UNIVERSIDADE D
COIMBRA

Carolina Ângelo Carminé

The Relationship Between the Economy and Air Transport in the European Union Countries and Southwest Regions

Dissertação de Mestrado Integrado em Engenharia Civil, na área de Especialização em Urbanismo, Transportes e Vias de Comunicação, orientada pelo Professor Doutor António José Pais Antunes e apresentada ao Departamento de Engenharia Civil da Faculdade de Ciências e Tecnologia da Universidade de Coimbra.

Maio de 2019

FCTUC DEPARTAMENTO DE ENGENHARIA CIVIL
FACULDADE DE CIÊNCIAS E TECNOLOGIA
UNIVERSIDADE DE COIMBRA

Carolina Ângelo Carminé

THE RELATIONSHIP BETWEEN THE ECONOMY AND AIR TRANSPORT IN THE EUROPEAN UNION COUNTRIES AND SOUTHWEST REGIONS

A RELAÇÃO ENTRE A ECONOMIA E TRANSORTE AÉREO NOS PAÍSES DA UNIÃO EUROPEIA E NAS REGIÕES DO SUDOESTE

Dissertação de Mestrado Integrado em Engenharia Civil, na área de Especialização em Urbanismo, Transporte e Vias de Comunicação, orientada pelo Professor Doutor António José Pais Antunes

Coimbra, Maio de 2019

ACKNOWLEDGMENTS

This dissertation marks the end of my academic track, and I would like to express the deepest appreciation to people who encouraged and supported me in carrying out this work.

Special thanks to Professor António Pais Antunes for the useful comments and remarks through all the learning process of this master thesis and for all his availability through these months of work.

To my family, colleagues and friends.

To all, a big thank you.

ABSTRACT

The analysis available about the relationship between economic activity and the number of air passengers clearly indicates that this number will grow in response to the increases on GDP, whereas in the case of cargo, the opposite happens, which means that the increase in volume (and value) of cargo anticipates the increases on GDP. The conclusions above are valid when one considers the world economy as a whole or the continental economies, but it is not clear that the same is true when considering national or regional economies. The purpose of this dissertation is to study the relationship between economy and air transport, passenger and cargo, in the European Union countries, with particular emphasis on Portugal, and in the NUTS 2 Europe Southwest regions (Spain, France and Italy, in addition to Portugal). One aspect that makes this topic especially interesting is that the evolution of the number of air passengers or the volume of air cargo in a country or region is supposed to be noticed, not only by the economy of the country and the region, but also by the economies of neighboring countries and regions - especially those with which the trade (of goods and services) is most intense. In particular, it is proposed to analyze whether the relationships are similar for all countries and regions or whether there are patterns depending on their characteristics. In order to carry out this study, we will use correlation and regression analysis techniques and clusters analysis methods, and the data available in the Eurostat database. The results showed that the relationship between the economy and air transport exists in some countries and regions, although the level of significance variates in each country and region without a geographic pattern.

Key words: Air Transport, GDP, Linear Regression, European Union.

RESUMO

As análises disponíveis sobre a relação entre o nível de atividade económica e o número de passageiros aéreos apontam claramente no sentido deste número crescer em resposta a aumentos do Produto Interno Bruto, enquanto que no caso da carga acontece o contrário, ou seja, o aumento do volume (e valor) da carga antecipa aumentos do PIB. As conclusões anteriores são válidas quando se considera a economia mundial como um todo ou as economias continentais, mas não é claro que o mesmo aconteça quando se consideram economias nacionais ou regionais. Assim, a presente dissertação tem por objeto o estudo da relação entre economia e transporte aéreo, de passageiros e de carga, nos países da União Europeia, com especial ênfase em Portugal, e nas regiões NUTS 2 do Sudoeste Europeu (Espanha, França e Itália, para além de Portugal). Um aspeto que torna este tema especialmente interessante decorre do facto da evolução do número de passageiros aéreos ou do volume de carga aérea de um país ou região ter em princípio a ver, não apenas com a economia do país e da região, mas também com a economia dos países e regiões vizinhos – sobretudo a daqueles com que o comércio (de bens e serviços) é mais intenso. Em particular, pretende-se analisar se as relações são semelhantes para todos os países e regiões ou se existem padrões em função das respetivas características. Para realizar este estudo, recorrer-se-á a técnicas de análise de correlação e regressão e a técnicas de análise de *clusters*, e aos dados disponíveis na base de dados do Eurostat. Os resultados mostraram que a relação entre a economia e transporte aéreo existe em alguns países e regiões embora o seu nível de significância varie entre países e regiões sem um padrão geográfico.

Palavras-chave: Transporte aéreo, PIB, Regressão Linear, União Europeia.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

AU Austria
CARGO Air Cargo
Coef Coefficient
FDI Foreign Direct Investment
FR France
IT Italy
EU European Union
GDP Gross Domestic Product
NUTS Nomenclature of Territorial Units for Statistics
PAX Air Passengers
PT Portugal
SP Spain
t Stat t Statistical
UK United Kingdom
y Year

1 INTRODUCTION

1.1 Context

The aviation industry has changed sharply over the years. If, in its earlier days, clients were primarily from the upper social classes, nowadays this industry carries millions of passengers in business, vacations, leisure or vacations and delivers cargo around the world. This is a fundamental activity for the global economy and countries' development. We can also say that air transport is the main reason why, in the modern world, mobility is easier and transport is faster.

Aviation's role in the global economy had a boost in recent times because of more reasonable air transport prices. For this reason, aviation is today an important activity not only because it generates wealth and jobs but it also shows great value in the logistics of several industries. When talking about the importance of air transport, we have to think outside its main purpose, transport, and notice its impact in others industries such as aircraft manufacturing, tourism and so on.

Gross Domestic Product (GDP) is an economic measure that can help us evaluating a national economy. GDP is the combination of wealth generated in a country in a given period (typically, one year). It is an economic aggregate that takes into account all economic activities. GDP is the sum of the consumption of durable and non-durable assets, capital investments, public administration expenses and commercial balance (i.e. the difference between the exports and imports of a country).

Studying the relationship between economic activity and air transport is one of the main goals of this dissertation. Previous studies, such as Ishutkina [2009], analyzed the relation between GDP and air transport starting at the country level, then for multi-country regions and, finally, in some case-studies involving different geographic and income categories. According to this study one of the conclusions was that:

> "*The air transport demand is found to be directly affected by exogenous demand shocks, economic downturns, political and economic sanctions, and the development of other transportation modes.*" [Ishutkina, 2009]

With the present work, we will study the relation between Gross Domestic Product and Air Passengers (GDP-PAX), i.e. we observe whether GDP has an effect in the number of air transported passengers of the following year.

The other relation we examined is between Air Cargo and Gross Domestic Product (CARGO-GDP); specifically, we want to see if the growth of air cargo transport is associated with an increase or decline of next year's GDP.

We will analyze these relations starting at the country level and then go to the regional level.

The first focus of our study is the European Union (EU) area. The European Union is a community created in 1945 with 28 Member States, nowadays. It aimed to create peace and the well-being of its citizens, promote an economic and social union between its countries and set a common currency (Euro), among other goals. EU's economy is the most productive in the world – it generated a GDP of 15,3 trillion Euros in 2017, more than that of the United States, according to Eurostat.

To have an overall view of Air Passenger, Air Cargo and GDP values in the EU and in the World, we took these data from The World Bank database, since 1990 to 2017.

The air passenger's evolution has a pretty linear growth. Between the dates we considered, in the European Union, air passengers show a growth of 273% (PAX_EU_{1990} = 205 million, PAX_EU_{2017} = 763.8 million). Ireland was the country in the EU that showed a greater variation, it grew approximately 148 million passengers. The European Union's air passengers in 2017, accounted for 19% of the World's total air passengers.

As for Air Cargo and GDP, we show their evolution in Figures 1.1 and 1.2.

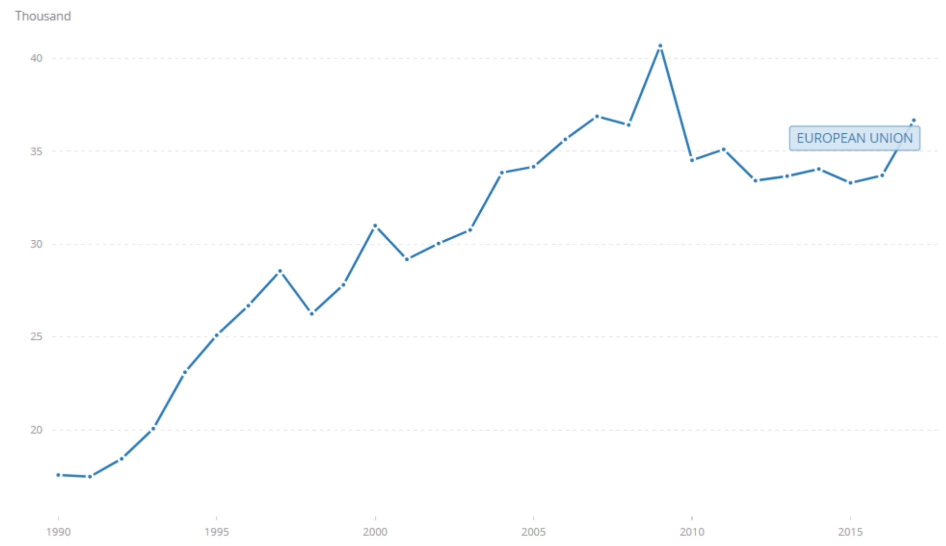


Figure 1.1 - Evolution of Air Cargo in the EU

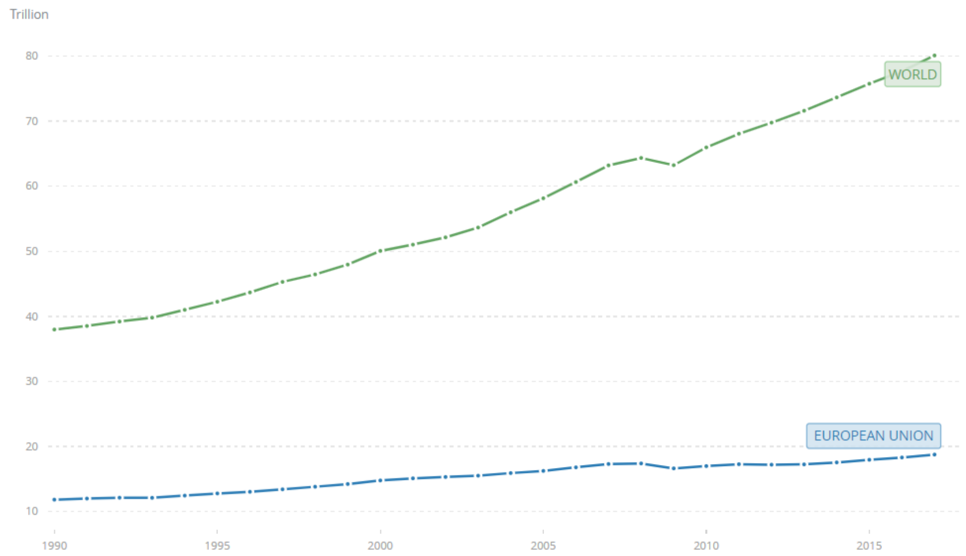


Figure 1.2 - Evolution of GDP in the World and the EU

In the first figure, we have the evolution of Air Cargo measured in thousands of metric tons times kilometers traveled. Air Cargo involves freight and postal traffic. As we can see in the evolution line, EU's growth is not constant over time - for example, in 2009 there is a spike. Air cargo grew from 17574 to 36672 (thousand) in this time window. In comparison to the World's Air Cargo, EU has a share of 17%, which is close to the percentage in Air Passengers (19%).

Figure 1.2 shows the evolution of the GDP in the World and in the European Union. This data is in constant 2010 U.S. dollars ($). "*Dollar figures for GDP are converted from domestic currencies using 2010 official exchange rates*" (World Bank). The GDP of the World grows 42 trillion euros whereas in the EU only grows 7 trillion. In 2017, the EU has 23% of the World's GDP.

Now, because our first sample of analysis is at country level, we collected the data for air passengers in 2017 from Eurostat, and we built a chart with all the Member States of the EU (Figure 1.3):

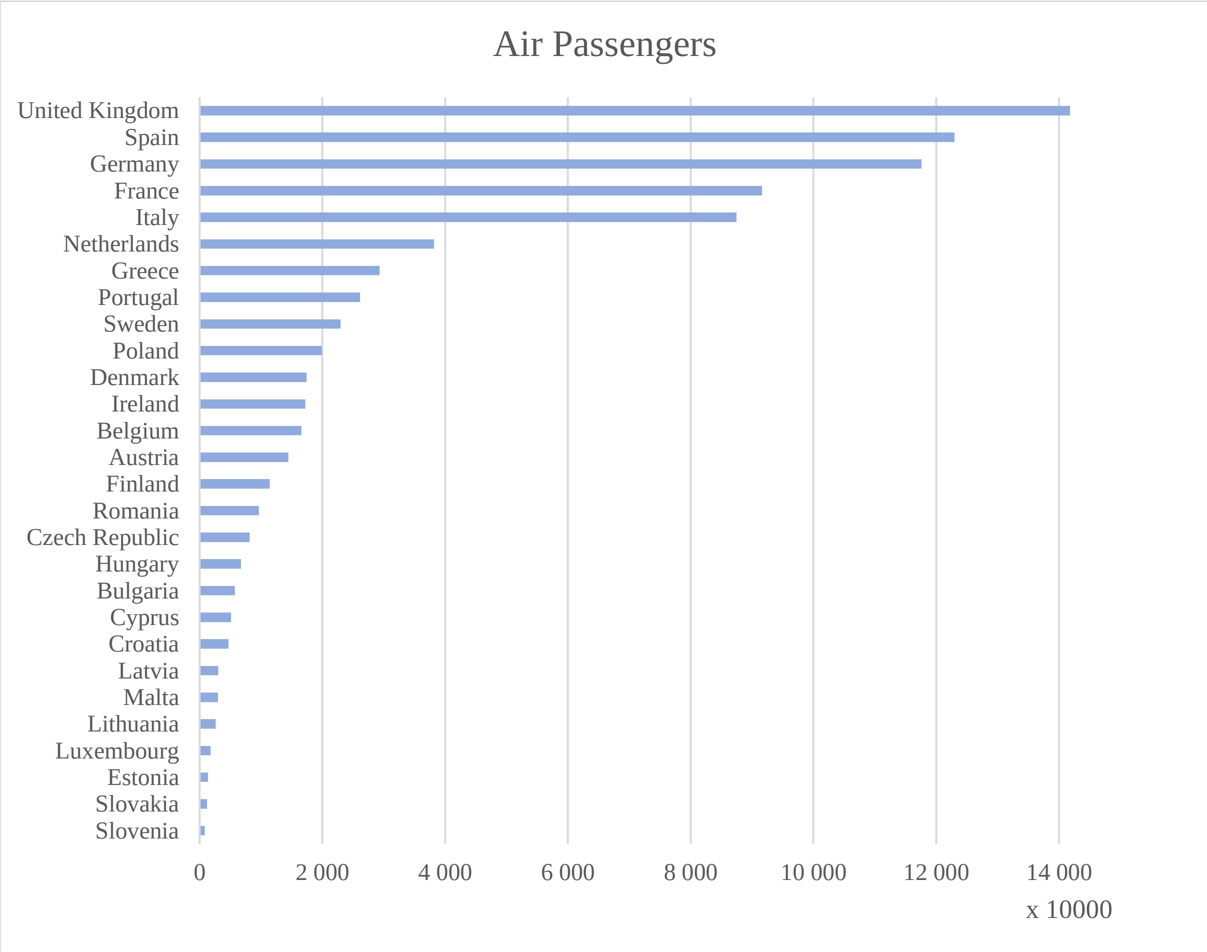


Figure 1.3 - Total Air Passengers in the EU in 2017

The United Kingdom has the highest number of air passengers, close to 142 million, while Slovenia has the lowest, approximately 850 thousand. We can also see a large variation of the PAX numbers for the different countries.

As for the cargo market, the EU in 2016 had the second largest share of global imports and exports. This was only possible due to the single market existing in the EU. Looking at the Figure 1.4 we can see the percentage of intra-EU trade and extra-EU trade for the different Member States.

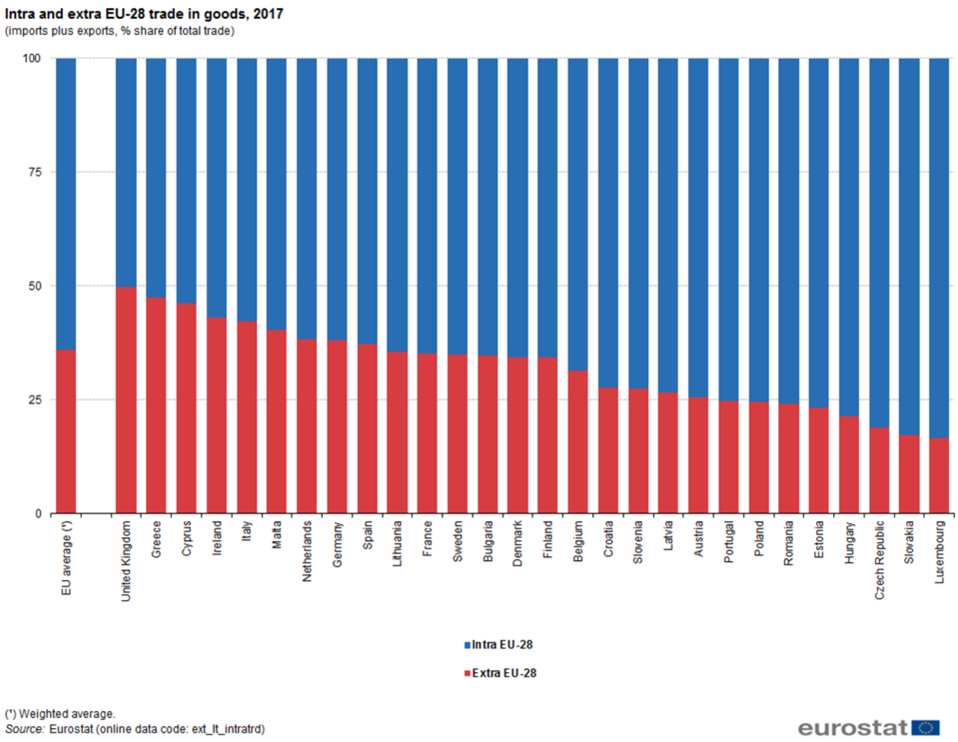


Figure 1.4 - Intra and extra-EU trade in cargo in 2017, (Eurostat)

As stated before, we will work at the country level and also with regions. These will be the Southwest regions of EU: France, Italy, Portugal and Spain. They are divided by NUTS II. NUTS stand for Nomenclature of Territorial Units for Statistics, which is a hierarchical territorial division system in regions. This terminology was created by Eurostat aiming to have a better division of the economic territory for statistics analysis concerning the nations of the European Union. There are three NUTS levels I, II and III - defined according to population, administrative and geographic criteria. The most recent classification of the NUTS is from 2015. The area we are going to analyze comprises a total of 74 NUTS regions.

