NOVA IMS Information Management School

MGI

Mestrado em Gestão de Informação
Master Program in Information Management

Influences of External Factors in Automotive Sales Forecasting

A Portuguese Case Study

Felipe Orestes Cuan Suárez Blanco

Dissertation presented as partial requirement for obtaining the Master's degree Information Management

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

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

INFLUENCES OF EXTERNAL FACTORS IN AUTOMOTIVE SALES FORECASTING

by

Felipe Orestes Cuan Suárez Blanco

Dissertation presented as partial requirement for obtaining the Master's degree in Information Management, with a specialization in Business Intelligence.

Advisor: Jorge Morais Mendes, PhD

February 2018

“It is often said there are two types of forecasts ... lucky or wrong!”

"Control" by Institute of Operations Management

ACKNOWLEDGEMENTS

Firstly, I would like to express my sincere gratitude to my advisor, Professor Jorge Mendes, for his continuous support during my research, for his patience, motivation and knowledge. His guidance helped me all the time while doing research and writing of this dissertation.

Besides my advisor, I would like as well to thank Eng. Ricardo de Almeida for inspiring me to write about this subject and also for sharing his comprehensive knowledge and wisdom. I hope to, one day, reach such level of expertise.

My sincere thanks also go to my family, my dearest Camila Steiner, José Mourinho, Fátima Bernardes, Rita Meneses, Francisco Lourenço and João Santos who, without their precious patience, knowledge and support, this dissertation would not have seen the light of day.

ABSTRACT

This study aims to contribute to a better understanding of the Portuguese automotive market, its behaviour and, additionally, attempt to create better sales prediction models, using linear regression models with ARIMA errors.

Through time series analysis, it will be studied how external factors such as socio-economic variables, consumer behaviour as well as market specific factors (e.g. gas prices, car rental services) affects the development of the automotive market. Furthermore, it will be tested if methodologies such as hierarchical forecasting on a bottom-up approach will bring better results compared to forecasting the market as a whole.

This understanding and forecasting techniques, if proven successful, can bring higher competitive advantage to companies in this sector, through more accurate planning and strategic management.

KEYWORDS

Car Sales, ARIMA, Hierarchical Forecasting, Regression with ARIMA Errors

INDEX

LIST OF FIGURES

LIST OF TABLES

1. INTRODUCTION

Success in market oriented companies depends not only on carrying effective management on daily basis but also on the ability to understand the intricate factors that shape the fabrics of the economy and, even better, to foresee and correctly assess changes in the future.

Due to socio-economic fluctuations and changing trends, this reality is present in many sectors of activity and is especially true in the automotive industry. The creation and development of the automotive vehicle has created one of the biggest changes in mankind history, allowing much faster exchange of goods and people, rocketing civilization into the modern age. However, as society develops further and further, influences such as globalization, car overproduction and the demand for more customized products (Holweg, 2008) leads to more competition among companies.

In Portugal, the automotive industry generated 18 billion euros in 2013 (ACAP, 2016a), through manufacturing, sales and exports which represents a total of 15% of the GDP. In 2016, auto loan reached 860 million euros just in the first five months of the year (Godinho, 2016).

Taking into consideration the Portuguese automotive market, which is comprised by 55 brands (ACAP, 2016b), the ability for these competing companies to generate reliable forecasts is of the utmost importance, which cannot be left only to intuitive guesses on market development. Nowadays, where advancements in data collection allow quality information to be much more accessible (AutoInforma, 2016), analysis and mathematical modelling of such data can create means to generate more accurate, and consequently, more reliable forecasts than ever before.

Having this in mind, the objective of this study is to allow further comprehension on what influences the development of the automotive market, which are the major factors in a macroeconomic scale and if there are other variables, such as social features or market specific characteristics that allows to correctly access and, if possible, predict short term future.

This type of short term forecasting, as of one or two months ahead, is especially important to companies/brands that rely on market share to fulfil their goals as this allows to plan ahead their operations, logistics and marketing schedules while having in account effects such as seasonality and vehicle lifecycles. Up to 95% of total yearly sales belong to companies/brands that are massified for the consumer. These companies tend to relate their objectives to market share presence and thus work intimately with the method explained before. The other 5% belong to luxury brands that target specially the aficionado and set their objectives by the total yearly sales, independent of time.

2. LITERATURE REVIEW

Time series analysis applied to sales forecast was first introduced by Lewandowsky (1974), concerning particularly the German automotive industry. This author, along the studies of Berkovec (1985), Dudenhöffer and Borscheid (2004) over automotive market modelling and sales forecasting established the foundations of time series analysis in this sector.

To accurately forecast, it is important to assess the data and determine if there are any consistent patterns, significant trends or seasonality and it is also important to find out if there are any outliers that can to be explained. While this analysis is done univariately, external factors - secondary time series parallel in time with the main time series – may have influence on the trend development, which adds another layer of complexity.

Concerning this last point, Brühl, Hülsmann, Friedrich and Reith (2009) analysed external factors of different natures, such as:

- Variables of global (national) economy
- Specific variables of the automotive market
- Variables of the consumer behaviour

Later on, this study was reviewed by Hülsmann, Borsheid, Friedrich and Reith (2012) enhancing the previous methodology.

In the Portuguese case, similar factors were analysed by Monteiro (2009) and Martins (2012). Monteiro concluded that unemployment and family income were factors that correlated with the market development while other factors such as tax load, end-of-life incentives or gas prices were not relevant. On the other hand, Martins concluded that both tax load and car-scrap bonus where relevant, opposing to unemployment rate which had little correlation with the market.

Faced with the disparity between these two studies and the lack of other literature regarding the influence of such factors in the Portuguese market, it would be interesting to determine which factors are indeed (if any) influencing the automotive market.

Also, to be noted is the influence of car rental services in the market development. These vehicles represent a significant portion of yearly sales – around 20% of the total market (ARAC, 2015), have a very stable seasonality and normally are subject to contracts of fixed volume of sales. At the moment, there are no studies available regarding this subject in Portugal.

Other matter with relevant interest is that in a sector regulated by the competition law, transparent and detailed data of each company/brand is available (AutoInforma, 2016). Such data includes daily and monthly vehicle sales by segment, category and model. Having such structured information, one can question if the analysis and modelling of such components will bring better results compared to analysing the car market as a whole.

Having such broad perspective is important for companies in this sector as it allows them to plan their stocks, both strategically and tactically for their near future and, at the same time, "keep an eye" on their competitors.

One way to analyse this information is using grouped/ hierarchical forecasting, where the various components (e.g. car model, segment, category time series) can influence the modelling at higher levels (e.g. vehicle market) (Hyndman, Ahmed, Athanasopoulos and Shang, 2010).

Classical approaches involve top-down or bottom-up methodologies, or a combination of both. The top-down method forecasts the aggregated time series and then disaggregates it based on historical proportions (Gross and Sohl, 1990). The bottom-up method involves forecasting each disaggregated series at the lowest levels and uses simple aggregation operations to obtain forecasts at higher levels.

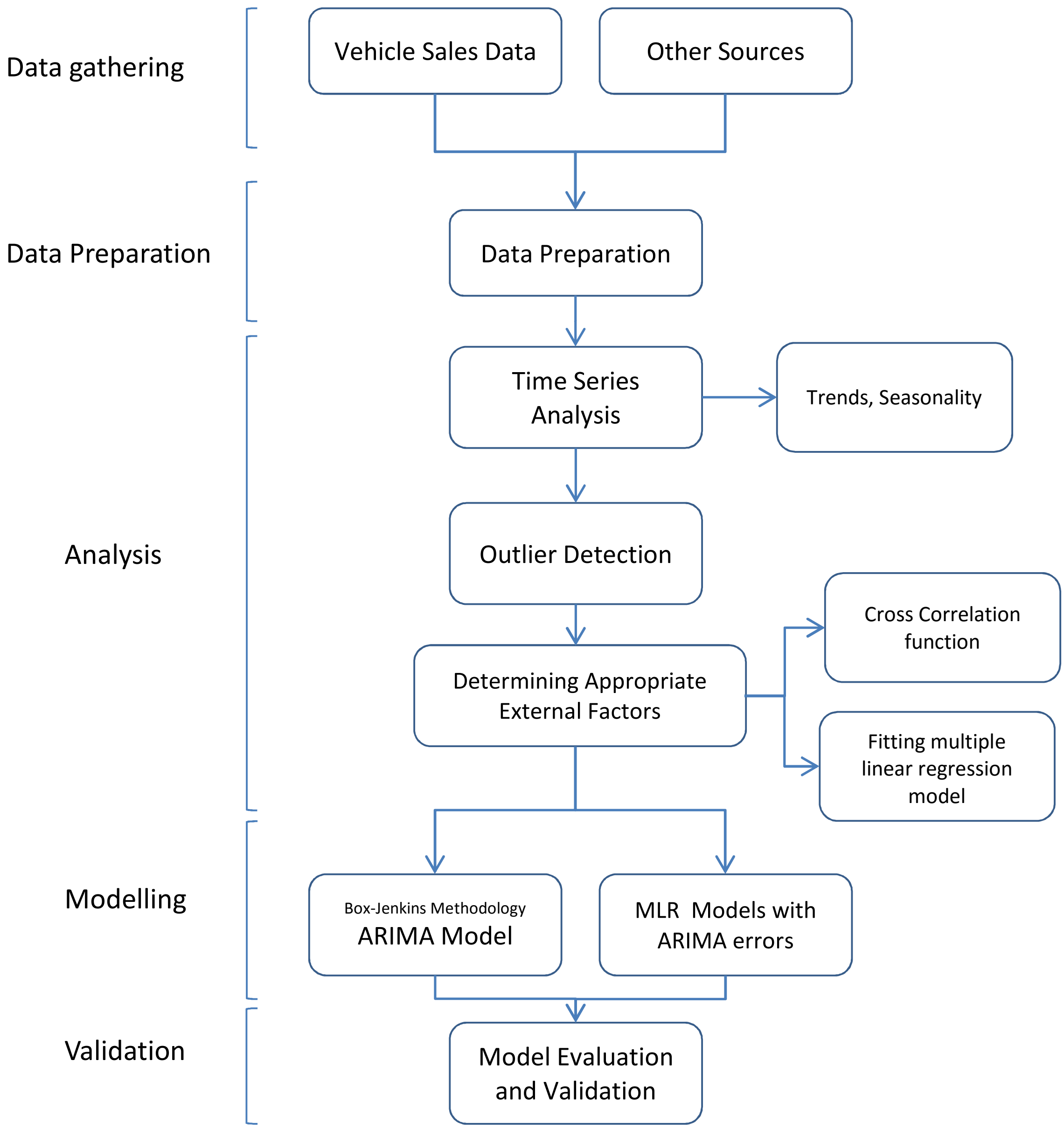
Having the broad perspective of the different aspects of the different aspects of the literature, this dissertation will focus on the following objectives:

O_1 – Is it possible to create more accurate predictions recurring to hierarchical forecasting compared with forecasting the market sales as a whole? If so, which type of segregation does bring better results: Car Rental or Vehicle Category?

O_2 – Which external factors influence the most when correlating with the different market subsets?

3. METHODS AND THEORETICAL FRAMEWORK

To follow the proposed objectives and further elaborate on this research, the subsequent schematic diagram describes, briefly, the major stages and the process workflow:



3.1. DATA GATHERING

3.1.1. Vehicle Sales

The main data series is comprised by new vehicle registrations in Portuguese territory. This information is gathered and managed by the Portuguese Automotive Trade Association (ACAP).

For this research was considered only light vehicles (i.e. gross weight less than 3500Kg and up to 9 seated places) in a time frame comprised between January 2006 to July 2016, with monthly and daily periodicity.

This data is divided by category, vehicle segment and brand. Vehicle category is defined by its purpose and usage, which can be for transporting passengers - Passenger Vehicle - or goods - Commercial Vehicle. Vehicle Segment adds another subdivision on passenger and commercial vehicles. It characterizes each vehicle according to its distance between wheel axis, engine power, price and bodywork.

3.1.2. Addicional Data and External Factors

To expand the complexity and enrich the mathematical models, additional data was gathered such as socio-economical indexes/indicators and specific automotive market data namely car rental statistics and oil prices.

Data regarding rental vehicles was gathered from the Portuguese Rent-a-Car Association (ARAC) which represents about 80% of all rental businesses. The available data is comprised by the number monthly vehicles delivered to rental companies, from January 2006 to July 2016, by brand.

Considering social-economic factors, this research will have in account factors previously used in studies such as Brühl et al (2009), Monteiro (2009), and Martins (2012). Additionally, other factors that were potentially interesting and relevant for this research were also considered.

A summary of the factors gathered are detailed in the table below:

Factor	Description	Type
R1	Unemployment rate of active population between 15 and 74 years old (%)	Social
R2	Private Consumption Indicator	Economic
R3	Economic Climate Indicator	Economic
R4	Expectations of unemployment evolution over the next 12 months	Social
R5	Expectations on economic climate over the next 12 months	Economic
R6	Consumer price index	Economic

Factor	Description	Type
R7	Crude Oil price - Brent (EUR)	Specific
R8	Nights spent at tourist accommodation establishments - Residents	Social
R9	Nights spent at tourist accommodation establishments - Non-Residents	Social
R10	Appreciation on economic climate of the last 12 months	Economic
R11	Economic Activity Indicator (%)	Economic
R12	Employment Index on Industry Sector	Social
R13	Employment Index on Commerce Sector	Social
R14	Employment Index on Service Sector	Social
R15	Savings Opportunity	Economic
R16	Expectations of equipment good acquisition in the next 12 months	Economic
R17	Expectations on sales prices over the next 3 months	Economic
R18	Number of weekday holidays per month	Calendar
R19	Number of workdays per month	Calendar
R20	Expectations on demand on services over the next 3 months	Economic
R21	Expectations on demand in commerce over the next 3 months	Economic
R22	Salary Index on Industry Sector	Economic
R23	Salary Index on Commerce Sector	Economic
R24	Salary Index on Service Sector	Economic

Table 1 – External Factors

3.1.3. Data Preparation

Even though all data is extracted from official and trustworthy sources, some data cannot be applied directly to the analysis as it needs validation and preparation. This task consists in verifying missing values and data discrepancies that may occur in the timeline.

There were no major issues aside from one or other missing value. To these specific cases it was applied values an interpolation function according to Stineman (Stineman, 1980).

Further transformations were required to adequate the data to their specific methodologies, such as taking in consideration for normalization, stationarity, heteroskedasticity and pre-whitening. These will be detailed in subsequent sections.

3.2. ANALYSIS

3.2.1. Time Series Analysis

An exploratory allows determining main characteristics of time series such as:

- Trends – measurements that, on average, tend to increase or decrease over time.
- Seasonality - regular repeating patterns of highs and lows related to calendar time such as seasons, quarters, months, and days of the week.

3.2.2. Choosing and Fitting Models

This research will focus in modelling using Box-Jenkins Methodology and Linear Regression with ARIMA errors. These methodologies were chosen due to their popularity in the industry and proven efficacy in forecasting.

3.2.2.1. Box-Jenkins Methodology

As proposed by George Box and Gwilym Jenkins (1970), this methodology involves identifying an appropriate an autoregressive integrated moving average (ARIMA) process that fits into the data, and then using the computed fitted model for forecasting.

The original Box-Jenkins modelling process involves an iterative three-stage process of model: selection, parameter estimation and model checking. Recent developments of this process, as in Makridakis, Wheelwright and Hyndman (1998) and add a preliminary stage of data preparation and a final stage of model application, as in forecasting and validation.

1. **Data preparation** involves cleaning and transformation of data, as mentioned in the previous section. Transformation of the data helps to stabilize parameters such as mean and variance in a time series and guarantee they do not change over time. Since stationarity is an assumption underlying ARIMA models, unit root tests allows checking this characteristic. Examples of unit root test used in this research are Augmented Dickey–Fuller test or Kwiatkowski–Phillips–Schmidt–Shin (KPSS) test.
   If a time series is non-stationary, transformations such as Box-Cox transformation and Differentiation are standard procedure to overcome such impediments.

2. **Model selection** in the Box-Jenkins methodology uses various tools such as analysing Auto-Correlation Function and Partial Auto-Correlation Function graphs on the transformed time series to try to identify potential ARIMA processes which might provide a good fit to the data.

3. **Parameter estimation** implies finding values for the model coefficients which provide the best fit to the data. To help in this process, there are measures that allow perceiving how well the model is fit. Examples of these are Akaike's Information Criterion and Bayesian Information Criterion.

4. **Model Checking** involves testing the assumptions of the model to identify any areas where the model is inadequate. This can be done by analysing the residuals of the model as it is expected to have the appearance of white noise. Again, unit root tests and ACF graphs allow checking these traits.
   If the model is found to be inadequate, it is necessary to go back to Step 2 and try to identify a better model.

5. **Forecasting**. Once the model has been selected, estimated and validated, it is then appropriate to use for forecasting. To measure forecast accuracy, last periods of the data should be left out of the modelling process to be used as testing data. In this study, accuracy of the model was done through cross-validation. This technique was also used by Hyndman (2014) and uses different training sets, each one containing one more observation than the previous one as shown in Figure 1 (The blue points are training sets, the red points are test sets and the grey points are ignored). The forecast accuracy measure is calculated on each test set and the results are averaged across all test set, adjusted to their different sizes. In this research, it was considered only a single forecast horizon, h, (next month) as it is the most crucial value to forecast. The accuracy measurement used was the mean absolute percentage error (MAPE).

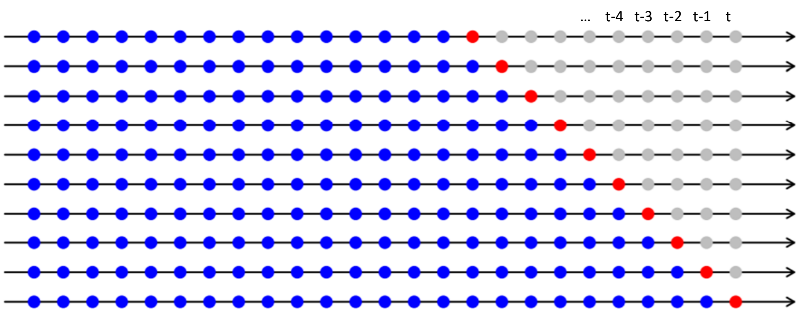


Figure 1 – Time series cross-validation.

3.2.2.2. Linear Regression models with ARIMA errors

To take in account other relevant information such as calendar effects or external factors, regression models can be applied in these situations. When using regression models with time series, it is common for the residuals to have a time series structure.

This procedure follows the subsequent steps:

1. Use ordinary least squares regression to estimate the model $y_t = \beta_0 + \beta_1 R1_t + \beta_2 Rk_t + \epsilon_t$ where y_t is a linear function of the k predictor variables ($R1_t$,..., Rk_t), and ϵ_t the residual series.
2. Examine the ARIMA structure (if any) of the sample residuals from the model in step 1.
3. If the residuals do have an ARIMA structure, use maximum likelihood to simultaneously estimate the regression model using ARIMA estimation for the residuals.
4. Examine the ARIMA structure (if any) of the sample residuals from the model in step 3. If white noise is present, then the model is complete. If not, continue to adjust the ARIMA model for the errors until the residuals are white noise.

An important consideration in estimating a regression with ARMA errors is that all variables in the model must first be stationary. A regression with ARIMA errors is the same as a regression model in differences with ARMA errors.

In this research, having so many factors/variables to dissect may cause an issue of overfitting the regression model. Thus, to determine and choose the best and most useful predictors it followed the best subset regression approach, where it is computed all possible models and displayed a smaller subset of variables (predictors) along with their adjusted R-squared and Mallows' C_p values. The idea is to understand which model presents the highest R^2 along with the lowest Mallows' C_p.

3.2.3. Relationships between the vehicle registries and external factors

To feed the correct variables/factors to the regression model it is necessary to assess which factors are indeed relevant for the regression model. One way to establish this relationship is to analyse the cross correlation function (CCF) between the main series (e.g. vehicle registries) and each of the exogenous factors as it identify which lags might be useful as predictors.

One major concern in the usage of CCF is that is affected by common trends that the two time series may have over time. A strategy to work around this difficulty is to use a pre-whitening procedure. This operation processes a time series to make it behave statistically as white noise thus reducing the presence of systematic information not relevant to prediction.

3.2.4. Hierarchical forecasting

As part of the proposed objectives for this study, it will be necessary to compare the accuracy of forecasts of the vehicle market as a whole, and the hierarchical methodology. The best model to apply will depend on the exploratory analysis and the relationships between main time series and the different external factors.

As for the hierarchical forecasting methodology, a multi-level hierarchy such as shown in figure 2 should be applied to the main time series, where level 0 is the complete aggregation series (i.e. total vehicle sales), level 1, the first level of disaggregation (i.e. Category: Passenger, Commercial, …) and level 2, the desegregation by brands in each of the correspondent categories.

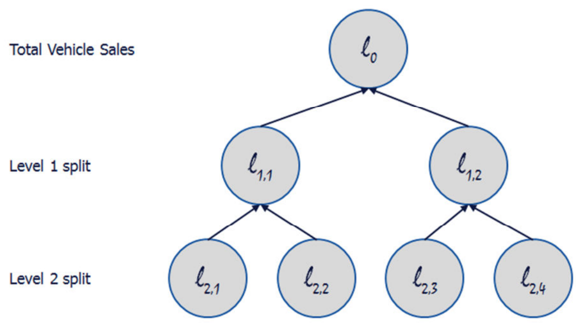


Figure 2 - Multi-level hierarchy

This study follows the theory of optimal combination forecast in hierarchical time series developed by Athanasopoulos et al (2007) and Hyndman et al (2010) which share similarities to approaches from Solomon and Weale (1991, 1993, 1996).

