

Mestrado em Gestão de Informação
Master Program in Information Management

SISTEMA DE BUSINESS INTELLIGENCE PARA APOIO AO PLANEAMENTO E GESTÃO DE INCENTIVOS

O caso Compete2020

Fernando Rodrigues Maia

Trabalho de Projeto apresentado como requisito parcial para
obtenção do grau de Mestre em Gestão de Informação

NOVA Information Management School
Instituto Superior de Estatística e Gestão de Informação
Universidade Nova de Lisboa

SISTEMA DE BUSINESS INTELLIGENCE
PARA APOIO AO PLANEAMENTO E GESTÃO DE INCENTIVOS

O CASO COMPETE2020

por

Fernando Rodrigues Maia

Trabalho de Projeto apresentado como requisito parcial para a obtenção do grau de Mestre em Gestão de Informação, Especialização em Gestão do Conhecimento e Business Intelligence

Orientador: Miguel de Castro Neto, PhD

Julho 2021

AGRADECIMENTO

Quando aos 50 anos avaliamos a possibilidade de fazer um mestrado e nos propomos a ter aulas com colegas da idade dos filhos, muitas dúvidas se levantam. Entrar nesse desafio e concluí-lo com sucesso só foi possível graças ao apoio direto ou indireto de várias pessoas.

Em primeiro lugar agradeço aos meus filhos. O Ricardo que ao frequentar a licenciatura de gestão de informação me deu a conhecer a NOVAIMS, os seus cursos e mestrados. O Rodrigo, cujo empenho no secundário serviu de exemplo ao pai. Em segundo lugar, mas não menos importante, agradeço à minha esposa Ana pelo apoio incondicional, pelo incentivo constante, por compreender e colmatar os momentos da minha menor disponibilidade para a família.

Agradeço também ao professor Miguel de Castro Neto, meu orientador desta tese, por toda a disponibilidade demonstrada, pela ajuda para ultrapassar os obstáculos que foram surgindo, por todos os contributos para melhorar esta tese, devido à sua vasta experiência e conhecimento, e ainda por acreditar sempre na minha capacidade de concluir estcom sucesso.

Quero agradecer ao COMPETE2020 pela possibilidade da realização desta tese num contexto de projeto real em curso na organização.

Por último agradeço a todos os meus colegas de mestrado e professores, que contribuíram para aumentar o meu conhecimento e tornaram estes 2 anos numa experiência enriquecedora.

RESUMO

A evolução vertiginosa das tecnologias de informação tem possibilitado às organizações melhorar as suas capacidades de funcionamento, tornando os seus processos mais ágeis, mais eficazes e mais eficientes. A informatização de praticamente todos os processos das organizações possibilita a existência de um manancial de dados que, além da parte operacional, poderão servir à ajuda nas tomadas de decisão da gestão intermédia e de topo. No entanto, os dados nem sempre se encontram disponíveis num formato compreensível aos gestores pelo que existe a necessidade de os transformar em informação útil.

Nos últimos 20 anos muitas das grandes organizações procuraram perceber como tirar partido dos seus dados operacionais desenvolvendo projetos de Business Intelligence, recorrendo às mais modernas ferramentas de BI e a recursos de IT especializados no tratamento de dados.

O projeto a que este relatório diz respeito decorreu no COMPETE2020, entidade pública responsável pela gestão fundos afetos ao desenvolvimento da competitividade das empresas nacionais. O projeto teve como âmbito a definição de um modelo de dados que servirá de suporte à futura solução de BI da organização, bem como a identificação e análise de potenciais ferramentas de BI para implementação da solução.

A futura ferramenta de BI tem como objetivo servir a gestão de topo e intermédia, nas decisões que permitam melhorar a sua capacidade de gestão da utilização dos fundos atuais e futuros para financiamento às empresas.

Como resultado final da solução de BI, passará a organização a ter disponível um conjunto de indicadores, inexistentes atualmente ou que exigem um trabalho manual repetitivo que não permite a disponibilização da informação em tempo útil pretendido.

PALAVRAS CHAVE

Business Intelligence, Data Warehouse, ETL, Modelo de Dados, Dashboards, Indicadores de Performance

ABSTRACT

The vertiginous evolution of information technology has enabled organizations to improve their operational capabilities, making their processes more agile, more effective and more efficient. The computerization of almost all organizational processes enables the existence of data that, in addition to the operational part, can help in the decision-making process of middle and top management. However, the data is not always available in a format that managers can understand, so there is a need to transform it into useful information.

In the last 20 years, many large organizations have tried to understand how to take advantage of their operational data by developing Business Intelligence projects, using the most modern BI tools and IT resources specialized in data processing.

The project to which this report refers took place in COMPETE2020, a public entity responsible for managing funds allocated to the development of the competitiveness of national companies. The project's scope was the definition of a data model that will support the future BI solution of the organization, as well as the identification and analysis of potential BI tools to implement the solution.

The future BI tool aims to serve top and middle management in the decisions that will improve their ability to manage the use of current and future funds for business financing.

As a final result of the BI solution, the organization will have available a set of indicators, that do not currently exist or that require repetitive manual work that does not allow the provision of information in the desired timely manner.

KEYWORDS

Business Intelligence, Data WareHouse, ETL, Modelo de Datos, Dashboards, Indicadores de Performance

ÍNDICE

1. INTRODUÇÃO	1
1.1 Apresentação resumida do programa Compete	1
1.2 Identificação das necessidades da organização	3
1.3 Questão e objetivo do trabalho.....	4
1.4 Abordagem metodológica	4
1.5 Estrutura da tese	5
2. REVISÃO DE LITERATURA.....	6
2.1 Informação de gestão nas organizações	6
2.2 Introdução histórica BI	7
2.3 Arquitetura de BI	10
2.4 ETL	14
2.5 Data Warehouse.....	17
2.6 Visualizações e DashBoards.....	21
3. ANÁLISE DE DADOS E REQUISITOS	24
3.1 Arquitetura de Sistemas Operacionais	24
Sistemas Operacionais	25
Repositório de dados das aplicações internas e externas	27
3.2 Quadros comunitários anteriores	28
3.3 Dificuldades Encontradas	28
3.4 Análise dos dados base do ReAD.....	29
3.5 Levantamento de Indicadores de negócios	30
4. MODELO DE DADOS E ARQUITETURA	32
4.1 Construção do Modelo Dimensional	32
Fact Tables	32
Dimension Tables.....	37
Diagrama de relações	40
4.2 Indicadores a desenvolver.....	40
4.3 Arquitetura proposta de BI.....	41
Staging Area	43
Processo ETL	43
Data Warehouse	44
Slowly Changing Dimensions	44
4.4 Apresentação da prova de conceito.....	45
4.5 Potencial de Análise Preditiva	54
4.6 Identificação dos requisitos da solução de BI.....	55
4.7 Análise de ferramentas de BI	56
5. CONCLUSÕES.....	60
5.1 Trabalho desenvolvido	60
5.2 Objetivos atingidos.....	61
5.3 Recomendações	61
6. BIBLIOGRAFIA	63
ANEXO I – Principais Tabelas do ReAD.....	65
ANEXO II - Principais Indicadores a Implementar	67
ANEXO III – Diagrama Relações do Modelo de Dados	72
ANEXO IV – ETL Detalhado.....	73

ÍNDICE DE FIGURAS

Figura 1: Evolução Funcionalidades de BI	10
Figura 2: Arquitetura Hub-and-Spoke, Inmon.....	11
Figura 3: Arquitetura Enterprise Data-Bus, Kimball	11
Figura 4: Fluxo de dados num sistema de BI.....	12
Figura 5: Fases de um processo de ETL.....	14
Figura 6: Tabela dimensão produto	18
Figura 7: Tabela factos vendas	18
Figura 8: Star Schema.....	20
Figura 9: Snowflake Schema	20
Figura 10: Tipos de ferramentas de BI (utilizador final).....	21
Figura 11: Fluxo de informação dos Sistemas de informação do Compete2020.....	25
Figura 12: Arquitetura de Sistemas de informação do Compete2020	27
Figura 13: Tabelas ReAD.....	29
Figura 14: exemplos Relatório ‘Sistema Incentivos’	30
Figura 15: Diagrama de Relações (ver detalhe no anexo III).....	40
Figura 16: Arquitetura solução BI.....	42
Figura 17: Dashboards Relatório mensal	47
Figura 18: Dashboard visão global	48
Figura 19: Dashboard realização por dimensão.....	49
Figura 20: Dashboard Análise Geográfica	50
Figura 21: Dashboard Prazos.....	51
Figura 22: Dashboard Projetos encerrados.....	52
Figura 23: Dashboard módulo operacional.....	53
Figura 24: Magic Quadrant for Analytics and Business Intelligence Platforms	55

ÍNDICE DE TABELAS

Tabela 1: Informação técnica SGO e ReAD (fonte: Departamento IT Compete2020)	28
Tabela 2: Descrição colunas de tabelas de factos.....	32
Tabela 3: FACT Operações.....	35
Tabela 4: FACT series.....	36
Tabela 5: Tabelas de Dimensões	39
Tabela 6: Indicadores	41

LISTA DE SIGLAS E ABREVIATURAS

AG – Autoridade de Gestão

BD – Base de Dados

BI – Business Intelligence

CRM – Customer Relationship Management

DM – Data Mart

DW – Data Warehouse

ERP – Enterprise Resource Planning

ETL - Extract Transform Load

FEDER – Fundo Europeu de Desenvolvimento Regional

FEEL - Fundos Europeus Estruturais e de Investimento

IAPMEI – Agência para a Competitividade e Inovação, I.P.

IT – Information Technology

KPI - Key Performance Indicator

NOVAIMS - Information Management School da Universidade Nova da Lisboa

OLAP – On-Line Analytical Processing

OT - Objetivos Temáticos

PT2020 – Portugal 2020

SA – Staging Area

SCD - Slow Changing Dimensions

SGO2020 – Sistema de Gestão Operacional

SIAC – Sistema de Apoio às Ações Coletivas

SIFSE – Sistema de Informação do Fundo Social Europeu

SI Audit 2020 – Sistema de Informação de Controlo e Auditoria

SI PT2020 – Sistema de Informação do Portugal 2020

TI – Tecnologias de Informação

1. INTRODUÇÃO

A atribuição de apoios às empresas, na forma de financiamentos reembolsáveis ou não reembolsáveis, tem como objetivo potenciar o desenvolvimento de empresas nacionais para operar no mercado global (IAPMEI, 2021). A correta atribuição destes apoios e a eficácia do seu acompanhamento são fatores fundamentais para uma boa aplicação das verbas do estado e essa é a missão do COMPETE2020.

Assim, à semelhança de qualquer outro negócio ou organização, é imprescindível efetuar uma eficaz gestão dos processos (projetos financiados) em curso e em todas as suas fases. Nos anos 50 William E. Deming disse *‘o que não se mede não se consegue gerir’* e isso é cada vez mais importante nos tempos atuais onde os negócios são cada vez mais rápidos. O elevado nível de informatização dos processos das organizações é gerador de uma cada vez maior quantidade de dados, que podem possibilitar medir e consequentemente gerir melhor uma organização.

William E. Deming (1986) disse ainda *‘Acredito em Deus, todos os outros devem apresentar dados e factos’*. O COMPETE2020 tem todas as suas operações suportadas em sistemas informáticos pelo que dispõe do recurso base para conseguir medir e gerir as suas operações, os dados. Com base neste ponto de partida e consciente da grande importância da necessidade bem gerir os apoios disponibilizados, decidiu a Administração do COMPETE2020 avaliar e estudar o melhor modo de transformar os dados existentes em informação de gestão útil à organização.

1.1 APRESENTAÇÃO RESUMIDA DO PROGRAMA COMPETE

O COMPETE 2020 tem como missão assegurar a seleção de projetos candidatos aos fundos disponíveis para melhorar a competitividade e a internacionalização da economia portuguesa, estando orientado sobretudo para as regiões menos desenvolvidas do Continente – Norte, Centro e Alentejo – tendo abrangência nacional nos projetos do Fundo de Coesão, formando com os Programas Operacionais Regionais do Continente uma rede diversificada de instrumentos de política pública com regras e objetivos comuns que cobre todo o território nacional (COMPETE2020, 2021).

O COMPETE 2020 encontra-se estruturado em 6 Eixos prioritários, tendo por base os Objetivos Temáticos (OT) estabelecidos no quadro regulamentar dos Fundos Europeus Estruturais e de Investimento (FEEI), devidamente alinhados com a Estratégia Europa 2020 (COMPETE2020, 2021):

- Eixo I: Reforço da investigação, do desenvolvimento tecnológico e da inovação (OT1);
- Eixo II: Reforço da competitividade das PME incluindo a redução de custos públicos de contexto (OT3 e OT2);
- Eixo III: Promoção da sustentabilidade e da qualidade do emprego e apoio à mobilidade dos trabalhadores (OT8);
- Eixo IV: Promoção de transportes sustentáveis e eliminação dos estrangulamentos nas principais redes de infraestruturas (OT7);
- Eixo V: Reforço da capacidade institucional das autoridades públicas e das partes interessadas e da eficiência da administração pública (OT11);
- Eixo VI: Assistência Técnica.

Para uma resposta adequada face aos objetivos estratégicos e específicos definidos para COMPETE 2020, perspetiva-se a mobilização de um conjunto diversificado de instrumentos de política pública, incluindo:

- Incentivos diretos ao investimento empresarial, sobretudo em I&I e qualificação de PME, primordialmente focalizados em estratégias de internacionalização;
- Apoios indiretos ao investimento empresarial, promovendo a capacitação e colaboração das empresas;
- Apoios à produção e difusão de conhecimento científico e tecnológico, reforçando quer as ligações internacionais do Sistema Nacional de I&I, quer o desenvolvimento das ligações, bem como de sinergias e de mecanismos eficazes de transferência de conhecimento e tecnologia, entre empresas, centros de I&D e o ensino superior;
- Apoios à formação empresarial, no sentido de capacitar os recursos humanos das empresas para os processos de inovação e internacionalização;
- Apoios a investimentos em infraestruturas de transportes, centrados na redução do tempo e custo de transporte para as empresas, sobretudo no âmbito da conectividade internacional;
- Apoios à modernização administrativa, visando a redução dos custos públicos de contexto;
- Capacitação dos serviços e dos trabalhadores em funções públicas, promovendo uma Administração Pública mais eficiente.

1.2 IDENTIFICAÇÃO DAS NECESSIDADES DA ORGANIZAÇÃO

Devido à amplitude da sua atuação, da complexidade técnica e institucional, da complexidade da estrutura de gestão, do volume de dados e informação, o controlo do Compete2020 implica, cada vez mais, um sistema de informação capaz de acompanhar a atividade operacional e ser uma ferramenta de apoio à gestão. Para dar resposta a estas necessidades o Compete2020 pretende implementar uma solução que permita extrair, tratar e organizar os dados dos sistemas operacionais e transformá-los em informação de gestão fiável e atempada de modo a responder a um conjunto de necessidades de gestão:

- Disponibilizar um conjunto de indicadores de gestão que permitam acompanhar as várias fases de um projeto financiado;
- Produção de informação e/ou relatórios relativos a períodos temporais;
- Definição de alertas de indicadores;
- Possibilitar a criação de novos indicadores;
- Possibilitar a integração de informação interna com dados de sistemas externos;
- Permitir a disponibilização de dados através de rotinas de extração/exportação;
- Permitir a análise de informação com a granularidade adequada às necessidades de cada utilizador;

A gestão dos projetos candidatados/apoiados é efetuada em todo o ciclo de vida do projeto, desde a candidatura até ao encerramento passando pelo acompanhamento da execução.

1.3 QUESTÃO E OBJETIVO DO TRABALHO

O sistema de informação do COMPETE2020 assegura o registo de toda a informação bem como a gestão operacional associada ao processo de candidatura à atribuição de apoios e seu acompanhamento posterior. Para efeitos de gestão são disponibilizados um conjunto de relatórios/consultas em formato html, excel e pdf que apresentam algumas limitações de flexibilidade e capacidade de resposta a indicadores de gestão, úteis a uma gestão eficaz e eficiente. Para ultrapassar estas limitações é necessário muitas vezes produzir reporting adicional com recurso a ferramentas como excel ou desenvolvimentos pela equipa de IT.

Com vista a melhorar a capacidade de gestão da organização, a administração do COMPETE2020 pretende implementar uma solução de Business Intelligence (BI) que possibilite a disponibilização de informação de gestão da atividade operacional.

Tendo em conta os objetivos do Compete2020 executou-se este trabalho para dar resposta à questão seguinte:

Será possível construir um modelo de BI, alimentado pelos sistemas operacionais do Compete2020, que permita a produção das análises e indicadores, quer de uma perspetiva operacional quer de gestão?

Relativamente a esta questão definiram-se 3 objetivos:

- 1) Criação do modelo de dados do data warehouse que possibilitará a futura implementação de um sistema de Business Intelligence;
- 2) Criação de um protótipo funcional, como prova de conceito, para apresentação de um conjunto de indicadores de gestão à administração;
- 3) Identificação e análise de potenciais ferramentas para implementação da futura solução de BI.

1.4 ABORDAGEM METODOLÓGICA

O desenvolvimento deste trabalho foi efetuado em várias etapas sequenciais de modo a garantir atempadamente a informação necessária a cada uma das etapas posteriores. Resumidamente foram executadas as etapas seguintes:

- 1) Levantamento, junto da equipa de IT e utilizadores chave, de informação técnica e funcional dos principais sistemas operacionais utilizados;
- 2) Reuniões com utilizadores chave para levantamento dos principais indicadores a implementar no futuro sistema de BI;
- 3) Desenho do modelo de dados do data warehouse e ETL respetivo;
- 4) Construção do protótipo funcional e apresentação à administração;

- 5) Pesquisa e análise comparativa de potenciais ferramentas para implementação da solução de BI;
- 6) Elaboração do relatório final a disponibilizar ao Compete2020.

1.5 ESTRUTURA DA TESE

A elaboração desta tese foi dividida em 5 capítulos. No primeiro capítulo apresenta-se a organização e as suas necessidades no âmbito do tema em análise nesta tese. É também identificada a principal questão a que se pretende dar resposta bem como os vários objetivos que esta tese propõe atingir. Apresenta-se ainda a metodologia seguida para desenvolvimento dos trabalhos bem como a estrutura da tese.

No segundo capítulo é efetuada uma revisão de literatura, dos principais temas a ter em conta no desenvolvimento deste trabalho, tendo em vista o adequado suporte aos resultados finais apresentados.

O terceiro capítulo descreve a metodologia e os trabalhos de levantamento realizados, junto dos principais intervenientes na organização, com vista ao melhor conhecimento possível da organização e suas necessidades.

No quarto capítulo são apresentados as principais análises efetuadas bem como os resultados finais atingidos, nomeadamente: modelo de dados, arquitetura de BI e protótipo construído para prova de conceito, bem como a avaliação de potenciais ferramentas para a sua implementação.

No quinto e último capítulo apresentam-se as principais conclusões e requisitos da solução de BI. Neste capítulo são também identificadas as principais recomendações futuras e fatores críticos de sucesso.

Por fim são apresentados um conjunto de anexos de suporte bem como as referências da literatura consultada neste trabalho.

2. REVISÃO DE LITERATURA

2.1 INFORMAÇÃO DE GESTÃO NAS ORGANIZAÇÕES

Desde os primórdios dos sistemas de informação até praticamente ao fim do século XX que as ferramentas utilizadas nas empresas tinham como principal objetivo automatizar e otimizar os processos chave das empresas (Davenport, 2010). Num tempo onde o avanço da tecnologia é cada vez mais rápido, o conhecimento é partilhado em rede e os negócios são globais, os negócios e os modelos de gestão tradicionais são postos em causa e requerem uma abordagem de gestão e ferramentas novas (Davenport et al., 2006). Estes fatores levaram a que no início do século XXI muitas empresas procurassem o melhor modo de utilizar a quantidade de dados existente para a gestão do negócio.

Diariamente tomamos um conjunto de decisões, tanto na vida pessoal como profissional, que se esperam sejam o mais acertadas possíveis. Do acerto dessas decisões poderá depender o sucesso ou insucesso de uma empresa sejam as decisões relativas a custos, vendas, recursos humanos, marketing, etc. (Larson, 2012). Se o objetivo de melhor analisar a informação é principalmente tomar melhores decisões e suportar as ações a tomar, então é importante que as decisões sejam claramente suportadas pela informação (Davenport, 2010).

Se é óbvio que o responsável máximo de uma organização tem que tomar decisões ao nível estratégico para o futuro da empresa, também é óbvio que a sua linha seguinte de gestão tem de ser capaz de tomar as decisões adequadas para o sucesso das áreas que coordenam (Larson, 2012). Para que as decisões sejam eficazes é importante que, em primeiro lugar, estejam definidos os objetivos para os quais se deve trabalhar. Em segundo, tem de haver uma forma de medir se um percurso escolhido está a avançar para ou a afastar-se desses objetivos e por fim é imprescindível que a informação baseada nessas medidas deva ser fornecida ao decisor de forma atempada (Larson, 2012).

A definição de objetivos deve ser clara e mensurável. Objetivos como ‘elevada satisfação do cliente’ ou ‘produtos de elevada qualidade’ sem que seja definido um método para efetuar esta medição tornam impossível avaliar se o objetivo foi ou não atingido. Assim, só com objetivos e respetivas métricas bem definidas, bem como o acesso atempado à informação, será possível aos gestores tomar decisões eficazes e com sucesso (Larson, 2012).

Uma empresa necessita de saber como está o negócio, analisar as suas operações e tomar decisões que ajudem a reduzir custos e aumentar vendas e lucros (Sherman, 2015), pelo que o conhecimento é fundamental para o seu crescimento e sucesso. O conhecimento vem da informação disponível e esta, por sua vez, vem dos dados. É função da equipa de BI obter e gerir os dados para dotar os vários setores da empresa com a informação de que necessitam para ganhar conhecimento; conhecimento que os ajuda a

tomar decisões informadas sobre cada passo dado. Esta informação é imprescindível para compreender os clientes, fornecedores, parceiros, concorrentes, etc..

Muitas empresas começam a utilizar ferramentas de BI e informação a nível departamental mas sem estarem incorporadas numa estratégia global da empresa. O desafio está em incorporar a análise de informação no suporte da estratégia de gestão da organização (Davenport & Harris, 2007). Há já suficientes evidências de que as decisões baseadas em análise de informação são mais acertadas do que as decisões tomadas com base na intuição (Davenport & Harris, 2007).

Um estudo de 2013 concluiu que as empresas que utilizam ferramentas analíticas superaram os seus concorrentes por larga margem: duas vezes mais provável que se encontrem no quartil superior do desempenho financeiro; cinco vezes mais provável tomarem decisões muito mais rapidamente do que os concorrentes; três vezes mais probabilidade de executar as decisões como pretendido e duas vezes mais provável que os dados sejam frequentemente utilizados na tomada de decisões (Pearson, 2013).

2.2 INTRODUÇÃO HISTÓRICA BI

Hans Peter Luhn apresentou, em 1958, a primeira definição de 'Business Intelligence'. Para Luhn 'Business' significava um conjunto de atividades realizadas para qualquer finalidade, já 'Intelligence' significava a capacidade de compreender as relações entre os factos apresentados, de modo a direcionar a ação para um objetivo desejado (Luhn, 1958). Ainda numa fase inicial do surgimento dos computadores Luhn antecipava já que a informação disponível teria de ser compreendida para poder ser utilizada na tomada de decisões para alcançar um objetivo.

Mais de 50 anos depois e quando se geram terabytes de dados em cada segundo que passa, o tema Business Intelligence está no topo das atenções das grandes (e não só) empresas. Brian Larson define 'Business Intelligence' como a disponibilização de informação útil e precisa aos decisores, nos momentos adequados, como suporte a tomadas de decisão eficazes (Larson, 2012).

Um dos grandes desafios dum projeto de BI é conseguir transformar dados em conhecimento. A coexistência de inúmeras aplicações operacionais com baixa (ou nenhuma) integração entre elas e a falta de informação histórica comparável são barreiras limitativas à produção de informação útil (Inmon, 2005).

Linhas de números que mostram vendas detalhadas ou dados de produção podem ser extremamente precisas, mas não são 'Business Intelligence' até serem colocadas num formato que possa ser facilmente compreendido por um decisor que precise de as utilizar (Larson, 2012).

Existe uma grande diferença entre dados e informação embora, por vezes, se misturem os termos. Os dados são em bruto e não organizados. Informação são dados que foram organizados, estruturados e

processados. A informação é o que utilizamos para obter conhecimento. Num projeto de Business Intelligence os dados são originados pelos sistemas operacionais, a informação é gerada pelo ETL (extração, transformação e carregamento) e o conhecimento resulta da aplicação da informação em relatórios, visualizações, dashboards, etc.. Um gestor utiliza o conhecimento para tomar as suas decisões (Sherman, 2015).

Uma solução de Business Intelligence é transversal a todas as funções dentro de uma organização, mas só acresce efetivamente valor se for utilizada eficazmente. Há uma correlação entre a boa utilização dum sistema de BI e o desempenho da organização, mas o acesso à informação não implica uma melhoria automática, o fator crítico de sucesso está na análise que as pessoas fazem dos dados e nas ações que consequentemente executam (Howson, 2014). O sistema BI permite que todos os níveis de uma organização consigam aceder a informação, interagir com ela e analisar dados que permitam gerir melhor o negócio, melhorar o desempenho, descobrir oportunidades e operar eficaz e eficientemente (Howson, 2014).

Para o sucesso de um BI não importa apenas a quantidade de dados que se possui mas sim se os utilizamos corretamente. Sem um plano de ação adequado, para produzir conhecimento com valor, os dados irão tornar-se basicamente num elefante-branco caro e inútil (Marr, 2017). Os dados isolados, dispersos por tabelas, relatórios ou outros documentos, sem o contexto do negócio acabam por ser pouco relevantes nas tomadas de decisão (Love et al., 2005). O valor acrescentado por uma solução de BI torna-se mais significativo quando as decisões estão diretamente ligadas à informação, mas nem todas as decisões são possíveis de associar à informação pelo que devem ser escolhidas aquelas que são críticas para o sucesso da organização (Davenport, 2010).

Assim como os negócios estão em constante mudança também as soluções de BI têm de estar desenhadas para acompanhar as mudanças (Kimball et al., 1998). Um sistema de BI deve apresentar informações de forma consistente pelo que deverá (Kimball & Ross, 2013):

- ser capaz de se adaptar às mudanças;
- apresentar a informação atempadamente;
- ser um repositório seguro de informação;
- servir como fonte fidedigna e de confiança para apoio à tomada de decisões;
- ser aceite pela comunidade de utilizadores;

Para que uma solução de BI possa disponibilizar informação a uma organização é imprescindível que os dados que o suportem sejam tratados de modo a conseguir que essa informação seja útil e de confiança. Os dados deverão assegurar os 5 C's ('clean', 'consistent', 'conformed', 'current', 'comprehensive') (Sherman, 2015).

- **Clean** – Os dados deverão ser limpos e validados para não conterem informação errada;

- **Consistent** – Toda a informação do BI terá de ser consistente. Não pode haver 2 valores para a mesma informação;
- **Conformed** – Toda a organização deve utilizar as mesmas dimensões para análise de informação semelhante;
- **Current** – A periodicidade de atualização de informação deve ser a adequada às necessidades das tomadas de decisão;
- **Comprehensive**- Os gestores deverão ter disponível todas a informação de que necessitam para a sua tomada de decisões, independentemente da proveniência e granularidade necessária.

No início de um projeto de BI existem um conjunto de questões às quais é importante saber responder de modo a evitar o insucesso do projeto (Sherman, 2015):

- Que problemas se pretendem resolver com o BI ou desafios querem ser abordados?
- Quem são os utilizadores?
- Quais são os benefícios que a organização irá previsivelmente obter?
- Houve anteriormente alguma iniciativa de implementar um BI que tenha falhado? Porquê?

A resposta a estas questões permitirá aumentar a probabilidade de sucesso do projeto de BI.

Outro dos fatores fundamentais para o sucesso é a existência de um apoio formal por parte da gestão de topo. Por último, mas não menos importante, é fundamental a participação no projeto dos recursos (equipa) mais válidos e capazes para o atingimento dos objetivos previstos, nomeadamente (Sherman, 2015):

- Colaboradores da organização que sejam os utilizadores futuros do sistema de BI;
- Colaboradores conhecedores do negócio e/ou dos sistemas operacionais;
- Colaboradores de IT conhecedores da infraestrutura técnica dos sistemas existentes;

Durante todo o decorrer do projeto é imprescindível manter o apoio da gestão de topo de modo a garantir os recursos necessários, financeiros e humanos, ao desenvolvimento do projeto, bem como intervir na resolução de problemas e outros imprevistos.

A implementação com sucesso de um sistema de BI irá dotar a empresa de uma ferramenta que ajudará e servirá de suporte às tomadas de decisão para a evolução da organização. Do mesmo modo é importante assegurar que o sistema de BI não é estático e consegue integrar novos dados disponíveis na organização, bem como tirar partido de novas funcionalidades disponibilizadas pela evolução das ferramentas existentes.

Pela análise da figura 1 poderá ver-se a evolução das funcionalidades de BI disponibilizadas ao longo do tempo (Howson, 2014) e que estão em constante evolução.