1.2 Objectives

Is there a relationship between the economy of a country and its air passengers of the following year? Does every country from the European Union reacts in the same way? Are air passengers of some countries influenced by the economy of others from EU? These are some questions that we will try to answer in this dissertation.

First, our objective is to understand whether there is relation between the economy and air transport of passengers and cargo in all members of the European Union. To accomplish this objective, we start by doing a linear regression between the GDP and Air Passengers, where the GDP data is lagged by one year with respect to Air Passenger data. To check whether not only a country own economy influences the passengers, we carry out a multiple linear regression where we add the GDP of the rest of EU's member states as a second independent variable. In case of cargo, the question goes on the other direction. Are the GDP of a country correlated with air cargo in the previous year? Here we are only interested in the economy of the country in question, so a simple linear regression is enough. Air Cargo lagged one year is the independent variable, and the country's GDP is the dependent variable.

The next question we address is: is there a geographic pattern in these relationships? To answer this, we will use a cluster analysis to join the countries in groups that have similar characteristics, so that we can find a pattern. The characteristics we considered are the significance (*t Stat*) - and magnitude (*Coefficient*/ *Elasticity*) of the relationship analyzed through the linear regressions.

Afterwards, we will take a closer look to Portugal's situation. Here we want to see if it is useful, in the particular case of Portugal, to look at a GDP with a larger lag, both for air passengers and cargo.

Finally, we want to observe if there is a relationship GDP-PAX in the European Southwest regions (France, Italy, Portugal and Spain). Here we will have these countries divided into smaller areas, based on NUTS II. The air cargo will not be studied in this case. The independent variables considered in this analysis are the GDP of the region, the GDP of the rest of the regions from the same country and the GDP of the rest of the countries from the EU, all lagged one year.

1.3 Dissertation Structure

This dissertation is organized into seven chapters. Below is a brief summary of each one of them:

In Chapter 1 is done an introduction with relevant information that helps to understand the issues at stake. The objectives and methodology are also presented in this first chapter. Lastly it is explained how the document is organized.

Chapter 2 shows some of the works on the subject with most relevance to our study. They consist of a doctoral thesis and some journal articles that have different methods of analyzing the subject in question.

Chapter 3 is key to the comprehension of this work. The main concepts and the methodology used in this dissertation is explained here.

Chapter 4 presents the first application of the methodology. The 28 member states of the European Union are the focus of this chapter. A correlation and regression analysis will be done to see if there is a relationship between GDP and Air passengers and also between Air cargo and GDP for each country. Afterwards using a cluster analysis these nations will be separated by groups having as a criterion the results from the regression analysis.

Chapter 5 contains the case study of Portugal. The analysis in this case is deeper than the one conducted for the other European Union countries.

Chapter 6 is dedicated to the NUTS II regions of the Southwest Europe, i.e. France, Italy, Portugal and Spain.

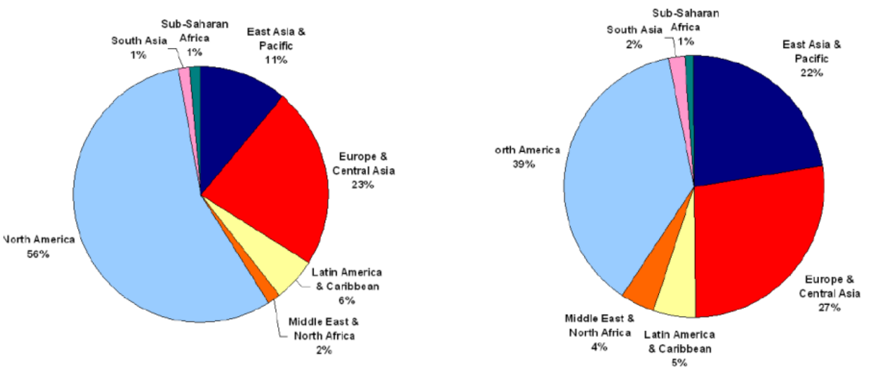
Finally, Chapter 7 summarizes the main conclusions about the influence of the economy on air passenger transport and about the influence of air cargo transport on the economy. Moreover, suggestions for further work on this subject are presented.

2 RELATED WORK

In this chapter, we intend to overview studies that were taken into account in our research. Similar studies on the relation between economy and air transport have been performed all over the world, though none of them specifically addressed the European Union.

"*Analysis of Air Transportation and Economic Development Data*" is one chapter of Ishutkina [2009] the doctoral thesis of Maryia Ishutkina, supervised by Prof. John Hansman (MIT). Here we are going to explore this chapter due to its importance to our work. It contains a study that involved 139 countries in a period time of 30 years (1975-2005) and, three types of analyses: firstly, a worldwide regional trends analysis; secondly, cross-sectional country analysis in the year of 2005; and thirdly, a GDP-PAX correlation analysis for the 30-year period. In all these analyses, the data used were the annual country level statistics.

By taking an overview of the trends of air passengers in large regions of the world (East Asia and Pacific, Latin America and Caribbean, North America, Sub-Saharan Africa, Europe and Central Asia, Middle East and North Africa and South Asian) the author found, as represented in Figure 2.1, higher values in North America, followed by Europe and Central Asia. However, through the years, East Asia and Pacific air passengers had a significant increase.



(a) Regional passenger shares in 1970. (b) Regional passenger shares in 2005.

Figure 2.1 - Changing patterns in regional shares (1970 to 2005) [Ishutkina, 2009]

Regarding worldwide trends, income (GDP) is one of the factors that can have more influence in people using air transportation, which is a relatively expensive mode. Therefore, high income countries have a greater percentage of air passengers, but these countries represent a low percentage of the world's population. They chose three other factors to take into account when talking about the requirements of countries' mobility: geographical location, country' size/population and distance to main economic markets. Most European countries offer good conditions in these respects, so they have lower need of mobility, whereas for New Zealand, Singapore, Bahrain and United Arab Emirates happens the opposite, because these countries are remote and/or small and/or distant to main markets. In isolated islands, air transport is very important for their economy, but the high values of air passengers may not be meaningful, because they correspond to passengers who are coming to the island, not the island population traveling outward.

In the country level analysis, Ishutkina studied the correlation coefficients, in the 30-year period, between air passengers and GDP. The correlation analysis determines the degree of linear association between those variables - if the correlation coefficient (R) is +1 those variables vary in the same direction in the exact same way; if the value is -1, the meaning is the opposite; and if the value is 0 the variables are independent. The results showed that 59% of the countries have a correlation coefficient higher than 0.8 meaning the relation GDP-PAX is strong.

Another analysis made by the author was about the growth of air passengers and GDP. In order to find this growth, she examined time series data and were able to divide the countries in four growth behavior groups, represented in the Figure 2.2.

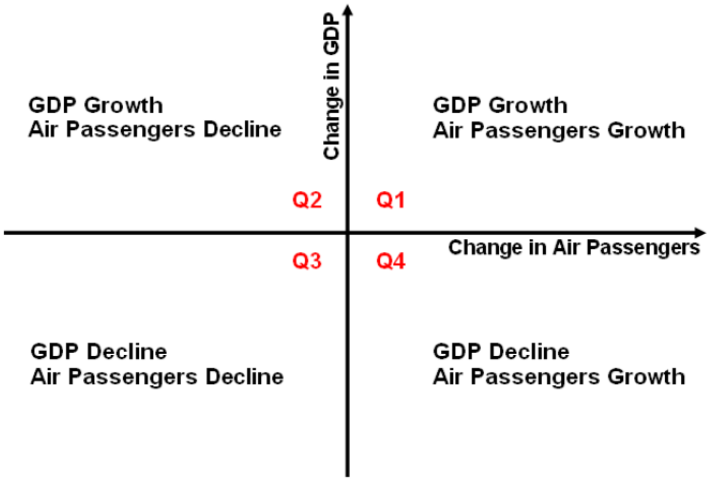


Figure 2.2 - Groups of growth behavior of Air Passengers and GDP [Ishutkina, 2009]

Then, to identify in which quadrant, in the previous figure, the countries were located they used yearly growth rates. The results show that 70% of the countries are in quadrants one and three, i.e. have mutual growth or decline with a positive correlation coefficient. The rest (30%), are in quadrant two and four, which means, they have opposite polarities and a negative correlation coefficient.

In a long-term analysis, they used longer time trends, instead of yearly growth rates, to identify continuous growth behavior, because some countries changed of quadrant in consecutive years which can lead to spurious results. Thus, a behavior was identified as continuous only when the annual growth rate stayed in the same quadrant for three or more years.

Stimulated and suppressed air transport development is a point also discussed in the analysis of growth behaviors in Ishutkina [2009]. It is evaluated by time trends and adjusted growth rates. They estimated by two types of growth rates, annual growth rate and average yearly growth rate. Using again the quadrants of Figure 2.2, she placed in the second and third quadrant the countries with suppressed air transport development, because this phenomenon happens when the growth rates are negative. The stimulated air transport development's countries are in the first and fourth, where the high growth behaviors are. They are considered high growth behaviors "*when the air passenger growth rates are higher when compared to the world's growth rates*".

The criteria she used to choose which countries to take for a detailed case-study were based on the results from the above analyses: difference growth behaviors (stimulated, suppressed and average), so that they would be able to identify the factor that influence these different behaviors; and after, to have a wider range, they chose different geographic and income category countries.

Profillidis & Botzoris [2015] did a research where they validate a causal relation between GDP and air passengers. In their paper, the level of this relation was measured by the degree of correlation between the number of air trips per 1,000 inhabitants and GDP per capita as given by the coefficient of determination (R^2). The sample used was worldwide and in multiple geographical areas of the world in a time period of 33 years (1980-2013). The coefficients of determination obtained for the World, Europe and Central Asia was R^2=0.96 meaning that the variables are closely related.

Specifically, for the year 2013, the authors studied the correlation between GDP per capita and air trips per 1,000 inhabitants of 53 countries. And the results showed that, if the GDP per capita grows 1%, then the number of worldwide air trips per capita will increase 1.47%.

In order to accomplish their second objective, which was to do an estimation of the number of air passengers for the years 2013 and 2030, they took into consideration fertility values, to predict the evolution of population. Another essential idea for this estimation is the assumption that the economic pattern and reasons for travelling would not have significant modifications in the future when compared to the past thirty years. From this analysis, it was shown that the annual growth rate of air passengers between 2020 and 2030, at a world level is 3.90%, and in Europe and Central Asia it would be 3.60%. They also realized with this study that the level of correlation would grow when the areas in study develop their air transport market [Profillidis & Botzoris, 2015].

Chi & Baek [2013] used economic growth and market shocks such as terrorist attacks, health issues, war and financial crises to analyze the short- and long- term effects these would have in air passengers and cargo transport demand in the United States. The market shocks were represented by dummy variables. These are artificial variables that enable to consider qualitative characteristics on regressions models. They are binary variables since they take one of two values, 0 or 1, to indicate the presence or the absence of a given characteristic. These variables represent something that does not have a numerical value or, if it has, those values do not have a numerical significance. With an autoregressive distributed lag (ARDL) model, they conducted a short and a long run analyses in the US.

Their results show that, - in the long run, both air passengers and freight services expand when the GDP is higher. In the opposite direction, air freight shows more impact on the increase of the US's income than air passengers. On the other hand, in the short term, only air passengers are susceptible to the growth of the economy. The market shocks have negative effects on air passengers but not as much as in air cargo transport [Chi & Baek, 2013].

Kasarda & Green [2005] describe a study of the relation between air cargo and GDP in 63 nations, chosen based on availability of data. They assumed that this relation existed and that while the two variables were interdependent, air cargo would drive the growth of the economy. Thus, to understand how they could improve air cargo towards an optimal effect on the economy, they studied three factors that had not been previously analyzed. The factors were: the effects of aviation liberalization, the customs' quality and the degree of corruption. Aviation liberalization was first evaluated by a correlation analysis, between air cargo, and economic indicators, such as GDP, FDI (Foreign Direct Investment) and trade, and, as they suspected, they are related. Customs was also important in the air cargo system, their policies are subjective which can be prejudicial. Lastly "*Corruption is a more complex issue that undoubtedly also impacts air cargo development and, to a broader extent, country competitiveness, foreign direct investment, and economic growth*.". To summarize, the results clearly suggest that a more liberalized aviation, a better quality of customs and low corruption would raise the economic development.

More studies on the subject were done, and they included other approaches such as Granger causality. This approach, created by Clive Granger allows us to understand whether a temporal series X helps to predict another series Y or whether it is the opposite that happens. Thus, if X contributes to predict Y then previous values of X will add significant information in the prediction of Y. To adopt this approach both series have to be stationary and both need to be lagged, this lag does not need to be the same for X and Y. There are four kinds of causal relation: causality - when Y is causing X; feedback - when Y is causing X and X is causing Y; instantaneous causality - when "*the current value of X is better "predicted" if the present value of Y is included in the "prediction" than if it is not*"; and causality lag which happens when the knowledge of the lagged values of Y does not help in the prediction of X [Granger 1969].

The Johansen test is one of recent co-integration tests. The co-integration tests have a great importance in the study of an economy. They allow us to find if there is an equilibrium or a relationship between the economic variables in the long run. The main advantage of this method compared with the one that existed before is that with the Johansen test we are able to identify how many co-integration vectors exist, while the others only were able to verify its the variables were co-integrated.

Chang & Chang [2009] study the air cargo transport relation with the national economy in Taiwan based on Granger causality. As mentioned in the previous paragraphs, before, this analysis relied in stationary time series, which means that the mean and variance were assumed not to change over time. After verifying if the variables are stationary, they tested the variables with the Johansen co-integration test, and if it were positive then there must exist one causal relation, in whichever direction. The final results showed that a bi-directional causality existed between them in the short- and long-term. This indicates that, the growth of air cargo transportation would cause GDP to increase, and the same would happen in the opposite direction. Thus, if any shock arises in the Air Cargo and GDP's system, this will have an effect on the long term.

Marazzo et al. [2010] investigated, in Brazil, the temporal links between air passengers and GDP. They made this research to learn how these variables are related and so that the future infrastructures, management and investments could be better planned. They adopted the co-integration of Johansen and Granger causality to do their analysis. They used the unit root and co-integration to detect if a long-term equilibrium between the growth of the economy and air transport demand exists; with the Granger causality and lag length, they concluded that these links exists and work in both directions. However, the effect provoked by the growth of passengers has on GDP variation is lower [Marazzo et al., 2010].

Rodríguez-Brindis et al. [2015] also used Granger causality and Johansen co-integration test to investigate the possible relation of causality between GDP and air passenger in Chile in the long run. Their final conclusion was that the causality relation is positive and bi-directional and that the elasticities are positive and show the existence of one co-integrated vector. The bi-directional causality implies that economy and air transport tend to grow in the same direction. Also, to evaluate the intensity of the relationship between the variables, they provoked a one percent shock in each variable to see the reaction of the other. The results showed that the increment in the number of air passengers is higher and continuous when it was provoked by a shock in the GDP.

3 METHODOLOGY

This chapter presents the approaches used to understand whether the economy has influence on air transport passenger and whether air cargo has influence on the economy. For that, we will use correlation and linear regression analysis, as well as cluster analysis.

3.1 Correlation Analysis

Correlation analysis is used to measure the degree of association between quantitative variables. This association does not account for the existence of causes and effects, i.e. the variables are not defined as dependent or independent. They are assumed as two random variables that vary together (or not). The relationship can be strong, weak or none. If the relation is strong means that knowing the variable *X* helps predicting *Y*, if it is weak, it does not help on the prediction of the other. Correlation coefficient can also be referred as *Multiple R*. The correlation coefficient is a value that ranges from -1 to 1. The Figure 3.1 is the representation of the coefficient range:

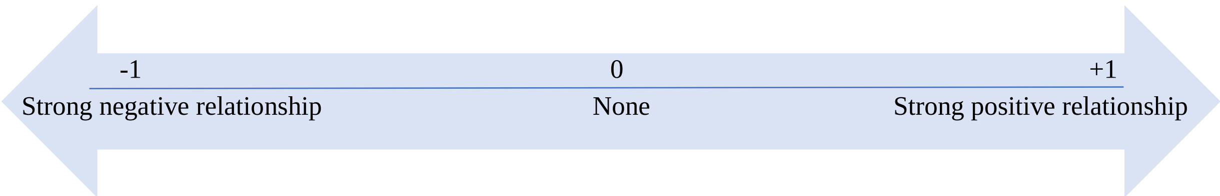


Figure 3.1 - Correlation range

As the coefficient gets closer to one, positive or negative, the relation is stronger. A coefficient of 1 indicates a perfect linear relation between the variables. If it is zero then there is no linear relationship. With a strong positive correlation ($r = +1$), we can say that if one variable has a large (rep. low) value the other has also a larger (rep. low) value. When the coefficient gets closer to $r = -1$ means that if one variable is high the other will be low. If the correlation coefficient is zero ($r = 0$), then knowing one variable does not tell anything about the value of the other. A strong correlation does not mean a causality relation, i.e. does not mean that one variable is caused by the other variable.

Since that with this analysis we cannot tell if a variable depends from the other, we will have to do a linear regression.

3.2 Linear Regression Analysis

Regression analysis consists in the realization of a statistical analysis that has the main goal of verify the existence of a relation between two or more variables. There are two types of linear regression:

- Simple regression, which has one dependent variable and one independent variable.
- Multiple regression, which has one dependent variable and two or more independent variables.

The dependent variable can also be called the response variable and the independent one the predictor variable. Next, we have a brief explanation of some of the output of these regressions, so we can have better understanding of the results that are in the next chapters.

The *t-Statistic* is the estimated coefficient divided by its standard error. It is used to test the hypothesis that the value of the coefficient is different from zero, i.e. it will validate if the independent variable belongs or not in the model.
The *p-value* is the probability of observing an equal *t-Stat* under the null hypothesis that the coefficient value is zero. A *p-value* of 0.05 or less, is usually the value at which the null hypothesis is rejected. If the *p-value* is greater than 5%, then the variable can be discarded from the model without having a big effect in the error measures. If *p-value* is lower than 0.05, the *t-Statistic* absolute value is equal or superior to 2, which means that the coefficient is significant.

After knowing if the *Coefficient* is significant, we want to see the intensity of that significance. For that that we analyze the estimated *Coefficient*. The *Coefficient* of a simple regression represents the variation of the dependent variable as function of the variation of one unit of the independent variable. In multiple regression, the *Coefficient* tells how much the dependent variable is expected to increase/decrease when that independent variable grows by one, while the other independent variables stay constant. The sign of the *Coefficient* tells us the direction of the result - if it is positive, the dependent variable will increase and if it is negative it will decrease.

