BMW PROJECT

Lotti Jacopo, Napoli Andrea, Palmentiero Luca

May 29, 2024

1 Introduction

We are a specialized business unit in forecasting and data analysis, directly tasked by BMW's Board to conduct forecasts for the next year. Our focus lies on identifying the market leader among the key competitors in the European automotive sector, particularly Mercedes and Audi, and forecasting the number of BMW car registrations in the upcoming year. We are committed to providing accurate and detailed forecasts that will support BMW's strategic decisions in the coming year. In addition, one of our main goal is to provide the Board with a comprehensive analysis of the macroeconomic variables that most significantly influence the number of new vehicle registrations for each country in which we are operating, thus offering a detailed and strategic overview for business decisions.

2 Problem statement 1

BMW's Board has requested the development of a predictive model to assess projected car registrations for the following year in the European countries where the company operates. This model integrates the most relevant key macroeconomic variables influencing car demand in the European market. Production planning is crucial to optimize resources and ensure market demand satisfaction. The ability to accurately anticipate future demand enables BMW to efficiently plan production, avoiding inventory excesses or shortages that could result in additional costs or missed opportunities. Furthermore, accurate planning allows the company to optimize its supply chain and production, ensuring a balanced distribution of resources and greater operational flexibility to promptly respond to market demand fluctuations.

2.1 Dataset

In order to build the models, we have decided to gather macroeconomic data for 29 European Countries from 1990 to 2021, as shown in Table 1. Data have been gathered from Eurostat and the European Central Bank Portal. Within the dataset, the statistical units consist of each single Country considered for each single Year of our interest (for example: Austria-2014, Austria-2015,...).

Variable Name	Description
`cy`	Country and Year of the values
`cr`	Number of car registration in that year and country
`pop`	Number of persons having their usual residence in a country on 1 January of the respective year
`gdp`	Gross Domestic Product (GDP)
`coe`	Compensation of employees
`fc`	Private consumption expenditure
`rd`	Number of fatalities caused by road accidents
`el`	Early leavers (Population aged 18-24) from education and training
`mar`	The crude marriage rate (ratio of the number of marriages to the average population in that year)
`le`	Life expectancy at birth
`fert`	Total fertility rate
`for`	Number of persons born abroad, who are usually resident in the reporting country
`or`	Old-age-dependency ratio
`dens`	Population density
`rgdp`	Real GDP per capita
`imp`	Import of goods and services
`exp`	Export of goods and services

Variable Name	Description
at	Air transport of passengers by country
eal1	Share of people with Less than primary, primary and lower secondary education (levels 0-2)
eal2	Share of people with Upper secondary and post-secondary non-tertiary education (levels 3 and 4)
eal3	Share of people with Tertiary education (levels 5-8)
hdi	Human development index
eyos	Expected years of schooling
myos	Mean years of schooling
gnipc	Gross National Income Per capita
gdi	Gender Development Index
gii	Gender Inequality Index
mmr	Maternal Mortality Ratio
abr	Adolescent Birth Ratio
co2	co2 production per capita
mfpc	Material footprint per capita

Table 1: Set of variables in the cross-section dataset

2.2 Method

It has been decided to proceed following two paths, the first one is to apply the logarithmic transformation to the macro economical variables which presents the higher variability (according to the coefficient of variation), in order to delete the differences that stands between different countries, the second one is to consider the data as it is ignoring these differences.
The evaluation of the best model fitted on the test take place using the Mean Squared Error (MSE), because this measure allows the comparison between the real value of the dependent variable and the predicted values obtained through OLS regression, not taking into consideration the different dimensions of the models. Between the models created, some has been created using cross-validation techniques, such as best subset selection, some using statistical techniques, such as Lasso regression, and other using some macroeconomic intuitions, in order to give a different perspective from the statistical selection of the software.
The models in order to be considered correctly specified have to satisfy the Anova test, a measure that test the null hypothesis of non-significance of all the regressors that has been omitted from the model, and the F-statistic, a measure that indicates the significance of model.
In order to start the analysis the dataset has been divided in two halves, train and test, then using the techniques previously mentioned the models has been created on the train dataset; firstly these models have to refuse the non-significance of the model, tested through the F-statistic, then the models that succeeded undergo the Anova test.
The model that overcome this first two test are correctly specified and can now be evaluated on their performance on the train: these evaluation take places using the MSE and the adjusted R^2; these four models emerges to be the best:

Model Type	Model	Specification
non-log	(I)	coe, fc, el, rgdp, imp, exp, at, eal1, eal2, eal3, eyos, gnipc, gdi, gii, mmr, co2, mfpc
non-log	(II)	pop, gdp, coe, fc, rd, co2
log	(III)	log(coe), log(fc), el, le, log(fr), or, rgdp, log(imp), log(exp), eal2, eal2, hdi, eyos, myos, gdi, gii, log(abr,) co2
log	(IV)	log(pop), log(gdp), log(coe), log(fr), or, log(dens), rgdp, log(imp), log(exp), eal1, eal2, hdi, eyos, myos, gdi, gii, mmr, log(abr,) co2

Table 2: **Model Specification**

2.3 Analysis

Once found the best models based on the train it is possible to recreate the same models on the test in order to validate on a different dataset what the train measures suggests. The performance of the models are evaluated using the MSE:

Model	MSE Train	AdjR2	MSE Test
(I)	14,150,595,302	0.9738	26,362,944,877
(II)	21,543,655,213	0.9643	36,328,891,512
(III)	63,901,656,667	0.8807	1.54962e+11
(IV)	61,217,163,482	0.8845	147,317,726,713

Table 3: **Model Evaluation**

As it is possible to see from Table 3 the best model, highlighted, appears to be the model (I), using the MSE Test metric, that confirms the computation of the MSE Train and the AdjR2 train, where the model (I) resulted the best one as well.

$$CR = \beta_0 + \beta_1 \texttt{coe} + \beta_2 \texttt{fc} + \beta_3 \texttt{el} + \beta_4 \texttt{rgdp} + \beta_5 \texttt{imp} + \beta_6 \texttt{exp} + \beta_7 \texttt{at} + \beta_8 \texttt{eal1} + \beta_9 \texttt{eal2} + \beta_{10} \texttt{eal3} + \beta_{11} \texttt{eyos} + \beta_{12} \texttt{gnipc} + \beta_{13} \texttt{gdi} + \beta_{14} \texttt{gii} + \beta_{15} \texttt{mmr} + \beta_{16} \texttt{co2} + \beta_{17} \texttt{mfpc} + \epsilon$$

For what concerns the interpretation of the model it is possible to say that intercept of the model stands at -5.327e+07, i.e. if all the values of all the regressors are 0 the number of car registration is negative, while among the regressors that affect the number of car registrations the most there are

- : `gdi`, each unitary increment of the measure lead to a positive increase of the `CR` by 4.936e+06 car
- `gii`, each unitary increment of the measure lead to a positive increase of the `CR` by 2.090e+06 car
- `eal1`, each unitary increment of the measure lead to a positive increase of the `CR` by 4.780e+05 car
- `eal2`, each unitary increment of the measure lead to a positive increase of the `CR` by 4.688e+05 car
- `eal3`, each unitary increment of the measure lead to a positive increase of the `CR` by 4.531e+05 car

It is interesting to notice how the `CR` increase of a similar value independently of which of the 3 `eal1`, `eal2`, or `eal3` increase, i.e. there is not a great difference of the incidence of the education level on the `CR`. All the previous variables have a positive correlation with the dependent variable as it is possible to expect.
On the other hand among the regressors that affect the number of car registration the least can be found:

- `at`, each unitary increment of the measure lead to a positive increase of the `CR` by 5.921e-04 car (note: the variable is not significant inside the model, having a p-value over 0.1)
- `fc`, each unitary increment of the measure lead to a positive increase of the `CR` by 9.208e-01 car
- `coe`, each unitary increment of the measure lead to a positive increase of the `CR` by 1.250 car
- `imp`, each unitary increment of the measure lead to a negative increase of the `CR` by -2.115 car
- `exp`, each unitary increment of the measure lead to a positive increase of the `CR` by 1.596 car

It is interesting to notice how the `imp` variable affects negatively the dependent variable while the `exp` variable affects positively.