4. PORTUGUESE VEHICLE MARKET

4.1. CHARACTERIZATION

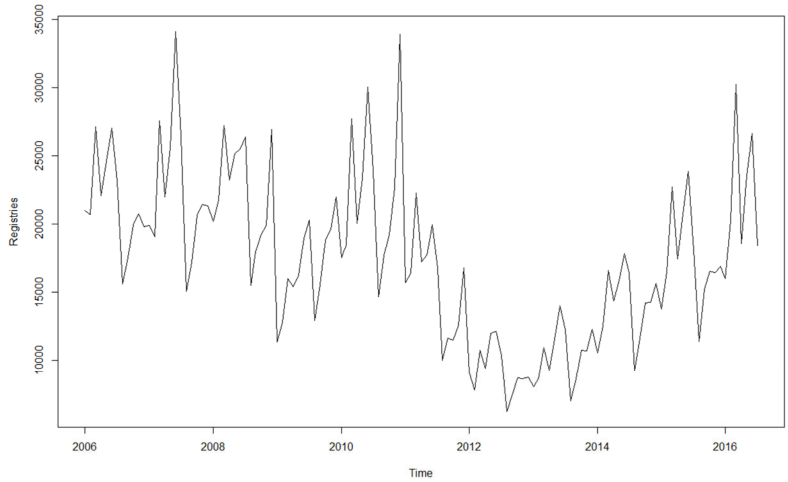


Figure 3 – Historical Time Series of total vehicle sales

The data gathered for this study revealed a very interesting period on automotive sales history as it can be observed in the figure above, between 2006 and 2009 there is a slight downward trend with occasional peaks before further reduction of sales. After 2011 and until 2013 the number of sales has plundered to numbers almost half of the previous years. Only after 2013 it is possible to see some recovery and a significant upward trend in the latter years.

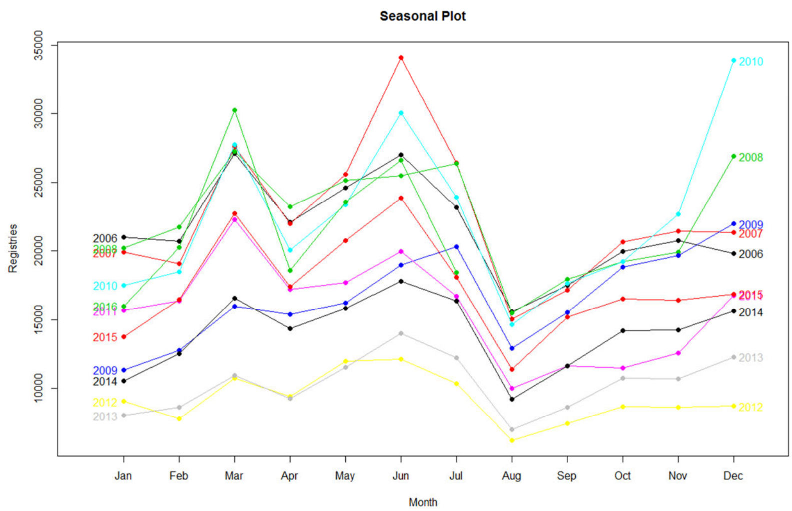


Figure 4 – Seasonality

Regarding seasonality, the series shows a pronounced seasonal behavior. Each quarter has upward trends, with different levels, being the second quarter the one with highest slope and the third with the lowest. This behavior is intimately related with the way sales objectives/goals of the major brands are specified.

Another way to analyze data is though heat maps, using different vehicle characteristics to extract further knowledge. For this analysis it was chosen to segregate by type, brand, segment, displacement and fuel type.

4.1.1. Passenger Market

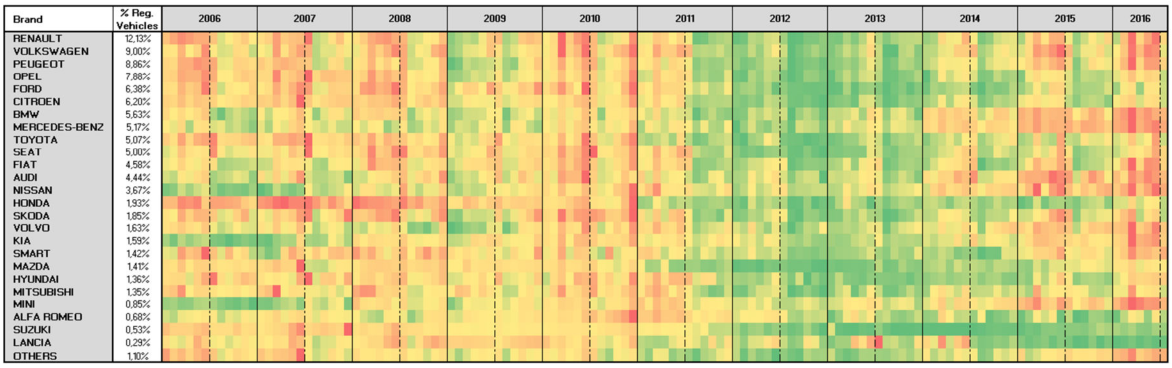


Figure 5 – Relative percentage of passenger vehicle sales, per brand (red is higher)

Comparing different passenger brands, it is interesting to see that before 2010 most of them had similar sales seasonality, reaching yearly highs around the middle of the year. However, after major sales drop during 2011-13 some brands haven’t recovered fully, with their sales dropping significantly. In recent years there has been some stabilization by the major players as the sales keep trending upward.

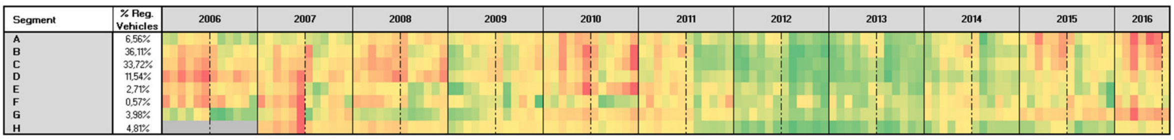


Figure 6 – Relative percentage of passenger vehicle sales, per segment
(**A** – Económico; **B** – Inferior; **C** - Médio-Inferior; **D** - Médio-Superior; **E** – Superior; **F** –Luxo; **G** – SUV; **H** – Monovolume)

Figure 7 – Relative percentage of passenger vehicle sales, per displacement

Figure 8 – Relative percentage of passenger vehicle sales, per fuel type

Comparing segment and displacement in the latter years, there is a tendency to choose vehicles with smaller displacement and lower segments. It should also be noted that there is an increasing preference of SUV's (Segment G) and electrical/hybrid vehicles are starting accelerate in sales, even though their market share is still very small.

4.1.2. Commercial Market

Figure 9 – Relative percentage of commercial vehicle sales, per brand

Figure 10 – Relative percentage of commercial vehicle sales, per segment

(**C** -Chassis-Cabina; **DTS** - Derivado Tecto Sobreelevado; **DV** - Derivado Van; **FM** - Furgão de Mercadorias; **FP** - Furgão Passageiros; **P4x2** - Pick-Up 4x2; **P4x4** - Pick-Up 4x4)

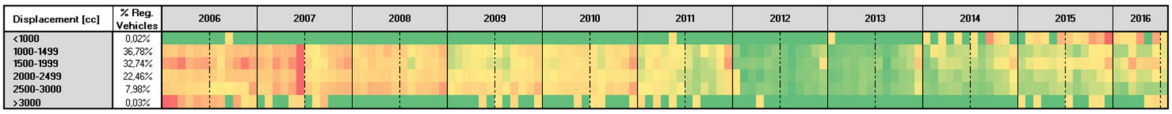


Figure 11 – Relative percentage of commercial vehicle sales, per displacement

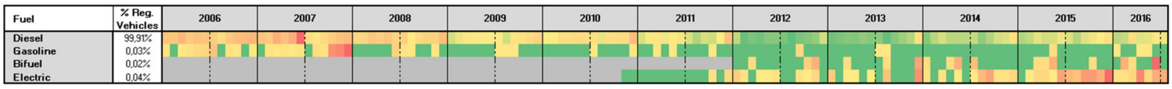


Figure 12 – Relative percentage of commercial vehicle sales, per fuel type

Observing commercial market, similar commentary as the passenger market can be concluded. In this case sales have diminished significantly after 2012 and smaller displacement engines are the preferred choice for newer vehicles. Also to be noted the downward trend for gasoline fuel source and an upward trend on bi-fuel and electrical, even though diesel is still prevailing as the main fuel source.

After this preliminary analysis and acquaintance of the market and the data, it is possible to start modelling.

4.2. MODELLING AND FORECASTING - TOTAL SALES

For the purpose of modelling the time series it will be analysed the time series of vehicle sales as a whole as first approximation. It is necessary to contemplate transformations to accommodate the initial conditions for ARIMA models. Such conditions/hypothesis consists in checking stationarity and heteroscedacity.

Following Guerrero (1993) methodology to test hetersocedacity, the time series will need a Box-Cox Transformation.

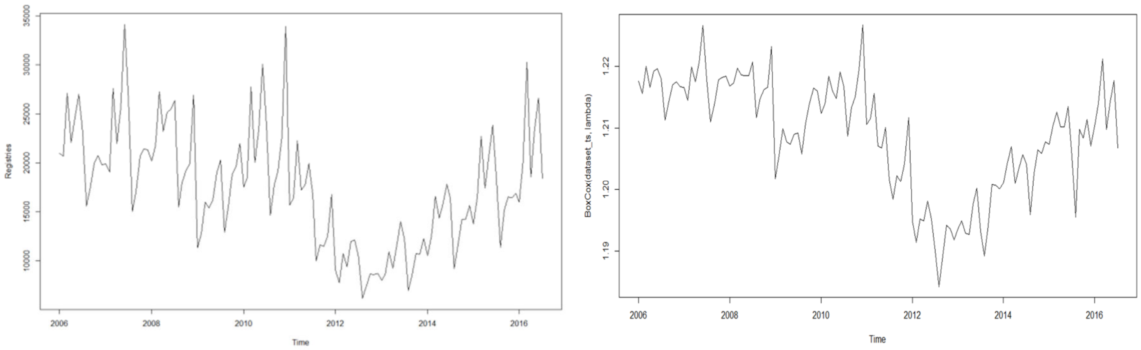


Figure 13 – Original series (left) vs Box-Cox Transformation (right)

To test for stationarity, it was used both Augmented Dickey-Fuller Test and KPSS Test for double validation.

	Null hypothesis	p-value
ADF Test	Series in non-stacionary	0,8432
KPSS Test	Series is stacionary	0,0100

Both tests confirm non-stacionarity and the need (and latter analysis the suficiency) for a first order differenciation.

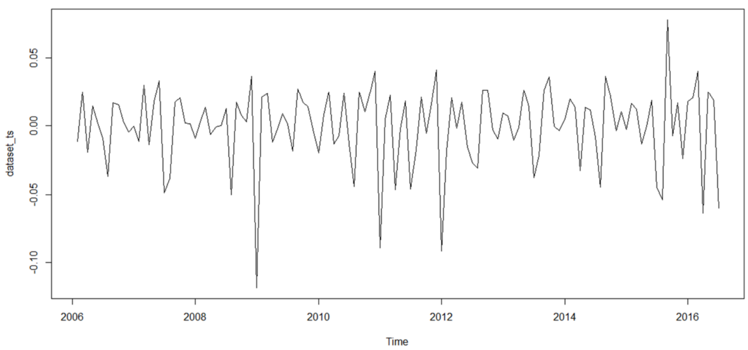


Figure 14 – Time series first order differentiation

Though these transformations, it is possible to evidence a few outliers, which will be considered later on in the modelling process.

Onward and on other modelling applications, these transformations were applied whenever it was deemed necessary.

4.2.1. ARIMA Estimation and Order Selection

To proceed with order estimation, auto-correlation graphs were created for the time series.

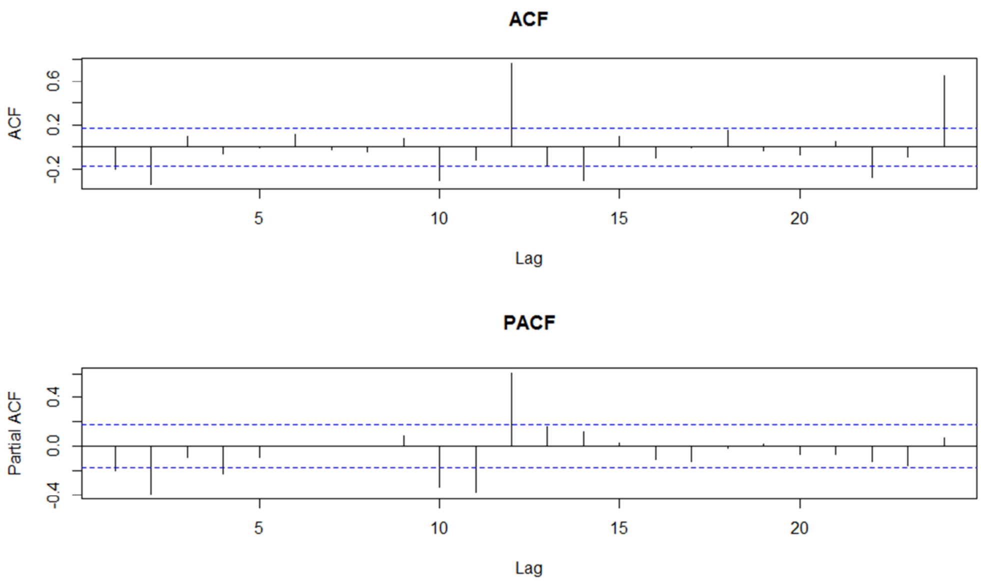


Figure 15 – Auto-Correlation Functions for total sales

The analysis of the auto-correlation functions shows evidence of high correlation of the second order on both ACF and PACF which in turn incline to a probable ARIMA model with (2,1,2) order as first approximation.

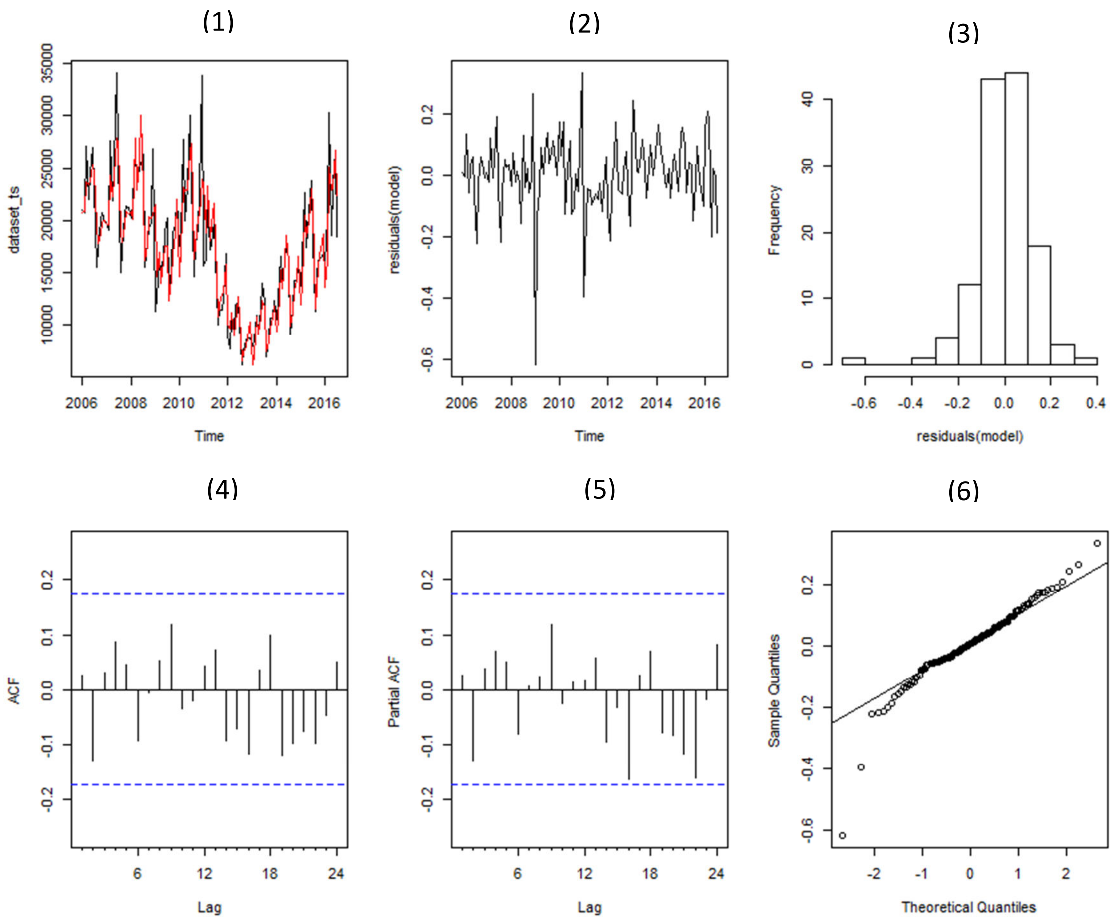
To determine the best model while having in account the preliminary ARIMA orders obtained through the auto-correlation graphs, it was calculated various models with permutations of the parameter orders based on the initial estimation. The process consisted on calculating models with parameters as permutations of orders in the neighborhood of the initial estimation.

The choice of the best model was done by choosing the top 10 models that minimize the Akaike Information Criteria (AIC) with general consensus of the other information criteria values (AICc and BIC) while also minimizing a 4-fold cross-validation error, having in account the values of unit root test for the residuals on different lags.

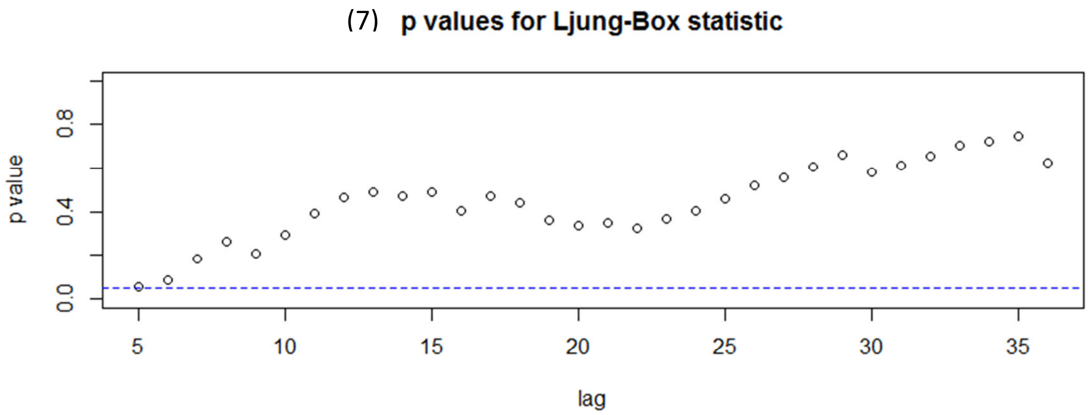
Non-Seasonal			Seasonal			Evaluation	Error (MAPE)	Diagnostic (Ljung-Box at lag)					
p	d	q	P	D	Q	AIC	k-fold Cross Validated	6	12	18	24	30	36
0	1	1	1	0	2	-132,423	**12,572%**	0,526	0,812	0,755	0,733	0,872	0,909
0	1	1	2	0	1	-132,632	**12,595%**	0,530	0,811	0,746	0,723	0,863	0,894
0	1	1	2	0	2	-132,873	**13,387%**	0,648	0,879	0,851	0,763	0,899	0,969
0	1	2	1	0	2	-131,850	**13,470%**	0,798	0,946	0,921	0,875	0,947	0,967
0	1	2	2	0	1	-131,942	**13,494%**	0,784	0,940	0,912	0,863	0,939	0,958
2	1	0	2	0	1	-132,633	**13,523%**	0,843	0,962	0,908	0,887	0,956	0,969
2	1	0	1	0	2	-132,459	**13,557%**	0,845	0,963	0,913	0,893	0,959	0,975
2	1	0	1	0	1	-132,159	**13,761%**	0,859	0,924	0,902	0,882	0,953	0,983
0	1	2	2	0	2	-132,103	**14,246%**	0,868	0,959	0,947	0,875	0,952	0,988
2	1	0	2	0	2	-132,539	**14,364%**	0,883	0,965	0,931	0,874	0,952	0,988

Table 2 – ARIMA models overview – Total Sales

After choosing the best adequate model though this process, extensive analysis was done to ensure the validity and correctness of the model. This includes (1) the predicted values, (2) the values of the residuals, (3) histogram of the residuals, (4, 5) ACF and PACF of the residuals, (6) Q-Q Plot and finally (7) unit root test – Ljung-Box statistics – over the residuals. In this latter test is important to remind the null hypothesis states that the data is independently distributed, i.e. there is no serial correlation in the observation.



(7)



As all the validation tests have passed, the ARIMA (0,1,1)(1,0,2) model is considered viable as a forecasting model.

4.3. Forecasting Total Sales with outliers

As it was mentioned before, there are outliers present on the window timeframe. Some are easily perceived, others are clearer after posterior transformations. To evidence the outliers in existence it was applied automatic outliers detection methodologies described by Cryer, J., Chan, K. (2012). Residuals are identified by fitting a loess curve for non-seasonal data and via a periodic STL decomposition for seasonal data. Residuals are labelled as outliers if they lie outside a specified quantile range of the residuals.