In this work, all variables are converted to natural logarithms, where the base is $e \approx 2.72$. Thus, it is called a log-log regression model. The advantage of the natural logarithms is that the interpretation of the *Coefficients* is straightforward. When both dependent and independent variables are log-transformed, the *Coefficients* represent the *Elasticity* of *Y* with respect to *X*, i.e. the *Coefficient* is the estimated percent change in your dependent variables for 1% change in the independent variable.

The coefficient of determination, also known by *R Square* (R^2), tells if the model is suitable to describe the phenomenon, or not. R^2 varies in the interval [0,1], it is a more suitable fit when it is closer to 1, while a value near zero indicates no linear relationship. It indicates the proportion of variance of the dependent variable that is accounted by the independent variable.
Adjusted R Square is a R^2 adapted for the number of terms in our model. *Adjusted* R^2 usually has a value lower than R^2 value. This happens because there is more than one independent variable. It gets lower when the number of independent variables grows. *Adjusted* R^2 is preferable than R^2 because when there are more than one independent variables, R^2 has a tendency of being inflated.

In Chapter 4 we will apply the linear regression to all countries of the European Union, each one at a time. For the air passenger analysis, we will use multiple regression. We will consider two independent variables: the GDP of the country in study and the GDP of the rest of the members of the EU. We use these two variables so that we can see not only if there is influence of the country own GDP on air passengers, but also if it has influence of the neighbor countries' GDP of this community. Both predictor variables are lagged one year from the air passengers. Below, is the regression equation of this analysis:

$$PAX = \beta_0 + \beta_1 \times GDP\_Country_{-1} + \beta_2 \times GDP\_Rest\_EU_{-1} \tag{3.1}$$

where:
PAX (number of air passengers), is the dependent variable.
β_0 is the regression constant.
β_1 & β_2 are the regression coefficients.
$GDP\_Country_{-1}, GDP\_Rest\_EU_{-1}$ (GDP of the country and GDP of the rest of the EU, with a one-year lag in relation to PAX), are the independent variables.

Air cargo is also studied in the fourth chapter. For air cargo evaluation, we will perform a simple linear regression between the country's GDP and air cargo from the previous year:

$$GDP\_Country = \beta_0 + \beta_1 \times Cargo_{-1} \qquad (3.2)$$

where:
$GDP\_Country$ (GDP of the Country), is the dependent variable.
β_0 is the regression constant.
β_1 is the regression coefficient.
$Cargo_{-1}$ (tons of air cargo, lagged one-year in relation to GDP), is the independent variable.

In Chapter 5 we study Portugal. Here, both for passenger and cargo analysis, we will use simple and multiple regression. In this chapter, there will be different independent variables than in the previous chapter. These new independent variables will take into account more than one year of difference between GDP and air passengers or cargo. By adding these variables, we are trying to see if, in case of Portugal, with more than one year of delay between the variables we will have different results to compare between them. Also, in air passengers, we will divide the EU Member States into two groups according to their socioeconomic relations with Portugal. The ones with more relations are: France, Germany, Italy, Spain and UK.

For the relation between GDP and air passengers, we did several analyses, maintaining the dependent variable (air passenger), and changing the independent variables. Specifically, we will have 10 different cases:

- Linear regression with Portugal's GDP.
- Multiple regression with Portugal's GDP lagged one and two years.
- Multiple regression with Portugal's and EU's GDP lagged one year.
- Multiple regression of Portugal's and Europe's GDP lagged one and two years.
- Linear regression with the sum of Portugal's GDP lagged one and two years.
- Multiple regression of the sum of the lagged years, of both GDPs.
- Linear regression of the rate of Portugal's GDP.
- Multiple regression with the rate of Portugal's and EU's GDP.
- Multiple regression with Portugal's GDP, GDP of the countries with more socioeconomic relations and GDP of the countries with less socioeconomic relations lagged one year to the passengers.
- Multiple regression with Portugal's GDP, GDP of the countries with more socioeconomic relations and GDP of the countries with less socioeconomic relations lagged two years to the passengers.

In case of air cargo in Portugal, we will consider three different cases:

- Linear regression between cargo lagged one year and Portugal's GDP.
- Multiple regression of cargo lagged one and two years.
- Linear regression with the sum of cargo lagged one and two years.

Finally, in Chapter 6, we will analyze the air passengers in the NUTS II Regions of Southwest Europe. Air cargo will not be treated in this chapter due to lack of data.

Using as dependent variable the air passengers in each region and as independent variables the GDP of that region, the GDP of the rest of the country and the GDP of the rest of the EU, all lagged one year, we performed a multiple regression considering the following equation:

$$PAX = \beta_0 + \beta_1 \times GDP\_Region_{-1} + \beta_2 \times GDP\_Rest\_Country_{-1} + \beta_3 \times GDP\_Rest\_EU_{-1} \quad (3.3)$$

where:
PAX, is the dependent variable.
β_0 is the regression constant.
β_1, & β_2& β_3 are the regression coefficients.
$GDP\_Region_{-1}, GDP\_Rest\_Country_{-1}, GDP\_Rest\_EU_{-1}$ are the independent variables.

In every case we will use data available at the "Eurostat" database, from 1993 until 2017 for Chapter 4 and 5. There are some gaps, but we will ignore those years that are missing. In Chapter 6 the data is from 2000 to 2016, with the exception of 2006.

The values for PAX used in our study are the passengers carried in air transport. The number of arrivals and departures should be very similar so we opted to use the departures values.

For cargo we considered the quantity (tonnage) of freight and mail on board by air transport by reporting country. Unfortunately, with these data, there are several years without information so the sample will be smaller than we hoped.

So that we can do the direct calculation of variation rates in volume, we can use chain series, which can convert the data into a chain volume series in a certain referenced year. The values for GDP we use are the chain volumes from 2010 measured in million euros. This means that the data is based on 2010 prices, linked over the years through fitting growth rates. In Chapter 6 the GDP utilized is at current market prices in million euro.

For performing the regressions, we used the SPSS (Statistical Package for the Social Sciences), which is a software that, in a simple way, provides multiple technics and statistics methods to organize and summarize databases, making them more understandable. It allows us to select and apply the most adequate statistic procedure to obtain the result for a question.

In order to clarify how the methodology was applied we will consider here the case of Austria. First, we present the simple regression equation for *Cargo* and *GDP*:

$$GDP\_AU = 8.83 + 0.31 \times Cargo_{-1} \quad (3.4)$$

8.83 is the expected value of the logged GDP without cargo.
0.31 is the percentage of variation of GDP for one percent variation of cargo.

$R^2 = 0.94$, which means that 94% of the variation of the *GDP* is explained by the regression.
$t\ Stat = 17.97\ (> 2)$ so, the growth of *Cargo* has a great significance in the growth of *GDP*.

Next, we show the figures that help the interpretation of results. Figure 3.2 shows the trendline that represents the relation CARGO-GDP.

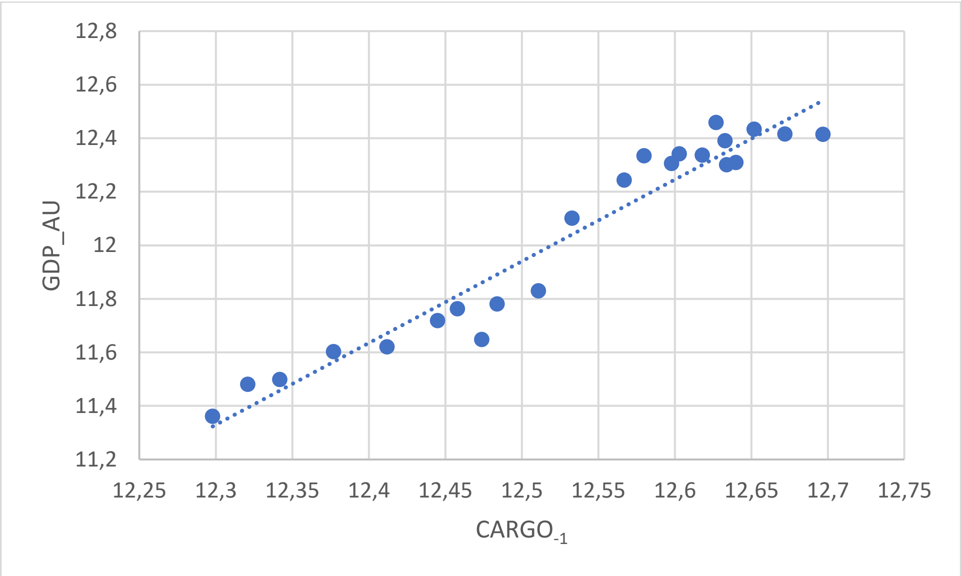


Figure 3.2 - Regression line of Austria’s CARGO-GDP

As we can see from Figure 3.2, points are generally close to the linear trendline, that is why the coefficient of determination is so high (0.94). Next, on Figure 3.3 we can observe the variation of the *GDP* and *Cargo* through the years we examined.

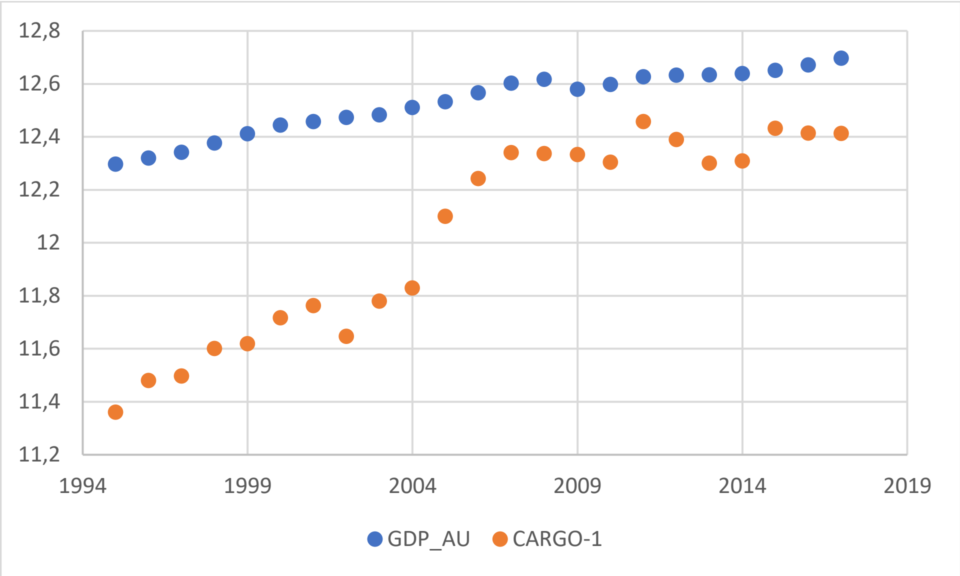


Figure 3.3 - Variation of Austria's GDP and Cargo

Now the multiple regression equation for *GDP* and *PAX*:

$$PAX = -10.95 + 3.22 \times GDP_{AU_{-1}} - 0.81 \times GDP\_Rest\_EU_{-1} \quad (3.5)$$

-10.95 is the mean value of the Passengers with zero GDP.
3.22 is percentage of passengers that grows when GDP of Austria increases one unit.
-0.81 is percentage of passengers that decreases when GDP of the rest of EU rises one unit.

$R^2 = 0.96$, which means that 96% of the variation of PAX is explained by the regression.

$t\ Stat\ (AU) = 3.09\ (> 2)$, the increase of Austria's *GDP* has some weight in its variation of air passengers.
$t\ Stat\ (EU) = -0.73$, it is between [-2,2], the *GDP* of the EU is not significant in the numbers of air passengers in Austria.

3.3 Cluster Analysis

The clusters analysis methods are a procedure of multivariate statistic used to organize entities into relatively homogeneous groups considering some of the features. This analysis is a technique of interdependence, i.e. there is no dependent or independent variable. Every variable is related positively or negatively with each other but none of them has a dependency relation with the others.

In the first level we can divide clustering methods into two types: hierarchical and nonhierarchical. In this work, we will use a hierarchical method. Hierarchical methods are also divided into two groups: agglomerative and divisive. We will use an agglomerative method. The cluster formation with an agglomerative hierarchal method is based on grouping the nearest cases according to the chosen distance measure, the Euclidean distance in our case, and at successive steps, similar clusters are merged until there is only one cluster.

Figure 3.4 is a representation of all the clustering methods that exist.

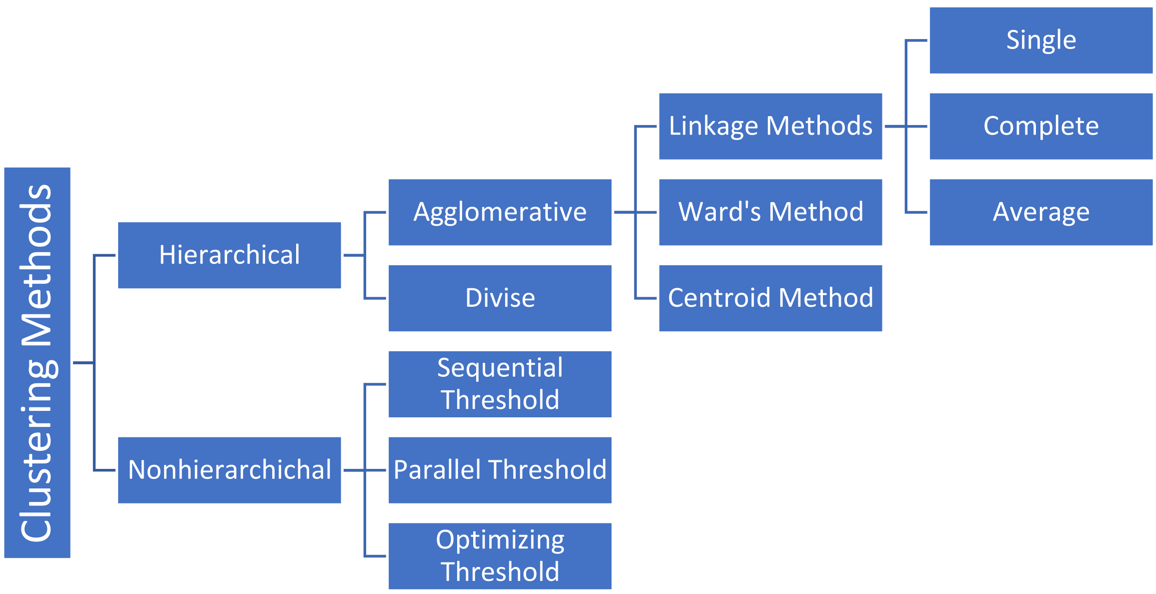


Figure 3.4 - Summary of Clustering Methods

The hierarchical grouping methods diverges on the way they calculate the distances between entities and groups, hence the choice of the method conditions the resulting classification. We will choose the centroid method where the distance between two clusters is calculated by making the sum of the distances between clusters means for all the variables.

With the results obtained from the previous regressions we will do a hierarchal cluster analysis with the help of the *SPSS software*, so that we can divide countries of the European Union, our entities, into smaller groups based on the coefficients and its statistical significance (*t Stat*).

When making this analysis, we have to standardize the data so that there is not any distortion in the results due to the different units of the values of the *Coefficients* and *t Stat*. So, we transform the values into standard values through the subtraction of the average and the division by the standard deviation. Thus, every vector will have an average of zero and a standard deviation of one. This standardization is called “Z scores”.

As a result, a dendrogram and an agglomeration schedule are built. A dendrogram is a graphic representation of the clustering process, it shows every step of the process since the break of all the entities until their incorporation into only one group. On the vertical axis, we have the name of the entities, in our case the EU countries. The horizontal axis has a scale that indicates the distance at which the clusters are grouped.

An agglomerative schedule is a table that shows the several stages of the clustering process, the cluster combination, the coefficients, the stage cluster first appears and the next stage where the first cluster appears. The coefficient is the Euclidean distance between the two cases being clustered.

In this analysis, we are able to choose the number of clusters we want to have. One way to choose this number is to analyze the column of the coefficients on the agglomeration schedule and look for a sudden jump. A good solution is the cluster number before that jump. The other way to select the cluster is to analyze the dendrogram and make a vertical cut at a distance of our choice, the number of lines it intersects is the number of clusters. Figures 3.5 and 3.6 exemplify the dendrogram and agglomerative schedule obtained for the Cargo and GDP cluster analysis.

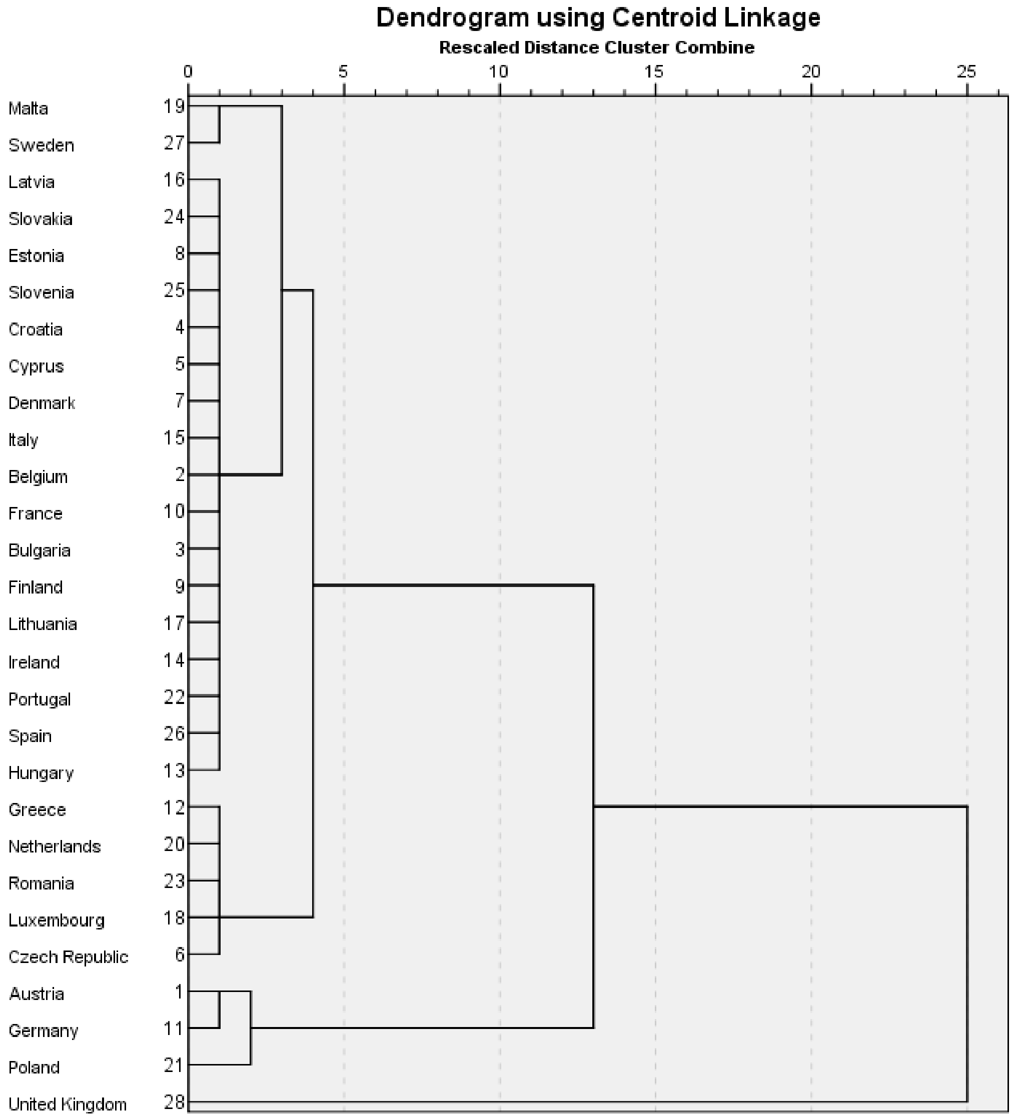


Figure 3.5 - Dendrogram of CARGO-GDP

Agglomeration Schedule

Stage	Cluster Combined		Coefficients	Stage Cluster First Appears		Next Stage
	Cluster 1	Cluster 2		Cluster 1	Cluster 2	
1	19	27	,004	0	0	24
2	14	22	,006	0	0	8
3	16	24	,006	0	0	9
4	4	5	,007	0	0	12
5	8	25	,011	0	0	9
6	2	10	,018	0	0	13
7	12	20	,023	0	0	16
8	14	26	,024	2	0	15
9	8	16	,027	5	3	14
10	3	9	,028	0	0	11
11	3	17	,027	10	0	13
12	4	7	,030	4	0	14
13	2	3	,071	6	11	21
14	4	8	,073	12	9	18
15	13	14	,075	0	8	21
16	12	23	,110	7	0	19
17	1	11	,143	0	0	23
18	4	15	,165	14	0	22
19	12	18	,254	16	0	20
20	6	12	,507	0	19	25
21	2	13	,566	13	15	22
22	2	4	,628	21	18	24
23	1	21	,674	17	0	26
24	2	19	1,653	22	1	25
25	2	6	2,088	24	20	26
26	1	2	7,992	23	25	27
27	1	28	16,403	26	0	0

Figure 3.6 - Agglomeration Schedule of GDP-Cargo

In this case, since we only have 27 countries and a big jump happens only on cluster 25 it is wiser to choose the final number of clusters from the cut on the dendrogram. For example, if we draw a vertical line at the cluster distance of 2 on the dendrogram we will intersect 5 lines, which means that we will have 5 clusters in this case.