3 Problem statement 2

In the European automotive industry, strategic planning is crucial for business success. Our goal is to forecast BMW's car registrations in 2020 across 15 European countries and we do so through two approaches.

1. First Approach: BMW Car Registrations Forecast.
   We examine the historical series of BMW registrations to understand the brand's past performance and inform our future decisions.

2. Second Approach: Total Registrations and BMW Market Share Forecast.
   We forecast the total car registrations in the 15 selected countries in 2020 and the BMW's market share. By simply multiplying the total registrations by our market share, we obtain an estimate of BMW car registrations.

Both approaches converge towards a single objective: to accurately plan raw material procurement (such as Steel, nickel, copper, aluminum, glass, plastic and rubber) and efficiently manage human resources. Simultaneously, we carefully consider the need for qualified personnel and the distribution of human resources to ensure that our team is adequately prepared to meet the challenges of the European automotive market. Ultimately, both approaches aim to equip our production to effectively respond to the changing demands of the market.

3.1 Dataset

The data set "Dataset_time_series" contains monthly data about BMW Car Registrations among 15 selected European countries, the Total Number of Car Registrations in Europe and the BMW Market Share. All these data refer to time span from 1990 to 2019, in order to avoid the use of strange data derived from the pandemic disease.
The data have been gathered from ACEA.com, the European Automobile Manufacturers' Association.

Variable Name	**Description**	**Type**
`Date`	Date (mm-yyyy)	Date
`Total_Month`	Total Car Registration in Europe per Month	Integer
`BMW_Month`	BMW Car Registrations per Month	Integer
`Market_Share_Month`	BMW Market Share per Month	Integer

Table 4: Set of variables in the time-series dataset

3.2 First Approach: BMW Car Registrations (BMW)

3.2.1 Method

The `BMW` time series has been split into training set (first 25 years, 1990-2014, 83,3% of the observation) to build the model and the test set (last 5 years, 2015-2019, 16,7% of the observation) to test it. Since dealing with monthly data and aiming to forecast BMW new car registrations for 2020, an h-step ahead forecast is planned with h = 12 months. Following some exploratory analysis, a decision was made to apply logarithmic transformation to the time series to stabilize it. Subsequently, regression of seasonality and trend on the time series revealed both to be significantly different from 0.
As shown in Figure 1, there's a strong seasonality, which can be also explained in economic terms: car registrations peak at quarter-end due to tax incentives, marketing strategies, and accounting deadlines. Conversely, August sees lower registrations owing to summer vacations, reduced business activity, and prioritization of holiday savings over major purchases.

