Group Members: Patrick Aquino, Jiahui Chen, Xiyu Chen, Ning Hu, Mengying Yu

Analysis EU Green House Consumption since 1990

ANLY511 – Final Project 2019

Group 1

Date: 12/9/19

Table of Contents

[1. Introduction 1](#_Toc26754178)

[2. Analysis and Statistical Methods 2](#_Toc26754179)

[**2.1 Dataset and Data Preparation** 2](#_Toc26754180)

[**2.2 Data Visualization** 2](#_Toc26754181)

[**2.3 Statistical Methods** 10](#_Toc26754182)

[3. Results 12](#_Toc26754183)

[4.Conclusion 20](#_Toc26754184)

[Reference: 21](#_Toc26754185)

[Appendix 21](#_Toc26754186)

# 1. Introduction

With the upcoming 2020 US Election, one of the most discussed topics during political debates is how will the government with the climate change issue. One of the main effects of climate change is the rising average temperature of the Earth, which is also called Global Warming. For centuries, since the industrialization era, people have blindly acknowledged the human involvement in the sudden rise of Earth’s temperature. However as more extreme and unusual climate catastrophes are happening—and more well-known people have become more willing of voicing their opinions—turning the other cheek method are placing wildlife and human lives in danger.

One of the biggest news this summer was the uncontrolled fire that has been devouring the Amazon forest. Most people outside of Brazil and its neighboring countries are unaware of this event until late summer, when in fact the Amazon forest has been on fire since the start of 2019 year—and currently is still on fire. Two summers ago, this problem was present in Western Europe. Mediterranean countries such as Spain, Portugal, Italy and Greece had suffered millions in damages due to sudden emergence of wildfires. Even in the very northern part of Europe, Scandinavian countries such as Sweden and Norway were also suffering from wildfires.

Wildfires are not the only ‘natural’ disasters that are ignited by Global Warming. Due to rise of temperature averages, hurricanes and typhoons are more prevalent in the Pacific Ocean. The rising of sea water levels is also linked from the melting of ice glaciers in both North and South poles. One study shows that if all the ice in the North and South poles would completely melt, the sea water levels will rise around 40-60 feet, which will submerge most populated coastal cities such as Los Angeles, Boston, Washington D.C., London, Paris, etc.

However, linking human causation to Global Warming is still being refuted. Out of every 100 scientists, 98 of them would say that humans have involvement in climate change, meanwhile the other 2 would deny it. Regardless, the hypothesis cannot be proven and can only be rejected. However, my group has the inner assumption that humans do in fact cause and are linked to Global Warming. We plan to examine the emission of GreenHouse Gas from different industries in the European Union since 1990 dataset. We will perform exploratory data analysis, conduct statistical tests and report results and conclusion respectively.

# 2. Analysis and Statistical Methods

### **2.1 Dataset and Data Preparation**

This project uses the the EU energy statistical country datasheets as dataset, which is from an official website of European Union - “ec.europa.eu.” This dataset covers from 1997 to 2017, including all EU member states and present in a comparative format long term time-series of the energy balances, electricity and heat generation, main energy indicators, cogeneration, transport fuels and greenhouse gases emissions. Besides studying the data of the European Union's 28 member countries, this project would focus on France, UK, Finland and Slovakia when conducting some statistical methods and data visualization.

This project extracts the data of CO2 emissions and Greenhouse Gas emissions from this dataset across 28 countries, around 60,000 data. The number of variables about CO2 emissions and Greenhouse Gas emissions is 75, such as “Fuel Combustion Activities”, “Energy Industries”, etc.