4 APPLICATION TO THE EU

In this chapter we analyze the relationship between GDP and air passengers and between air cargo and GDP in all countries of the European Union.

The countries that we will work with are members of the European Union. The European Union was born in 1945. Its main goal, as mentioned before, was to create a peaceful Europe. It started with six countries and throughout the years others joined. Now, in 2018, there are 28 members. Figure 4.1 is a map of today's European Union.

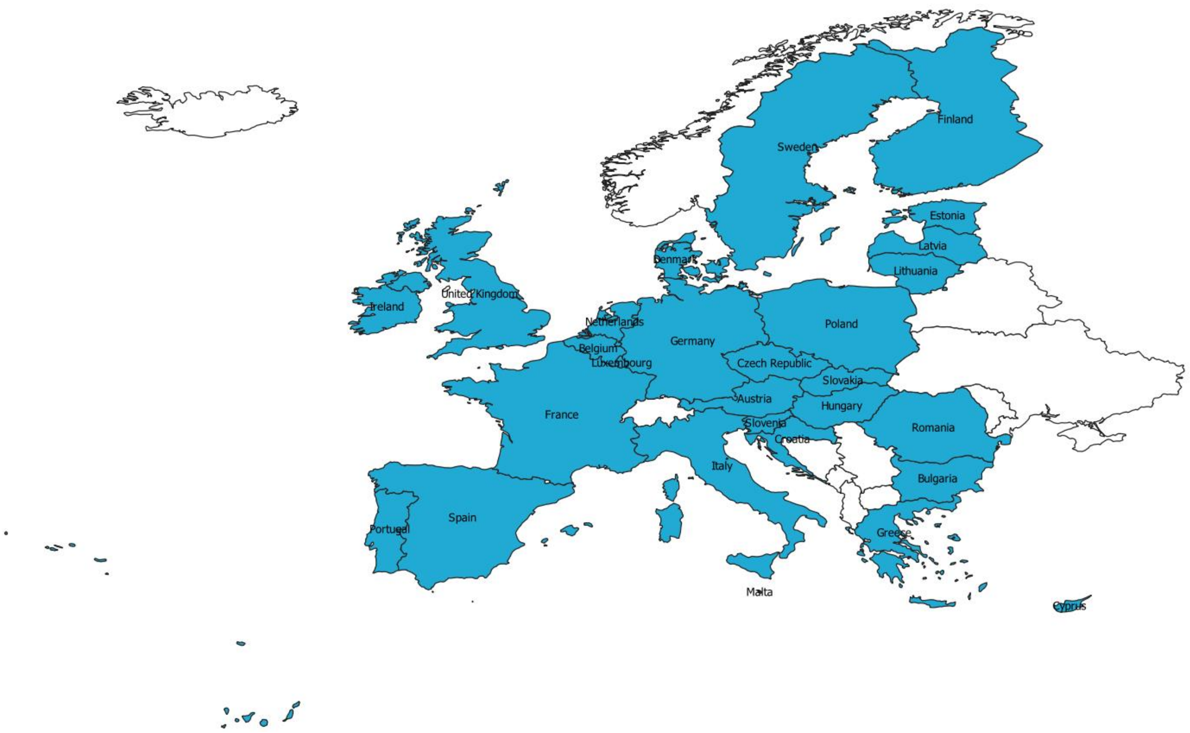


Figure 4.1 - European Union Member States

4.1 Passengers

As explained in Chapter 3, we run the multiple linear regression model on *SPSS*, for each country. The values collected from each country are presented in Table 4.1.

Table 4.1 - Results for Air Passenger and GDP Regression

		Country's GDP		Country's GDP		
COUNTRY	*Multiple R*	*Coefficient*	*t Stat*	*Coefficient*	*t Stat*	*Adjusted R^2*
Austria	0.98	3.22	3.1	-0.81	-0.73	0.97
Belgium	0.87	7.95	3.45	-6.41	-2.67	0.74
Bulgaria	0.88	1.03	1	4.14	1.97	0.73
Croatia	0.96	-2.52	-4.75	6.6	8.17	0.91
Cyprus	0.88	-0.95	-3	2.94	5.68	0.74
Czech Republic	0.88	0.59	0.39	2.07	0.63	0.73
Denmark	0.86	-1.07	-0.55	3.32	2.16	0.69
Estonia	0.85	-0.15	-0.16	5.14	2.34	0.67
Finland	0.89	0.049	0.048	2.34	1.79	0.76
France	0.94	1.35	3.09	0.81	4.51	0.88
Germany	0.97	2.19	7.55	0.35	4.05	0.93
Greece	0.90	-0.13	-0.49	2.47	7.47	0.79
Hungary	0.95	2.75	3.04	0.94	0.79	0.89
Ireland	0.97	0.84	2.81	0.6	0.73	0.93
Italy	0.91	-0.39	-0.33	3.21	6.62	0.81
Latvia	0.89	-1.48	-1.68	10.11	4.95	0.75
Lithuania	0.94	0.39	0.22	9.2	1.78	0.86
Luxembourg	0.94	5.41	4.89	-7.39	-2.87	0.86
Malta	0.98	2.13	9.44	-1.3	-2.46	0.95
Netherlands	0.94	-0.7	-0.58	2.87	2.18	0.87
Poland	0.96	1.71	2.71	5.88	2.53	0.91
Portugal	0.97	-1.78	-3.29	4.36	11.45	0.94
Romania	0.98	2.03	1.77	4.3	1.28	0.95
Slovakia	0.87	-2.81	-1.59	16.66	2.82	0.72
Slovenia	0.73	1.13	1.26	0.45	0.35	0.45
Spain	0.94	1.6	1.8	1.43	1.18	0.87
Sweden	0.94	2.96	2.29	-0.7	-0.31	0.86
United Kingdom	0.97	1.54	7.55	0.28	2.48	0.93

In this table we can see that the results variate from country to country. The coefficient of correlation (*Multiple R*) has great values, all close to 1, thus we can say that the variables in study have a strong positive relationship between them.

When looking at the *Adjusted* R^2, with the exception of Slovenia, all of the countries have a value higher than 65% (Slovenia's R^2 = 45%). Austria has the highest *adjusted* R^2, meaning it is the country which the model has a more suitable fit.

Afterwards, we introduce the values of *Coefficient* and *t Stat* on *QGis software* and were able to build maps so that we can have a clearer view of the results. Represented in following figures are the 28 countries with different colors.

Figure 4.2 and 4.3 represent maps with the *Coefficient*. In red are the countries with a negative *Coefficient*, which means that the percentage of air passengers will decrease with the growth of the GDP. The countries in green represent a positive *Coefficient* meaning that when the GDP grows, the percentage of air transport passengers will also grow. A more intense color, whether is red or green, represents a higher absolute value of the *Coefficient*.

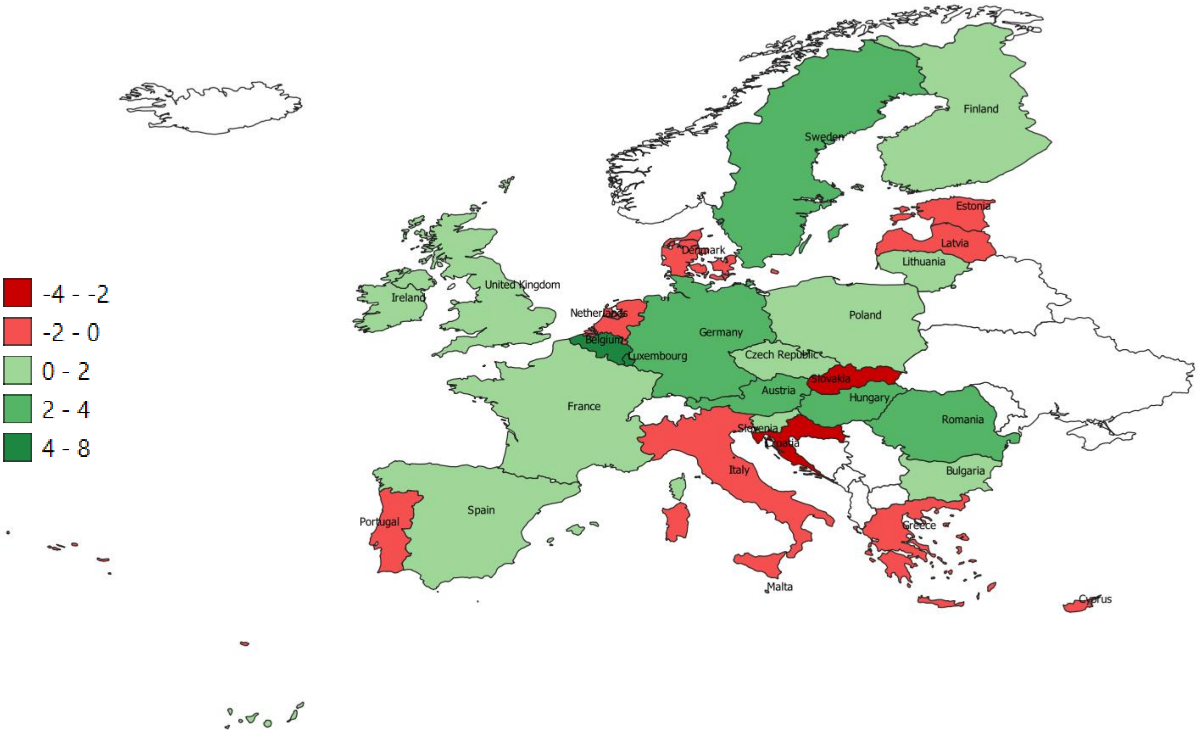


Figure 4.2 - Map of Europe with *Coefficient (Elasticity)* results for GDP of Country

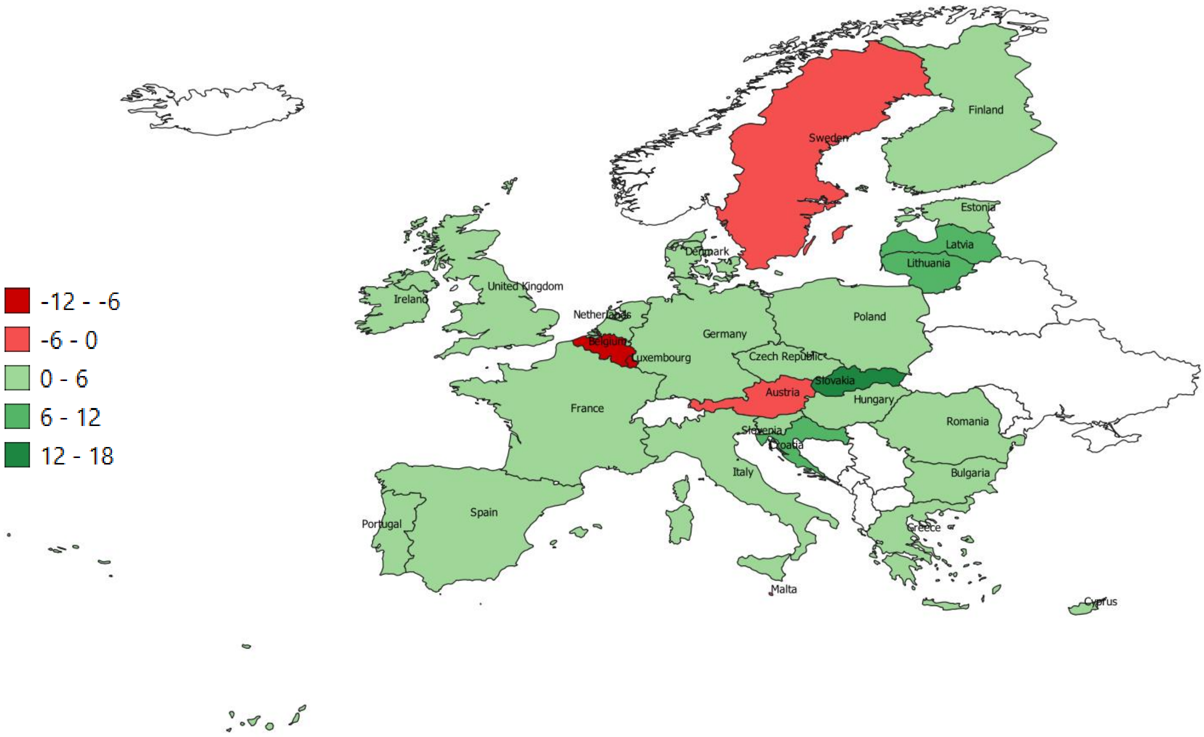


Figure 4.3 - Map of Europe with *Coefficient (Elasticity)* results for GDP of EU

Figure 4.2 represent the *Coefficient (Elasticity)* when analyzed the country's GDP. It has more countries with a negative value and have a lower range of percentages than Figure 4.3, which represent the *Coefficient* of the GDP of the others Member States of the EU.

In Figure 4.3, although most of the countries are in the same interval, 0 to 6%, the percentage of growth goes up to 18%. This means that, when taking in consideration the economy of the other EU countries, air passengers would have a big growth.

With the *Coefficient* we could see what is the percentage of passenger that will increase or decline with the growth of the GDP, for all the countries. Now, the next step is to analyze the *t Sta*t results, so that we can understand which countries have a significance relationship between GDP and air passengers.

In figure 4.4 and 4.5, the various colors depend from the different values of *t Stat*. In yellow are the countries which do not have a significant correlation GDP-PAX, that is, their *t Stat* values are between -2 and 2. With the color red are the countries with a negative value and ≤2. Finally, in green, are the countries which have a positive value and are equal or higher than two. As darker as the colors get, whether is green or red, the greater the absolute value of *t Stat* will be.