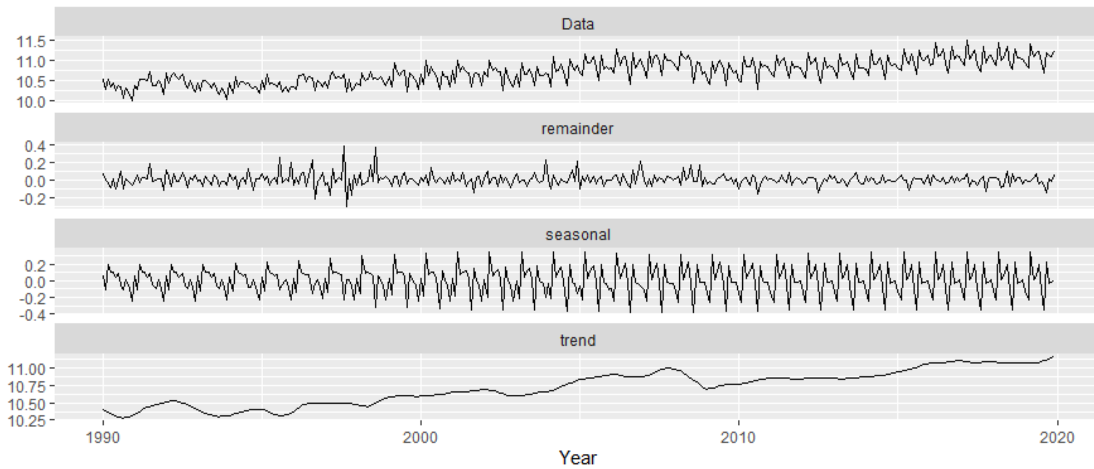

Figure 1: Modern Decomposition of the `log(BMW)` time series.

A Dickey-Fuller test was then conducted to examine non-stationarity. Rejecting the null hypothesis at a 1% confidence level (Dickey-Fuller = -5.6362), it was concluded that the time series is stationary. Afterward, the process went on by creating various ARIMA models, noticing initially unconvincing autocorrelation function (acf) and partial autocorrelation function (pacf). Further investigation into the seasonal component was pursued, leading to an initial differentiation that improved both acf and pacf.

The decision was then made to estimate all possible SARIMA(p,0,q)(P,1,Q) models with p, q, P, Q between 0 and 2. The four best models selected were:

1. log(BMW) = SARIMA(1,0,2)(0,1,1) (based on BIC nd AIC)
2. log(BMW) = SARIMA(1,0,2)(2,1,2) (according to MSE)
3. log(BMW) = SARIMA(2,0,2)(0,1,1) (according to MAE)
4. log(BMW) = SARIMA(0,0,0)(0,1,0).

Continuing in the same manner, ARMA(p,0,q) models with deterministic trend and seasonality were estimated in all possible combinations with p and q between 0 and 2. The four best models chosen were:

1. log(BMW) = ARMA(2,2) (according to BIC, MSE, and AIC)
2. log(BMW) = ARMA(0,0)
3. log(BMW) = AR(1)
4. log(BMW) = MA(1)

Subsequently, these eight selected models were estimated in the training set.

3.2.2 Analysis

We now have 8 models that we are going to test out of sample. We need to remember what our goal is, which is to predict the next 12 months. So wanting to predict 12 observations, and having 60 out-of-sample observations, we are going to make 49 predictions for each model. To compare the various models we use the classical MSE, and also a measure we have called it TOTMSE, which is the square of the difference between the sum of the predicted and the sum of the actual values.

Model	Description	Mse	TotMSe
model1	ARMA(2,2) with deterministic trend and seasonalities	0.012137219	0.2291949
model2	AR(1) with deterministic trend and seasonalities	0.012562148	0.2927581
model3	MA(1) with deterministic trend and seasonalities	0.012545757	0.2961513
model4	SARIMA(1,0,2)(2,1,2)	0.005291181	0.2967404
model5	Deterministic trend and seasonalities	0.012510918	0.3005084
model6	SARIMA(1,0,2)(0,1,1)	0.005380777	0.3063123
model7	SARIMA(2,0,2)(0,1,1)	0.005390541	0.3074772
model8	SARIMA(0,0,0)(0,1,0)	0.007634504	0.5304835

Table 5: Model descriptions for BMW and corresponding MSE and Total MSE values

As shown in Table 4, the best model according to MSE is the SARIMA(1,0,2)(2,1,2), and the best one according to TOTMSE is the ARMA(2,2) with deterministic trend and seasonalities.

3.3 Second Approach: Total Car Registration (tcr) and BMW's Market Share (ms)

3.3.1 Method

The `tcr` and the `ms` time series were treated exactly the same way as the `BMW` one, and so they have been split into training set (first 25 years) to build the model and the test set (last 5 years) to test it. In the previous approach, an h-step ahead forecast was performed, where h = 12 months for both variables, `tcr` and `ms`. After conducting exploratory analysis, it was decided to apply the logarithm to the time series to stabilize them. Subsequently, regression analysis was conducted to model the seasonality and trend components of the time series. In figure 2 it's possible to notice that for the `tcr` variable, only the seasonalities were found to be significant.