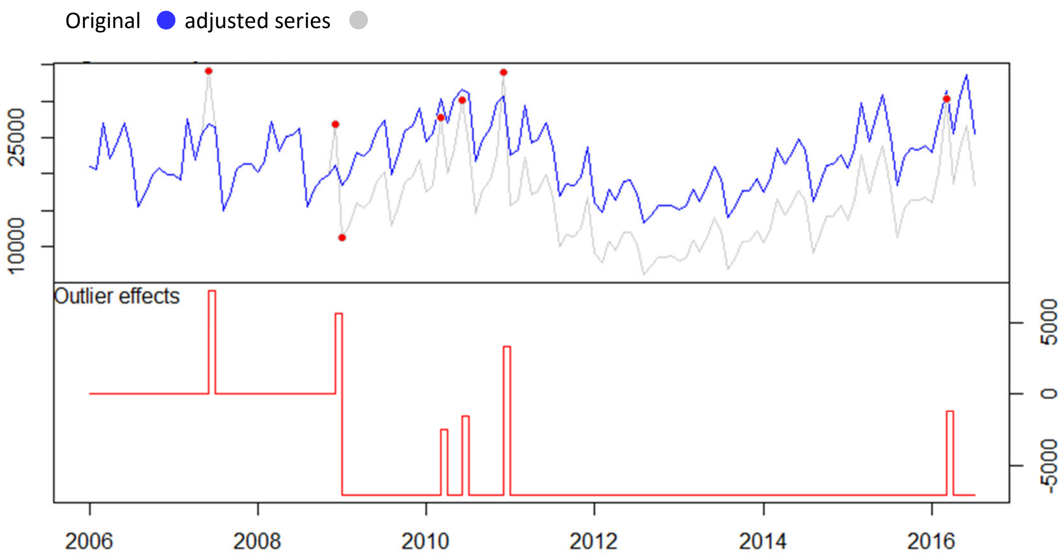


Figure 16 – Outlier detection - Total Sales

From the results of the outlier detection it is possible to summary the findings:

	Date	Type	Observations
AO18	Jun 2007	Additive	Increase over CO_2 emission tax + Increase on vehicle tax on July 2007
AO36	Dec 2008	Additive	Increase over CO_2 emission tax on January 2009
LS37	Jan 2009	Level Shift	Increase over CO_2 emission tax
AO51	Mar 2010	Additive	Tax incentive to end-of-life vehicle on April 2010
AO54	Jun 2010	Additive	VAT increase from 20% to 21% on July 2010
AO60	Dec 2010	Additive	VAT increase from 21% to 23% on Jan 2011
AO123	Mar 2016	Additive	Increase on vehicle tax on April 2016

Table 3 – Outliers – Total Sales

It is interesting to find that most outliers detected coincide with major economic events, mostly related with automotive incentives or taxation. There is a pattern that every time an increase of taxation or incentive is applied, the previous month has a substantial sales increase. Such events normally appear in the broadcast news and the general public takes action, sometimes anticipating their purchases. Also these events are many times exploited by automotive stands to further entice the sales.

To determine if it is possible to create a better model for forecasting in the light of this new information, a multiple linear regression was applied and analysed the significance of the different coefficients. Dummy variables were created to accommodate these outliers.

For this approach, the following regression equation was analysed and the correspondent coefficients calculated.

$$y_t = \beta_0 + \beta_1\, AO18_t + \beta_2\, AO36_t + \beta_3\, LS37_t + \beta_4\, AO51_t + \beta_5\, AO54_t + \beta_6\, AO60_t + \beta_7\, AO123_t + \epsilon_t \qquad (1)$$

	β	s_β	t value	Pr(>|t|)	Significance
β_0	21683,1	754,2	28,751	2,00E-16	***
AO18	12418,9	4461,8	2,783	6,26E-03	**
AO36	5208,9	4461,8	1,167	2,45E-01	
LS37	-6706,7	889,4	-7,541	1,02E-11	***
AO51	12742,6	4422,8	2,881	4,70E-03	**
AO54	15100,6	4422,8	3,414	8,75E-04	***
AO60	18904,6	4422,8	4,274	3,89E-05	***
AO123	15300,6	4422,8	3,460	7,52E-04	***

Table 4 – Regression coefficients – Total Sales with outliers

Significance codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

As per the table above, not all the outliers are significant for the regression equation. To maximize the model quality and optimize the number of significant regressors, an all subset model selection algorithm by Calcagno and Mazancourt (2010) was applied, which selects the best equations maximizing the significance of the included variables and information criteria.

Since the exhaustive screening of all variable permutations is extremely computational demanding, the best equation candidates were selected through genetic algorithm. Even though is a sub-optimal method, it can bring overall good results.

Aside from information criteria, another way to check the importance of each variable is with a graphical representation of each model and its importance over the model adjusted R^2 and Mallow's C_p. Each row represents a model and each dark colored block represents the inclusion of such variable in the model. The objective is the find a model that maximizes R^2 and minimizes C_p. A variable that is more persistent over the various models should be preferred over a less persistent one.

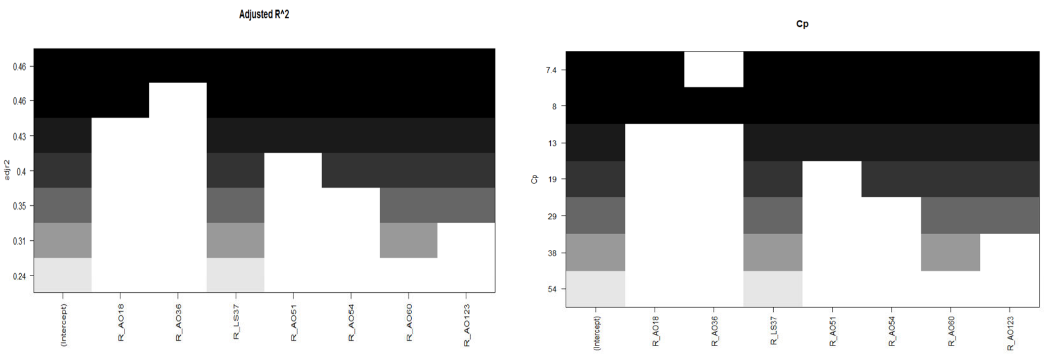


Figure 17 – Subset Model diagrams – Adjusted R^2 and Mallow's C_p for Total Sales

Having in account the various criteria for variable selection, the best overall variables are comprised in the follow regression equation.

$$yt = \beta_0 + \beta_1\, AO18_t + \beta_2\, AO51_t + \beta_3\, AO54_t + \beta_4\, AO60_t + \beta_5\, AO123_t + \epsilon_t \quad (2)$$

	β	s_β	t value	Pr(>|t|)	Significance
β_0	16943,4	487,3	34,793	2,00E-16	***
AO18	17159,0	5401,1	3,177	1,89E-03	**
AO51	10776,1	5401,5	1,995	4,83E-02	*
AO54	13134,5	5401,3	2,432	1,65E-02	*
AO60	16938,7	5401,7	3,136	2,15E-03	**
AO123	13334,5	5401,2	2,469	1,50E-02	*

Table 5 – Final Regression coefficients – Total Sales with outliers

Once the regression model has been selected, it is estimated the orders of the ARIMA parameters, this time regarding the residuals of the regression model, and subsequently calculated the various permutations around the preliminary parameters.

Non-Seasonal			Seasonal			Evaluation	Error (MAPE)	Diagnostic (Ljung-Box at lag)					
p	d	q	P	D	Q	AIC	k-fold Cross Validated	6	12	18	24	30	36
2	1	2	1	0	2	-154,855	**8,983%**	0,931	0,995	0,997	0,979	0,997	0,993
2	1	2	2	0	2	-161,409	**9,483%**	0,995	1,000	1,000	0,982	0,998	0,999
2	1	0	1	0	2	-155,007	**9,745%**	0,669	0,951	0,959	0,902	0,973	0,959
2	1	0	2	0	2	-161,786	**9,754%**	0,693	0,976	0,990	0,942	0,988	0,995
2	1	1	2	0	2	-159,982	**9,755%**	0,735	0,983	0,993	0,952	0,991	0,997
1	1	2	2	0	2	-158,458	**10,502%**	0,507	0,922	0,970	0,909	0,976	0,989
2	1	2	1	0	1	-155,297	**10,553%**	0,764	0,887	0,982	0,961	0,992	0,993
2	1	0	1	0	1	-155,045	**10,721%**	0,562	0,810	0,903	0,852	0,958	0,963
0	1	2	2	0	2	-159,893	**10,843%**	0,457	0,899	0,963	0,895	0,970	0,984
0	1	1	2	0	2	-160,112	**11,294%**	0,200	0,663	0,818	0,767	0,907	0,937

Table 6 – ARIMA models overview – Total Sales with outliers

This development has conducted into better accuracy results compared to the first ARIMA estimation, without outliers.

4.4. Forecasting Total Sales with External Factors

To have a better understanding on the influences of external factors over Portuguese vehicle sales, it was applied cross-correlation function to determine if there is any relationship at a lagged time and thus see it is a useful predictor of sales.

Since the direct application of CCF methodologies are affected by the time series structures and any “in common” trends that both series may have over time. It was needed to apply methodologies to minimize these disturbances such as pre-whitening.

As consequence of this process, the correlation values become much more subtle, without any high correlation values, which imply a much more careful observation and choice to not be confounded with not correlated factors.

4.4.1. Factors with some evidence of Correlation

From the various factors analysed, the ones that has shown higher correlation value (positive correlation in grades of red, negative correlation in grades of blue) are summarized below.

Factors	Lags												
	t-12	t-11	t-10	t-9	t-8	t-7	t-6	t-5	t-4	t-3	t-2	t-1	t
R1	0,06	-0,02	-0,09	-0,07	-0,05	-0,26	-0,21	-0,11	-0,22	-0,24	-0,27	-0,09	-0,1
R2	-0,22	-0,11	-0,05	-0,01	0,04	0,02	0,04	0,04	0,08	0,14	0,21	0,25	0,24
R3	-0,03	-0,03	0,06	0,11	0,12	0,04	-0,06	-0,02	0,01	0,17	0,18	0,22	0,27
R4	0,06	-0,07	-0,02	0,15	-0,08	-0,11	0,04	0,07	-0,02	-0,06	0,01	-0,17	-0,25
R5	-0,19	0	-0,11	-0,09	0,04	0,12	-0,01	-0,21	-0,19	0,06	-0,04	0,04	0,17
R6	-0,03	0,01	-0,01	-0,16	-0,06	0,06	-0,06	0,04	0,11	-0,01	0,11	0,06	-0,24
R7	0,04	-0,03	-0,11	-0,28	-0,05	-0,02	-0,09	-0,01	0,11	0,13	0,03	0,02	-0,08
R10	-0,09	0,09	-0,03	-0,02	0,01	0,05	-0,05	-0,21	-0,2	0,08	0,09	0,25	0,16
R11	-0,14	0,01	0,00	-0,01	-0,03	0,02	0,2	0,15	0,18	0,28	0,33	0,33	0,22
R15	-0,05	0,12	-0,08	-0,05	0,08	0,01	0,05	-0,19	-0,04	-0,01	0,06	-0,09	0,2
R16	-0,1	0,1	-0,08	0,07	0,18	-0,13	-0,04	-0,03	-0,05	-0,06	0,09	-0,03	0,04
R17	0,01	0,07	0,11	0,11	0,1	-0,07	-0,08	-0,01	-0,08	0,19	0,05	0,35	0,05
R18	0,11	-0,08	0,20	-0,11	-0,07	0,01	0,04	-0,12	0,10	0,06	-0,11	-0,05	0,21
R19	-0,03	0,13	-0,28	0,13	0,10	-0,13	0,06	0,10	-0,17	0,10	0,04	-0,21	0,18

R1 - Unemployment rate; **R2** - Private Consumption Indicator; **R3** - Economic Climate Indicator; **R4** - Expectations of unemployment over next 12 months; **R5** - Expectations on economic climate over the next 12 months; **R6** - Consumer price index; **R7** - Crude Oil price; **R10** - Appreciation on economic climate of the last 12 months; **R11** - Economic Activity Indicator; **R15** - Savings Opportunity; **R16** - Expectation of equipment good acquisition in the next 12 months; **R17** - Expectation on sales prices over the next 3 months; **R18** - Number of weekday holidays per month; **R19** - Number of workdays per month

Figure 18 – Cross Correlation table – Total Sales

Observing the different factors, it is interesting to note that most of these are factors that indicate the economic status of the country. The only social factor present in unemployment rate which is most correlated at lag 2. Also to be noted is how oil prices affect vehicle sales 9 months earlier.

4.4.2. Spurious Correlations

While analysing the CCF table, there are some factors featured as high positive correlative values such as R17 - expectations on sales prices. Intuitively and empirically it is not expected for automotive sales to go up when sales prices are going up, therefore this correlation will be considered as spurious.

4.5. Modelling and Forecasting using Regression with Arima Errors

From the factors that has shown the highest correlation, a regression model was created using them as predictors at respective lags.

	β	s_β	t value	Pr(>|t|)	Significance
β_0	68102,6	146393,4	0,465	6,43E-01	
R1_lag7	-7874,1	2630,0	-2,994	3,49E-03	**
R2_lag1	353,0	145,4	2,427	1,70E-02	*
R3_lag1	270,8	440,9	0,614	5,40E-01	
R4_lag0	4333,7	3579,3	1,211	2,29E-01	
R5_lag3	14651,3	13104,1	1,118	2,66E-01	
R6_lag0	9,2	3,1	2,986	3,58E-03	**
R7_lag9	-60,3	20,0	-3,015	3,28E-03	**
R10_lag1	12718,2	12995,9	0,979	3,30E-01	
R11_lag2	117405,1	216340,5	0,543	5,89E-01	
R15_lag0	28627,5	22800,2	1,256	2,12E-01	
R16_lag8	27633,1	24106,2	1,146	2,54E-01	
R18_lag0	701,8	432,2	1,624	1,08E-01	
R19_lag0	15,5	14,0	1,109	2,70E-01	
AO18	12758,0	3264,6	3,908	1,72E-04	***
AO36	3090,4	3348,5	0,923	3,58E-01	
LS37	-2000,8	1300,5	-1,538	1,27E-01	
AO51	8248,1	3281,9	2,513	1,36E-02	*
AO54	10260,9	3215,9	3,191	1,91E-03	**
AO60	14439,2	3351,8	4,308	3,95E-05	***
AO123	9351,3	3256,2	2,872	5,01E-03	**

Table 7 –Regression coefficients – Total Sales with outliers and external factors

From this initial regression equation, it is possible to see that many on these factors are not significant for the model.

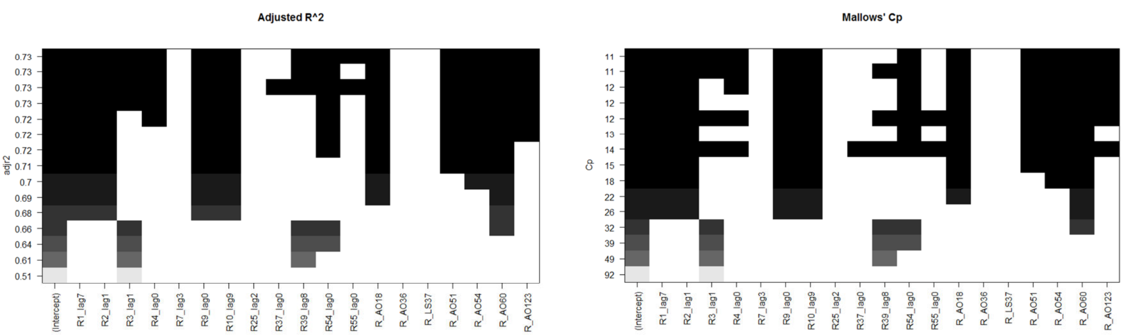


Figure 19 – Subset Model diagrams

	β	s_β	t value	Pr(>|t|)	Significance
β_0	40481,1	20169,4	2,007	4,98E-02	*
R1_lag7	-11496,2	2635,7	-4,362	3,05E-05	***
R3_lag1	1682,9	294,7	5,711	1,07E-07	***
R4_lag0	10708,4	2827,1	3,788	2,55E-04	***
R6_lag0	7,2	2,2	3,301	1,32E-03	**
R7_lag9	-82,4	19,0	-4,339	3,32E-05	***
R16_lag8	55650,2	21655,1	2,570	1,16E-02	*
AO18	10997,7	3052,4	3,603	4,84E-04	***
AO36	6078,1	3063,4	1,984	4,99E-02	*
AO51	8375,3	3041,1	2,754	6,95E-03	**
AO54	11594,0	3014,1	3,847	2,07E-04	***
AO60	16685,7	3032,8	5,502	2,71E-07	***
AO123	10191,6	3087,1	3,301	1,32E-03	**

Table 8 –Final Regression coefficients – Total Sales with outliers and external factors

$$y_t = \beta_0 + \beta_1\, R1_{t-7} + \beta_2\, R3_{t-1} + \beta_3\, R4_t + \beta_4\, R6_t + \beta_5\, R7_{t-9} + \beta_6\, R16_{t-8} + \beta_7 AO18_t + \beta_8 AO36_t + \beta_9 AO51_t + \beta_{10}\, AO54_t + \beta_{11}\, AO60_t + \beta_{12}\, AO123_t + \epsilon_t \qquad (3)$$

From this general model it is possible to observe the inverse relationship towards employment rates and oil prices; and the direct relationship on economic climate, employment expectations of the following year and expectation on equipment acquisition. There is also some effect on consumer prices but its effect is very small when compared with other factors. This may be due to the slow changing behaviour of this factor. The model above was used to create the forecasting model and was chosen the top 10 models that minimized the Akaike Information Criteria and sorted by lowest MAPE.

Non-Seasonal			Seasonal			Evaluation	Error (MAPE)	Diagnostic (Ljung-Box at lag)					
p	d	q	P	D	Q	AIC	k-fold Cross Validated	6	12	18	24	30	36
2	1	0	1	0	1	-177,438	**8,266%**	0,984	0,977	0,997	0,989	0,998	0,997
0	1	2	1	0	1	-178,981	**8,524%**	1,000	0,977	0,998	0,989	0,997	0,998
2	1	2	2	0	0	-180,877	**8,635%**	0,985	0,939	0,989	0,994	0,998	0,998
2	1	2	1	0	1	-182,228	**8,954%**	0,979	0,956	0,989	0,987	0,997	0,998
0	1	2	2	0	0	-177,792	**9,102%**	1,000	0,959	0,996	0,991	0,997	0,997
1	1	1	1	0	1	-185,248	**9,156%**	0,892	0,882	0,967	0,966	0,991	0,992
1	1	1	1	0	2	-183,272	**9,318%**	0,884	0,877	0,966	0,959	0,990	0,991
0	1	2	2	0	1	-177,935	**9,335%**	0,999	0,974	0,999	0,988	0,998	0,997
2	1	2	1	0	2	-180,372	**9,420%**	0,978	0,954	0,989	0,981	0,997	0,997
2	1	2	2	0	1	-181,336	**10,082%**	0,975	0,941	0,989	0,975	0,995	0,995

Table 9 – ARIMA models overview – Total Sales with outliers and external factors

Through the analysis of the previous table it can be observed that there is an increase in accuracy compared with both previous models.

5. HIERARCHICAL FORECASTING - VEHICLE MARKET AS A SUM OF PARTS

To test the initial hypothesis that forecasting multiple subsets of the data has better accuracy than forecasting the data as a whole; this chapter will focus on the analysis of different segments of the automotive market.

5.1. Forecasting Total Sales in a one Level Hierarchy: Car Rental + Free Market

The first level split to be analysed will be between Car Rental vehicle acquisition and the Free Market which is comprised by the general public and companies.

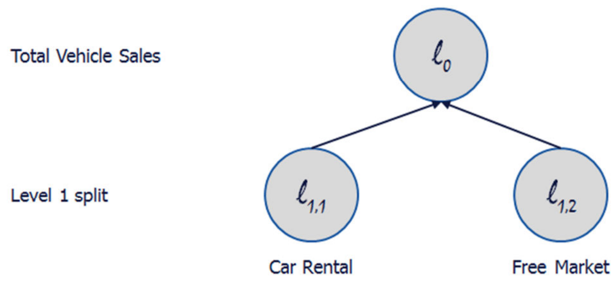


Figure 20 – One level hierarchical forecasting: Car Rental + Free Market

5.1.1. Forecasting Car Rental

As commented in the first chapter, Portuguese car rental market represents 20% of the total market and over the years has shown a rather coherent and predictable seasonality. Such behaviour can be favourable for forecasting models as its accuracy should be much higher.