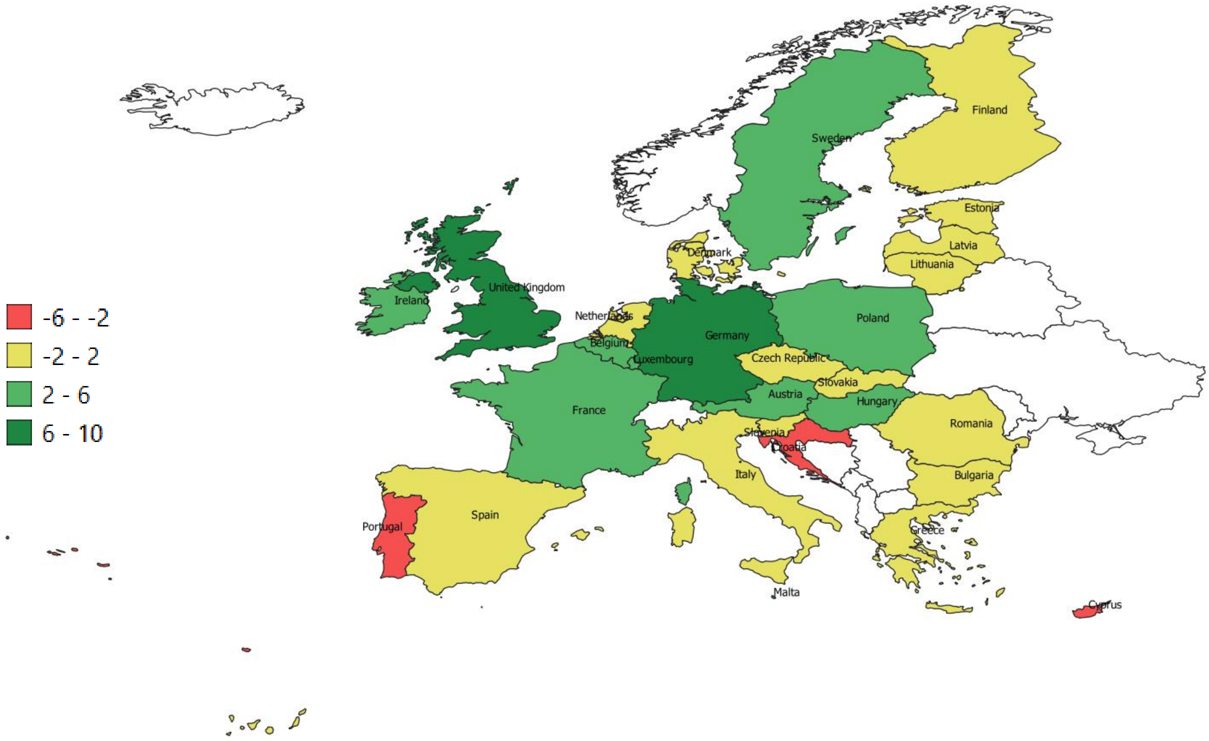


Figure 4.4 - Map of Europe with *t Stat* results for GDP of Country

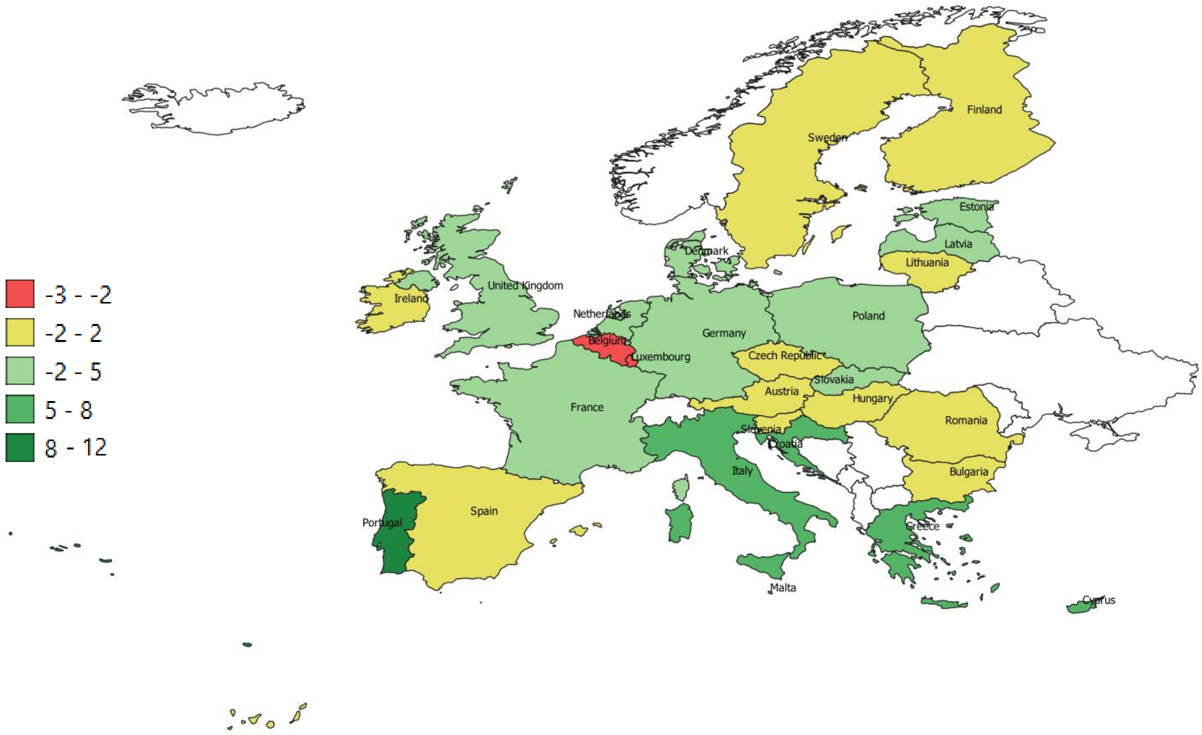


Figure 4.5 - Map of Europe with *t Stat* results for GDP of EU

In Figure 4.4, which has the *t Stat* results for the GDP of the country in analysis, we see that 14 members of the EU do not have a significant relation between air passengers and their country's GDP, meaning that a high national economy is not a main reason for their population to travel. On the other hand, Germany, Malta and United Kingdom have a very strong GDP_Country-PAX relation.

With the *t Stat* values of the GDP of the rest of the countries from the EU, Figure 4.5 shows some differences from the previous map. However, there are 7 countries which remain in the interval [-2,2] of *t Stat*, they are: Bulgaria, Czech Republic, Finland, Lithuania, Romania, Slovenia and Spain. The total of not significant relationships decreases to 12 countries from the first case. We also can notice that *t Stat* values are higher in this case, which means that the level of significance of the relation is superior than in the first case. So, we can conclude that people in the EU travel more when the economy of others countries is better than their own.

The following step in air passenger's analysis is to gather the countries of the European Union into small groups with similar characteristics. For that we will perform an agglomerative hierarchical cluster analysis on SPSS. The resulting groups will help us learn if there is a geographic pattern on the countries in study, in terms of economy and air passengers. We will have two cases: the relation Country's GDP-PAX and EU's GDP-PAX.

- **GDP_Country-PAX**

After analyzing the agglomeration schedule (Appendix A) we should have 24 clusters, because is the cluster before a big gap. However, this a very large number of clusters and it does not help in our objective, so we choose the number of clusters by cutting the dendrogram (Appendix B) at the distance 2.5. Thus, we will have 9 clusters, that are represented in Figure 4.6.

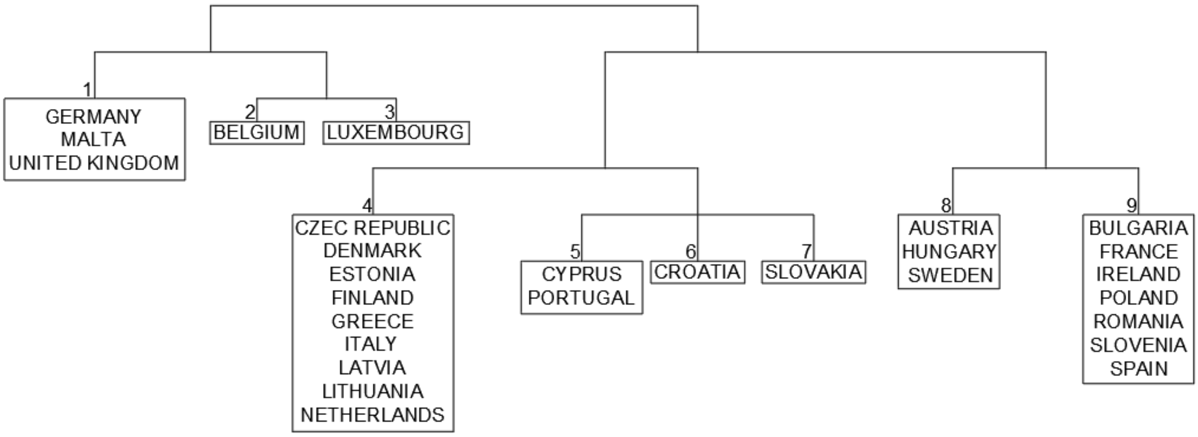


Figure 4.6 - Dendrogram GDP_Country-PAX

This previous figure displays the dendrogram we built with the hierarchical relation between our nine clusters. As we can see, when analyzing the GDP of the county, there are some countries that have their own cluster, because their characteristics are different and cannot be grouped with others. As for the cluster with more than one country, like cluster 4 or 9 there is no clearly geographic pattern observed.

- **GDP_EU-PAX**

Similar as it happen above, we should choose the number of clusters by cutting the dendrogram also at the distance 2.5, because if we use the other alternative (analyze the agglomerative schedule) we would have a very large number of clusters. The agglomerative schedule and the original dendrogram are in Appendices C and D, respectively. Figure 4.7 is the dendrogram we built.

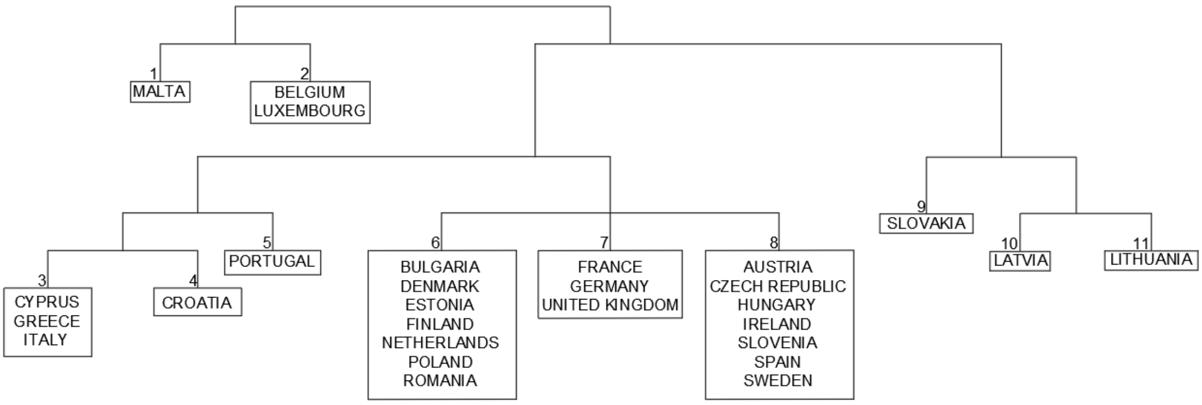


Figure 4.7 - Dendrogram GDP_EU-PAX

The dendrogram in this figure shows that we have 11 clusters and six of them only have one country. This tells us that the characteristics we chose to do this analysis, *t Stat* and *Coefficient* from the regression of the EU's GDP, are not very similar from country to country and for that reason there are many countries in an individual cluster. Hence, we cannot find a pattern we were looking for.

4.2 Cargo

Now, doing the Air Cargo analysis we see the Table 4.4, with the results of the linear regression.

Table 4.2 - Results for Air Cargo and GDP Regression

COUNTRY	*Multiple R*	*Coefficient*	*t Stat*	R^2
Austria	0.97	0.31	17.97	0.94
Belgium	0.84	0.17	5.78	0.70
Bulgaria	0.92	0.22	6.5	0.84
Croatia	0.40	0.13	1.14	0.16
Cyprus	0.26	0.11	0.97	0.07
Czech Republic	0.75	0.6	4.08	0.56
Denmark	0.39	0.075	1.13	0.15
Estonia	0.54	0.1	2.38	0.29
Finland	0.85	0.18	6.87	0.72
France	0.83	0.2	5.42	0.69
Germany	0.96	0.26	16.24	0.92
Greece	0.89	0.44	6.84	0.80
Hungary	0.73	0.29	3.97	0.53
Ireland	0.59	0.29	2.54	0.35
Italy	0.19	0.016	0.78	0.04
Latvia	0.56	0.12	2.53	0.31
Lithuania	0.92	0.21	7.53	0.85
Luxembourg	0.86	0.54	7.77	0.74
Malta	0.06	-0.11	-0.23	0.00
Netherlands	0.86	0.47	6.32	0.74
Poland	0.99	0.32	21.4	0.98
Portugal	0.58	0.31	2.58	0.34
Romania	0.85	0.39	5.41	0.73
Slovakia	0.57	0.14	2.62	0.33
Slovenia	0.61	0.074	2.52	0.37
Spain	0.55	0.26	2.44	0.30
Sweden	0.19	-0.11	-0.58	0.04
United Kingdom	0.63	1.28	3.04	0.40

Conversely to air passenger's analysis, in cargo analysis, the *Multiple R* values are not as good. As for R^2, many of the percentages are low, which is not a very positive result. Figure 4.8 and 4.9 will help to have a better perspective of these results.

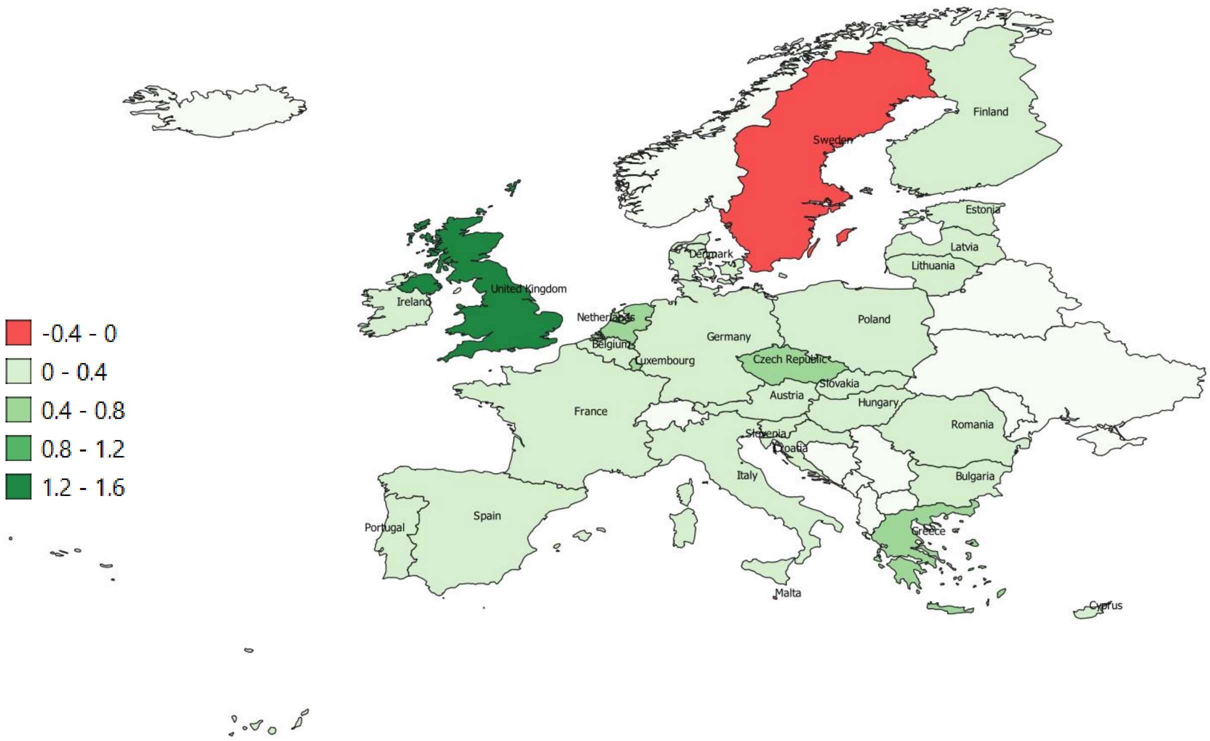


Figure 4.8 - Map of Europe with *Coefficient* results for Cargo

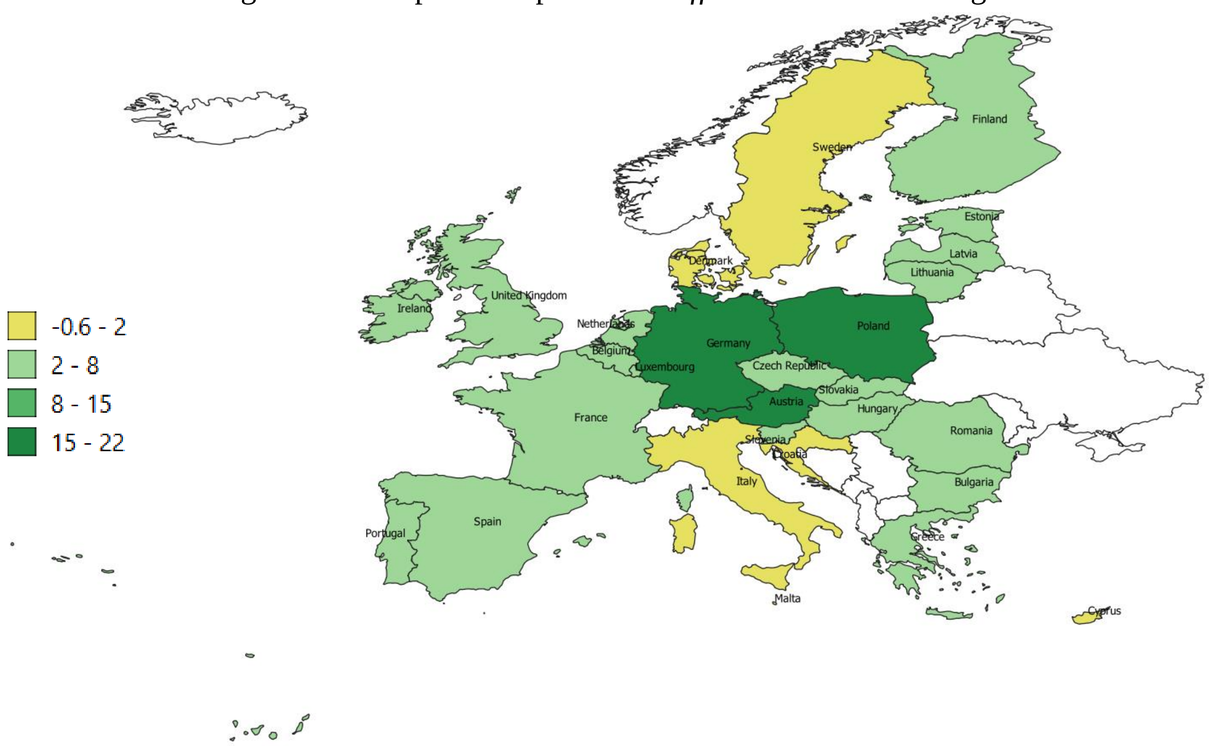


Figure 4.9 - Map of Europe with *t Stat* results for Cargo

The scale of colors is the same as the one used in air passengers, bearing in mind that now we are looking for the growth of GDP in function of air cargo. In Figure 4.8 is represented *Coefficients* values and in Figure 4.9 the *t Stat* of the CARGO-GDP relation.

Figure 4.8 shows that there is not a big variation of the C*oefficient* value, it only goes from -0.15 to 1.30 which means that when the ton of cargo grows 1%, the growth of GDP is not higher than 1.28%, (on UK's case). We can also see that only two countries have a negative value of C*oefficient*, Malta and Sweden.

From Figure 4.9 we can see that only in 5 countries the GDP has an independent development from the air cargo. However, the rest have a GDP growth influenced by evolution of air cargo, which means that cargo is very important for the enrichment of each country economy. In case of Austria, Germany and Poland, air cargo transportation has a great significance on their GDP. None of them have a negative and lower than 2 value of *t Stat*.

Here, in air cargo analysis, we have to notice that because of the deficit of data the results can be a few far from the reality.

In order to find smaller groups with similar characteristics of Cargo and GDP, next we have the cluster analysis for Air Cargo.

- **CARGO-GDP**

In this case, like it happened in air passenger cluster analysis, we choose the number of clusters by doing a vertical cut on the dendrogram, at 2.5 distance, which results in 5 clusters. The output information of this analysis is in Appendices E and F. Figure 4.10 is the dendrogram built by us.

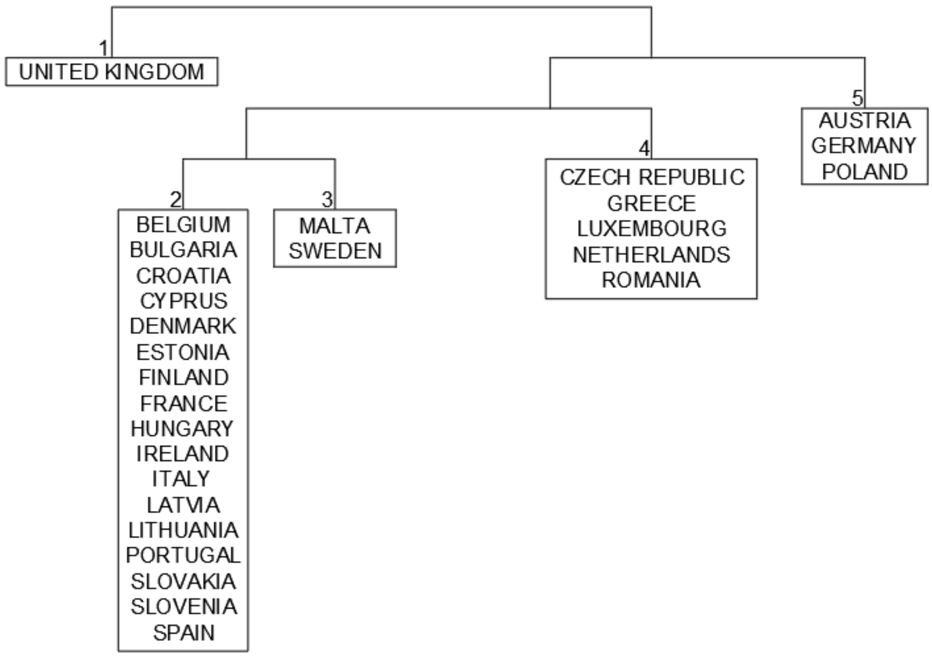


Figure 4.10 - Dendrogram CARGO-GDP

From Figure 4.10 we see that cluster 2 has a reasonable size. It contains 17 countries, so the features in analysis should be very similar in this group. Unfortunately, as it happened in the previous cluster analyses, we could not find a geographic pattern.

