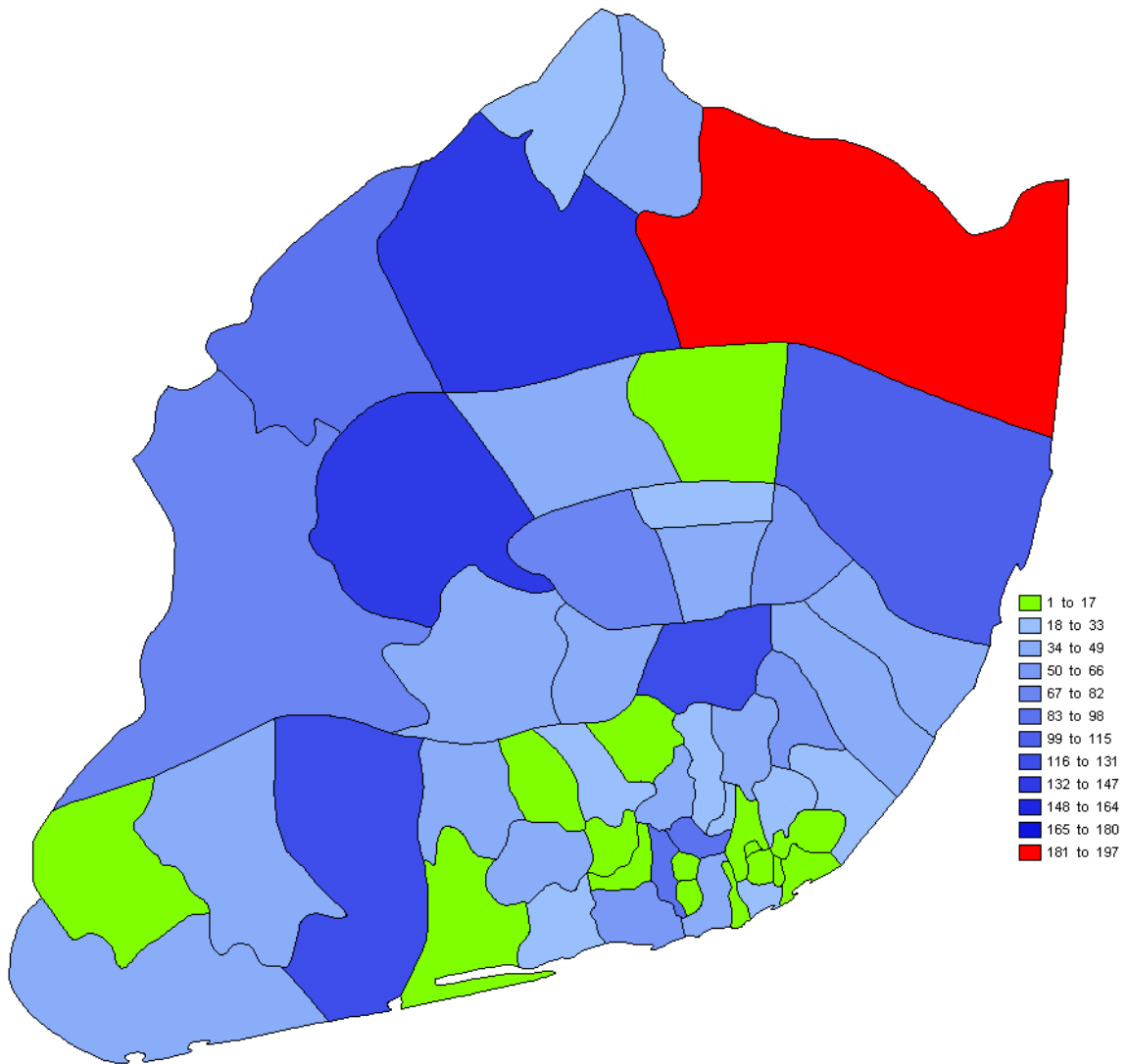


Figura 46 - Mapa dos valores PREVISTOS para o mês de Julho

Freguesia com maior densidade de crimes em Julho:
Santa Maria dos Olivais (196,93).

Como era de esperar, ambos os gráficos são basicamente, idênticos, tendo em conta a previsão efectuada, o que, mais uma vez, nos indica, a boa adequação do método utilizado no modelo preditivo para efectuar a previsão (rede neuronal com uma camada escondida).

7. TRABALHOS FUTUROS

Existem actualmente, diversos projectos, que têm vindo a melhorar a relação da polícia com o cidadão, alguns inovadores, outros apenas melhorias do que já existe de forma dispersa. Dos diversos projectos em curso, destacam-se alguns que embora ainda não possam ser avaliados, realçam no entanto a linha de orientação deste trabalho.

7.1 Projecto SIGpol

Sendo o uso interligado e articulado das tecnologias de informação um factor essencial da eficácia das forças policiais, foram lançados projectos inovadores como o SIGpol (que usa a informação georeferenciada para registar a criminalidade e apoiar a acção policial).⁵⁵

O projecto SIGpol resulta de um acordo celebrado entre o ISCPSI e o Centro Nacional de Informação Geográfica (CNIG), com vista a criar um registo automatizado de ocorrências de natureza policial, para a área de Lisboa, tendo por base um Sistema de Informação Geográfica (SIG), foi pensado em torno de três factores chave:

- Elaboração de expediente policial;
- Recolha de dados e análise estatística;
- Apoio à decisão das patrulhas.

O grande objectivo deste projecto é facilitar, ao máximo, a elaboração das peças de expediente, por parte dos operacionais que trabalham nas esquadras. Para tal, foi criada uma interface gráfica amigável, para o utilizador, que o guia em todos os passos que são necessários efectuar.

⁵⁵ in http://www.inst-informatica.pt/servicos/informacao-e-documentacao/biblioteca-digital/sociedade-da-informacao-1/2005-e-anos-anteriores/Portugal_na_SI.pdf (adaptado)

Quando todos os dados estiverem inseridos no computador, o agente manda imprimir o documento que depois segue os trâmites legais normais.

Os dados referentes à localização da ocorrência policial são armazenados de forma a serem georeferenciados através de um SIG, para permitir uma posterior análise estatística da distribuição espácio-temporal da criminalidade.

Com este sistema poder-se-á analisar sistematicamente todas as variáveis e verificar as horas e os locais mais propícios à prática criminal.

O SIGpol permitirá, ainda, análises mais complexas, mas automatizadas através de algoritmos matemáticos desenvolvidos pela Investigação Operacional, que apoiarão as decisões do Comandante da Esquadra na criação de patrulhas, que privilegiem as áreas onde seja mais susceptíveis a ocorrência de actividades criminosas.

Este projecto encontra-se já numa fase de experimentação inicial, numa esquadra, a fim de se verificar a fiabilidade do sistema. Para a evolução do SIGpol estão previstas várias alternativas como a ligação em rede e a compatibilização com outras aplicações que possam entretanto surgir.

Embora este projecto seja ambicioso e necessário, a verdade é que 10 anos após a decisão da sua implementação, ainda não se viram resultados práticos e muito menos a sua extrapolação para todas as esquadras.

7.2 Esquadra virtual

A "esquadra virtual" é uma das inovações resultantes do Sistema Estratégico de Informação, Gestão e Controlo Operacional (SEI), que vai ligar todos os serviços da PSP a uma espécie de "cérebro electrónico".⁵⁶

A sua função é gerir todas as informações em conjunto e em tempo real, desde ocorrências e participações, até missões em grandes eventos, passando por outras actividades, como as licenças de porte de armas, o controlo da ocupação de celas ou a elaboração de relatórios e estatísticas.

⁵⁶ in http://revistadeimprensa.forumpsp.net/dossier_021.htm (adaptado)

É um modelo para interligar os cidadãos entre si e estes com os diferentes organismos públicos e privados que compõem o tecido social das cidades, é a construção de uma *Intranet* urbana que normalmente utiliza os mesmos circuitos da Internet e pode ser acedida através desta.

Esta esquadra virtual poderá ser uma extensão da esquadra real no ciberespaço, em que a população, da área geográfica da jurisdição desta, poderá recorrer, para resolver, de forma expedita, os seus problemas.

Muitos dos serviços e informações da esquadra virtual têm de ser capazes de orientar a sua acção específica, para a resolução dos problemas da comunidade.

Na esquadra virtual existirá a preocupação de transmitir informação com interesse directo a uma comunidade específica, por exemplo:

- Os índices de criminalidade da área;
- Os tipos de crimes mais comuns e as medidas de prevenção situacional a tomar pela população;
- Divulgação dos projectos de segurança que estão a ser desenvolvidos especificamente para aquela comunidade;
- Divulgação das principais solicitações não criminais colocadas à polícia;
- Explicação da forma como os recursos humanos e materiais da esquadra real são empenhados.

7.3 Cooperação policial ao nível das tecnologias da informação

O combate à criminalidade só será eficaz se existir uma forte cooperação entre as diversas forças da ordem nacionais.

A interligação das tecnologias de informação das polícias portuguesas poderá tornar-se algo complicado, para os problemas da cooperação policial ao nível nacional.⁵⁷

⁵⁷ RASI (2008:41) "Em 2008, foi constituída a Rede Europeia dos Serviços Tecnológicos de Polícia, cujo objectivo é o de promover uma melhor divulgação e troca de experiências entre serviços competentes dos Estados-membros encarregues da investigação e da promoção de novas tecnologias em matéria de segurança quando estas possam servir à protecção de pessoas e bens e serem adaptadas às necessidades operacionais."

E na realidade esta ligação trará grandes vantagens, no entanto é necessário ter atenção aos limites a respeitar e a alguns obstáculos a ultrapassar.

O aproveitamento conjunto dos meios tecnológicos por parte das várias polícias poderá trazer duas grandes vantagens:

A primeira e a rentabilização dos meios existentes, os recursos humanos, materiais e financeiros das polícias, estes, são sempre reduzidos para o nível de solicitações a que estas estão sujeitos. É de todo o interesse encontrar formas de maximizar o aproveitamento dos recursos existentes.

A segunda é o aumento da eficácia policial graças à rápida permuta de informações de âmbito policial e à criação de canais de comunicação modernos e eficientes.

A permuta electrónica de informações sobre as actividades criminais, os suspeitos da sua prática ou simplesmente dos métodos utilizados é essencial e poderá vir a desempenhar um importante papel no combate à criminalidade.

Cada organização policial deverá criar, no âmbito das suas competências, bases de dados com informações policiais e preocupar-se em implementar mecanismos que permitam o acesso por parte das outras organizações policiais.

A utilização de meios modernos de comunicação, como o *email*, é outra forma de aumentar a eficácia policial. Desta forma, seria possível às polícias comunicarem directamente enviando e recebendo esclarecimentos no mais curto espaço de tempo possível, ao contrário do que ocorre actualmente em que a comunicação é feita através de correio institucional tradicional, que é demasiado demorado e ineficaz.

Para garantir um bom funcionamento da cooperação policial, no âmbito das tecnologias de informação é necessário estar atento para alguns obstáculos que são necessários ultrapassar, nomeadamente:

- A estrutura de cooperação;
- O desfasamento tecnológico;
- A cultura organizacional.

7.4 A GEO-Polícia

É um novo conceito de mútua cooperação estratégica entre as polícias dos 30 países de língua portuguesa e espanhola. GEO é uma marca com finalidade altruísta, institucional e comercial da fundação Geolíngua e acima de tudo, um conceito e uma meta a atingir.⁵⁸

A fundação Geolíngua assume a paternidade da criação de um polícia com a marca “GEO”, para que o mesmo possa obter os benefícios financeiros, decorrentes da comercialização da referida marca representada pelo seguinte logotipo.



Figura 47 – Logotipo da GEO-Polícia

A fonte de rendimento para os salários, formação profissional e equipamentos para a GEO-Polícia, virá dos lucros obtidos com a venda dos produtos e serviços com a marca GEO, de propriedade de direito e de facto, da fundação Geolíngua.⁵⁹

Tendo em conta que a proposta principal da fundação é viabilizar o bilinguismo de uma forma democrática e de uma maneira natural, a língua portuguesa, pelo facto de entender 90% do Espanhol é a única língua do mundo que preenche este requisito.

⁵⁸ Texto adaptado da Fundação Geolíngua, pode ser consultado em www.geopress.org

⁵⁹ In http://www.agal-gz.org/pdf/20080526_pgl_tratado_tordesilhas_II.pdf

A Fundação Geolíngua possui como objectivos: 1. Social - promover acções humanitárias, formativas e de intercâmbio nas áreas da Educação, Saúde e Segurança junto dos países e comunidades de língua oficial portuguesa e espanhola, visando a valorização e continuidade dos laços históricos com esses países, numa perspectiva de progresso social; 2. Sociocultural - promover a auto-estima e o diálogo bilingue, ou seja, através da língua materna e uma segunda língua de comunicação entre os povos em geral, a partir dos países e comunidades de língua portuguesa e espanhola, tendo em vista baptizar de língua “Geolíngua” o resultado do acordo ortográfico dos 8 países de língua oficial portuguesa, visando desta forma a constituição de direito e de facto, uma “GEO-Comunidade”, real, virtual e interactiva.

Um GEO-Polícia que comunica em língua portuguesa é um polícia naturalmente bilingue e extremamente útil para poder comunicar com as 700 milhões de iberófonos que representam a metade do mundo. Sabendo que uma comunidade constitui-se pela comunhão de indivíduos e culturas sendo a segurança pública uma das formas mais depuradas de expressão popular, a institucionalização de um GEO-polícia, apresenta-se como um estandarte da Iberofonia, através de acções de endomarketing e endoeconomia.

Perante os factos, a GEO-polícia, dependerá única e exclusivamente da vontade e da cumplicidade do verdadeiro “chefe” de todos os governos e multinacionais do mundo civilizado, o eleitor e consumidor que somos todos nós.

CONCLUSÕES

Ao utilizarmos um modelo preditivo supervisionado, sabemos *a priori* que iremos ter resultados optimistas, pois estamos a trabalhar com uma variável *target* da qual já sabemos o resultado. No entanto, face aos dados disponíveis, foi o melhor método encontrado.

Os resultados do modelo preditivo foram considerados muito bons. Face ao valor real de referência, as previsões foram 97.95 % correctas. Tal resultado foi obtido utilizando uma rede neuronal com uma camada escondida e com os dados particionados em 70 % para o conjunto de treino e 30 % para o conjunto de validação.

Verificou-se também que, tanto a árvore de decisão como a rede neuronal com duas camadas escondidas, produziam resultados com um erro quadrático médio, cerca de dez vez superior em relação aos outros métodos, pelo que foram menosprezados no processo de selecção.

O facto de estarmos a utilizar o modelo com apenas 53 registos, levantou alguns problemas na capacidade do modelo em evitar as correlações espúrias e houve necessidade de algum tratamento na configuração dos parâmetros dos métodos, para evitar introduzir “ruído” que de alguma forma pudesse prejudicar os resultados obtidos.

Constatou-se que, pequenas alterações na forma como o conjunto de treino e conjunto de validação eram divididos, produziam erros quadráticos médios completamente diferentes.

Após algum “rolar” do modelo, veio a verificar-se que a rede neuronal com uma camada escondida, era o método que lidava melhor com valores elevados de *input* e o que apresentava um erro médio menor, sendo por isso o mais adequado nas previsões, no entanto, a regressão também apresentou bons resultados, sendo no entanto mais adequado para problemas lineares.

Quanto às questões de investigação inicialmente formuladas, passamos a descrever as conclusões obtidas:

- Qual a freguesia onde se regista maior ocorrência criminal *per capita*?

A freguesia de Santa Justa, é a que tem o maior índice de criminalidade, embora não seja a mais populosa, é no entanto onde se concentra a maior criminalidade especialmente o crime de furto;

Tal justifica-se por ser a zona da cidade com mais estabelecimentos de diversão nocturna (bairro alto) e pelo facto de ser muito visitada por turistas.

- Qual a correlação entre nível populacional e criminalidade?

Não foi detectado que o nível de população tivesse alguma correlação com o nível de criminalidade, de facto, as freguesias mais populosas, também têm maior número de ocorrências criminais, mas se fizermos uma análise conjunta dessa criminalidade, observamos que a dimensão geográfica e o nível populacional também são elevados, o que origina rácios baixos nos índices de criminalidade.

- Qual o tipo de crime mais praticado?

O crime de furto e roubo é o mais praticado. Os números revelam que a percentagem deste tipo de crime é esmagador em relação aos restantes. Mesmo tendo em conta que foram agregados o crime de furto e roubo, as ocorrências são elevadas. Outra característica é o facto de os maiores índices destes crimes se registarem em freguesias com baixa densidade populacional (baixa pombalina) o que prova que os delinquentes se deslocam da periferia (onde se localizam os bairros sociais e maior taxa populacional) para o centro da cidade para cometerem ilícitos criminais. A secção de turismo da PSP (que se encontra localiza no Rossio) tem um elevado número de registo de furtos, tal deve-se, sobretudo ao facto de ser uma zona frequentada por muitos turistas e local de actuação de “carteiristas”.

Houve à partida, uma restrição na elaboração deste trabalho, restrição essa que já era previsível. Foi o facto de estarmos a trabalhar com uma amostra da realidade, ou seja, de toda a criminalidade, apenas usamos, aquela que foi denunciada á PSP no Concelho de Lisboa e no período compreendido entre 1 de Janeiro e 31 de Julho de 2007.

Pelo facto de estarmos a utilizar apenas parte dos dados, não nos permite fazer extrapolações para outras regiões, por envolverem outras realidades sociais e económicas.

Apesar deste facto, veio a constatar-se que os dados eram suficientes, o problema foi a qualidade dos mesmos.

Os dados iniciais destaca-se pela má qualidade. De facto, as expectativas iniciais de fazer uma georeferenciação exacta das ocorrências criminais, tiveram mesmo de ser alteradas, porque os dados recebidos não possuíam as qualidades nem as variáveis necessárias para executar esta missão, houve por isso uma necessidade de reformular os objectivos iniciais e limitar a georeferenciação ao nível da freguesia.

Na verdade a qualidade dos dados era tão má, que perdeu-se um tempo precioso na sua adequação ao trabalho. Os dados continham inúmeros erros, omissões, discrepâncias e redundância, cada registo de ocorrência continha várias variáveis duplicadas.

Houve também necessidade de converter os dados para valores numéricos pois estes encontravam-se em formato nominal, para tal, teve de haver uma compreensão profunda, correcção e codificação dos dados.

Eram inúmeras as situações de registos sem o campo do local da ocorrência preenchido, tal foi em parte colmatado pelo local de registo (freguesia da esquadra).

Os registos da secção de turismo não puderam ser trabalhados de igual forma como os restantes, porque não possuíam o local da ocorrência e agregar esses registos à freguesia onde se encontra a secção de turismo iria enviesar a distribuição geográfica das ocorrências.

A cautela, limitamo-nos a elaborar a estatística descritiva desses dados e não foram considerados para efeitos de mapeamento e tratamento pelo sistema de informação geográfica,

Ao optar-se por agregar os dados ao nível de freguesia, trouxe uma nova visão sobre a problemática pois encontraram-se outros factores que influenciam a criminalidade, tais como os aspectos demográficos e urbanísticos que de algum modo estavam correlacionados com os índices de criminalidade.

Foi então decidido adicionar variáveis do Censo 2001, como o índice populacional e urbanístico por freguesia.

A ausência de um conhecimento sustentado sobre a realidade sócio-criminal, tem por consequência o impedimento da criação de condições para uma reflexão serena, estudada, estratégica e eficaz quanto aos resultados obtidos. Deverá observar-se como estratégias para o futuro, algumas medidas:

- Deverá ter-se em atenção as mudanças sociais, endógenas ou induzidas pelo exterior que influenciam directamente o número de ocorrências criminais e que podem contribuir para explicar certos fenómenos nas variações observadas;
- As variações da criminalidade devem ser enquadradas em ciclos mais dilatados onde se possam observar os padrões existentes;
- Deve-se distinguir, dentro da criminalidade participada, as ocorrências que devem ser tomadas como indicadores de eficiência do desempenho (positivas e negativas);
- Deve-se dar especial prioridade aos indicadores mais preocupantes;
- Qualquer variação nos índices criminais, deve ser interpretado e merecer uma reflexão de enquadramento, avaliado-se o seu significado, com recurso a modelos estatísticos e às novas tecnologias;
- As áreas de jurisdição policial deveriam estar enquadradas com as áreas administrativas, para facilitar análises geográficas e evitar conflitos de registos.

Quanto ao sistema de informação onde são registadas as ocorrências, deveria sofrer melhoramentos que permitissem controlar a qualidade e fornecer ao utilizador os dados agrupados por factos e dimensões.

Tal só será possível se a actual base de dados for convertida para uma *datawarehouse* devidamente configurada para o tipo de utilização e adaptada a outras congéneres de modo a permitir a interoperabilidade de todo o sistema:

- A base de dados de ocorrências criminais, deveria permitir filtrar o *input* dos dados e garantir assim a qualidade dos mesmos;
- Todos os utilizadores do sistema deveriam receber a necessária formação de acordo com o perfil de utilização.

É de realçar o envolvimento do governo nas prossecução de políticas de segurança interna que tenham por objectivo o melhoramento dos meios e metodologias das polícias para a redução da criminalidade:

“A modernização, o desenvolvimento social e económico, os direitos, liberdades e garantias e o exercício da cidadania plena só são possíveis com segurança para todos.” (RASI 2008:17)

“A segurança interna, a paz pública e a prevenção da criminalidade são missões absolutamente prioritárias. Eliminar os factores de insegurança, prevenir o crime e perseguir os seus autores são tarefas [obrigatórias] – tarefas da comunidade e para a comunidade, que a todos dizem respeito, a todos beneficiam e requerem uma perspectiva integrada.” (RASI 2008:17)

Um trabalho deste tipo, abre um campo inovador para o futuro. Mas ainda temos algum “caminho” a percorrer, até termos “no terreno”, polícias capazes, bem preparados e com equipamento adequado para que possam recolher informação de forma correcta, fidedigna e essencialmente útil, na elaboração de estratégias de redução da criminalidade e diminuição do sentimento de insegurança junto dos cidadãos.