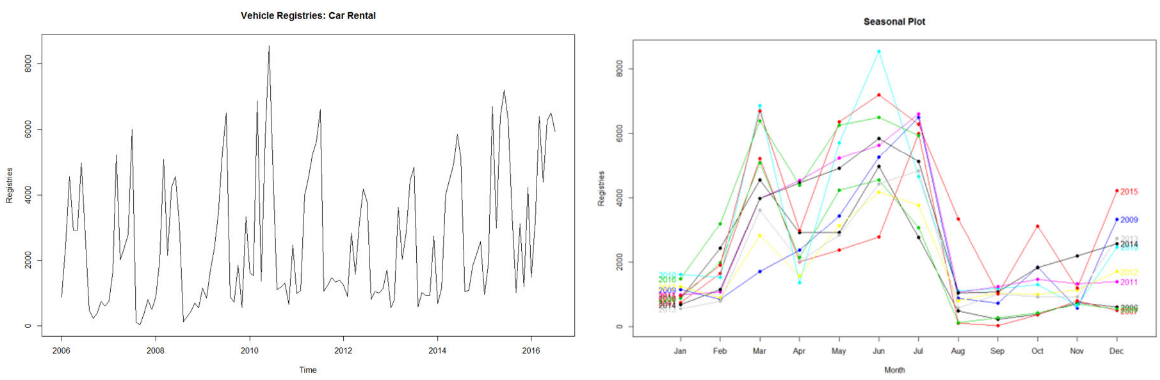


Figure 21 – Car Rental time series and seasonal plot

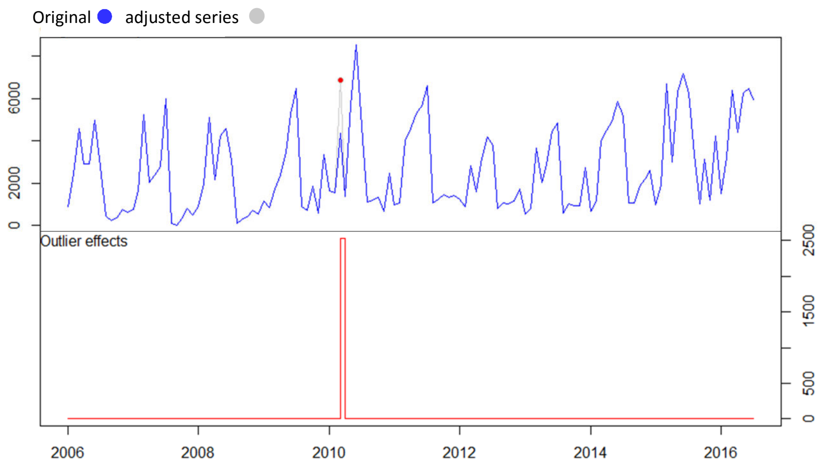


Figure 22 – Outlier detection – Car Rental

	Date	Type	Observations
AO51	Mar 2010	Additive	Tax incentive to end-of-life vehicle on April 2010

Table 10 – Outliers – Car Rental

Regarding major outliers, the only significant outlier detected coincides with the government announcement of tax incentives to end-to-life vehicles. This market normally is a high volume fleet acquisition with established contracts and dates so it is to be expected not to have any major and unexpected fluctuations as with the general public.

Best regression model with outliers

$$y_t = \beta_0 + \beta_1\, AO51_t + \epsilon_t \qquad (4)$$

	β	s_β	t value	Pr(>|t|)	Significance
β_0	2592,7	178,8	14,502	2,00E-16	***
AO51	4267,3	2014,8	2,118	3,62E-02	*

Table 11 – Regression model - Car Rental with outliers

Regarding to external factors, the ones that show higher correlation are in fact the ones related with tourism and demand on service activities.

External Factor Evaluation

Factors	Lags												
	t-12	t-11	t-10	t-9	t-8	t-7	t-6	t-5	t-4	t-3	t-2	t-1	t
R8	0,00	-0,14	-0,06	-0,17	-0,10	-0,07	-0,05	0,06	0,16	0,20	0,32	0,19	0,12
R9	0,08	-0,03	-0,02	-0,20	-0,22	-0,10	-0,17	0,00	0,14	0,23	0,28	0,15	0,19
R20	-0,08	-0,10	-0,07	0,00	-0,03	-0,02	0,00	-0,04	0,09	0,15	0,25	0,37	0,32

R8 - Nights spent at tourist accommodation establishments – Residents; **R9** - Nights spent at tourist accommodation establishments - Non-Residents; **R20** – Expectations on demand on services over the next 3 months

Table 12 – Cross Correlation table – Car Rental

It is interesting to observe that the tourism anticipates sales of car rental, especially resident tourists. Quite possibly it refers to sales of next touristic season and not direct reaction of the increase in tourism rate. More interesting is the expectation on the increase of activity in the service sector and vehicle acquisition may be a way of preparation for this expected increase. There is no evidence any social factors correlated with this data.

Best regression model with outliers and external factors

$$y_t = \beta_0 + \beta_1 R8_{t-2} + \beta_2 R20_{t-1} + \beta_3 AO51_t + \epsilon_t \qquad (5)$$

	β	s_β	t value	Pr(>|t|)	Significance
β_0	-19115,3	8821,1	-2,167	3,22E-02	*
R8_lag2	3479,7	1122,6	3,100	2,41E-03	**
R20_lag1	6726,9	540,2	12,453	2,00E-16	***
AO51	3410,3	1365,3	2,498	1,38E-02	*

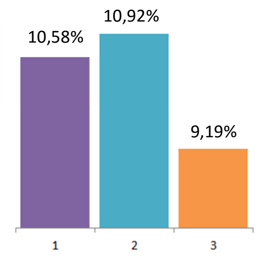
Table 13 – Regression model - Car Rental with outliers and external factors

It would be expected that non-resident tourism would have higher impact on car rental sales but is actually the residents that are significant.

Forecast Evaluation

Below are summarized the three best models of each criterion: standard ARIMA, regression model with ARIMA errors with outliers as predictors and regression model with ARIMA errors with outliers and external factors as predictors.

Type	ARIMA Parameters						Error (MAPE)	Diagnostic (Ljung-Box at lag)
	Non-Seasonal			Seasonal				
	p	d	q	P	D	Q		
(1) Standard	0	1	1	0	1	1		
(2) Reg. with outliers	0	1	1	1	1	2		
(3) Reg. with predictors and outliers	0	1	1	0	1	1		



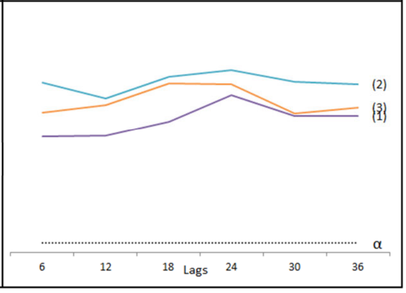


Table 14 – ARIMA models overview – Car Rental

From these three models the one with external factors show higher accuracy.

5.1.2. Forecasting Free Market

The free market, comprised the general public and enterprises, normally does not have scheduled planning as the car rental agreements so its trends tend to be much more unexpected and therefore more difficult to forecast.

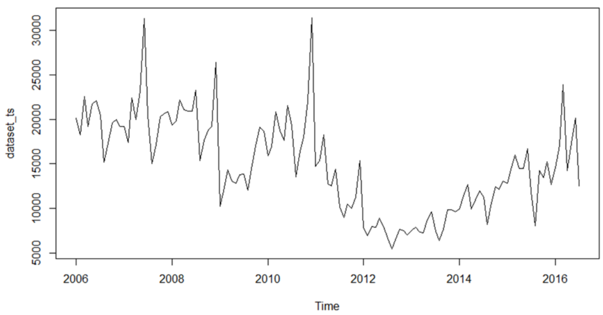


Figure 23 – Free Market time series

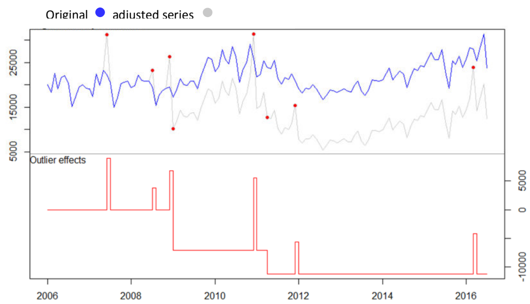


Figure 24 – Outlier detection – Free Market

	Date	Type	Observations
AO18	Jun 2007	Additive	Changes on CO2 emission tax + Increase on vehicle tax on July 2007
AO31	Jul 2008	Additive	VAT decrease from 21% to 20% on July 2008
AO36	Dec 2008	Additive	Alteration over CO2 emission tax on January 2009
LS37	Jan 2009	Level Shift	Alteration over CO2 emission tax
AO60	Dec 2010	Additive	VAT increase from 21% to 23% on Jan 2011
LS64	Apr 2011	Level Shift	Portugal announces arrangement for monetary aid from IMF
AO72	Dec 2011	Additive	Increase on vehicle tax on Jan 2012
AO123	Mar 2016	Additive	Increase on vehicle tax on April 2016

Table 15 – Outliers – Free Market

Through the outlier detection it is possible to observe that the general public is much more susceptible to announcements and changes of economic policies. As appeared in the outlier detection of the total sales data, every time there would be an increase on taxes, the previous month is subjected to an overgrowth in sales. It is also interesting in this series the capture of other outliers such as decrease of taxes and specially, in April 2011, the coincidence with the announcement of the monetary aid from IMF. This surely implies that any major economic news is indeed influencer of vehicle sales.

Best regression model with outliers

$$y_t = \beta 0 + \beta_1\, AO18_t + \beta_2\, AO36_t + \beta_3\, LS37_t + \beta_4\, AO60_t + \beta_5\, LS64_t + \beta_6\, AO123_t + \epsilon_t \qquad (6)$$

	β	s_β	t value	Pr(>|t|)	Significance
β_0	19710,5	511,0	38,574	2,00E-16	***
AO18	11613,5	3023,0	3,842	1,97E-04	***
AO36	6633,5	3023,0	2,194	3,01E-02	*
LS37	-3547,4	776,2	-4,570	1,19E-05	***
AO60	15250,9	3036,3	5,023	1,79E-06	***
LS64	-5271,0	694,5	-7,589	7,64E-12	***
AO123	13003,9	3003,1	4,330	3,11E-05	***

Table 16 – Regression model – Free Market with outliers

External Factor evaluation

Taking in consideration external factor contribution, the ones that have major effect over the free market as described on the following table.

	Lags												
Factors	t-12	t-11	t-10	t-9	t-8	t-7	t-6	t-5	t-4	t-3	t-2	t-1	t
R1	-0,07	-0,04	0,00	-0,06	-0,19	-0,32	-0,29	-0,22	-0,23	-0,22	-0,26	-0,19	-0,11
R2	-0,24	-0,18	-0,14	-0,09	-0,03	0,00	0,05	0,07	0,13	0,17	0,22	0,26	0,25
R11	-0,16	-0,02	-0,04	-0,02	0,02	0,08	0,19	0,19	0,23	0,31	0,35	0,32	0,28
R14	0,05	0,01	0,02	0,10	0,15	0,20	0,24	0,32	0,23	0,12	0,04	-0,01	0,11
R24	0,26	-0,05	-0,25	0,04	0,07	0,11	0,23	0,14	-0,18	-0,24	-0,02	0,20	0,32

R1 - Unemployment rate; **R2** - Private Consumption Indicator; **R11** - Economic Activity Indicator; **R14** - Employment Index on Service Sector; **R24** – Salary index on Service Sector

Table 17 – Cross Correlation table – Free Market

Best regression model with outliers and external factors

$$y_t = \beta_0 + \beta_1\, R2_{t-1} + \beta_2\, AO18_t + \beta_3\, AO60_t + \beta_4\, AO72_t + \beta_5\, AO123_t + \epsilon_t \qquad (7)$$

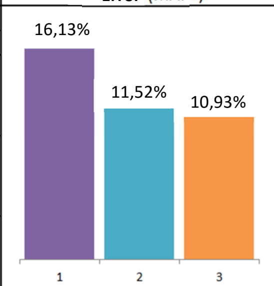
	β	s_β	t value	Pr(>|t|)	Significance
β_0	-15710,0	2443,0	-6,433	3,04E-09	***
R2_lag1	457,3	38,6	11,836	2,00E-16	***
AO18	12150,0	2492,0	4,878	3,51E-06	***
AO60	15750,0	2478,0	6,354	4,45E-09	***
AO72	7252,0	2540,0	2,855	5,12E-03	**
AO123	10250,0	2487,0	4,121	7,18E-05	***

Table 18 – Regression model – Free Market with outliers and external factors

For the model above, the factor that influences the most is the private consumption indicator in a positive relationship.

Forecast Evaluation

	ARIMA Parameters						Error (MAPE)	Diagnostic (Ljung-Box at lag)
Type	Non-Seasonal			Seasonal				
	p	d	q	P	D	Q		
(1) Standard	2	1	0	2	0	1		
(2) Reg. with outliers	2	1	1	1	0	2		
(3) Reg. with predictors and outliers	1	1	2	1	0	1		



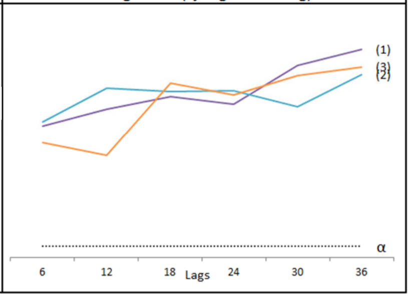


Table 19 – ARIMA models overview – Free Market

As expected, there is an increase of accuracy when adding outliers and, furthermore, when adding external factors, even though very slightly. Analyzing the error itself, it was expected to be higher that car rental. Also it has to be considered that the Free Market series is comprised by two subsets, passenger and commercial, with different behaviors.

5.1.3. Results

	Free Market (112101)		RAC (011011)		Total			
horizon	TrueVal	FcVal	TrueVal	FcVal	TrueVal	FcVal	Abs Error	Per Error
t-3	14181	14375	4391	4218	18572	18593	21	0,113%
t-2	17303	14615	6247	6313	23550	20928	2622	11,134%
t-1	20134	17745	6488	7461	26622	25206	1416	5,319%
t	12500	14371	5936	6932	18436	21303	2867	15,551%
							MAPE	8,029%

Table 20 – Forecast aggregation: Car Rental + Free Market

Aggregating the best forecasting models for car rental and free market into a one level hierarchy has a MAPE of 8,029% which is slightly better when compared with the original forecast on vehicle sales as a whole.

5.2. Forecasting Total Sales in a One Level Hierarchy: Passenger + Commercial

Another type of hierarchical split that can be analysed is, naturally, between passenger and commercial vehicles. Passenger vehicles are normally associated with the general public while commercial with business to business acquisitions and both are expected to have different trends.

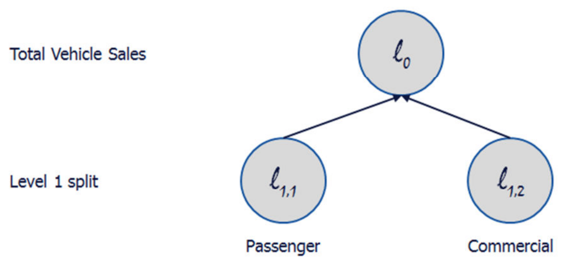


Figure 25 – One level hierarchical forecasting: Passenger + Commercial vehicles

5.2.1. Forecasting Passenger Market

Passenger marker is comprised by approximately 80% of total vehicle sales and, as mentioned before, targets the public in general. In the latter years there has been an upward trend in sales with consistent seasonal behaviour which will favour the forecast models.

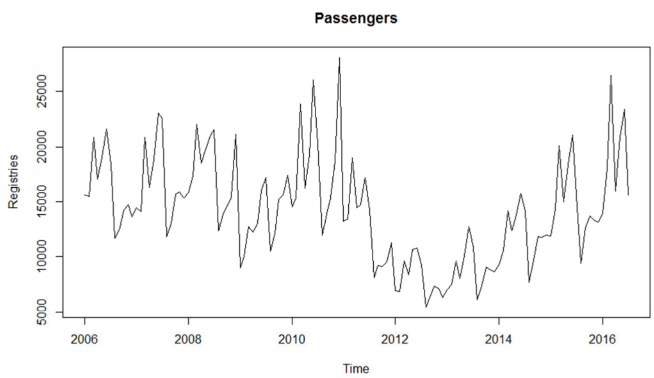


Figure 26 – Passenger Market time series

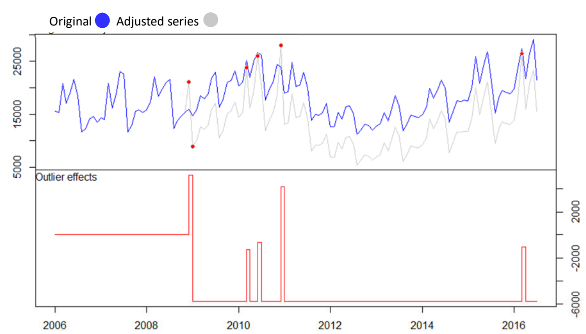


Figure 27 – Outlier detection – Passenger Market

	Date	Type	Observations
AO36	Dec 2008	Additive	Increase over CO2 emission tax on January 2009
LS37	Jan 2009	Level Shift	Increase over CO2 emission tax
AO51	Mar 2010	Additive	Tax incentive to end-of-life vehicle on April 2010
AO54	Jun 2010	Additive	VAT increase from 20% to 21% on July 2010
AO60	Dec 2010	Additive	VAT increase from 21% to 23% on Jan 2011
AO123	Mar 2016	Additive	Increase on vehicle tax on April 2016

Table 21 – Outliers – Passenger Market

From the outliers detection it is also clear that passenger vehicle sales are sensitive to vehicle taxation and incentives. These outliers are similar from the ones that were captured previously.

Best regression model with outliers

$$y_t = \beta_0 + \beta_1 \, LS37_t + \beta_2 \, AO51_t + \beta_3 \, AO54_t + \beta_4 \, AO60_t + \beta_5 \, AO123_t + \epsilon_t \qquad (8)$$

	β	s_β	t value	Pr(>|t|)	Significance
β_0	16941,9	641,2	26,421	2,00E-16	***
LS37	-4409,2	762,4	-5,783	5,86E-08	***
AO51	11303,4	3869,4	2,921	4,16E-03	**
AO54	13496,4	3869,4	3,488	6,79E-04	***
AO60	15585,4	3869,4	4,028	9,86E-05	***
AO123	13926,4	3869,4	3,599	4,64E-04	***

Table 22 – Regression model – Passenger Market with outliers

External Factor evaluation

	Lags												
Factors	t-12	t-11	t-10	t-9	t-8	t-7	t-6	t-5	t-4	t-3	t-2	t-1	t
R1	0,02	-0,04	-0,09	-0,07	-0,07	-0,31	-0,20	-0,15	-0,23	-0,24	-0,29	-0,12	-0,09
R3	-0,04	-0,02	0,06	0,10	0,10	0,02	-0,03	0,00	0,08	0,21	0,26	0,26	0,32
R18	0,11	-0,08	0,20	-0,11	-0,07	0,01	0,04	-0,12	0,10	0,06	-0,11	-0,05	0,21

R1 - Unemployment rate; **R3** - Economic Climate Indicator; **R18** – Number of weekday holidays per month

Table 23 – Cross Correlation table – Passenger Market

Best regression model with outliers and external factors

$$y_t = \beta_0 + \beta_1 \, R3_{t-1} + \beta_2 \, R18_t + \beta_3 \, AO51_t + \beta_4 \, AO54_t + \beta_5 \, AO60_t + \beta_6 \, AO123_t + \epsilon_t \qquad (9)$$

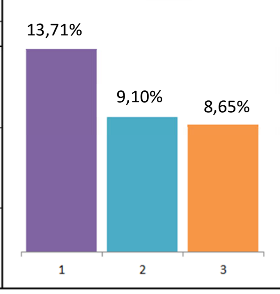
	β	$s_β$	t value	Pr(>|t|)	Significance
$β_0$	38818,1	6062,4	6,403	3,77E-09	***
R3_lag1	871,9	137,0	6,366	4,50E-09	***
R18_lag0	720,4	326,4	2,207	2,94E-02	*
AO51	8883,9	3053,0	2,910	4,37E-03	**
AO54	10380,9	3054,0	3,399	9,40E-04	***
AO60	12526,6	3053,6	4,102	7,83E-05	***
AO123	11447,1	3031,9	3,776	2,58E-04	***

Table 24 – Regression model – Passenger Market with outliers and external factors

From this model, it seems that even though unemployment rate is positively correlated with passenger vehicle sales, it was not significant for the regression model, which can be concluded that passenger market is mainly influenced by economic factors.

Forecast Evaluation

	ARIMA Parameters						Error (MAPE)	Diagnostic (Ljung-Box at lag)
	Non-Seasonal			Seasonal				
Type	p	d	q	P	D	Q		
(1) Standard	0	1	1	1	0	1		
(2) Reg. with outliers	2	1	0	1	0	2		
(3) Reg. with predictors and outliers	0	1	2	1	0	1		



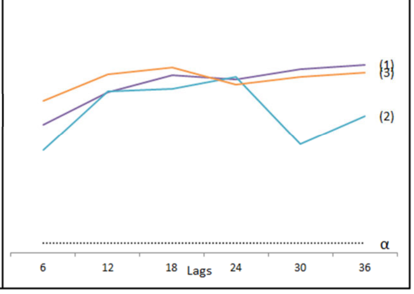


Table 25 – ARIMA models overview – Passenger Market

From this summary of models, similar deduction as the previous models can be extracted. Models with outliers and external factors tend to have higher accuracy than a standard ARIMA approach.

5.2.2. Forecasting Commercial Market

Commercial market complements the previous study, with 20% of the total market.