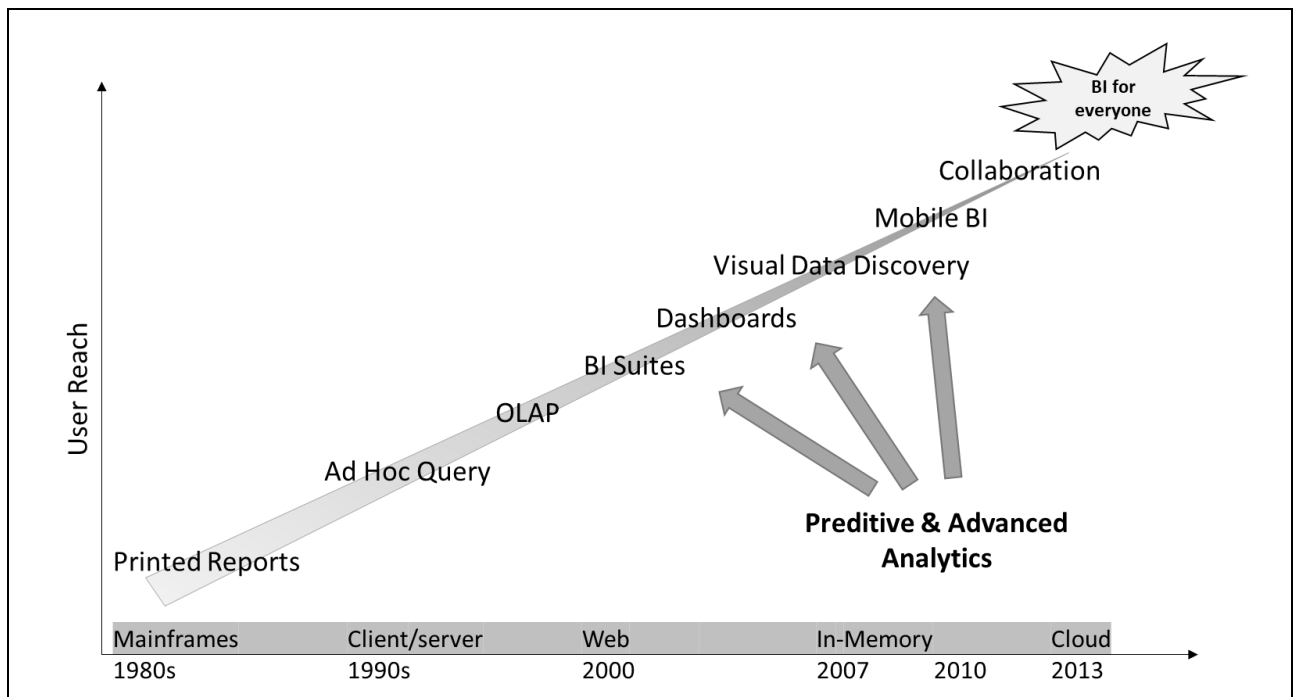


Figura 1: Evolução Funcionalidades de BI (fonte Howson, 2014)

2.3 ARQUITETURA DE BI

A arquitetura de um sistema de BI é como o motor de um carro, um elemento imprescindível, poderoso mas que os utilizadores, como os condutores, não precisam de perceber (Howson, 2014).

O ponto de partida são os sistemas operacionais, são eles os principais geradores de dados para o BI. Os utilizadores dos sistemas de produção, logística, vendas, contabilidade, CRM, etc., registam todas as transações que permitem controlar e gerir os vários setores de uma organização. São estes dados, residentes em sistemas integrados, ou não, e dados de outras fontes: ficheiros, dados externos à organização, logs de website/ecommerce, dados de redes sociais, etc., que permitem a criação de informação estruturada para os utilizadores do BI. Se os dados não forem lidos dos sistemas operacionais então não poderão ser analisados ou se contiverem erros irão originar informação errada.

A maior parte das organizações que inicia o desenvolvimento de um sistema de BI, começa normalmente por extrair os dados dos sistemas operacionais, através de um processo de ETL (Extract, Transform, Load), para uma base de dados com informação relativa a um setor da empresa, designada por Datamart (Kimball et al., 1998). Se os dados contemplam vários setores da empresa poderão dar origem a vários Datamarts ou se a base de dados é única designa-se por Data Warehouse (Inmon, 2005). Com a evolução das soluções de BI existem várias variantes de arquitetura ajustadas às necessidades de cada organização.

Inmon e Kimball defenderam abordagens diferentes para a arquitetura da solução de BI.

Inmon (2005) parte de uma abordagem top-down, criando um modelo de dados EDW (Enterprise Data Warehouse) conforme se apresenta na figura 2, já Kimball (1998) parte de uma abordagem bottom-up, orientada para os requisitos do negócio, para desenhar os datamarts conforme figura 3.

Para Inmon a base da sua arquitetura consiste na recolha dos dados, já devidamente tratados, para um único Data Warehouse. ETLs secundários extraem informação do Data Warehouse para os Datamarts. Para Kimball o Data Warehouse é dispensável, o ETL assegura a extração e transformação dos dados diretamente para os Datamarts.

Inmon defende que o Data Warehouse deverá estar desenhado na terceira forma normalizada mas os datamarts não, já Kimball usa modelos dimensionais nos datamarts sem necessidade de normalização.

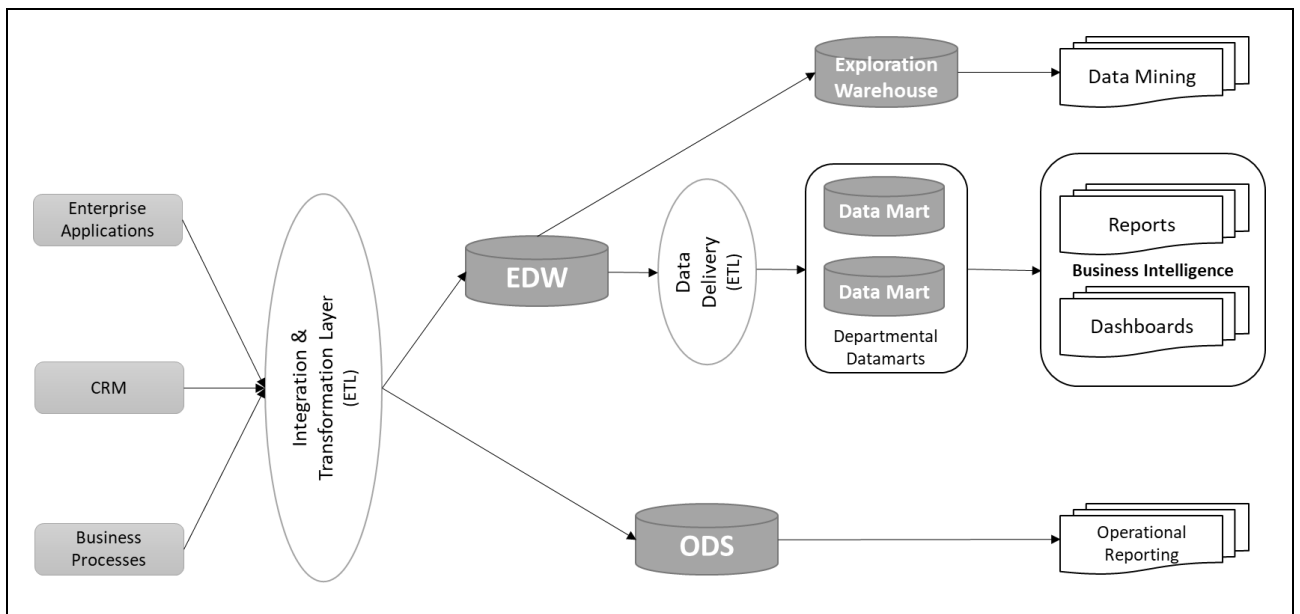


Figura 2: Arquitetura Hub-and-Spoke de Inmon (fonte Sherman, 2015)

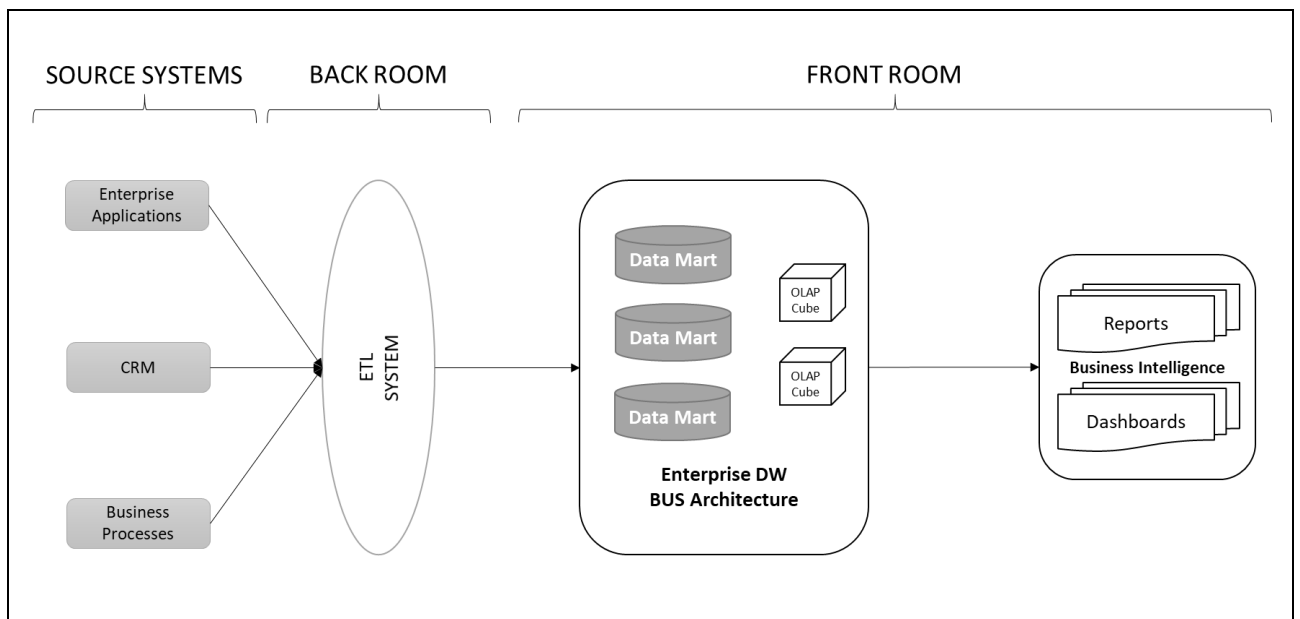


Figura 3: Arquitetura Enterprise Data-Bus de Kimball (fonte Sherman, 2015)

Atualmente a arquitetura Hub-and-Spoke de Inmon é considerada como a melhor prática. Existem no entanto também modelos híbridos, com Data Warehouses na terceira forma normalizada e Datamarts desenhados com modelos dimensionais segundo Kimball (Sherman, 2015; Kimball & Ross, 2013).

Independentemente da solução de armazenamento de dados utilizada podemos considerar que a arquitetura de uma sistema de BI contempla 4 grandes fases de fluxo de dados/informação (Sherman, 2015):

- Dados são criados e atualizados nos sistemas operacionais;
- Dados dos sistemas operacionais são limpos, transformados e integrados;
- Dados são carregados no repositório de informação: Datamarts e Data Warehouse;
- Dados são acedidos pelas ferramentas de BI para reporting e análise.

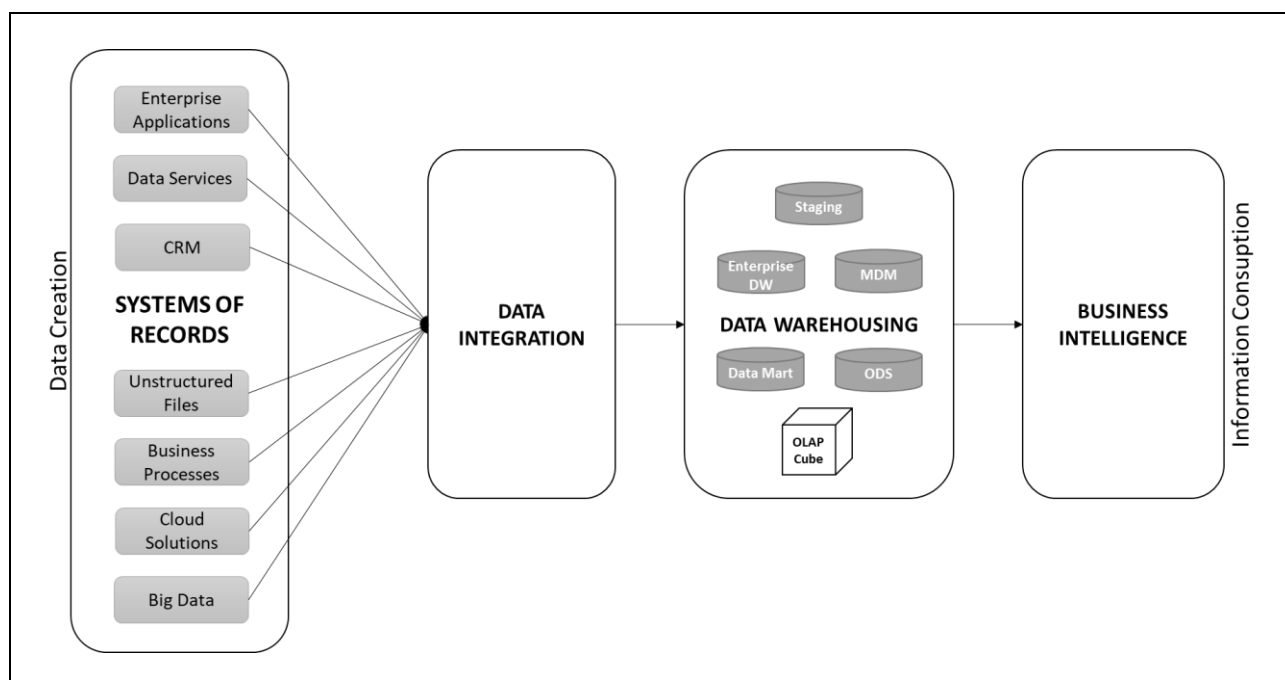


Figura 4: Fluxo de dados num sistema de BI (fonte Sherman, 2015)

Um levantamento detalhado dos dados disponíveis nos sistemas operacionais ou outras fontes de dados, internas ou externas, utilizadas na organização, é imprescindível para poder construir um BI efetivamente útil. Só com este levantamento é possível avaliar e desenhar o ETL adequado à produção de informação útil para os utilizadores.

O processo de ETL é um dos principais consumidores de tempo, especialmente se houver muitas fontes de dados. Numa fase inicial, das soluções de BI, todo o processo de ETL era desenvolvido em código, programado especificamente para a solução. Com o aparecimento de soluções mais evoluídas de BI passaram a estar incluídas ferramentas específicas para esta função (Howson, 2014). No entanto a

diversidade dos dados ou os processos da organização têm impacto na complexidade do ETL e nem mesmo as ferramentas de BI mais evoluídas contemplam todas as necessidades à sua implementação. É importante desconfiar e saber avaliar ferramentas e consultores que se apresentam como possuidores de soluções globais (Sherman, 2015).

Uma das dúvidas recorrentes em projetos de construção de soluções de BI é a necessidade, ou não, de produzir um data warehouse (ou datamarts). Em projetos de pequena dimensão ou de natureza mais operacional é seguida por vezes esta abordagem, com o acesso das ferramentas de visualização e consulta do BI a acederem diretamente aos dados dos sistemas operacionais. No entanto esta prática pode ter impacto relevante nos sistemas operacionais se as consultas forem de alguma complexidade e tratarem volumes grandes de dados. Para projetos de maior dimensão e com objetivos mais focados em indicadores de gestão é importante a construção de um data warehouse, nomeadamente quando:

- Se pretenda efetuar análises transversais ou interfuncionais, que implicam normalmente a utilização de dados em sistemas diferentes;
- Se pretenda efetuar análises temporais ou hierárquicas com informação agregada;
- Se pretenda reporting ou análises rápidas aos dados.

Com o crescimento exponencial dos dados produzidos nas organizações, seja pelo volume, seja pela velocidade a que são produzidos ou pela variedade de dados, é cada vez mais frequente a utilização de ferramentas de Big Data integradas na solução de BI (Howson, 2014). Estas ferramentas permitem dotar as soluções de BI da capacidade de tratar quantidades massivas de dados sejam eles gerados em tempo real ou dados arquivados.

Por fim, o último elemento de uma solução de BI é o front-end que permite aos utilizadores poderem efetuar as análises à informação disponível e com isso obter os elementos que os ajudem nas suas decisões. Um data warehouse pode ser muito poderoso e completo, mas se as ferramentas disponibilizadas aos utilizadores não forem capazes de permitir o acesso intuitivo e fácil à informação então todo o trabalho será em vão (Howson, 2014). As capacidades técnicas das ferramentas são importantes, mas também a facilidade de utilização e interface intuitiva são imprescindíveis ao sucesso e adesão dos utilizadores. Quanto maior for a facilidade de aprendizagem de utilização das ferramentas e a capacidade de obtenção da informação pretendida maior será a probabilidade de sucesso da solução de BI.

Nos últimos 20 anos foram sendo disponibilizadas, pelos vários fabricantes, um conjunto de ferramentas de front-end, que permitem aos vários tipos de utilizadores aceder de acordo com as suas necessidades à informação de que precisam, no momento em que dela necessitam.

Entre as várias ferramentas podemos destacar: Reporting, dashboards, scorecards, consultas aos dados, análise com excel, BI móvel, análise preditiva e prescritiva, etc.

Nem todas as ferramentas são para todos os utilizadores. A cada utilizador devem ser disponibilizadas as ferramentas apropriadas à sua necessidade e devem evoluir à medida que a capacidade e requisitos dos utilizadores evoluem. A intuitividade das ferramentas bem como a facilidade de utilização são importantes, mas não deve ser desprezada a formação adequada aos utilizadores para que possam tirar partido da plenitude de funcionalidades para aceder a mais e melhor informação.

2.4 ETL

O processo de ETL (Extract, Transform, Load) consiste na extração dos dados dos sistemas operacionais e após transformação popular o data warehouse ou datamarts com informação num formato utilizável pelas ferramentas de visualização e análise de informação (Ballard et al., 1998).

O desenvolvimento do processo de ETL é a componente mais complexa e demorada na construção de uma solução de BI (Kimball & Caserta, 2004; Howson, 2014; Kimball & Ross, 2013). Os requisitos do negócio, a multiplicidade de fontes de dados e janelas temporais para processar os dados são alguns dos principais fatores a ter impacto na duração desta tarefa (Kimball & Ross, 2013) que pode representar 60 a 70% do tempo total do projeto (Kimball & Caserta, 2004). O desenho do processo de ETL tem como base o modelo de dados do data warehouse previamente definido (Kimball & Ross, 2013).

As 3 fases dum processo de ETL são: a extração dos dados, a transformação dos dados e o carregamento, após transformação, para o data warehouse ou datamarts.

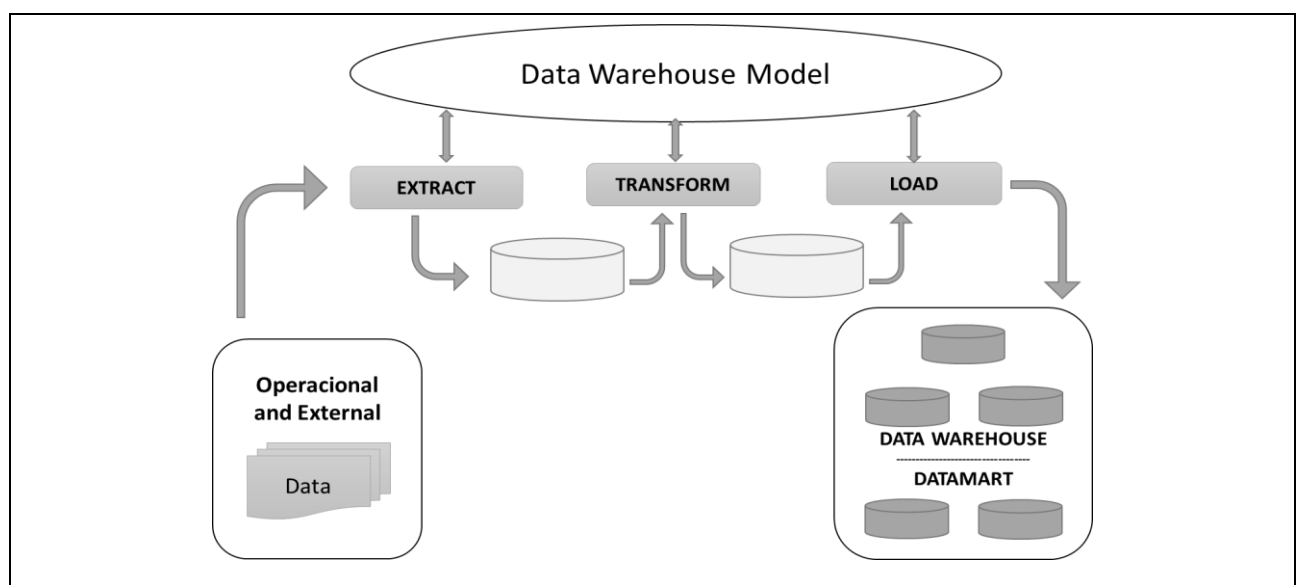


Figura 5: Fases de um processo de ETL (fonte Ballard et al., 1998)

A primeira fase consiste na extração dos dados a partir dos sistemas operacionais ou das fontes originais dos dados. Normalmente são guardados numa área de trabalho (staging área) que pode ser temporária ou permanente.

A segunda fase é a transformação dos dados, onde através de vários procedimentos se procura melhorar a qualidade dos dados seja através de processos de limpeza de potenciais erros, complemento de informação através de cruzamento de fontes ou melhoria da consistência/integridade dos dados.

A terceira fase consiste no carregamento dos dados para o data warehouse (ou datamarts) para que possam ser utilizados pelas várias ferramentas de BI, disponibilizadas aos utilizadores finais para consulta de informação de negócio.

Estas 3 fases do processo de ETL necessitam de ser geridas pelo que existem técnicas e ferramentas utilizadas pela equipa de desenvolvimento (Kimball & Ross, 2013).

Nos primórdios do desenvolvimento de BIs, o processo de ETL era totalmente desenvolvido através de programação à medida de cada projeto. Atualmente os vários fabricantes, de software de desenvolvimento de BI, oferecem já ferramentas que permitem implementar a maior parte das funções do ETL dum modo relativamente rápido, eficaz e de fácil manutenção à posteriori. As ferramentas modernas de ETL permitem extrair os dados de centenas de aplicações, base de dados ou ficheiros de dados com formatos variados.

Uma das primeiras decisões a tomar é o modo como é feita a recolha dos dados, total ou incremental?

Os fatores fundamentais para esta decisão são o volume de dados e a janela temporal para o efetuar. Se o volume é pouco significativo e a janela temporal suficiente, poderá efetuar-se sempre a recolha total dos dados. Já se o volume for muito significativo ou a janela temporal for demasiado pequena, deverá proceder-se apenas à extração dos dados novos (Kimball & Ross, 2013). Num processo de recolha incremental a atualização de dados previamente carregados carece de uma análise e implementação cuidada.

A multiplicação dos sistemas na organizações, quer de gestão, operacional ou de controlo tem feito crescer o volume de dados gerados bem como a velocidade com que são gerados, trazendo dificuldades acrescidas aos sistemas de ETL. Como resposta a estes desafios têm surgido ferramentas de BIG DATA, que podem funcionar integradas no processo de ETL.

Depois de extraídos os dados base, dos sistemas fonte, é necessário proceder à sua transformação para que se possa acrescentar valor à informação a disponibilizar aos utilizadores. Nesta fase os dados são limpos de potenciais erros, são complementados com mais informação e/ou são tratados de modo a apresentar informação coerente e consistente.

A principal fonte de erros dos dados são os próprios sistemas operacionais. Um mau desenho, má utilização ou falta de formação/sensibilização dos utilizadores leva a que existam frequentemente erros nos dados ou dados em falta. De entre os problemas mais frequentes salientam-se:

- Preenchimento dos mesmos dados com formatos diversos (por ex. datas, números, códigos postais, nome de localidades, etc.;
- Tipos de dados incorretos (por ex. campos de texto para escrever números ou datas);
- Falta de indicação de regras de preenchimento;
- Campos vazios por falta de informação no momento de preenchimento ou por não ser importante naquele momento.

O processo de limpeza procurará corrigir o máximo possível destes problemas de modo a que a informação final possa ser útil aos utilizadores.

Outro dos passos da transformação é o conseguir a integridade dos dados e a sua coerência. Pelo facto dos dados serem recolhidos de sistemas diferentes e alguns deles terem informação redundante é necessário garantir a sua coerência. Por outro lado a junção de dados dispersos por vários sistemas permite enriquecer a informação final a disponibilizar (Kimball & Ross, 2013).

O passo final, já depois dos dados devidamente transformados, é o carregamento dos dados para o data warehouse que alimentará as ferramentas de análise e visualização de informação pelos utilizadores.

Este carregamento de dados consiste principalmente no preenchimento das tabelas de dimensões e factos que constituem o data warehouse. O correto preenchimento destas tabelas é fundamental para ter um data warehouse fiável, escalável e de fácil manutenção.

As tabelas de dimensões, embora normalmente mais pequenas que as tabelas de factos, são mais numerosas e fundamentais para o funcionamento do data warehouse. Quanto mais forte for a interligação das dimensões com as tabelas de factos, maior será a capacidade dos utilizadores de obterem informação variada e útil, que não obteriam dos sistemas operacionais. As tabelas de factos têm normalmente maior volume de dados, são elas que contêm os dados para o cálculo de todas as métricas pretendidas pelos utilizadores (Kimball & Ross, 2013).

Uma vez que não é possível garantir que um sistema de ETL apenas trate de dados novos, é imprescindível ter um sistema de ETL preparado para atualização de dados existentes, principalmente nas tabelas de dimensões. Uma atualização de qualquer atributo de uma tabela de dimensões pode ser executada de vários modos, sendo o resultado final dependente do modo escolhido pelo que é importante saber exatamente o que se pretende fazer. Esta implementação designa-se por SCD (Slow Changing Dimensions) e caracteriza-se por ter 3 opções de atualização: type 1, type 2 ou type 3.

Resumidamente a implementação destas 3 opções é a seguinte:

- type 1, os dados antigos são substituídos pelos novos;

- type 2, é adicionada uma nova linha na tabela de dimensões sendo que os factos anteriores a esta alteração ficam associados à linha anterior e os novos à nova linha. Mantem-se o registo de datas em cujo intervalo o registo é válido;

- type 3, é criado um novo atributo (coluna) para manter a informação anterior mas a corrente substitui o atributo anterior;

Existem ainda mais opções de gestão de SCD, type 4 a 7 mas saem foram do âmbito deste trabalho (Kimball & Ross, 2013).

Por último é importante referir que um sistema de ETL tem de ter implementada uma rotina de atualização de dados de acordo com as necessidades da organização, mas também ajustada às questões técnicas muito dependentes do volume e disponibilidade dos sistemas operacionais. Um sistema de ETL com funcionamento desadequado pode ter impacto no desempenho dos sistemas operacionais ou apresentar informação errada ou incoerente.

A gestão de todo o sistema de ETL deve garantir a fiabilidade da informação, a sua disponibilidade bem como a sua capacidade de evolução perante as necessidades da empresa (Kimball & Caserta, 2004). Um sistema de BI não é estático, deve acompanhar as necessidades da organização e o processo de ETL é o seu coração.

2.5 DATA WAREHOUSE

A primeira questão que se põe aquando do desenho de um data warehouse é qual o modelo a utilizar. Existem dois modelos que são a base da maioria das soluções de BI: o modelo relacional defendido por Inmon e o modelo dimensional defendido por Kimball. O modelo relacional tem como base a distribuição dos dados por um conjunto de tabelas, com relações devidamente estabelecidas entre si e devidamente normalizadas na terceira forma normal (Inmon, 2005). Já o modelo dimensional pressupõe a existência apenas de tabelas de dimensões e tabelas de factos, sendo que a normalização se aplica apenas às tabelas de factos (Sherman, 2015).

O diagrama de relações de um modelo dimensional tem muito menos tabelas e relações que um modelo relacional, é muito mais simples e a sua interpretação é mais fácil. O modelo relacional está mais orientado para sistemas operacionais que necessitam de tratar os dados (criar, modificar e atualizar) em tempo real, já no modelo dimensional a atualização só ocorre no carregamento dos dados e o principal objetivo é o desempenho na consulta dos dados pelos utilizadores (Sherman, 2015). No modelo relacional não existe informação redundante pelo que a atualização dos dados contém pouco volume de informação e ocorre instantaneamente (Howson, 2014), já as consultas de informação são menos eficientes pelo facto dos dados se encontrarem dispersos por várias tabelas. No modelo dimensional acontece o oposto, existe muita redundância o que torna a atualização mais demorada e o volume de dados maior, no entanto a

consulta de informação (o principal objetivo do BI) é muito mais eficiente permitindo uma maior capacidade de resposta à consulta de informação dos utilizadores, seja ela agregada, filtrada ou detalhada.

A facilidade de interpretação do modelo de dados dimensional e o seu melhor desempenho tornaram-no no modelo mais utilizado no desenho dos data warehouse das soluções de BI (Sherman, 2015) (Kimball & Ross, 2013).

Factos e Dimensões

Os dados utilizados numa solução de BI podem ser divididos em 4 categorias: medidas, dimensões, atributos e hierarquias; (Larson, 2012). Estas 4 categorias são a base do desenho do modelo dimensional.

As medidas são os factos que a organização quer conhecer e medir, podem ser vendas, compras, inventário, etc. são a base e o principal propósito de um BI. Os factos podem ser analisados a nível individual, agregados ou calculados. É esta informação que se quer conhecer para servir de suporte às tomadas de decisão (Kimball & Ross, 2013; Sherman, 2015).

As dimensões permitem estabelecer o contexto das medidas que se estão a analisar (Kimbal & Ross, 2013). Não se pretende saber apenas o valor total de uma medida qualquer. Numa organização poderá ser necessário analisar as vendas, por exemplo, por ano, ou por região, ou por departamento, etc. As várias dimensões associadas aos factos respetivos possibilitarão várias visões de análise da mesma informação.

Os factos e as dimensões são suportados em tabelas distintas. Num data warehouse existem normalmente uma ou mais tabelas de factos e várias tabelas de dimensões. Normalmente as tabelas de dimensões têm menor quantidade de registos do que as tabelas de factos, que contêm todo o detalhe dos dados que se pretendem medir.