### **2.2 Data Visualization**

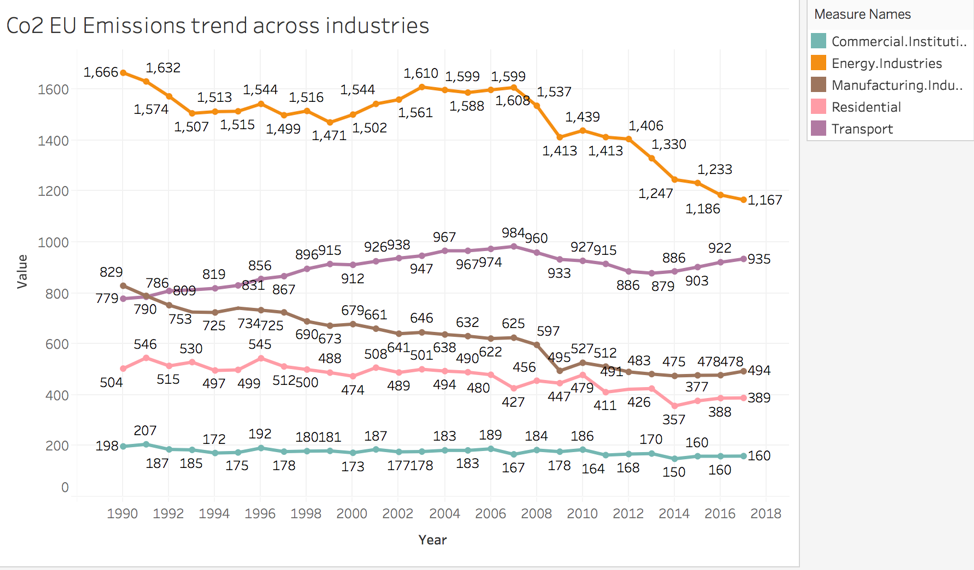


Figure 1: This is the overview of Co2 emission across the 5 major industries in EU 28 countries. It shows us that from 1990 to 2018, Energy industry has a relatively large decrease in CO2 emission, while those of Manufacturing and Residential industries are relatively smaller. On the contrary, the level of CO2 emission in Transport industry seems to rise throughout 1990 to 2007. Though it drops for a certain degree during 2008 to 2013, it rises again from 2014 to 2018. Commercial Industry seems to have no significant change in CO2 emission.

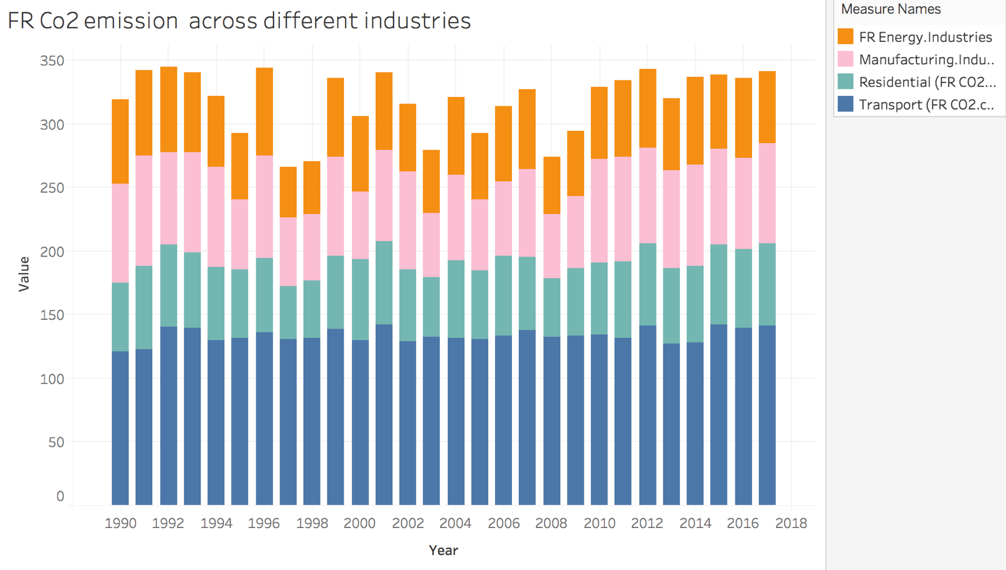
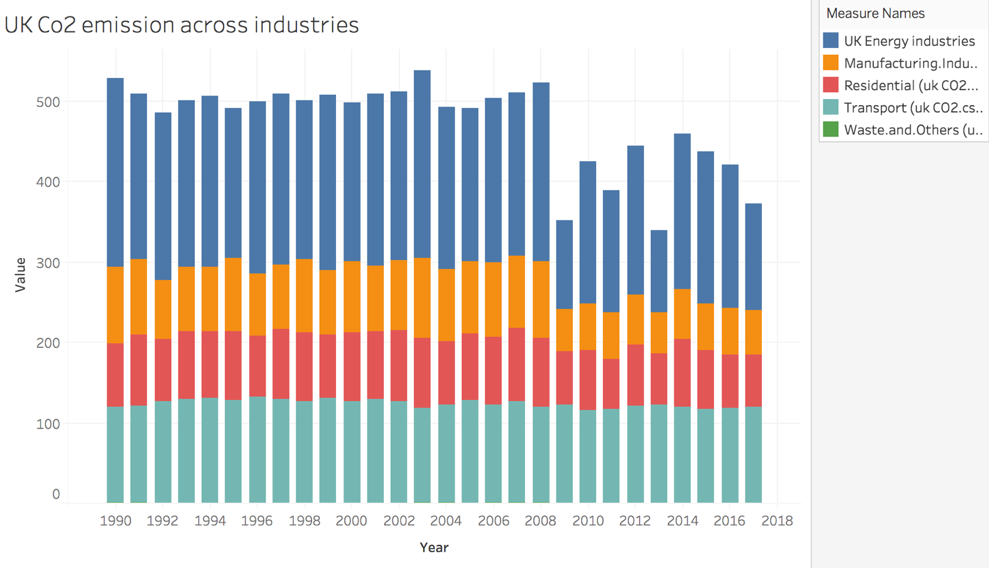