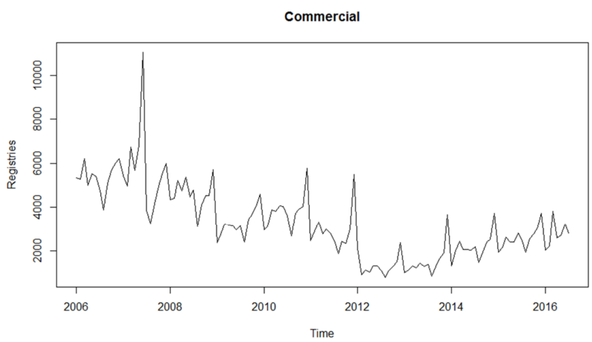


Figure 28 – Commercial Market time series

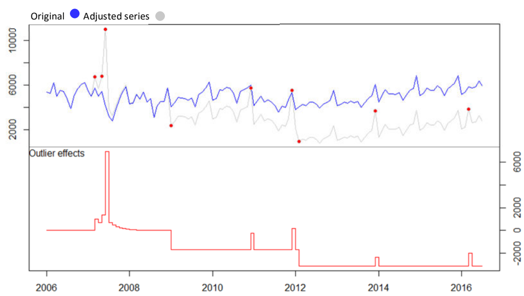


Figure 29 – Outlier detection – Commercial Market

	Date	Type	Observations
TC15	Mar 2007	Temp. Change	
TC17	May 2007	Temp. Change	Antecipation on tax increase
AO18	Jun 2007	Additive	Increase on CO_2 emission tax + Increase on vehicle tax on July 2007
LS37	Jan 2009	Level Shift	Alteration over CO_2 emission tax
AO60	Dec 2010	Additive	VAT increase from 21% to 23% on Jan 2011
AO72	Dec 2011	Additive	Increase on vehicle tax on Jan 2012
LS74	Fev 2012	Level Shift	
AO96	Dec 2013	Additive	
AO123	Mar 2016	Additive	Increase on vehicle tax on April 2016

Table 26 – Outliers – Commercial Market

Again, most of the outliers detected are influenced by economic events.

Best regression model with outliers

$$y_t = \beta_0 + \beta_1\, AO18_t + \beta_2\, LS37_t + \beta_3\, AO60_t + \beta_4\, AO72_t + \beta_5\, LS74_t + \beta_6\, AO123_t + \epsilon_t \qquad (10)$$

	β	s_β	t value	Pr(>|t|)	Significance
β_0	5063,8	130,4	38,831	2,00E-16	***
AO18	6016,2	782,4	7,689	4,54E-12	***
LS37	-1912,2	184,4	-10,369	2,00E-16	***
AO60	2611,4	782,4	3,338	1,13E-03	**
AO72	2358,4	782,4	3,014	3,14E-03	**
LS74	-1175,7	168,0	-6,997	1,61E-10	***
AO123	1842,1	778,7	2,366	1,96E-02	*

Table 27 – Regression model – Commercial Market with outliers

On the analysis of cross-correlations with the different external factors, it was found that there was no significant correlation with any factor which implies that the outlier analysis is rather enough or there were other external factors that were not taken in account. Having in account the accuracy of the prediction model using only outliers, it can be considered that this model is rather good and viable.

Forecast Evaluation

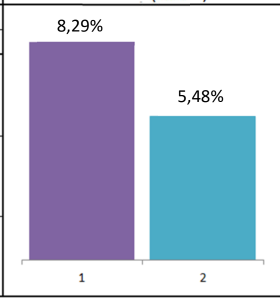
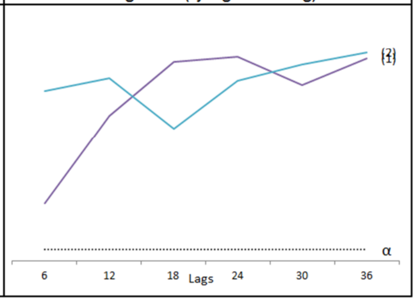
Type	ARIMA Parameters						Error (MAPE)	Diagnostic (Ljung-Box at lag)
	Non-Seasonal			Seasonal				
	p	d	q	P	D	Q		
(1) Standard	0	1	2	0	1	1		
(2) Reg. with outliers	1	1	0	0	1	2		
(3) Reg. with predictors and outliers								

Table 28 – ARIMA models overview – Commercial Market

Having a more predictable behavior, the forecasting accuracy has greatly increased when compared with the previous subsets.

5.2.3. Results

	Passenger (012101)		Commercial (110012)		Total			
horizon	TrueVal	FcVal	TrueVal	FcVal	TrueVal	FcVal	Abs Error	Per Error
t-3	15978	17238	2594	2441	18572	19679	1107	5,961%
t-2	20851	19530	2699	2699	23550	22229	1321	5,609%
t-1	23369	23560	3253	2821	26622	26381	241	0,905%
t	15632	18690	2804	2881	18436	21571	3135	17,005%
							MAPE	7,370%

Table 29 – Forecast aggregation: Passenger + Commercial Market

Compared with the previous hierarchical aggregation, this one provides a better average accuracy. Partially due to Free Market encompassing two different behaviors, which in this model are separated. In the light of this information it is compelling to determine if car rental separated from the passenger and commercial market would bring even better results.

5.3. Forecasting Total Sales in a two Level Hierarchy

To further enhance the study, it will be analyzed if a two level hierarchy will bring significant accuracy improvement when compared with the previous models. For the second level split it will be considered between passenger car rental and passenger free market as an attempt of bringing together the knowledge acquired before. From the commercial category, there is no significant sales on car rental, thus it will be considered the same model as previously.

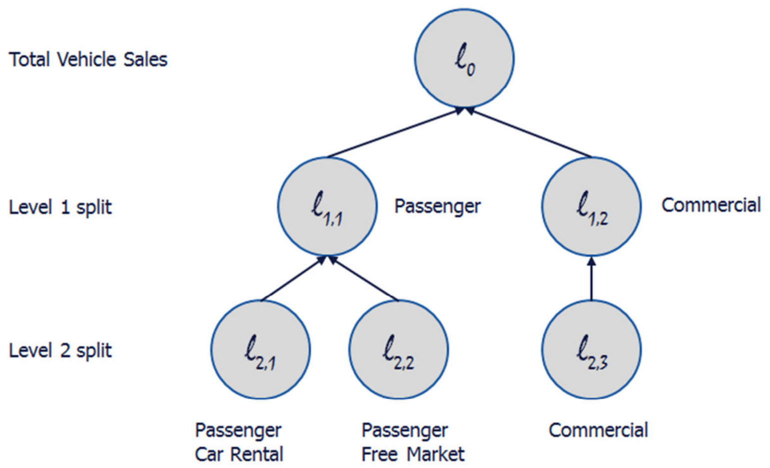


Figure 30 – Two levels hierarchical forecasting

5.3.1. Forecasting Passenger Car Rental

From the car rental standpoint, passenger vehicles comprise the majority of the sales, so it is expected for this subset to behave in terms of forecasting to the car rental market.

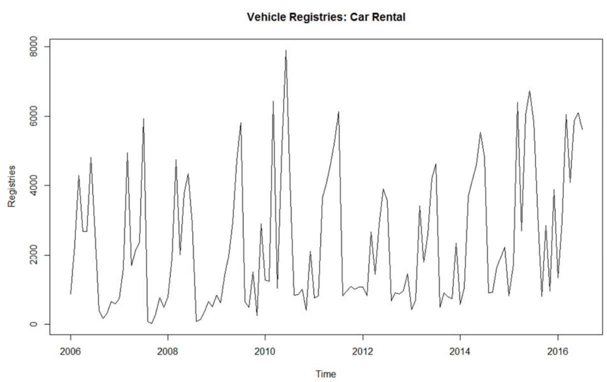


Figure 31 – Passenger Car Rental time series

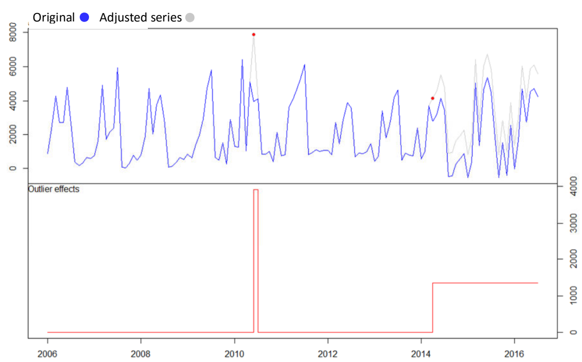


Figure 32 – Outlier detection – Passenger Car Rental

	Date	Type	Observations
AO54	Jun 2010	Additive	VAT increase from 20% to 21% on July 2010
LS100	Apr 2014	Level Shift	

Table 30 – Outliers – Passenger Car Rental

Best regression model with outliers

$$y_t = \beta_0 + \beta_1\ AO54_{\ t} + \beta_2\ LS100_t + \epsilon_t \qquad (11)$$

	β	s_β	t value	Pr(>|t|)	Significance
β_0	1979,3	178,3	11,101	2,00E-16	***
AO54	5911,7	1774,1	3,332	1,14E-03	**
LS100	1613,8	378,2	4,267	3,90E-05	***

Table 31 – Regression model – Passenger Car Rental with outliers

As the data is subset further and further, there are less outliers detected that have significant impact over the series.

External Factor evaluation

	Lags												
Factors	t-12	t-11	t-10	t-9	t-8	t-7	t-6	t-5	t-4	t-3	t-2	t-1	t
R8	0,01	-0,14	-0,06	-0,18	-0,10	-0,07	-0,04	0,06	0,15	0,20	0,32	0,20	0,12

R8 – Nights spent at tourist accommodation establishments - Residents

Table 32 – Cross Correlation table – Passenger Car Rental

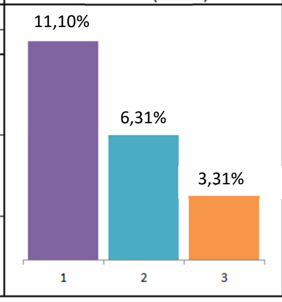
Best regression model with outliers and external factors

$$y_t = \beta_0 + \beta_1\ R8_{t-2} + \beta_2\ AO54 + \beta_3\ LS100_t + \epsilon_t \qquad (12)$$

	β	s_β	t value	Pr(>|t|)	Significance
β_0	150,9	13,4	11,287	2,00E-16	***
R8_lag2	-0,1	0,0	-5,460	2,58E-07	***
AO54	97,0	34,7	2,794	6,05E-03	**
LS100	35,9	7,4	4,837	3,92E-06	***

Table 33 – Regression model – Passenger Car Rental with outliers and external factors

Type	ARIMA Parameters: Non-Seasonal p	d	q	Seasonal P	D	Q
(1) Standard	0	1	1	0	1	1
(2) Reg. with outliers	2	1	2	2	1	1
(3) Reg. with predictors and outliers	2	1	1	1	1	1



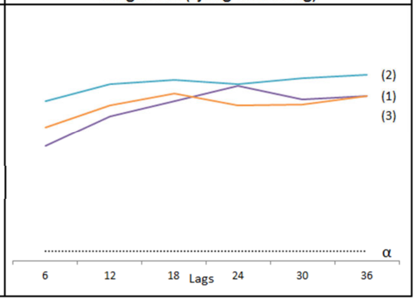


Table 34 – ARIMA models overview – Passenger Car Rental

As expected, the forecast accuracy is very high when forecasting with the regression aggregated with outliers and external factors.

5.3.2. Forecasting Passenger Free Market

The Passenger free market has the biggest weight in this split, in terms of volume, and also has the less consistent behavior, so it would be expected to have lower prediction accuracy.

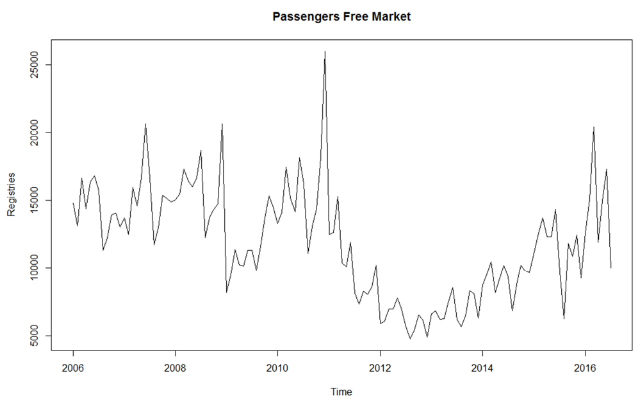


Figure 33 – Passenger Free Market time series

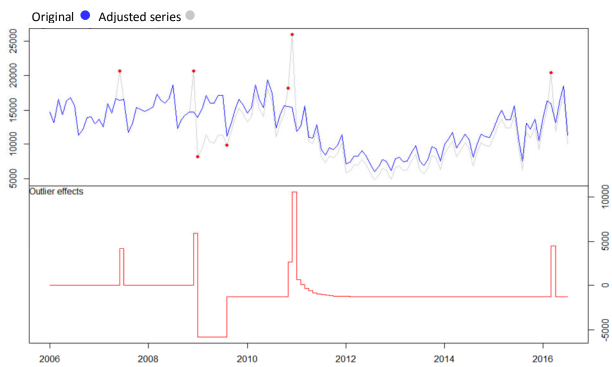


Figure 34 – Outlier detection – Passenger Free Market

	Date	Type	Observations
AO18	Jun 2007	Additive	Increase on CO_2 emission tax + Increase on vehicle tax on July 2007
AO36	Dec 2008	Additive	Increase over CO_2 emission tax on January 2009
LS37	Jan 2009	Level Shift	Increase over CO_2 emission tax
LS44	Aug 2009	Level Shift	
TC59	Nov 2011	Temp. Change	Antecipation on VAT increase
AO60	Dec 2011	Additive	VAT increase from 21% to 23% on Jan 2011
AO123	Mar 2016	Additive	Changes on vehicle tax on April 2016

Table 35 – Outliers – Passenger Free Market

Similar to the previous models, most of the outliers are regarded for vehicle taxation.

Best regression model with outliers

$$y_t = \beta_0 + \beta_2\ AO18_t + \beta_3\ AO36_t + \beta_3\ LS37_t + \beta_3\ TC59_t + \beta_3\ AO60_t + \beta_3\ AO123_t + \epsilon_t \qquad (13)$$

	β	s_β	t value	Pr(>|t|)	Significance
β_0	14792,9	495,9	29,830	2,00E-16	***
AO18	5865,1	2933,8	1,999	4,79E-02	*
AO36	5865,1	2933,8	1,999	4,79E-02	*
LS37	-4513,3	1200,2	-3,761	2,64E-04	***
TC59	7775,5	2456,0	3,166	1,96E-03	**
AO60	10563,1	3339,8	3,163	1,98E-03	**
AO123	10427,9	2910,3	3,583	4,93E-04	***

Table 36 – Regression model – Passenger Free Market with outliers

External Factor evaluation

	Lags												
Factors	t-12	t-11	t-10	t-9	t-8	t-7	t-6	t-5	t-4	t-3	t-2	t-1	t
R1	-0,05	0,02	0,07	0,01	-0,11	-0,27	-0,24	-0,21	-0,23	-0,26	-0,27	-0,21	-0,12
R13	-0,01	-0,19	-0,04	0,20	0,22	0,14	0,23	0,33	0,07	-0,03	0,18	0,22	0,08
R14	0,02	-0,06	-0,06	0,02	0,11	0,17	0,23	0,36	0,28	0,16	0,07	0,02	0,11

R1 - Unemployment rate; **R13** – Employment Index on Commerce Sector; **R14** - Employment Index on Service Sector

Table 37 – Cross Correlation table – Passenger Free Market

Best regression model with outliers and external factors

$$y_t = \beta_0 + \beta_1\ R14_{t-5} + \beta_2\ TC59_t + \epsilon_t \qquad (14)$$

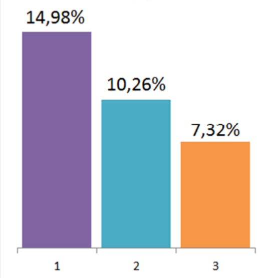
	β	s_β	t value	Pr(>|t|)	Significance
β_0	5900,6	1876,1	3,145	1,31E-02	**
R14_lag5	103,3	27,3	3,783	2,47E-04	***
TC59	12573,6	2228,1	5,643	1,20E-07	***

Table 38 – Regression model – Passenger Free Market with outliers and external factors

Particularly interesting in this model is the fact that is based mostly on social factors contrary of the most models found up until now which are based on economic factors. The unemployment rate is inversely correlated with vehicle sales as expected and it is positively correlated with employment index. However, both unemployment rate and employment index are correlated between themselves as employment index is a subset of the unemployment rate. Thus on the calculations were considered only the index. Other factors did not appear to have major correlation or significance over the regression model.

Type	ARIMA Parameters					
	Non-Seasonal			Seasonal		
	p	d	q	P	D	Q
(1) Standard	0	1	1	2	0	2
(2) Reg. with outliers	2	1	1	2	0	2
(3) Reg. w/ predictors and outliers	2	1	2	2	0	2

Error (MAPE)



Diagnostic (Ljung-Box at lag)

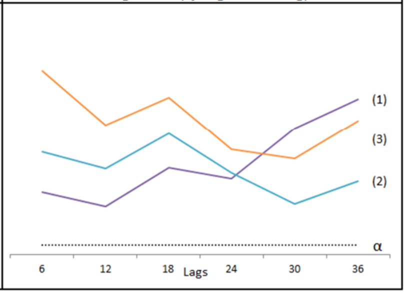


Table 39 – ARIMA models overview – Passenger Free Market

The quality of the results go accordingly with the previous models which states that the error measure is smaller on regression models with outliers and external factors.

5.3.3. Total Market Results

	Pass Car Retal (211111)		Pass Free Market (212202)		Commercial (110012)		Total			
horizon	TrueVal	FcVal	TrueVal	FcVal	TrueVal	FcVal	TrueVal	FcVal	Abs Error	Per Error
t-3	4076	3869	11902	12623	2594	2441	18572	18933	361	1,944%
t-2	5886	5473	14965	13598	2699	2699	23550	21770	1780	7,558%
t-1	6095	6276	17274	17351	3253	2821	26622	26448	174	0,654%
t	5597	6214	10035	11413	2804	2881	18436	20508	2072	11,239%
									MAPE	5,349%

Table 40 – Forecast aggregation:
Passenger Car Rental + Passenger Free Market + Commercial Market

This hierarchy has shown the best average results when compared with other models. Also to be noted is the fact that needs less external factors to provide good results. One plausible explanation for this happening is the fact that by subsetting the data, there is less variability of the different series, allowing better prediction values.

6. CONCLUSIONS

6.1. Summary of Findings

To wrap up, it can be found below a graph that summarizes the error measures of the different models analyzed in this study.

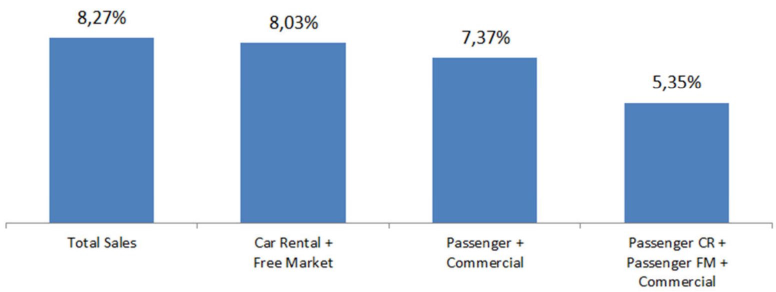


Figure 35 – Error measure (MAPE) of best models

It can me concluded that higher level hierarchy forecasting may provide better results as subsetting the series provide more “clean” data with less variability. Also to be noted that regression models with arima errors perform better overall mainly by taking in account the various outliers present in the series as they have significant impact. Also a few factors, mainly economic do have influence on the behavior of the series.

6.2. Conclusion

From this study it is possible to derive a few conclusions that are important to better understand and model how Portuguese automotive market behaves.

It was shown that indeed is possible to create more accurate predictions by subsetting the data and recurring to hierarchical forecasting when compared with forecasting the market sales as a whole and a segregation based on vehicle category has proven to have slight overall advantage when considering a one level split. Better results were found when these two segregation types were aggregated using a two level split.

Regarding to outlier events, they do have impact on forecasting capabilities of the regression models and it is interesting that almost all are derived from economical events, especially on vehicle taxation. Most of these events are announced through mass media, which in turn causes general public to be aware and take action, whether anticipating or delaying their buy. These situations are also exploited by car dealerships that create a sense of urgency and thus mobilize even further the general public.

On the subject of other external factors, the ones that play significant influence on modelling the market behaviour are from economic nature, such as economic climate indicator. Factors of social nature such as employability do have correlation with sales but many times are not significant as a prediction variable. On car rental market, it is interesting to find that tourism and expectations of demand in services are strong components for prediction. However, as commercial market, some behaviour is dependent on pre-arranged business contracts which cannot be predicted easily.

6.3. FUTURE WORK

In terms of future work and having account the work developed in this study, it will be interesting to evaluate if further segregations, such as subsetting the data by the different brands will bring even better results. Even if does not bring significant improvements it may bring remarkable insights for competitive analysis for data driven decision making.

7. ADDENDUM

After the study conclusion and before submission, it was possible to gather new data of the main features and thus allow retesting and further validate the robustness of the prediction models created when applied over a time span.

The new data is comprised of a timeframe between August 2016 to December 2017 which, when aggregated with the previous information, allows 12 years of historical data and subsequently 144 data points.

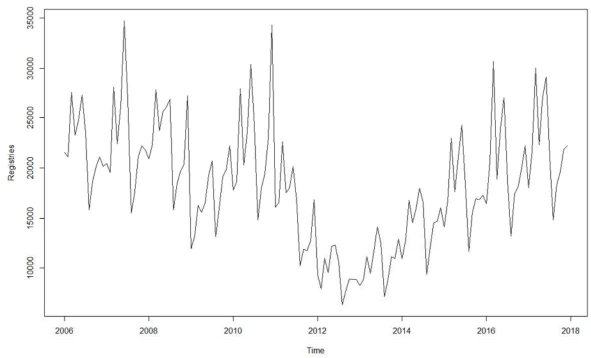


Figure 36 – Historical Time Series of total vehicle sales

From the analysis of the new time series, it appears that 2017 it a year that follows the same upward trend and seasonality as the years before but with a very slight deceleration.