No entanto, assiste-se actualmente um levantar de consciência em relação á sinistralidade rodoviária e pretende-se fazer o levantamento geográfico para posterior detecção dos *hot-spots* a fim de tomar as devidas medidas preventivas.

Esperamos que tais medidas se alastrem e se generalizem no registo preciso e georeferenciado da criminalidade a fim de se evoluir para uma polícia moderna e mais eficaz.

Para terminar, fica o apelo do feito em 24 de Novembro de 2009 no primeiro Congresso Nacional de Segurança e Defesa:

“A elaboração de uma estratégia nacional de Segurança e Defesa não pode ser apenas um problema dos peritos. É também um assunto do povo. E é a todos os portugueses que pedimos o contributo para a optimização dos recursos existentes.”

REFERÊNCIAS BIBLIOGRÁFICAS

- ABRAHAMSON, E. & Rosenkopf, L., - ***Institutional and Competitive Bandwagons: Using Mathematical Modeling as a Tool to Explore Innovation Diffusion***, in *Academy of Management Review*, Nr. 3, 1993.
- AINSWORTH, P. - ***Offender Profiling and Crime Analysis***. Devon, UK: William Publishing, 2001.
- ALMEIDA, M. - ***Vitimação e insegurança no concelho de Lisboa***, Gabinete de Estudos e Planeamento do Ministério da Justiça, Lisboa, 1998.
- ALTER, S. - ***Information Systems: A Management Perspective***, Addison Wesley, 1992.
- ANSELIN, L. - ***Spatial Dependence in Linear Regression Models With an Application to Spatial Econometrics***. in *Handbook of Applied Economics Statistics*. Berlim, Germany: Ringer-Verlag, 1998.
- ANSELIN, L.; *et al.* - ***Spatial Analyses of Crime***. in *Criminal Justice: Vol. 4, Measurement and analysis of crime and justice*. Washington, USA : D. Duffee National Institute of Justice, 2000.
- ARMSTRONG, M. - ***Basic Linear Geostatistics***. Berlim, Germany: Springer, 1998.
- AZEVEDO, M. - ***Teses Relatórios e Trabalhos Escolares: Sugestões para Estruturação da Escrita***, 2.^a Edição, Lisboa, Universidade Católica Editora, 2001.
- BAÇÃO, F. – ***Introdução ao Data Mining***. Lisboa, ISEGI-UNL, 2006.
- BAILEY, T. - ***Interactive spatial data analysis***. Massachusetts, USA: Addison Wesley, 1995.
- BAYLEY, D. - ***Police for the future***, Oxford University Press, New York, pp. 1-15, 1994.
- BERNASCO, W. & Nieuwbeerta, P. – ***How do Residential Burglars Select Target Areas ?*** – *British Journal of Criminology* – 2005 vol. 44, 296-315
- BLOCK, R. - ***Community, Environment, and Violent Crime***. in *Criminology*. 17:1979.
- BLOCK, R. - ***Space, place and crime***, New York, USA: Criminal Justice Press, 1995.

- BOWERS, K. & Hirschfield, A. - ***Exploring links between crime and disadvantage in north-west England: an analysis using GIS***, International journal of geographical information science, 1999, vol. 13 no 2, 159-184 ISSN 1365-8816
- BRANTINGHAM, P. - ***Environmental Criminology***, London: Sage, 1981.
- BRANTINGHAM, P. - ***Mobility, notoriety and crime: A study of crime patterns in urban nodal points***. in Journal of Environmental Systems, 1982.
- BRASSEL, K. *et al.* - ***The Buffalo Crime Mapping System: Design strategy for the display and analysis of spatially referenced crime data*** – GIS laboratory, Amherst, New York, 78-85, 1977
- BRATTON, W. - ***Cutting Crime and Restoring Order: What America Can Learn from New York's Finest***, Heritage Foundation Policy and Research Analysis, Prelecção 573, 1996.
- BRATTON, W. - ***Great Expectations: How Higher Expectations for Police Departments Can Lead to a Decrease in Crime***, Washington, DC, Robert H. Langworthy Ed., 1999.
- BRATTON, W. & Knobler, P. - ***Turnaround: How America's Top Cop Reversed the Crime Epidemic***, Nova Iorque, NY, Random House, 1998.
- BROWN, L. - ***Community Policing: A Practical Guide for Policy Officials***, in Perspectives on Policing. Washington, D.C., National Institute of Justice and John F. Kennedy School of Government (Harvard University), 1989.
- CABENA, P. *et al.* – ***Discovering Data Mining, From Concept to Implementation***. New Jersey, USA: Prentice Hall PTR, (s. d.).
- CANTER, P. - ***GIS and Crime Analysis in Baltimore County***, Maryland, 2005.
- CARTER, D. & Radelet, L. – ***The Police and the Community***, prentice hall, 1999, ISBN 0-13-619677-2
- CHAINEY, S. - ***GIS and Crime Mapping***: John Wiley and Sons, 2005.
- CHAPIN, F. & R. Brail. - ***Human Activity Systems in the Metropolitan United States***. Environment and behaviour. 1(2): 1969.
- CLARKE, R. - ***Situational Crime Prevention: Its Theoretical Basis and Practical Scope***, in Crime and Justice: An Annual Review of Research, eds. Michael Tonry and Norval Morris, Chicago, IL, University of Chicago Press, 1983.
- CLIFFORD, N. & T. Hope. - ***Strategic Planning Model for Community Safety and Housing Market Renewal***. Community Safety Journal. 3: 2004.
- COHEN, L. & M. Felson. - ***Social Change and Crime Rates***. American Sociological Review. 44: 1979.

- COLLEEN, McCue. - ***Data Mining and Predictive Analysis: Intelligence Gathering and Crime Analysis.*** Oxford, UK: Butterworth-Heinemann, 2007.
- CRESWELL, J. - ***Research Design: Qualitative, Quantitative and Mixed Methods,*** Thousand Oaks, Ca: Sage, 2003.
- DAVENPORT, T. - ***Process Innovation: Reengineering Work Through Information Technology,*** Boston, Harvard Business School Press, 1993.
- DIAS, F. & Andrade, M. - ***Criminologia: O homem delinquente e a sociedade criminógena,*** Coimbra, Coimbra Editora, 1984.
- ECK, J. *et al.* - ***Mapping Crime: Understanding Hot Spots.*** USA: National Institute of Justice, 2005.
- ECK, J. *et al.* - ***Problem Solving: Problem-Oriented Policing in Newport News,*** Washington, D.C., Police Executive Research Forum. 1987.
- EISNER, M. - ***Modernization, Self-control and Lethal Violence,*** British Journal of Criminology, London, 41, 2001.
- ERICKSON, R. - ***Promoting security: the Division of expert knowledge in Policing*** in MIYAZAWA, K. & MIYAZAWA, S., *in Crime prevention in the urban community,* Kluwa Law and Taxation Publishers, Boston, pp. 11-55, 1995.
- ESBENSEN, F. & D. Huizinga. - ***Juvenile victimization and delinquency.*** Youth & Society. 22(2): 1991.
- ESTEVES, A. - ***A criminalidade na cidade de Lisboa: Uma geografia da insegurança,*** n.º4, Edições Colibri, Lisboa, 1999.
- FALLETI, F. - ***Criminalité Internationale : Planète Criminelle, Le Crime, Phénomène Social du Siècle?,*** Paris, PUF, 1998.
- FARRELLI, G. & Ken P. - ***Once Bitten, Twice Bitten: Repeat Victimization and its Implications for Crime Prevention.*** Crime Prevention Unit Series, no. 46. London, England, Home Office. 1993.
- FELSON, M & R.V. Clarke. - ***Opportunity makes the thief: practical theory for crime prevention.*** Police Research Series, no. 98. London, Home Office, Policing and Reducing Crime Unit. 1998.
- FERNANDES, A. - ***Métodos e regras para elaboração de trabalhos académicos e científicos,*** 2ª ed., Porto, Porto Editora, 1995.
- FERNANDES, F. – “Sun Tzu. A arte (e a ciência) da polícia” in *politeia, s.d.*
- FERREIRA, E. - ***Crime e insegurança em Portugal: Padrões e tendências 1985-1996,*** Celta Editora, Oeiras, 1998.

- FERRET, J. - **Évaluer La Police Dite de Proximité ? Certitudes et Incertitudes Tirées des Expériences Françaises**, in Les Cahiers de La Sécurité Intérieure, Revue Trimestrielle de Sciences Sociales, n. 53, Paris, Institut des Hautes Études de La Sécurité Intérieure, 2003.
- GEFEN, D. *et al.* - **Structural Equation Modeling and Regression: Guidelines of Research Practice**. Communications of the Association for Information Systems. 4(7). 1-78. 1990.
- GEORGES, D. - **The geography of crime and violence: a spacial and ecological perspective**: 1978.
- GIDDENS, A. - **As consequências da Modernidade**, Oeiras, 3.^a Edição, Celta Editora, 1996.
- GLOWACKI *et al.* - **Pratiques de Sécurité Urbaine**, Forum Européen pour la sécurité urbaine, Paris, p.12, 1996.
- GOLDSTEIN, H. - **Problem-Oriented Policing**, Nova Iorque, NY, McGraw-Hill, 1990.
- GOLDSTEIN, H. - **Problem-Oriented Policing: the rationale, the concept and reflections on its implementation**, Paper prepared for the Conference on Crime Control in Theory and Practice, Robinson, College, Cambridge University, July 10, Police Research Group, Home Office, England, 1995.
- GONÇALVES, L. - **Evolução Estatística do Crime em Portugal: características e tendências**, Gabinete de Estudos e Planeamento do Ministério da Justiça, Lisboa, 1999.
- GONZALES, A. *et al.* - **Mapping Crime: Understanding Hot Spots**. Washington: National Institute of Justice, 2005.
- GOOTMAN, E. - **A Police Department's Growing Allure: Crime Fighters From Around the World Visit for Tips** in The New York Times, 24 de Outubro, Nova Iorque, 2000.
- GROFF, E. & T. McEwen. - **Disaggregating the Journey to Homicide. in GIS and Crime Analysis**. F. Wang (ed.). Hershey, Pennsylvania. Idea Group. 2005.
- GROFF, E.; La Vigne & Nancy G. – **Forecasting The Future of Predictive Crime Mapping**. (n. 1.) : Crime Prevention Studies, Volume 13, pp. 29-57, 2001.
- GUINEL, J. - **Espionagem na Internet: As guerras no Ciberespaço**, Lisboa, Difusão Cultural, pg. 52, 1997.
- GURR, Ted Robert - **Historical Trends in Violent Crime: an Overview**, in Ted Robert Gurr (org)., Violence in América, Londres, Sage Publications, 1988.
- HARRIES, K. - **Mapping crime: principle and practice**. Washington, DC, U.S. National Institute of Justice. 1999.
- HARRIS, R. *et al.* - **Geodemographics GIS and Neighbourhood Targeting**. Sussex, England: John Wiley & Sons, 2005.

- HAYSLETT & McCall, K. - ***Neighbourhoods, land use, and robbery rates: a test of routine activity theory.*** Dallas, Texas. University of Texas at Dallas. Position paper. 2002.
- HENDERSON, L. - ***The Baltimore CitiStat Program: Performance and Accountability, Managing for Results Series,*** IBM Endowment for The Business of Government, 2003.
- HIRSCHFIELD, A. & Bowers, K. - ***Mapping and Analysing Crime Data – Lessons from Research and Practice,*** London, UK: Taylor & Francis, 2001.
- HOUGH, J. & P. Mayhew. -***The British Crime Survey.*** Home Office research studies. no. 76 London, England. 1983.
- HUITEMS, C. - ***E Deus criou a Internet,*** Lisboa, Publicações Dom Quixote pg. 129, 1995.
- JOÃO, Paulo - ***Biometrics Systems for Identification and Security.*** Politeia - revista do ISCPST, ano III nº 1 Jan-Jun 2006, Editora Almedina, 115-124, Set 2006, ISSN 1646-0367
- JONATHAN, J. *et al.* – ***Predicting the geo-temporal variations of crime and disorder.*** Glamorgan, UK: Elsevier, 2003.
- JONHNSON, S., *et al.* - ***Prospective Crime Mapping in Operational Context: Final Report.*** London: UCL Jill Dando Institute of Crime Science, 2007.
- KATE, B.; Shane, J. & Ken, P. – ***Prospective Hot-Spotting – The Future of Crime Mapping ?.*** in The British Journal of Criminology, vol. 44 num. 5: Oxford University Press, 2004.
- KAWACHI, K. & R. G. Wilkinson. - ***Crime: social disorganization and relative deprivation,*** Social Science & Medicine. 48: 1999.
- KELLING, G. & C. Coles. - ***Fixing Broken Windows: Restoring Order and Reducing Crime in Our Communities,*** New York, New York, Touchstone. 1998.
- KELLING, G., - ***Measuring what matters: a new way of thinking about crime and public order,*** in City Journal, Vol. 2 No. 2, Nova Iorque, NY, 1992.
- KELLING, G., & Bratton, W. - ***Declining Crime Rates: Insider's Views of the New York City Story,*** in Journal of Criminal Law and Criminology Nr. 88, Nova Iorque, 1998.
- KENNEDY, D. - ***Computer-aided police dispatching in Houston,*** Texas, John F. Kennedy School of Government case program, Cambridge, MA, Harvard University, 1990.
- KENNEDY, S. & D.R. Forde. - ***Homicide from East to West: a test of the impact of culture and economic inequality on regional trends of violent crime in Canada,*** Discussion Paper no 17. University of Alberta. Centre for Criminological Research. 1988.

- KNUTSSON, J. – ***Problem Oriented Policing - from innovation to mainstream***, crime prevention studies, volume 15. Devon, UK: Willian Publishing, 2003.
- KOVÁCS, Zsolt L. – ***Redes Neurais Artificiais: Fundamentos e Aplicações***. São Paulo: Livraria da Física, 2002. ISBN 85-88325-14-4
- KRAHN, H. & Gartrell, J. W. - ***Explaining the link between income inequality and homicide rates: A cross-national study***. Discussion Paper no 4. University of Alberta. Centre for Criminological Research. 1984.
- LAND, McCall & Cohen, L. E. - ***Structural covariates of homicide rates: Are there any invariances across time and social space?***. American Journal of Sociology. 95: 1990.
- LANGLOIS, A. & Kitchen, P. - ***Identifying and measuring dimensions of urban deprivation in Montreal: an analysis of the 1996 census data***. Urban Studies. 38(1): 119–139. 2001.
- LANGWORTHY, R. - ***Measuring What Matters: Proceedings from the Police Research Institute Meetings***, Washington, DC, National Institute of Justice, 1999.
- LAURITSEN, S. & Laub, J. H. -***The link between offending and victimization among adolescents***. Criminology. 29(2): 1991.
- LeBEAU, J. -***The journey to rape: geographic distance and the rapist's method of approaching the victim***.- Journal of Police Science and Administration. 15 (2):1987.
- LIU, H. & Brown, D. – ***Criminal Incident Prediction Using a Point-Pattern-Based Density Model***. International journal of forecasting 19, 2003, Elsevier 603-622
- LOCH, A. - ***A importância do mapeamento da criminalidade utilizando-se tecnologia de SIG para auxiliar a Segurança Pública no combate à violência***: 2002.
- LONGLEY, P. - ***GIS and Science***. John Wiley and Sons: 2005.
- LOUREIRO, M. - ***A política de Segurança Interna***, Ministério da Administração Interna, Lisboa, pp. 17-27, 1995.
- LOURENÇO, N. & Lisboa, M. - ***Violência, Criminalidade e Sentimento de Insegurança***, Separata da Revista “textos”, Lisboa, CEJ – Ministério da Justiça, N.º 2, 1998.
- MACHADO, F. & Cabral, J. - ***Segurança e Vitimação na cidade de Lisboa***, Universidade Católica Portuguesa, Observatório de Segurança de Lisboa, Lisboa, 2000.
- MANN, P. – ***Métodos de Investigação Sociológica***. Rio de Janeiro: Zahar Editores S.A., 1983.
- MANN, D. & Sutton, M. - ***Netcrime: More Changes in the Organization of Thieving***, British Journal of Criminology, Vol. 38, n.º 2, Spring, pg.202, 1998.

- MANNHEIM, Herman – **Criminologia Comparada**. London: Fundação Calouste Gulbenkian, 1985. volume 1 e 2
- MAPLE, J. - ***The Crime Fighter: Putting the Bad Guys Out of Business***, Nova Iorque, Doubleday, 1999.
- MATHIEU, J. - ***L'Insécurité : Que sais je?***, 1ª edição, Puf, Paris, pp. 9-20; 107-125, 1995.
- MATOS, João - **Fundamentos de Informação Geográfica**. Lisboa: Lidel, 2001. ISBN 972-7-57185-9
- MCDONALD, P. *et al.* - ***Managing Police Operations: Implementing the NYPD Crime Control Model Using COMPSTAT***, Belmont, CA, Wadsworth Publishing Co., 2001.
- MONET, Jean-Claude – ***Polices et Sociétés en Europe***. Paris: La Documentation Française, 1993.
- MOORE, S. - ***Investigating Crime and Deviance***, Sociology in action, Collins Educational, London, pp. 183-204, 1996.
- MORENOFF, J. & Robert J. – ***Violent crime and the spatial dynamics of neighbourhood transition: Chicago 1970-1990***. Social forces. 76, 1:1997.
- MORENOFF, J. *et al.* - ***Neighbourhood inequality, collective efficacy and the spatial dynamics of urban violence***. Research Report, no.00-451. Population Studies Centre. Institute for Social Research. University of Michigan. 2001.
- MURRAY, T. A. *et al.* – ***Exploratory spatial data analysis techniques for examining urban crime*** – British Journal of Criminology, spring 2001 vol. 41 309-329, ISSN 0007-0955
- NATH, V. S. – ***Crime pattern detection using data mining*** – oracle corporation. 41-44, 2006.
- OBERWITTLER, D. - ***Re-Balancing routine activity and social disorganization theories in the explanation of urban violence: A new approach to the analysis of spatial crime patterns based on population at risk***. Journal of Quantitative Criminology. *in the press.*(s. d.)
- OLIVEIRA, J. - ***As Políticas Públicas de Segurança e os Modelos de Policiamento***, Tese de Mestrado Inédita, ISCTE, Departamento de Sociologia, 2001.
- OSBORNE, D. & Wernicke, S. – ***Introduction to Crime Analysis: Basic Resources for Criminal Justice Practice***. Binghamton, New York: Routledge, 2003. ISBN 0-7890-1868-3
- PICCA, G. - ***La Criminologie***, Presses Universitaires de France, Paris, pp. 71-92, 1983.
- QUIVY, R. & Campenhoudt, L. - **Manual de Investigação em Ciências Sociais**, 2ª ed., Lisboa, Gradiva Publicações, 1998.

- RASCÃO, J. – **Análise Estratégica: Sistemas de Informação para a Tomada de Decisão Estratégica**. Lisboa: Edições Sílabo, 2001. ISBN 972-618-258-1
- RATCLIFFE, J. - **A temporal constraint theory to explain opportunity-based spatial offending patterns**. Journal of Research in Crime and Delinquency. 2006.
- RATCLIFFE, J. - **Aoristic Crime Analysis**. International Journal of GIS: 2000.
- RATCLIFFE, J. - **Chasing Ghosts? Police Perception of High Crime Areas**: British Journal of Criminology, 2001.
- RENGERT, F. -**The journey to crime**. in Bruinsma, G. et al. (eds.). Punishment, Places and Perpetrators: developments in criminology and criminal justice research. Portland, Oregon. William Publishing. 2004.
- RICO, J. & Salas, L. - **Inseguridad Ciudadana y Policía**, Editorial Tecnos S.A., Madrid, 1988.
- ROCHA et al. - **A (in)segurança dos lisboetas – Inquérito à população de Lisboa sobre comportamentos preventivos e percepção de segurança**, Edição não revista e policopiada, Escola Superior de Polícia, Lisboa, 1997.
- ROCHÉ, S. - **Insécurité et Libertés**. Paris: Seuil, 1994.
- ROCHÉ, S. - **Le sentiment d'insécurité : Sociologie d'aujourd'hui**, Puf, Paris, 1993.
- ROCHÉ, S. - **Sociologie politique de l'insécurité : Sociologie d'aujourd'hui**, Puf, Paris, pp. 1-11; 142-147, 1998.
- RONCEK, W. & Maier, P. A. - **Bars, blocks, and crimes revisited: linking the theory of routine activities to the empiricism of hot spots**. Criminology. 29: 1991.
- SACCO, V. - **Evaluating Satisfaction** in Brodeur, J. - How to recognize good policing: Problems and Issues, Sage Publications, London. pp. 123-137, 1998.
- SAMPSON et al. - **Spatial (Dis)Advantage and Homicide in Chicago Neighborhoods**. in Spatially Integrated Social Science. Michael F. Goodchild and Donald G. Janelle (eds.): Oxford University Press. 2004.
- SAMPSON, R. & Janet, L. - **Violent victimization and offending: Individual, situational, and community-level risk factors**. in Understanding and Preventing Violence. A.J. Reiss and J.A. Roth (eds.). Washington, D.C. National Academy Press. 1994.
- SAMPSON, R. & Steven, R. - **Systematic Social Observation of Public Spaces: A New Look at Disorder in Urban Neighbourhoods**. American Journal of Sociology. 105: 1999.