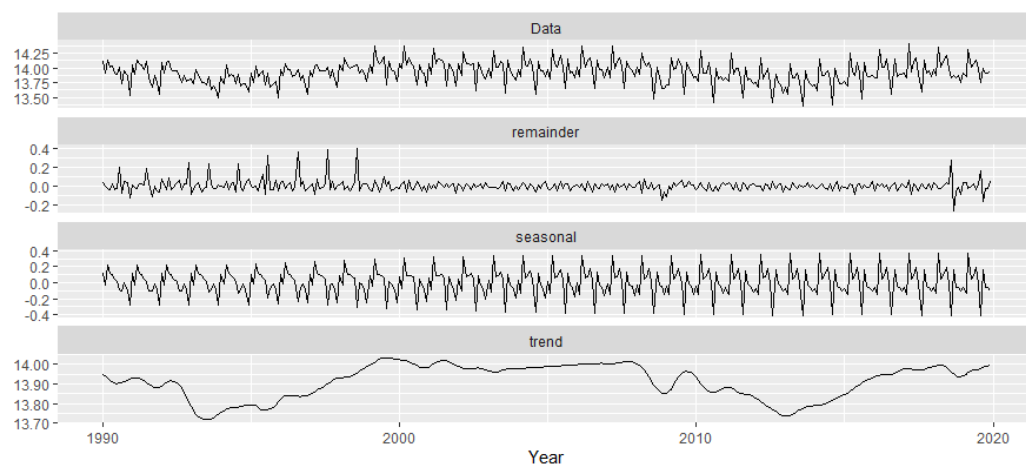

Figure 2: Modern Decomposition of the `log(tcr)` time series.

Instead, the results for the `ms` variable indicated that both trend and seasonalities were found to be significant, as shown in Figure 3.

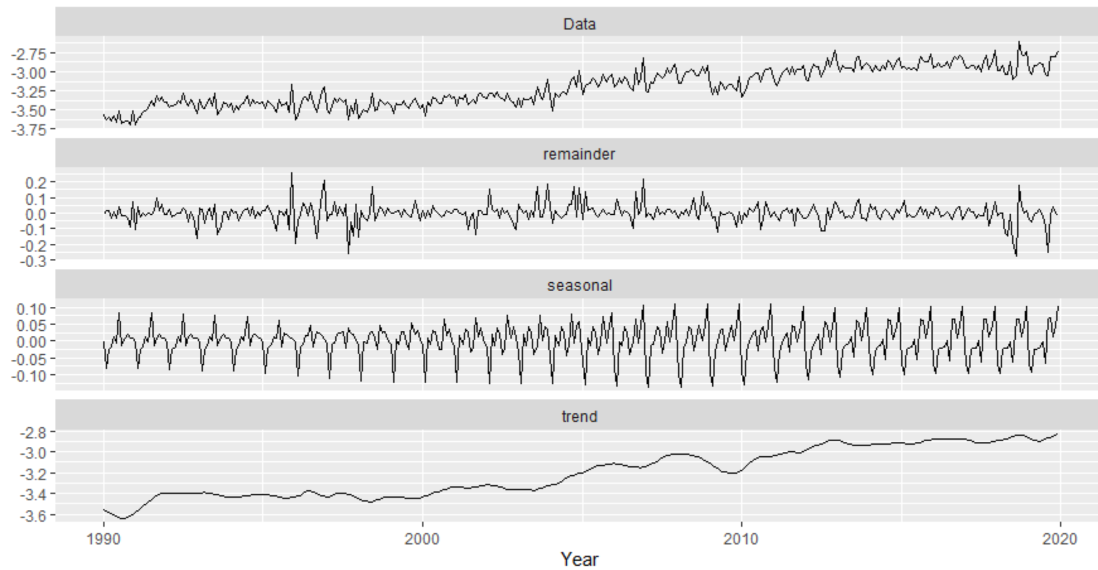

Figure 3: Modern Decomposition of the `log(ms)` time series.