Considering the overall series there are no significant outliers but observing only car rental sales, 2017 was a rather unusual year (figure 37), with very high activity during the second quarter and peaking in June. This peak will be considered as an outlier. Also, sales of commercial vehicles had an uncommon seasonality behavior on the last quarter (figure 38) there were no significant outliers to be considered.

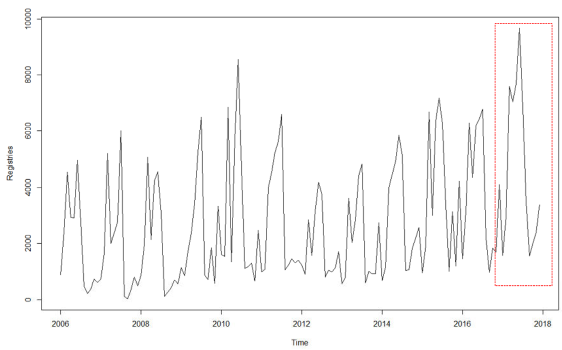


Figure 37 – Historical Time Series of total Car Rental

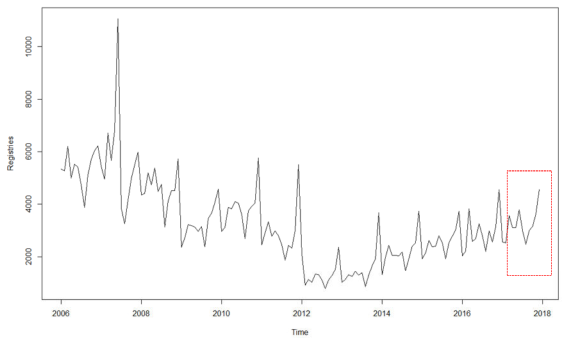


Figure 38 – Historical Time Series of total Car Rental

7.1. MODEL COMPARISON (USING UNCHANGED PARAMETERS)

By applying the same parameters of previously calculated models, the accuracy results can be summarized in the graphs below, showing only the best models.

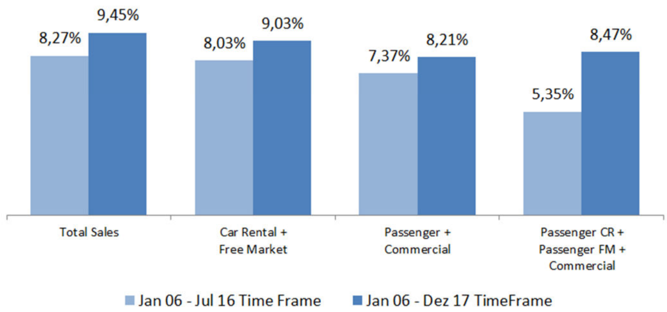


Figure 38 – Error (MAPE) model comparison using the same parameters

Overall the accuracy of the prediction models has worsened a little bit. The best model in evidence is the hierarchical forecasting with passenger and commercial split. Forecasting with subset data still shows better results when compared with forecasting with whole data. In terms of robustness it is not clear that increasing hierarchical levels improve the accuracy of the forecast.

These results illustrate the importance of updating models with regular frequency as some parameters or even factors may not be as significant anymore.

7.2. MODEL PARAMETER RE-ESTIMATION

In order to truly assess the quality of the models, the ARIMA models were updated through re-estimation of the parameters, using the same methodologies for model estimation and validation. The external factors did not exhibit significant changes regarding their respective regression models.

When forecasting vehicle sales as a whole, the model re-estimation has proven to be effective as accuracy rate has further increased.

	ARIMA Parameters						Error (MAPE)
	Non-Seasonal			Seasonal			
Type	p	d	q	P	D	Q	
(1) Standard	2	1	0	2	0	1	
(2) Reg. with outliers	1	1	2	1	0	2	
(3) Reg. w/ predictors and outliers	0	1	1	2	0	2	

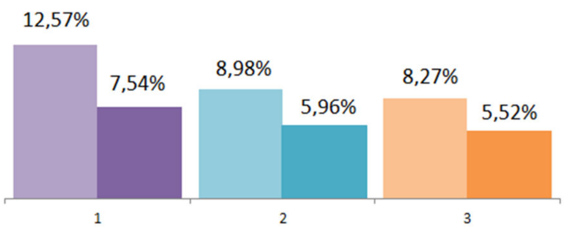


Figure 39 – Error (MAPE) for total vehicle sales
(left bar: Jan 06 – Jul 16 Time Frame; right bar: Jan 06 – Dez 17 Time Frame)

When comparing the best forecasting aggregate model in this study - two level hierarchical split - the error values of the different models can be summarized in the graph below.

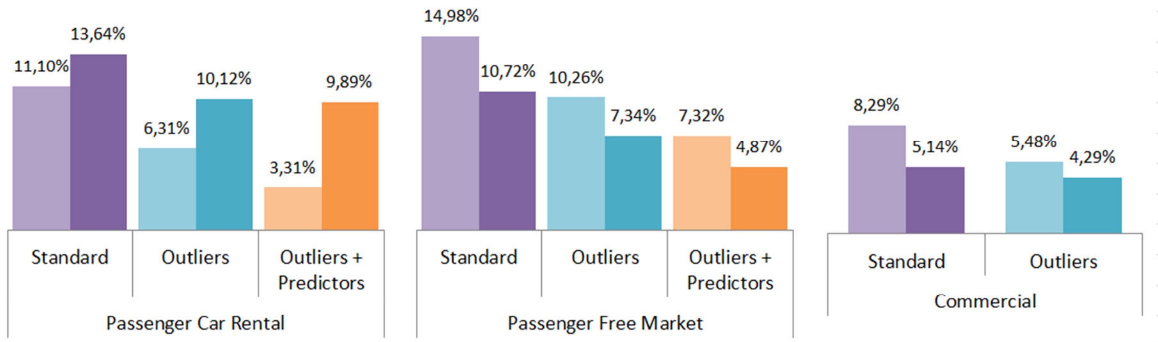


Figure 40 – Error (MAPE) for two level hierarchical split

As mentioned earlier, due to the unusual behavior of the car rental sales in 2017, the prediction models have worse results when compared with the previous time frame. However, re-estimating the ARIMA parameters captured well the behavior for passenger free market and commercial market, increasing even more the forecast accuracy. The forecast aggregation of these models have a 4,61% of MAPE.

In conclusion, passenger free market and commercial market models prevail as good long term forecast models as long there is periodic re-estimation of the model parameters. On the other hand, car rental which is more influenced by business to business agreements needs to be monitored closer for more accurate predictions.

8. BIBLIOGRAPHY

ACAP - Associação Do Comércio Automóvel De Portugal (2016a). *Quem somos*. Retrieved October 28, 2016 from http://www.acap.pt/

ACAP - Associação Do Comércio Automóvel De Portugal (2016b). *Estatísticas do Sector Automóvel, edição 2016*. ACAP

ARAC - Associação dos Industriais de Aluguer de Automoveis sem Condutor (2016). *Estatísticas*. Retrieved October 28, 2016 from http://www.arac.pt/

Athanasopoulos, G., Ahmed, R., Hyndman, R. (2007) *Hierarchical Forecast for Australian domestic tourism*. Department of Econometrics and Business Statistics Working Paper Series 12/07

AutoInforma (2016) *Estatísticas.* Retrieved October 28, 2016 from http://www.autoinforma.pt

Berkovec, J. (1985) *New Car Sales and Used Car Stocks: A Model for the Automobile Market*. The RAND Journal of Economics 2, 195–214

Brühl, B., Hülsmann, M., Friedrich, C. and Reith, D. (2009). *A Sales Forecast Model for the German Automobile Market Based on Time Series Analysis and Data Mining Methods*. Advances in Data Mining. Applications and Theoretical Aspects, 9th Industrial Conference, ICDM 2009, Leipzig, Germany

Calcagno, V., Mazancourt, C., *glmulti: An R Package for Easy Automated Model Selection with (Generalized) Linear Models*. Journal of Statistical Software, Vol. 34 Issue 12.

Cryer, J., Chan, K. (2008), *Time Series Analysis with applications in R – second edition*, Springer-Verlag New York

Dudenhöffer, F., Borscheid, D. (2004) *Automobilmarkt-Prognosen: Modelle und Methoden.*
In: Ebel, B., Hofer, M.B., Al-Sibai, J. (2004) *Automotive Management – Strategie und Marketing in der Automobilwirtschaft*, pp. 192–202, Springer, Berlin, Heidelberg

Godinho, Raquel (2016, July 17) Banca empresta mais de cinco milhões de euros por dia para a compra de carro. *Jornal de Negócios*. Retrieved October 22, 2016 from http://www.jornaldenegocios.pt/

Gross, C., Sohl, J. (1990). *Disaggregation methods to expedite product line forecasting*. Journal of Forecasting 9(3), 233-254

Guerrero, V.M. (1993), *Time-series analysis supported by power transformations*, Journal of forecasting Vol. 12, 37-48

Holweg (2008) *The Evolution in the Automotive Industry*. In: Parry, G., Graves, A. (2008*). Build To Order* (pp 13-34). Springer, London

Hülsmann, M., Borscheid, D., Friedrich, C. and Reith,D. (2012). *General Sales Forecast Models for Automobile Markets and their Analysis*. Transactions on Machine Learning and Data Mining Vol. 5, No.2, 65-86.

Hyndman, Rob J., Ahmed, R. A., Athanasopoulos, G., Shang (2010). *Optimal combination forecasts for hierarchical time series*. Department of Econometrics and Business Statistics Working Paper Series 09/07

Lewandowsky, R. (1974). *Prognose - und Informationssysteme und ihre Anwendungen*. De Gruyter, Berlin.

Makridakis, S., Wheelwright, S., Hyndman, R. (1998) *Forecasting: Methods and Applications, Third edition*. John Wiley and Sons

Martins, A. (2012). *Procura Automóvel, Estimação e Previsão* (Master's Dissertation, Instituto Superior de Economia e Gestão, Lisbon, Portugal). Retrieved October 22, 2016 from https://www.repository.utl.pt/

Monteiro, C. (2009). *Análise do Sector Automóvel em Portugal* (Master's Dissertation, instituto Politécnico de Bragança, Bragança, Portugal). Retrieved October 22, 2016 from https://bibliotecadigital.ipb.pt/

Solomou, S., Weale, M. (1991). *Balanced estimates of UK GDP 1870-1913*. Explorations in Economic History 28 (1), 54–63.

Solomou, S., Weale, M. (1993). *Balanced estimates of national accounts when measurement errors are autocorrelated: the UK, 1920-38*. Journal of the Royal Statistical Society: Series A 156 (1), 89–105.

Solomou, S., Weale, M. (1996). *UK national income, 1920-1938: the implications of balanced estimates*. The Economic History Review 49 (1), 101–115

Stineman, R. W. (1980) *A Consistently Well Behaved Method of Interpolation.* Creative Computing , volume 6, number 7, p. 54-57.

9. APPENDICES

9.1 External Factors Detailed

Factor	Description (ENG)	Description (PT)	Type	URL
R1	Unemployment rate of active population between 15 and 74 years old (%)	Taxa de desemprego (%) da população ativa com idade entre 15 e 74 anos	Social	https://www.ine.pt/xportal/xmain?xpid=INE&xpgid=ine_indicadores&indOcorrCod=0007976&contexto=bd&selTab=tab2
R2	Private Consumption Indicator	Indicador Coincidente do Consumo Privado	Economic	https://www.bportugal.pt/Mobile/BPStat/GraficoBoletimEstatistico.aspx?img=3&BEIndID=826845&SW=1269
R3	Economic Climate Indicator	Indicador de clima económico	Economic	https://www.ine.pt/xportal/xmain?xpid=INE&xpgid=ine_indicadores&indOcorrCod=0001222&contexto=bd&selTab=tab2
R4	Expectations of unemployment evolution over the next 12 months	Perspectiva sobre evolução do desemprego nos próximos 12 meses	Social	https://www.ine.pt/xportal/xmain?xpid=INE&xpgid=ine_indicadores&indOcorrCod=0001166&contexto=bd&selTab=tab2
R5	Expectations on economic climate over the next 12 months	Perspectiva sobre a situação económica do país nos próximos 12 meses	Economic	https://www.ine.pt/xportal/xmain?xpid=INE&xpgid=ine_indicadores&indOcorrCod=0001169&contexto=bd&selTab=tab2
R6	Consumer price index	Índice de preços no consumidor (IPC, Base - 2012)	Economic	https://www.ine.pt/xportal/xmain?xpid=INE&xpgid=ine_indicadores&indOcorrCod=0008352&contexto=bd&selTab=tab2
R7	Crude Oil price - Brent (EUR)	Cotação do barril de petróleo - Brent (EUR)	Specific	https://pt.investing.com/commodities/brent-oil
R8	Nights spent at tourist accommodation establishments - Residents		Social	http://appsso.eurostat.ec.europa.eu/nui/show.do?dataset=tour_occ_nim&lang=en
R9	Nights spent at tourist accommodation establishments - Non-Residents		Social	http://appsso.eurostat.ec.europa.eu/nui/show.do?dataset=tour_occ_nim&lang=en
R10	Appreciation on economic climate of the last 12 months	Apreciação sobre a situação económica do país nos últimos 12 meses	Economic	https://www.ine.pt/xportal/xmain?xpid=INE&xpgid=ine_indicadores&indOcorrCod=0001170&contexto=bd&selTab=tab2
R11	Economic Activity Indicator (%)	Indicador de atividade económica	Economic	https://www.ine.pt/xportal/xmain?xpid=INE&xpgid=ine_indicadores&indOcorrCod=0007756&contexto=bd&selTab=tab2
R12	Employment Index on Industry Sector	Índice de emprego na indústria - bruto	Social	https://www.ine.pt/xportal/xmain?xpid=INE&xpgid=ine_indicadores&indOcorrCod=0009461&contexto=bd&selTab=tab2
R13	Employment Index on Commerce Sector	Índice de emprego no comércio a retalho	Social	https://www.ine.pt/xportal/xmain?xpid=INE&xpgid=ine_indicadores&indOcorrCod=0009285&contexto=bd&selTab=tab2
R14	Employment Index on Service Sector	Índice de emprego nos serviços - bruto	Social	https://www.ine.pt/xportal/xmain?xpid=INE&xpgid=ine_indicadores&indOcorrCod=0007709&contexto=bd&selTab=tab2
R15	Savings Opportunity	Oportunidade de realização de poupança	Economic	https://www.ine.pt/xportal/xmain?xpid=INE&xpgid=ine_indicadores&indOcorrCod=0001176&contexto=bd&selTab=tab2
R16	Expectations of equipment good acquisition in the next 12 months	Perspectiva sobre aquisição de bens de equipamento nos próximos 12 meses	Economic	https://www.ine.pt/xportal/xmain?xpid=INE&xpgid=ine_indicadores&indOcorrCod=0001164&contexto=bd&selTab=tab2
R17	Expectations on sales prices over the next 12 months	Perspectiva sobre a evolução dos preços nos próximos 12 meses	Economic	https://www.ine.pt/xportal/xmain?xpid=INE&xpgid=ine_indicadores&indOcorrCod=0001167&contexto=bd&selTab=tab2
R18	Number of weekday holidays per month	Número de feriados em dia de semana	Calendar	
R19	Number of workdays per month	Número de dias úteis por mês sem feriados	Calendar	
R20	Expectations on demand on services over the next 3 months	Perspectivas sobre a procura nos próximos 3 meses (Saldo de respostas extremas) dos serviços	Economic	https://www.ine.pt/xportal/xmain?xpid=INE&xpgid=ine_indicadores&indOcorrCod=0001200&contexto=bd&selTab=tab2
R21	Expectations on demand in commerce over the next 3 months	Perspectivas sobre a procura nos próximos 3 meses (Saldo de respostas extremas) do comércio	Economic	https://www.ine.pt/xportal/xmain?xpid=INE&xpgid=ine_indicadores&indOcorrCod=0001155&contexto=bd&selTab=tab2
R22	Salary Index on Industry Sector	Índice de remunerações na indústria - bruto	Economic	https://www.ine.pt/xportal/xmain?xpid=INE&xpgid=ine_indicadores&indOcorrCod=0009459&contexto=bd&selTab=tab2
R23	Salary Index on Commerce Sector	Índice de remunerações no comércio a retalho - bruto	Economic	https://www.ine.pt/xportal/xmain?xpid=INE&xpgid=ine_indicadores&indOcorrCod=0009284&contexto=bd&selTab=tab2
R24	Salary Index on Service Sector	Índice de remunerações nos serviços - bruto	Economic	https://www.ine.pt/xportal/xmain?xpid=INE&xpgid=ine_indicadores&indOcorrCod=0007723&contexto=bd&selTab=tab2

Table I – External Factors Explained

9.2 Car Rental Forecasting Detailed

Non-Seasonal			Seasonal			Evaluation	Error (MAPE)	Diagnostic (Ljung-Box at lag)					
p	d	q	P	D	Q	AIC	k-fold Cross Validated	6	12	18	24	30	36
0	1	1	0	1	1	1040,088	**10,579%**	0,622	0,627	0,705	0,846	0,738	0,737
0	1	1	0	1	2	1041,436	**11,540%**	0,672	0,734	0,827	0,930	0,873	0,867
0	1	2	0	1	1	1040,824	**12,107%**	0,716	0,636	0,720	0,843	0,802	0,779
0	1	1	1	1	1	1041,188	**12,115%**	0,690	0,775	0,868	0,950	0,907	0,903
1	1	1	0	1	1	1040,709	**12,147%**	0,721	0,636	0,725	0,843	0,803	0,777
0	1	1	2	1	1	1042,433	**12,309%**	0,681	0,808	0,893	0,959	0,901	0,907
0	1	2	0	1	2	1042,175	**13,187%**	0,779	0,748	0,837	0,936	0,920	0,907
1	1	1	0	1	2	1042,066	**13,255%**	0,784	0,748	0,842	0,937	0,920	0,905
0	1	2	1	1	1	1041,893	**13,857%**	0,805	0,797	0,882	0,959	0,948	0,940
1	1	1	1	1	1	1041,777	**13,929%**	0,813	0,800	0,887	0,960	0,950	0,940

Table II – TOP 10 ARIMA models overview – Car Rental

9.3 Car Rental Forecasting with Outliers Detailed

Non-Seasonal			Seasonal			Evaluation	Error (MAPE)	Diagnostic (Ljung-Box at lag)					
p	d	q	P	D	Q	AIC	k-fold Cross Validated	6	12	18	24	30	36
0	1	1	1	1	2	1041,416	**10,920%**	0,916	0,830	0,944	0,984	0,921	0,907
0	1	1	0	1	1	1038,168	**11,333%**	0,914	0,769	0,894	0,957	0,843	0,820
0	1	1	0	1	2	1040,107	**11,628%**	0,919	0,790	0,909	0,968	0,872	0,852
2	1	1	0	1	1	1041,859	**11,742%**	0,912	0,740	0,878	0,943	0,842	0,811
0	1	2	0	1	1	1040,005	**11,904%**	0,906	0,741	0,875	0,948	0,855	0,827
0	1	1	1	1	1	1040,059	**11,908%**	0,924	0,807	0,921	0,975	0,893	0,875
1	1	1	0	1	1	1039,983	**11,963%**	0,904	0,736	0,872	0,946	0,856	0,826
0	1	1	2	1	1	1040,400	**12,058%**	0,912	0,882	0,970	0,994	0,931	0,929
0	1	2	1	1	1	1041,854	**12,740%**	0,922	0,791	0,912	0,973	0,916	0,896
1	1	1	1	1	1	1041,825	**12,835%**	0,921	0,789	0,911	0,973	0,918	0,899