5 APPLICATION TO PORTUGAL

Portugal is the most southwest country of the European Union and it has become a member of this community in 1986. We chose Portugal to do a specific study because it was one of the countries which both *t Stat* and *Coefficients* changed signs when looking at its own GDP and at the GDP of Europe. In this chapter we will evaluate the relation GDP-PAX and CARGO-GDP in this specific country.

5.1 Passengers

Table 5.1 is a summary table of the results from the several linear regressions performed between GDP and air passengers in Portugal:

Table 5.1 - Results for PAX-GDP Regression in Portugal

Independent variable	*t Stat*	*Coefficient*
GDP PT lagged 1 y	3.88	3.65
GDP PT lagged 1 y	-3.60	-2.12
GDP EU lagged 1 y	11.75	4.43
GDP PT lagged 1 y	0.06	0.23
GDP PT lagged 2 y	0.98	2.90
GDP PT lagged 1 y	0.64	1.11
GDP PT lagged 2 y	-1.97	-3.19
GDP EU lagged 1 y	1.52	2.78
GDP EU lagged 2 y	1.07	1.96
GDP PT lag 1 + lag 2	4.09	3.42

Independent variable	*t Stat*	*Coefficient*
GDP PT lag 1 + lag 2	-3.91	-2.28
GDP EU lag 1 + lag 2	11.40	4.53
ΔGDP-PT	1.72	0.94
ΔGDP PT	1.39	1.20
ΔGDP EU	-0.40	-0.47
GDP PT lagged 1 y	-3.30	-1.62
GDP EU+ lagged 1 y	1.01	2.60
GDP EU- lagged 1 y	1.05	1.66
GDP PT lagged 2 y	-3.21	-2.20
GDP EU+ lagged 2 y	0.37	1.28
GDP EU- lagged 2 y	1.22	2.51

The previous table show the results from the regressions in the same order as they were mentioned in chapter 3. When considering the GDP of Portugal as an individual variable the results are very different than when we take into account the Europe's GDP, both *t Stat* and *Coefficient*, with Portugal's GDP going from a positive value to a negative one.

The regression with the GDP's rates does not have significant values, just like the regressions where we divided the countries which has more or less socioeconomic relations with Portugal. Both of these showed no advantages in analyze them.

The case where the *Coefficient* has the highest value is when we consider the sum of Europe's GDP lagged one and two years from the air passengers. Thus, we can say that the population of Portugal take a longer time to react to an economic growth, and also the Europe's economy has a higher influence in the number of air trips than Portugal's economy.

5.2 Cargo

For air cargo analysis in Portugal, Table 5.2 as the results of the regressions.

Table 5.2 - Results for CARGO-GDP Regression in Portugal

	t Stat	*Coefficient*
CARGO lagged 1 y	3.18	0.35
CARGO lagged 1 y	2.56	0.32
CARGO lagged 2 y	0.52	0.06
CARGO PT lagged 1 + lagged 2	2.78	0.38

In every variable, with the exception of cargo with lagged two years, *t Stat* has a significant quality. Also, the *Coefficient* are all positive, which means that when cargo rises the GDP of the next years will also increase, however, it would happen with a low percentage.

6 APPLICATION TO SOUTHWEST

In the Southwest Regions of the European Union there are a total of 98 airports, however, only some regions of this area will be studied due to the absence of information from the "Eurostat" database. This is expected because not all regions have an airport and others do not have enough information for us to make our analysis. Although France has a higher area, Spain is the country which has more airports (35). Below, Figure 6.1 presents a map with the name of the regions that will be studied in this chapter.

Figure 6.1 - Map of Europe's Southwest Regions

Hence, considering 46 regions, we executed the linear regression and the results are presented in Appendix G.

After, using the output of the *t Stat* and the *Coefficients* we built the next maps on *QGis*. The color yellow represents *t Stat* values between [-2,2], green represents positive values and red negative values. A more intense color represents a higher value of *t Stat* or *Coefficient*.

There are three maps of each parameter because there are three different independent variables in the multiple regression. Figure 6.2, 6.3 and 6.4 have *t Stat* values.

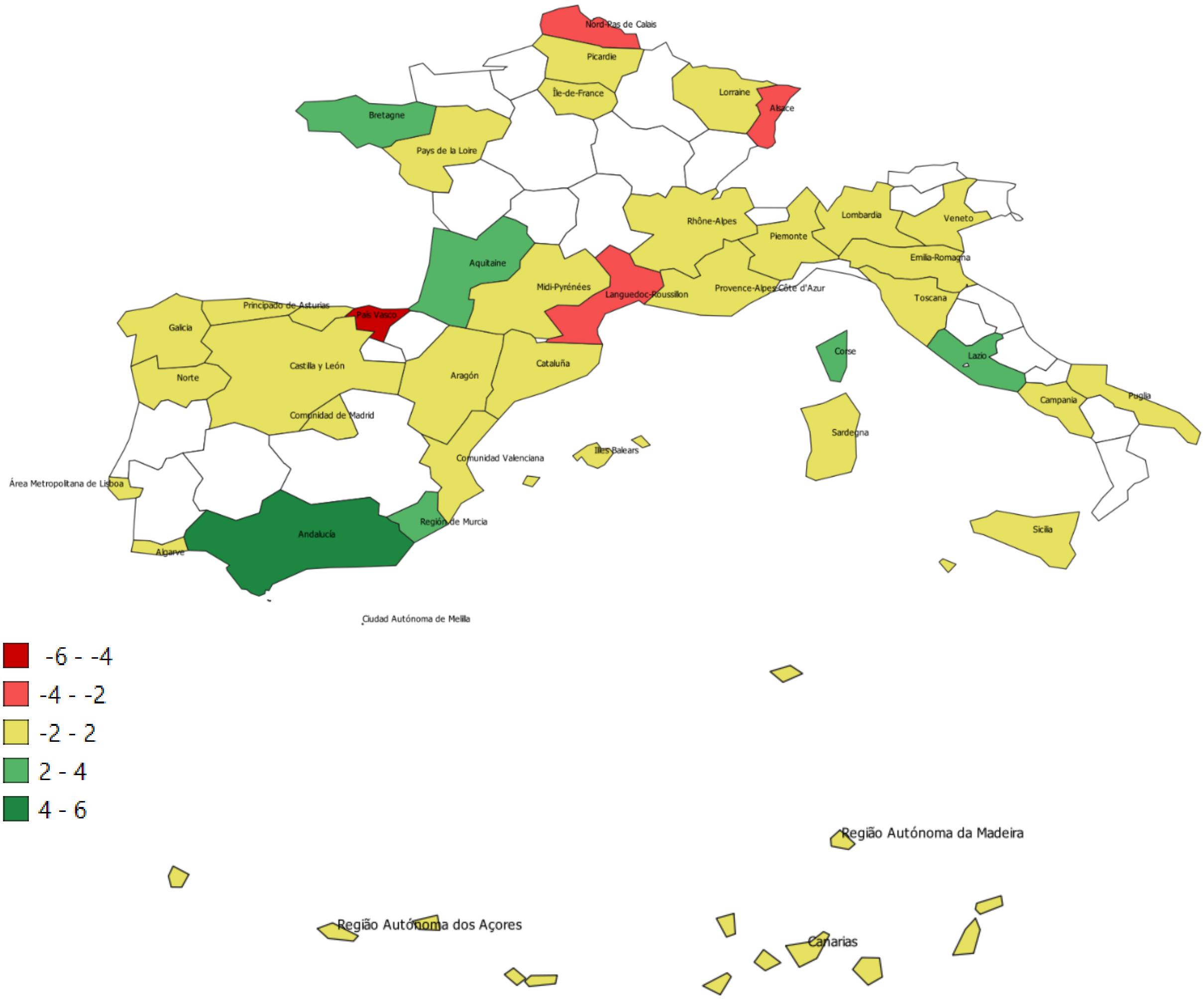


Figure 6.2 - Map of the Southwest Regions of EU with *t Stat* results for the Region's GDP

In the case of the Region's GDP, *t Stat* values have a range that goes from negative six to positive six. The area in yellow is predominant in this map, which means that the GDP of each region does not influence the number of air passengers. The positive influenced regions are localized in the south of Spain, in France and in Italy.

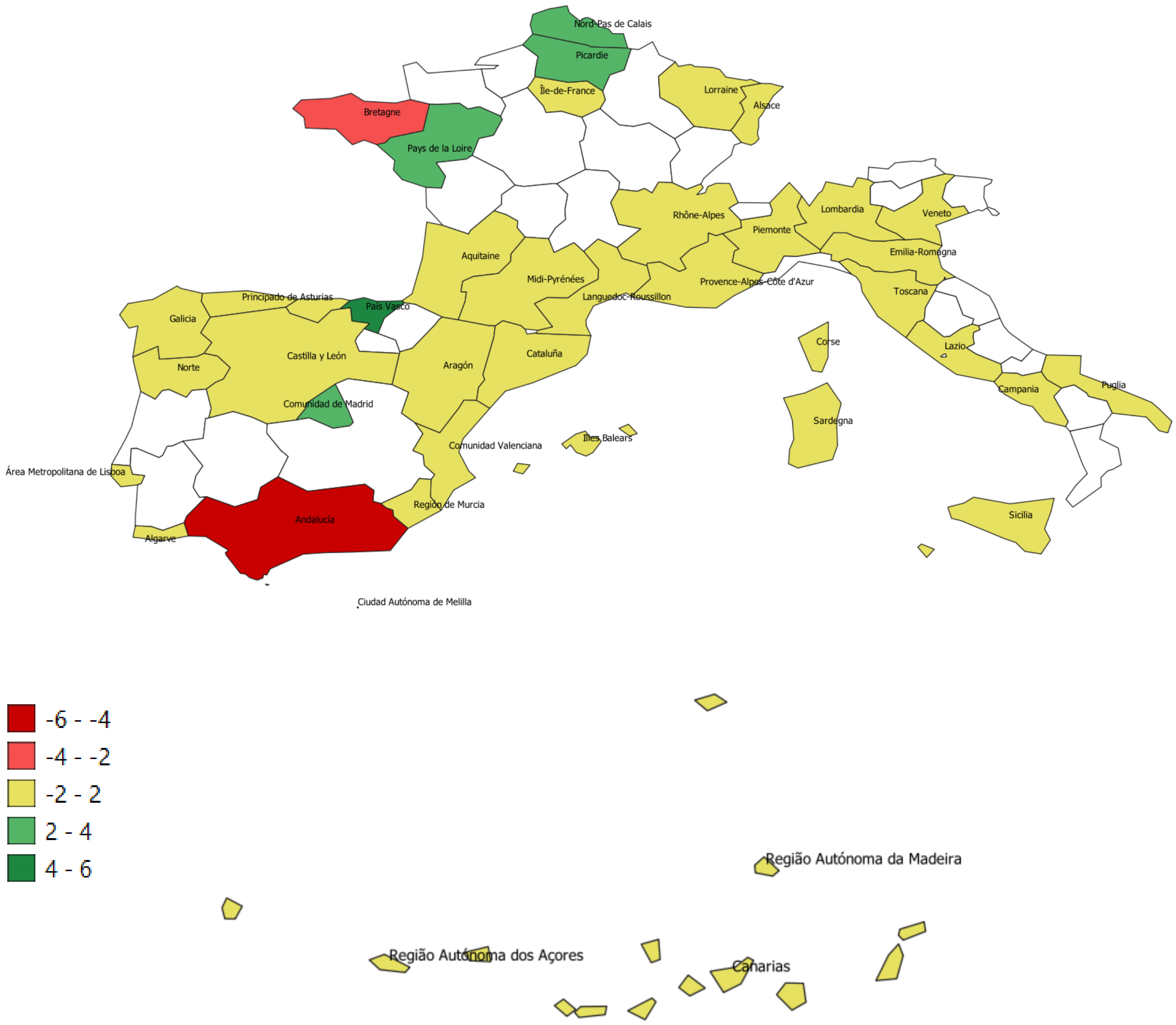


Figure 6.3 - Map of the Southwest Regions of EU with *t Stat* results for the Country's GDP

In figure 6.3, once again, most of the regions are represented with yellow. Conversely of what happens in the first map, Andalucía air passengers grow in a different direction than when observing the growth of the GDP of Spain. The regions where the Country's GDP has significance, positive and negative, in the number of air passengers are only situated in Spain and France. Portugal and Italy are in the interval of non-significance.

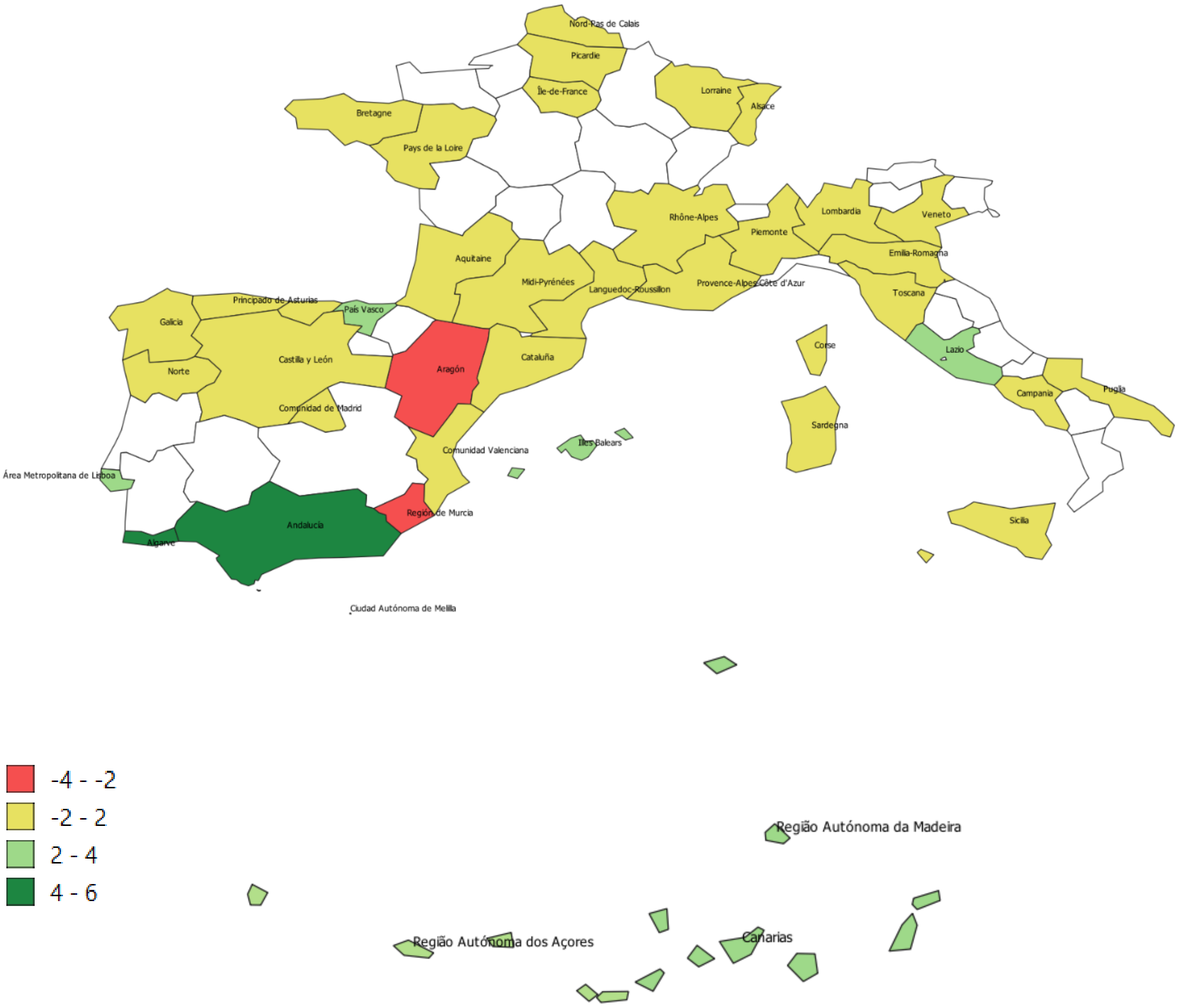


Figure 6.4 - Map of the Southwest Regions of EU with *t Stat* results for the EU's GDP

This map shows that the south of the Iberian Peninsula is the area of the southwest regions where the GDP of the European Union has a greater significance in air passengers, with a lower value but still with relevance, the islands of Portugal and Spain, Portugal's capital, Lazio and País Vasco are represented with a positive value of *t Stat*. One possible explanation for this is that these areas have very high tourist attraction. Aragon and Murcia in Spain are the only regions with a negative value.

Next, Figure 6.5, 6.6 and 6.7 have the maps with the representation of the *Coefficient*. As said before, the parameter will tell us the percentage of air passengers that will increase or decline with the growth of 1% of the GDP.