<table border="1"><thead><tr><th>Product Dimension</th></tr></thead><tbody><tr><td>Product Key (PK)</td></tr><tr><td>SKU Number (Natural Key)</td></tr><tr><td>Product Description</td></tr><tr><td>Brand Name</td></tr><tr><td>Category Name</td></tr></tbody></table>	Product Dimension	Product Key (PK)	SKU Number (Natural Key)	Product Description	Brand Name	Category Name	<table border="1"><thead><tr><th>Retail Sales Facts</th></tr></thead><tbody><tr><td>Date Key (FK)</td></tr><tr><td>Product Key (FK)</td></tr><tr><td>Store Key (FK)</td></tr><tr><td>Promotion Key (FK)</td></tr><tr><td>Customer Key (FK)</td></tr><tr><td>Clerk Key (FK)</td></tr><tr><td>Transaction #</td></tr><tr><td>Sales Dollars</td></tr><tr><td>Sales Units</td></tr></tbody></table>	Retail Sales Facts	Date Key (FK)	Product Key (FK)	Store Key (FK)	Promotion Key (FK)	Customer Key (FK)	Clerk Key (FK)	Transaction #	Sales Dollars	Sales Units
Product Dimension																	
Product Key (PK)																	
SKU Number (Natural Key)																	
Product Description																	
Brand Name																	
Category Name																	
Retail Sales Facts																	
Date Key (FK)																	
Product Key (FK)																	
Store Key (FK)																	
Promotion Key (FK)																	
Customer Key (FK)																	
Clerk Key (FK)																	
Transaction #																	
Sales Dollars																	
Sales Units																	

Figura 6: Tabela dimensão produto,
(fonte Kimball & Ross, 2013)

Figura 7: Tabela factos vendas,
(fonte Kimball & Ross, 2013)

Uma tabela de factos contem campos de dois tipos: chaves e medidas, as chaves permitem relacionar com a dimensão respetiva e as medidas que são os factos da organização que se pretendem analisar. Existe também, quase sempre, uma chave do tipo data que situa o facto no tempo.

Existem 3 tipos de factos que importa distinguir:

Aditivos – Que se podem agregar e somar em todas as dimensões, vendas por exemplo.

Semiaditivos – Podem ser somados em algumas dimensões, mas não todas, tal como saldos de contas bancárias;

Não aditivos – Não podem ser somados, preços unitários ou temperaturas por exemplo.

Uma tabela de dimensões contem uma chave e um conjunto de atributos relativos à dimensão. O atributo principal é normalmente a descrição associada a cada um dos registos da dimensão sendo os restantes informação complementar que possibilitará, aos utilizadores, efetuar análises mais detalhadas da informação (Larson, 2012). Na figura anterior ‘Brand Name’ e ‘Category Name’ são atributos da dimensão ‘Product’.

Por fim, uma hierarquia é uma estrutura de 2 ou mais níveis dentro de uma dimensão. Um exemplo de hierarquia poderá ser ‘City / Region / Country’. A utilização de hierarquias permite aos utilizadores efetuarem análises com vários graus de granularidade e efetuar ‘drill-up’ or ‘drill-down’ da informação (Larson, 2012).

É comum dizer-se que uma solução de BI será tanto melhor quanto maior for o detalhe dos seus atributos e a capacidade analítica dependerá dos níveis das suas hierarquias (Kimball & Ross, 2013).

Star Schema vs Snowflake Schema

As tabelas de factos e dimensões são armazenadas em bases de dados relacionais usando um de dois esquemas de relações: star schema ou snowflake schema.

Numa organização star schema a tabela de factos liga-se através de uma chave a cada uma das dimensões possibilitando assim vários contextos de análise das medidas existentes na tabela de factos.

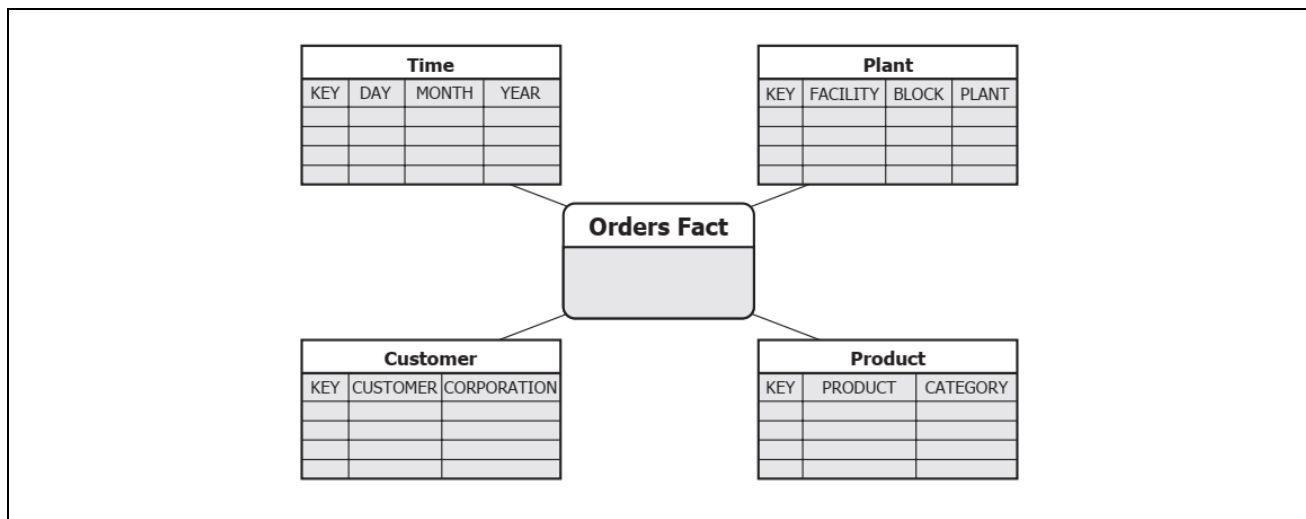


Figura 8: Star Schema, (fonte Howson, 2014)

A simplicidade desta organização tem benefícios significativos na performance da consulta aos dados. As consultas são simples, envolvem poucas tabelas e todas elas se relacionam apenas com a tabela de factos, obtendo-se os resultados bastante rápidos devido à excelente capacidade dos motores de base de dados para lidar com este tipo de consultas.

No star schema as dimensões não estão normalizadas contendo todos os atributos e hierarquias numa única tabela. Já no snowflake schema as dimensões são normalizadas, resultando num desdobramento de tabelas em função das hierarquias existentes (Sherman, 2015). Esta normalização tem a vantagem de eliminar dados redundantes e utiliza-se, por vezes, em data warehouses com dimensões de elevado volume de registos ou hierarquias com muitos níveis. Neste esquema as consultas envolvem mais tabelas, inter-relacionadas, resultando em consultas bastante mais pesadas com impacto na performance para os utilizadores.

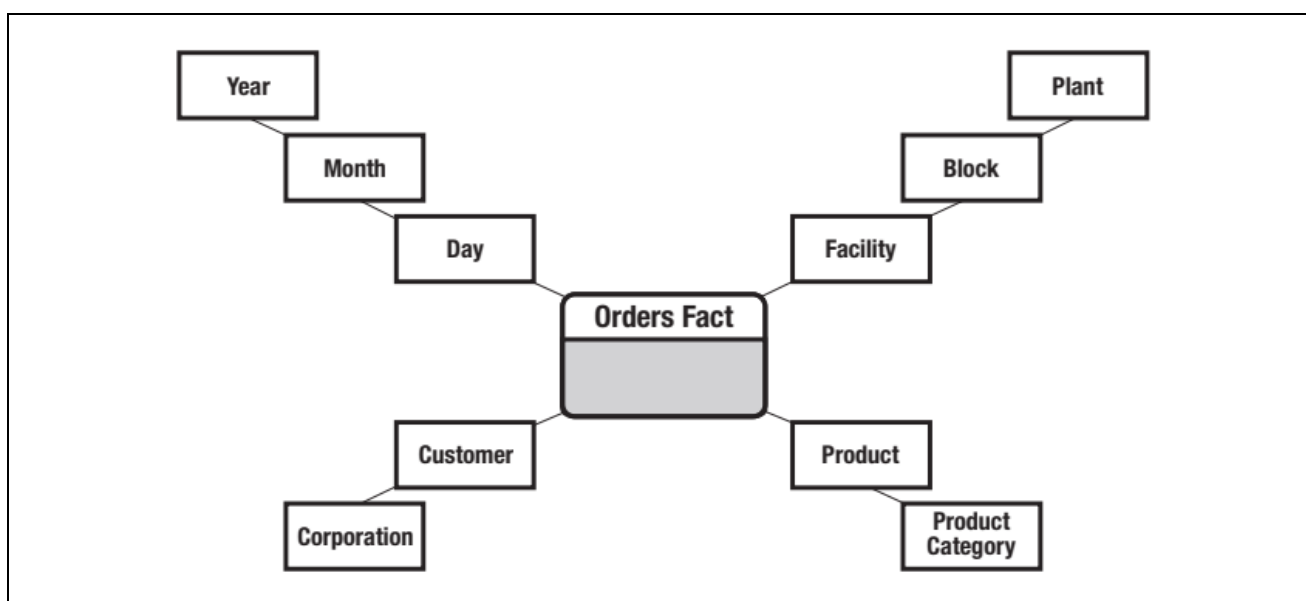


Figura 9: Snowflake Schema, (fonte Howson, 2014)

Pela simplicidade de interpretação e pelas vantagens de desempenho utiliza-se, maioritariamente, star schema no modelo dos data warehouses (Sherman, 2015; Kimball & Ross, 2013; Larson, 2012; Howson, 2014).

2.6 VISUALIZAÇÕES E DASHBOARDS

A construção do ETL e o modelo de data warehouse são a base de uma solução de BI mas as ferramentas de front-end são aquelas que os utilizadores finais vão utilizar e permitirão tirar partido de todo o trabalho anteriormente efetuado. Existe um conjunto variado de ferramentas para tirar partido dos dados, umas mais simples outras mais complexas, mais estruturadas ou mais flexíveis mas todas elas com o objetivo de dar aos utilizadores a informação mais adequada à sua necessidade. Existem vários tipos de utilizadores numa organização pelo que dificilmente haverá uma única ferramenta adequada a todos (Sherman, 2015), pelo que é importante avaliar necessidades, competências e encontrar as melhores soluções. Um data warehouse pode ser perfeito, mas se os utilizadores não conseguirem extrair dele a informação que necessitam para apoio às suas decisões, de pouco servirá e rapidamente estará condenado ao fracasso (Howson, 2014).

Desde o início das primeiras soluções de BI que as ferramentas para o utilizador têm evoluído e outras têm aparecido, permitindo modos diferentes de explorar a informação existente. Resumidamente podemos identificar e descrever algumas das principais ferramentas existentes de acordo com a figura seguinte.

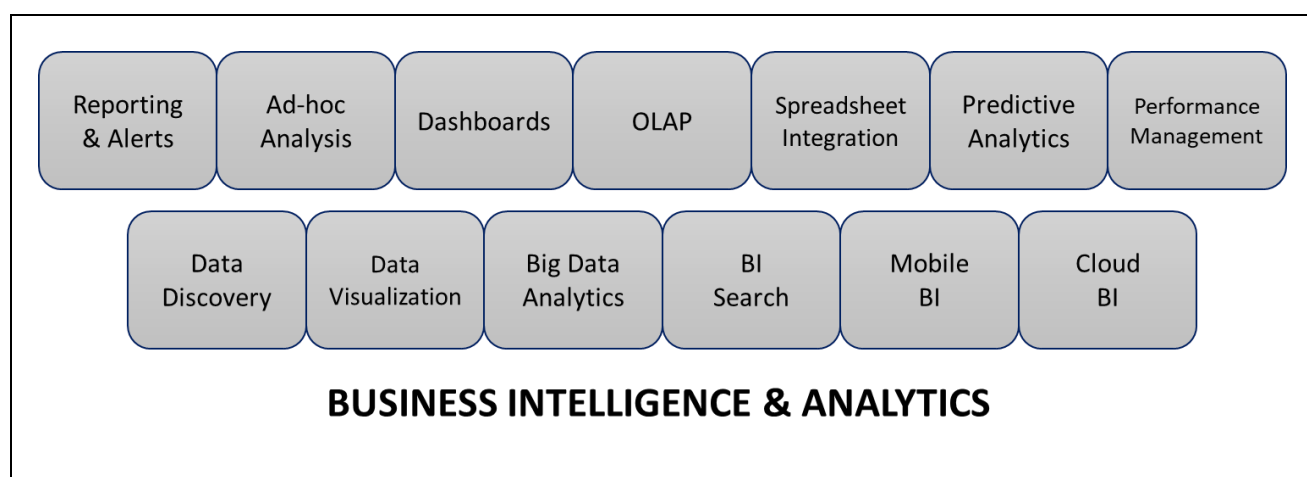


Figura 10: Tipos de ferramentas de BI (utilizador final), (fonte Sherman, 2015)

Reporting

É a forma mais tradicional e mais difundida para disponibilização de informação nas empresas. Podem, normalmente, ser obtidos on-line ou gerados off-line e enviados via e-mail, periodicamente ou despoletados por acontecimentos nos dados. Estáticos nos primórdios do BI mas atualmente já com capacidades de filtragem e/ou drill-down.

Ad-hoc Analysis

Relatórios personalizados onde o próprio utilizador escolhe os dados que quer visualizar. Tem a vantagem do utilizador não depender do IT para construir relatórios adequados à sua necessidade. Pode servir como base exploratória de dados.

Dashboards

Vários (poucos) painéis numa única página/écran. Combinação de gráficos e tabelas de dados dos principais indicadores da empresa ou departamento. Possibilidade de drill-down para obter mais detalhe na informação. Pode ser atualizado periodicamente ou por ação do utilizador. É habitual o uso de cores para mais facilmente salientar dados positivos ou negativos.

OLAP Analysis

Abordagem mais estruturada para obtenção de relatórios ad hoc, em que os utilizadores podem selecionar combinações de dimensões e medidas para a sua análise. Permite possibilidades de consulta:

- Drill-down: obtenção de informação mais detalhada;
- Drill-across: análise através de múltiplas dimensões;
- Slice and dice: usar drill-down e drill-across simultaneamente.

Este tipo de análise estava normalmente associada à construção de cubos OLAP implementados em sistemas próprios, mas atualmente já não é um requisito uma vez que a capacidade de BDs relacionais, tabulares ou em-memória já permite este tipo de funcionalidades.

Spreadsheet Integration

A maior parte dos BIs atuais permite a integração dos dados com folhas de cálculo onde os utilizadores poderão explorar os dados. Tem a vantagem de ser uma ferramenta que os utilizadores já conhecem. Normalmente é possível o refrescamento dos dados a partir da sua fonte.

Predictive Analytics

Diversas ferramentas para utilização dos dados históricos para construção de modelos preditivos e/ou determinar tendências nos dados.

Performance Management

Ferramentas para ajuda na otimização, gestão e medição da performance de gestão: planeamento, orçamento e consolidação financeira.

Data Discovery

É uma evolução das análises OLAP tendo como base a utilização de ferramentas de perceção visual que possibilitam uma navegação/descoberta mais intuitiva dos dados.

Data Visualization

Consiste na representação gráfica da informação e dos dados, utilizando elementos visuais como gráficos e mapas, possibilitando ver e compreender tendências, outliers e padrões nos dados (Tableau, 2020).

Big Data Analytics

Conjunto de metodologias e ferramentas para analisar e tratar dados que não podem ser tratados pelas tecnologias tradicionais, quer seja pelo seu elevado volume, pela sua velocidade de geração e análise em tempo-real quer pela variedade do tipo de dados.

BI Search

Funcionalidade recente em alguns sistemas que permite ao utilizador escrever em linguagem ‘natural’ uma questão com a informação que se pretende obter. A ferramenta ‘interpreta’ a questão, analisa os dados e produz um resultado sobre a forma de uma tabela ou gráfico.

Mobile BI / Cloud BI

Refere-se à possibilidade de utilizar as funcionalidades de BI através de dispositivos móveis ou residente na cloud.

Nenhuma solução de BI tem todas as ferramentas atrás referidas nem nenhuma organização as utiliza na totalidade. É importante perceber quais as que se adequam aos objetivos da organização bem como à capacidade/conhecimento dos utilizadores.

3. ANÁLISE DE DADOS E REQUISITOS

3.1 ARQUITETURA DE SISTEMAS OPERACIONAIS

O sistema de informação do COMPETE2020 assegura, de forma fiável e validada, a recolha, registo e armazenamento de todos os elementos necessários para monitorização, avaliação, gestão financeira, verificação e auditoria. A sua operacionalização, em todas as dimensões e ao longo de todo o ciclo de vida das operações, permite criar a capacidade de disponibilização de toda a informação que a AG necessita para cumprimento das suas responsabilidades, perante entidades nacionais e comunitárias.

O sistema de informação do COMPETE2020 caracteriza-se por ser um sistema de rede, comum e centralizado, onde toda a informação requerida e produzida é integralmente registada, com validação prévia à entrada no sistema. Deste modo, o sistema de informação é a única base de informação do Programa, quer para efeitos de relacionamento com as autoridades nacionais, quer com a Comissão Europeia. O processo de recolha de dados é suportado em ferramentas eletrónicas, as quais dialogam diretamente com o sistema de informação.

A existência de formulários eletrónicos prevê e assegura a recolha dos dados da responsabilidade do beneficiário, interagindo, quando aplicável, com bases de dados disponíveis na Administração Pública, designadamente no que diz respeito aos elementos de identificação e caracterização do candidato ou beneficiário. Esta recolha de dados junto do beneficiário é realizada através de diferentes tipologias de formulários, com especificidades próprias do sistema de apoio em causa, designadamente o formulário de candidatura e o formulário de pedido de pagamento, este último complementado com informação adicional requerida aquando do pedido de pagamento final da operação, nomeadamente informação correspondente a resultados da operação.

Para além dos formulários eletrónicos, a AG dispõe também de ferramentas informáticas que suportam e normalizam todas as análises nas diferentes fases do ciclo de vida da operação, as quais são incorporadas no sistema de informação do COMPETE 2020, no caso dos projetos financiados FEDER e FC.

Todavia, tendo em conta a arquitetura global do PT2020, o sistema de informação do COMPETE2020 é parte integrante do SI PT 2020.

Nos termos do Acordo de Parceria, o SI PT 2020 assegura a informação, relativa ao conjunto dos FEEI através do Balcão 2020, do repositório geral de dados e do registo de dívidas, assim como, no caso dos fundos de coesão, a integração da informação relativa a pagamentos e recuperações, controlo e auditoria.

O Balcão 2020 constitui, assim, o ponto de acesso aos Programas Operacionais financiados pelos Fundos Europeus Estruturais e de Investimento (FEEI), para todas as entidades que pretendam candidatar a financiamento as suas operações, assegurando a interoperabilidade entre os intervenientes. O sistema de informação da AG assume funcionalidades que asseguram os processos de análise, de gestão dos

programas e a interoperabilidade com o Balcão2020 e a ADCoesão, nas suas vertentes de coordenação, certificação, pagamentos e auditoria.

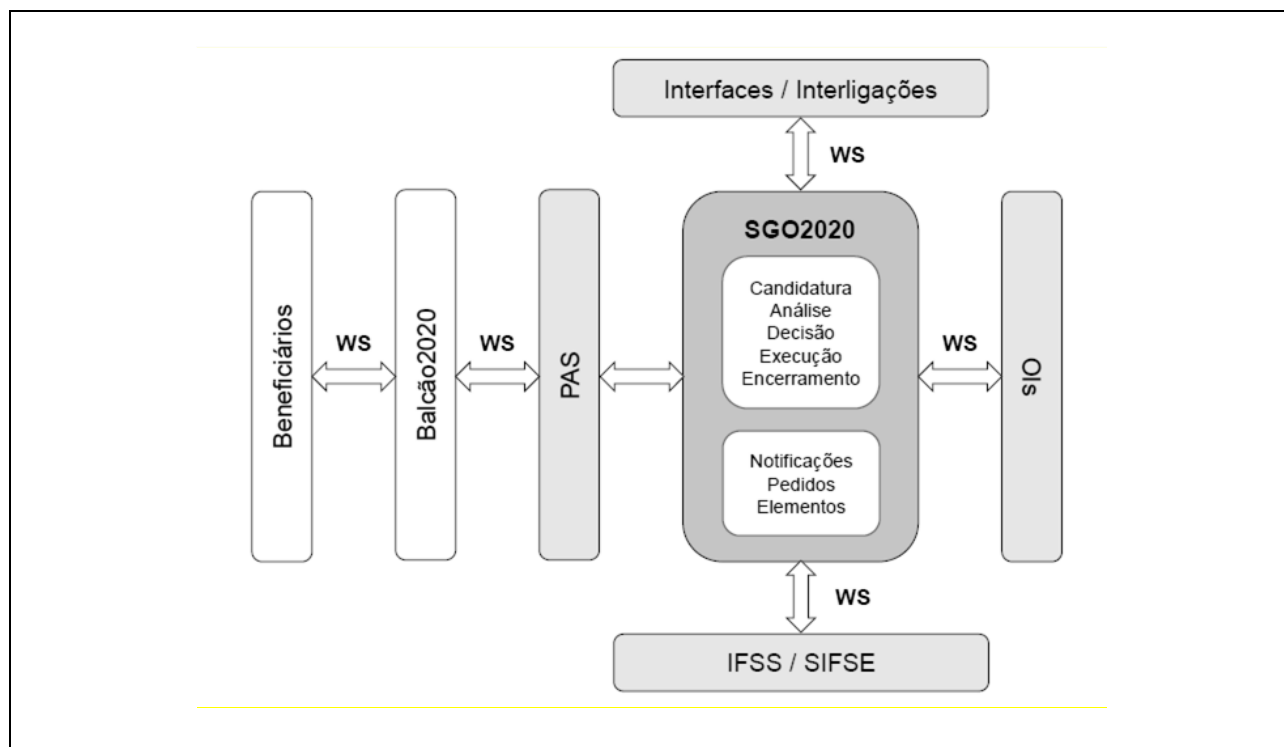


Figura 11: Fluxo de informação dos Sistemas de informação do Compete2020

Sistemas Operacionais

A Arquitetura de Sistemas de Informação do Compete2020 é constituída por um conjunto de aplicações internas e externas. Foi solicitado ao Departamento de IT do Compete2020 um conjunto de informação sobre todas as aplicações utilizadas, mas a informação disponibilizada foi apenas relativo ao SGO e ReAD, pelo que o levantamento aqui descrito apresenta algumas limitações em termo de detalhe.

A principal aplicação operacional é o SGO, onde é efetuado o registo de toda a informação relativa aos projetos candidatados, bem como o seu acompanhamento. Nesta aplicação são registadas as candidaturas, decisões de aprovação, acompanhamento de execução, pagamentos e encerramento. Conta com cerca de 660 utilizadores ativos e os seus dados estão distribuídos entre base de dados SQL, ficheiros XML e ficheiros anexados (pdf, doc, etc.).

O SIFSE é um sistema externo, centralizado e que se destina a gerir os projetos apoiados pelo FSE.

O PAS destina-se ao acesso por parte dos beneficiários a registo e consulta de informação.

Existe uma aplicação de registo de faturas para controlo de todas as faturas relativas aos projetos aprovados e pagamentos respetivos.

São ainda utilizados sistemas externos, via acesso remoto ou utilização de webservice, para obtenção/validação de informação complementar, nomeadamente:

- ADC / Balcão 2020– dados dos beneficiários, situação de dividas
- IAPMEI – Certificação PME,
- InformaDB, outros....
- Balcão 2020- candidaturas fundo coesão,
- SIAUDIT,
- SPTD,
- Assiduidade interna (SGE)

O ReAD é uma solução de reporting/consulta de informação para produção de relatórios e análises de dados com base em informação recolhida de SGO e SIFSE. No ReAD não é registada qualquer informação, os seus dados são extraídos periodicamente dos Sistemas SGO e SIFSE e armazenados numa base de dados SQL.

A análise dos dados possibilitou verificar que é guardado um registo mensal, por projeto, que contém um conjunto significativo de dados de cada projeto, incluindo beneficiário, datas, valores e várias classificações. É com base nessa informação que é emitido o relatório mensal do Compete2020 e disponibilizadas consultas (via SGO) para pesquisa e análise de informação. Esta informação pode ser também exportada para Excel de modo a servir de base de trabalho para análises específicas pelos vários departamentos do Compete2020.

Tendo em conta que a solução de BI a implementar irá complementar o reporting mensal, era imprescindível que todos os dados apresentados pelos dois sistemas fossem coerentes. O facto de não existir documentação técnica, nem conhecimento interno sobre as estruturas de dados utilizadas no SGO, impossibilitou que se utilizasse este sistema para extrair os dados para o desenho do modelo de dados do data warehouse, pelo que se optou por considerar a base de dados do ReAD com fonte de dados base para o ETL da futura solução de BI.

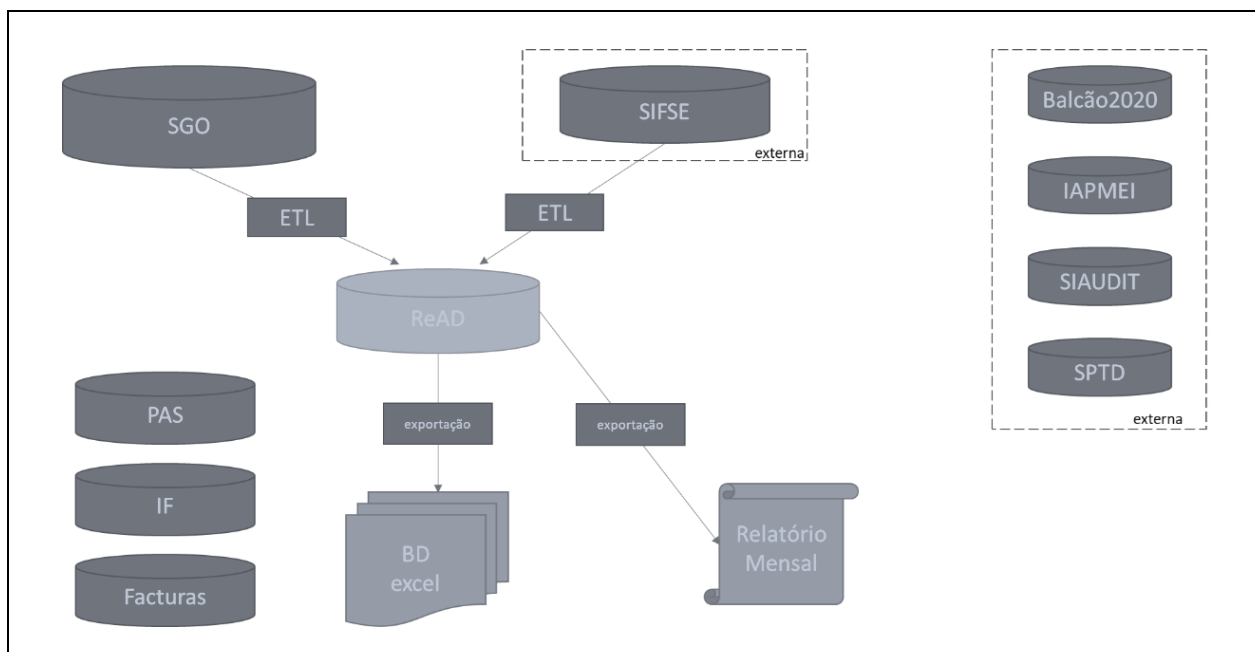


Figura 12: Arquitetura de Sistemas de informação do Compete2020

Repositório de dados das aplicações internas e externas

As aplicações externas têm os seus próprios repositórios de dados e não existe acesso direto nem conhecimento da sua estrutura. O acesso a estes sistemas é feito de modo remoto ou via web services. Dos vários sistemas externos utilizados apenas o SIFSE alimenta a base de dados do ReAD, via web services, mas não foi disponibilizada documentação técnica sobre o procedimento.

Relativamente às aplicações internas, todas têm um servidor próprio. Neste servidor além da base de dados SQL específica de cada aplicação, estão também armazenados ficheiros de dados se eventualmente a aplicação os utiliza, nomeadamente (pdf, xls, xml, doc, etc.).

De todas as aplicações apenas foi disponibilizada informação do SGO e READ.

Nome Aplicação	SGO	ReAD
Descrição Macro das funcionalidades	Sistema Operacional – Decisão, Execução e Encerramento de candidaturas	Produção de Relatórios e Análises de Dados com base em informação recolhida de SGO e SIFSE
Número de Utilizadores e serviços/ departamentos	660 utilizadores ativos Cerca de 15 Organismos	Os mesmos do SGO. (o acesso a este sistema é feito a partir do SGO)
Tipo de Dados (Base Dados, Ficheiros: texto, XML, xls, etc)	Informação em BD, ficheiros estruturados - xml Anexos - outros formatos, pdf, doc, etc	Informação em BD.
Dimensão: Número de tabelas BD	253	178

Ou Número de ficheiros			
Volume de dados	Atualmente 2.5 TB Base de dados aprox. 1 GB (gz)		1.5 GB
Integração com outros sistemas: dados importados	Web services ADC / IAPMEI		Web services SGO para importação de dados
Integração com outros sistemas: dados exportados	Web services disponíveis para todas AG e OI		Exportação para Excel

Tabela 1: Informação técnica SGO e ReAD (fonte: Departamento IT Compete2020)

3.2 QUADROS COMUNITÁRIOS ANTERIORES

Sob a responsabilidade do Compete2020 existem também dados relativos a programas anteriores residentes em servidores próprios. Não foi disponibilizada qualquer informação sobre estes sistemas, pelo que os seus dados não foram considerados no desenho do modelo de dados nem na definição da futura ferramenta de Business Intelligence. A integração futura destes dados na solução carece de uma análise detalhada, com possível impacto em toda a solução de BI. Ressalvando as questões legais quanto ao período em que os sistemas anteriores terão de estar disponíveis para consulta, poderá o BI apresentar-se como uma alternativa para consulta de informação histórica sem exigência da manutenção em estado operacional dos servidores antigos.