- SAMPSON, R. *et al.* - **Neighbourhoods and violent crime: a multilevel study of collective efficacy**. Science. 277: 1997.
- SAMPSON, R. - **The community context of violent crime**. in Sociology and the public agenda. W.J. Wilson (ed.). Newbury Park, California. Sage Publications. 1993.
- SAVOIE, J. - **Geocoding Crime Data: Feasibility Study on Collecting Data from Police Forces**. Canadian Center for Justice Statistics. Statistics Canada. 2005.
- SHAW, C. & McKay, H. - **Juvenile Delinquency in Urban Areas**. Chicago, Illinois. University of Chicago Press. 1942.
- SHERMAN, L. *et al.* - **Hot Spots of Predatory Crime: Routine Activities and the Criminology of Place**. Criminology. 27, 1: 1989.
- SHERMAN, L. W. - **Policing for Crime Prevention, University of Maryland**, Department of Criminology and Criminal Justice (Ed.), Preventing Crime: What Works, What Doesn't, What's Promising, Washington, DC, Office of Justice Programs, 1994.
- SILVERMAN, E. - **NYPD Battles Crime: Innovative Strategies in Policing**, Boston, Northeastern University Press, 1999.
- SKOGAN, W. - **Measuring What Matters: Crime, Disorder, Fear**, in LANGWORTHY, R. Measuring What Matters: Proceedings from the Police Research Institute Meetings, Washington, DC, National Institute of Justice, 1999.
- SOUTH, J. K. & Crowder, K. D. - **Residential mobility between cities and suburbs: Race suburbanization, and back-to-the-city moves**. Demography. 34: 1997.
- SPARROW, M. *et al.* - **Beyond 911: A New Era for Policing**, Nova Iorque, NI, Basic Books, 1992.
- TANG, Zhao & MacLennan, Jamie – **Data Mining With SQL Server 2005**. Indianapolis: Wiley Publishing, 2005. ISBN 978-0-471-26261-3
- TROJANOVIC, R. & Moore, M. - **Corporate Strategies for Policing**, in Perspectives on Policing, Washington, DC, National Institute of Justice, 1988.
- TROJANOWICZ, R. & Bucqueroux, B. - **Community Policing: a Contemporary Perspective**, Cincinnati, OH, Anderson Publishing Co, 1990.
- TURNER, S. - **Delinquency and distance**. in M. E. Wolfgang & T. Sellin (eds.), Delinquency: Selected Studies. New York. John Wiley & Sons. 1969.
- VALENTE, M. – **Regime Jurídico da Investigação Criminal: comentado e anotado** – 3ª edição revista e aumentada, editora Almedina, Coimbra, Abril 2006, ISBN 972-40-2802-X

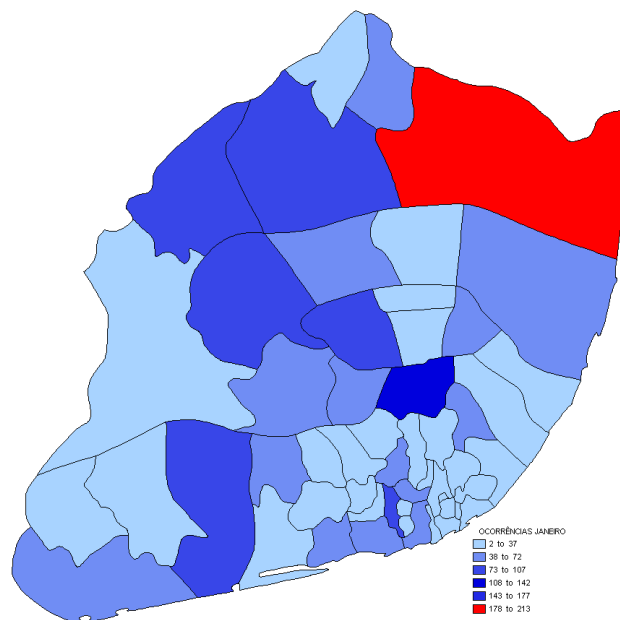
- VAUJOUR, J. - ***La sécurité du citoyen : Que sais je?***, Puf, Paris, pp. 3-21; 120-126, 1980.
- VARAJÃO, J. – **A Arquitectura da Gestão de Sistemas de Informação**. Lisboa: FCA editora, 1998. ISBN 972-722-140-8
- VARAJÃO, J. & Amaral, L. – **Planeamento de Sistemas de Informação**. Lisboa: FCA editora, 2000. ISBN 972-722-193-9
- VIEIRA, S. - **Introdução à BioEstatística**, 4ª edição, Editora Campus, Rio de Janeiro, pp. 51-74, 1996.
- VLOKOFF, V. - **Pequena história da Desinformação: Do Cavalo de Tróia à Internet**, Notícias Editorial, Lisboa, 2000.
- WALLER, I. - ***The Police: first in aid?*** in Lurigio *et al.* - Victims of crime: problems, policies and programs, vol. 25, Sage Publications, London, pp. 139-156, 1990.
- WANG, F. - ***GIS and Crime Analysis***, London, UK: Idea Group Publishing, 2004.
- WEISBURD, D., *et al.* - ***Reforming to Preserve: CompStat and Strategic Problem-Solving in American Policing***, in Criminology and Public Policy 2, Washington, DC, 2002.
- WIKSTROM, P. - ***Urban Crime, Criminals and Victims***. New York, New York. Springer. 1991.
- WILES, P. & Costello, A. - ***The road to nowhere: the evidence for travelling criminals***. Home Office Research Study, no 207. Research, London, England. Home Office, Development and Statistics Directorate. 2000.
- WILSON, J. & Kelling, G. - ***La Police et La Sécurité du Voisinage : Les Vitres Cassées***, in Normandeau, André (org.), Une Police Professionnelle de Type Communautaire, Montreal, Méridien, Tome I. 1998.
- WOSTPHAL, Christopher & Blaxton, Teresa – ***Data Mining Solutions: Methods and tools for solving real-world problems***. New York: Wiley Publishing, 1998.

APÊNDICES

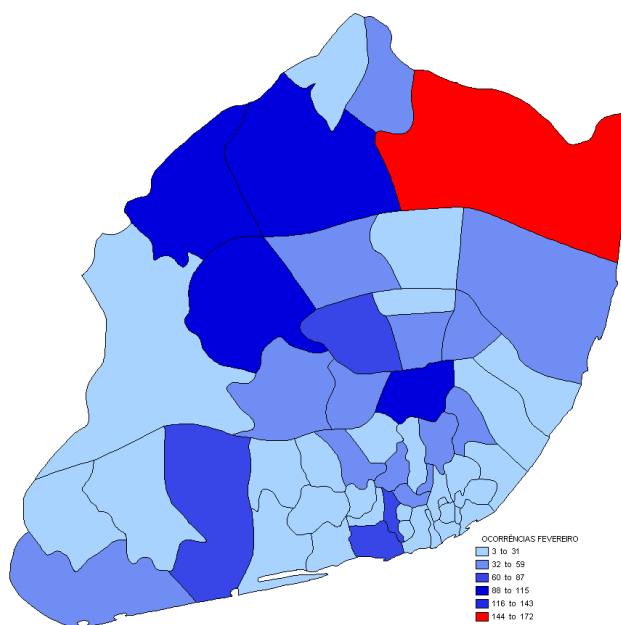
Apêndice 1 - Cronograma de actividades

Actividades	2008						2009												
	J	A	S	O	N	D	J	F	M	A	M	J	J	A	S	O	N	D	
Revisão bibliográfica	▶	▶	▶	▶	▶	▶	▶	▶	▶	▶	▶	▶	▶	▶	▶	▶	▶		
Revisão crítica da literatura	▶	▶	▶	▶	▶	▶	▶	▶	▶	▶	▶	▶	▶	▶	▶	▶	▶		
Aplicação norma NP405:1-4	▶	▶	▶	▶	▶	▶	▶	▶	▶	▶	▶	▶	▶	▶	▶	▶	▶		
Obtenção dados 2007	▶	▶																	
Preparação guião entrevista		▶																	
Recolha e pré-processamento dados			▶	▶	▶	▶					▶	▶	▶	▶	▶				
Pedido e elaboração prefácio			▶																
Entrevista semi-estruturada			▶																
Reunião com orientador			▶					▶			▶					▶	▶	▶	
Pedido de <i>software</i> específico							▶	▶	▶										
Licenças de software										▶	▶	▶	▶	▶	▶				
Preparação do modelo preditivo SAS										▶	▶	▶	▶	▶	▶	▶	▶	▶	
Conclusões preliminares												▶	▶	▶	▶	▶	▶		
Análise descritiva dos dados recolhidos																			
Conclusões e recomendações													▶	▶	▶	▶	▶		
Entrega orientadores										▶				▶				▶	
Correcção final															▶	▶	▶		
Entrega final ISEGI																		▶	
Verificação e <i>errata</i>																			▶
Preparação apresentação																			▶

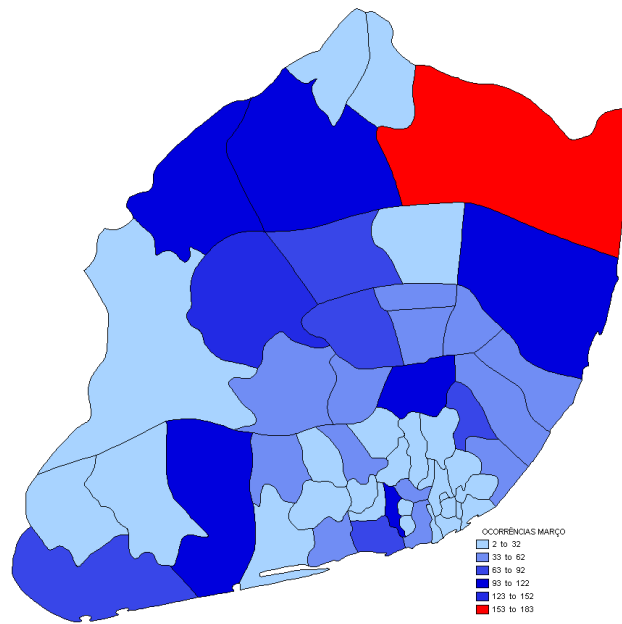
Apêndice 2 – Mapas Temáticos



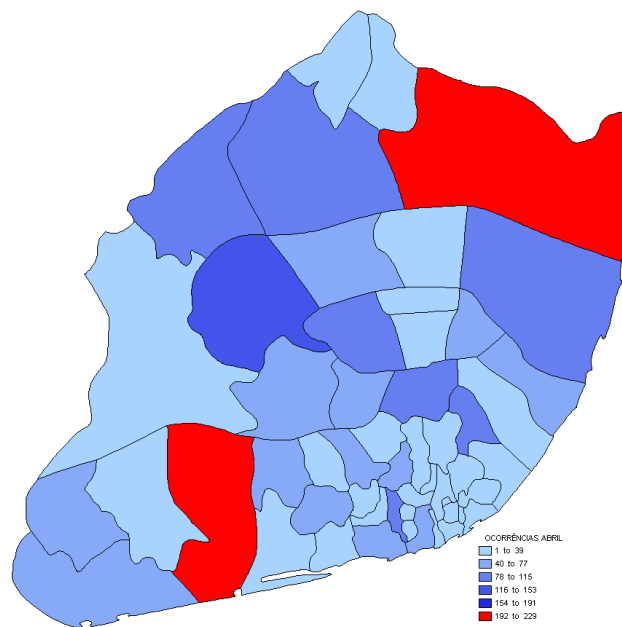
Freguesia com maior densidade de crimes em Janeiro:
Santa Maria dos Olivais (213).



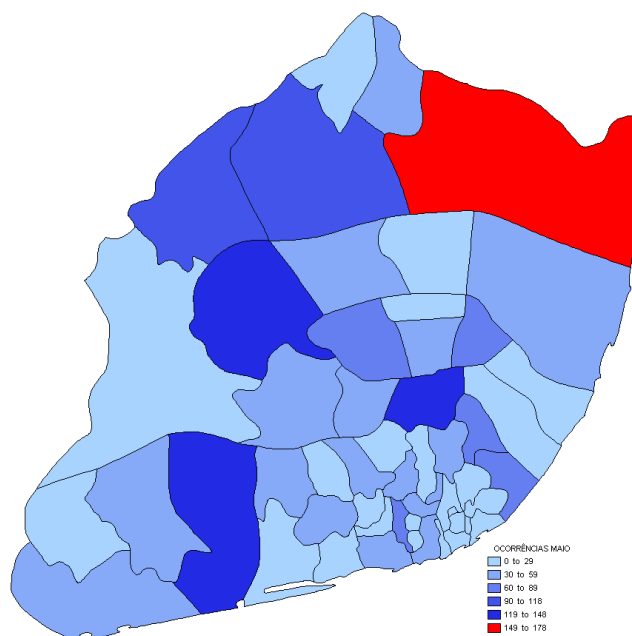
Freguesia com maior densidade de crimes em Fevereiro:
Santa Maria dos Olivais (172).



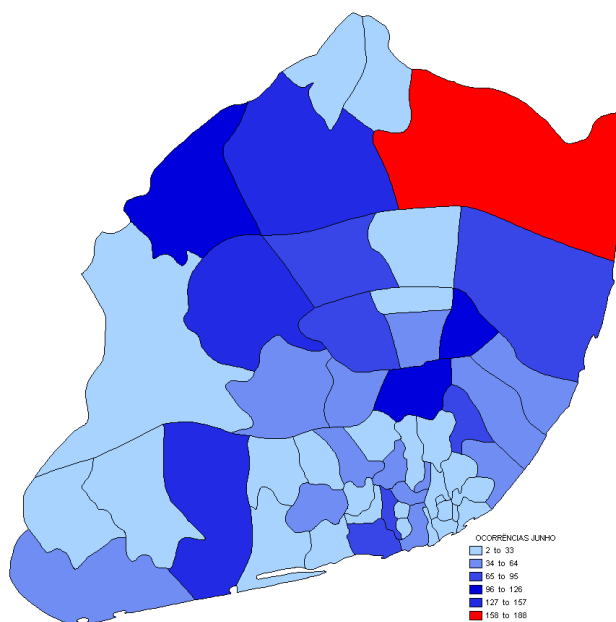
Freguesia com maior densidade de crimes em Março:
Santa Maria dos Olivais (183).



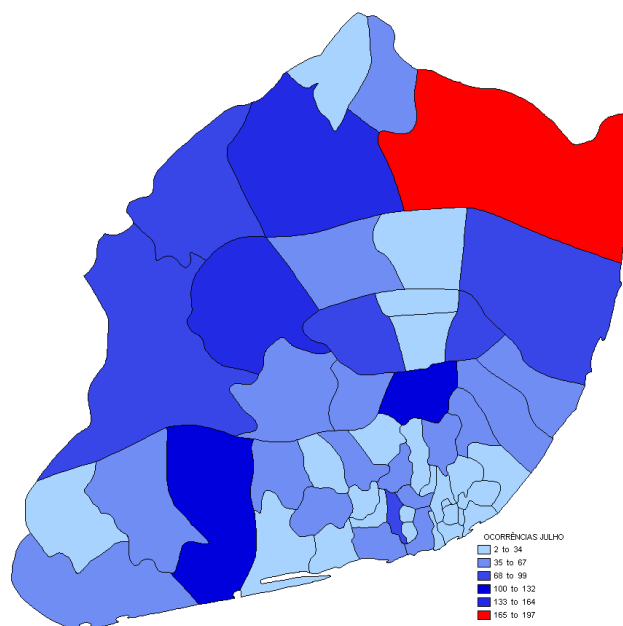
Freguesias com maior densidade de crimes em Abril:
Santa Maria dos Olivais (199) e Alcântara (229).



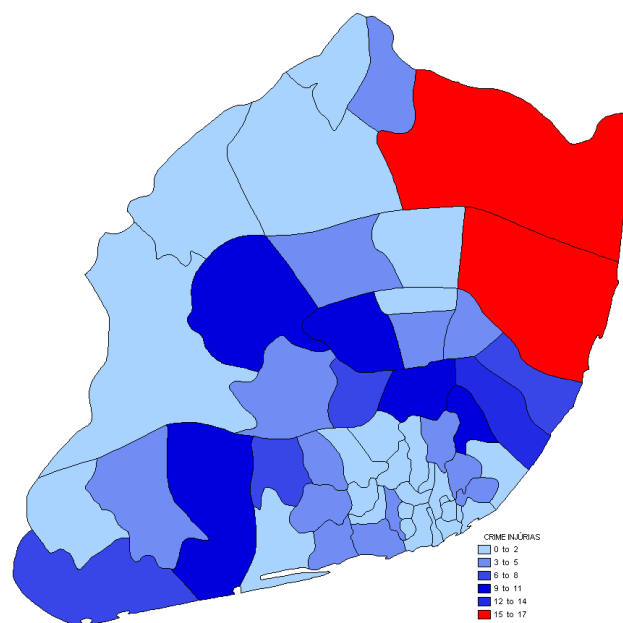
Freguesia com maior densidade de crime em Maio:
 Santa Maria dos Olivais (178).



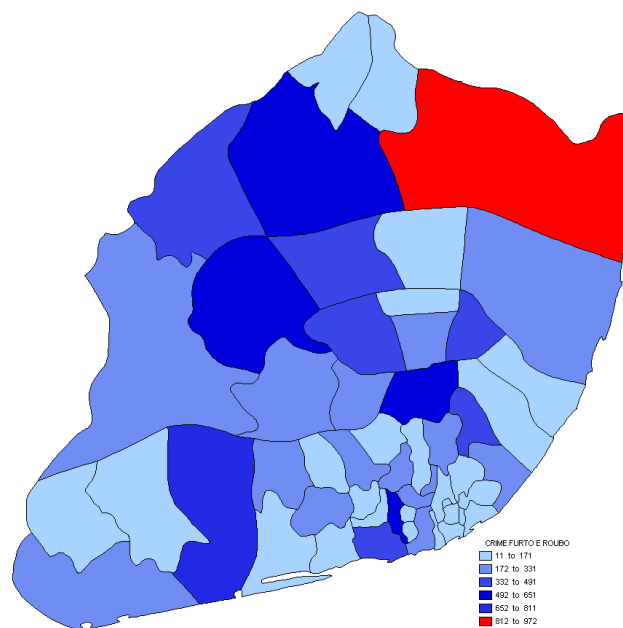
Freguesia com maior densidade de crimes em Junho:
 Santa Maria dos Olivais (188).



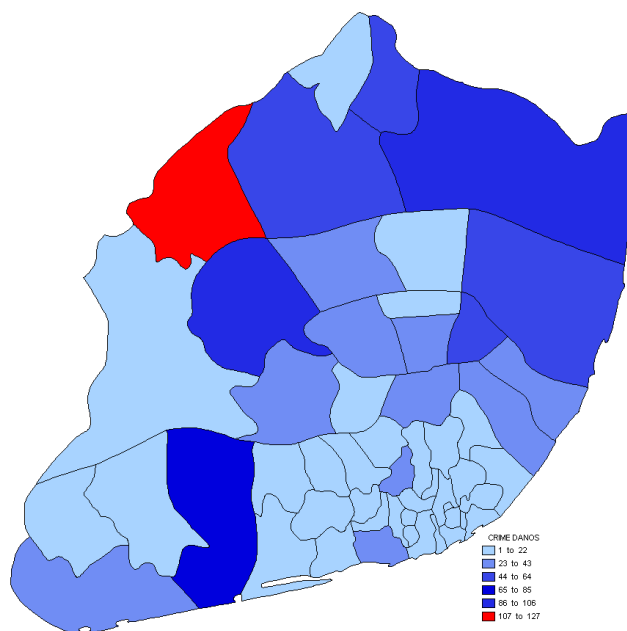
Freguesia com maior densidade de crimes em Julho:
Santa Maria dos Olivais (197).



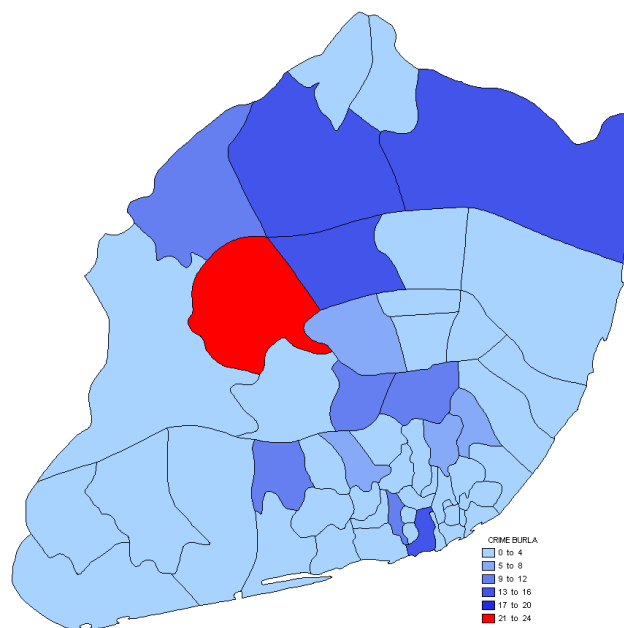
Freguesias com maior densidade de crimes de injúria:
Santa Maria dos Olivais (17) e Marvila (15).



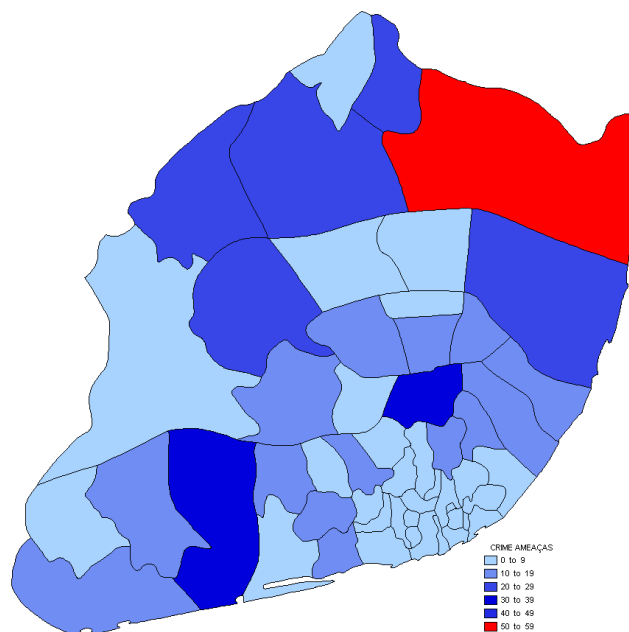
Freguesia com maior densidade de crimes de furto e roubo:
Santa Maria dos Olivais (972).