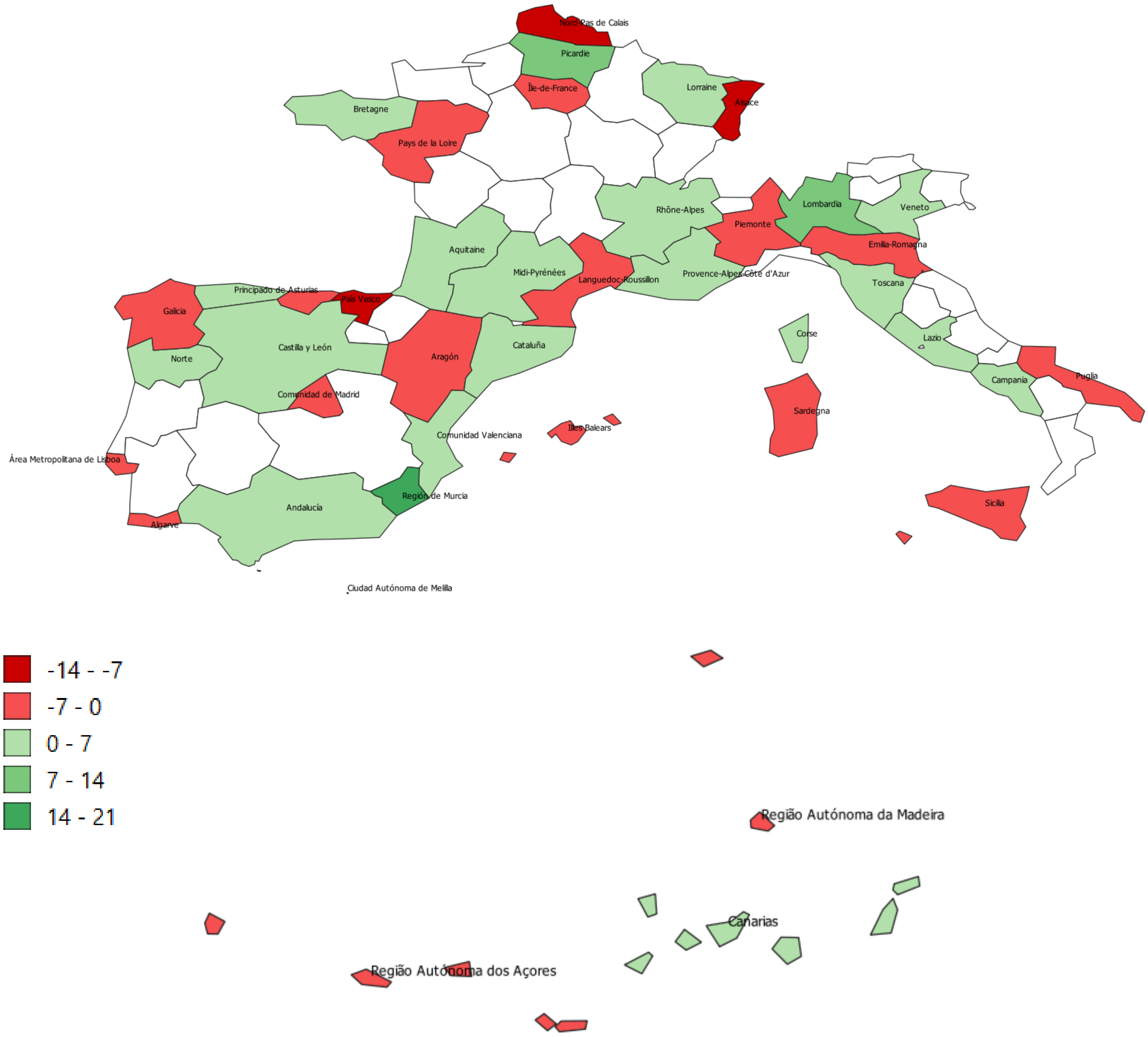


Figure 6.5 - Map of the Southwest Regions of EU with the *Coefficient* results for the Region's GDP

The *Coefficient* of Region's GDP-PAX is represented in the previous picture, i.e. the percentage of air passengers that will variate when the GDP of each region grows. Although in this map the scale is wider, most of the regions have the *Coefficient* between the interval zero to seven, which means that the variation is not high.

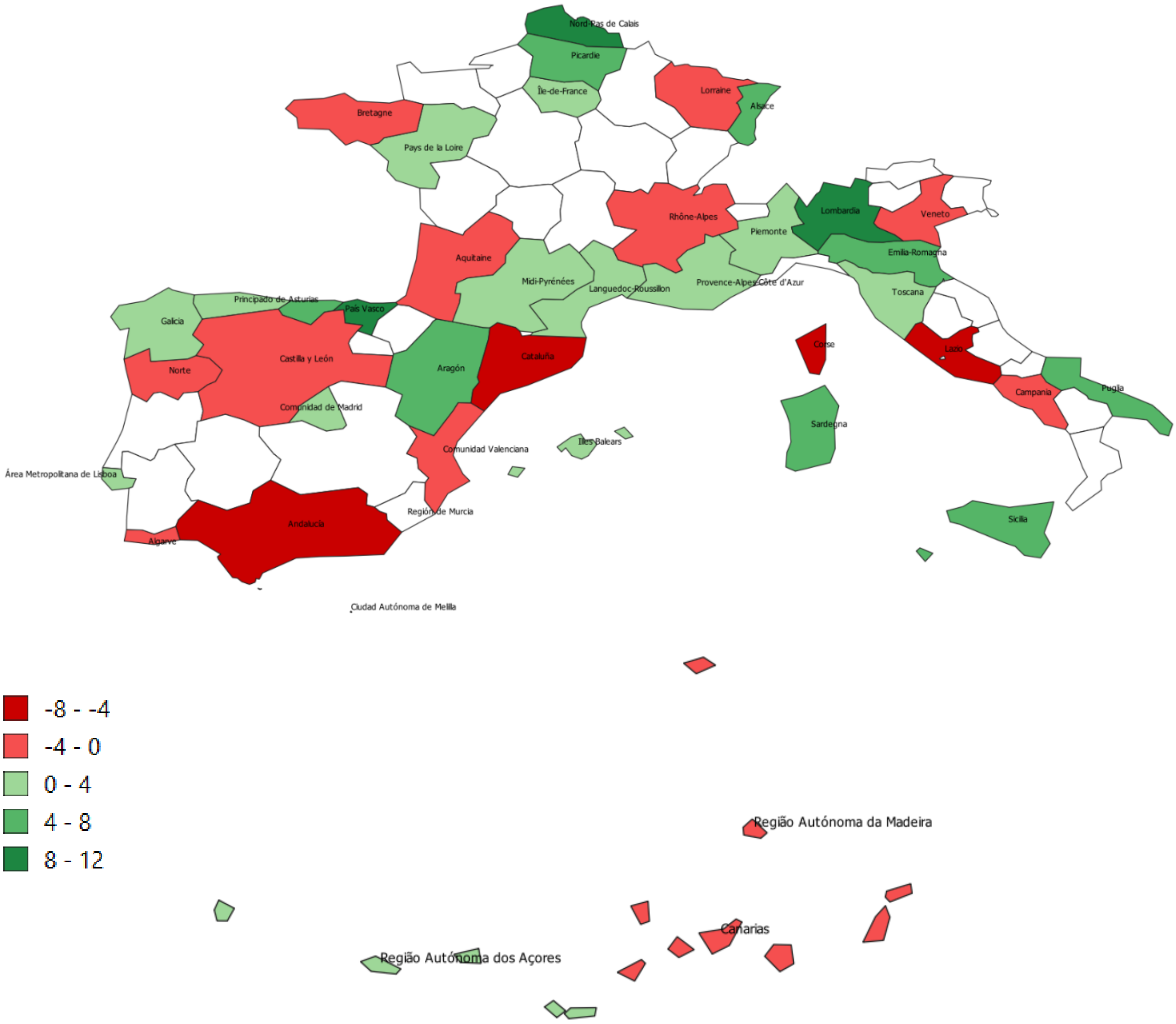


Figure 6.6 - Map of the Southwest Regions of EU with the *Coefficient* results for the Country's GDP

This is the map where the *Coefficients* have a greater variation, i.e. there is not an interval of *Coefficients* values where many of countries are inserted. As is happened in the *t Stat* analysis for the GDP of the Country, Andalucía has a negative value. The regions with a stronger Country's GDP-PAX relation are located in the north of Spain, France and Italy.

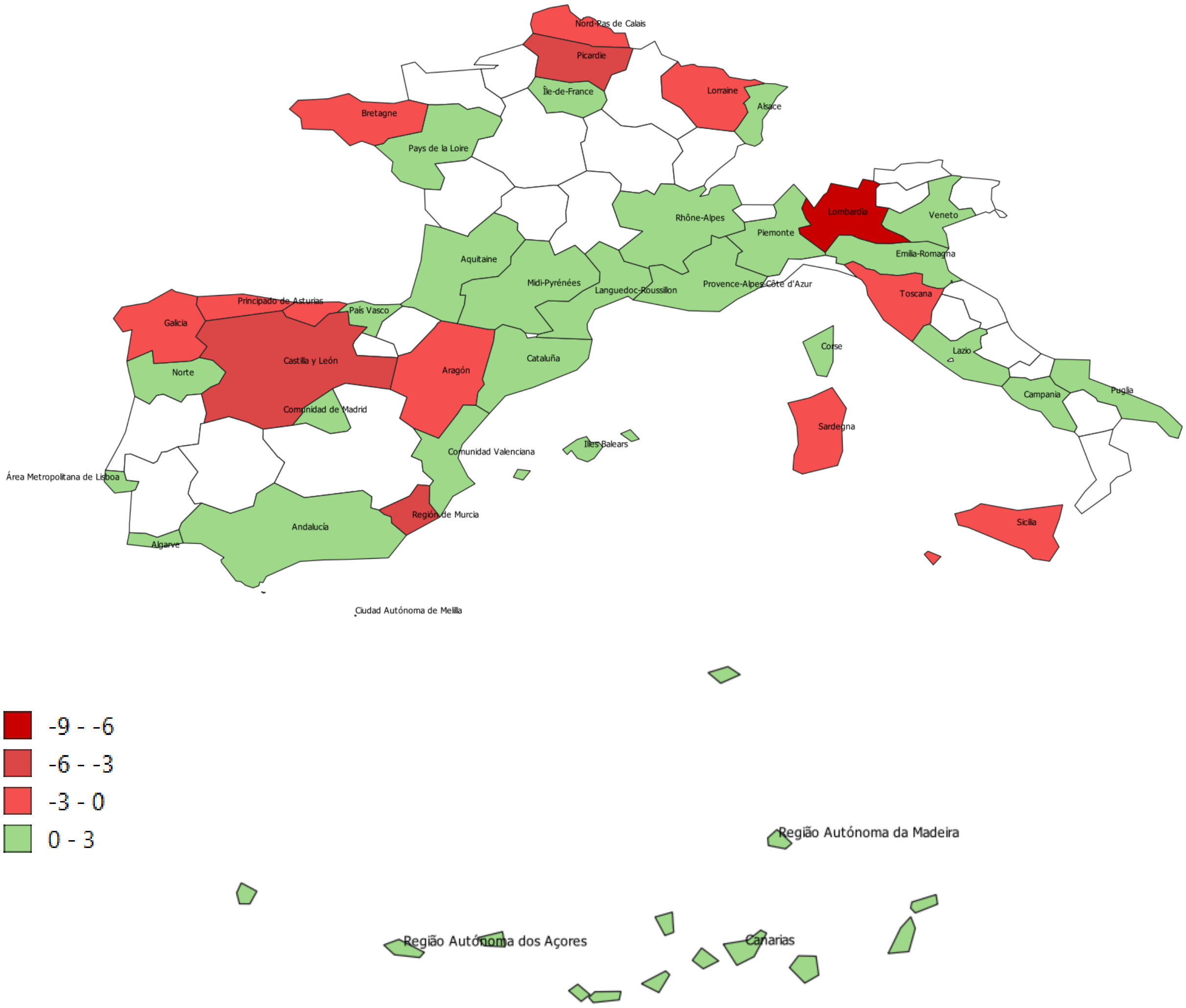


Figure 6.7 - Map of the Southwest Regions of EU with the *Coefficient* results for the EU's GDP

More than half of the regions in this map have a positive value of *Coefficients*, however these percentages are poor. The number of passengers will only increase until 3% when the GDP of the rest of the countries from EU grows one unit. While with negative values the percentages are higher, meaning that when the GDP of Europe increases, the number of people traveling by air transport declines.

7 CONCLUSIONS

In this dissertation we did a study focused in finding whether there is a relationship between the economy and air transport in the European Union, including air passengers and air cargo.

The first part of this work included a brief introduction of the theme, previous studies which were related to our work and a description of our methodology. The method adopted to do this analysis was linear regression. The values of *t Stat* and *Coefficient* were used to evaluate, respectively, the significance of the relationship and the percentage of variation of the dependent variable. At country level analysis we also applied a cluster analysis in order to group the countries which had similar characteristics, in order to find a geographical pattern.

The second part is dedicated to the applications of the methodology. Initially, we studied the relation at country level, then the individual case of Portugal, and at last the Southwest Regions of Europe. Our work is different when compared to others before. We took in consideration lagged variables and also the GDP of neighbor nations or regions in order to see if they have any influence on air passengers. In the case of air passengers, we wanted to know if the growth of GDP has any influence in the number of air passengers of the following year. For the air cargo analysis want to find if there is a relationship in the opposite direction, i.e. we want to see the reaction of the GDP when the cargo of the previous year increases. In air cargo we only analyzed the economy of each country.

Overall, most of the questions we mentioned in the objectives were answered with the results of the applications. The results showed that the relation between the economy and air transport exists, although it variates from country to country.

In many cases of the air passenger analysis, we saw that, when looking at the *t Stat* of the country's GDP, its influence on air passengers was not significant, while the GDP of the other countries of the EU was, and the same would happen on the other direction. However, we can say that when considering other countries economy both the significance and the percentage of variation have higher values than when we consider the economy of the country in study. As for the geographic pattern of this relationships in the EU we were not able to find it. The cluster analysis gave us groups with similar characteristics but geographically they were not related.

With air cargo analysis we could see that there is a relationship between air cargo and GDP. This relation is not so strong in every country, i.e. *t Stat* values are almost all above 2 but the percentage of variation, represented by the *Coefficient,* is not very high. The cluster analysis also did not show a geographic pattern.

The analysis of the particular case of Portugal showed that by adding lagged GDP of the EU from one and two years the relation GDP-PAX is stronger. The air cargo regression which had better results was the one with the sum of lagged one and two years of cargo in relation to the GDP.

In the European Southwest Regions, where we only analyzed air passengers, we had as independent variables the GDP of the Region, Country and EU, the GDP of the European Union. The EU's GDP is the one which has a greater significant in the number of air passengers for more regions.

Some suggestion to further work is a Granger Causality analysis. When applying the Granger approach, it will possible to determine if there is a causal relationship between the variables we studied. In other words, this method will say whether the growth in the economy will help predict the number of air passengers, or the same but in the opposite direction. Also, whether the growth in air cargo will predict the growth on the economy.

BIBLIOGRAPHIC REFERENCE

Chang, Y. and Chang, Y. (2009), "Air cargo expansion and economic growth: Finding the empirical link", *Journal of Air Transport Management*, 15, 264-266.

Chi, J. and Baek, J. (2013), "Dynamic relationship between air transport demand and economic growth in the United States: A new look", *Transport Policy*, 29, 257-260.

EUROSTAT @. https://ec.europa.eu/eurostat. Access February 2019.

Granger, C. (1969), "Investigating Causal Relations by Econometric Models and Cross-spectral Methods", *The Econometric Society*, 37, 424-438.

INE @. https://www.ine.pt/xportal/xmain?xpgid=ine_main&xpid=INE . Access October 2018

Ishutkina, M. (2009), "Analysis of the Interaction Between Air Transportation and Economic Activity: A Worldwide Perspective", Doctoral Dissertation, Department of Aeronautics & Astronautics of Massachusetts Institute of Technology, USA.

Kasarda, J. and Green, J. (2005), "Air cargo as an economic development engine: A note on opportunities and constraints", *Journal of Air Transport Management*, 11, 459-462.

Marazzo, M, Scherre, R. and Fernandes, E. (2010), "Air transport demand and economic growth in Brazil: A time series analysis", *Transportation Research* Part E, 46, 261-269.

Profillidis, V. and Botzoris, G. (2015), "Air passenger transport and economic activity", *Journal of Air Transport Management*, 49, 23-27.

Rodríguez -Brindis, M., Mejía-Alzate, M. and Zapata-Aguirre, S. (2015), "La causalidade entre el crecimiento económico y la expansión del transporte aéreo: un análisis empírico para Chile", *Revista de Economía*, 18, 127-144.

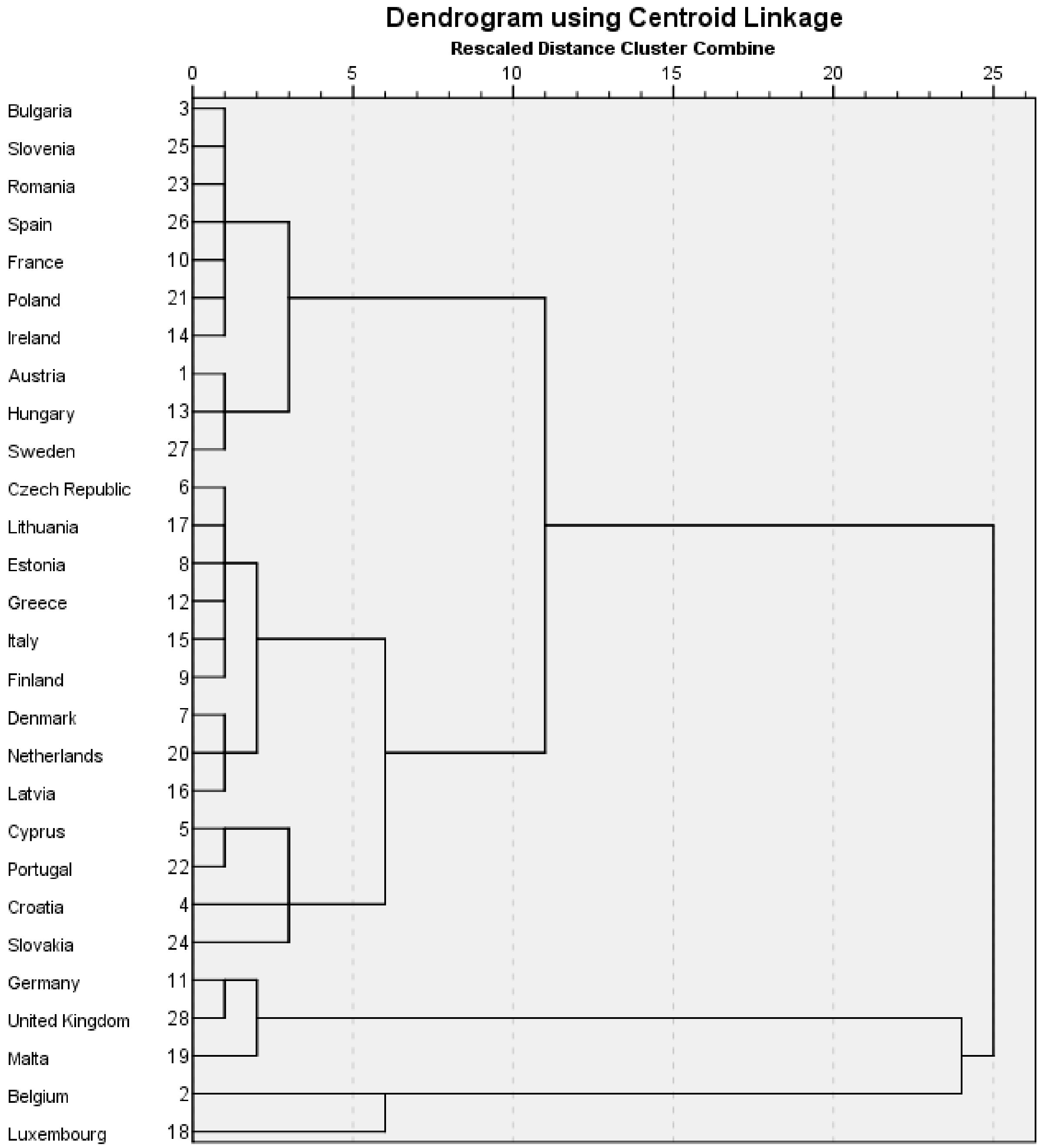
THE WORLD BANK @. https://data.worldbank.org/. Access February 2019.