3.3 DIFICULDADES ENCONTRADAS

No decurso do projeto foram várias as dificuldades com que a equipa se confrontou e que condicionaram o decorrer dos trabalhos, quer na duração quer na obtenção dos resultados pretendidos. A falta de informação relativamente às estruturas de dados das aplicações operacionais não permitiu fazer um desenho do modelo de dados de acordo com as melhores práticas e com isso o resultado final ficou aquém do desejável.

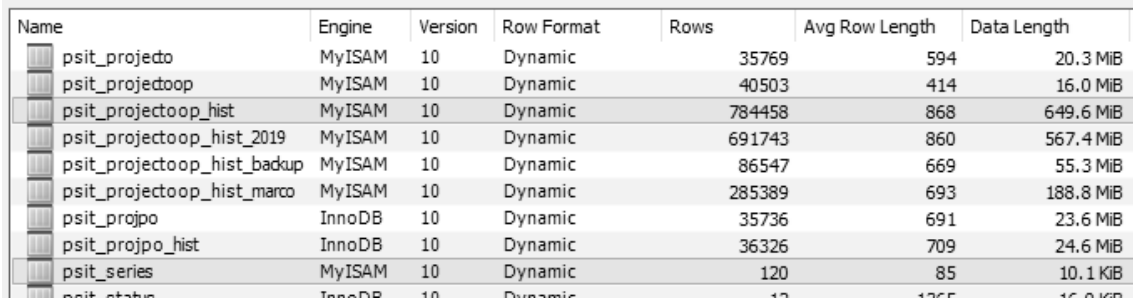
Numa solução de Business Intelligence é importante ter acesso aos dados base dos sistemas transacionais (neste caso seriam principalmente o SGO e SIFSE) e conhecer a estrutura de dados e seu significado.

Relativamente ao SGO foi disponibilizada uma cópia completa da base de dados, no entanto não existe nenhuma informação técnica sobre os dados nem diagrama de relações das tabelas da base de dados. Em face das limitações existentes não foi possível desenhar o modelo de dados com base nas tabelas do SGO.

Em último recurso decidiu-se que os dados a usar seriam os mesmos que as ferramentas internas de reporting utilizavam (ReAD). Este caminho permitiria desenhar um modelo de dados e um protótipo de BI cuja coerência poderia ser validada com o reporting existente, mas tinha como limitação não ter os dados necessários para dar resposta a muitos dos indicadores identificados na fase de levantamento. Não sendo esta a solução final pretendida, entendeu-se mesmo assim que seria o melhor caminho a seguir de modo a atingir parte dos objetivos pretendidos e demonstrar a validade e utilidade da solução de BI.

3.4 ANÁLISE DOS DADOS BASE DO READ

A base de dados do ReAD está alojada num servidor MySQL e é constituída por várias tabelas (99) mas as mais relevantes para a solução de BI são a `psit_projectoop_hist` e a `psit_series`.



Name	Engine	Version	Row Format	Rows	Avg Row Length	Data Length
psit_projecto	MyISAM	10	Dynamic	35769	594	20.3 MiB
psit_projectoop	MyISAM	10	Dynamic	40503	414	16.0 MiB
psit_projectoop_hist	MyISAM	10	Dynamic	784458	868	649.6 MiB
psit_projectoop_hist_2019	MyISAM	10	Dynamic	691743	860	567.4 MiB
psit_projectoop_hist_backup	MyISAM	10	Dynamic	86547	669	55.3 MiB
psit_projectoop_hist_marco	MyISAM	10	Dynamic	285389	693	188.8 MiB
psit_projpo	InnoDB	10	Dynamic	35736	691	23.6 MiB
psit_projpo_hist	InnoDB	10	Dynamic	36326	709	24.6 MiB
psit_series	MyISAM	10	Dynamic	120	85	10.1 KiB
psit_status	InnoDB	10	Dynamic	12	125	1.0 KiB

Figura 13: Tabelas ReAD

A tabela `psit_projectoop_hist` contém 95 campos de dados (anexo I). Durante o projeto não foi disponibilizado qualquer documento técnico sobre o conteúdo de cada campo. Apenas foi disponibilizado um ficheiro associando as colunas do relatório ‘Ponto de situação mensal’, os respetivos campos da tabela SQL e a as regras de filtragem. Existe na tabela um conjunto adicional de campos que podem ser utilizados para futuros indicadores, mas carecem de uma adequada caracterização e interpretação junto da equipa de desenvolvimento.

A tabela `psit_projectoop_hist` contém o registo mensal do ponto de situação de cada projeto. Existe um registo de cada projeto que apresenta o ‘estado’ do projeto no final do mês. Do ponto de vista financeiro todos os valores apresentados refletem sempre o valor acumulado no final de cada mês. Esta tabela continua, à data de 5/3/2020, 784.458 registos sendo o número médio mensal de 46.000 registos. O primeiro mês com registos é outubro de 2018.

A tabela `psit_series` contém o valor mensal acumulado dos valores mais importantes para o reporting do Compete2020, nomeadamente valor de investimento candidatado, incentivo: aprovado, contratado, pago e executado bem como valor de pagamentos: certificados, não certificados e reembolsados. Esta tabela contém 120 registos, um por mês, desde maio de 2015. Não foi disponibilizada qualquer informação sobre o modo de cálculo desta tabela. Tentou-se replicar estes dados usando os registos da tabela `psit_projectoop_hist` mas os valores obtidos foram ligeiramente diferentes para a maioria dos meses. Num projeto futuro será importante analisar o porquê destas diferenças. Para este projeto tomou-se como correta a informação da `psit_series` e para obtenção dos valores mensais trabalhou-se com as diferenças entre os meses.

3.5 LEVANTAMENTO DE INDICADORES DE NEGÓCIOS

Atualmente a gestão do Compete2020 tem ao seu dispor dois relatórios mensais (um para projetos Compete2020 e outro para projetos Sistema Incentivos) em formato excel/pdf. Os relatórios apresentam o ponto de situação mensal relativamente ao acompanhamento dos projetos apoiados. São relatórios estáticos com alguns dos indicadores mais importantes para a gestão, mas sem qualquer possibilidade de filtragem/análise por dimensão.



Figura 14: exemplos Relatório 'Sistema Incentivos'

Por outro lado, no ReAD é possível efetuar um conjunto significativo de pesquisas/filtragens de dados, mas o resultado obtido é sob a forma de listagens/exportações para excel.

Como consequência, da pouca versatilidade das ferramentas utilizadas atualmente, muitas das necessidades da gestão obrigam à exportação e tratamento de dados em excel para obtenção de indicadores específicos necessários.

A solução de BI prevista tem como objetivo dar resposta à maior parte dos indicadores necessários atualmente, bem como estar preparada para dar resposta a futuros indicadores. Os dados operacionais disponíveis e o modelo de dados desenhado são cruciais para atingir os objetivos propostos.

Na primeira parte do projeto efetuaram-se um conjunto de reuniões com os vários departamentos de modo a poder identificar os indicadores respetivos necessários. Com o apoio da equipa do Compete2020 foram identificados os mais importantes e descritos no anexo II.

O anexo II inclui 187 indicadores, sendo que 111 não podem ser implementados com os dados atuais disponíveis ou carecem de informação complementar para melhor serem quantificados. Dos restantes 76 indicadores, 37 foram implementados no protótipo e 39 não foram implementados, mas a informação disponível possibilita a sua implementação num projeto futuro.

4. MODELO DE DADOS E ARQUITETURA

4.1 CONSTRUÇÃO DO MODELO DIMENSIONAL

Tal como já referido anteriormente a falta de informação sobre os dados dos sistemas transacionais obrigou a que se utilizassem algumas tabelas do sistema ReAD para construir o modelo de dados. O modelo desenvolvido contempla duas tabelas de factos (Fact Tables) e vinte e nove tabelas de dimensões (Dim Tables). Optou-se por um modelo STAR SCHEMA por ser a mais indicada nesta solução.

A tabela **FACT operacoes** tem como base a tabela **psit_projectoop_hist** e contém uma granularidade ao nível dos dados mensais acumulados de cada projeto. A tabela **FACT series** tem como base a tabela **psit_series** e contém o total acumulado mensal para vários indicadores (não calculáveis através **FACT operacoes**).

Fact Tables

- **FACT operacoes**

Esta tabela contém uma chave por cada dimensão considerada no modelo e os campos relevantes para todos os indicadores pretendidos. Foram criadas algumas métricas para serem usadas na obtenção dos indicadores pretendidos. Para efeitos de protótipo mantiveram-se na tabela também as descrições relativas às dimensões, mas deverão ser eliminados na sua implementação (ver campo observações).

Campo	Nome do campo
Campo Base /Calculado	Indicação se é uma cópia simples do dado base da tabela (nome do campo) ou calculado
Expressão	Regra de filtragem ou expressão de cálculo
Tipo	Tipo de dados
Observações	Observações relevantes à compreensão e interpretação dos dados

Tabela 2: Descrição colunas de tabelas de factos

Estrutura da tabela **FACT operações** abaixo.

Tabela base: **psit_projectoop_hist**

Campo	Campo Base /Calculado	Expressão	Tipo	Observações
FK_ag			INT	eliminar: PO Finan.
FK_agrupamentos			INT	eliminar: Agrupamentos Setoriais
FK_aviso			INT	eliminar: aviso
FK_Beneficiario			INT	eliminar: NIF, Beneficiário
FK_cae			INT	eliminar: an_cae
FK_cgrupo			INT	eliminar:cod_grupo, Grupo
FK_dimensao			INT	eliminar: Dimensão
FK_divisao			INT	
FK_dominio			INT	

FK_dst_anl			INT	eliminar: DST / ANL
FK_efice			INT	eliminar: Classificação Indústria EFICE
FK_Eixo			INT	
FK_enei			INT	eliminar: ENEI
FK_fundo			INT	eliminar: fundo
FK_gai			INT	eliminar: Grandes Agrupamentos Industriais
FK_instrumento			INT	eliminar: instrumento
FK_itec			INT	eliminar: Instrumento
FK_medida			INT	eliminar: Medida
FK_nutii			INT	eliminar: NUTS II (trab)
FK_nutiii			INT	eliminar: NUTS III (trab)
FK_organismo			INT	eliminar: organismo
FK_OT			INT	
FK_PI			INT	eliminar: PI
FK_setor			INT	eliminar: Sector
FK_subclasse			INT	
FK_TI			INT	
FK_tic			INT	eliminar: Intensidade Tecnológica e de Conhecimento
FK_TO			INT	
FK_transacionavel			INT	eliminar: Transacionável
INV MIL	cnd_invest		DEC	
Apoiado?		if [dec_data] = null then null else [apoiado]	TXT	
Nº Proj Aprov.		if ([apoiado] = "Sim" and [anulado]="") then 1 else 0	INT	
Com TA?	contratado		TXT	
Com Pagamentos?		if [pagamentos] > 0 then 1 else 0	INT	
NUTS III (trab)		if Text.Contains([an_lst_nute_iii], ",") then "Multirregiões ou n.a./n.d." else if (Text.Contains([an_lst_nute_iii], "RA Açores") or Text.Contains([an_lst_nute_iii], "RA Madeira") or Text.Contains([an_lst_nute_iii], "Estrangeiro") or ([an_lst_nute_iii]= "")) or ([an_lst_nute_iii]= null) then "Multirregiões ou n.a./n.d." else [an_lst_nute_iii]	TXT	
NUTS II (trab)		if Text.Contains([an_lst_nute_ii], ",") then "Multirregiões ou n.a./n.d." else if (Text.Contains([an_lst_nute_ii], "RA Açores") or Text.Contains([an_lst_nute_ii], "RA Madeira") or Text.Contains([an_lst_nute_ii], "Estrangeiro") or ([an_lst_nute_ii]= "")) or ([an_lst_nute_ii]= null) then "Multirregiões ou n.a./n.d." else [an_lst_nute_ii]	TXT	
Municípios de Baixa Densidade?		"t_baixad", "tbc_parn" => "baixad", "codigo" => "an_lst_conc"	TXT	
FLAG Aprovados		Nº Proj Aprov.	INT	
FLAG Contratados		if ((([Com TA?]="1") and ([FLAG Aprovados]=1)) then 1 else 0	INT	
FLAG Pagamentos		if ([Com Pagamentos?]= 1) then 1 else 0	INT	
FLAG execução		if ([Exec]Fundo" > 0 and [FLAG Aprovados]=1) then 1 else 0	INT	
FLAG Encerrado		if [enc_dt_fim] <> null then 1 else 0	INT	
FLAG desp certificada		if [desp_t] <> 0 then 1 else 0	INT	
VAL Inv Cand		[#[Cand]Investimento]/1000	DEC	
VAL Inv Elegivel		[#[Dec]Elegivel]*[FLAG Aprovados]/1000	DEC	
VAL Aprov Incentivo		[#[Dec]Incentivo]*[FLAG Aprovados]/1000	DEC	
VAL Cont Incentivo		[#[Dec]Incentivo]*[FLAG Contratados]/1000	DEC	
VAL Pagamentos		[Pagamento Total]*[FLAG Pagamentos]/1000	DEC	
VAL Execução Fundo		[#[Exec]Fundo]*[FLAG execução]/1000	DEC	
VAL desp certificada		[desp_t]*[FLAG desp certificada]/1000	DEC	

INV_MILHÕES		[INV MIL]/1000	DEC	
Concelho Princ		1º elemento de [concelho(s)]	TXT	if null then "ND"
Concelho Sec		restante de [concelho(s)]	TXT	
anomes	anomes		INT	
dt_ref	dt_ref		DATE	
dt_act	dt_act		DATE	
Código de Operação	nup		TXT	
[Exec]Fundo	execucao		DEC	
NºProj	nproj		INT	
NIF	nif		INT	
Beneficiário	nome		TXT	
PI	pi		TXT	
Medida	medida_d		TXT	
Aviso	aviso		TXT	
PO Finan.	ag_d		TXT	
Organismo	oi_d		TXT	
[Cand]Investimento	cnd_invest		DEC	
Pagamento Total	pagamentos		DEC	
DST/ ANL	anulado		TXT	
ENEI	enei_d		TXT	
Dimensão	dimensao		TXT	if "" or null then "ND"
Sector		join an_cae(psit_projectoop_hist) com codigo(tbc_pam_cae) >="setor"	TXT	if null or "" then "ND ou NE"
Agrupamentos Setoriais		join an_cae(psit_projectoop_hist) com codigo(tbc_pam_cae) >="agrup_sectorial"	TXT	if null or "" then "ND ou NE"
Grandes Agrupamentos Industriais		join an_cae(psit_projectoop_hist) com codigo(tbc_pam_cae) >="agrup_industrial"	TXT	if null or "" then "ND ou NE"
Classificação Indústria EFICE		join an_cae(psit_projectoop_hist) com codigo(tbc_pam_cae) >="efice"	TXT	if null or "" then "ND ou NE"
Intensidade Tecnológica e de Conhecimento		join an_cae(psit_projectoop_hist) com codigo(tbc_pam_cae) >="itc"	TXT	if null or "" then "ND ou NE"
mes_sequencial		ano*12+mês	INT	
fundo	fundo		TXT	
[Dec]Elegível		if (([enc_data] <> "") and ([anulado]="")) then [enc_eleg] else [dec_eleg]	DEC	
[Dec]Incentivo		if (([enc_data] <> "") and ([anulado]="")) then [enc_incent] else [dec_incent]	DEC	
cnd_data	cnd_data		DATE	
dec_dt_p	dec_dt_p		DATE	
nt_dt_p	nt_dt_p		DATE	
dec_data	dec_data		DATE	
dt_contrato	dt_contrato		DATE	
dec_dt_inicio	dec_dt_inicio		DATE	
dec_dt_fim	dec_dt_fim		DATE	
enc_dt_fim	enc_dt_fim		DATE	
enc_dt_inicio	enc_dt_inicio		DATE	
dec_invest	dec_invest		DEC	
enc_invest	enc_invest		DEC	
enc_eleg	enc_eleg		DEC	
enc_incent	enc_incent		DEC	
dec_eleg_original	dec_elg		DEC	

dec_incent_original	dec_incent		DEC	
contratado	contratado		TXT	
sistema	sistema		TXT	
instrumento	instrumento		TXT	
Prazo Análise		[dec_dt_p]-[cnd_data]	INT	dias calendário
Prazo Decisão		[dec_data]-[cnd_data]	INT	dias calendário
Prazo Contratação		[dt_contrato]-[nt_dt_p]	INT	dias calendário
Atraso Início Investimento		[enc_dt_inicio]-[dec_dt_inicio]	INT	dias calendário
Atraso Conclusão		[enc_dt_fim]-[dec_dt_fim]	INT	dias calendário
Duração Candidatura		[dec_dt_fim]-[dec_dt_inicio]	INT	dias calendário
Duração Projecto Encerrado		[enc_dt_fim]-[enc_dt_inicio]	INT	dias calendário
Cod grupo		1- SI, 2 - Compete, 3 - ambos, 0-outros	TXT	
Grupo		1- SI, 2 - Compete, 3 - ambos, 0-outros	TXT	
Dotacao	Dotacao		DEC	
desp_t	desp_t		TXT	
ot	ot		TXT	
an_cae	an_cae		TXT	
Concelho		if (((Concelho Sec) = null) and ((Concelho Princ) <>"")) then [Concelho Princ] else "Multiconcelho ou ND"	TXT	
Transacionavel		join an_cae(psit_projectoop_hist) com codigo(tbc_pam_cae) >="transac"	TXT	
dimensao_d		join dimensao(psit_projectoop_hist) com cod_TipoPromotor(tbc_dimensao) >="dimensao_d"	TXT	if null then "Projetos Conjuntos"
tipo empresa		if [Dimensão] = 4 then "Grande Empresa" else "PME"	TXT	
INV aprovado	Calculated Column	'FACT Operacoes'[VAL Inv Elegível]*'FACT Operacoes'[FLAG Aprovados]	DEC	
INV contratado	Calculated Column	'FACT Operacoes'[VAL Inv Elegível]*'FACT Operacoes'[FLAG Contratados]	DEC	
INV pago	Calculated Column	'FACT Operacoes'[VAL Inv Elegível]*'FACT Operacoes'[FLAG Pagamentos]	DEC	
INV executado	Calculated Column	'FACT Operacoes'[VAL Inv Elegível]*'FACT Operacoes'[FLAG execução]	DEC	
INV certificado	Calculated Column	'FACT Operacoes'[VAL desp certificada]*'FACT Operacoes'[FLAG desp certificada]	DEC	
INV Encerrado	Calculated Column	'FACT Operacoes'[enc_invest]*'FACT Operacoes'[FLAG Encerrado]/1000	DEC	
Prazo_Decisão	Calculated Column	VAR DELTA=30 VAR A=CALCULATE(INT(COUNTROWS ('Datas')/DELTA), DATESBETWEEN ('Datas'[Dia], 'FACT Operacoes'[cnd_data], 'FACT Operacoes'[dec_data]), 'Datas'[Dia útil] , ALL ('FACT Operacoes')) RETURN IF(A>9,"300+",FORMAT((A+1)*DELTA,"000"))	TXT	número de projetos por cada período de 30 dias
Prazo_Contratação	Calculated Column	VAR DELTA=20 VAR A=CALCULATE(INT(COUNTROWS ('Datas')/DELTA), DATESBETWEEN ('Datas'[Dia], 'FACT Operacoes'[nt_dt_p], 'FACT Operacoes'[dt_contrato]), 'Datas'[Dia útil] , ALL ('FACT Operacoes')) RETURN IF(A>9,"200+",FORMAT((A+1)*DELTA,"000"))	TXT	número de projetos por cada período de 20 dias
Atraso_Início_Investim ento	Calculated Column	VAR DELTA=30 VAR A=INT('FACT Operacoes'[Atraso Início Investimento]/DELTA) RETURN IF(A>9,"300+",FORMAT((A+1)*DELTA,"000"))	TXT	número de projetos por cada período de 30 dias
Atraso_Conclusão	Calculated Column	VAR DELTA=30 VAR A=INT('FACT Operacoes'[Atraso Conclusão]/DELTA) RETURN IF(A>9,"300+",FORMAT((A+1)*DELTA,"000"))	TXT	número de projetos por cada período de 30 dias
Duração_Candidatura	Calculated Column	VAR DELTA=60 VAR A=INT('FACT Operacoes'[Duração Candidatura]/DELTA) RETURN IF(A>9,"600+",FORMAT((A+1)*DELTA,"000"))	TXT	número de projetos por cada período de 60 dias
Duração_Projecto_Enc errado	Calculated Column	VAR DELTA=60 VAR A=INT('FACT Operacoes'[Duração Projecto Encerrado]/DELTA) RETURN IF(A>9,"600+",FORMAT((A+1)*DELTA,"000"))	TXT	número de projetos por cada período de 60 dias
Prazo_Análise	Calculated Column	VAR DELTA=30 VAR A=CALCULATE(INT(COUNTROWS ('Datas')/DELTA), DATESBETWEEN ('Datas'[Dia], 'FACT Operacoes'[cnd_data], 'FACT Operacoes'[dec_dt_p]), 'Datas'[Dia útil] , ALL ('FACT Operacoes')) RETURN IF(A>9,"300+",FORMAT((A+1)*DELTA,"000"))	INT	dias úteis

Tabela 3: FACT Operações

- **FACT series**

Esta tabela não tem ligação às dimensões por não haver informação que as ligue. Contém apenas alguns indicadores por período temporal (mês)

Tabela base: psit_series

Campo	Campo Base / Calculada	Expressão	Tipo	Observações
anomes	anomes		INT	
ano	anomes		INT	
mes	mês		INT	
mes_sequencial		[ano]*12+[mes]	INT	
mes_sequencial_anterior		mes_sequencial -1	INT	
MAnt Inv Cand		Inv Cand do mês anterior	DEC	
MAnt Inc Aprovado		Inc Aprovado do mês anterior	DEC	
MAnt Inc Contratado		Inc Contratado do mês anterior	DEC	
MAnt Inc Pago		Inc Pago do mês anterior	DEC	
MAnt Inc Executado		Inc Executado do mês anterior	DEC	
Inv Cand Mês		[Inv Cand]-[MAnt Inv Cand]	DEC	
Inc Aprovado Mês		[Inc Aprovado] -[MAnt Inc Aprovado]	DEC	
Inc Contratado Mês		[Inc Contratado]-[MAnt Inc Contratado]	DEC	
Inc Pago Mês		[Inc Pago]-[MAnt Inc Pago]	DEC	
Inc Executado Mês		[Inc Executado]-[MAnt Inc Executado]	DEC	
anomesext	anomes		TXT	
Inv Cand	inv_cand		DEC	
Inc Aprovado	inc_aprovado		DEC	
Inc Contratado	inc_contratado		DEC	
Inc Pago	inc_pago		DEC	
Inc Executado	inc_executado		DEC	
Adtos cert.	pag_c		DEC	
Adtos não cert.	pag_nc		DEC	
Reembolso	pag_r		DEC	
grupo	grupo		DEC	
Dotacao		join grupo(psit_series) com Grupo(DIM Grupos) >="dotacao"	DEC	

Tabela 4: FACT series

Dimension Tables

Um dos principais objetivos da gestão do Compete2020 é flexibilidade e variedade de análises possíveis de efetuar com os dados existentes. Deste modo foram definidas 29 dimensões que permitirão analisar de modo detalhado múltiplas visões dos dados.

Abaixo discriminam-se todas as tabelas, bem como as chaves de relação com a tabela **FACT Operacoes**.

DIM Beneficiario:

Column Name	Tipo	Relação
BK_Beneficiario	INT	FACT Operacoes.FK_Beneficiario
NIF	INT	
Beneficiário	TEXT	

DIM OT:

Column Name	Tipo	Relação
BK_OT	INT	FACT Operacoes.FK_OT
OT		

DIM PI:

Column Name	Tipo	Relação
BK_PI	INT	FACT Operacoes.FK_PI
PI	TEXT	
filtro	TEXT	

DIM Eixo POCI:

Column Name	Tipo	Relação
BK_Eixo	INT	FACT Operacoes.FK_Eixo
Eixo POCI		

DIM TI:

Column Name	Tipo	Relação
BK_TI	INT	FACT Operacoes.FK_TI
TI	TEXT	

DIM TO:

Column Name	Tipo	Relação
BK_TO	INT	FACT Operacoes.FK_TO
TO	TEXT	

DIM Instrumento:

Column Name	Tipo	Relação
BK_instrumento	INT	FACT Operacoes.FK_instrumento
instrumento	TEXT	

DIM Medida:

Column Name	Tipo	Relação
BK_medida	INT	FACT Operacoes.FK_medida
Medida	TEXT	

DIM Aviso:

Column Name	Tipo	Relação
BK_aviso	INT	FACT Operacoes.FK_aviso
aviso	TEXT	

DIM Fundo:

Column Name	Tipo	Relação
BK_fundo	INT	FACT Operacoes.FK_fundo
fundo	TEXT	

DIM Organismo:

Column Name	Tipo	Relação
BK_organismo	INT	FACT Operacoes.FK_organismo
Organismo	TEXT	

DIM DST ANL:

Column Name	Tipo	Relação
BK_dst_anl	INT	FACT Operacoes.FK_dst_anl
DST/ ANL	TEXT	

DIM ENEI:

Column Name	Tipo	Relação
BK_enei	INT	FACT Operacoes.FK_enei
ENEI	TEXT	

DIM Subclasse:

Column Name	Tipo	Relação
BK_subclasse	INT	FACT Operacoes.FK_subclasse
Subclasse	TEXT	

DIM Divisao:

Column Name	Tipo	Relação
BK_divisao	INT	FACT Operacoes.FK_divisao
Divisão	TEXT	

DIM Sector:

Column Name	Tipo	Relação
BK_setor	INT	FACT Operacoes.FK_setor
Sector	TEXT	

DIM Agrupamentos Setoriais:

Column Name	Tipo	Relação
BK_agrupamentos	INT	FACT Operacoes.FK_agrupamentos
Agrupamentos Setoriais	TEXT	

DIM Grandes AI:

Column Name	Tipo	Relação
BK_gai	INT	FACT Operacoes.FK_gai
Grandes Agrupamentos Industriais	TEXT	

DIM EFICE:

Column Name	Tipo	Relação
BK_efice	INT	FACT Operacoes.FK_efice
Classificação Indústria EFICE	TEXT	

DIM Transacionavel:

Column Name	Tipo	Relação
BK_transacionavel	INT	FACT Operacoes.FK_transacionavel
Transacionável	TEXT	

DIM Intensidade TeC:

Column Name	Tipo	Relação
BK_itec	INT	FACT Operacoes.FK_itec
Intensidade Tecnológica e de Conhecimento	TEXT	

DIM TIC:

Column Name	Tipo	Relação
BK_tic	INT	FACT Operacoes.FK_tic
TIC	TEXT	

DIM Domínio Intervencao:

Column Name	Tipo	Relação
BK_dominio	INT	FACT Operacoes.FK_dominio
Domínio Intervenção	TEXT	

DIMREAD tbc_autoridadegestao:

Column Name	Tipo	Relação
BK_ag	INT	FACT Operacoes.FK_ag
ag	TEXT	
AutoridadeGestao	TEXT	

DIM Grupos:

Column Name	Tipo	Relação
BK_cgrupo	INT	FACT Operacoes.FK_cgrupo
Cod_grupo	INT	
Grupo	TEXT	
Dotacao	INT	

DIMREAD tbc_pam_cae:

Column Name	Tipo	Relação
BK_cae	INT	FACT Operacoes.FK_cae
codigo	TEXT	
subclasse	TEXT	
divisao	TEXT	
setor	TEXT	
agrup_sectorial	TEXT	
agrup_industrial	TEXT	
efice	TEXT	
transac	TEXT	
itc	TEXT	
tic	TEXT	

DIMREAD tbc_dimensao:

Column Name	Tipo	Relação
BK_dimensao	INT	FACT Operacoes.FK_dimensao
cod_TipoPromotor	INT	
TipoPromotor	TEXT	

DIM NUTS II:

Column Name	Tipo	Relação
BK_nutii	INT	FACT Operacoes.FK_nutii
NUTS II (trab)		

DIM NUTS III:

Column Name	Tipo	Relação
BK_nutiii	INT	FACT Operacoes.FK_nutiii
NUTS III (trab)		

Tabela 5: Tabelas de Dimensões

Diagrama de relações

Da implementação das Fact tables e Dim tables no data warehouse e respetivas relações entre tabelas resultará o diagrama de relações semelhante à figura seguinte: (anexo III)

Existem ligeiras diferenças entre o diagrama abaixo e o modelo atrás apresentado pelo facto do diagrama ter sido criado através da plataforma PowerBI utilizada para desenvolver o protótipo.