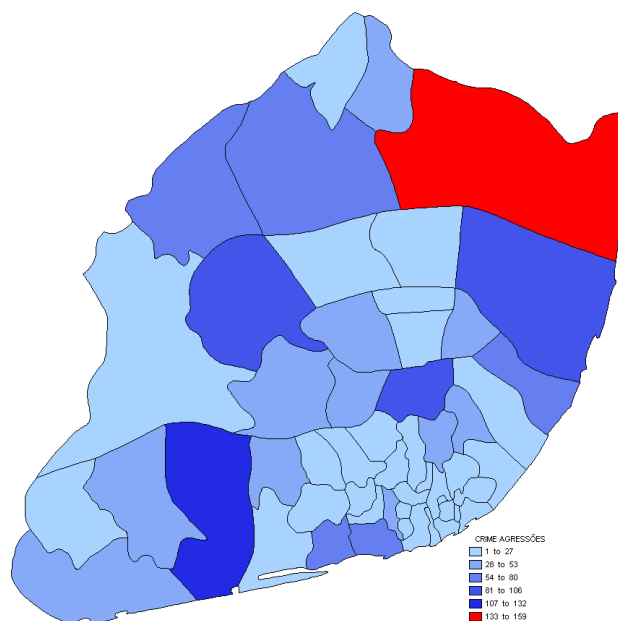
Freguesia com maior densidade do crime de danos:
Carnide (127).



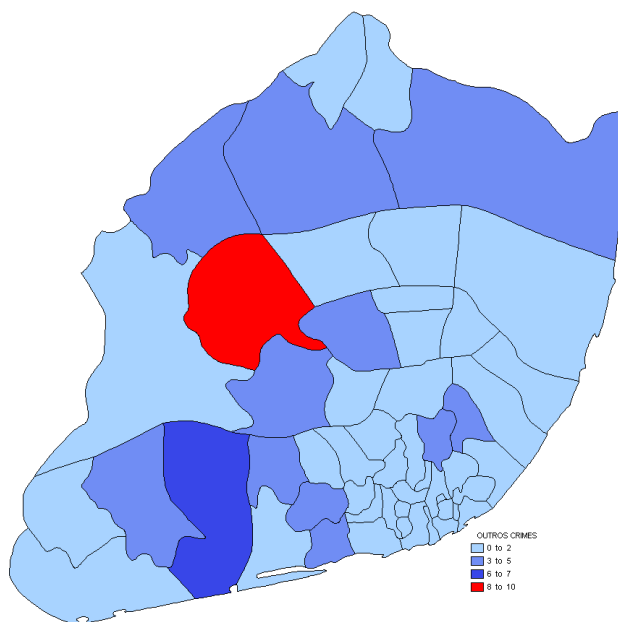
Freguesia com maior densidade do crime de burla:
 São Domingos de Benfica (24).



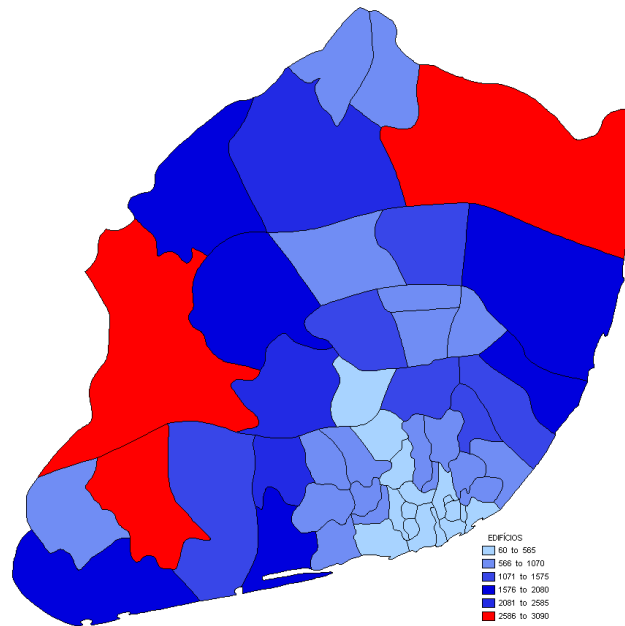
Freguesia com maior densidade do crime de ameaças:
 Santa Maria dos Olivais (59).



Freguesia com maior densidade do crime de agressões:
 Santa Maria dos Olivais (159).

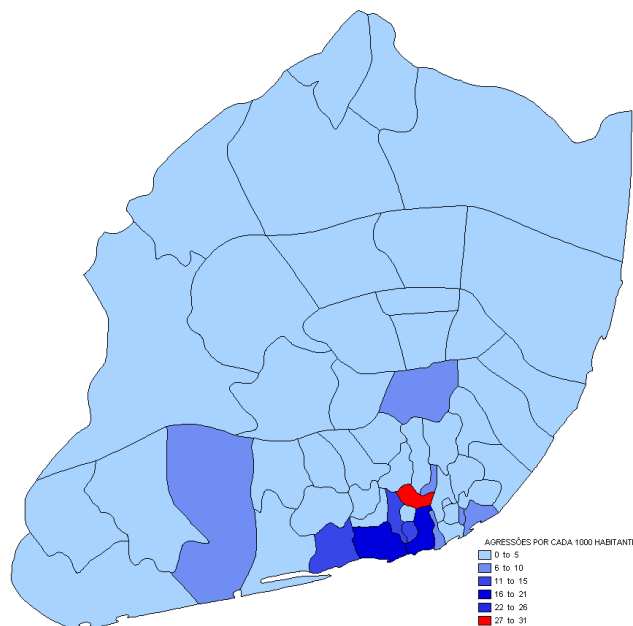


Freguesia com maior densidade de outros crimes:
 São Domingos de Benfica (10).



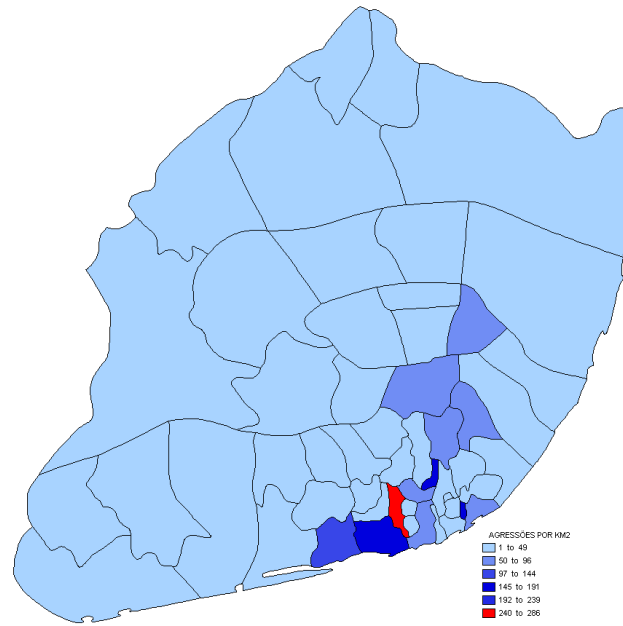
Freguesias com maior densidade de edifícios:

Santa Maria dos Olivais (3.055), Benfica (2.826) e Ajuda (3.090).

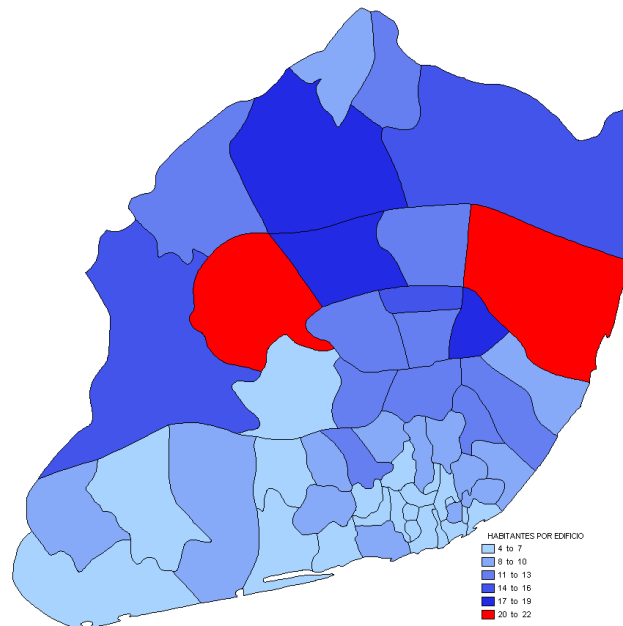


Freguesia com maior densidade do crime de agressões por cada mil habitantes:

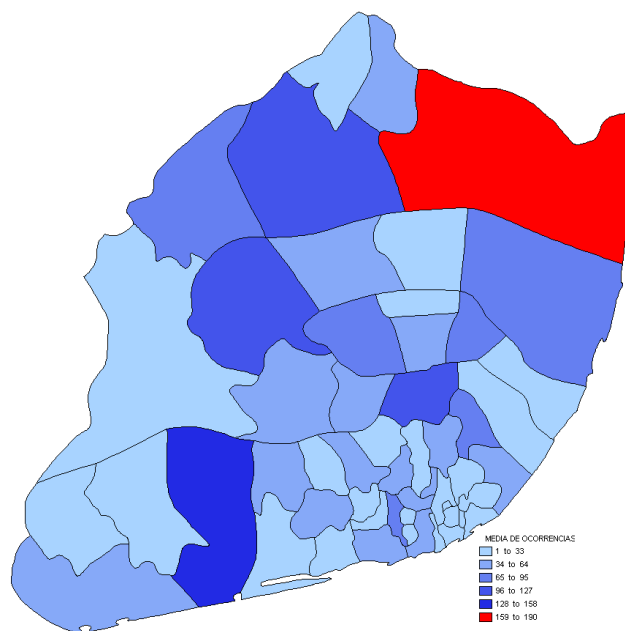
Santa Justa (31,43).



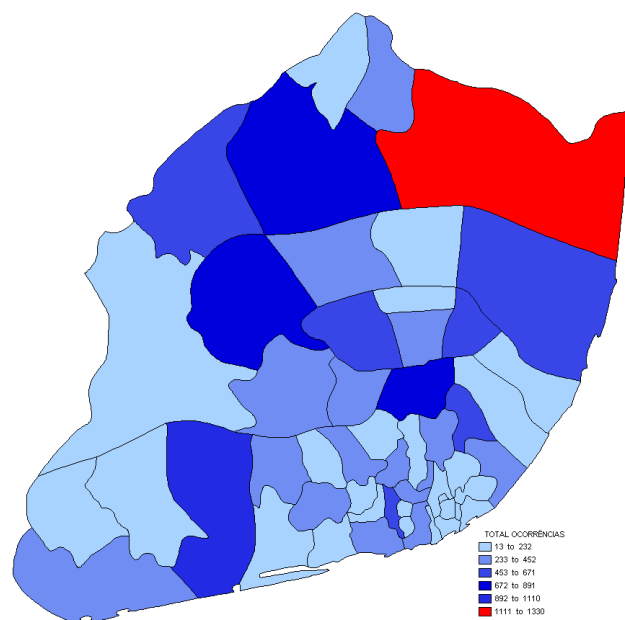
Freguesia com maior densidade do crime de agressões pela área:
Encarnação (287).



Freguesias com maior densidade populacional por edifício:
Marvila (22) e São Domingos de Benfica (20).



Freguesia com maior densidade da média mensal de ocorrências:
Santa Maria dos Olivais (190).

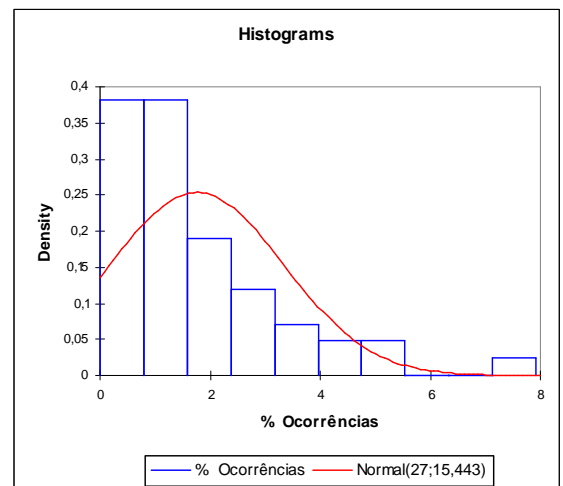
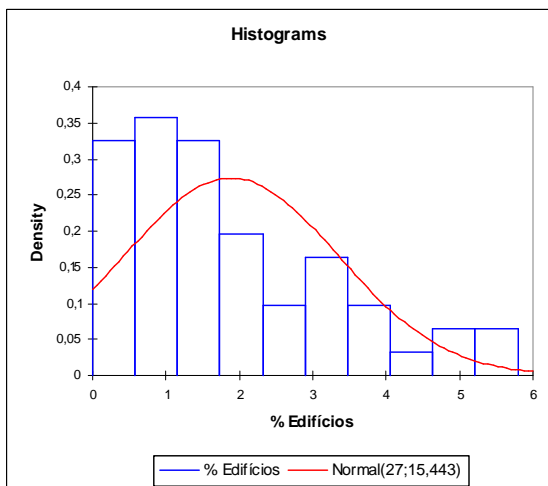
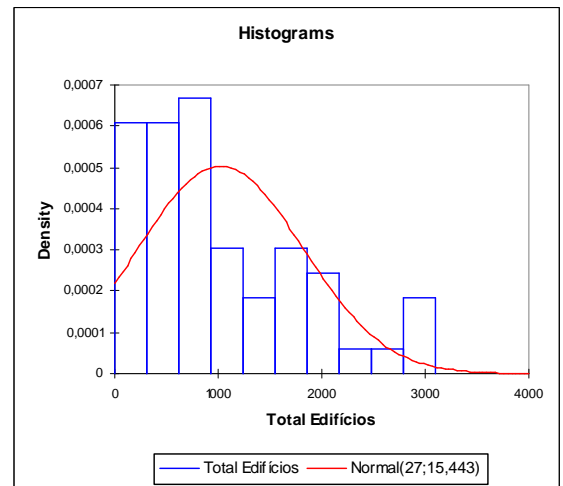
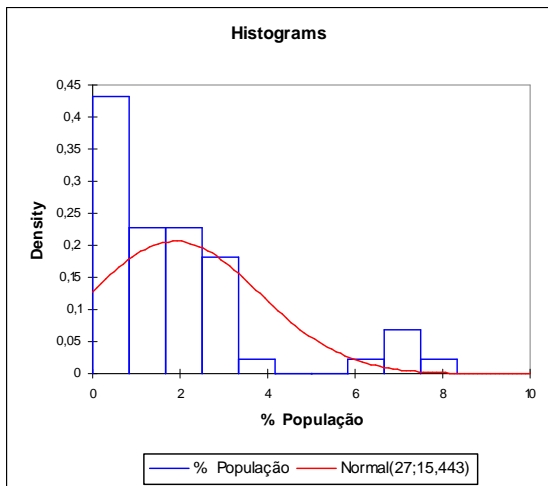
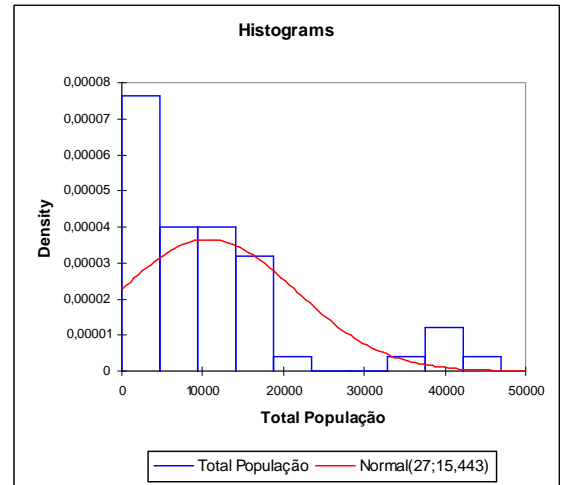
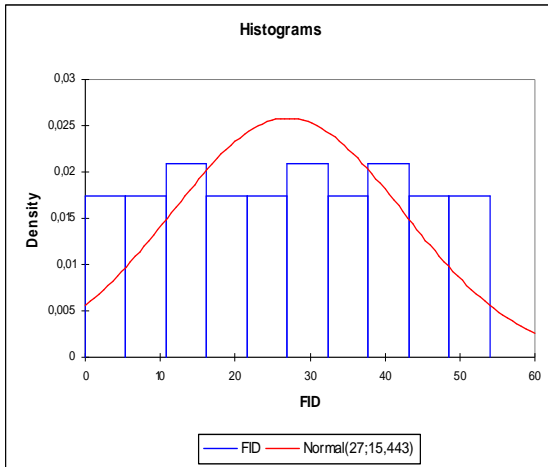


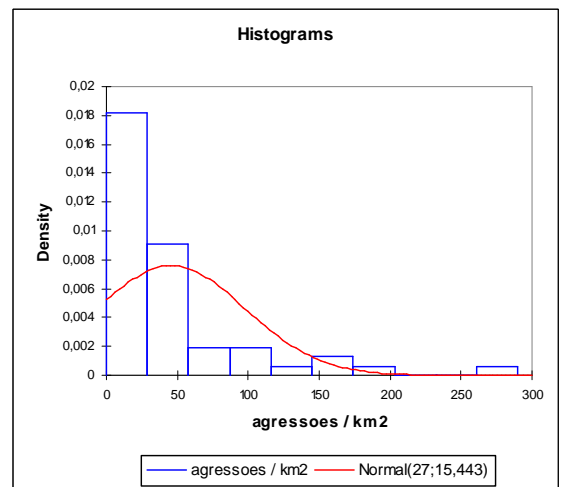
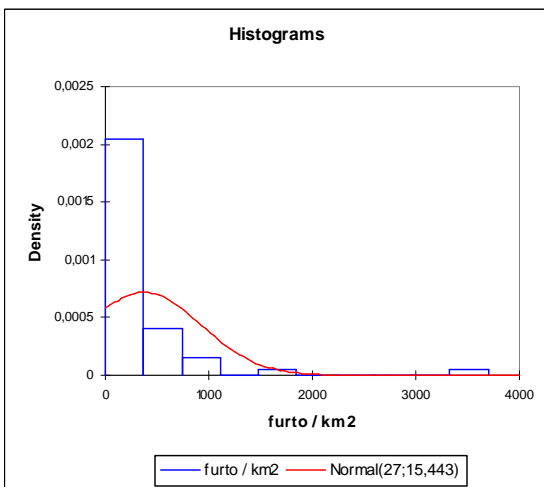
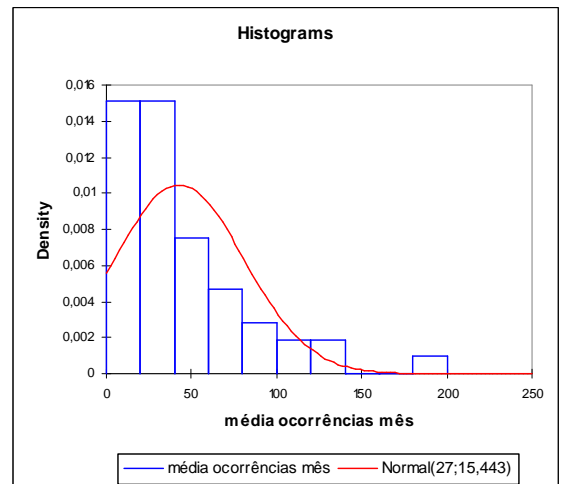
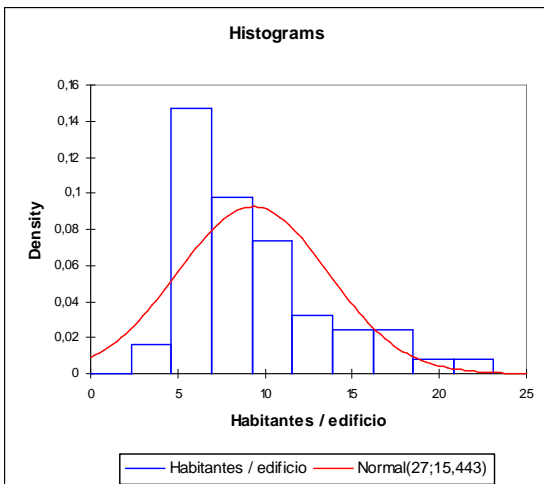
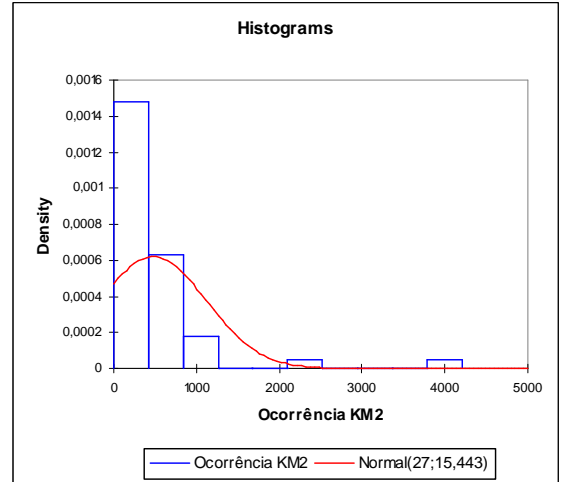
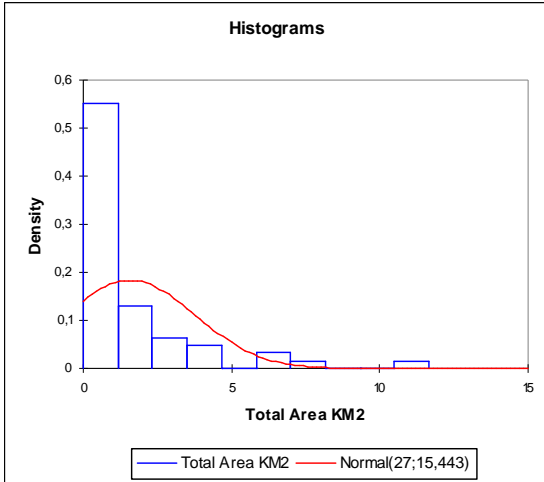
Freguesia com maior densidade de ocorrências:
Santa Maria dos Olivais (1.330).