Figure 2: Among various industries in France, Transportation seems to take up a relatively larger portion of the overall CO2 emission, while other industries seem to share a relatively similar portion. Over the course of years shown in this figure, there seems to be no significant change in the amount of CO2 emissions in the given industries.

Figure 3: Among various industries in UK, Energy industry seems to take up a relatively larger portion of the overall CO2 emissions, and the second largest portion is transportation. Over the course of years shown in this figure, it seems that CO2 emission in Manufacturing and Energy industry dropped a little from 2009 to 2018. There is no visible change in transportation.

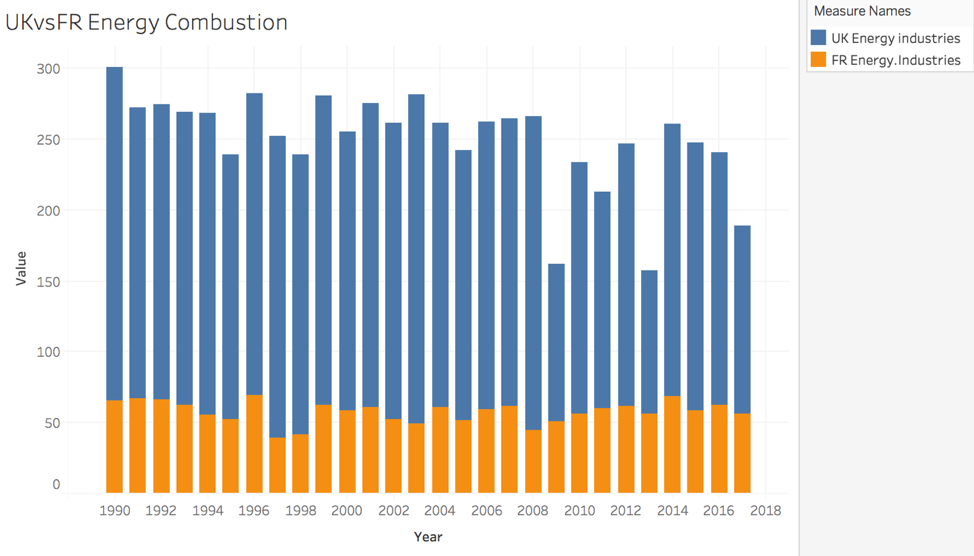


Figure 4: UK versus FR CO2 emission in energy industry

Figure 4 shows that UK has much more CO2 emission than that of FR in Energy Industry.