Table III – TOP 10 ARIMA models overview – Car Rental with outliers

9.4 Car Rental Forecasting with External Factors

	Lags												
Factors	-12	-11	-10	-9	-8	-7	-6	-5	-4	-3	-2	-1	0
R1	-0,03	-0,02	-0,03	-0,05	-0,04	-0,07	-0,12	-0,13	-0,14	-0,13	-0,15	-0,11	-0,11
R2	-0,09	-0,08	-0,06	-0,04	-0,02	0,00	0,01	0,03	0,06	0,09	0,12	0,14	0,15
R3	-0,07	-0,06	-0,04	-0,02	0,02	0,06	0,06	0,06	0,08	0,11	0,15	0,19	0,21
R4	0,07	0,01	0,04	0,00	-0,01	-0,06	-0,11	-0,04	-0,07	-0,11	-0,08	-0,13	-0,15
R5	-0,11	-0,17	-0,20	-0,03	-0,11	-0,03	0,12	-0,02	0,00	0,16	0,10	0,11	0,17
R6	0,09	0,10	0,11	0,01	0,00	0,02	0,00	-0,03	-0,05	-0,07	-0,12	-0,11	-0,08
R7	-0,11	-0,07	-0,10	-0,12	-0,15	-0,06	0,00	0,03	0,14	0,22	0,14	0,16	0,14
R8	0,00	-0,14	-0,06	-0,17	-0,10	-0,07	-0,05	0,06	0,16	0,20	0,32	0,19	0,12
R9	0,08	-0,03	-0,02	-0,20	-0,22	-0,10	-0,17	0,00	0,14	0,23	0,28	0,15	0,19
R10	-0,08	-0,12	-0,05	0,04	-0,03	0,02	0,07	-0,01	0,04	0,15	0,13	0,15	0,17
R11	-0,04	-0,01	0,05	0,10	0,11	0,14	0,17	0,19	0,20	0,18	0,18	0,16	0,11
R12	0,06	0,08	0,07	0,08	0,09	0,10	0,11	0,10	0,12	0,12	0,07	0,08	0,04
R13	0,07	0,07	0,04	0,11	0,12	0,14	0,19	0,19	0,16	0,16	0,15	0,11	0,08
R14	0,09	0,12	0,11	0,17	0,14	0,16	0,22	0,23	0,25	0,22	0,16	0,14	0,09
R15	0,03	-0,05	0,02	0,11	0,01	0,05	0,14	0,00	0,03	0,11	0,05	0,07	0,02
R16	-0,08	-0,11	-0,11	-0,09	-0,07	0,00	0,04	-0,07	0,08	0,00	0,05	0,14	0,16
R17	0,07	0,12	0,19	0,12	0,19	0,19	0,12	0,18	0,14	0,08	0,12	0,11	0,03
R18	0,12	-0,18	0,07	-0,07	0,01	-0,03	0,07	-0,17	-0,01	0,04	0,01	-0,02	0,10
R19	0,03	0,02	-0,13	0,09	-0,06	0,04	0,03	0,13	-0,01	0,08	-0,06	0,03	-0,07
R20	-0,08	-0,10	-0,07	0,00	-0,03	-0,02	0,00	-0,04	0,09	0,15	0,25	0,37	0,32
R21	-0,08	-0,11	-0,13	-0,03	0,01	-0,08	0,06	0,11	0,13	0,15	0,21	0,17	0,27
R22	0,06	-0,02	-0,06	0,05	0,01	0,02	0,11	0,21	0,07	-0,05	0,09	-0,01	0,09
R23	0,17	-0,01	0,02	0,09	0,07	0,05	0,19	0,10	0,08	-0,02	-0,09	0,06	-0,04
R24	0,05	0,09	-0,08	0,14	0,05	0,08	0,17	0,24	0,14	0,03	-0,04	0,09	0,06

Table IV – CCF Table on External Factors vs Car Rental

Non-Seasonal			Seasonal			Evaluation	Error (MAPE)	Diagnostic (Ljung-Box at lag)					
p	d	q	P	D	Q	AIC	k-fold Cross Validated	6	12	18	24	30	36
0	1	1	0	1	1	1012,783	**9,193%**	0,751	0,793	0,908	0,904	0,748	0,780
1	1	1	0	1	1	1011,637	**10,049%**	0,930	0,892	0,952	0,934	0,893	0,900
1	1	1	0	1	2	1013,626	**10,305%**	0,931	0,890	0,951	0,931	0,887	0,897
1	1	1	1	1	1	1013,620	**10,443%**	0,931	0,890	0,950	0,929	0,884	0,895
1	1	2	0	1	1	1013,137	**10,598%**	0,948	0,912	0,961	0,960	0,938	0,945
0	1	2	0	1	1	1011,139	**10,604%**	0,947	0,911	0,961	0,960	0,938	0,946
0	1	2	0	1	2	1013,113	**10,977%**	0,948	0,910	0,960	0,958	0,934	0,944
0	1	2	1	1	1	1013,100	**11,121%**	0,949	0,910	0,959	0,957	0,932	0,943
2	1	1	0	1	1	1012,325	**11,608%**	0,975	0,955	0,979	0,986	0,978	0,981
2	1	2	0	1	1	1013,100	**13,922%**	1,000	0,989	0,989	0,988	0,977	0,982

Table V – TOP 10 ARIMA models overview – Car Rental with external factors

9.5 Free Market Forecasting

Non-Seasonal			Seasonal			Evaluation	Error (MAPE)	Diagnostic (Ljung-Box at lag)					
p	d	q	P	D	Q	AIC	k-fold Cross Validated	6	12	18	24	30	36
2	1	0	2	0	1	-599,066	**16,126%**	0,582	0,659	0,715	0,678	0,852	0,922
2	1	0	1	0	1	-600,777	**16,128%**	0,614	0,649	0,729	0,659	0,850	0,934
2	1	1	1	0	1	-599,844	**16,659%**	0,802	0,795	0,873	0,739	0,899	0,963
0	1	1	1	0	2	-599,037	**17,229%**	0,318	0,387	0,489	0,479	0,713	0,821
0	1	1	1	0	1	-600,364	**17,379%**	0,298	0,314	0,442	0,410	0,658	0,803
0	1	1	2	0	1	-599,196	**17,519%**	0,315	0,402	0,508	0,497	0,725	0,813
2	1	2	2	0	2	-599,163	**18,282%**	0,999	0,977	0,977	0,926	0,974	0,995
2	1	2	1	0	1	-601,781	**18,866%**	1,000	0,979	0,984	0,941	0,979	0,995
2	1	2	1	0	2	-600,103	**18,917%**	1,000	0,984	0,984	0,943	0,978	0,994
2	1	2	2	0	1	-600,217	**19,161%**	1,000	0,986	0,984	0,944	0,978	0,993

Table VI – TOP 10 ARIMA models overview – Free Market

9.6 Forecasting Free Market with Outliers

Non-Seasonal			Seasonal			Evaluation	Error (MAPE)	Diagnostic (Ljung-Box at lag)					
p	d	q	P	D	Q	AIC	k-fold Cross Validated	6	12	18	24	30	36
2	1	1	1	0	2	-654,766	**11,517%**	0,603	0,750	0,736	0,741	0,670	0,810
2	1	0	1	0	2	-654,077	**12,025%**	0,321	0,469	0,438	0,519	0,512	0,705
2	1	1	2	0	1	-654,348	**12,259%**	0,498	0,669	0,646	0,622	0,510	0,699
2	1	0	2	0	1	-653,831	**12,749%**	0,277	0,419	0,379	0,440	0,403	0,619
2	1	2	1	0	1	-660,722	**13,053%**	0,951	0,887	0,883	0,899	0,916	0,973
2	1	2	1	0	0	-657,453	**13,055%**	0,975	0,967	0,955	0,971	0,961	0,991
2	1	2	1	0	2	-661,336	**13,375%**	0,814	0,858	0,890	0,901	0,940	0,979
2	1	2	2	0	1	-662,747	**13,642%**	0,893	0,911	0,921	0,943	0,956	0,988
2	1	2	2	0	0	-658,730	**13,677%**	0,989	0,957	0,936	0,948	0,931	0,986
2	1	2	2	0	2	-660,789	**13,724%**	0,905	0,913	0,921	0,943	0,955	0,988

Table VII – TOP 10 ARIMA models overview – Free Market with outliers

9.7 Forecasting Free Market with External Factors

Factors	Lags												
	-12	-11	-10	-9	-8	-7	-6	-5	-4	-3	-2	-1	0
R1	-0,07	-0,04	0,00	-0,06	-0,19	-0,32	-0,29	-0,22	-0,23	-0,22	-0,26	-0,19	-0,11
R2	-0,24	-0,18	-0,14	-0,09	-0,03	0,00	0,05	0,07	0,13	0,17	0,22	0,26	0,27
R3	-0,06	-0,04	0,03	0,06	0,08	0,03	0,05	0,07	0,11	0,14	0,14	0,14	0,17
R4	0,04	0,00	-0,02	0,06	-0,13	-0,02	0,04	0,01	-0,03	-0,15	-0,13	-0,05	-0,10
R5	-0,20	-0,03	-0,08	-0,12	0,00	0,02	-0,11	-0,11	-0,08	0,07	-0,02	-0,02	0,04
R6	0,22	0,08	-0,01	-0,11	-0,19	-0,04	0,19	0,16	0,10	-0,03	-0,06	0,05	0,07
R7	0,04	-0,01	-0,11	-0,16	-0,03	-0,03	-0,02	0,07	0,11	0,07	0,08	0,06	-0,01
R8	-0,18	-0,15	-0,01	0,08	0,08	-0,04	0,01	0,07	0,07	0,06	0,09	0,01	-0,23
R9	-0,12	-0,06	0,03	-0,01	-0,02	-0,02	0,03	0,12	0,20	0,08	-0,03	-0,07	-0,14
R10	-0,12	0,07	-0,01	-0,04	-0,02	-0,01	-0,08	-0,14	-0,11	0,07	0,04	0,05	0,05
R11	-0,16	-0,02	-0,04	-0,02	0,02	0,08	0,19	0,19	0,23	0,31	0,35	0,32	0,28
R12	0,07	0,07	0,09	0,09	0,10	0,08	0,10	0,13	0,14	0,08	0,08	0,03	0,09
R13	0,05	-0,07	-0,01	0,15	0,19	0,15	0,15	0,18	0,05	-0,01	0,12	0,13	0,10
R14	0,05	0,01	0,02	0,10	0,15	0,20	0,24	0,32	0,23	0,12	0,04	-0,01	0,11
R15	-0,09	0,10	0,00	-0,16	0,01	-0,06	-0,01	-0,16	0,01	-0,04	0,05	-0,12	0,09
R16	-0,11	0,06	-0,01	0,10	0,08	-0,09	-0,12	-0,03	-0,09	0,04	0,08	-0,03	0,03
R17	0,03	0,02	0,01	0,10	0,07	0,07	0,02	0,03	0,15	0,11	0,05	0,14	0,17
R18	0,09	0,02	-0,04	-0,04	-0,11	0,04	0,09	0,03	-0,12	0,01	-0,11	0,12	0,11
R19	-0,08	-0,04	0,01	-0,01	0,12	-0,08	0,02	0,02	0,03	-0,04	0,08	-0,16	0,08
R20	-0,02	0,05	-0,12	-0,11	-0,12	-0,15	-0,11	0,07	0,05	0,17	0,04	0,12	0,18
R21	-0,12	-0,02	-0,02	-0,13	-0,14	-0,06	-0,01	0,00	0,06	0,05	0,04	0,11	0,12
R22	0,13	-0,08	-0,17	0,05	0,07	0,06	0,11	0,11	-0,16	-0,17	0,04	0,18	0,12
R23	0,15	-0,02	-0,16	0,04	0,04	0,08	0,14	0,11	-0,18	-0,16	0,06	0,13	0,18
R24	0,26	-0,05	-0,25	0,04	0,07	0,11	0,23	0,14	-0,18	-0,24	-0,02	0,20	0,32

Table VIII – CCF Table on External Factors vs Free Market

Non-Seasonal			Seasonal			Evaluation	Error (MAPE)	Diagnostic (Ljung-Box at lag)					
p	d	q	P	D	Q	AIC	k-fold Cross Validated	6	12	18	24	30	36
1	1	2	1	0	1	-610,118	**10,934%**	0,508	0,451	0,773	0,722	0,807	0,845
0	1	2	1	0	2	-609,128	**11,155%**	0,430	0,379	0,713	0,640	0,746	0,747
0	1	2	2	0	1	-609,145	**11,241%**	0,432	0,384	0,719	0,647	0,750	0,752
1	1	1	1	0	2	-608,983	**11,248%**	0,399	0,349	0,690	0,618	0,721	0,726
2	1	1	1	0	1	-609,424	**11,250%**	0,630	0,564	0,835	0,746	0,850	0,853
0	1	1	1	0	1	-612,236	**11,329%**	0,397	0,341	0,688	0,616	0,718	0,743
0	1	1	1	0	2	-610,459	**11,398%**	0,347	0,318	0,663	0,590	0,669	0,689
1	1	1	1	0	1	-610,702	**11,429%**	0,447	0,367	0,710	0,633	0,759	0,770
0	1	2	1	0	1	-610,820	**11,442%**	0,474	0,391	0,728	0,648	0,776	0,784
0	1	1	2	0	1	-610,459	**11,455%**	0,350	0,322	0,668	0,596	0,674	0,694

Table IX – TOP 10 ARIMA models overview – Free Market with external factors

9.8 Forecasting Passenger Market

Non-Seasonal			Seasonal			Evaluation	Error (MAPE)	Diagnostic (Ljung-Box at lag)					
p	d	q	P	D	Q	AIC	k-fold Cross Validated	6	12	18	24	30	36
0	1	1	1	0	1	-324,988	**13,706%**	0,677	0,847	0,937	0,913	0,970	0,992
0	1	1	1	0	2	-324,281	**13,931%**	0,596	0,874	0,945	0,905	0,972	0,991
0	1	1	2	0	1	-325,241	**14,083%**	0,596	0,886	0,949	0,907	0,975	0,987
2	1	0	1	0	1	-325,703	**14,501%**	0,980	0,987	0,994	0,986	0,997	1,000
0	1	2	1	0	1	-324,023	**14,810%**	0,861	0,937	0,979	0,956	0,986	0,997
2	1	0	2	0	1	-325,958	**14,835%**	0,958	0,994	0,996	0,986	0,998	0,999
2	1	0	1	0	2	-325,101	**14,862%**	0,958	0,993	0,996	0,987	0,998	1,000
0	1	2	2	0	1	-324,783	**14,988%**	0,866	0,981	0,991	0,968	0,993	0,997
2	1	1	2	0	1	-324,389	**15,034%**	0,987	0,998	0,998	0,988	0,998	0,999
0	1	2	1	0	2	-323,939	**15,069%**	0,868	0,978	0,991	0,969	0,993	0,998

Table X – TOP 10 ARIMA models overview – Passenger Market

9.9 Forecasting Passenger Market with Outliers

Non-Seasonal			Seasonal			Evaluation	Error (MAPE)	Diagnostic (Ljung-Box at lag)					
p	d	q	P	D	Q	AIC	k-fold Cross Validated	6	12	18	24	30	36
2	1	0	1	0	2	-400,852	**9,097%**	0,543	0,853	0,867	0,929	0,575	0,721
2	1	1	1	0	2	-398,861	**9,117%**	0,548	0,856	0,869	0,929	0,582	0,728
2	1	0	2	0	2	-399,426	**10,533%**	0,655	0,887	0,913	0,952	0,682	0,810
2	1	1	2	0	2	-397,423	**10,553%**	0,660	0,890	0,914	0,953	0,688	0,814
1	1	2	1	0	2	-395,591	**10,645%**	0,261	0,601	0,707	0,847	0,510	0,649
0	1	2	1	0	2	-395,727	**11,317%**	0,193	0,531	0,697	0,842	0,478	0,612
2	1	2	1	0	2	-400,410	**11,369%**	0,884	0,973	0,943	0,974	0,746	0,880
1	1	2	2	0	2	-394,205	**12,038%**	0,330	0,642	0,777	0,888	0,612	0,747
0	1	2	2	0	2	-394,287	**12,580%**	0,231	0,552	0,746	0,862	0,543	0,681
2	1	2	2	0	2	-400,280	**13,085%**	0,928	0,951	0,962	0,990	0,905	0,976

Table XI – TOP 10 ARIMA models overview – Passenger Market with outliers

9.10 Forecasting Passenger Market with External Factors

	Lags												
Factors	-12	-11	-10	-9	-8	-7	-6	-5	-4	-3	-2	-1	0
R1	0,02	-0,04	-0,09	-0,07	-0,07	-0,31	-0,20	-0,15	-0,23	-0,24	-0,29	-0,12	-0,09
R2	-0,18	-0,11	-0,07	-0,03	0,01	0,02	0,04	0,05	0,08	0,12	0,16	0,20	0,20
R3	-0,04	-0,02	0,06	0,10	0,10	0,02	-0,03	0,00	0,08	0,21	0,26	0,26	0,32
R4	0,03	-0,04	0,01	0,08	-0,06	-0,07	0,01	0,07	-0,01	-0,08	0,00	-0,11	-0,17
R5	-0,17	-0,01	-0,16	-0,06	0,03	0,10	0,00	-0,23	-0,20	0,08	-0,04	0,05	0,15
R6	-0,03	0,03	0,02	-0,12	-0,04	0,02	-0,05	0,02	0,08	-0,03	0,08	0,03	-0,19
R7	-0,01	0,01	-0,10	-0,22	-0,06	0,00	-0,06	0,03	0,13	0,12	0,04	0,01	-0,06
R8	-0,09	-0,11	0,21	-0,08	-0,09	0,03	0,22	-0,18	0,20	-0,11	0,05	0,11	-0,01
R9	-0,03	0,00	0,07	0,02	-0,01	0,07	-0,05	-0,06	0,21	-0,13	-0,03	0,08	0,07
R10	-0,04	0,04	-0,06	0,00	0,00	0,02	-0,02	-0,18	-0,14	0,08	0,06	0,17	0,08
R11	-0,09	0,01	-0,01	0,02	0,00	0,04	0,15	0,13	0,15	0,22	0,23	0,23	0,15
R12	0,02	0,02	0,02	0,02	0,03	0,02	0,02	0,03	0,03	0,02	0,02	0,03	0,06
R13	0,08	0,04	0,10	0,15	0,13	0,09	0,07	0,12	0,12	0,17	0,16	0,09	0,20
R14	0,07	0,10	0,12	0,12	0,19	0,13	0,08	0,13	0,15	0,17	0,17	0,11	0,20
R15	0,00	0,05	-0,08	-0,04	0,08	-0,03	0,05	-0,18	-0,07	0,03	0,02	-0,07	0,14
R16	-0,09	0,06	-0,08	0,11	0,17	-0,09	-0,05	-0,02	-0,09	0,00	0,08	0,05	0,01
R17	0,03	0,03	0,09	0,12	0,08	0,00	-0,08	0,09	0,20	0,10	0,02	0,10	0,09
R18	0,11	-0,08	0,20	-0,11	-0,07	0,01	0,04	-0,12	0,10	0,06	-0,11	-0,05	0,21
R19	-0,03	0,13	-0,28	0,13	0,10	-0,13	0,06	0,10	-0,17	0,10	0,04	-0,21	0,18
R20	-0,07	-0,06	-0,07	0,01	0,07	-0,08	-0,11	-0,05	-0,07	0,10	0,08	0,22	0,10
R21	-0,07	-0,07	-0,07	-0,06	0,03	0,00	0,01	-0,09	-0,01	-0,06	0,12	0,18	0,07
R22	0,04	0,07	-0,03	0,02	0,04	0,03	0,01	0,07	-0,03	0,01	0,01	0,04	0,10
R23	0,07	0,00	0,01	0,06	0,05	0,01	0,01	0,05	0,01	0,08	0,06	0,01	0,06
R24	0,11	0,10	-0,07	0,10	0,02	0,04	0,02	0,05	0,00	0,03	0,02	-0,02	0,19

Table XII – CCF Table on External Factors vs Passenger Market

Non-Seasonal			Seasonal			Evaluation	Error (MAPE)	Diagnostic (Ljung-Box at lag)					
p	d	q	P	D	Q	AIC	k-fold Cross Validated	6	12	18	24	30	36
0	1	2	1	0	1	-333,707	**8,650%**	0,801	0,941	0,978	0,889	0,929	0,951
2	1	0	2	0	2	-333,371	**8,661%**	0,725	0,941	0,984	0,766	0,815	0,899
1	1	1	1	0	1	-333,712	**8,732%**	0,741	0,908	0,962	0,878	0,921	0,942
0	1	2	1	0	2	-332,590	**8,836%**	0,782	0,937	0,973	0,860	0,899	0,899
0	1	1	1	0	1	-333,146	**8,938%**	0,442	0,750	0,880	0,835	0,898	0,924
0	1	2	2	0	2	-335,092	**8,947%**	0,739	0,930	0,982	0,806	0,848	0,909
1	1	2	2	0	2	-333,050	**9,055%**	0,738	0,935	0,984	0,807	0,848	0,913
1	1	1	1	0	2	-332,534	**9,073%**	0,686	0,883	0,945	0,829	0,878	0,874
0	1	2	2	0	1	-332,567	**9,178%**	0,796	0,942	0,976	0,873	0,909	0,899
0	1	1	2	0	2	-334,165	**9,636%**	0,362	0,692	0,885	0,723	0,787	0,871

Table XIII – TOP 10 ARIMA models overview – Passenger Market with external factors

9.11 Commercial Market

Non-Seasonal			Seasonal			Evaluation	Error (MAPE)	Diagnostic (Ljung-Box at lag)					
p	d	q	P	D	Q	AIC	k-fold Cross Validated	6	12	18	24	30	36
0	1	2	0	1	1	827,289	**8,287%**	0,254	0,639	0,875	0,895	0,773	0,889
1	1	1	0	1	1	827,288	**8,405%**	0,255	0,639	0,874	0,895	0,774	0,890
0	1	1	0	1	1	825,670	**8,465%**	0,258	0,656	0,883	0,907	0,774	0,892
0	1	1	0	1	2	826,962	**8,786%**	0,348	0,763	0,935	0,947	0,821	0,908
0	1	1	1	1	1	827,042	**8,793%**	0,336	0,753	0,933	0,945	0,816	0,904
0	1	2	1	1	2	828,286	**9,083%**	0,525	0,806	0,921	0,897	0,832	0,950
1	1	1	1	1	2	828,274	**9,175%**	0,527	0,806	0,920	0,896	0,831	0,950
0	1	1	1	1	2	826,580	**9,189%**	0,515	0,812	0,925	0,904	0,838	0,952
0	1	1	2	1	2	828,311	**9,666%**	0,488	0,801	0,887	0,832	0,790	0,929
0	1	1	2	1	0	828,236	**11,133%**	0,537	0,909	0,909	0,863	0,917	0,908