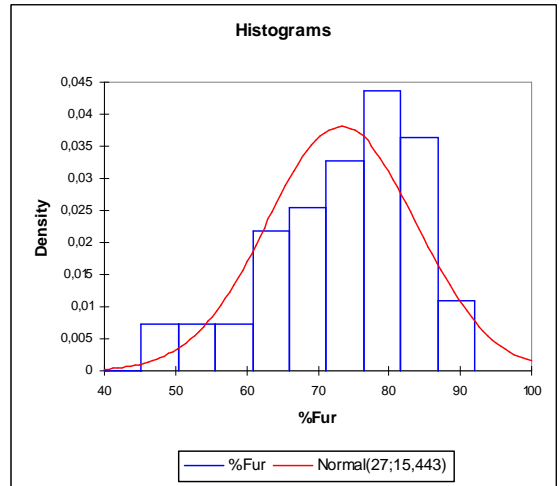
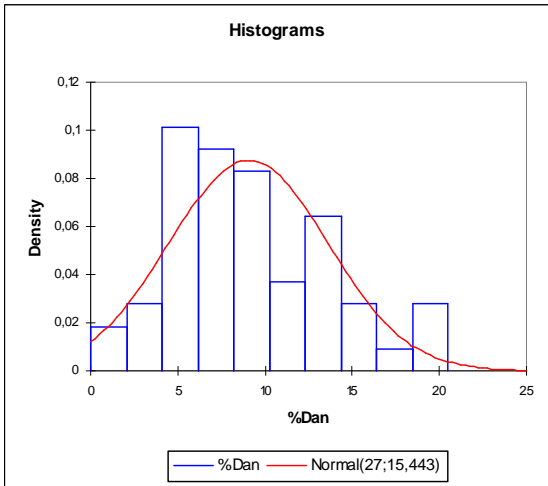
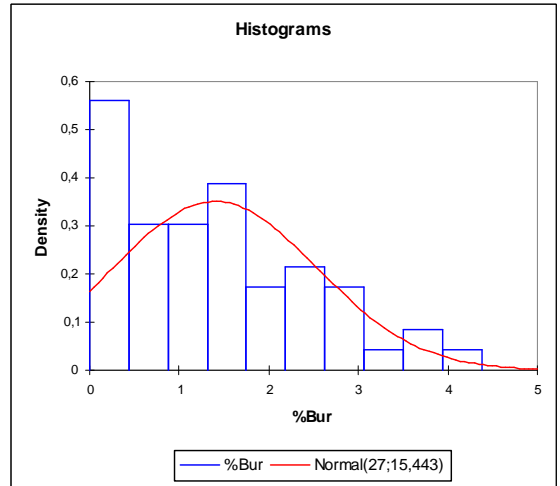
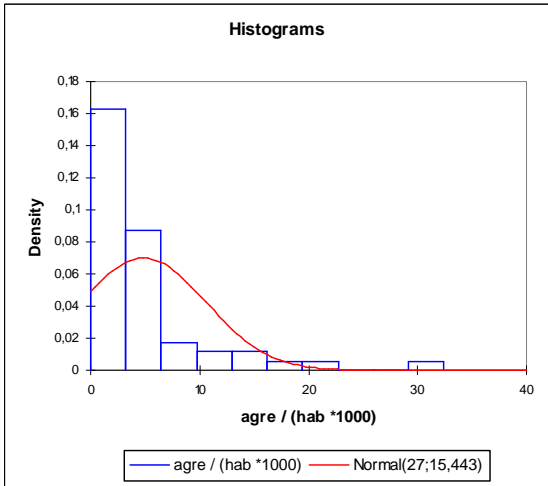
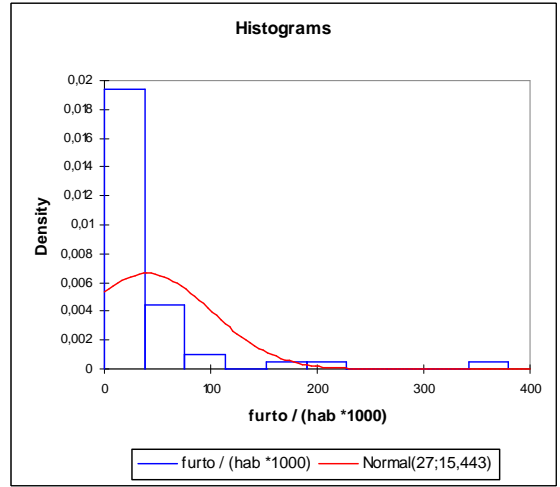
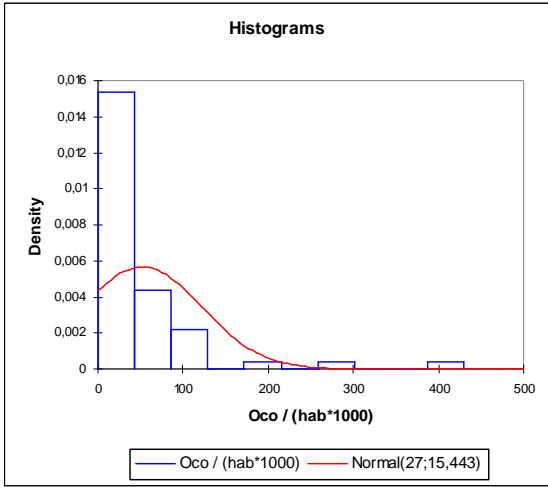
Apêndice 3 – Tabela com os dados agregados

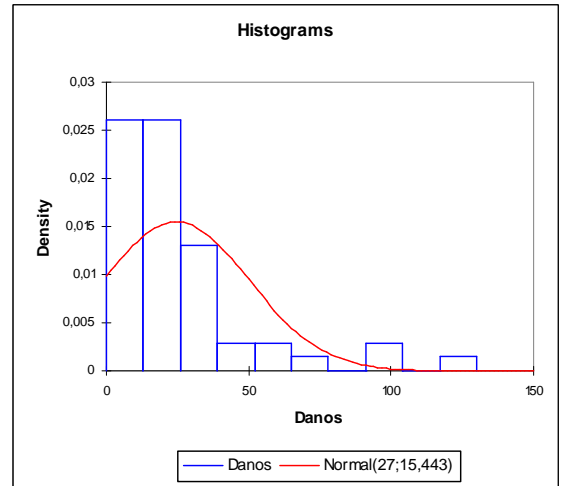
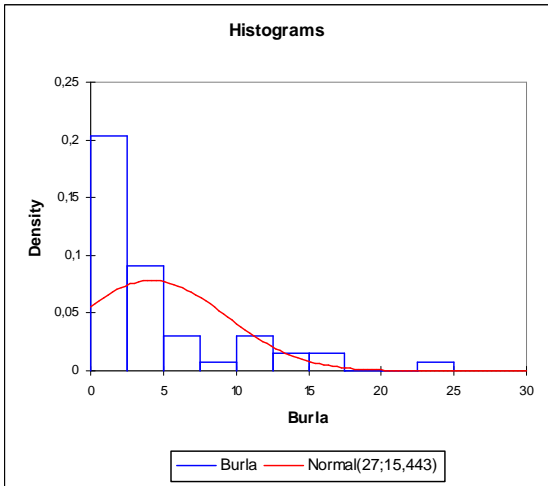
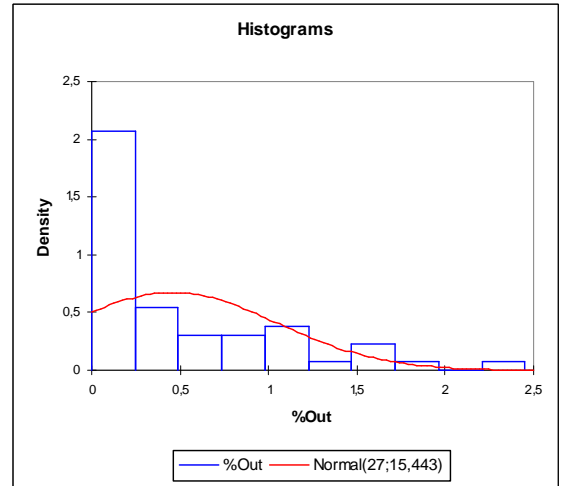
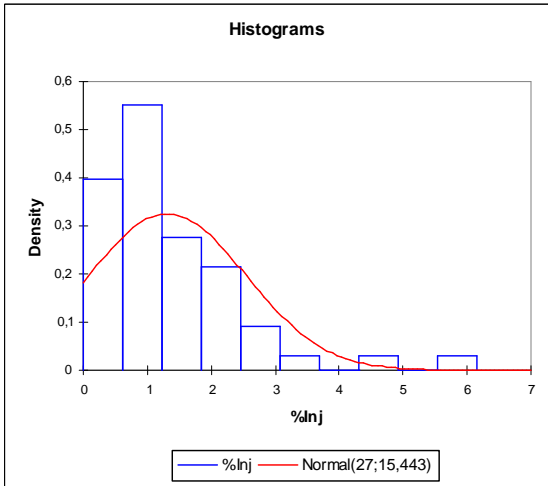
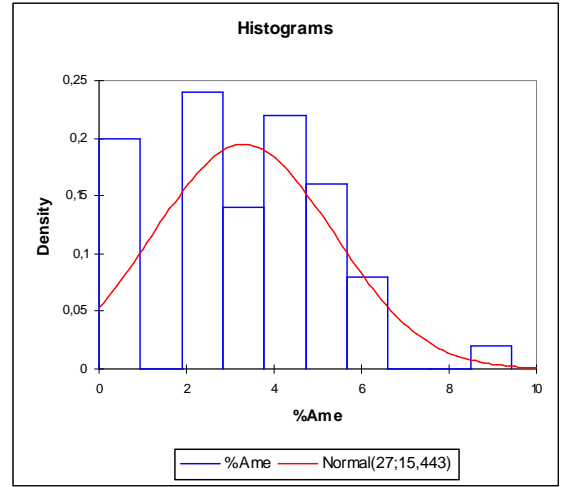
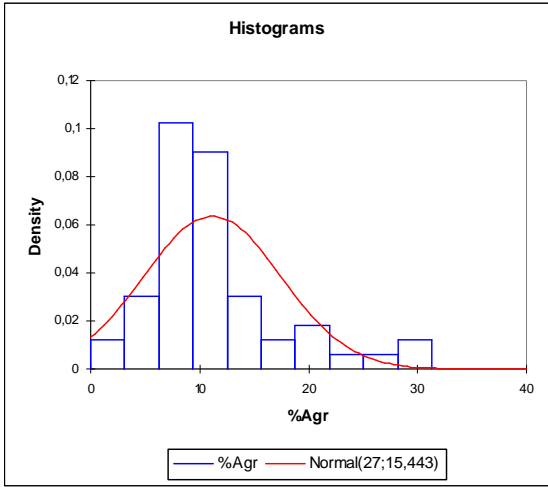
FD	Freguesia	% População	% Edifícios	Total Ocorrências	% Ocorrências	Total Area KM2	Habitantes KM2	Ocorrência KM2	Habitantes / edifício	média ocorrências mês	rturo / km2	agressões / km2	Occ / (hab / 1000)	rturo / (hab / 1000)	aggr / (hab / 1000)	% Bur	% Dan	% Fur	% Agr	% Ame	% Inj	% Out	Burta	Danos	Furto	Agressões	Armaças	Injúrias	Outro	Jan	Fev	Mar	Abr	Mai	Jun	Jul			
1	Ajudá	17568	318	300	589	197	1,16	3,15	5701	63	6	28	35	15	10,97	6,07	2,62	0,5	10,2	55,3	23,9	6,6	2,0	1,5	1	20	109	47	13	4	3	9	26	24	21	43	27	43	
2	Alcantara	14443	256	1575	290	894	4,39	3290	204	9	128	151	24	6,90	45,90	7,6	7,6	74,2	12,0	4,0	0,8	4	88	663	107	36	9	7	95	84	101	229	124	142	119				
3	Alto do Pina	10253	170	633	115	455	2,66	882	12504	555	18	23	401	50	44,38	32,97	4,00	0,9	12,5	73,6	9,0	2,9	1,1	0,0	4	57	335	41	13	5	0	40	36	60	57	75	103	84	
4	Alvalade	9620	171	633	117	455	2,66	882	12504	555	18	23	401	14	16,94	13,93	8,83	2,5	6,1	82,2	4,9	2,2	0,6	4	2	1	10	134	8	4	2	1	16	24	38	23	21	15	25
5	Arcozelo	9644	171	633	117	455	2,66	882	12504	555	18	23	401	10	14,72	9,88	1,66	2,1	12,7	66,9	11,3	6,3	0,7	0,0	3	18	95	16	9	1	0	16	18	19	20	19	26	24	
6	Arjos	9738	172	1000	184	251	1,49	0,48	20288	623	10	38	358	63	25,76	17,68	3,68	2,4	8,8	69,5	12,0	5,6	1,2	6	22	172	30	14	4	3	35	45	32	31	34	28	46		
7	Beato	14241	252	1687	344	230	1,35	1,41	10100	163	8	33	79	47	16,15	7,86	4,63	1,3	13,5	48,7	28,7	4,3	3,5	0,0	3	31	112	66	10	8	0	24	29	37	45	25	35	35	
8	Benfca	41368	7,33	2626	521	222	1,31	7,94	5270	28	15	32	25	2	5,37	4,08	0,34	0,5	14,1	91,0	6,3	0,0	0,9	0,0	1	3	202	14	0	2	0	19	16	32	35	18	28	74	
9	Campo Grande	11148	1,97	627	115	422	2,48	2,44	4569	173	18	60	137	11	37,85	30,05	2,42	3,8	7,3	79,4	6,4	1,9	0,9	0,2	16	31	335	27	8	4	1	44	59	63	72	54	74	56	
10	Camplide	19989	3,36	1843	339	666	3,66	4,02	4724	163	10	94	100	19	34,55	21,70	4,00	1,5	19,4	63,8	11,6	3,8	0,3	0,6	10	172	42	76	25	2	4	83	104	95	88	98	100	87	
11	Camde	587	0,10	113	0,21	38	0,21	0,85	11740	720	5	8	600	20	61,33	51,11	1,70	0,0	13,9	83,3	2,8	0,0	0,0	0	5	30	1	0	0	0	2	5	9	4	2	12	2		
12	Castelo	10519	1,86	876	1,61	246	1,45	1,70	6162	145	12	35	78	21	23,41	12,86	3,43	0,0	19,5	54,1	14,6	9,3	1,6	0,8	0	49	133	36	23	4	2	42	32	29	31	48	26	38	
13	Charneca de Jesus	4319	0,76	532	0,98	113	0,65	0,54	7988	209	8	16	163	28	26,16	20,38	3,47	0,0	6,2	77,9	13,3	2,7	0,0	0	7	88	15	3	0	0	21	15	18	21	11	17	10		
14	Encarnação	3182	0,56	464	0,85	629	3,70	0,15	21213	4193	7	90	3827	287	19,67	170,96	13,51	1,7	3,3	86,5	6,8	0,6	0,3	11	21	544	43	4	4	2	88	84	95	88	82	95	97		
15	Enfca	6900	1,23	704	1,30	150	0,88	0,34	20471	441	10	21	344	29	21,55	16,81	1,44	0,0	12,0	76,0	6,7	0,7	0,0	0	18	117	10	1	4	0	22	11	21	22	24	28	22		
16	Graça	8670	1,54	1008	1,86	260	1,53	0,72	13042	361	9	37	296	25	39,99	24,57	2,08	0,4	3,5	81,9	5,9	5,0	1,2	1,2	1	9	213	18	13	3	3	34	24	31	52	38	42	38	
17	Lagoa	37693	6,68	2088	3,85	777	4,57	6,28	6002	124	18	111	98	9	20,61	16,42	1,49	1,7	7,7	79,7	8,9	3,5	0,5	0,5	13	80	619	54	27	0	4	88	103	111	104	103	130	138	
18	Madalena	3807	0,7	74	0,14	42	0,25	0,11	3465	382	5	6	236	36	10,63	68,42	10,53	2,4	19,0	61,9	9,5	4,8	2,4	0,0	1	6	26	4	2	1	0	8	8	4	2	6	8	6	
19	Madrices	341	0,06	60	0,11	39	0,23	0,10	3410	390	6	6	280	40	114,37	76,25	11,73	2,6	15,4	66,7	10,3	5,1	0,0	0,0	1	6	26	4	2	0	0	4	4	5	6	8	4	8	
20	Manilá	38767	6,87	1753	3,23	623	3,08	6,29	6163	83	21	75	62	16	13,49	8,46	2,88	0,2	9,2	62,7	19,1	5,5	2,9	0,4	1	48	328	100	29	15	2	48	57	111	79	69	90	79	
21	Menes	5093	0,91	884	1,63	120	0,71	0,30	18977	470	6	18	280	40	23,56	16,49	2,36	3,3	12,5	70,0	10,0	4,2	0,0	0,0	4	15	84	12	5	0	0	23	21	21	17	10	18	10	
22	Nossa Senhora de Fatima	15391	2,71	1184	2,18	560	3,29	1,87	8177	298	13	80	243	22	36,62	26,88	2,15	1,3	5,0	81,1	7,5	2,7	1,6	0,9	7	28	454	42	15	9	5	78	71	77	104	85	69	76	
23	Pena	8668	1,07	797	1,47	94	0,55	0,49	12584	192	8	13	135	31	15,49	10,88	2,47	2,1	7,0	70,2	16,0	2,1	2,1	0,0	2	7	86	15	2	2	0	22	17	12	13	9	2	19	
24	Peña de França	13722	2,43	1321	2,43	601	2,95	0,86	20791	759	10	72	594	79	36,51	26,57	3,79	1,0	4,2	78,2	10,8	3,8	1,8	0,6	5	21	392	62	19	9	3	52	40	83	88	80	92	86	
25	Prazeres	890	0,16	132	0,24	48	0,28	0,08	11000	600	7	538	13	33	22,79	16,42	1,72	1,1	17,2	72,0	7,5	0,0	2,2	0,0	0	3	43	1	0	0	0	0	13	16	20	11	15	11	8
26	Sacramento	4091	0,72	652	1,20	93	0,56	0,21	19433	443	6	13	319	33	22,79	16,42	1,72	1,1	17,2	72,0	7,5	0,0	2,2	0,0	0	1	16	67	7	0	0	0	9	7	3	10	1	10	8
27	Santa Catarina	5680	1,04	657	1,21	255	1,50	0,57	10281	447	9	38	363	42	43,62	36,32	4,10	0,8	5,1	81,2	9,4	2,4	0,6	0,4	2	13	207	24	6	2	1	25	15	38	36	62	48	30	
28	Santa Engracia	7270	1,29	633	1,53	127	0,75	0,62	11726	205	9	18	123	44	17,47	10,45	3,71	1,6	12,6	59,8	21,3	2,4	2,4	0,0	2	16	76	27	3	3	0	16	16	28	22	10	22	11	
29	Santa Isabel	700	0,12	148	0,27	299	1,76	0,24	2917	146	5	43	1086	92	42,74	37,51	31,43	1,3	20,8	68,0	7,4	0,7	0,3	0,3	4	6	263	22	2	1	1	45	38	43	34	38	61	40	
30	Santa Maria de Belem	9796	1,73	1901	3,50	399	2,35	3,39	2878	118	5	57	90	12	40,90	31,26	4,10	0,5	6,3	76,4	10,0	3,0	1,5	0,3	2	33	305	40	12	6	1	40	59	84	65	46	60	45	
31	Santa Maria dos Olivais	46410	8,22	3025	5,63	1330	7,82	10,66	4354	125	15	190	91	15	28,66	20,94	3,43	1,2	7,7	73,1	12,0	4,4	1,3	0,4	16	102	972	169	59	17	5	213	172	183	199	178	188	197	
32	Santiago	867	0,15	168	0,31	13	0,08	0,06	14283	217	5	2	183	17	15,17	12,84	1,17	0,0	7,7	84,6	7,7	0,0	0,0	0,0	0	1	11	1	0	0	0	2	3	2	1	0	2	3	
33	Santo Condestavel	17553	3,11	2554	4,70	298	1,76	1,01	17299	296	7	43	215	36	17,03	12,36	1,98	3,0	5,0	72,6	11,7	4,0	2,3	1,3	9	15	217	35	12	7	4	40	29	49	56	53	29	43	
34	Santo Estevão	2047	0,36	398	0,66	88	0,52	0,18	11372	489	6	13	306	83	42,99	26,97	7,33	2,3	12,5	62,5	17,0	3,4	2,3	0,0	2	11	55	15	3	2	0	6	15	16	10	7	24	10	
35	Santos-o-Velho	4013	0,71	695	1,28	222	1,31	0,51	7889	435	6	32	251	122	55,32	31,90	15,45	0,0	5,9	57,7	27,9	5,4	1,8	0	13	128	62	12	3	4	41	28	43	34	23	28	25		
36	Sao Cristovao e Sao Louren	1612	0,29	209	0,38	31	0,18	0,08	20190	398	8	4	325	25	19,23	16,13	1,24	0,0	9,7	83,9	6,5	0,0	0,0	0,0	0	3	26	2	0	0	0	5	8	5	5	1	2	5	
37	Sao Domingos de Benfca	33678	5,93	1657	3,05	880	5,18	4,30	7832	205	20	126	146	21	26,13	18,68	2,64	2,1	10,8	71,5	10,1	2,7	1,0	1,1	24	95	629	89	24	9	10	91	98	144	147	127	138	135	
38	Sao Francisco Xavier	8101	1,43	1032	1,90	143	0,84	2,10	3968	68	8	20	89	3	17,65	15,66	0,74	2,1	4,2	86,3	4,2	0,0	0,0	0,0	4	24	123	21	12	12	2	26	23	36	25	13	39	36	
39	Sao Jose	10703	3,02	1453	2,68	188	1,16	1,56	18944	127	12	28	79	13	11,80	7,20	1,23	2,0	12,1	62,1	10,6	6,1	6,1	1,0	4	24	123	21	12	12	2	26	23	36	25	13	39		