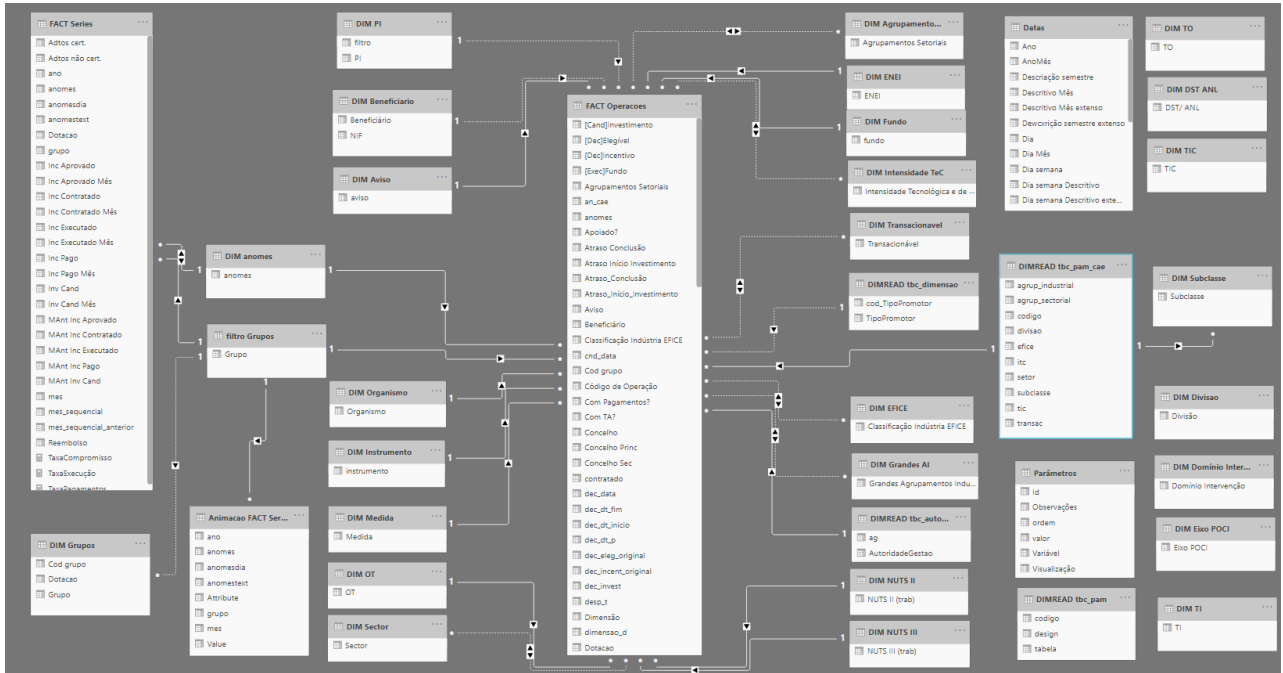


Figura 15: Diagrama de Relações (ver detalhe no anexo III)

4.2 INDICADORES A DESENVOLVER

Para dar respostas à necessidade de obtenção de alguns indicadores foram criadas as medidas discriminadas na tabela seguinte:

Medida	Tabela	Descrição	Expressão
%Dotação2020 Execução	FACT Operacoes	Percentagem de execução do fundo da dotação prevista	$\text{divide}(\text{sum}(\text{FACT Operacoes}[\text{VAL Execução Fundo}]), \text{divide}(\text{sum}(\text{FACT Operacoes}[\text{Dotacao}]), \text{count}(\text{FACT Operacoes}[\text{Dotacao}])))$
%Dotação2020 Pagamentos	FACT Operacoes	Percentagem de pagamentos da dotação prevista	$\text{divide}(\text{sum}(\text{FACT Operacoes}[\text{VAL Pagamentos}]), \text{divide}(\text{sum}(\text{FACT Operacoes}[\text{Dotacao}]), \text{count}(\text{FACT Operacoes}[\text{Dotacao}])))$
Tx Exec IncAprovado	FACT Operacoes	Taxa de inv elegível final vs elegível original	$\text{sum}(\text{FACT Operacoes}[\text{enc_incent}]) / \text{sum}(\text{FACT Operacoes}[\text{dec_incent_original}])$
Tx Exec InvElegível	FACT Operacoes	Taxa de incentivo aprovado final vs original aprovado	$\text{sum}(\text{FACT Operacoes}[\text{enc_eleg}]) / \text{sum}(\text{FACT Operacoes}[\text{dec_eleg_original}])$
Tx Exec InvAprovado	FACT Operacoes	Taxa de investimento final vs investimento original	$\text{sum}(\text{FACT Operacoes}[\text{enc_invest}]) / \text{sum}(\text{FACT Operacoes}[\text{dec_invest}])$
%Dotação2020 Incentivo	FACT Operacoes	Percentagem de incentivo aprovado da dotação prevista	$\text{divide}(\text{sum}(\text{FACT Operacoes}[\text{VAL Aprov Incentivo}]), \text{divide}(\text{sum}(\text{FACT Operacoes}[\text{Dotacao}]), \text{count}(\text{FACT Operacoes}[\text{Dotacao}])))$
%Dotação2020 DespCert	FACT Operacoes	Percentagem da despesa certificada da dotação prevista	$\text{divide}(\text{sum}(\text{FACT Operacoes}[\text{VAL desp certificada}]), \text{divide}(\text{sum}(\text{FACT Operacoes}[\text{Dotacao}]), \text{count}(\text{FACT Operacoes}[\text{Dotacao}])))$
%IncCont	FACT Operacoes	Taca de execução do fundo	$\text{sum}(\text{FACT Operacoes}[\text{VAL Execução Fundo}]) / \text{sum}(\text{FACT Operacoes}[\text{VAL Cont Incentivo}])$
TxCont	FACT Operacoes	Taxa de incentivo contratado	$\text{sum}(\text{FACT Operacoes}[\text{VAL Cont Incentivo}]) / \text{sum}(\text{FACT Operacoes}[\text{VAL Aprov Incentivo}])$

TxPag	FACT Operacoes	Taxa de incentivo pago	sum('FACT Operacoes'[VAL Pagamentos])/sum('FACT Operacoes'[VAL Cont Incentivo])
Tx Proj Encerrados	FACT Operacoes	Taxa de projetos encerrados	sum('FACT Operacoes'[FLAG Encerrado])/calculate(sum('FACT Operacoes'[FLAG Contratados]),ALLEXCEPT('FACT Operacoes','DIM anomes'[anomes],'filtro Grupos'[Grupo]))
TaxaCompromisso	FACT Series	Taxa de compromisso	sum('FACT Series'[Inc Aprovado])/sum('FACT Series'[Dotacao])
TaxaPagamentos	FACT Series	Taxa de Pagamentos	sum('FACT Series'[Inc Pago])/sum('FACT Series'[Dotacao])
TaxaExecução	FACT Series	Taxa de Execução	sum('FACT Series'[Inc Executado])/sum('FACT Series'[Dotacao])

Tabela 6: Indicadores

4.3 ARQUITETURA PROPOSTA DE BI

Ao longo deste relatório já foram referidas as várias condicionantes deste projeto que não permitiram o desenho de um modelo de dados tendo como ponto de partida as aplicações operacionais. A alternativa encontrada foi a de usar algumas tabelas da base de dados do ReAD.

A base de dados do ReAD é já ela própria resultante de um processo de ETL desenvolvido e controlado pelo departamento de IT / parceiro externo. Com esta decisão tornou-se possível garantir a uniformidade e coerência dos dados produzidos pela futura solução de BI e com isso comparar os resultados obtidos com os dos relatórios mensais de ponto de situação.

Esta opção tem, no entanto, algumas limitações que se discriminam a seguir:

- Impossibilidade de verificar e garantir que os dados usados são os indicados para o resultado pretendido;
- Impossibilidade de garantir a coerência e integridade dos dados;
- Impossibilidade de garantir a coerência temporal da informação;
- Falta de controlo (pela solução BI) da periodicidade e frequência dos momentos de extração de dados;
- Solução de BI fica dependente e poderá ter problemas futuros decorrentes de alterações não previstas e/ou não informadas na estrutura das tabelas ou base de dados;
- Eventuais necessidades futuras de dados adicionais ficarão sempre dependentes do conhecimento do departamento de IT / parceiro e da disponibilidade / custo para as desenvolver;

Deste modo, embora não tivesse sido possível neste projeto, a solução de BI futura deverá implementar um sistema da ETL tendo como fonte os dados base do SGO/outros em vez da base de dados do ReAD.

Numa arquitetura teórica e ideal, a solução preconizada deverá seguir a solução da figura seguinte:

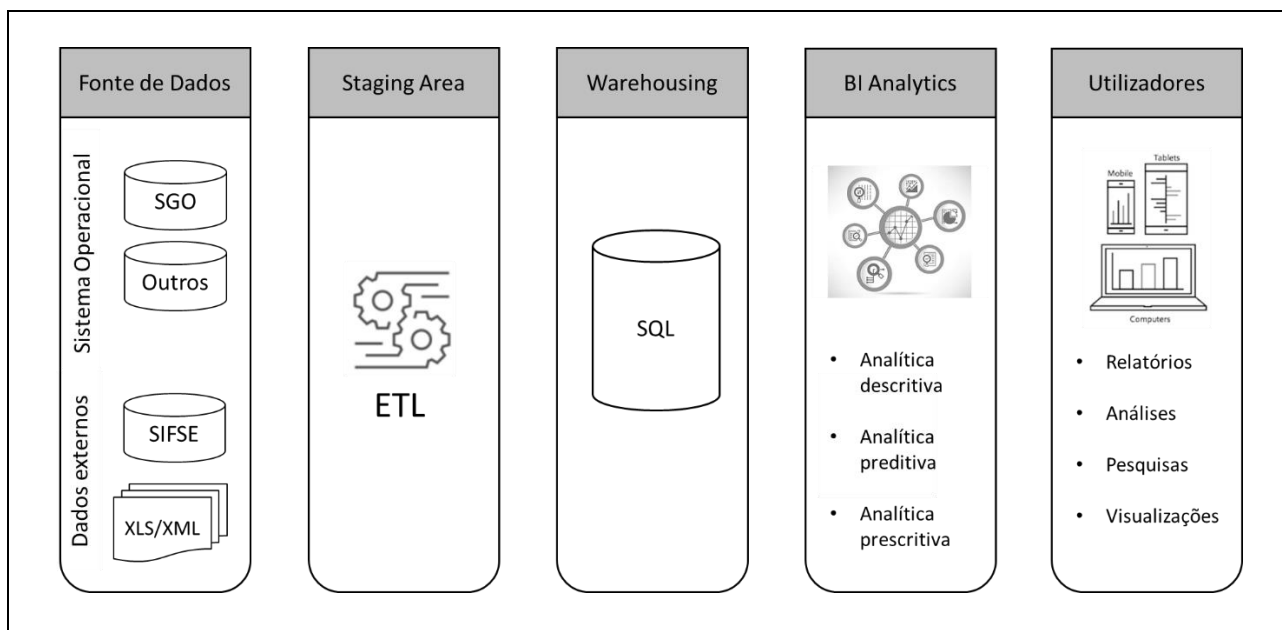


Figura 16: Arquitetura solução BI

Fontes de dados

As fontes de dados que alimentam o BI deverão ser todas as estruturas de dados (base de dados, ficheiros xml, excel, pdf ou outros) dos sistemas operacionais utilizados no Compete2020, nomeadamente (SGO, SIFSE, faturas, etc.), bem como dos sistemas externos que fornecem dados relevantes para o acompanhamento e análise efetuados.

Staging Area

Consiste num repositório de dados (base de dados) de trabalho para onde são descarregados todos os dados extraídos dos sistemas operacionais e efetuadas todas as tarefas de tratamento de dados, nomeadamente limpeza, verificação de integridade, cálculo de métricas, etc. e construção das tabelas de fatos e dimensões. Esta área é também responsável pelo controlo da informação histórica e carregamento dos novos dados (ou atualizados) para o data warehouse. A estrutura de dados a utilizar é a do modelo proposto.

Warehousing

Consiste no repositório de dados final para utilização da aplicação frontend de BI. A estrutura de dados a utilizar é a do modelo proposto que consiste num repositório integrado de dados, orientado ao negócio do Compete2020, não volátil e variando ao longo do tempo.

BI Analytics

Aplicação disponibilizada ao utilizador final que permitirá efetuar todas as análises/visualizações dos indicadores necessários à gestão bem como extrair a informação para análises específicas, podendo

integrar três tipos de analítica: descritiva (o que aconteceu); preditiva (o que pode acontecer); prescritiva (que ações tomar para que o que acontecer corresponda aos nossos objetivos)

Utilizadores

Qualquer colaborador da organização que utilize a ferramenta de BI para análise dos dados, com perfis de acesso devidamente segregados garantindo segurança e privacidade.

Staging Area

Para a staging área sugere-se a criação de uma base de dados usando o MySQL, uma vez que é o SGBDR usado nas restantes aplicações.

Processo ETL

O processo de ETL preconizado no projeto Compete2020 contempla os 3 passos fundamentais para uma solução de BI: Extração, Tratamento e Carregamento (Loading)

Os dados deverão ser extraídos dos sistemas transacionais utilizados no Compete2020 para a gestão operacional dos apoios concedidos e respetivo acompanhamento. A periodicidade da extração poderá ser diária, semanal ou mensal, em função do detalhe dos indicadores pretendidos e da dinâmica de monitorização. A ferramenta indicada para este procedimento poderá variar em função da solução global de BI escolhida, mas as funções serão semelhantes.

Os dados deverão ser extraídos, se possível, em períodos offline para garantir a coerência e integridade da informação e evitar sobrecarregar os sistemas. Da informação recebida conclui-se que a maioria da informação necessária estará na base de dados do SGO e SIFSE, em ficheiros XML e excel complementares, bem como na base de dados do sistema de faturas. Poderá ainda haver dados relevantes disponíveis em sistemas externos acessíveis via webservices.

Em função do volume e modo como os dados são registados/armazenados deverá ser avaliado se a extração deverá ser global ou incremental. Nesta fase não existem dados suficientes para identificar qual a melhor opção.

No processo de tratamento dos dados deverão ser asseguradas um conjunto de tarefas de modo a produzir informação corretamente estruturada para o data warehouse, nomeadamente:

- Limpeza e tratamento de campos vazios/nulos;
- Limpeza, tratamento e uniformização de dados semelhantes (maiúsculas, minúsculas, abreviaturas, expressões resumidas, etc.);

- Validação e atribuição de tipos de dados de acordo com os indicadores pretendidos;
- Construção das tabelas relativas às dimensões pretendidas;
- Construção das tabelas de fatos, com a granularidade adequada ao detalhe da informação pretendida;
- Cálculo de métricas necessárias à obtenção dos indicadores previstos;

O processo de carregamento de dados para o data warehouse deverá ser efetuado com a periodicidade e frequência a definir. O detalhe do ETL construído para a elaboração do protótipo está descrito no anexo IV.

Data Warehouse

Para o Data Warehouse sugere-se a criação de uma base de dados usando o MySQL uma vez que é o SGBDR usado nas restantes aplicações. A estrutura da base de dados deverá corresponder ao modelo atrás definido.

Slowly Changing Dimensions

No levantamento dos requisitos para a implementação do BI não foi identificada nenhuma necessidade de manter algum tipo de histórico relativamente a alguma da informação, pelo que é de optar pela atualização de dados já existentes, quer a nível das tabelas de dimensões quer das tabelas de factos. Preconiza-se que seja utilizada uma solução tipo 1 ou seja substituição dos dados atualizados.

4.4 APRESENTAÇÃO DA PROVA DE CONCEITO

Para apresentação da prova de conceito optou-se pela ferramenta PowerBI. Para além de reconhecida internacionalmente como líder de mercado, foi escolhida porque possibilita apresentar resultados rápidos e sem necessidade de uma preparação exaustiva do data warehouse. Deste modo foi construído um protótipo focado em três grandes objetivos:

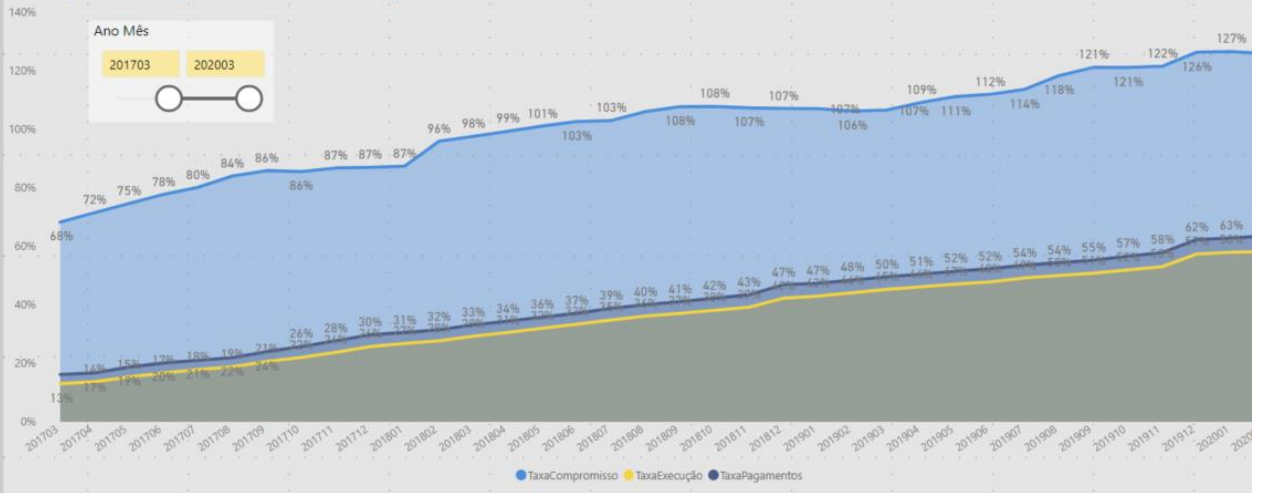
- Responder a todos os indicadores disponíveis atualmente no Relatório Mensal, com validação da coincidência de resultados;
- Demonstração de potencial para apresentação de novos indicadores;
- Demonstração de funcionalidades de análise e exploração dos dados: filtragem, agrupamento, pesquisa, visualização, exportação de dados, etc., com impacto significativo na produtividade dos utilizadores.

Para desenvolvimento do protótipo fez-se a ligação do PowerBI diretamente à base de dados do ReAD para extração dos dados. Utilizando as capacidades nativas do PowerBI efetuou-se todo o tratamento e transformação dos dados bem como a construção das tabelas de factos e dimensões. Embora o protótipo seja totalmente funcional recomenda-se, se este for utilizado num ambiente produtivo, que as tabelas de factos e dimensões sejam implementadas no data warehouse, num servidor de base de dados, devido ao volume de dados e potencial de crescimento. O powerBI deverá ser utilizado principalmente para desenvolvimento das visualizações.

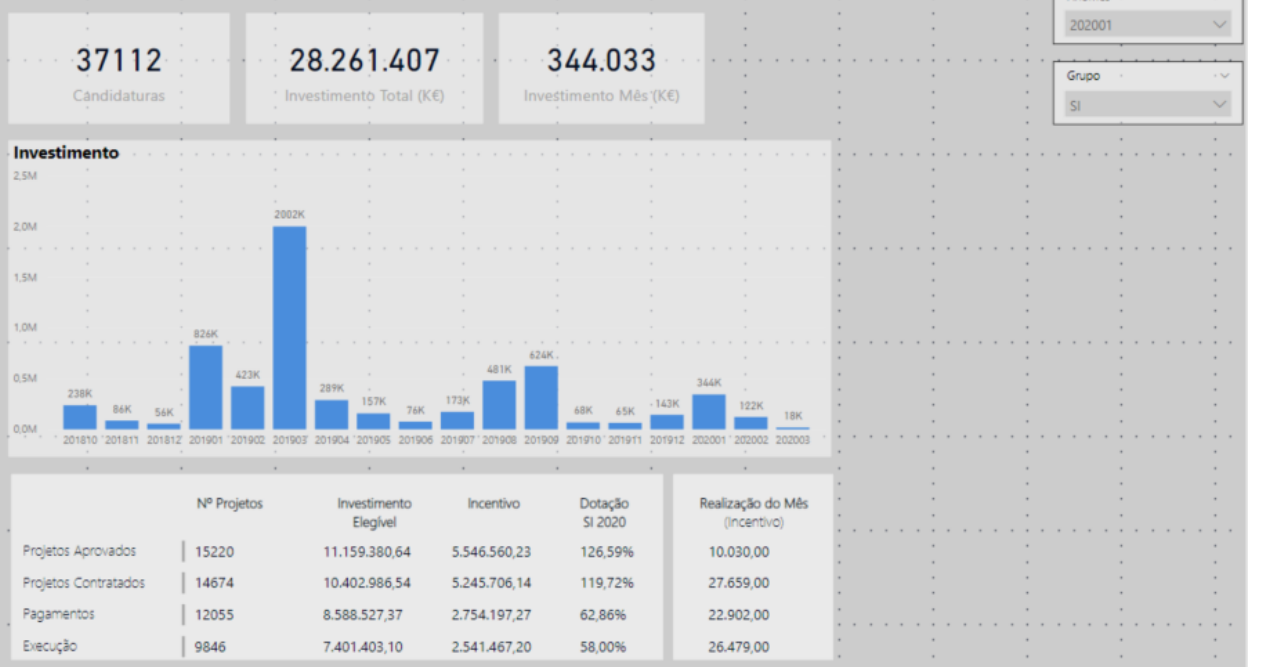
Num primeiro passo desenvolveram-se todas as visualizações que procuravam replicar a informação do relatório mensal. Confirmados os resultados avançou-se numa segunda fase para a construção de um conjunto de visualizações que permitissem dar resposta a muitos dos indicadores identificados na fase de levantamento. Tendo em conta as limitações dos dados disponíveis nesta fase implementou-se uma parte significativa dos identificados inicialmente.

As visualizações seguintes foram executadas à semelhança dos relatórios mensais existentes, replicam as apresentações de valores e gráficos, sem funcionalidades adicionais implementadas.

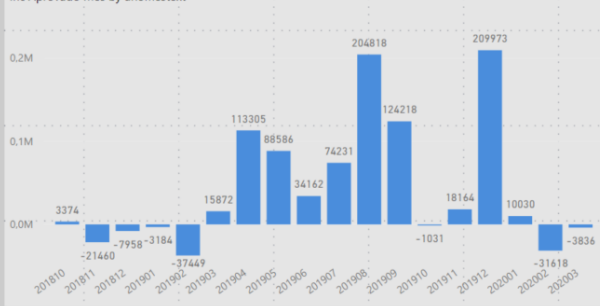
Evolução do Compromisso face à Dotação Indicativa dos Sistemas de Incentivos



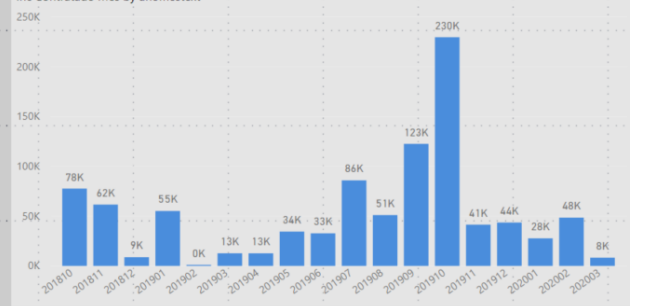
1. Realização Mensal



Inc Aprovado Mês by anomestext



Inc Contratado Mês by anomestext



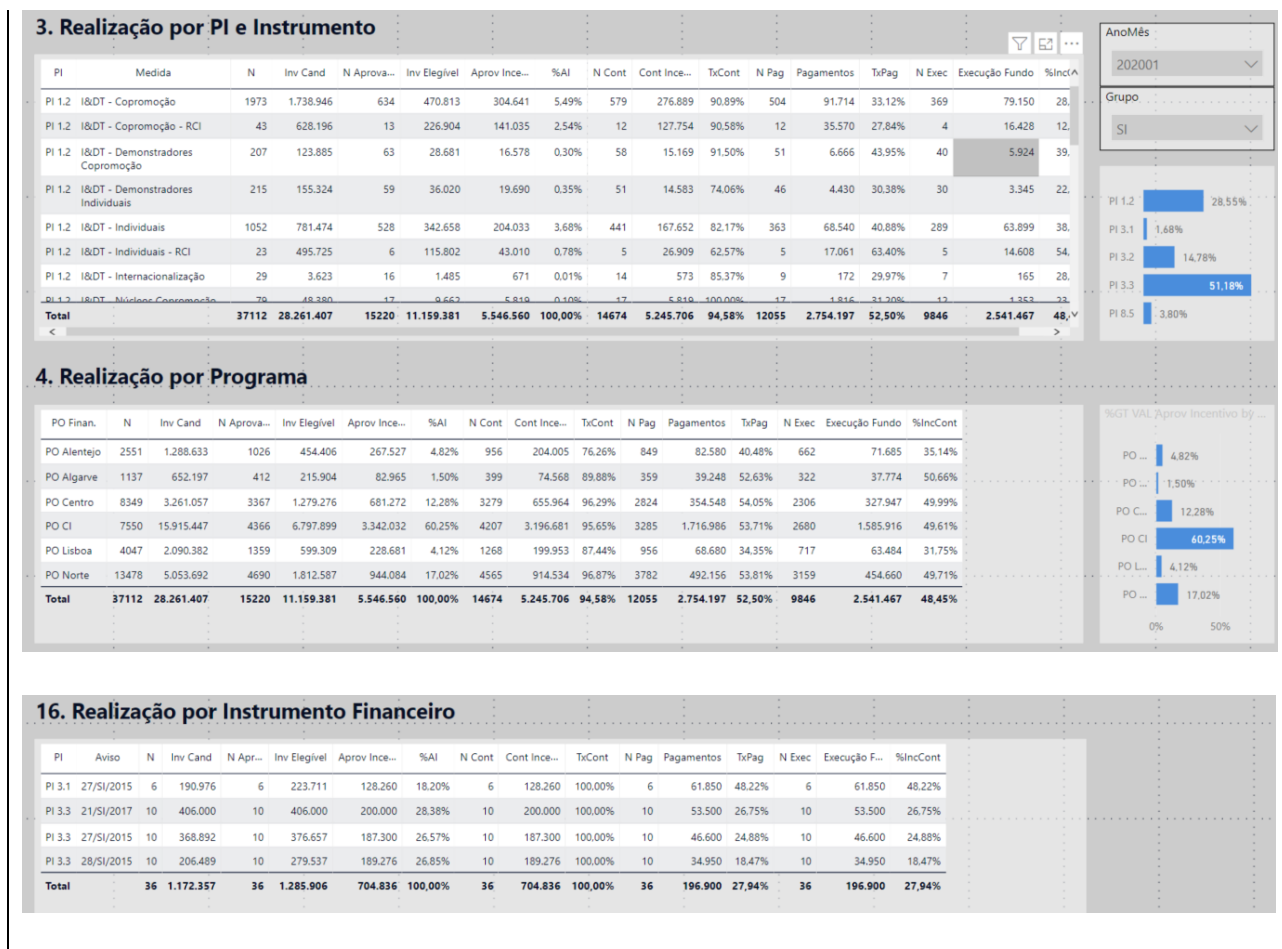


Figura 17: Dashboards Relatório mensal

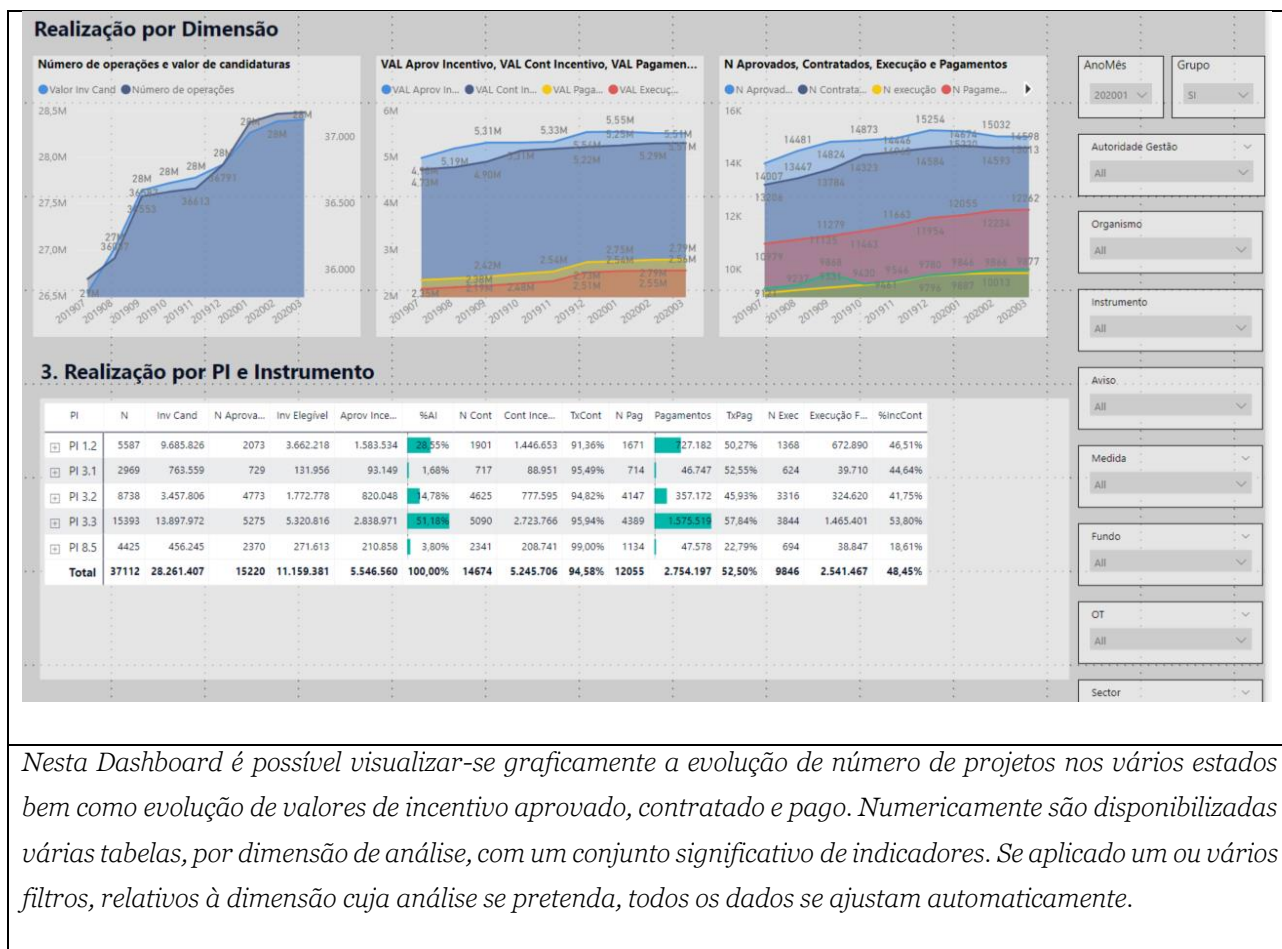
A seguir apresentam-se um conjunto de visualizações, algumas variantes das anteriores, melhorando a capacidade de visualização e fundamentalmente adicionando capacidades de pesquisa e filtragem de informação de modo a potenciar as visualizações segundo as dimensões identificadas pelo Compete2020.