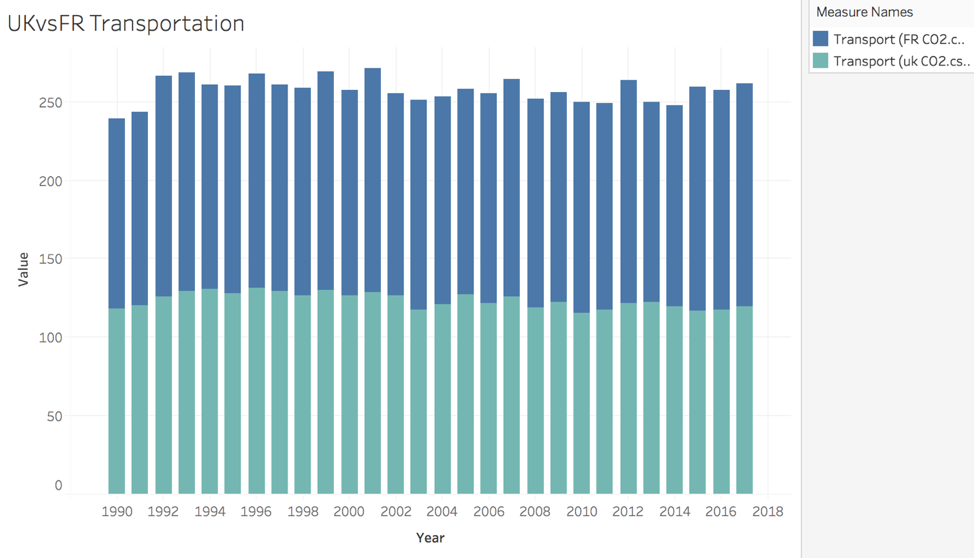


Figure 5: UK versus FR CO2 emission in Transportation. It seems that both UK and France have a similar amount of CO2 emission in Transportation Industry.

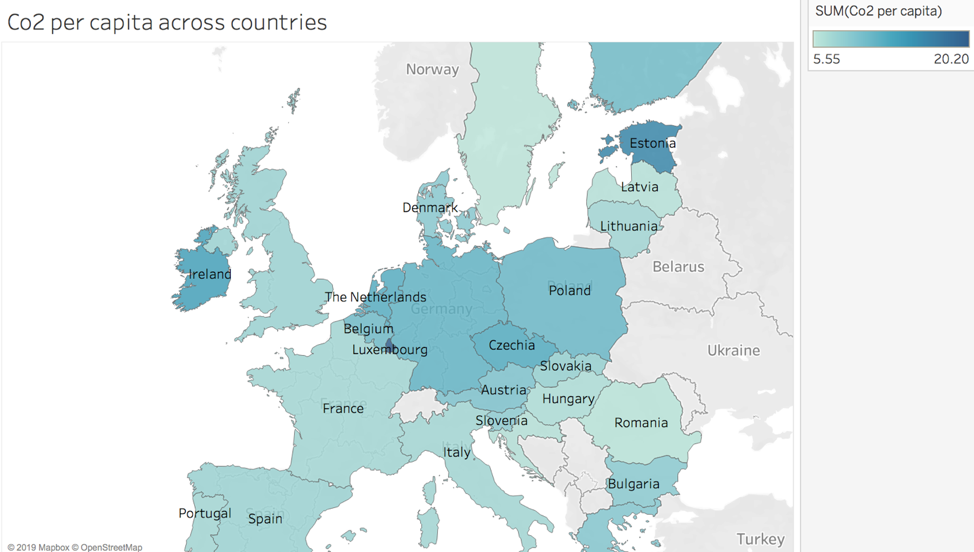


Figure 6: CO2 emissions per capita across EU28 countries. Among these countries, Estonia, Ireland and Poland have relatively higher CO2 emission while Romania and Hungary have relatively less.