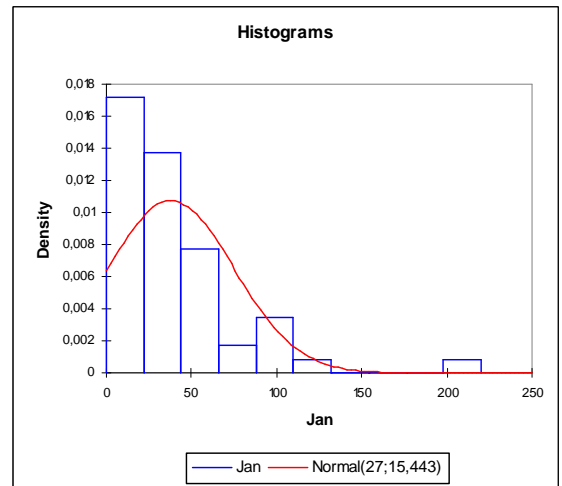
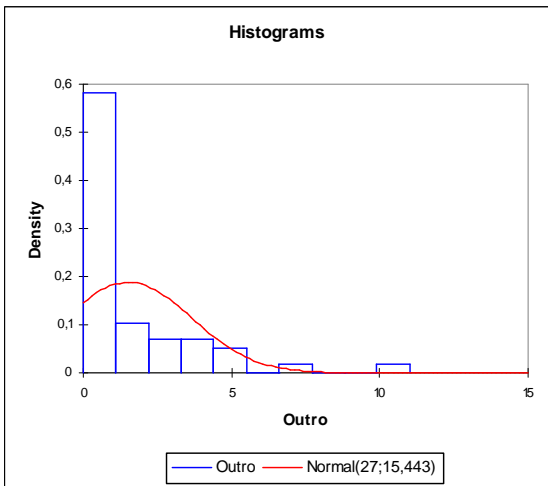
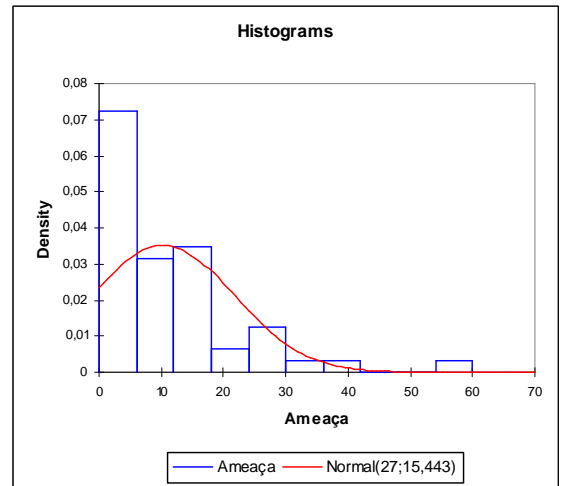
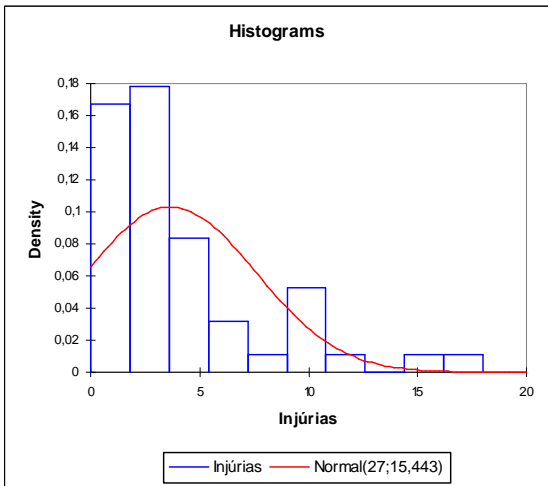
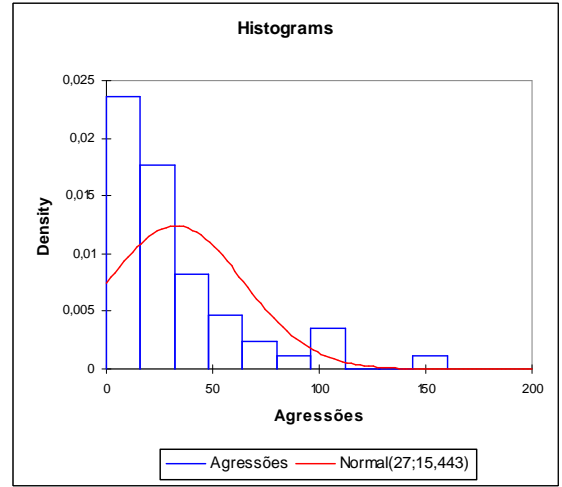
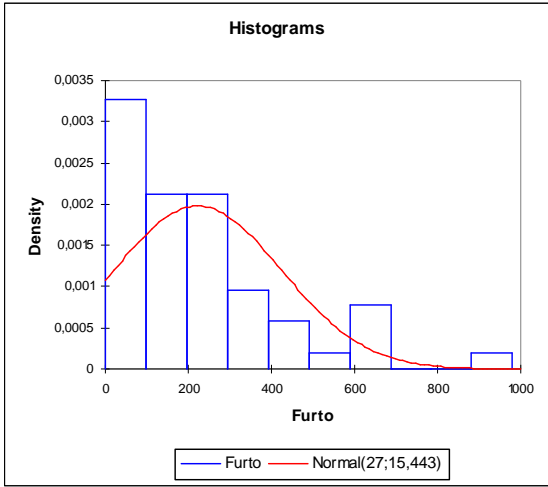
Apêndice 4 –Histogramas das variáveis

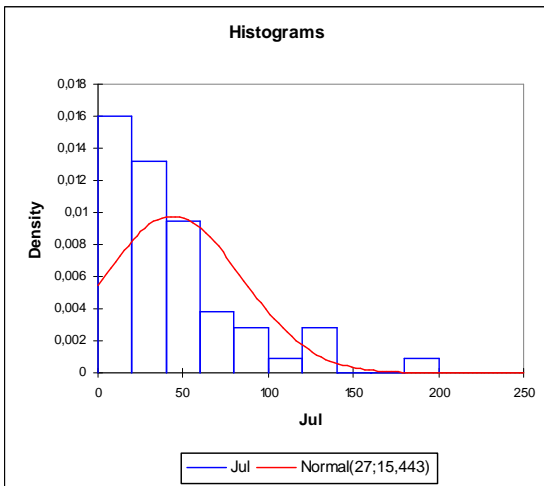
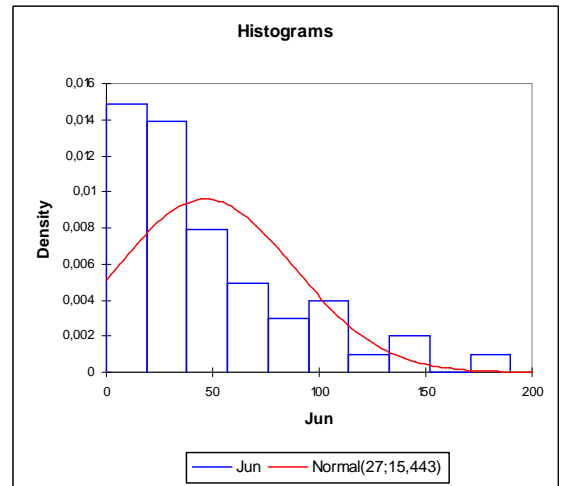
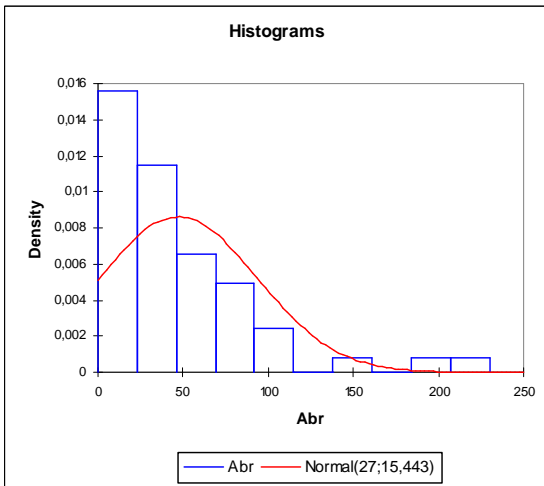
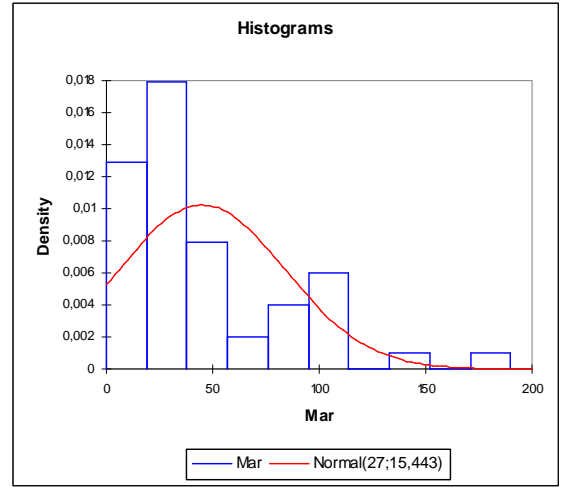
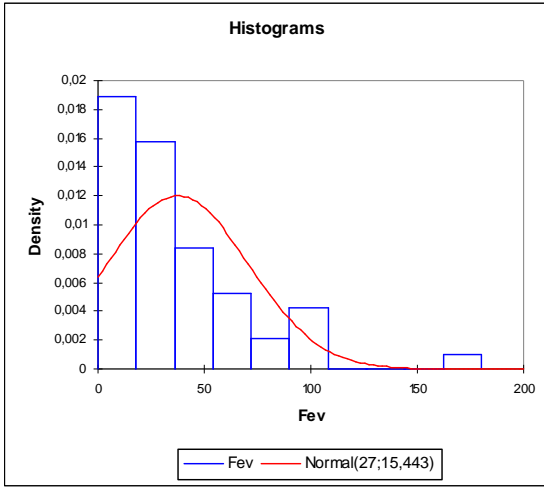




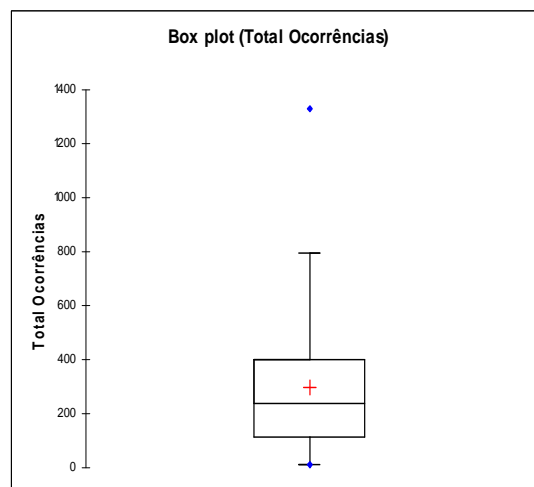
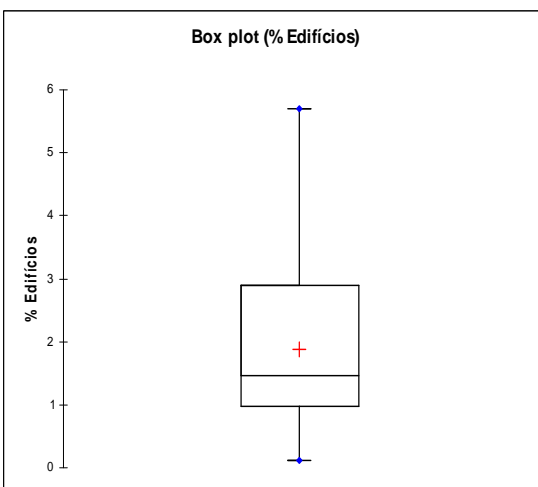
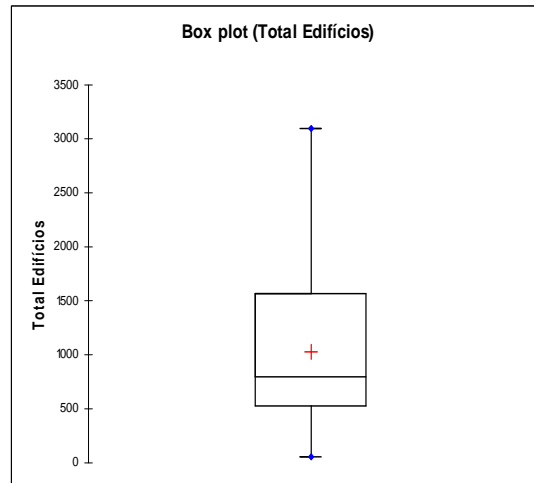
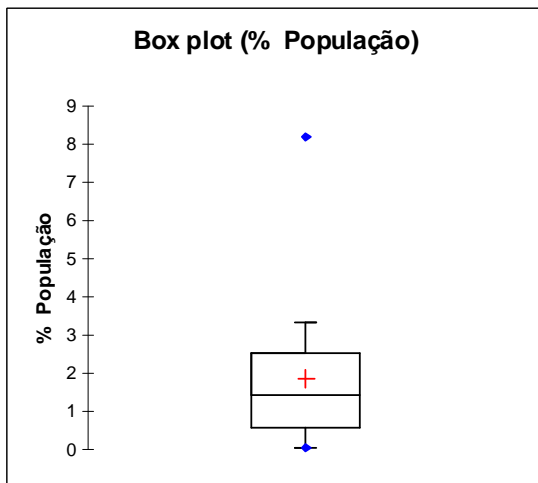
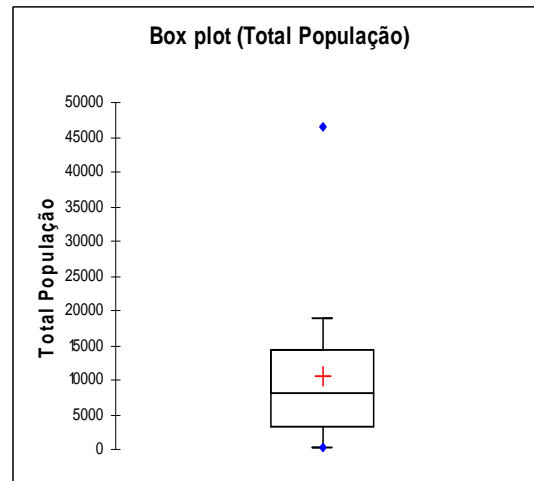
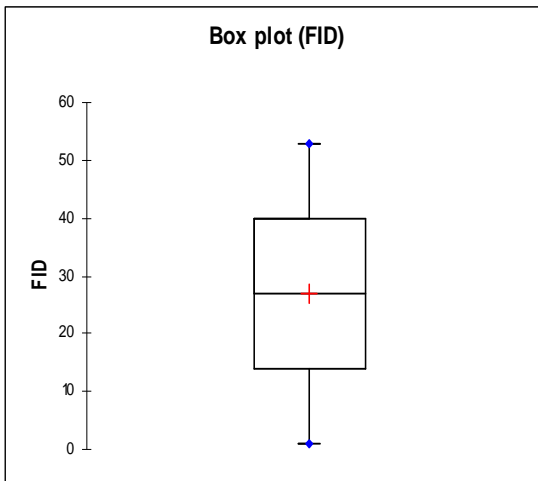


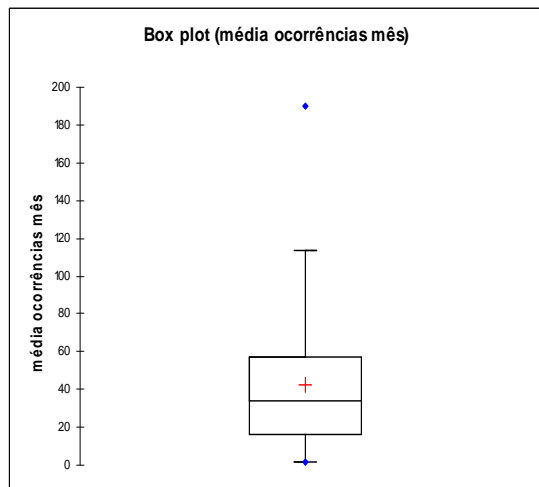
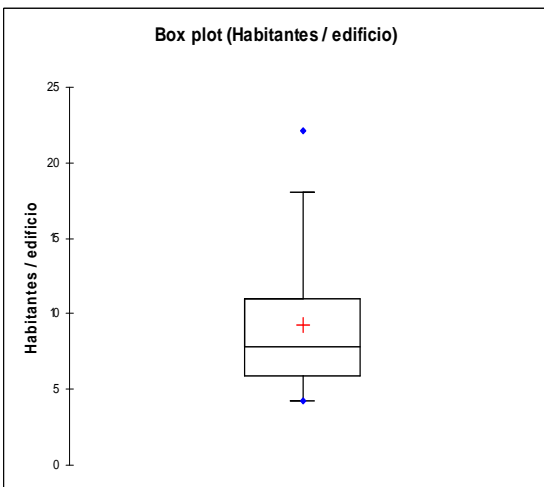
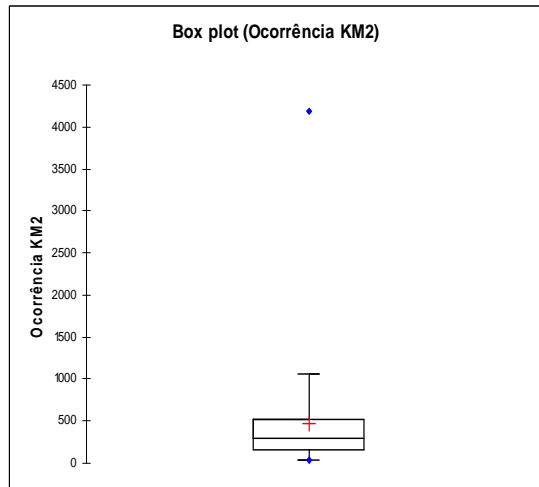
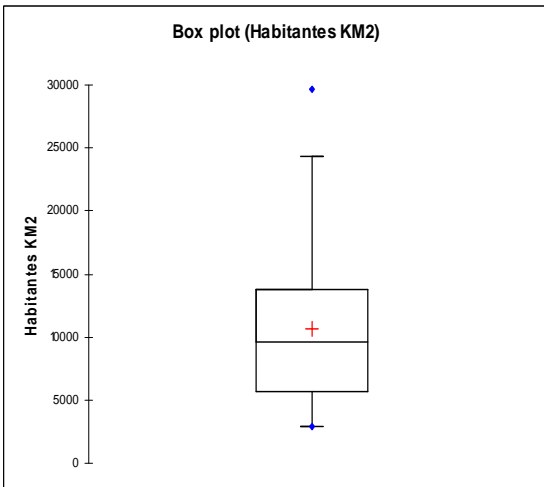
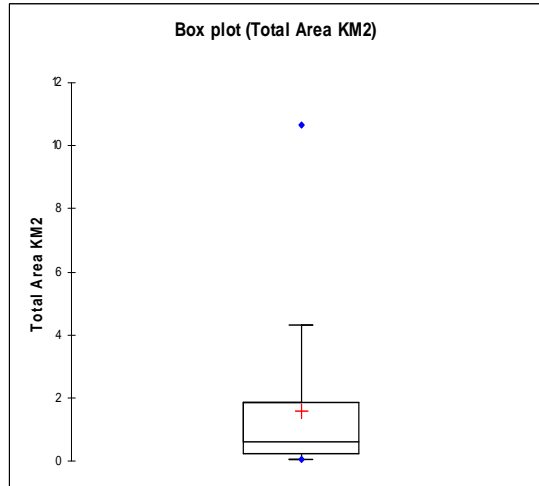
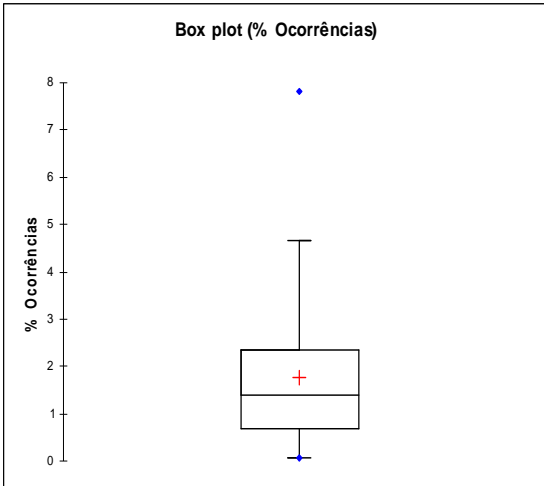