APPENDIX A - AGGLOMERATIVE SCHEDULE GDP_ COUNTRY-PAX

Agglomeration Schedule

Stage	Cluster Combined		Coefficients	Stage Cluster First Appears		Next Stage
	Cluster 1	Cluster 2		Cluster 1	Cluster 2	
1	3	25	,008	0	0	15
2	8	12	,010	0	0	4
3	6	17	,010	0	0	13
4	8	15	,012	2	0	6
5	7	20	,026	0	0	16
6	8	9	,027	4	0	13
7	23	26	,035	0	0	15
8	10	21	,038	0	0	12
9	1	13	,042	0	0	10
10	1	27	,057	9	0	21
11	11	28	,080	0	0	19
12	10	14	,091	8	0	17
13	6	8	,106	3	6	18
14	5	22	,139	0	0	20
15	3	23	,143	1	7	17
16	7	16	,183	5	0	18
17	3	10	,191	15	12	21
18	6	7	,321	13	16	23
19	11	19	,347	11	0	26
20	4	5	,494	0	14	22
21	1	3	,533	10	17	25
22	4	24	,622	20	0	23
23	4	6	1,283	22	18	25
24	2	18	1,418	0	0	26
25	1	4	2,533	21	23	27
26	2	11	5,742	24	19	27
27	1	2	6,012	25	26	0

APPENDIX B - DENDROGRAM GDP_COUNTRY-PAX

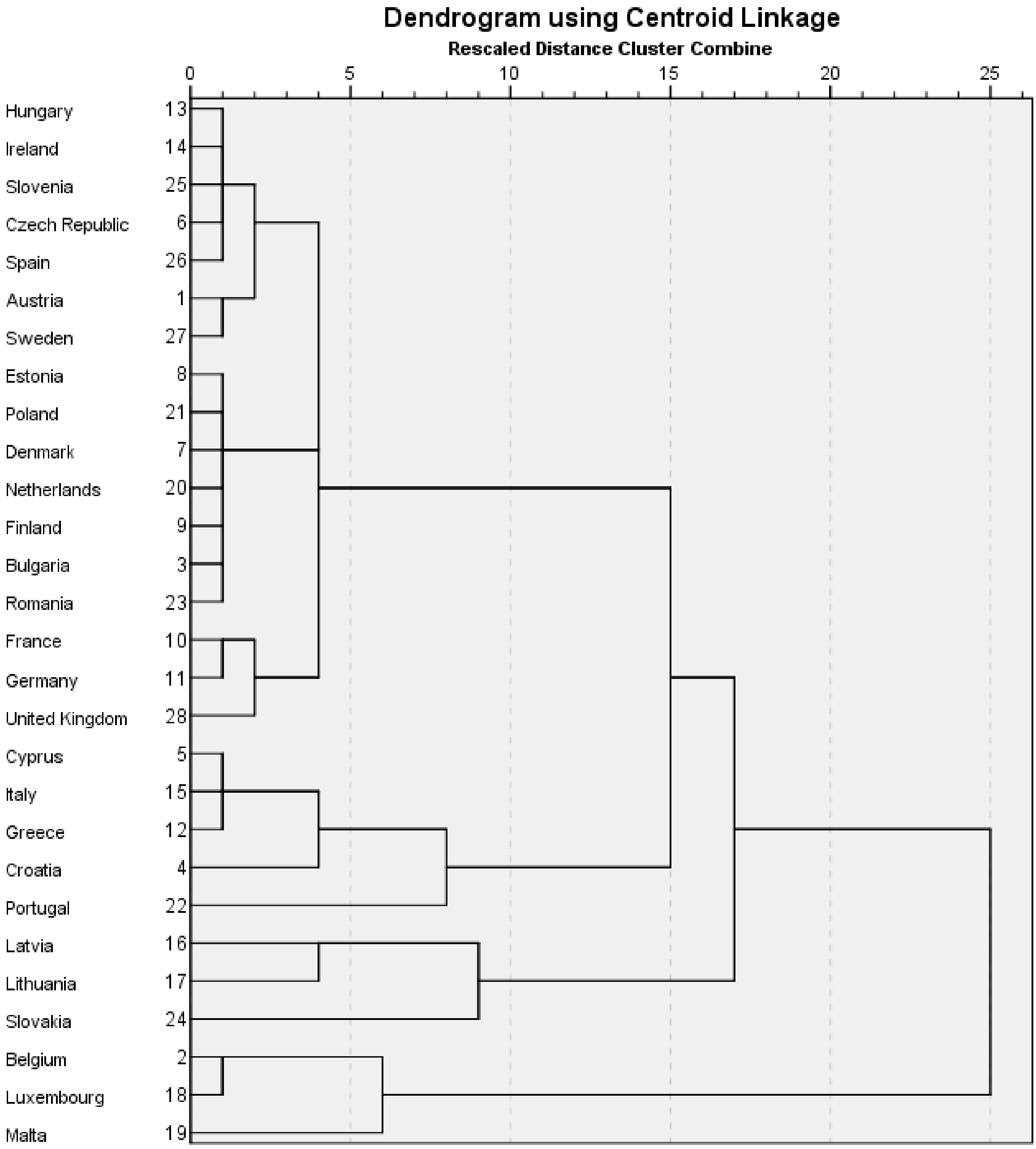


APPENDIX C - AGGLOMERATIVE SCHEDULE GDP_EU-PAX

Agglomeration Schedule

Stage	Cluster Combined		Coefficients	Stage Cluster First Appears		Next Stage
	Cluster 1	Cluster 2		Cluster 1	Cluster 2	
1	13	14	,006	0	0	4
2	7	20	,009	0	0	7
3	1	27	,017	0	0	17
4	13	25	,020	1	0	11
5	8	21	,029	0	0	15
6	10	11	,029	0	0	16
7	7	9	,040	2	0	13
8	3	23	,045	0	0	13
9	6	26	,047	0	0	11
10	2	18	,048	0	0	22
11	6	13	,062	9	4	17
12	5	15	,085	0	0	14
13	3	7	,103	8	7	15
14	5	12	,177	12	0	20
15	3	8	,235	13	5	18
16	10	28	,302	6	0	19
17	1	6	,303	3	11	18
18	1	3	,795	17	15	19
19	1	10	,712	18	16	25
20	4	5	,869	0	14	23
21	16	17	,961	0	0	24
22	2	19	1,454	10	0	27
23	4	22	1,845	20	0	25
24	16	24	2,289	21	0	26
25	1	4	3,746	19	23	26
26	1	16	4,233	25	24	27
27	1	2	6,430	26	22	0

APPENDIX D - DENDROGRAM GDP_EU-PAX

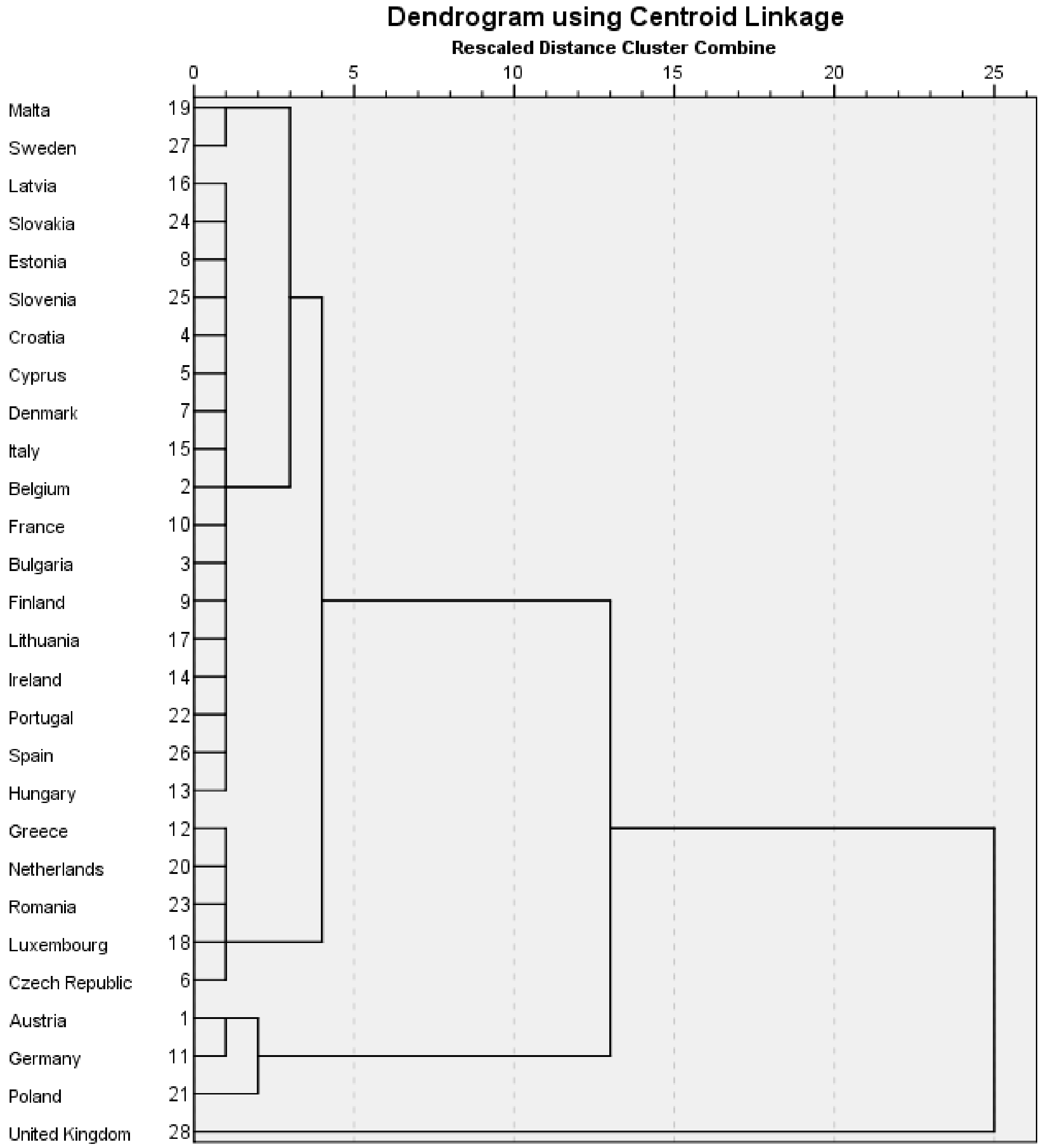


APPENDIX E - AGGLOMERATIVE SCHEDULE GOODS-GDP_COUNTRY

Agglomeration Schedule

Stage	Cluster Combined		Coefficients	Stage Cluster First Appears		Next Stage
	Cluster 1	Cluster 2		Cluster 1	Cluster 2	
1	19	27	,004	0	0	24
2	14	22	,006	0	0	8
3	16	24	,006	0	0	9
4	4	5	,007	0	0	12
5	8	25	,011	0	0	9
6	2	10	,018	0	0	13
7	12	20	,023	0	0	16
8	14	26	,024	2	0	15
9	8	16	,027	5	3	14
10	3	9	,028	0	0	11
11	3	17	,027	10	0	13
12	4	7	,030	4	0	14
13	2	3	,071	6	11	21
14	4	8	,073	12	9	18
15	13	14	,075	0	8	21
16	12	23	,110	7	0	19
17	1	11	,143	0	0	23
18	4	15	,165	14	0	22
19	12	18	,254	16	0	20
20	6	12	,507	0	19	25
21	2	13	,566	13	15	22
22	2	4	,628	21	18	24
23	1	21	,674	17	0	26
24	2	19	1,653	22	1	25
25	2	6	2,088	24	20	26
26	1	2	7,992	23	25	27
27	1	28	16,403	26	0	0

APPENDIX F - DENDROGRAM GOODS-GDP_COUNTRY



APPENDIX G - RESULTS OF THE MULTIPLE REGRESSION FOR THE SOUTHWEST REGIONS

FRANCE - REGION		*t Stat*	*Coef*	*Adjusted R^2*
Île de France	GDP_Region	-0.461	-0.143	0.907
	GDP_FR	0.900	0.625	
	GDP_EU	0.766	0.398	
Picardie	GDP_Region	1.744	7.298	0.929
	GDP_FR	2.071	6.790	
	GDP_EU	-1.989	-5.725	
Nord - Pas-de-Calais	GDP_Region	-3.574	-9.126	0.851
	GDP_FR	3.591	10.034	
	GDP_EU	-0.327	-0.477	
Lorraine	GDP_Region	1.845	2.014	0.705
	GDP_FR	-0.497	-0.528	
	GDP_EU	-1.022	-1.369	
Alsace	GDP_Region	-2.311	-7.509	0.790
	GDP_FR	1.771	4.403	
	GDP_EU	1.620	2.634	
Pays de la Loire	GDP_Region	-0.878	-1.893	0.904
	GDP_FR	2.272	2.310	
	GDP_EU	1.278	1.992	
Bretagne	GDP_Region	2.697	4.539	0.487
	GDP_FR	-2.192	-3.191	
	GDP_EU	-0.319	-0.528	
Aquitaine	GDP_Region	2.132	2.501	0.907
	GDP_FR	-1.256	-1.477	
	GDP_EU	0.256	0.234	
Midi-Pyrénées	GDP_Region	0.670	0.487	0.921
	GDP_FR	0.462	0.413	
	GDP_EU	0.073	0.043	
Rhône-Alpes	GDP_Region	0.811	1.182	0.959
	GDP_FR	-0.365	-0.574	
	GDP_EU	1.074	0.505	
Languedoc-Roussillon	GDP_Region	-3.085	-3.326	0.586
	GDP_FR	1.780	3.368	
	GDP_EU	0.719	0.669	

Provence-Alpes-Côte d'Azur	GDP_Region	0.291	0.267	0.862
	GDP_FR	0.155	0.201	
	GDP_EU	0.587	0.403	
Corse	GDP_Region	2.165	2.311	0.882
	GDP_FR	-1.770	-5.074	
	GDP_EU	1.981	2.732	

SPAIN - REGION		*t Stat*	*Coef*	*Adjusted R^2*
Galicia	GDP_Region	-0.165	-0.918	0.700
	GDP_SP	0.379	2.213	
	GDP_EU	-0.780	-0.433	
Principado de Asturias	GDP_Region	0.009	0.024	0.655
	GDP_SP	0.599	1.823	
	GDP_EU	-0.977	-1.323	
Cantabria	GDP_Region	-0.325	-2.335	0.883
	GDP_SP	0.735	5.667	
	GDP_EU	-0.514	-1.124	
País Vasco	GDP_Region	-4.038	-13.174	0.837
	GDP_SP	4.295	12.021	
	GDP_EU	3.389	2.024	
Aragón	GDP_Region	-0.418	-2.629	0.811
	GDP_SP	0.898	6.133	
	GDP_EU	-2.168	-2.083	
Comunidad de Madrid	GDP_Region	-1.791	-2.616	0.789
	GDP_SP	2.528	3.470	
	GDP_EU	0.527	0.286	
Castilla y León	GDP_Region	0.643	6.059	0.628
	GDP_SP	-0.227	-2.078	
	GDP_EU	-1.621	-3.619	
Cataluña	GDP_Region	1.242	6.657	0.883
	GDP_SP	-1.145	-5.796	
	GDP_EU	1.135	0.602	
Comunidad Valenciana	GDP_Region	0.989	1.863	0.808
	GDP_SP	-0.672	-1.388	
	GDP_EU	1.618	1.031	
Illes Balears	GDP_Region	-1.648	-2.628	
	GDP_SP	1.460	1.776	0.828

	GDP_EU	3.517	1.744	
Andalucía	GDP_Region	4.878	6.783	0.872
	GDP_SP	-4.586	-7.192	
	GDP_EU	4.336	1.508	
Región de Murcia	GDP_Region	2.124	21.956	0.811
	GDP_SP	-1.685	-19.658	
	GDP_EU	-2.316	-3.131	
Ciudad Autónoma de Melilla	GDP_Region	-0.273	-0.367	0.656
	GDP_SP	0.453	0.444	
	GDP_EU	1.390	0.837	
Canarias	GDP_Region	0.955	1.579	0.536
	GDP_SP	-1.258	-1.888	
	GDP_EU	3.733	1.122	

ITALY - REGION		*t Stat*	*Coef*	*Adjusted R^2*
Piemonte	GDP_Region	-0.517	-0.653	0.748
	GDP_IT	0.888	1.241	
	GDP_EU	0.674	0.294	
Lombardia	GDP_Region	0.376	8.747	0.322
	GDP_IT	0.552	11.338	
	GDP_EU	-1.020	-9.016	
Veneto	GDP_Region	1.095	2.789	0.924
	GDP_IT	-0.181	-0.418	
	GDP_EU	0.216	0.126	
Emilia-Romagna	GDP_Region	-1.129	-4.727	0.889
	GDP_IT	1.532	4.348	
	GDP_EU	1.726	2.919	
Toscana	GDP_Region	0.628	4.823	0.896
	GDP_IT	0.281	1.843	
	GDP_EU	-0.600	-1.053	
Lazio	GDP_Region	2.299	3.721	0.955
	GDP_IT	-1.833	-4.885	
	GDP_EU	3.130	2.129	
Campania	GDP_Region	0.597	0.642	0.848
	GDP_IT	-0.308	-0.435	
	GDP_EU	1.861	1.034	

Puglia	GDP_Region GDP_IT GDP_EU	-0.846 1.130 1.992	-6.145 5.852 2.461	0.851
Sicilia	GDP_Region GDP_IT GDP_EU	-0.323 0.821 -0.305	-1.345 5.276 -0.721	0.775
Sardegna	GDP_Region GDP_IT GDP_EU	-0.247 1.045 -0.451	-1.001 5.528 -0.534	0.874

PORTUGAL - REGION		*t Stat*	*Coef*	*Adjusted R^2*
Norte	GDP_Region GDP_PT GDP_EU	1.443 -1.948 1.968	3.884 -3.247 2.406	0.927
Algarve	GDP_Region GDP_PT GDP_EU	-0.770 -0.278 6.094	-0.623 -0.286 1.867	0.865
Área Metropolitana de Lisboa	GDP_Region GDP_PT GDP_EU	-1.156 0.689 2.382	-2.078 1.884 2.105	0.931
Região Autónoma dos Açores	GDP_Region GDP_PT GDP_EU	-1.302 1.182 2.627	-3.217 3.631 2.088	0.691
Região Autónoma da Madeira	GDP_Region GDP_PT GDP_EU	-0.238 -0.436 2.785	-0.163 -0.576 1.161	0.594