A Dickey Fuller test was then performed to test for non-stationarity, and the null hypothesis was rejected at a 1% confidence level for both `tcr` (Dickey-Fuller = -5.3732) and `ms` (Dickey-Fuller = -4.0215), confirming that the time series are stationary. Following the initial investigation into ARIMA and ARMA models, it was observed that the autocorrelation function (acf) and partial autocorrelation function (pacf) were not convincing. Therefore, the seasonal component was differentiated, which resulted in an improvement in both acf and pacf.
Subsequently, the estimation of all possible SARIMA(p,0,q)(P,1,Q) models was carried out, with p, q, P, and Q ranging between 0 and 2. The four best models selected were:

1. log(tcr) = SARIMA(1,0,2)(0,1,1) (according to BIC and AIC)
2. log(tcr) = SARIMA(1,0,2)(2,1,2) (according to MSE and MAE)
3. log(tcr) = SARIMA(1,0,1)(1,1,1)
4. log(tcr) = SARIMA(0,0,0)(0,1,0).

Similarly, ARMA(p,0,q) models with deterministic seasonality were estimated in all possible combinations with p and q between 0 and 2. The four best models chosen were:

1. log(tcr) = ARMA(2,2) (according to BIC, MSE, and AIC)
2. log(tcr) = ARMA(0,0)
3. log(tcr) = AR(1)
4. log(tcr) = MA(1)

The process was the same for the MS time series, the four best SARIMA(p,0,q)(P,1,Q) models were:

1. log(MS) = SARIMA(1,0,1)(0,1,1) (according to BIC)
2. log(MS) = SARIMA(1,0,2)(2,1,2) (according to MSE)
3. log(MS) = SARIMA(1,0,2)(1,1,1) (according to AIC)
4. log(MS) = SARIMA(0,0,0)(0,1,0).

And the best four ARMA(p,0,q) models with deterministic seasonality and trend were:

1. log(MS) = ARMA(1,1) (according to BIC and AIC)

2. log(MS) = ARMA(1,2) (according to MSE)
3. log(MS) = ARMA(0,0)
4. log(MS) = AR(1)

3.3.2 Analysis

With 8 models selected for out-of-sample testing, our objective remains predicting the next 12 months. Given 60 out-of-sample data points and the goal of predict 12 observations, there will be 49 predictions for each model. These models will be evaluated using classical Mean Squared Error (MSE) and a measure we've termed TOTMSE, which represents the square of the difference between the sum of the predicted and the sum of the actual values.

Model	Description	Mse	TotMSe
model1	ARMA(2,2) with deterministic trend and seasonalities	0.008807635	0.2083681
model2	SARIMA(1,0,1)(1,1,1)	0.006597129	0.2586430
model3	SARIMA(1,0,2)(2,1,2)	0.006961772	0.2980900
model4	SARIMA(1,0,2)(0,1,1)	0.007273964	0.3314089
model5	AR(1) with deterministic trend and seasonalities	0.009971475	0.3424518
model6	SARIMA(0,0,0)(0,1,0)	0.007440133	0.3518921
model7	MA(1) with deterministic trend and seasonalities	0.009971378	0.3547937
model8	Deterministic trend and seasonalities	0.009927713	0.3629886

Table 6: Model descriptions and corresponding MSE and Total MSE values for `tcr`

Model	Description	Mse	TotMSe
model1	SARIMA(0,0,0)(0,1,0)	0.005970963	0.1008611
model2	SARIMA(1,0,1)(0,1,1)	0.005483325	0.1492419
model3	SARIMA(1,0,2)(2,1,2)	0.005495004	0.1703999
model4	SARIMA(1,0,2)(1,1,1)	0.005747209	0.1793806
model5	ARMA(1,1) with deterministic trend and seasonalities	0.007596951	0.2084906
model6	ARMA(1,2) with deterministic trend and seasonalities	0.007719205	0.2183378
model7	AR(1) with deterministic trend and seasonalities	0.008953511	0.4087127
model8	Deterministic trend and seasonalities	0.009177279	0.4568828