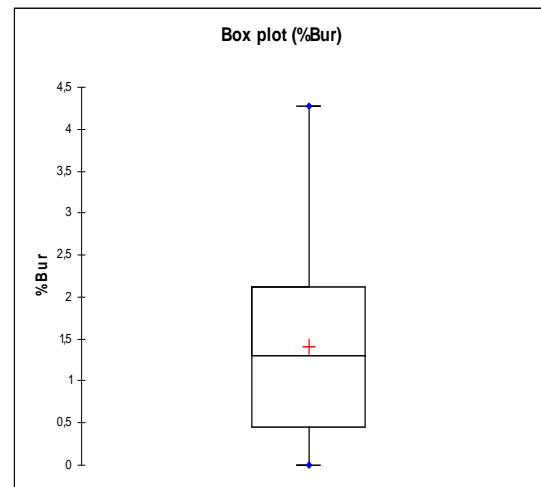
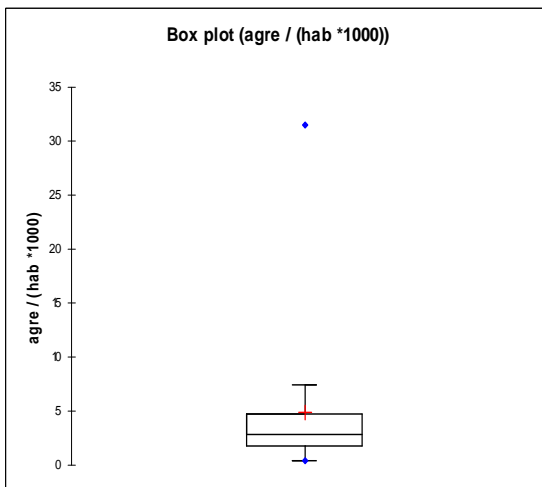
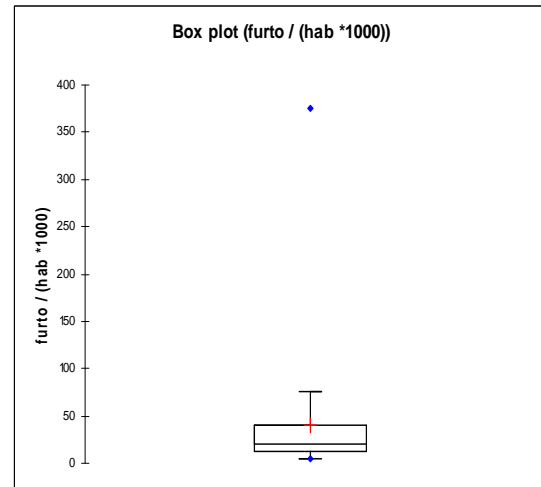
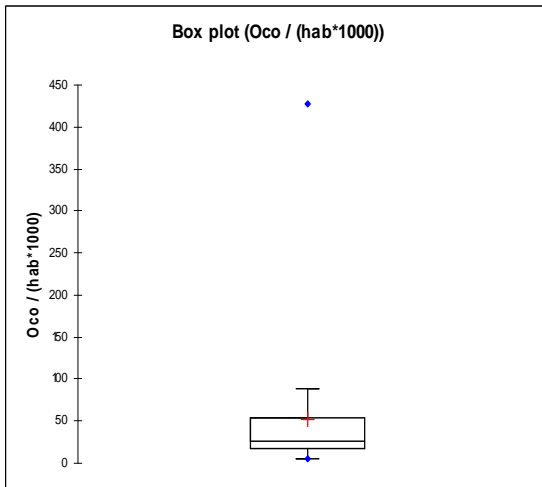
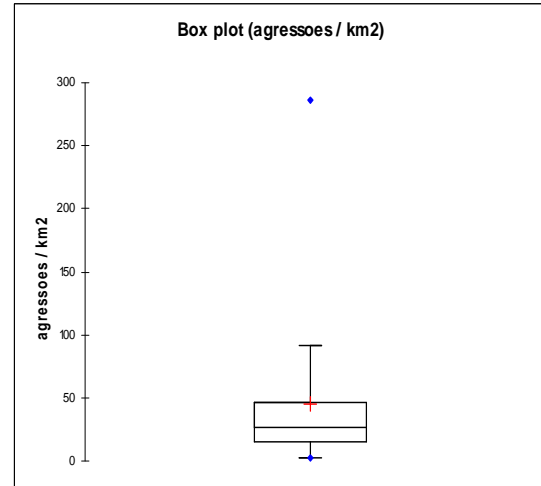
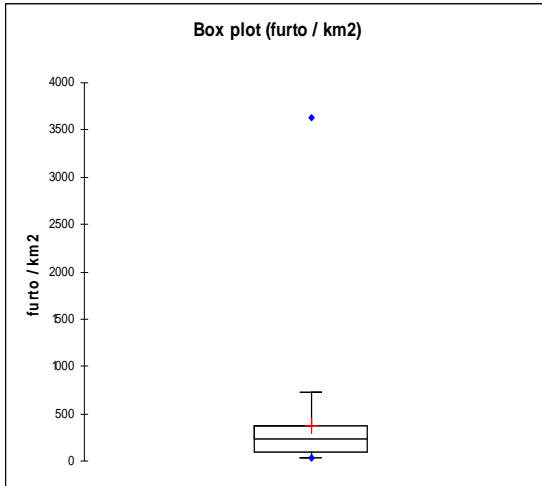


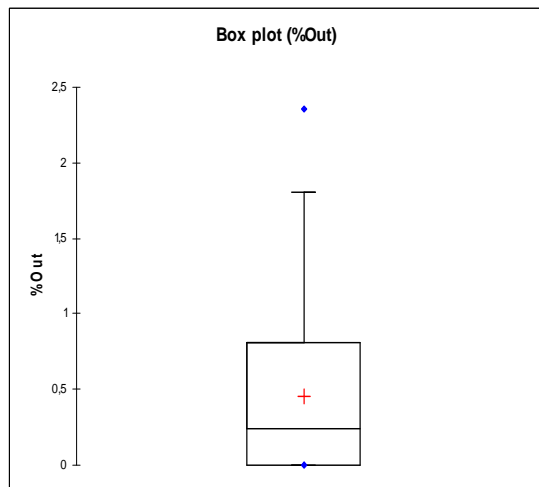
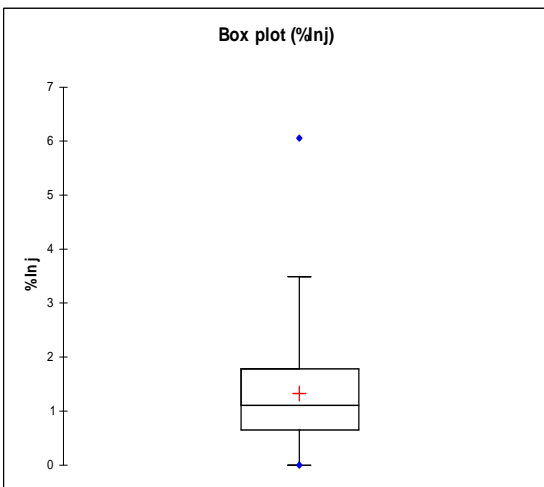
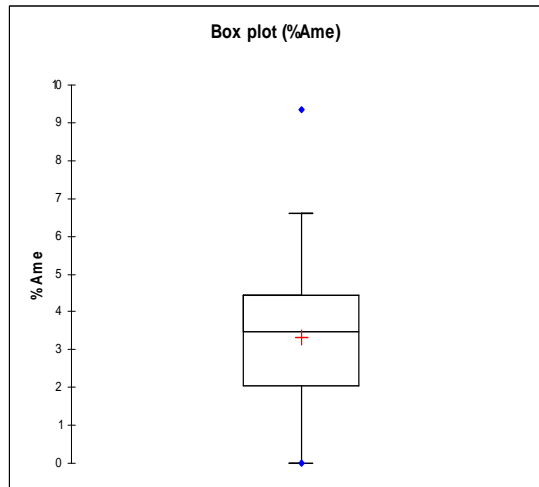
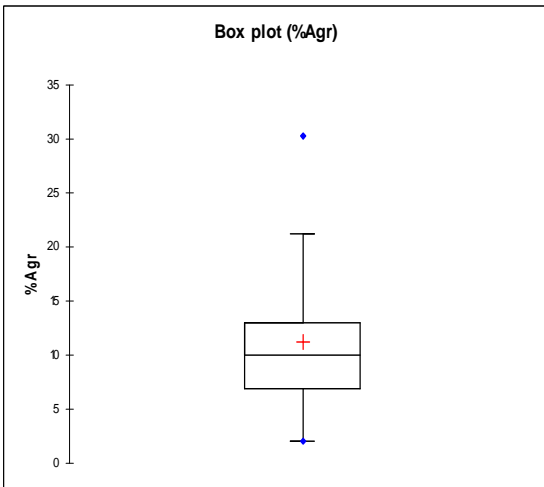
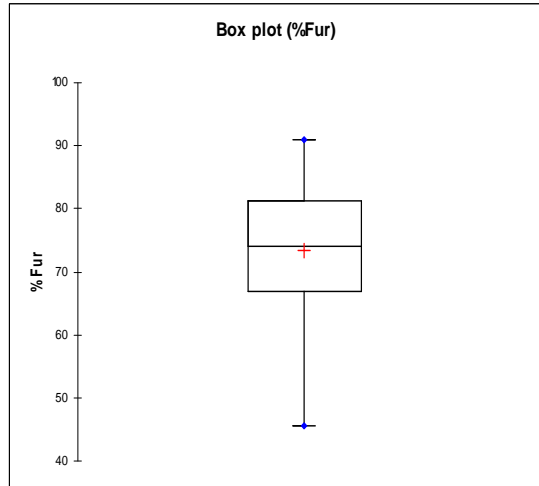
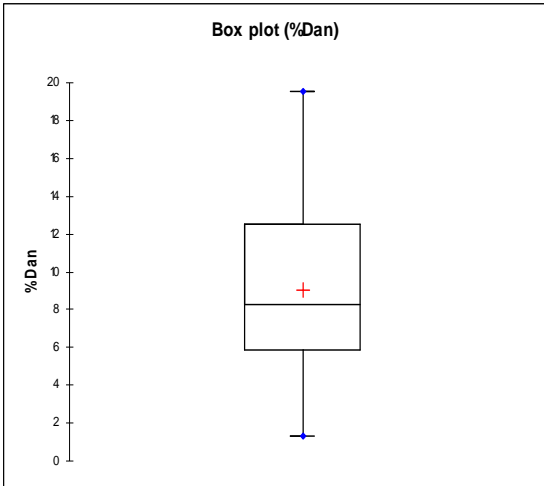


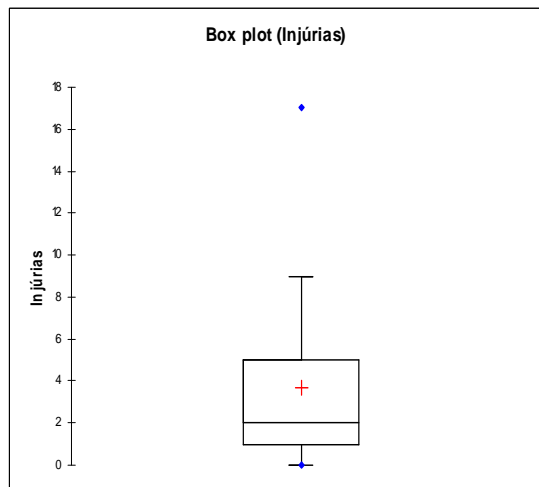
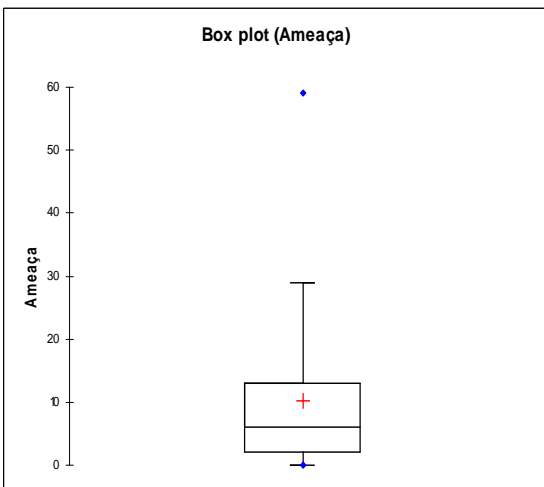
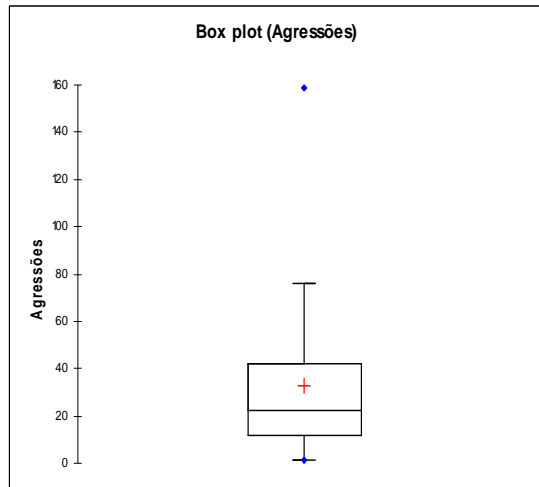
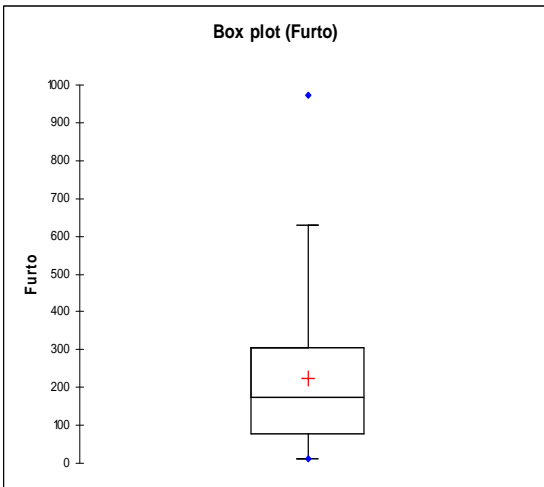
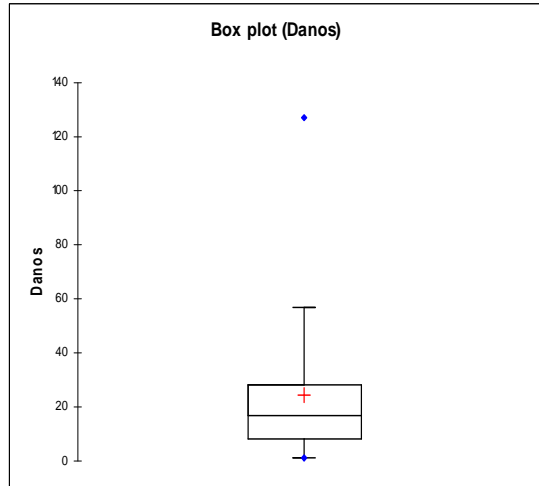
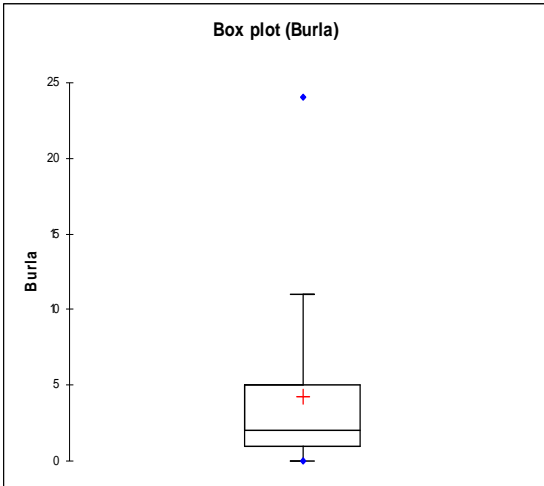
Apêndice 5 – Box-plot das variáveis

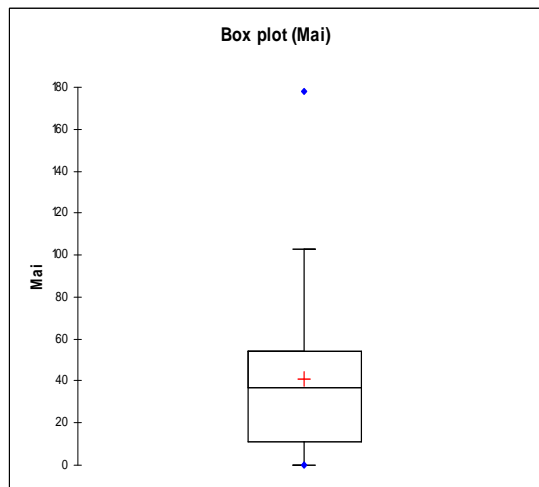
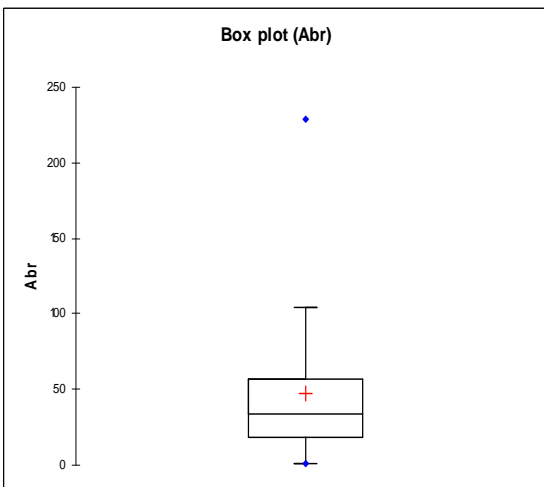
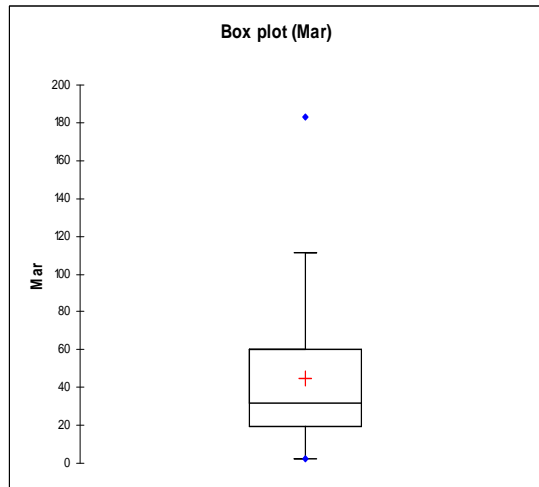
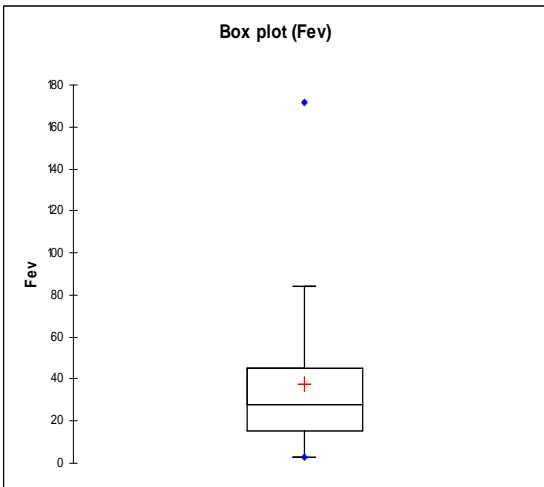
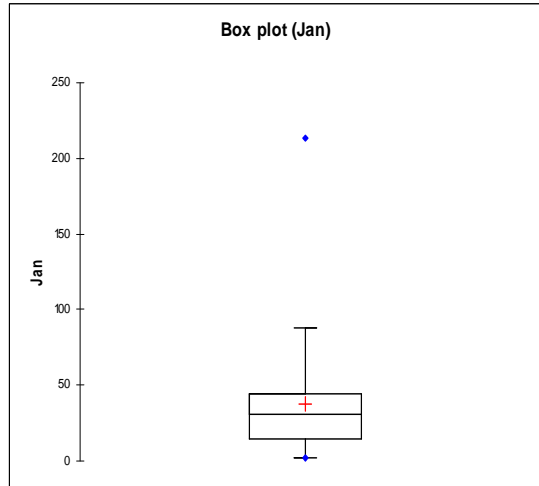
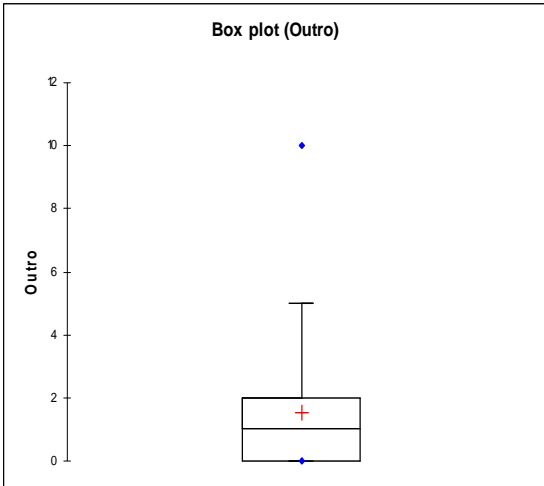


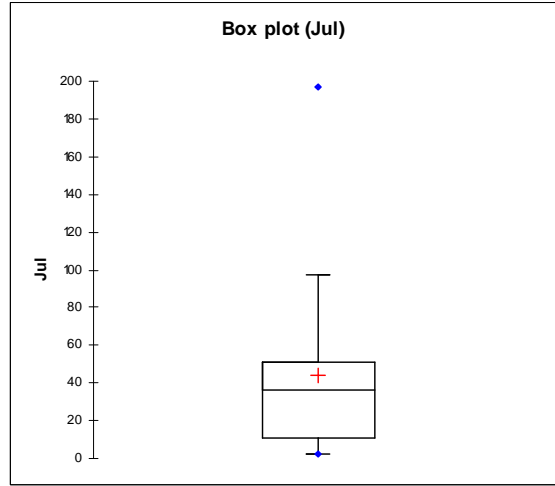
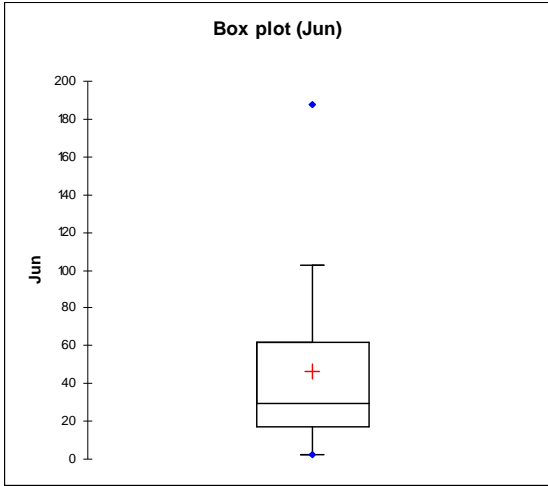