Nesta Dashboard é possível obter informação sobre número de projetos candidatados (à data do mês selecionado e grupo de projetos) bem como número de projetos aprovados, contratados, com pagamentos, com execução e com despesa certificada. Além dos número e projetos estão disponíveis os valores de Investimento elegível, do Incentivo aprovado e a percentagem da dotação já realizada. Gráficamente mostra-se a evolução dos vários tipos de investimento. Dois gráficos mostram ainda o número de projetos encerrados bem como os anulados ou que desistiram.

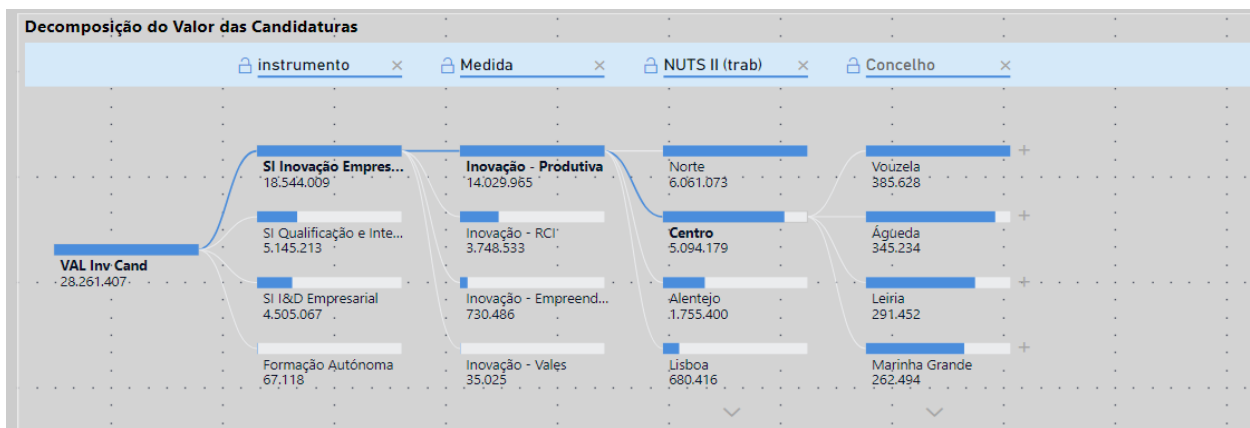
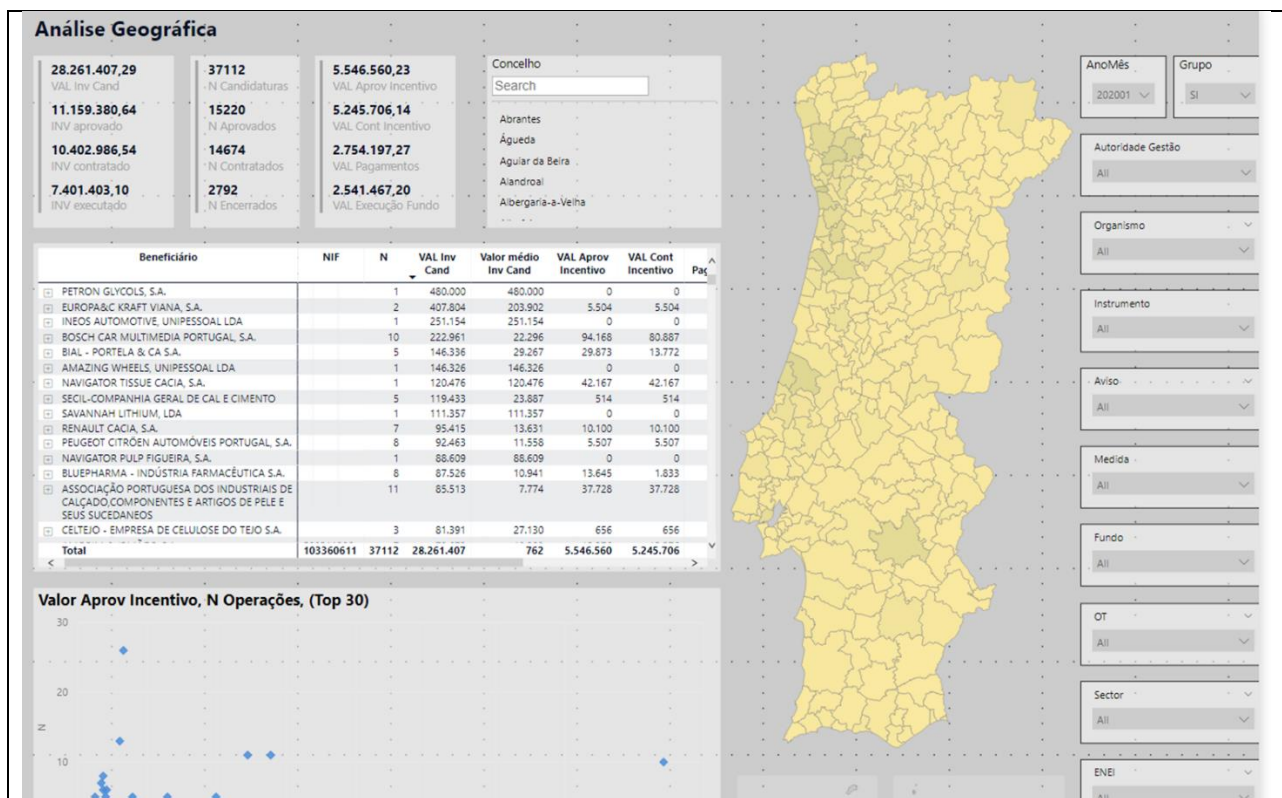
Se aplicado um ou vários filtros, relativos à dimensão cuja análise se pretenda, todos os dados se ajustam automaticamente.

Figura 18: Dashboard visão global



Nesta Dashboard é possível visualizar-se graficamente a evolução de número de projetos nos vários estados bem como evolução de valores de incentivo aprovado, contratado e pago. Numericamente são disponibilizadas várias tabelas, por dimensão de análise, com um conjunto significativo de indicadores. Se aplicado um ou vários filtros, relativos à dimensão cuja análise se pretenda, todos os dados se ajustam automaticamente.

Figura 19: Dashboard realização por dimensão



Nesta Dashboard é possível visualizar-se geograficamente a distribuição do número de projetos nos vários estados bem como os vários valores de incentivo.

Na tabela é possível consultar informação variada dos promotores, relativamente à região selecionada, tal como número de projetos, valor candidatado total e médio por projeto, etc.

Graficamente é possível avaliar o Top30 dos promotores relativamente a incentivo aprovado e número de projetos. Numa tabela está também disponível informação detalhada de cada projeto bem como analisar a decomposição dos valores de investimento segundo várias dimensões.

Se aplicados filtros, relativos à dimensão cuja análise se pretenda, todos os dados se ajustam automaticamente.

Figura 20: Dashboard Análise Geográfica

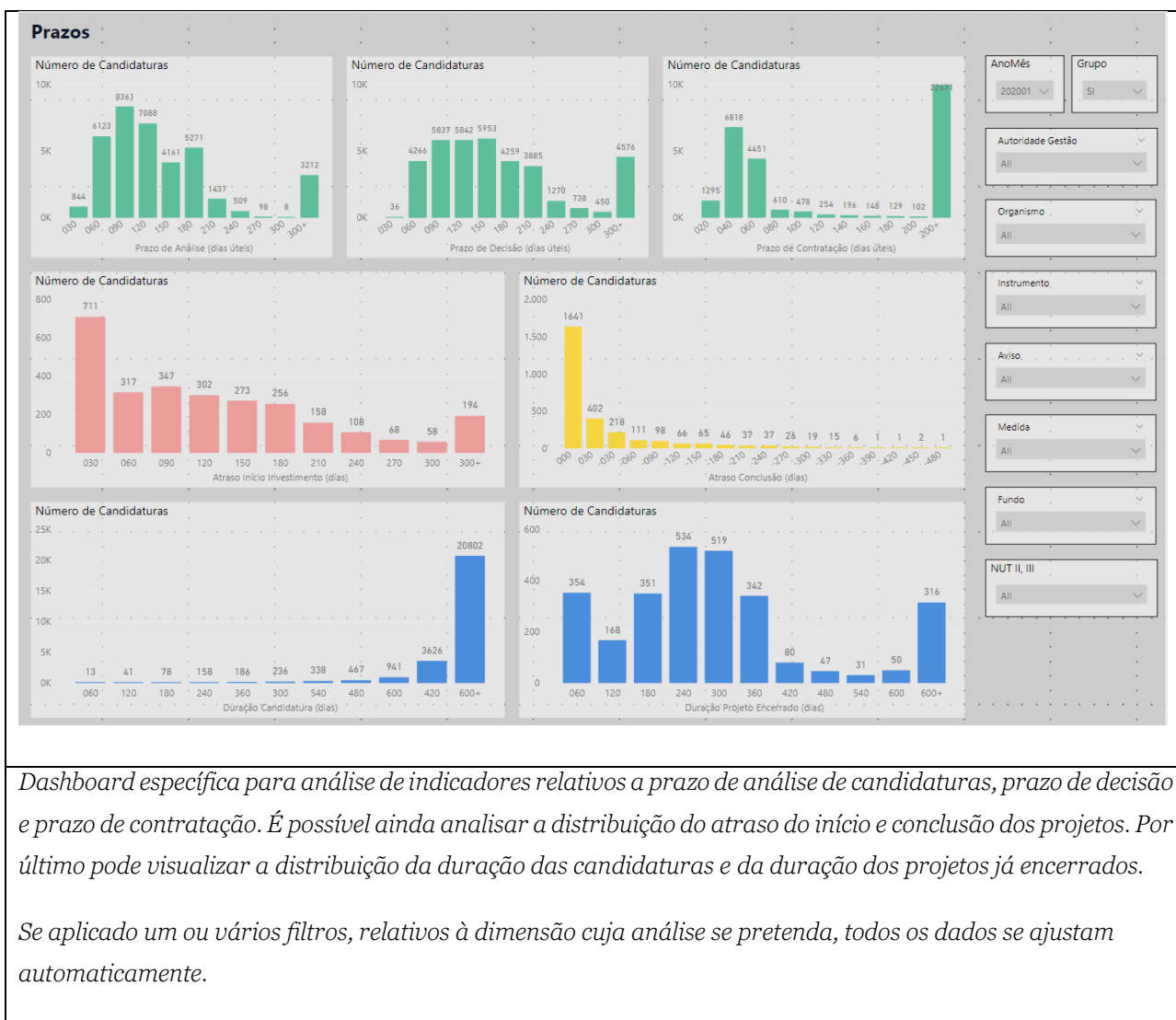
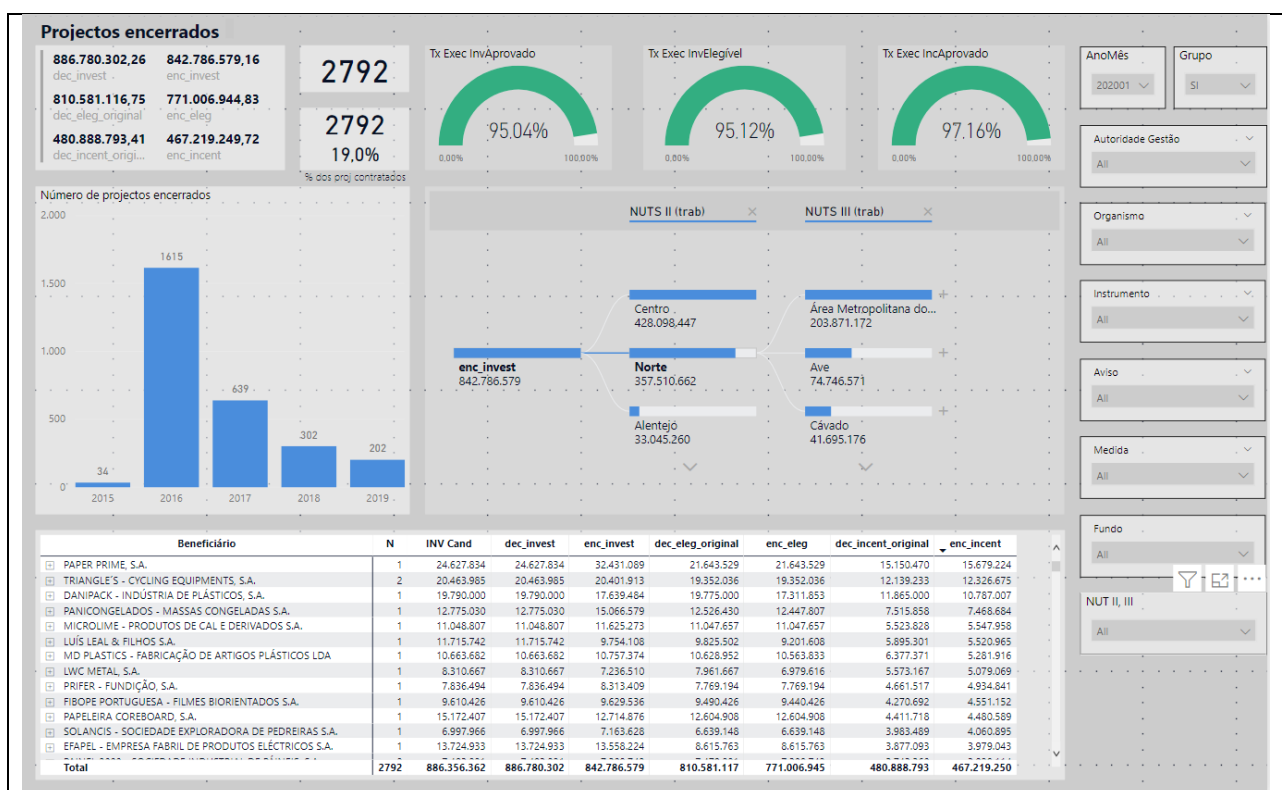


Figura 21: Dashboard Prazos

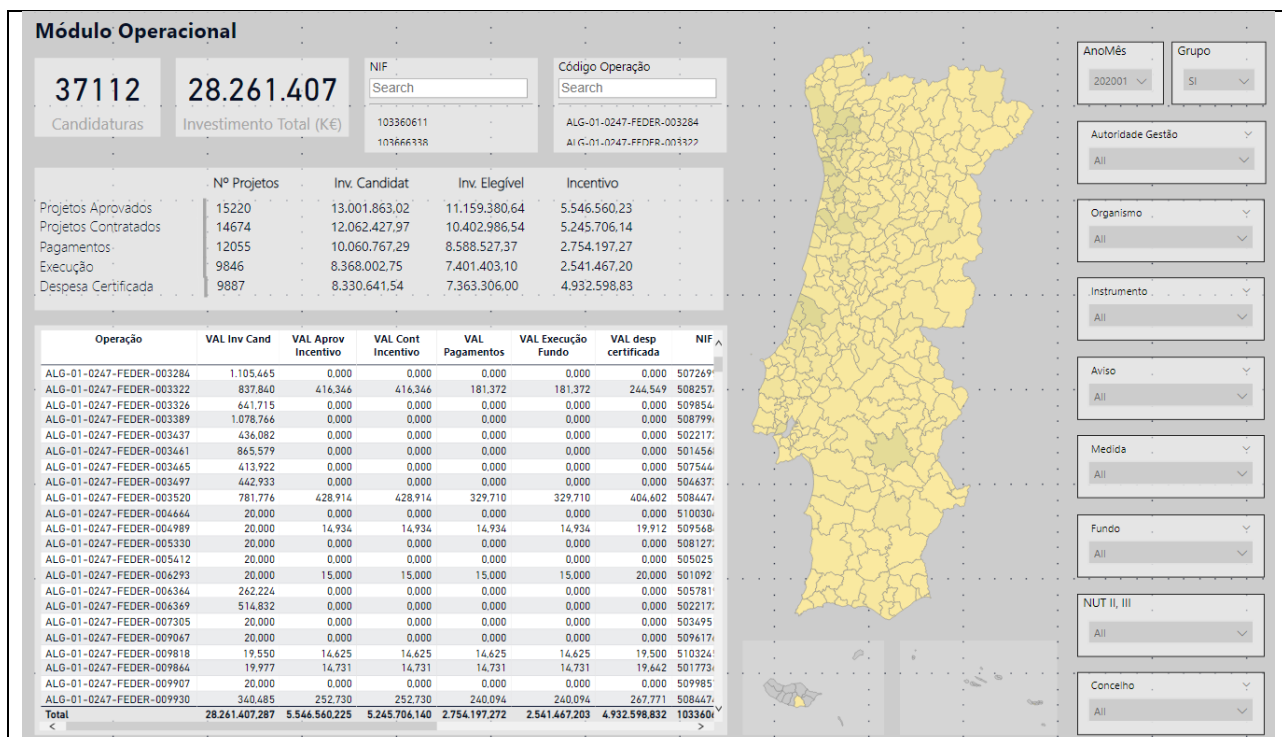


Dashboard específica para análise de indicadores relativos a projetos já encerrados, nomeadamente número de projetos, taxas de execução, distribuição do investimento, etc.

Complementarmente à informação consolidada é disponibilizado detalhe de informação por Beneficiário.

Se aplicado um ou vários filtros, relativos à dimensão cuja análise se pretenda, todos os dados se ajustam automaticamente.

Figura 22: Dashboard Projectos encerrados



Dashboard dirigida a utilizadores mais operacionais permitindo efetuar pesquisas por região, beneficiário ou até código de operação além da aplicação de filtros das várias dimensões.

São apresentados vários indicadores relativos a número de projetos e valores de investimento/Incentivo.

Complementarmente à informação consolidada é disponibilizado detalhe de informação por operação.

Se aplicado um ou vários filtros, relativos à dimensão cuja análise se pretenda, todos os dados se ajustam automaticamente.

Figura 23: Dashboard módulo operacional

Os exemplos atrás apresentados não são a totalidade dos implementados no protótipo, pelo que é importante a sua visualização e teste. A maioria das visualizações desenvolvidas permitem a exportação dos dados para um ficheiro excel para poderem ser efetuadas análises específicas caso seja necessário.

4.5 POTENCIAL DE ANÁLISE PREDITIVA

A existência de um histórico de dados de volume relevante pode permitir ao Compete2020 fazer estudos de análise preditiva e com isso tomar decisões sobre as estratégias mais adequadas na gestão das candidaturas e seu acompanhamento. Não tendo sido feito nenhum trabalho detalhado, devido às dificuldades no acesso à informação base durante o projeto, podem enumerar-se algumas análises que podem ser feitas e resultados possíveis de ser obtidos.

- Analisar e prever taxas de afetação dos fundos, de pagamentos, de execução, etc. em função dos projetos candidatados e sua tipologia;
- Analisar e prever evolução do número de projetos aprovados em função das suas características;
- Analisar e prever prazos e tempos de resposta em função do número, tipo, valor e características das candidaturas;
- Analisar e prever, em cada momento, a necessidade de aumentar ou reduzir o 'ritmo' de afetação dos apoios.

Se aos dados existentes se juntarem dados dos sistemas operacionais poderá ser possível:

- Avaliar pontos de estrangulamento nas várias fases do projeto e estimar impacto futuro;
- Analisar principais fatores condicionantes da aprovação de projetos e com isso estimar número de projeto/volume de investimento candidatado ou volume de incentivo aprovado;
- Analisar e prever cenários.

Com dados externos dos beneficiários ou mercado poderá ser possível:

- Avaliar o impacto dos apoios concedidos e prever impactos futuros numa determinada região, tipo de atividade, dimensão de empresa etc.;
- Analisar e antecipar potenciais empresas/setores com incorreta utilização de fundos.

Este tipo de análises carece de um estudo aprofundado do volume e qualidade dos dados disponíveis. Não se resume à utilização de uma ferramenta de BI, mas deverá ser implementado com um projeto com recursos adequados de data science.

4.6 IDENTIFICAÇÃO DOS REQUISITOS DA SOLUÇÃO DE BI

Tendo em conta os objetivos de médio/longo prazo do Compete2020 entendeu-se que para este projeto deveriam ser tidas as contas as ferramentas que têm uma boa capacidade de visualização de dados bem como potencial para implementar funcionalidades analíticas e com isso aumentar significativamente a utilidade desta solução.

Os relatórios da Gartner são dos mais considerados relativamente ao uso da tecnologia nos negócios pelo que apresentamos aqui o gráfico relativo ao “Magic Quadrant for Analytics and Business Intelligence Platforms”, produzido em fevereiro de 2020 e que representa o posicionamento das principais ferramentas presentes no mercado.

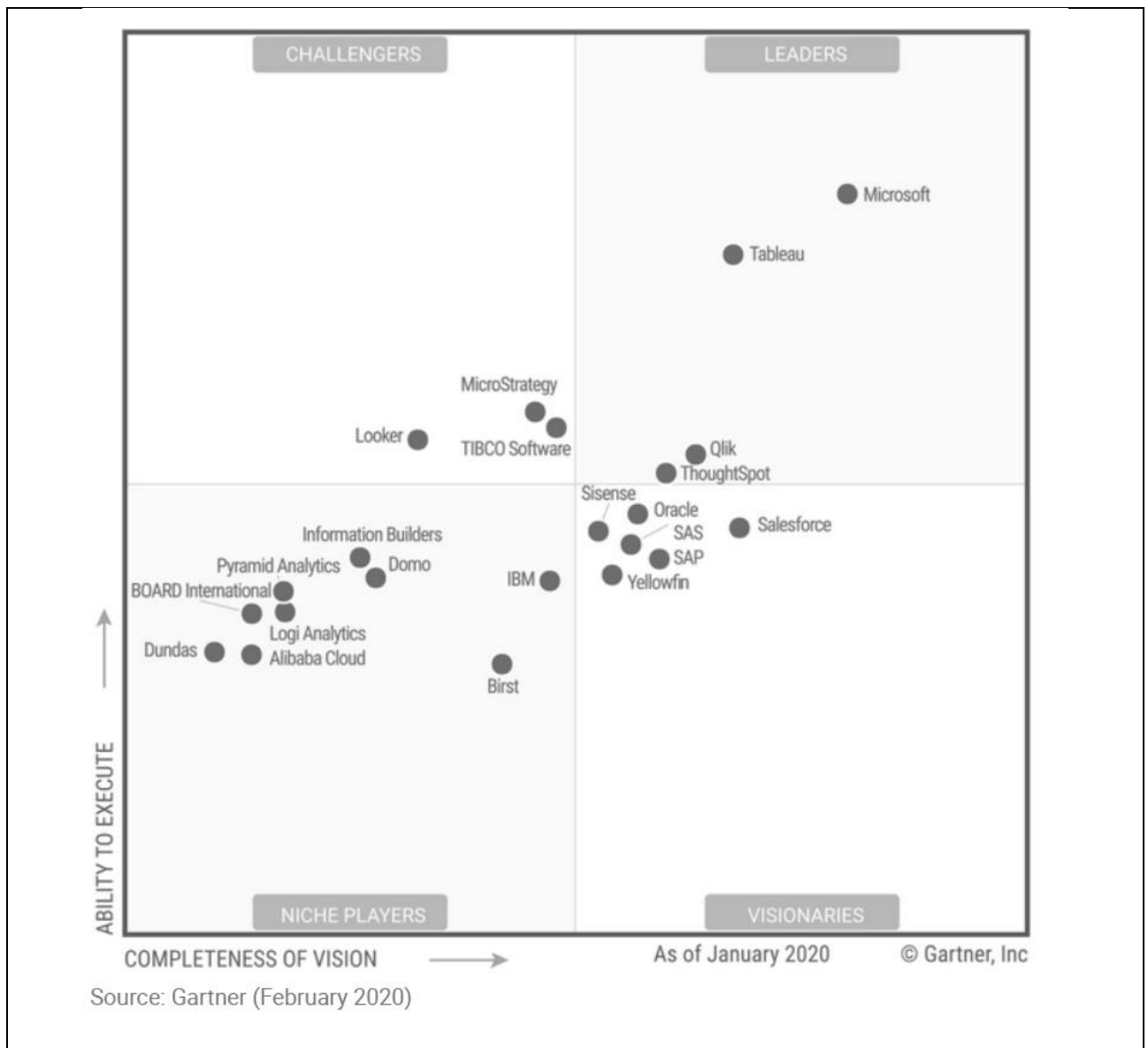


Figura 24: Magic Quadrant for Analytics and Business Intelligence Platforms (fonte Gartner, 2020)

Pela análise do gráfico podemos concluir que os fabricantes/ferramentas melhor colocados estão presentes no espaço designado por 'leaders'. Segundo a análise da Gartner os 'leaders' são os que apresentam uma melhor capacidade de visão quanto às tendências futuras, mas são também capazes de, no presente, disponibilizar ferramentas adequadas ao estado e necessidades atuais das empresas.

Tendo em conta as necessidades do Compete2020 consideram-se como mais relevantes fatores tais como:

- Rápida implementação, de modo a dotar a organização, num curto espaço de tempo, de informação e indicadores úteis à gestão;
- Escalabilidade, capacidade de evoluir em função do número de utilizadores, volumes de dados e funcionalidades;
- Adequada relação preço/funcionalidades;
- Solução ajustada aos conhecimentos/capacidades dos utilizadores do Compete2020 com vista à futura manutenção e implementação de funcionalidades;
- Capacidade de ligar a fontes de dados diversas;
- Capacidade de tratamento dos dados;
- Soluções de visualização e filtragem de dados;
- Capacidade de análise de tendências e disponibilização de funcionalidades analíticas avançadas.

As 3 soluções melhor classificadas pela Gartner: Microsoft, Qlik e Tableau dão resposta a todos os requisitos atrás pelo que se concentrou a análise nestas ferramentas/fabricantes.

4.7 ANÁLISE DE FERRAMENTAS DE BI

Microsoft

A Microsoft é, neste momento, a empresa líder em ferramentas de BI. Possui uma solução escalável que se ajusta às necessidades das empresas. A versão base, gratuita, contempla a solução PowerBI desktop, instalável em cada um dos utilizadores e que permite ligar a dados locais e externos. A partilha faz-se distribuindo o ficheiro *.pbix* pelos utilizadores. Num ambiente de office 365 poderão ser utilizadas as funcionalidades de partilha dentro da rede interna da organização, mas existem custos de licenciamento em função das funcionalidades e tipo/número de utilizadores. A versão PRO do PowerBI permite a partilha dentro e fora da rede da organização.

A versão base do PowerBI possui já algumas ferramentas analíticas e visualizações para análise preditiva. Com o Microsoft Azure toda a solução pode ser implementada na cloud e ferramentas adicionais com capacidade analítica, big data, ML e AI estão disponíveis.

Relativamente aos requisitos mais importantes tendo em conta os objetivos do Compete2020 salientam-se os seguintes pontos dignos de registo:

Pontos Fortes:

- Ausência de custo de licenciamento, ou baixo custo, para a implementação inicial;
- Ferramenta de uso intuitivo, seja para quem utiliza ou produz as dashboards, muito à semelhança de outras ferramentas Microsoft. Dificuldade de aprendizagem reduzida e rápida obtenção de resultados;
- Universo muito elevado de utilizadores e parceiros que promovem a rápida inclusão de novas visualizações e funcionalidades;
- Comunidade muito grande de utilizadores com elevada disponibilidade de informação de suporte, aprendizagem, exemplos, etc.;
- Contínuo crescimento das possibilidades de ligação a informação de sistemas de terceiros potenciando o acesso a novos dados;
- Produto com uma aposta forte da Microsoft em fazê-lo crescer e disponibilizar funcionalidades acrescidas;
- Capacidade de integração com restantes aplicações Microsoft

Pontos Fracos:

- A alternativa ao Azure ou Power BI Pro, com recurso a servidor local (Power BI report server), tem menos funcionalidades disponíveis nomeadamente na área analítica;
- A experiência prática de muitos utilizadores tem evidenciado algumas limitações de velocidade nas conexões a dados externos (via acesso 'direct query') acabando por optar por importar todos os dados para o Power BI;
- O uso de Power BI Desktop com reports e dashboards com muita informação requiere computadores com capacidade significativa de memória e CPUs com bom desempenho.

Qlik

Qlik está presente no mercado há cerca de 20 anos. Tem sido um dos líderes, mas nos últimos anos tem perdido alguma dinâmica. Nos últimos dois anos tem investido na reestruturação do produto, através de aquisições de empresas específicas, com o objetivo de reforçar as capacidades analíticas. É reconhecido

como um dos produtos mais versáteis do ponto de vista de distribuição pelos utilizadores, podendo funcionar em várias plataformas cloud, ou com servidores próprios ou ainda através de SaaS.

Pontos Fortes:

- Facilidade de distribuição e manutenção da solução pelos utilizadores;
- Boa capacidade analítica;
- Forte capacidade de configuração de visualizações.

Pontos Fracos:

- Tem algum grau de dificuldade de utilização para utilizadores com poucos conhecimentos técnicos;
- A elevada capacidade de configuração torna-se, por vezes, um entrave à produtividade e facilidade de utilização;
- Preço;
- Alguma instabilidade na integração dos vários produtos fruto das alterações na empresa nos últimos anos.

Tableau

Em 2003 foi lançada primeira versão deste produto e tem sido um dos principais produtos no mercado na área de BI e Análise de dados. Tendo em conta os custos de licenciamento e de implementação é utilizado maioritariamente por empresas de grande dimensão em projetos com volumes significativos de dados ou com retorno significativo.

Pontos Fortes:

- Elevada capacidade no tratamento de grandes volumes de dados;
- Excelência nas funcionalidades disponíveis para acesso a fonte de dados e desenvolvimento de dashboards;
- Escalabilidade da solução.

Pontos Fracos:

- Ferramenta mais técnica e menos intuitiva de uso mais indicado para analistas de dados do que para utilizadores não avançados;
- Preço elevado e com muitos adicionais em função de funcionalidades necessárias;

- Implementação de projetos carece de afetação de recursos experientes e tem normalmente durações e custos significativos.