Table 7: Model descriptions and corresponding MSE and Total MSE values for `ms`

As shown in Table 6, the best model for `tcr` according to MSE is the SARIMA(1,0,1)(1,1,1), while the best one according to TOTMSE is the ARMA(2,2) with deterministic trend and seasonalities. Moreover, as we can see in Table 7, the best model for `ms` according to MSE is the SARIMA(1,0,1)(0,1,1), while the best one according to TOTMSE is the SARIMA(0,0,0)(0,1,0) .

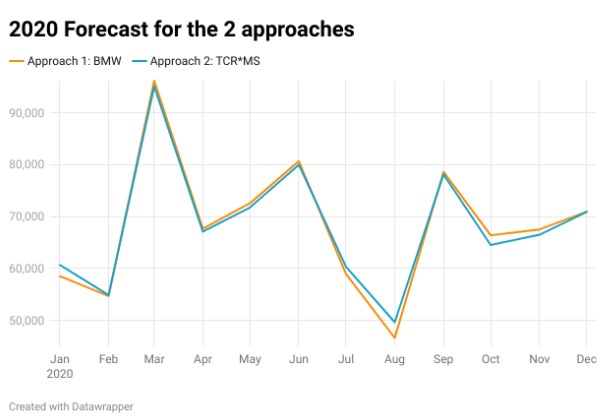

Figure 4: Comparison between the 2 Approaches for 2020 forecast.

Figure 4 compares the predictions obtained from the best models (selected according to the MSE) of the two different approaches: it's possible to see that these two models slightly differ from each other and consequently it's possible to rely on the consistency of our estimation.

4 Problem statement 3

Our business analytics team has been committed directly by BMW's Board to predict whether BMW will be the market leader among the three major European premium brands - Mercedes, Audi, and BMW - for all months of 2020. This prediction is based on a historical dataset of car registrations in Europe from 1990 to 2019 for each of the mentioned brands. This approach is grounded in the understanding that the top performer among various contenders possesses higher market power than its direct competitors, that is reflected in the pricing strategy.
Therefore, our analysis aims to provide an accurate and reliable forecast on whether BMW will emerge as the market leader in different months of 2020, thereby providing BMW with a solid foundation for its pricing and business strategic decisions.

4.1 Dataset

As shown in Table 8, it has been collected data from ACEA.com, the European Automobile Manufacturers' Association regarding monthly car registrations of BMW and the 2 main competitors (Audi and Mercedes) and then they have been compared. After that, in order to proceed with our assessments, for each month we built a binary variable (`Leader`) that returns "1" when BMW is the market leader and "0" when it isn't.

Variable Name	Description	Type
`Date`	Date (mm-yyyy)	Date
`BMW`	BMW Monthly Car Registration in Europe	Integer
`Mercedes`	Mercedes Monthly Car Registration in Europe	Integer
`Audi`	Audi Monthly Car Registration in Europe	Integer
`Leader_Binary`	European Car Market Leader as a categories (1: BMW, 0: Others)	Integer

Table 8: Set of variables in the classification dataset

4.2 Method

Since the variable `leader` is a time series object, a few exploratory analyses were performed (referencing the R code provided with the paper). These analyses revealed that the time series exhibits several significant seasonalities but not a significant trend. To verify this observation, the autocorrelation function (ACF) and partial autocorrelation function (PACF) were evaluated, and both showed significant autocorrelation only for the first lags. Subsequently, a Dickey-Fuller Test was conducted to assess the non-stationarity of the time series. The test produced a value of -4.7597, which leads to the rejection of the null hypothesis of non-stationarity ($H_0 : \beta_1 = 1$), indicating that `leader` is stationary. Since the dependent variable is binary, several Probit models were specified. These models take into account the seasonal component (S_t), the lag value of a period (Y_{t-1}), and the lag value of a period (Y_{t-12}). Specifically, considering the dependent variable `leader` as distributed with a Bernoulli distribution with probability p_t ($Leader_t \sim \text{Be}(p_t)$), it is known that $p_t = \Phi(X_t'\beta)$, where Φ is the cumulative distribution function of the standard normal distribution. If the probability is specified as $p_t = \Phi(\pi_t)$, then the models can be written as $\pi_t = \beta_0 + \beta_1 X_1 + \ldots + \beta_n X_n$, where the vector of explanatory variables X_t includes the aforementioned components. The models specified are as follows:

model 1: $Leader_t = \Phi(\beta_0 + \beta_1 Leader_{t-1} + \epsilon_t)$
model 2: $Leader_t = \Phi(\beta_0 + \beta_1 Leader_{t-12} + \epsilon_t)$
model 3: $Leader_t = \Phi(\beta_0 + \beta_1 S_t + \epsilon_t)$
model 4: $Leader_t = \Phi(\beta_0 + \beta_1 S_t + \beta_2 Leader_{t-12} + \epsilon_t)$
model 5: $Leader_t = \Phi(\beta_0 + \beta_1 S_t + \beta_2 Leader_{t-1} + \epsilon_t)$

To compare the various models, Accuracy and Precision measurements are used. These measurements are defined as follows:

		Predicted	
		N (0)	**P (1)**
Observed	**N (0)**	True Negative (TN)	False Positive (FP)
	P (1)	False Negative (FN)	True Positive (TP)

Table 9: Confusion Matrix

$$\textbf{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \qquad \textbf{Precision} = \frac{TP}{TP + FP}$$

Formula 1 and Formula 2: Accuracy and Precision Measurements.

In particular, accuracy is defined as the ratio of the number of correct predictions (both positive and negative) to the total number of predictions, as shown in Formula 1. It is a simple and intuitive metric that measures the overall ability of the model to correctly classify instances. However, accuracy can be misleading when the data is imbalanced, meaning one class is much more frequent than the other. The `leader` time series in fact has 63 values of 1 (17.5% of the total 360 observations) and 297 values of 0 (82.5%). For this reason, the Precision metric was used, which is the ratio of true positives (TP) to the sum of true positives and false positives (FP), as shown in Formula 2, i.e. the total values that are predicted to be positive. This metric measures the quality of the model's positive predictions, i.e., how many of those classified as positive are actually positive. Precision is particularly useful in scenarios where the cost of false positives is high. In this case, if it is predicted that the company will be the market leader, a specific pricing policy will be implemented, which would have very negative effects if it turns out to be a false positive.

4.3 Analysis

There are now 5 models to be tested out of sample. It is important to remember the goal, which is to predict if BMW will be the leader among these 3 brands in the next 12 months. To predict 12 observations and having 60 out-of-sample observations, 49 predictions will be made for each model.

Model	**Description**	**Accuracy**	**Precision**
Model 1	Leader_t $= \Phi(\beta_0 + \beta_1 \text{Leader}_{t-1} + \epsilon_t)$	0.7194	0.0000
Model 2	Leader_t $= \Phi(\beta_0 + \beta_1 \text{Leader}_{t-12} + \epsilon_t)$	0.8231	0.3649
Model 3	Leader_t $= \Phi(\beta_0 + \beta_1 S_t + \epsilon_t)$	0.9490	0.8488
Model 4	Leader_t $= \Phi(\beta_0 + \beta_1 S_t + \beta_2 \text{Leader}_{t-12} + \epsilon_t)$	0.9235	0.7558
Model 5	Leader_t $= \Phi(\beta_0 + \beta_1 S_t + \beta_2 \text{Leader}_{t-1} + \epsilon_t)$	0.8282	0.5791

Table 10: Model descriptions with their accuracy and precision values.

By looking at the Table 10, the best model, according to both the Accuracy and the Precision measures, is the Model 3, i.e. the one that takes into account the seasonal component.