Anexo 1 - Modelo de dados CRISP do SPSS

Business Understanding	Data Understanding	Data Preparation	Modelling	Evaluation	Deployment
<p>Determine Business Objectives Background Business Objectives Business success criteria</p> <p>Assess Situation Inventory of resources Requirements, assumptions and constraints Risks and contingencies Terminology Costs and benefits</p> <p>Determine Data Mining Goals Data Mining goals Data Mining success criteria</p> <p>Produce Project Plan Project Plan Initial assignment of tools and techniques</p>	<p>Collect Initial Data Initial data collection report</p> <p>Describe Data Data description report</p> <p>Explore Data Data exploration report</p> <p>Verify Data Quality Data quality report</p>	<p>Data Set Data set description</p> <p>Select Data Rationale for inclusion and exclusion</p> <p>Clean Data Data cleansing report</p> <p>Construct Data Derived attributes Generated records</p> <p>Integrate Data Merged data</p> <p>Format Data Reformatted data</p>	<p>Select Modelling Techniques Modelling technique Modelling assumptions</p> <p>Generate Text Design Text design</p> <p>Build Model Parameter settings Models Model description</p> <p>Assess Model Model assessment Revised parameter settings</p>	<p>Evaluate Results Assessment of Data Mining results with Business Success criteria Approved models</p> <p>Review Process Review of process</p> <p>Determine Next Steps List of possible actions Decision</p>	<p>Plan Deployment Deployment plan</p> <p>Plan Monitoring and Maintenance Monitoring and maintenance plan</p> <p>Produce Final Report Final report Final presentation</p> <p>Review Project Experience documentation</p>

(in www.spss.com adaptado)

Anexo 2 - Abordagens de investigação

	Quantitativa	Qualitativa
Paradigma	Hipotético – dedutivo	Holístico - interpretativo
Dados	Representados em termos: Numéricos, Quantitativos, Estruturados e não valorativos	Representados de forma verbal: Qualitativos, Com maior riqueza de detalhe
Papel do investigador	Observador Distância objectiva	Interpretador da realidade Imerso no contexto
Abordagem	Positivista Experimental Estudos confirmatórios	Interpretativa Não experimental Estudos exploratórios
Análise	Estatística Inferências a partir de amostras Teste de hipóteses e teorias	Conteúdo ou caso Padrão a partir dos próprios Dados Hermenêutica e fenomenologia

(Dias 2000, adaptado)

Anexo 3 - Metodologias de crime mapping

Method	Unit of Analysis	Data Requirements	Software Requirements	Skill Requirements	Advantages	Disadvantages
Past Hot Spots	Hot Spot	Address level data	GIS, possibly spatial statistical software or raster-based GIS to identify hot spots	Basic GIS skills required Knowledge of spatial statistical software or Raster-based mapping helpful	Easy to compute and understand	Assumes hot spots persist over time Uncertain what level of temporal aggregation is appropriate Does not inform prevention strategies
Repeat Victimization	Address		Spreadsheet GIS helpful but not necessary	Basic spreadsheet and statistical knowledge		Assumes hot spots correspond with hot spots Does not inform prevention efforts
Random Walk	Precinct or other areal unit	Area level data	Spreadsheet or statistical program to compute forecast and desktop GIS to aggregate data and display results	Basic spreadsheet and statistical knowledge Basic desktop GIS knowledge	Easy to compute and understand Very sensitive to changes	Does not use a series of historical figures Does not account for seasonality
Naive Lag 12	Precinct or some other areal unit					Statistical Program that does classical decomposition and seasonal indices Desktop GIS to aggregate data and display results
Classical Decomposition by Individual Areal Unit	Precinct or some smaller areal unit				District level seasonality measure had mixed effectiveness at increasing accuracy Pooled seasonality measure across the entire jurisdiction increases homogeneity of data that, in turn improves prediction accuracy	
Classical Decomposition pooled for entire Jurisdiction						
Exponential Smoothing			Statistical program that does exponential smoothing Desktop GIS to aggregate data and display results	Knowledge of forecasting techniques Basic desktop GIS knowledge	Use of exponential smoothing allows statistic to take into account changes over time Most accurate in forecasting small or moderate changes in crime	Requires specialized software to compute statistics Requires knowledge of forecasting techniques Does not forecast large and/or sudden changes in crime well
Linear Regression		Grid level data	Statistical program able to compute a spatial lag Desktop GIS to aggregate data and display results	Intermediate knowledge of statistics Intermediate desktop GIS knowledge	Can incorporate sudden changes in crime rate in forecasts Can be used to forecast crime changes that are significantly different from normal variation	Requires specialized software to compute statistics Requires intermediate knowledge of statistics and software to implement the model
Point Process Model	Police precincts	Census data at block group level Reported crime data at precinct level	C plus, programming language	Advanced knowledge of programming to create customized model Advanced knowledge of algorithms / processes underlying point process models	Indicates which independent variables are good predictors	Difficult to replicate Use of large areal units limits precision Narrows areas for future deployment rather than targeting specific areas for interventions
Artificial Neural Network		Address level call-for-service data	Requires custom programming and significant computing power	Advanced knowledge of neural networks programming	Potentially more robust than other methods	Very data intensive Requires significant computing power and time Requires a high degree of expertise No statistical tests of significance Unable to determine which inputs have predictive power
Raster GIS Models	Grid Cells	Point of area level data	Raster GIS	Intermediate to advanced knowledge of statistics and forecasting Advanced GIS knowledge	Variables used in the model are theory-driven Great visualization tool both for indicator variables and crime predictions Ability to include special relationships in model	Requires specialized GIS software Can require advanced knowledge of statistics and forecasting Model was better at predicting low risk areas than high risk ones

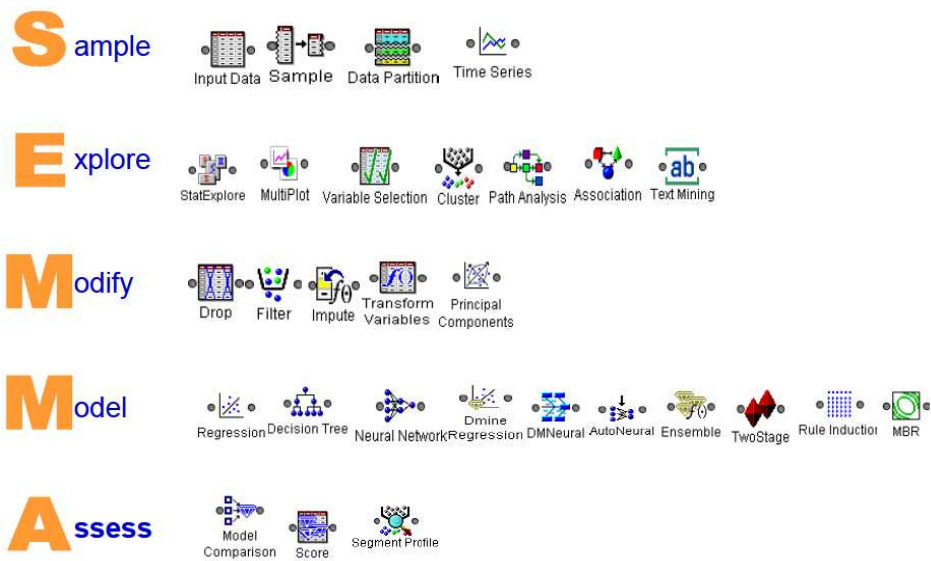
(Groff, Elizabeth et al. 2001)

Anexo 4 - Comparação entre os modelos de data mining

Algorithms	Class. & Regr. Trees	Linear Regression	Multilayer Neural Net	K-Nearest Neighbors	Radial Basis Fns.	Naive Bayes	Rule Induction	Logistic Regression	TimeSeries	Sequential. Rules	K-Means	Association Rules	Kohonen Net
Enterprise Miner (SAS)	X	X	X		X			X	X	X	X	X	
Clementine (SPSS)	X	X	X				X				X	X	X

(in <http://www.cytel.com>)

Anexo 5 - Metodologia SEMMA do Enterprise Miner



(in <http://www.gusasbrasil.com.br>)

Anexo 6 – Pequena História das Freguesias de Lisboa⁶⁰

A Paróquia ou Freguesia Eclesiástica

Das instituições destinadas ao exercício do culto tem o primeiro e principal lugar a freguesia ou paróquia (eclesiástica). Nos tempos mais antigos, à circunscrição territorial, da cidade ou do campo, em que viviam indivíduos todos sujeitos espiritualmente à mesma autoridade eclesiástica, pároco, pastor ou sura, chamava-se colação (do latim *collatio*). A circunscrição territorial ou distrito da colação tinha a sua sede num templo ou igreja matriz. Mais tarde passaram a denominar, e ainda hoje chamam, *parrochia* ou *parochia* (paróquia), tanto a igreja matriz ou sede paroquial, como a sua circunscrição territorial ou colação. Com o mesmo significado, e simultaneamente, usou-se e usa-se a expressão freguesia, aplicada tanto ao distrito territorial, como à igreja matriz; aos moradores da freguesia dá-se indiferentemente a designação de paroquianos ou de fregueses. As freguesias agregaram-se e desagregaram-se ao longo do tempo, originando aquilo a que se chama, árvore genealógica descendente.

Antes da reconquista Cristã da cidade em 1147

Nos dois ou três primeiros séculos da era cristã, durante a dominação romana, o cristianismo teve grande expansão na Península Ibérica, como se infere das perseguições movidas pelos imperadores romanos, do martírio de vários prelados, e da assistência de muitos bispos em vários concílios celebrados em terras da Península. Foi em Lisboa muito violenta a perseguição aos cristãos no tempo do imperador romano Diocleciano, sendo tradição corrente que nessa época (ano 307) foram martirizados em Lisboa os irmãos Veríssimo, Máxima e Júlia, patronos da igreja de Santos-o-Velho. Não nos ficou, porém, notícia alguma de prelados da igreja lisbonense durante o período do domínio sarraceno, mas é natural que os houvesse, como sucedia noutras terras de Portugal sujeitas aquele domínio (Beja, Braga, Coimbra, Lamego, Porto, Viseu).

⁶⁰ Fonte: Câmara Municipal de Lisboa (adaptado)

Da Conquista Cristã até ao fim do Século XIV

Tomada Lisboa, foi por D. Afonso Henriques investido nas funções de prelado do novo ou restabelecido bispado, um sacerdote inglês D. Gilberto, que vinha na armada com os cruzados.

Não há documentos dessa época que nos digam que igrejas paroquiais existiam na cidade conquistada, mas sabe-se que D. Gilberto fundou, dois templos, que foram respectivamente as sedes ou os predecessores das paróquias dos Mártires e de S. Vicente e ainda, segundo o arcebispo D. Rodrigo da Cunha a de Santa Justa. Outros autores atribuem ainda ao reinado de D. Afonso Henriques a criação da freguesia de Santa Maria Maior no templo da Sé (em 1170). Passados 17 anos sobre a conquista, um documento menciona a existência da freguesia de Santa Maria Madalena (1164) e São Pedro (de Alfama) em 1191. Um documento que existia em 1668 no cartório do Convento de São Vicente-de-Fora, e referente a um sínodo realizado em 1191 pelo bispo D. Soeiro Anes na Sé Catedral de Lisboa, menciona como existentes nesse ano as seguintes 6 igrejas colegiadas, sedes de freguesias: Extra-muros da cerca moura: São Vicente, Nossa Senhora dos Mártires e Santa Justa e Intra-muros da cerca moura: Santa Cruz da Alcáçova, São Bartolomeu, São Martinho e São Jorge.

Verifica-se assim que no final do século XII, havia em Lisboa, com maior probabilidade, subordinadas ao bispado de Lisboa, as seguintes 10 freguesias: São Vicente, Nossa senhora dos Mártires, Santa Justa, Santa Maria da Sé, Santa Maria Madalena, Santa Cruz do Castelo, São Bartolomeu, São Martinho, São Jorge e São Pedro (de Alfama).

No Arquivo Nacional da Torre do Tombo guarda-se um pergaminho datado da era 1247, (anos de Cristo 1209 ou 1229), reinados de D. Afonso II ou D. Sancho II, que além das atrás citadas 10 freguesias paroquiais de Lisboa e arrabaldes, menciona mais as 13 seguintes: São Julião, Santa Marinha do Outeiro, São Lourenço, São Nicolau, Santo André, Santo Estevão, São Miguel, Santa Maria de Alcamim (São Cristóvão), São Mamede, São João (Baptista ou da Praça), São Tomé (do Penedo), São Jacob (São Tiago) e São Salvador (da Mata).

Destas 23 freguesias três, Mártires, Santa. Justa e Santo Estevão, possuíam extensíssimas áreas, cujos limites confinavam com freguesias do Termo de Lisboa.

Encravados entre as freguesias, havia em Lisboa alguns tratos de território isentos da jurisdição eclesiástica. Eram as judiarias ou bairros israelitas e a mouraria.

Quando em 1496, acabou a sua existência política, foram os primeiros distribuídos pelas freguesias limítrofes, e o ultimo incorporado na de Santa. Justa, onde estava encravado.

Durante mais de três séculos a estrutura paroquial da cidade manteve-se estacionária, conforme nos dá conta Cristóvão Rodrigues de Oliveira, no seu *Summário* começado a elaborar em 1551, dá notícia de 24 freguesias, as 23 já existentes mais a do Loreto acabada de criar nesse ano.

Se supusermos que a criação das 23 freguesias se efectuou com um ritmo muito rápido, durante o primeiro século depois da conquista, como se poderá explicar que durante os 3 séculos seguintes essa cadência tivesse desaparecido, não havendo necessidade de criação de novas freguesias, apesar de o povoado se ter expandido muito para além das primitivas muralhas, e de a população haver aumentado consideravelmente. Provavelmente a resposta estará no facto de muitas dessas 23 freguesias, já existirem desde há muito tempo, como consequência do processo lento e gradual do aumento da população cristã.

Por isso as datas que os escritores, citam como sendo as das fundações das igrejas paroquiais mencionadas, são colhidas em documentos autênticos, provam simplesmente que tais igrejas já existiam nessas datas.

Meado e 2ª metade do século XVI

Durante o século XV não consta que se tenha dado qualquer modificação paroquial da cidade, com excepção da que resultou da extinção das comunas judaicas e da mouraria, a que já nos referimos. A urbanização continuou a alargar-se para além do núcleo constituído pela população das zonas limitadas pela cerca fernandina, mas foi principalmente depois do começo do século XVI, quando as conquistas, a navegação e o comércio com o oriente trouxeram para o Reino riquezas e comodidades da vida então ainda não sonhadas – que tiveram como consequência natural o aumento da população de Lisboa – que originou a necessidade de criação de novas freguesias. Foram principalmente as extensas paróquias periféricas da cidade, Nossa Senhora dos Mártires, Santa Justa e Santo Estevão, que, por desdobramentos sucessivos, forneceram o maior número de freguesias criadas no meado e na 2ª metade do século XVI.

As freguesias então criadas foram 12: Nossa Senhora do Loreto (depois Nossa Senhora Encarnação), Nossa Senhora da Ajuda, Santa Catarina (do Monte Sinai), Anjos, Sant'Ana (depois Nossa Senhora da Pena), São Paulo, Santos-o-Velho, São José (d'entre as hortas), Nossa Senhora da Conceição (depois Conceição Nova), Santa Engrácia, Trindade (depois Santíssimo Sacramento) e São Sebastião (da Mouraria, depois Nossa Senhora do Socorro).

É como podemos ver nesta altura, mais precisamente em 20 de Novembro de 1567, destacada da freguesia de Santa Justa, que é criada a Freguesia de São José.

Do Século XVII até ao terramoto de 1755

Neste período de 155 anos são criadas apenas 3 freguesias: São Sebastião da Pedreira, Nossa Senhora das Mercês e Santa Isabel Rainha de Portugal.

De referir que a freguesia de Santa Isabel Rainha de Portugal, tem a sua origem no desmembramento das freguesias de Santos, São Sebastião da Pedreira, Santa Catarina e São José, em 15 de Maio de 1741.

2ª metade do século XVIII

Desmoronadas muitas casas, confundidos os limites paroquiais da cidade Baixa pelo terramoto de 1755, e reedificada em seguida a cidade segundo novo plano regular, tornou-se necessário proceder a uma nova distribuição e delimitação das freguesias, o que foi feito em 8 de Abril de 1770. Quase todas as paróquias que permaneceram no seu local primitivo sofreram, modificação nos limites dos seus distritos. Criaram-se pelo mesmo tempo as três seguintes freguesias: Nossa Senhora da Lapa, Santa Joana (depois Coração de Jesus) e Senhor Jesus da Boa Morte.

Em 22 de Janeiro de 1780, sob proposta do Cardeal Patriarca, uma nova distribuição de paróquias foi aprovada e confirmada pelo alvará régio de 19 de Abril de 1780. Esta divisão é a que fundamentalmente ainda está em vigor.

No século XIX

Em 28 de Dezembro de 1833 criou-se a primeira freguesia sem invocação religiosa, que foi a de Belém. Por portaria de 26 de Outubro de 1835, foi permitido ao bispo da diocese anexar paróquias, nos termos do direito canónico. As anexações foram as seguintes: Santa Marinha a Santo André (31 de Maio 1835), Salvador a São Tomé (17 Outubro de 1836) e São Martinho a São Tiago (17 Outubro de 1836).

Estes agrupamentos, e o adição da freguesia de Belém, fixou em 38 o número de freguesias existentes nos meados do século XIX. Mais tarde anexaram-se: Salvador e São Tomé a São Vicente (1 Fevereiro 1856), São João da Praça a Santa Maria Maior ou Sé (24 de Dezembro 1885) e São Lourenço a São Cristóvão (16 de Junho de 1886).

Em 18 de Julho de 1885, foram incorporadas no território do município algumas freguesias, ou partes de freguesias, que nessa época pertenciam aos Concelhos de Belém e dos Olivais: São Pedro de Alcântara, Nossa Senhora da Ajuda, Nossa Senhora de Belém, Benfica, Carnide, São Bartolomeu do Beato, Charneca, Ameixoeira, Lumiar, Olivais e Campo Grande.

O território do município de Lisboa foi de novo ampliado em 22 de Junho de 1886, acrescentando-lhe 2 freguesias do concelho dos Olivais – Sacavém e Camarate- que foram no entanto desanexadas em 20 de Setembro de 1895, passando então para o concelho de Loures. Vê-se que Lisboa contava, no fim do século XIX, 42 freguesias.

No Século XX

Neste século, em consequência da urbanização de terrenos e campos ao norte e ao poente do núcleo de maior densidade populacional, onde se construíram novos e amplos bairros servidos por belas avenidas, originou-se a necessidade de novas paróquias, que foram criadas com territórios destacados de outras que, pela sua vastidão e população, já não podiam satisfazer às conveniências de administração dos Sacramentos, e à comodidade dos seus fregueses.

Segundo o código Administrativo de 31 de Dezembro de 1940, o concelho de Lisboa dividia-se administrativamente em 43 freguesias, distribuídas por 4 bairros. Pelo Decreto-Lei nº 42142 de 17 de Fevereiro de 1959, a cidade de Lisboa passou a estar dividida nas actuais 53 freguesias.

Século XXI

Actualmente discute-se a criação da Freguesia do Oriente, que, a concretizar-se, será a 54ª freguesia, irá englobar a zona da antiga Expo98, tem cerca de 10.000 habitantes e uma área de 3,3 km², no entanto tem tido o voto desfavorável da autarquia de Loures que reivindica aquele espaço como pertencente ao seu Concelho.

Anexo 7 – Legislação e documentos de referência

Lei 109/09	15 Set	Aprova a Lei do Cibercrime
Lei 51/07	31 Ago	Lei de Política Criminal
Lei 48/07	29 Ago	Código Processo Penal
Lei 59/07	4 Set	Código Penal
Lei 17/06	23 Maio	Lei-quadro de Política Criminal
Decreto-lei 305/02	13 Dez	Altera a Lei 21/2000 de 10 Agosto
Lei 21/00	10 Ago	Lei da Organização da Investigação Criminal
Lei 5/99	27 Jan	Lei de Organização e Funcionamento da PSP
Lei 60/98	28 Ago	Estatuto do Ministério Público
Lei 67/98	26 Out	Lei da Protecção de Dados Pessoais
Decreto Regulamentar 4/95	31 Jan	Protecção de dados pessoais face à informática
Lei 2/95	25 Jan	Protecção de dados pessoais no sector da polícia
Lei 5/95	21 Fev	Lei da segurança dos dados pessoais
Lei 109/91	17 Ago	Lei da Criminalidade Informática
Lei 20/87	12 Jan	Lei de Segurança Interna
Portaria 290/87	8 Abr	Quadro de Pessoal com Função Policial
Decreto Regulamentar 240/82	16 Out	Ratifica o Decreto-lei 410/1982
CRP		Constituição da República Portuguesa
PGR 2008		Relatório Anual da Procuradoria Geral da República
RASI 2007 e 2008		Relatório Anual de Segurança Interna
NP405-4	2002	Norma Portuguesa para documentos electrónicos
NP405-3	2000	Norma Portuguesa para documentos não publicados
NP405-2	1998	Norma Portuguesa para materiais não livro
NP405-1	1994	Norma Portuguesa para documentos impressos
CENSO 2001		Censo Nacional efectuado pelo INE