O levantamento dos requisitos efetuado no âmbito do projeto, bem como o conhecimento obtido sobre os sistemas existentes permitem concluir que a solução que melhor se ajusta às necessidades do Compete2020 é o PowerBI. De entre os vários fatores que permitem chegar a esta conclusão destacamos:

- O Compete2020 já utiliza algumas ferramentas da Microsoft com qual o PowerBI se integra;
- Custo zero no licenciamento, numa fase inicial, sendo que custos de licenciamento futuro serão em função de funcionalidades que a organização ache oportuno implementar;
- Ferramenta de fácil aprendizagem e fácil utilização, ajustada às capacidades dos recursos humanos do Compete2020, sem necessidade ou com pouca necessidade de envolvimento de recursos externos;
- Protótipo implementado demonstra a viabilidade da solução.

5. CONCLUSÕES

5.1 TRABALHO DESENVOLVIDO

A organização Compete2020 pretende melhorar as suas capacidades de análise e acompanhamento dos fundos por si geridos e para isso procedeu à execução de um projeto de criação do modelo BI para implementação de uma solução futura.

Durante o decorrer do projeto a equipa deparou-se com dificuldades na obtenção de informação sobre os sistemas operacionais e suas estruturas de dados de suporte que obrigaram a opções mais limitadas, mas no final do projeto foi possível apresentar um modelo de dados, bem como um protótipo de dashboards de análise e visualização de dados

Relativamente ao modelo de dados foram desenvolvidos dois modelos. No primeiro modelo chegou-se à conclusão que o conhecimento sobre a informação existente não era suficiente para garantir a fiabilidade dos resultados. A alternativa encontrada foi usar a base de dados do ReAD como base de trabalho para o BI. A solução adotada não corresponde à opção ideal, que seria o acesso aos dados base dos sistemas operacionais (SGO e outros sistemas), apresentando várias limitações devidamente elencadas nos capítulos respetivos.

No âmbito do projeto foi produzido um protótipo funcional que demonstra as potencialidades da estratégia proposta e que permite já dar resposta a todos os indicadores do relatório mensal atualmente em utilização, tendo ainda sido desenvolvidos dashboards para um conjunto de novos indicadores identificados durante a execução do projeto.

No projeto de implementação do BI e caso se consiga o acesso e informação dos sistemas operacionais, seria vantajoso utilizar estes dados em vez dos do ReAD. Além de ultrapassar as limitações identificadas poderá o Compete2020 elaborar um conjunto de indicadores adicionais, bem como efetuar análises preditivas que poderão ajudar na própria operacionalidade da organização.

Quanto às ferramentas de BI foram identificadas, avaliadas as suas funcionalidades e efetuado um estudo comparativo. Apresentadas as opções à organização concluiu esta que o PowerBI seria a solução mais indicada.

5.2 OBJETIVOS ATINGIDOS

No âmbito do projeto foram definidos 3 objetivos que se pretendiam atingir:

- 1) Criação do modelo de dados;
- 2) Construção de um protótipo funcional;
- 3) Identificação e avaliação de potenciais ferramentas de BI;

Todos os objetivos foram atingidos embora alguns deles com algumas limitações.

O modelo de dados foi desenhado e todas as indicações e orientações para evolução futura foram especificados no relatório para futura implementação da solução de BI.

O protótipo funcional foi desenvolvido, testado e apresentado à administração. Conta-se brevemente que seja disponibilizado a um universo restrito de utilizadores para início de utilização.

Ressalve-se que o protótipo, sendo totalmente funcional, necessitará de um projeto de implementação para entrada em produção pois contém soluções implementadas e limitações próprias de um protótipo. A futura solução deverá seguir o modelo de arquitetura de BI descrito e respetiva implementação.

Relativamente à ferramenta de BI a utilizar no futuro concluiu-se que o PowerBI seria a melhor opção principalmente pelas suas capacidades técnicas, facilidade de utilização e custo de licenciamento.

5.3 RECOMENDAÇÕES

No âmbito do projeto foi produzido um protótipo funcional que demonstra as potencialidades de uma solução de BI e que permite já dar resposta a todos os indicadores do relatório mensal atualmente em utilização, tendo ainda sido desenvolvidos dashboards para um conjunto de novos indicadores identificados e caracterizados durante a execução do projeto. Ressalve-se que o protótipo, sendo totalmente funcional, necessitará de um projeto de implementação para entrada em produção pois contém soluções implementadas e limitações próprias de um protótipo. A futura solução deverá seguir o modelo de arquitetura de BI descrito e respetiva implementação.

Relativamente ao modelo de dados a alternativa encontrada foi usar a base de dados do ReAD como base de trabalho para o BI. A solução adotada não corresponde à melhor opção, que seria o acesso aos dados base do SGO e outros sistemas, apresentando várias limitações devidamente elencadas nos capítulos respetivos.

No projeto de implementação do BI e caso se consiga o acesso e informação dos sistemas operacionais, seria vantajoso utilizar estes dados em vez dos do ReAD. Além de ultrapassar as limitações identificadas poderá o Compete2020 elaborar um conjunto de indicadores adicionais, bem como efetuar análises preditivas que poderão ajudar na própria operacionalidade da organização.

Relativamente à ferramenta de BI a utilizar sugere-se a o PowerBI principalmente pelas suas capacidades técnicas, facilidade de utilização e custo de licenciamento.

O sucesso na implementação do projeto de BI dependerá de vários fatores: organizacionais, tecnológicos e humanos, sendo que todos eles são fatores críticos para o sucesso do projeto.

Do ponto de vista organizacional é imprescindível que a administração do Compete2020 seja a dinamizadora do projeto e que os departamentos participantes, quer do ponto de vista técnico quer funcional, estejam alinhados na procura da melhor solução para a organização.

Do ponto de vista tecnológico é importante garantir que:

- As funcionalidades da ferramenta escolhida consigam dar resposta aos requisitos do projeto;
- O licenciamento da ferramenta se ajuste ao modelo de utilização da organização;
- A complexidade técnica da ferramenta é compreensível e gerível pela organização ou se é necessário a intervenção e futura dependência de um parceiro externo.
- A capacidade é ajustada às necessidades e a escalabilidade é a adequada ao crescimento da solução;
- A organização tem, ou conseguirá ter, todos os recursos de hardware e infraestrutura de comunicações e datacenter necessários ao bom funcionamento da solução de BI.

Do ponto de vista de recursos humanos é importante garantir que:

- São envolvidos os recursos que detêm o conhecimento funcional e técnico necessários à implementação da solução;
- Existe a disponibilidade de tempo para colaborar no projeto;
- Existem recursos humanos capazes de receber o conhecimento funcional e técnico da ferramenta a utilizar e são capazes de garantir a manutenção e evolução futura.

6. BIBLIOGRAFIA

- Ballard, C., Herreman, D., Schau, D., Bell, R., Kim, E., & Valencic, A. (1998). Data Modeling Techniques for Data Warehousing. San Jose: IBM Corp.
- Compete2020 (2021). https://www.compete2020.gov.pt/publicacoes/detalhe/Brochura_Compete2020. Obtido a 20/7/2021.
- Davenport, T. (2010). Business intelligence and organizational decisions. IGI Publishing
DOI: 10.4018/jbir.2010071701
- Davenport, T., Leibold, M. & Voelpel, S, (2006). Strategic Management in the Innovation Economy. Wiley.
- Davenport, T., Harris, J. (2007). Competing on Analytics_ The New Science of Winning. Harvard Business School Publishing Corporation.
- Davenport, T. (2013). Analytics 3.0.pdf
<https://hbr.org/2013/12/analytics-30>. Obtido a 11/7/2021
- Deming, W. E. (1986). Out of the Crisis. MIT Press.
- Walton, M. (1986). The Deming Management Method. The Berkeley Publishing Group.
- Howson, C. (2014). Successful Business Intelligence (2nd ed.). McGraw-Hill.
- IAPMEI (2021). <https://www.iapmei.pt/PRODUTOS-E-SERVICOS/Incentivos-Financiamento/Sistemas-de-Incentivos.aspx>. Obtido a 26/7/2021.
- Inmon, W. H. (2005). Building the data warehouse (4th ed.). John Wiley & Sons.
- Kimball R., Reeves L., Ross M. & Thornthwaite W. (1998). The Data Warehouse Lifecycle Toolkit. Wiley Publishing, Inc.
- Kimball, R. & Ross, M. (2013). The Data Warehouse Toolkit: The Definitive Guide to Dimensional Modeling, (3rd ed.). John Wiley & Sons, Inc.
- Kimball, R. & Caserta, J. (2004). The data warehouse ETL toolkit: practical techniques for extracting, cleaning, conforming, and delivering data. Wiley Publishing, Inc.
- Larson, B. (2012). Delivering Business Intelligence with Microsoft® SQL Server® 2012, (3rd ed.). McGraw-Hill
- Love, P., Fong, P. & Irahni, Z., (Eds.). (2005). Management of Knowledge in Project Environments. Elsevier Butterworth-Heinemann
- Luhn, H. (1958). A Business Intelligence System. IBM Journal, 6, 314-319.
<https://ieeexplore.ieee.org/document/5392644/authors#authors>
<http://dx.doi.org/10.1147/rd.24.0314>
- Marr, B. (2017). Data Strategy, How to profit from a world of big data, analytics and the internet of things. Kogan Page.

Pearson, T., (2013). https://www.bain.com/contentassets/25c167a5149c42168994338f9dc99ffe/bain_brief_big_data_the_organizational_challenge2.pdf

Sherman, R. (2015). Business Intelligence Guidebook, From Data Integration to Analytics. Elsevier.

Tableau (2020). What Is Data Visualization?

<https://www.tableau.com/learn/articles/data-visualization> . Obtido a 9/1/2021.

ANEXO I – PRINCIPAIS TABELAS DO READ

Tabela: **psit_projectoop_hist**

Coluna:	Tipo
anomes	char(6)
nup	varchar(45) PK
dt_ref	date PK
dt_act	datetime
nproj	int
ag	char(2)
fundo	varchar(45)
nif	char(9)
nome	varchar(150)
design	varchar(250)
resumo	varchar(5000)
aviso	char(15)
fase	varchar(45)
ot	varchar(45)
pi	varchar(45)
ti	varchar(45)
toto	varchar(45)
instrumento	varchar(120)
sistema	char(5)
medida	char(10)
oi	varchar(45)
cnd_invest	decimal(16,2)
cnd_data	date
apoiado	varchar(45)
org_data	date
dec_dt_p	date
nt_dt_p	date
dec_dt_f	date
nt_dt_f	date
dec_dt_aprov	date
dec_data	date
dt_contrato	date
dec_dt_inicio	date
dec_dt_fim	date
dec_invest	decimal(16,2)
dec_eleg	decimal(16,2)
dec_incent	decimal(16,2)
dec_desp_pub	decimal(16,2)
dec_incent_r	decimal(16,2)
dec_incent_nr	decimal(16,2)
dec_isencaio	decimal(16,2)
enc_data	date
enc_dt_inicio	date
enc_dt_fim	date

Tabela: **psit_series**

Coluna:	Tipo
anomes	char(6) PK
ano	smallint
mes	smallint
grupo	varchar(45) PK
inc_aprovado	decimal(16,2)
inc_contratado	decimal(16,2)
inc_pago	decimal(16,2)
inc_executado	decimal(16,2)
inv_cand	decimal(16,2)
pag_c	decimal(16,2)
pag_nc	decimal(16,2)
pag_r	decimal(16,2)

enc_invest	decimal(16,2)
enc_eleg	decimal(16,2)
enc_dpub	decimal(16,2)
enc_incent	decimal(16,2)
enc_incent_r	decimal(16,2)
enc_incent_nr	decimal(16,2)
enc_isencaio	decimal(16,2)
contratado	char(1)
pagamentos	decimal(16,2)
execucao	decimal(16,2)
ex_eleg	decimal(16,2)
ex_dpub	decimal(16,2)
ex_nac	decimal(16,2)
pag_c	decimal(16,2)
pag_nc	decimal(16,2)
pag_r	decimal(16,2)
anulado	varchar(45)
enei_d	char(250)
an_cae_prom	varchar(45)
an_lst_conc	varchar(250)
an_lst_nute_iii	varchar(250)
an_lst_nute_ii	varchar(250)
an_cae	varchar(45)
ag_d	varchar(45)
oi_d	varchar(120)
medida_d	varchar(120)
eixo	char(4)
eixo_d	varchar(80)
setor	char(2)
setor_d	varchar(45)
ag_setor	char(2)
ag_setor_d	varchar(45)
dimensao	smallint
estado_fse	varchar(120)
aprovado	char(1)
desp_t	decimal(16,2)
dom_interv	varchar(45)
an_tema_fse	varchar(45)
cod_act_econ	varchar(45)
cod_dim_territ	varchar(45)
forma_financ	varchar(45)
defunct	smallint
dt_defunct	date
exclusao	smallint
apr_i_invest	decimal(16,2)
apr_i_eleg	decimal(16,2)
apr_i_incent	decimal(16,2)
pag_incent_r	decimal(16,2)
execucao_sifse	decimal(16,2)
ex_eleg_sifse	decimal(16,2)
ex_dpub_sifse	decimal(16,2)

ANEXO II - PRINCIPAIS INDICADORES A IMPLEMENTAR

Tipo de Indicadores	Indicadores	Métrica	Implementado?	Possível?	Observações
Programa	Projetos Aprovados	N.º de projetos	Sim		
Programa	Projetos Contratados	N.º de projetos	Sim		
Programa	Projetos com pagamentos	N.º de projetos	Sim		
Programa	Projetos com execução	N.º de projetos	Sim		
Programa	Projetos encerrados	N.º de projetos	Sim		
Programa	Projetos Aprovados	Incentivo	Sim		
Programa	Projetos Contratados	Incentivo	Sim		
Programa	Projetos com pagamentos	Incentivo	Sim		
Programa	Projetos com execução	Incentivo	Sim		
Programa	Projetos encerrados	Incentivo	Sim		
Programa	Projetos Aprovados	% Dotação	Sim		
Programa	Projetos Contratados	% Dotação	Sim		
Programa	Projetos com pagamentos	% Dotação	Sim		
Programa	Projetos com execução	% Dotação	Sim		
Programa	Projetos encerrados	% Dotação	Sim		
Programa	Projetos Aprovados	% Compromisso	Sim		
Programa	Projetos Contratados	% Compromisso	Sim		
Programa	Projetos com pagamentos	% Compromisso	Sim		
Programa	Projetos com execução	% Compromisso	Sim		
Programa	Projetos encerrados	% Compromisso	Sim		
Pagamentos (incluindo proj. anulados)	Pagamentos Totais	N.º de projetos	Sim		inclui projetos anulados/desistidos
Pagamentos (incluindo proj. anulados)	Adiantamentos Certificáveis	N.º de projetos	Sim	Sim	
Pagamentos (incluindo proj. anulados)	Adiantamentos Não Certificáveis	N.º de projetos	Não	Sim	
Pagamentos (incluindo proj. anulados)	Pagamentos de Despesas	N.º de projetos	Não	Sim	
Pagamentos (incluindo proj. anulados)	Pagamentos Totais	Incentivo	Sim	Sim	
Pagamentos (incluindo proj. anulados)	Adiantamentos Certificáveis	Incentivo	Sim	Sim	
Pagamentos (incluindo proj. anulados)	Adiantamentos Não Certificáveis	Incentivo	Não	Sim	
Pagamentos (incluindo proj. anulados)	Pagamentos de Despesas	Incentivo	Não	Sim	
Pagamentos (incluindo proj. anulados)	Pagamentos Totais	% Pagamentos	Sim	Sim	
Pagamentos (incluindo proj. anulados)	Adiantamentos Certificáveis	% Pagamentos	Sim	Sim	
Pagamentos (incluindo proj. anulados)	Adiantamentos Não Certificáveis	% Pagamentos	Não	Sim	
Pagamentos (incluindo proj. anulados)	Pagamentos de Despesas	% Pagamentos	Não	Sim	
Execução (inclui projetos anulados)	Projetos com execução	N.º de projetos	Sim	n.a.	
Execução (inclui projetos anulados)	Projetos incluídos no último PCD à CE	N.º de projetos	n.a.	n.a.	
Execução (inclui projetos anulados)	Projetos a incluir no último PCD à CE	N.º de projetos	n.a.	n.a.	
Execução (inclui projetos anulados)	Projetos com execução	Incentivo	Sim	n.a.	
Execução (inclui projetos anulados)	Projetos incluídos no último PCD à CE	Incentivo	n.a.	n.a.	
Execução (inclui projetos anulados)	Projetos a incluir no último PCD à CE	Incentivo	n.a.	n.a.	
Execução (inclui projetos anulados)	Projetos com execução	% Dotação	Sim	n.a.	
Execução (inclui projetos anulados)	Projetos incluídos no último PCD à CE	% Dotação	n.a.	n.a.	
Execução (inclui projetos anulados)	Projetos a incluir no último PCD à CE	% Dotação	n.a.	n.a.	
Execução (inclui projetos anulados)	Projetos com execução	% Execução	Sim	n.a.	
Execução (inclui projetos anulados)	Projetos incluídos no último PCD à CE	% Execução	n.a.	n.a.	
Execução (inclui projetos anulados)	Projetos a incluir no último PCD à CE	% Execução	n.a.	n.a.	

Indicadores de execução	Projetos aprovados sem contrato	N.º de projetos	Não	Sim	
Indicadores de execução	Projetos Contratados sem pagamento	N.º de projetos	Não	Sim	
Indicadores de execução	Projetos apenas com adiantamentos	N.º de projetos	Não	Sim	
Indicadores de execução	Projetos com pagamento de despesas sem pedido final	N.º de projetos	Não	Não	Incluir no ReAD o pedido final
Indicadores de execução	Projetos com pagamento final sem encerramento	N.º de projetos	Não	Não	Incluir no ReAD o pedido final
Indicadores de execução	Projetos encerrados	N.º de projetos	Sim	Sim	
Indicadores de execução	Projetos aprovados sem contrato	Investimento	Não	Sim	
Indicadores de execução	Projetos Contratados sem pagamento	Investimento	Não	Sim	
Indicadores de execução	Projetos apenas com adiantamentos	Investimento	Não	Sim	
Indicadores de execução	Projetos com pagamento de despesas sem pedido final	Investimento	Não	Não	Incluir no ReAD o pedido final
Indicadores de execução	Projetos com pagamento final sem encerramento	Investimento	Não	Não	Incluir no ReAD o pedido final
Indicadores de execução	Projetos encerrados	Investimento	Sim	Sim	
Indicadores de execução	Projetos aprovados sem contrato	Investimento elegível	Não	Sim	
Indicadores de execução	Projetos Contratados sem pagamento	Investimento elegível	Não	Sim	
Indicadores de execução	Projetos apenas com adiantamentos	Investimento elegível	Não	Sim	
Indicadores de execução	Projetos com pagamento de despesas sem pedido final	Investimento elegível	Não	Não	Incluir no ReAD o pedido final
Indicadores de execução	Projetos com pagamento final sem encerramento	Investimento elegível	Não	Não	Incluir no ReAD o pedido final
Indicadores de execução	Projetos encerrados	Investimento elegível	Sim	Sim	
Indicadores de execução	Projetos aprovados sem contrato	Incentivo	Não	Sim	
Indicadores de execução	Projetos Contratados sem pagamento	Incentivo	Não	Sim	
Indicadores de execução	Projetos apenas com adiantamentos	Incentivo	Não	Sim	
Indicadores de execução	Projetos com pagamento de despesas sem pedido final	Incentivo	Não	Não	Incluir no ReAD o pedido final
Indicadores de execução	Projetos com pagamento final sem encerramento	Incentivo	Não	Não	Incluir no ReAD o pedido final
Indicadores de execução	Projetos encerrados	Incentivo	Sim	Sim	
Indicadores de execução	Projetos aprovados sem contrato	Pagamentos	Não	Sim	
Indicadores de execução	Projetos Contratados sem pagamento	Pagamentos	Não	Sim	
Indicadores de execução	Projetos apenas com adiantamentos	Pagamentos	Não	Sim	
Indicadores de execução	Projetos com pagamento de despesas sem pedido final	Pagamentos	Não	Não	Incluir no ReAD o pedido final
Indicadores de execução	Projetos com pagamento final sem encerramento	Pagamentos	Não	Não	Incluir no ReAD o pedido final
Indicadores de execução	Projetos encerrados	Pagamentos	Sim	Sim	
Indicadores de execução	Projetos aprovados sem contrato	Execução	Não	Sim	
Indicadores de execução	Projetos Contratados sem pagamento	Execução	Não	Sim	
Indicadores de execução	Projetos apenas com adiantamentos	Execução	Não	Sim	
Indicadores de execução	Projetos com pagamento de despesas sem pedido final	Execução	Não	Não	Incluir no ReAD o pedido final
Indicadores de execução	Projetos com pagamento final sem encerramento	Execução	Não	Não	Incluir no ReAD o pedido final
Indicadores de execução	Projetos encerrados	Execução	Sim	Sim	
Indicadores de execução	Projetos aprovados sem contrato	% Dotação	Não	Sim	
Indicadores de execução	Projetos Contratados sem pagamento	% Dotação	Não	Sim	
Indicadores de execução	Projetos apenas com adiantamentos	% Dotação	Não	Sim	
Indicadores de execução	Projetos com pagamento de despesas sem pedido final	% Dotação	Não	Não	Incluir no ReAD o pedido final
Indicadores de execução	Projetos com pagamento final sem encerramento	% Dotação	Não	Não	Incluir no ReAD o pedido final
Indicadores de execução	Projetos encerrados	% Dotação	Não	Não	
Indicadores de execução	Projetos aprovados sem contrato	% Compromisso	Não	Sim	
Indicadores de execução	Projetos Contratados sem pagamento	% Compromisso	Não	Sim	
Indicadores de execução	Projetos apenas com adiantamentos	% Compromisso	Não	Sim	
Indicadores de execução	Projetos com pagamento de despesas sem pedido final	% Compromisso	Não	Não	Incluir no ReAD o pedido final

Indicadores de execução	Projetos com pagamento final sem encerramento	% Compromisso	Não	Não	Incluir no ReAD o pedido final
Indicadores de execução	Projetos encerrados	% Compromisso	Não	Não	
Análise de candidaturas - resumo	Candidaturas em fase de análise - concursos	N.º de projetos	Não	Não	O ReAD não tem esta informação
Análise de candidaturas - resumo	Candidaturas em fase de análise - convite	N.º de projetos	Não	Não	O ReAD não tem esta informação
Análise de candidaturas - resumo	Candidaturas em fase de análise - concursos	Investimento	Não	Não	O ReAD não tem esta informação
Análise de candidaturas - resumo	Candidaturas em fase de análise - convite	Investimento	Não	Não	O ReAD não tem esta informação
Análise de candidaturas - resumo	Candidaturas em fase de análise - concursos	%Candidaturas	Não	Não	O ReAD não tem esta informação
Análise de candidaturas - resumo	Candidaturas em fase de análise - convite	%Candidaturas	Não	Não	O ReAD não tem esta informação
Parecer do organismo	Projetos com parecer do organismo em atraso	N.º de projetos	Não	Não	O ReAD não tem esta informação
Parecer do organismo	Projetos com parecer do organismo em atraso	Investimento	Não	Não	O ReAD não tem esta informação
Parecer do organismo	Projetos com parecer do organismo em atraso	% sem parecer	Não	Não	O ReAD não tem esta informação
Alegações, reclamações, recursos e pedidos de ajuste	Projetos com alegações contrárias fora de prazo	N.º de projetos	Não	Não	O ReAD não tem esta informação
Alegações, reclamações, recursos e pedidos de ajuste	Projetos com reclamações fora de prazo	N.º de projetos	Não	Não	O ReAD não tem esta informação
Alegações, reclamações, recursos e pedidos de ajuste	Projetos com recursos ou pedidos de ajuste pendentes	N.º de projetos	Não	Não	O ReAD não tem esta informação
Alegações, reclamações, recursos e pedidos de ajuste	Projetos com alegações contrárias fora de prazo	Incentivo	Não	Não	O ReAD não tem esta informação
Alegações, reclamações, recursos e pedidos de ajuste	Projetos com reclamações fora de prazo	Incentivo	Não	Não	O ReAD não tem esta informação
Alegações, reclamações, recursos e pedidos de ajuste	Projetos com recursos ou pedidos de ajuste pendentes	Incentivo	Não	Não	O ReAD não tem esta informação
Alegações, reclamações, recursos e pedidos de ajuste	Projetos com alegações contrárias fora de prazo	Investimento	Não	Não	O ReAD não tem esta informação
Alegações, reclamações, recursos e pedidos de ajuste	Projetos com reclamações fora de prazo	Investimento	Não	Não	O ReAD não tem esta informação
Alegações, reclamações, recursos e pedidos de ajuste	Projetos com recursos ou pedidos de ajuste pendentes	Investimento	Não	Não	O ReAD não tem esta informação
Alegações, reclamações, recursos e pedidos de ajuste	Projetos com alegações contrárias fora de prazo	% pendentes	Não	Não	O ReAD não tem esta informação
Alegações, reclamações, recursos e pedidos de ajuste	Projetos com reclamações fora de prazo	% pendentes	Não	Não	O ReAD não tem esta informação
Alegações, reclamações, recursos e pedidos de ajuste	Projetos com recursos ou pedidos de ajuste pendentes	% pendentes	Não	Não	O ReAD não tem esta informação
Bolsa descativação incentivo	Projetos com contratação em atraso	N.º de projetos	Não	Não	O ReAD não tem esta informação
Bolsa descativação incentivo	Projetos com prazo ultrapassado para apresentação pedido final	N.º de projetos	Não	Não	O ReAD não tem esta informação
Bolsa descativação incentivo	Projetos com metas de execução em incumprimento	N.º de projetos	Não	Não	O ReAD não tem esta informação
Bolsa descativação incentivo	Projetos com adiantamentos sem justificação	N.º de projetos	Não	Não	O ReAD não tem esta informação
Bolsa descativação incentivo	Projetos com contratação em atraso	Incentivo	Não	Não	O ReAD não tem esta informação
Bolsa descativação incentivo	Projetos com prazo ultrapassado para apresentação pedido final	Incentivo	Não	Não	O ReAD não tem esta informação
Bolsa descativação incentivo	Projetos com metas de execução em incumprimento	Incentivo	Não	Não	O ReAD não tem esta informação
Bolsa descativação incentivo	Projetos com adiantamentos sem justificação	Incentivo	Não	Não	O ReAD não tem esta informação
Bolsa descativação incentivo	Projetos com contratação em atraso	%Incluídos BDI	Não	Não	O ReAD não tem esta informação
Bolsa descativação incentivo	Projetos com prazo ultrapassado para apresentação pedido final	%Incluídos BDI	Não	Não	O ReAD não tem esta informação
Bolsa descativação incentivo	Projetos com metas de execução em incumprimento	%Incluídos BDI	Não	Não	O ReAD não tem esta informação
Bolsa descativação incentivo	Projetos com adiantamentos sem justificação	%Incluídos BDI	Não	Não	O ReAD não tem esta informação
Bolsa de Recuperação	Projeto aprovado sem termo de aceitação nos 90 dias após a data de notificação da decisão		Não	Sim	Identificar datas a utilizar

Bolsa de Recuperação	Projeto contratado há mais de 12 meses, sem início de execução		Não	Sim	Identificar datas a utilizar
Bolsa de Recuperação	Projeto contratado há mais de 12 meses, com execução inferior a 10%		Não	Sim	Identificar datas a utilizar
Bolsa de Recuperação	Projeto iniciados, com execução maior ou igual a 10% e sem acréscimos de execução nos últimos 12 meses há mais de 12 meses, sem início de execução		Não	Sim	Identificar datas a utilizar
Operação limpeza	Projetos contratados há mais de 6 meses com execução inferior a 10%	N.º de projetos	Não	Sim	Identificar datas a utilizar
Operação limpeza	Projetos contratados há mais de 6 meses com execução inferior a 10%	Incentivo	Não	Sim	Identificar datas a utilizar
Operação limpeza	Projetos contratados há mais de 6 meses com execução inferior a 10%	% contratado + 6 meses	Não	Sim	Identificar datas a utilizar
Pedidos de Pagamento	Pedidos de pagamento em análise há mais de 1 ano	N.º de projetos	Não	Não	Não existe informação detalhada dos PP
Pedidos de Pagamento	Pedidos de pagamento em análise entre 6 meses e 1 ano	N.º de projetos	Não	Não	Não existe informação detalhada dos PP
Pedidos de Pagamento	Pedidos de pagamento em análise há menos de 6 meses	N.º de projetos	Não	Não	Não existe informação detalhada dos PP
Pedidos de Pagamento	Pedidos de pagamento em análise há mais de 1 ano	Incentivo	Não	Não	Não existe informação detalhada dos PP
Pedidos de Pagamento	Pedidos de pagamento em análise entre 6 meses e 1 ano	Incentivo	Não	Não	Não existe informação detalhada dos PP
Pedidos de Pagamento	Pedidos de pagamento em análise há menos de 6 meses	Incentivo	Não	Não	Não existe informação detalhada dos PP
Pedidos de Pagamento	Pedidos de pagamento em análise há mais de 1 ano	% em análise	Não	Não	Não existe informação detalhada dos PP
Pedidos de Pagamento	Pedidos de pagamento em análise entre 6 meses e 1 ano	% em análise	Não	Não	Não existe informação detalhada dos PP
Pedidos de Pagamento	Pedidos de pagamento em análise há menos de 6 meses	% em análise	Não	Não	Não existe informação detalhada dos PP
Encerramento	Projetos com pedido de pagamento final sem encerramento	N.º de projetos	Não	Não	Não existe informação se fez o pedido final
Encerramento	Projetos com pedido de pagamento final sem encerramento	Incentivo	Não	Não	Não existe informação se fez o pedido final
Encerramento	Projetos com pedido de pagamento final sem encerramento	% com pedido final	Não	Não	Não existe informação se fez o pedido final
Tempo global de análise de candidaturas			Sim		
Tempos médios de análise e desvio face aos prazos (dias úteis)	Tempo médio de análise de candidaturas - Parecer OI		Não	Não	Não existe informação necessária
Tempos médios de análise e desvio face aos prazos (dias úteis)	Tempo médio de análise de candidaturas - Decisão		Não	Não	Não existe informação necessária
Tempos médios de análise e desvio face aos prazos (dias úteis)	Tempo médio de análise de candidaturas - Resposta a reclamações		Não	Não	Não existe informação necessária
Tempos médios de análise e desvio face aos prazos (dias úteis)	Tempo médio de análise de candidaturas - Resposta a recursos		Não	Não	Não existe informação necessária
Tempos médios de análise e desvio face aos prazos (dias úteis)	Tempo médio de análise de candidaturas - Resposta a pedidos de ajuste		Não	Não	Não existe informação necessária
Tempos médios de análise e desvio face aos prazos (dias úteis)	Tempo médio da contratação		Não	Sim	
Tempos médios de análise e desvio face aos prazos (dias úteis)	Tempo médio para apresentação do 1.º pedido de pagamento		Não	Não	Não existe informação necessária
Tempos médios de análise e desvio face aos prazos (dias úteis)	Tempo médio de análise dos pedidos de pagamento		Não	Não	Não existe informação necessária
Tempos médios de análise e desvio face aos prazos (dias úteis)	Tempo médio de encerramento após apresentação do pedido de pagamento final		Não	Não	Não existe informação necessária
Tempos médios de análise e desvio face aos prazos (dias úteis)	Tempo médio de análise de candidaturas - Parecer OI		Não	Não	Não existe informação necessária
Tempos médios de análise e desvio face aos prazos (dias úteis)	Tempo médio de análise de candidaturas - Decisão		Não	Não	Não existe informação necessária
Tempos médios de análise e desvio face aos prazos (dias úteis)	Tempo médio de análise de candidaturas - Resposta a reclamações		Não	Não	Não existe informação necessária
Tempos médios de análise e desvio face aos prazos (dias úteis)	Tempo médio de análise de candidaturas - Resposta a recursos		Não	Não	Não existe informação necessária