Table XIV – TOP 10 ARIMA models overview – Commercial Market

9.12 Forecasting Commercial Market with Outliers

Non-Seasonal			Seasonal			Evaluation	Error (MAPE)	Diagnostic (Ljung-Box at lag)					
p	d	q	P	D	Q	AIC	k-fold Cross Validated	6	12	18	24	30	36
1	1	0	0	1	2	679,492	**5,481%**	0,748	0,804	0,582	0,790	0,863	0,916
1	1	0	1	1	1	680,080	**5,830%**	0,755	0,788	0,561	0,747	0,826	0,901
1	1	0	2	1	1	679,438	**5,911%**	0,709	0,835	0,683	0,839	0,911	0,932
1	1	1	2	1	1	681,437	**5,920%**	0,707	0,835	0,685	0,840	0,911	0,933
1	1	0	0	1	1	678,876	**6,361%**	0,754	0,742	0,533	0,665	0,762	0,869
2	1	0	0	1	1	680,876	**6,361%**	0,754	0,742	0,533	0,665	0,762	0,870
1	1	1	0	1	1	680,876	**6,361%**	0,754	0,742	0,533	0,665	0,762	0,869
1	1	0	1	1	2	679,959	**6,463%**	0,779	0,805	0,654	0,837	0,912	0,952
1	1	0	2	1	2	681,354	**6,665%**	0,755	0,812	0,691	0,841	0,920	0,946
1	1	0	2	1	0	679,528	**6,669%**	0,692	0,783	0,645	0,758	0,877	0,888

Table XV – TOP 10 ARIMA models overview – Commercial Market with outliers

9.13 Forecasting Commercial Market with External Factors

	Lags												
Factors	-12	-11	-10	-9	-8	-7	-6	-5	-4	-3	-2	-1	0
R1	-0,06	-0,10	-0,09	-0,09	-0,13	-0,17	-0,16	-0,11	-0,10	-0,11	-0,11	-0,09	-0,07
R2	-0,20	-0,17	-0,13	-0,10	-0,06	-0,03	0,00	0,02	0,06	0,09	0,13	0,17	0,20
R3	-0,06	-0,02	0,02	0,09	0,10	0,07	0,00	0,02	0,02	0,05	0,08	0,12	0,18
R4	0,03	-0,03	0,00	-0,02	-0,06	-0,02	-0,01	-0,06	0,00	-0,04	-0,11	-0,12	-0,16
R5	-0,15	-0,02	-0,06	-0,09	0,03	0,02	0,03	-0,04	-0,05	-0,04	-0,03	-0,01	0,10
R6	0,19	0,14	0,07	0,07	-0,05	0,02	-0,03	0,02	0,02	-0,02	-0,04	0,02	-0,05
R7	0,23	0,00	-0,08	-0,17	0,03	-0,06	-0,01	-0,07	0,10	0,01	0,05	0,01	-0,03
R8	-0,05	-0,11	-0,02	0,10	0,15	-0,06	0,04	-0,05	0,09	0,00	0,09	-0,04	-0,10
R9	-0,10	0,00	0,06	0,11	0,07	0,01	0,01	0,06	0,13	0,02	-0,02	-0,10	-0,12
R10	-0,10	0,12	0,03	-0,05	-0,01	0,01	-0,04	0,01	-0,08	0,05	0,03	0,15	0,18
R11	-0,09	-0,02	-0,01	-0,06	-0,01	0,00	0,09	0,09	0,14	0,17	0,24	0,23	0,19
R12	0,09	0,10	0,09	0,13	0,10	0,10	0,05	0,10	0,07	0,09	0,05	0,09	0,04
R13	0,11	0,03	0,08	0,15	0,18	0,15	0,06	0,09	0,07	0,07	0,11	0,10	0,14
R14	0,06	0,12	0,14	0,23	0,16	0,18	0,17	0,15	0,14	0,15	0,13	0,15	0,09
R15	-0,13	0,20	-0,08	-0,22	-0,13	0,00	-0,05	-0,02	-0,03	0,01	0,01	0,07	0,13
R16	-0,13	0,15	0,07	0,03	0,04	-0,02	-0,06	0,06	-0,08	-0,06	-0,01	-0,08	0,08
R17	0,05	0,12	-0,08	0,20	0,11	0,09	0,02	0,08	0,10	0,19	0,06	0,15	0,10
R18	0,07	-0,03	-0,01	0,05	-0,06	-0,02	-0,01	0,03	0,03	-0,06	0,02	-0,06	0,09
R19	-0,08	-0,03	0,03	-0,07	0,04	0,07	-0,07	0,06	0,00	-0,05	0,03	-0,02	-0,04
R20	-0,02	0,03	-0,01	-0,05	-0,02	-0,10	-0,05	0,00	0,06	0,03	0,02	0,10	0,07
R21	-0,12	0,11	-0,11	0,05	-0,08	0,06	0,02	0,02	-0,10	-0,04	-0,02	0,05	0,17
R22	0,11	-0,01	-0,06	-0,01	0,09	0,05	0,09	-0,02	-0,05	-0,09	0,03	0,08	0,15
R23	0,11	0,03	-0,04	0,03	0,06	0,03	0,08	0,00	-0,04	-0,06	0,07	0,02	0,18
R24	0,08	0,02	-0,03	-0,07	0,10	0,08	0,12	0,03	-0,04	-0,15	0,06	0,03	0,21

Table XVI – CCF Table on External Factors vs Commercial Market

9.14 Forecasting Passenger Car Rental

Non-Seasonal			Seasonal			Evaluation	Error (MAPE)	Diagnostic (Ljung-Box at lag)					
p	d	q	P	D	Q	AIC	k-fold Cross Validated	6	12	18	24	30	36
0	1	1	0	1	1	1012,796	**11,098%**	0,605	0,761	0,841	0,924	0,853	0,870
0	1	1	0	1	2	1014,308	**11,918%**	0,633	0,832	0,911	0,964	0,927	0,933
0	1	1	1	1	1	1014,172	**12,318%**	0,639	0,848	0,926	0,971	0,940	0,945
2	1	1	0	1	1	1015,085	**12,566%**	0,763	0,820	0,884	0,936	0,910	0,906
0	1	2	0	1	1	1013,494	**12,773%**	0,726	0,793	0,863	0,928	0,902	0,901
1	1	1	0	1	1	1013,238	**12,979%**	0,747	0,803	0,872	0,930	0,911	0,907
0	1	2	0	1	2	1014,996	**13,895%**	0,760	0,859	0,924	0,970	0,960	0,958
1	1	1	0	1	2	1014,740	**14,137%**	0,784	0,869	0,930	0,972	0,966	0,962
0	1	2	1	1	1	1014,835	**14,372%**	0,772	0,878	0,939	0,978	0,971	0,969
1	1	1	1	1	1	1014,570	**14,647%**	0,798	0,889	0,946	0,980	0,976	0,973

Table XVII – TOP 10 ARIMA models overview – Passenger Car Rental

9.15 Forecasting Passenger Car Rental With Outliers

Non-Seasonal			Seasonal			Evaluation	Error (MAPE)	Diagnostic (Ljung-Box at lag)					
p	d	q	P	D	Q	AIC	k-fold Cross Validated	6	12	18	24	30	36
2	1	2	2	1	1	1009,289	**6,311%**	0,844	0,931	0,956	0,932	0,963	0,982
2	1	2	1	1	1	1008,559	**7,577%**	0,817	0,908	0,931	0,910	0,953	0,978
2	1	1	1	1	1	1008,966	**8,236%**	0,588	0,770	0,892	0,879	0,918	0,955
2	1	2	0	1	2	1008,947	**8,297%**	0,809	0,906	0,933	0,893	0,938	0,969
0	1	1	1	1	1	1009,138	**8,709%**	0,238	0,495	0,679	0,745	0,806	0,885
2	1	1	0	1	2	1008,977	**8,812%**	0,606	0,770	0,884	0,854	0,894	0,940
2	1	1	0	1	1	1007,452	**9,093%**	0,626	0,773	0,886	0,818	0,853	0,919
2	1	2	0	1	1	1007,663	**9,259%**	0,803	0,891	0,928	0,837	0,887	0,941
0	1	1	0	1	1	1008,400	**9,852%**	0,266	0,456	0,641	0,650	0,665	0,760
1	1	1	0	1	1	1009,053	**9,922%**	0,367	0,497	0,627	0,544	0,630	0,738

Table XVIII – TOP 10 ARIMA models overview – Passenger Car Rental with outliers

9.16 Forecasting Passenger Car Rental With External Factors

	Lags												
Factors	-12	-11	-10	-9	-8	-7	-6	-5	-4	-3	-2	-1	0
R1	-0,03	-0,03	-0,04	-0,06	-0,06	-0,10	-0,14	-0,15	-0,16	-0,14	-0,17	-0,13	-0,13
R2	-0,11	-0,09	-0,07	-0,05	-0,03	-0,01	0,01	0,02	0,06	0,09	0,12	0,15	0,16
R3	-0,08	-0,07	-0,06	-0,03	0,01	0,05	0,06	0,06	0,08	0,12	0,16	0,20	0,22
R4	0,08	0,02	0,05	0,01	0,00	-0,05	-0,11	-0,05	-0,08	-0,12	-0,10	-0,15	-0,18
R5	-0,12	-0,18	-0,23	-0,07	-0,14	-0,06	0,08	-0,03	-0,01	0,14	0,10	0,12	0,18
R6	0,09	0,10	0,11	0,01	0,01	0,02	0,00	-0,03	-0,04	-0,06	-0,11	-0,11	-0,09
R7	-0,12	-0,08	-0,11	-0,12	-0,14	-0,06	-0,01	0,03	0,13	0,22	0,16	0,17	0,14
R8	0,01	-0,14	-0,06	-0,18	-0,10	-0,07	-0,04	0,06	0,15	0,20	0,32	0,20	0,12
R9	0,06	-0,02	-0,01	-0,14	-0,16	-0,07	-0,11	0,00	0,10	0,16	0,20	0,11	0,13
R10	-0,08	-0,13	-0,07	0,00	-0,04	0,01	0,04	-0,03	0,03	0,14	0,13	0,16	0,17
R11	-0,06	-0,02	0,04	0,10	0,11	0,14	0,18	0,20	0,21	0,19	0,19	0,18	0,13
R12	0,07	0,10	0,09	0,11	0,12	0,12	0,14	0,13	0,15	0,14	0,09	0,09	0,05
R13	0,08	0,08	0,06	0,12	0,13	0,16	0,20	0,20	0,18	0,17	0,17	0,13	0,11
R14	0,07	0,09	0,09	0,13	0,12	0,13	0,17	0,17	0,18	0,17	0,13	0,11	0,08
R15	0,04	-0,05	0,02	0,10	0,02	0,05	0,13	-0,01	0,03	0,10	0,04	0,06	0,02
R16	-0,09	-0,13	-0,13	-0,11	-0,09	-0,02	0,02	-0,08	0,08	0,00	0,06	0,15	0,17
R17	0,06	0,11	0,20	0,14	0,21	0,20	0,15	0,20	0,17	0,11	0,14	0,13	0,05
R18	0,11	-0,18	0,08	-0,08	0,01	-0,03	0,09	-0,18	-0,01	0,03	0,01	-0,02	0,09
R19	0,03	0,03	-0,13	0,09	-0,06	0,04	0,02	0,14	-0,02	0,08	-0,06	0,03	-0,06
R20	-0,07	-0,09	-0,06	-0,02	-0,03	-0,02	0,00	-0,03	0,07	0,11	0,18	0,27	0,24
R21	-0,07	-0,10	-0,11	-0,05	-0,01	-0,07	0,03	0,07	0,09	0,11	0,15	0,13	0,20
R22	0,07	0,00	-0,04	0,06	0,02	0,04	0,11	0,22	0,09	-0,03	0,09	-0,01	0,09
R23	0,17	0,01	0,04	0,10	0,08	0,06	0,19	0,10	0,08	0,00	-0,07	0,06	-0,03
R24	0,04	0,07	-0,05	0,11	0,04	0,07	0,13	0,18	0,11	0,04	-0,01	0,06	0,05

Table XIX – CCF Table on External Factors vs Passenger Car Rental Market

Non-Seasonal			Seasonal			Evaluation	Error (MAPE)	Diagnostic (Ljung-Box at lag)					
p	d	q	P	D	Q	AIC	k-fold Cross Validated	6	12	18	24	30	36
2	1	1	1	1	1	1004,396	**3,314%**	0,704	0,819	0,884	0,820	0,826	0,868
2	1	2	1	1	1	1004,451	**3,315%**	0,867	0,914	0,922	0,841	0,879	0,916
1	1	1	1	1	1	1004,455	**3,446%**	0,436	0,522	0,655	0,630	0,690	0,737
0	1	1	1	1	1	1003,324	**3,510%**	0,363	0,580	0,713	0,723	0,760	0,802
2	1	1	0	1	1	1002,924	**3,671%**	0,720	0,821	0,889	0,781	0,768	0,822
0	1	1	0	1	2	1003,494	**3,675%**	0,367	0,581	0,710	0,708	0,740	0,779
2	1	2	0	1	1	1003,233	**3,834%**	0,864	0,906	0,922	0,779	0,801	0,851
1	1	1	0	1	1	1003,614	**4,001%**	0,423	0,464	0,623	0,516	0,520	0,546
0	1	2	0	1	1	1003,839	**4,030%**	0,406	0,479	0,635	0,552	0,550	0,576
0	1	1	0	1	1	1002,357	**4,062%**	0,355	0,521	0,664	0,628	0,617	0,650

Table XX – TOP 10 ARIMA models overview – Passenger Car Rental with external factors

9.17 Forecasting Passenger Free Market

Non-Seasonal			Seasonal			Evaluation	Error (MAPE)	Diagnostic (Ljung-Box at lag)					
p	d	q	P	D	Q	AIC	k-fold Cross Validated	6	12	18	24	30	36
0	1	1	2	0	2	-1285,171	**14,978%**	0,338	0,263	0,468	0,410	0,677	0,835
2	1	0	2	0	2	-1284,884	**15,908%**	0,457	0,397	0,655	0,533	0,783	0,907
2	1	1	1	0	1	-1286,143	**16,586%**	0,762	0,640	0,858	0,615	0,843	0,928
2	1	0	1	0	1	-1286,595	**16,601%**	0,483	0,422	0,693	0,526	0,776	0,874
2	1	2	2	0	2	-1287,005	**17,630%**	0,999	0,930	0,981	0,895	0,976	0,994
2	1	2	1	0	2	-1287,053	**18,079%**	0,999	0,953	0,988	0,912	0,980	0,993
2	1	2	2	0	1	-1286,980	**18,119%**	1,000	0,953	0,989	0,911	0,980	0,993
2	1	2	1	0	1	-1288,740	**18,126%**	1,000	0,943	0,987	0,898	0,976	0,992
0	1	1	1	0	1	-1286,219	**18,218%**	0,283	0,211	0,428	0,320	0,580	0,722
2	1	2	2	0	0	-1285,467	**19,335%**	0,999	0,948	0,988	0,874	0,972	0,979

Table XXI – TOP 10 ARIMA models overview – Passenger Free Market

9.18 Forecasting Passenger Free Market With Outliers

Non-Seasonal			Seasonal			Evaluation	Error (MAPE)	Diagnostic (Ljung-Box at lag)					
p	d	q	P	D	Q	AIC	k-fold Cross Validated	6	12	18	24	30	36
2	1	1	2	0	2	-1318,263	**10,259%**	0,554	0,463	0,651	0,439	0,276	0,396
2	1	2	2	0	2	-1320,263	**10,274%**	0,991	0,898	0,948	0,778	0,709	0,847
2	1	2	1	0	2	-1321,755	**11,799%**	0,988	0,875	0,924	0,702	0,619	0,764
2	1	1	1	0	2	-1319,478	**12,120%**	0,494	0,415	0,588	0,353	0,200	0,281
2	1	1	1	0	1	-1318,536	**12,510%**	0,366	0,262	0,437	0,189	0,099	0,163
2	1	2	2	0	1	-1321,213	**12,917%**	0,981	0,840	0,897	0,625	0,531	0,689
2	1	0	1	0	2	-1318,757	**13,047%**	0,228	0,206	0,365	0,215	0,123	0,164
2	1	0	1	0	1	-1318,645	**13,066%**	0,232	0,179	0,333	0,153	0,090	0,146
2	1	1	2	0	1	-1318,646	**13,165%**	0,417	0,336	0,513	0,255	0,128	0,191
2	1	2	1	0	1	-1321,180	**13,990%**	0,966	0,757	0,826	0,504	0,423	0,604

Table XXII – TOP 10 ARIMA models overview – Passenger Free Market with outliers

9.19 Forecasting Passenger Free Market With External Factors

Factors	Lags												
	-12	-11	-10	-9	-8	-7	-6	-5	-4	-3	-2	-1	0
R1	-0,05	0,02	0,07	0,01	-0,11	-0,27	-0,24	-0,21	-0,23	-0,26	-0,27	-0,21	-0,12
R2	-0,22	-0,16	-0,10	-0,05	0,01	0,04	0,06	0,08	0,12	0,16	0,21	0,25	0,26
R3	-0,07	-0,04	0,05	0,05	0,05	0,00	0,05	0,11	0,16	0,20	0,19	0,17	0,20
R4	0,05	0,00	0,01	0,14	-0,11	-0,04	0,10	0,04	-0,05	-0,16	-0,14	-0,03	-0,08
R5	-0,19	-0,03	-0,06	-0,16	-0,04	0,02	-0,14	-0,14	-0,07	0,03	0,01	-0,01	0,05
R6	0,18	0,05	-0,08	-0,17	-0,23	-0,06	0,20	0,22	0,11	-0,03	-0,01	0,10	0,05
R7	-0,01	-0,04	-0,12	-0,20	-0,11	-0,07	-0,05	0,07	0,10	0,07	0,11	0,07	0,02
R8	-0,18	-0,16	-0,01	0,05	0,00	-0,06	0,02	0,10	0,09	0,08	0,14	0,03	-0,18
R9	-0,09	-0,05	-0,02	-0,06	-0,07	-0,05	0,02	0,12	0,22	0,12	0,01	-0,03	-0,11
R10	-0,13	0,05	-0,01	-0,08	-0,03	0,01	-0,06	-0,18	-0,12	0,02	0,05	0,04	0,02
R11	-0,11	-0,03	-0,03	-0,01	0,02	0,06	0,13	0,13	0,15	0,21	0,22	0,20	0,17
R12	0,05	0,04	0,05	0,05	0,08	0,06	0,08	0,14	0,15	0,09	0,09	0,05	0,12
R13	-0,01	-0,19	-0,04	0,20	0,22	0,14	0,23	0,33	0,07	-0,03	0,18	0,22	0,08
R14	0,02	-0,06	-0,06	0,02	0,11	0,17	0,23	0,36	0,28	0,16	0,07	0,02	0,11
R15	-0,06	0,11	0,00	-0,10	0,03	0,04	-0,02	-0,18	-0,01	-0,01	0,03	-0,16	0,04
R16	-0,10	0,06	-0,04	0,12	0,07	-0,15	-0,14	-0,06	-0,09	0,03	0,09	-0,05	0,01
R17	0,02	-0,02	-0,01	0,12	0,10	0,10	0,00	0,00	0,21	0,18	0,07	0,19	0,23
R18	0,07	0,10	-0,02	-0,10	-0,12	0,09	0,14	-0,02	-0,17	-0,01	-0,06	0,14	0,17
R19	-0,10	-0,05	-0,04	0,03	0,13	-0,15	0,03	0,08	0,00	0,02	0,05	-0,18	0,04
R20	0,02	0,04	-0,07	-0,08	-0,09	-0,15	-0,12	0,02	0,05	0,07	0,05	0,10	0,20
R21	-0,08	-0,04	-0,04	-0,18	-0,19	-0,11	-0,05	-0,04	0,08	0,06	0,07	0,13	0,12
R22	0,08	-0,13	-0,14	0,07	0,07	0,02	0,13	0,14	-0,16	-0,19	0,06	0,22	0,07
R23	0,09	-0,06	-0,14	0,04	0,04	0,06	0,16	0,15	-0,19	-0,18	0,08	0,17	0,12
R24	0,15	-0,09	-0,17	0,04	0,06	0,04	0,15	0,13	-0,13	-0,17	0,00	0,18	0,18

Table XXIII – CCF Table on External Factors vs Passenger Free Market

Non-Seasonal			Seasonal			Evaluation	Error (MAPE)	Diagnostic (Ljung-Box at lag)					
p	d	q	P	D	Q	AIC	k-fold Cross Validated	6	12	18	24	30	36
2	1	2	2	0	2	-1234,760	**7,323%**	0,986	0,694	0,841	0,567	0,516	0,719
2	1	1	2	0	2	-1231,292	**7,951%**	0,507	0,233	0,411	0,213	0,116	0,225
2	1	0	2	0	2	-1231,414	**7,986%**	0,368	0,173	0,311	0,165	0,097	0,176
2	1	2	1	0	2	-1235,464	**10,267%**	0,984	0,620	0,736	0,384	0,346	0,485
2	1	2	1	0	1	-1235,850	**10,326%**	0,947	0,571	0,715	0,340	0,305	0,459
2	1	2	2	0	1	-1235,038	**10,930%**	0,977	0,602	0,725	0,356	0,318	0,450
2	1	1	1	0	2	-1232,103	**11,345%**	0,483	0,210	0,332	0,142	0,071	0,116
2	1	2	1	0	0	-1230,402	**13,877%**	0,996	0,688	0,724	0,351	0,312	0,350
2	1	1	1	0	0	-1228,410	**13,903%**	0,616	0,330	0,392	0,140	0,073	0,076
2	1	2	2	0	0	-1230,405	**13,992%**	0,978	0,578	0,651	0,250	0,220	0,278

Table XXIV – TOP 10 ARIMA models overview – Passenger Free Market with external factors