Tempos médios de análise e desvio face aos prazos (dias úteis)	Tempo médio de análise de candidaturas - Resposta a pedidos de ajuste		Não	Não	Não existe informação necessária
Tempos médios de análise e desvio face aos prazos (dias úteis)	Tempo médio da contratação		Não	Sim	
Tempos médios de análise e desvio face aos prazos (dias úteis)	Tempo médio para apresentação do 1.º pedido de pagamento		Não	Não	Não existe informação necessária
Tempos médios de análise e desvio face aos prazos (dias úteis)	Tempo médio de análise dos pedidos de pagamento		Não	Não	Não existe informação necessária
Tempos médios de análise e desvio face aos prazos (dias úteis)	Tempo médio de encerramento após apresentação do pedido de pagamento final		Não	Não	Não existe informação necessária
Cumprimento dos prazos previstos	% projetos com parecer do organismos fora do prazo previsto face ao total		Não	Não	Não existe informação necessária
Cumprimento dos prazos previstos	% projetos com comunicação de decisão fora do prazo previsto face ao total		Não	Não	Não existe informação necessária
Cumprimento dos prazos previstos	% projetos contratados fora do prazo previsto face ao total		Não	Não	Não existe informação necessária
Cumprimento dos prazos previstos	% projetos com parecer do organismos fora do prazo previsto face ao total		Não	Não	Não existe informação necessária
Cumprimento dos prazos previstos	% projetos com comunicação de decisão fora do prazo previsto face ao total		Não	Não	Não existe informação necessária
Cumprimento dos prazos previstos	% projetos contratados fora do prazo previsto face ao total		Não	Não	Não existe informação necessária
Cumprimento dos prazos previstos	% projetos com parecer do organismos fora do prazo previsto face ao total		Não	Não	Não existe informação necessária
Cumprimento dos prazos previstos	% projetos com comunicação de decisão fora do prazo previsto face ao total		Não	Não	Não existe informação necessária
Cumprimento dos prazos previstos	% projetos contratados fora do prazo previsto face ao total		Não	Não	Não existe informação necessária
Alegações contrárias	% de alegações contrárias apresentadas face ao total de projetos não elegíveis		Não	Não	Não existe informação necessária
Alegações contrárias	% de alegações contrárias com razão para p promotor face ao total de alegações contrárias		Não	Não	Não existe informação necessária
Alegações contrárias	% de alegações contrárias apresentadas face ao total de projetos não elegíveis		Não	Não	Não existe informação necessária
Alegações contrárias	% de alegações contrárias com razão para p promotor face ao total de alegações contrárias		Não	Não	Não existe informação necessária
Alegações contrárias	% de alegações contrárias apresentadas face ao total de projetos não elegíveis		Não	Não	Não existe informação necessária
Alegações contrárias	% de alegações contrárias com razão para p promotor face ao total de alegações contrárias		Não	Não	Não existe informação necessária
Pedidos de Pagamento	% de pedidos de pagamento com análise < 6 meses face ao total		Não	Não	Não existe informação necessária
Pedidos de Pagamento	% de pedidos de pagamento com análise entre 6 meses e 1 ano face ao total		Não	Não	Não existe informação necessária
Pedidos de Pagamento	% de pedidos de pagamento com análise > 1 ano face ao total		Não	Não	Não existe informação necessária
Pedidos de Pagamento	% de pedidos de pagamento com análise < 6 meses face ao total		Não	Não	Não existe informação necessária
Pedidos de Pagamento	% de pedidos de pagamento com análise entre 6 meses e 1 ano face ao total		Não	Não	Não existe informação necessária
Pedidos de Pagamento	% de pedidos de pagamento com análise > 1 ano face ao total		Não	Não	Não existe informação necessária
Pedidos de Pagamento	% de pedidos de pagamento com análise < 6 meses face ao total		Não	Não	Não existe informação necessária
Pedidos de Pagamento	% de pedidos de pagamento com análise entre 6 meses e 1 ano face ao total		Não	Não	Não existe informação necessária
Pedidos de Pagamento	% de pedidos de pagamento com análise > 1 ano face ao total		Não	Não	Não existe informação necessária

ANEXO IV – ETL DETALHADO

	Source = MySQLDatabase("192.168.1.113", "wd_read", [ReturnSingleDatabase=true, HierarchicalNavigation=true]),
	wd_read = Source({Schema="wd_read"})[Data],
Ler dados de 'psit_ProjectoOP_Hist'	psit_projectoop_hist1 = wd_read[{Name="psit_ProjectoOP_Hist"}][Data],
Filtro limitador de recolha de dados	#"Filtered Rows1" = Table.SelectRows(psit_projectoop_hist1, each [anomes] >= "201907"),
Criar campo auxiliar 'ConjuntoProj' (1 SI, 0 outros)	#"Added Custom9" = Table.AddColumn(#"Filtered Rows1", "ConjuntoProj", each Int8.From([(sistema] = "d" or [sistema] = "e" or [sistema] = "g" or [sistema] = "i" or [sistema] = "q") and ([ag] <> "81"))),
Criar campo auxiliar 'Compete' (2 Compete, 0 outros)	#"Added Conditional Column8" = Table.AddColumn(#"Added Custom9", "Compete", each if [ag] = "01" then 2 else 0),
Criar campo auxiliar 'Cod grupo' (2 Compete, 1 SI, 3 ambos)	#"Added Custom10" = Table.AddColumn(#"Added Conditional Column8", "Cod grupo", each [ConjuntoProj]+[Compete]),
Criar campo 'INV MIL' = 'cnd_invest'	#"Duplicated Column" = Table.DuplicateColumn(#"Added Custom10", "cnd_invest", "INV MIL"),
Criar campo 'Apoiado?' = (if [dec_data] = null then null else [apoiado])	#"Added Conditional Column" = Table.AddColumn(#"Duplicated Column", "Apoiado?", each if [dec_data] = null then null else [apoiado]),
Criar campo 'Nº Proj Aprov.' = (if ([apoiado] = "Sim" and [anulado]="") then 1 else 0)	#"Added Conditional Column1" = Table.AddColumn(#"Added Conditional Column", "Nº Proj Aprov.", each if ([apoiado] = "Sim" and [anulado]="") then 1 else 0),
Criar campo "Com TA?" = [contratado]	#"Added Custom" = Table.AddColumn(#"Added Conditional Column1", "Com TA?", each [contratado]),
Criar campo "Com Pagamentos?" = (if [pagamentos] > 0 then 1 else 0)	#"Added Conditional Column2" = Table.AddColumn(#"Added Custom", "Com Pagamentos?", each if [pagamentos] > 0 then 1 else 0),
Substituir valores NULL por "" em {"an_lst_conc", "an_lst_nute_iii", "an_lst_nute_ii"}),	#"Replaced Value4" = Table.ReplaceValue(#"Added Conditional Column2", null, "", Replacer.ReplaceValue, {"an_lst_conc", "an_lst_nute_iii", "an_lst_nute_ii"}),
Criar campo 'NUTS III (trab)' em função de 'an_lst_nute_iii'	#"Added Conditional Column3" = Table.AddColumn(#"Replaced Value4", "NUTS III (trab)", each if Text.Contains([an_lst_nute_iii], ",") then "Multirregiões ou n.a./n.d." else if (Text.Contains([an_lst_nute_iii], "RA Açores") or Text.Contains([an_lst_nute_iii], "RA Madeira") or Text.Contains([an_lst_nute_iii], "Estrangeiro") or ([an_lst_nute_iii] = "" or ([an_lst_nute_iii] = null))) then "Multirregiões ou n.a./n.d." else [an_lst_nute_iii]),
Criar campo 'NUTS II (trab)' em função de 'an_lst_nute_iii'	#"Added Conditional Column4" = Table.AddColumn(#"Added Conditional Column3", "NUTS II (trab)", each if Text.Contains([an_lst_nute_iii], ",") then "Multirregiões ou n.a./n.d." else if (Text.Contains([an_lst_nute_iii], "RA Açores") or Text.Contains([an_lst_nute_iii], "RA Madeira") or Text.Contains([an_lst_nute_iii], "Estrangeiro") or ([an_lst_nute_iii] = "" or ([an_lst_nute_iii] = null))) then "Multirregiões ou n.a./n.d." else [an_lst_nute_iii]),
	#"Merged Queries" = Table.NestedJoin(#"Added Conditional Column4", {"an_lst_conc"}, #"DIMREAD tbc_concelho_baixadensidade", {"codigo"}, "READ tbc_concelho_baixadensidade", JoinKind.LeftOuter),
Criar campo 'Municípios de baixa densidade?' com base na tabela 'READ tbc_concelho_baixadensidade'	#"Expanded READ tbc_concelho_baixadensidade" = Table.ExpandTableColumn(#"Merged Queries", "READ tbc_concelho_baixadensidade", {"design"}, {"Municípios de Baixa Densidade?"}),
Preencher campo 'Municípios de baixa densidade?' com "ND ou NE" se for NULL	#"Replaced Value" = Table.ReplaceValue(#"Expanded READ tbc_concelho_baixadensidade", null, "ND ou NE", Replacer.ReplaceValue, {"Municípios de Baixa Densidade?"}),
	#"Merged Queries1" = Table.NestedJoin(#"Replaced Value", {"an_cae"}, #"DIMREAD tbc_pam_cae", {"codigo"}, "READ tbc_pam_cae", JoinKind.LeftOuter),
Criar campos com base na tabela "READ tbc_pam_cae", {"setor", "agrup_sectorial", "agrup_industrial", "efice", "itc"}	#"Expanded READ tbc_pam_cae" = Table.ExpandTableColumn(#"Merged Queries1", "READ tbc_pam_cae", {"setor", "agrup_sectorial", "agrup_industrial", "efice", "itc"}),
Preencher campos {"setor", "agrup_sectorial", "agrup_industrial", "efice", "itc"} com "ND ou NE" se NULL	#"Replaced Value1" = Table.ReplaceValue(#"Expanded READ tbc_pam_cae", null, "ND ou NE", Replacer.ReplaceValue, {"READ tbc_pam_cae.setor", "READ tbc_pam_cae.agrup_sectorial", "READ tbc_pam_cae.agrup_industrial", "READ tbc_pam_cae.efice", "READ tbc_pam_cae.itc"}),
	#"Renamed Columns" = Table.RenameColumns(#"Replaced Value1", {"READ tbc_pam_cae.setor", "Sector"}, {"READ tbc_pam_cae.agrup_sectorial", "Agrupamentos Setoriais"}, {"READ tbc_pam_cae.agrup_industrial", "Grandes Agrupamentos Industriais"}, {"READ tbc_pam_cae.efice", "Classificação Indústria EFICE"}, {"READ tbc_pam_cae.itc", "Intensidade Tecnológica e de Conhecimento"}),
Preencher campos {"Sector", "Agrupamentos Setoriais", "Grandes Agrupamentos Industriais", "Classificação Indústria EFICE", "Intensidade Tecnológica e de Conhecimento"} com "ND ou NE" se NULL	#"Replaced Value2" = Table.ReplaceValue(#"Renamed Columns", null, "ND ou NE", Replacer.ReplaceValue, {"Sector", "Agrupamentos Setoriais", "Grandes Agrupamentos Industriais", "Classificação Indústria EFICE", "Intensidade Tecnológica e de Conhecimento"}),
Transformar tipo de dados de 'enc_data' para TEXT	#"Changed Type2" = Table.TransformColumnTypes(#"Replaced Value2", {"enc_data", type text}),
Preencher 'enc_data'="" se NULL	#"Replaced Value3" = Table.ReplaceValue(#"Changed Type2", null, "", Replacer.ReplaceValue, {"enc_data"}),
Calculo auxiliar de1 (valor final elegível)	#"Added Conditional Column5" = Table.AddColumn(#"Replaced Value3", "de1", each if ((([enc_data] <> "") and ([anulado]="")) then [enc_eleg] else [dec_eleg]),
Calculo auxiliar de1 (valor final incentivo)	#"Added Conditional Column6" = Table.AddColumn(#"Added Conditional Column5", "di1", each if ((([enc_data] <> "") and ([anulado]="")) then [enc_incent] else [dec_incent]),
Guardar valor original de 'dec_eleg' em 'dec_eleg_original' e "dec_incent" em "dec_incent_original"	#"Renamed Columns4" = Table.RenameColumns(#"Added Conditional Column6", {"dec_eleg", "dec_eleg_original"}, {"dec_incent", "dec_incent_original"}),
Renomear campos {"de1" para "dec_eleg"} e {"di1" para "dec_incent"}	#"Renamed Columns3" = Table.RenameColumns(#"Renamed Columns4", {"de1", "dec_eleg"}, {"di1", "dec_incent"}),
Alterar tipos de dados: {"dec_incent", type number}, {"dec_eleg", type number}	#"Changed Type3" = Table.TransformColumnTypes(#"Renamed Columns3", {"dec_incent", type number}, {"dec_eleg", type number}),
Renomear campos:	#"Renamed Columns1" = Table.RenameColumns(#"Changed Type3", {"nup", "Código de Operação"}, {"nproj", "NºProj"}, {"ni", "NIE"}, {"nome", "Beneficiário"}, {"pi", "PI"}, {"medida", "Cmedida"}, {"medida_d", "Medida_d"}, {"ag_d", "PO Finan."}, {"oi_d", "Organismo"}, {"cnd_invest",

	"[Cand]Investimento"}, {"dec_eleg", "[Dec]Elegivel"}, {"dec_incent", "[Dec]Incentivo"}, {"execucao", "[Exec]Fundo"}, {"pagamentos", "Pagamento Total"}, {"anulado", "DST/ANL"}, {"enei_d", "ENEI"}, {"dimensao", "Dimensao"}, {"an_lst_conc", "Concelho(s)"}, {"aviso", "Aviso"}}, {"Removed Other Columns" = Table.SelectColumns("#Renamed Columns1", {"anomes", "dt_ref", "dt_act", "Codigo de Operacao", "Nº Proj Aprov.", "Com TA2", "Com Pagamentos"}, {"Exec]Fundo", "Nº Proj", "NIF", "Beneficiario", "PI", "fundo", "instrumento", "Medida", "Aviso", "PO Finan.", "Organismo"}, {"Cand]Investimento", "INV MIL", "Apoiado"}, {"Dec]Elegivel", "[Dec]Incentivo", "Pagamento Total", "DST/ANL", "ENEI", "Dimensao", "NUTS III (trab)", "NUTS II (trab)", "Municipios de Baixa Densidade", "Sector", "Agrupamentos Setoriais", "Grandes Agrupamentos Industriais", "Classificacao Industria EFICE", "Intensidade Tecnologica e de Conhecimento", "Concelho(s)", "cnd_data", "dec_dt_p", "nt_dt_p", "dec_data", "dt_contrato", "dec_dt_inicio", "dec_dt_fim", "enc_dt_fim", "enc_dt_inicio", "dec_invest", "enc_invest", "enc_eleg", "enc_incent", "dec_eleg_original", "dec_incent_original", "contratado", "sistema", "Cod grupo", "desp_t", "ot", "an_cae"});
Remover campos não utilizados	
Alterar tipo de dados	"#_Changed Type" = Table.TransformColumnTypes("#Removed Other Columns", {"anomes", Int64.Type}, {"Nº Proj Aprov.", Int64.Type}, {"Com Pagamentos", Int64.Type}, {"NIF", Int64.Type}, {"Com TA2", type text}, {"Apoiado?", type text}, {"NUTS III (trab)", type text}, {"NUTS II (trab)", type text});
Criar campo "FLAG Aprovados" = {"#Nº Proj Aprov."} (1 ou 0)	"#_Added Custom" = Table.AddColumn("#_Changed Type", "FLAG Aprovados", each {"#Nº Proj Aprov."});
Criar campo "FLAG Contratados"	"#_Added Conditional Column" = Table.AddColumn("#_Added Custom", "FLAG Contratados", each if ({("#Com TA2")=1}) and ({#"FLAG Aprovados"}=1)) then 1 else 0);
Criar campo "FLAG Pagamentos"	"#_Added Conditional Column1" = Table.AddColumn("#_Added Conditional Column", "FLAG Pagamentos", each if ({#"Com Pagamentos"} = 1) then 1 else 0);
Criar campo "FLAG execução"	"#_Added Custom1" = Table.AddColumn("#_Added Conditional Column1", "FLAG execução", each if ({#"[Exec]Fundo"} > 0 and {"#FLAG Aprovados"}=1) then 1 else 0);
Criar campo "FLAG desp certificada"	"#_Added Conditional Column9" = Table.AddColumn("#_Added Custom1", "FLAG desp certificada", each if ({#desp_t} <> 0 then 1 else 0);
Criar campo "FLAG DSTANL"	"#_Added Conditional Column11" = Table.AddColumn("#_Added Conditional Column9", "FLAG DSTANL", each if ({#"DST/ANL"} = "" then 0 else 1);
Alterar tipo de dados	"#_Changed Type" = Table.TransformColumnTypes("#_Added Conditional Column11", {"FLAG execução", Int64.Type}, {"FLAG Pagamentos", Int64.Type}, {"FLAG Contratados", Int64.Type}, {"FLAG Aprovados", Int64.Type}, {"FLAG desp certificada", Int64.Type}, {"FLAG DSTANL", Int64.Type});
Criar campo "VAL Inv Cand" = {"#Cand]Investimento"}/1000	"#_Added Custom2" = Table.AddColumn("#_Changed Type", "VAL Inv Cand", each {"#Cand]Investimento"}/1000);
Criar campo "VAL Inv Elegivel" = {"#Dec]Elegivel"}*[FLAG Aprovados]/1000	"#_Added Custom3" = Table.AddColumn("#_Added Custom2", "VAL Inv Elegivel", each {"#Dec]Elegivel"}*[FLAG Aprovados]/1000);
Criar campo "VAL Aprov Incentivo" = {"#Dec]Incentivo"}*[FLAG Aprovados]/1000	"#_Added Custom4" = Table.AddColumn("#_Added Custom3", "VAL Aprov Incentivo", each {"#Dec]Incentivo"}*[FLAG Aprovados]/1000);
Criar campo "VAL Cont Incentivo" = {"#Dec]Incentivo"}*[FLAG Contratados]/1000	"#_Added Custom5" = Table.AddColumn("#_Added Custom4", "VAL Cont Incentivo", each {"#Dec]Incentivo"}*[FLAG Contratados]/1000);
Criar campo "VAL Pagamentos" = {"#_Added Custom5", "VAL Pagamentos", each [Pagamento Total]*[FLAG Pagamentos]/1000	"#_Added Custom6" = Table.AddColumn("#_Added Custom5", "VAL Pagamentos", each [Pagamento Total]*[FLAG Pagamentos]/1000);
Criar campo "VAL Execução Fundo" = {"#Exec]Fundo"}*[FLAG execução]/1000	"#_Added Custom7" = Table.AddColumn("#_Added Custom6", "VAL Execução Fundo", each {"#Exec]Fundo"}*[FLAG execução]/1000);
Criar campo "VAL desp certificada" = {"#_Added Custom7", "VAL desp certificada", each [desp_t]*[FLAG desp certificada]/1000	"#_Added Custom11" = Table.AddColumn("#_Added Custom7", "VAL desp certificada", each [desp_t]*[FLAG desp certificada]/1000);
Alterar tipo de dados	"#_Changed Type1" = Table.TransformColumnTypes("#_Added Custom11", {"VAL Execução Fundo", type number}, {"VAL Pagamentos", type number}, {"VAL Cont Incentivo", type number}, {"VAL Aprov Incentivo", type number}, {"VAL Inv Elegivel", type number}, {"VAL Inv Cand", type number}, {"VAL desp certificada", type number});
Criar campo "INV_MILHÕES" = {"#_Changed Type1", "INV_MILHÕES", each [INV MIL]/1000	"#_Added Custom8" = Table.AddColumn("#_Changed Type1", "INV_MILHÕES", each [INV MIL]/1000);
Alterar tipo de dados	"#_Changed Type2" = Table.TransformColumnTypes("#_Added Custom8", {"INV_MILHÕES", type number});
Criar campos {"Concelho(s).1", "Concelho(s).2"} (primeiro concelho da lista, restantes)	"#_Split Column by Delimiter" = Table.SplitColumn("#_Changed Type2", "Concelho(s)", Splitter.SplitTextByEachDelimiter({"", QuoteStyle.Csv, false}, {"Concelho(s).1", "Concelho(s).2"});
Alterar tipo de dados	"#_Changed Type3" = Table.TransformColumnTypes("#_Split Column by Delimiter", {"Concelho(s).1", type text}, {"Concelho(s).2", type text});
Criar campo "Concelho Princ" = "Concelho(s).1"	"#_Renamed Columns1" = Table.RenameColumns("#_Changed Type3", {"Concelho(s).1", "Concelho Princ"});
Atualizar "Concelho Princ" = "ND" se NULL	"#_Replaced Value1" = Table.ReplaceValue("#_Renamed Columns1", null, "ND", Replacer.ReplaceValue, {"Concelho Princ"});
Criar campo "Concelho Sec" = "Concelho(s).2"	"#_Renamed Columns2" = Table.RenameColumns("#_Replaced Value1", {"Concelho(s).2", "Concelho Sec"});
Calculo Auxiliar	"#Duplicated Column1" = Table.DuplicateColumn("#_Renamed Columns2", "anomes", "anomes - Copy");
Calculo Auxiliar	"#Integer-Divided Column" = Table.TransformColumns("#Duplicated Column1", {"anomes - Copy", each Number.IntegerDivide(..., 100), Int64.Type});
Calculo Auxiliar	"#Added Custom1" = Table.AddColumn("#Integer-Divided Column", "anomes_seq", each {"#anomes - Copy"}*12+{"#anomes - Copy"}*100);
Calculo Auxiliar	"#Removed Columns" = Table.RemoveColumns("#Added Custom1", {"anomes - Copy"});
Criar campo "anomes_seq" com cálculos auxiliares anteriores	"#Renamed Columns2" = Table.RenameColumns("#Removed Columns", {"anomes_seq", "mes_sequential"});
Alterar tipo de dados	"#Changed Type1" = Table.TransformColumnTypes("#Renamed Columns2", {"mes_sequential", Int64.Type});
Criar campo "Prazo Análise" = [dec_dt_p]-[cnd_data]	"#Added Custom2" = Table.AddColumn("#Changed Type1", "Prazo Análise", each [dec_dt_p]-[cnd_data]);
Criar campo "Prazo Decisão" = [dec_data]-[cnd_data]	"#Added Custom3" = Table.AddColumn("#Added Custom2", "Prazo Decisão", each [dec_data]-[cnd_data]);
Criar campo "Prazo Contratação" = [dt_contrato]-[nt_dt_p]	"#Added Custom4" = Table.AddColumn("#Added Custom3", "Prazo Contratação", each [dt_contrato]-[nt_dt_p]);

Criar campo "Atraso Inicio Investimento" = [enc_dt_inicio]-[dec_dt_inicio]	##"Added Custom5" = Table.AddColumn(#"Added Custom4", "Atraso Inicio Investimento", each [enc_dt_inicio]-[dec_dt_inicio]),
Criar campo "Atraso Conclusão" = [enc_dt_fim]-[dec_dt_fim]	##"Added Custom6" = Table.AddColumn(#"Added Custom5", "Atraso Conclusão", each [enc_dt_fim]-[dec_dt_fim]),
Criar campo "Duração Candidatura" = [dec_dt_fim]-[dec_dt_inicio]	##"Added Custom7" = Table.AddColumn(#"Added Custom6", "Duração Candidatura", each [dec_dt_fim]-[dec_dt_inicio]),
Criar campo "Duração Projecto Encerrado" = [enc_dt_fim]-[enc_dt_inicio]	##"Added Custom8" = Table.AddColumn(#"Added Custom7", "Duração Projecto Encerrado", each [enc_dt_fim]-[enc_dt_inicio]),
Alterar tipo de dados	##"Changed Type4" = Table.TransformColumnTypes(#"Added Custom8",{{"Duração Projecto Encerrado", Int64.Type}, {"Duração Candidatura", Int64.Type}, {"Atraso Conclusão", Int64.Type}, {"Atraso Inicio Investimento", Int64.Type}, {"Prazo Contratação", Int64.Type}, {"Prazo Decisão", Int64.Type}, {"Prazo Análise", Int64.Type}}),
Criar campo "FLAG Encerrado" = se [enc_dt_fim] <> null then 1 else 0	##"Added Conditional Column7" = Table.AddColumn(#"Changed Type4", "FLAG Encerrado", each if [enc_dt_fim] <> null then 1 else 0),
Alterar tipo de dados	##"Changed Type5" = Table.TransformColumnTypes(#"Added Conditional Column7",{{"FLAG Encerrado", Int64.Type}}),
Join 'DIM Grupos'	##"Merged Queries2" = Table.NestedJoin(#"Changed Type5", {"Cod grupo"}, #"DIM Grupos", {"Cod grupo"}, "DIM Grupos", JoinKind.LeftOuter),
Selecionar campo 'Dotacao' de 'DIM Grupos'	##"Expanded DIM Grupos" = Table.ExpandTableColumn(#"Merged Queries2", "DIM Grupos", {"Grupo", "Dotacao"}, {"Grupo", "Dotacao"}),
Criar campo "Concelho"	##"Added Custom12" = Table.AddColumn(#"Expanded DIM Grupos", "Concelho", each if ([Concelho Sec] = null) and ([Concelho Princ] <> "") then [Concelho Princ] else "Multiconcelho ou ND"),
Join "READ tbc_pam_cae"	##"Merged Queries4" = Table.NestedJoin(#"Added Custom12", {"an_cae"}, #"DIMREAD tbc_pam_cae", {"codigo"}, "READ tbc_pam_cae", JoinKind.LeftOuter),
Selecionar campo 'transac' de 'READ tbc_pam_cae'	##"Expanded READ tbc_pam_cae1" = Table.ExpandTableColumn(#"Merged Queries4", "READ tbc_pam_cae", {"transac"}, {"Transacionavel"}),
Join "READ tbc_dimensao"	##"Merged Queries5" = Table.NestedJoin(#"Expanded READ tbc_pam_cae1", {"Dimensao"}, #"DIMREAD tbc_dimensao", {"cod_TipoPromotor"}, "READ tbc_dimensao", JoinKind.LeftOuter),
Selecionar 'dimensao_d' = "TipoPromotor"	##"Expanded READ tbc_dimensao" = Table.ExpandTableColumn(#"Merged Queries5", "READ tbc_dimensao", {"TipoPromotor"}, {"dimensao_d"}),
Atualizar 'dimensao_d' = "Projetos Conjuntos" se NULL	##"Replaced Value6" = Table.ReplaceValue(#"Expanded READ tbc_dimensao", null, "Projetos Conjuntos", Replacer.ReplaceValue, {"dimensao_d"}),
Criar campo 'tipo empresa' = (se [Dimensao] = 4 then "Grande Empresa" else "PME")	##"Added Conditional Column10" = Table.AddColumn(#"Replaced Value6", "tipo empresa", each if [Dimensao] = 4 then "Grande Empresa" else "PME"),
Atualizar 'ENEI' = "ND" se NULL	##"Replaced Value7" = Table.ReplaceValue(#"Added Conditional Column10", "", "ND", Replacer.ReplaceValue, {"ENEI"}),
Atualizar 'ENEI' = "ND" se VAZIO	##"Replaced Value8" = Table.ReplaceValue(#"Replaced Value7", null, "ND", Replacer.ReplaceValue, {"ENEI"}),
Atualizar 'Beneficiário' para maiúsculas	##"Uppercased Text" = Table.TransformColumns(#"Replaced Value8",{{"Beneficiário", Text.Upper, type text}}),
Atualizar 'Instrumento' substituindo 'Sistema de Apoio' por 'SA'	##"Replaced Value5" = Table.ReplaceValue(#"Uppercased Text", "Sistema de Apoio", "SA", Replacer.ReplaceText, {"instrumento"}),
Atualizar 'Instrumento'	##"Replaced Value9" = Table.ReplaceValue(#"Replaced Value5", "Regime de Apoio às", "RA", Replacer.ReplaceText, {"instrumento"}),
Atualizar 'Instrumento'	##"Replaced Value10" = Table.ReplaceValue(#"Replaced Value9", "Sistema de Incentivos à", "SI", Replacer.ReplaceText, {"instrumento"}),