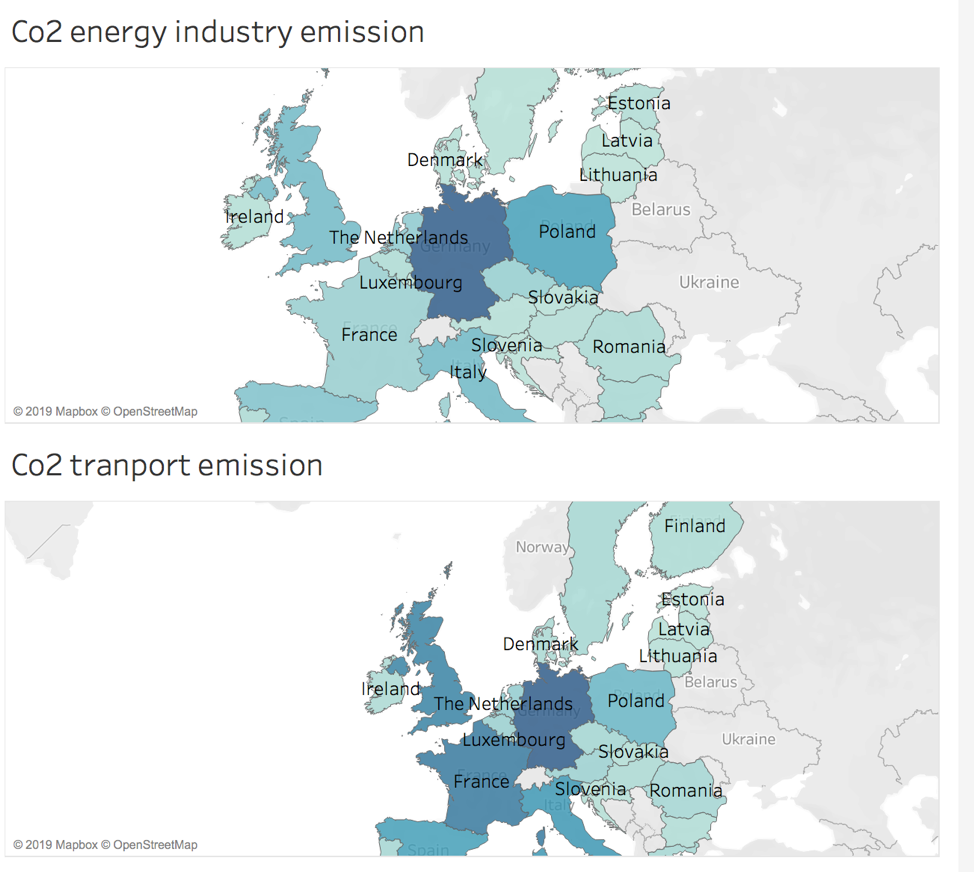


Figure 7: Comparison between energy and transportation industry across the EU's 28 countries.

Similarity between these two maps: Germany and Poland have higher CO2 emission in both industries. However, France and Ireland have higher CO2 emissions in transportation while energy industry does not.

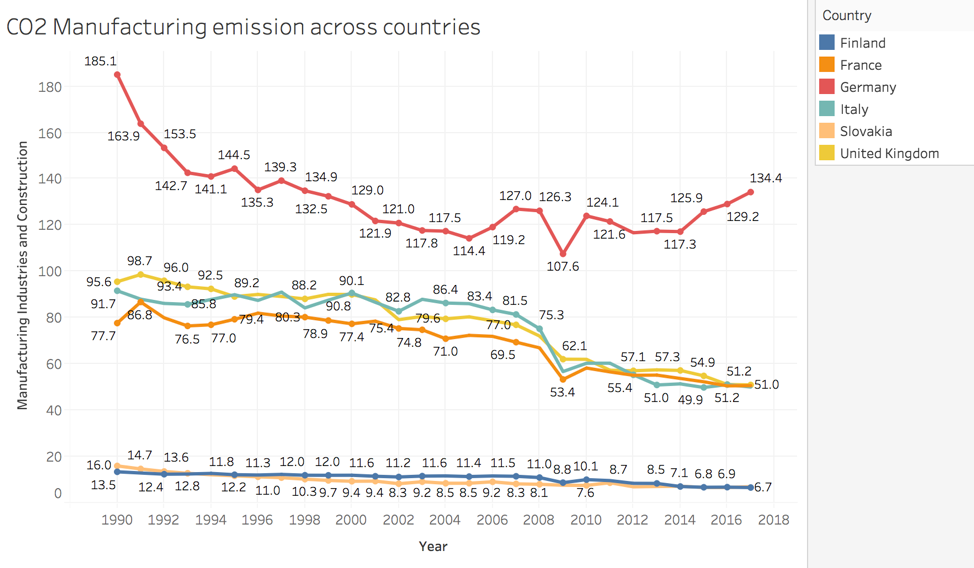


Figure 8: CO2 emission trend in manufacturing industries

This figure shows that Finland and Slovakia have similar trend in controlling CO2 emissions. France, UK and Italy seem to be reducing their emissions yearly while Germany’s CO2 emissions went up after 2014.

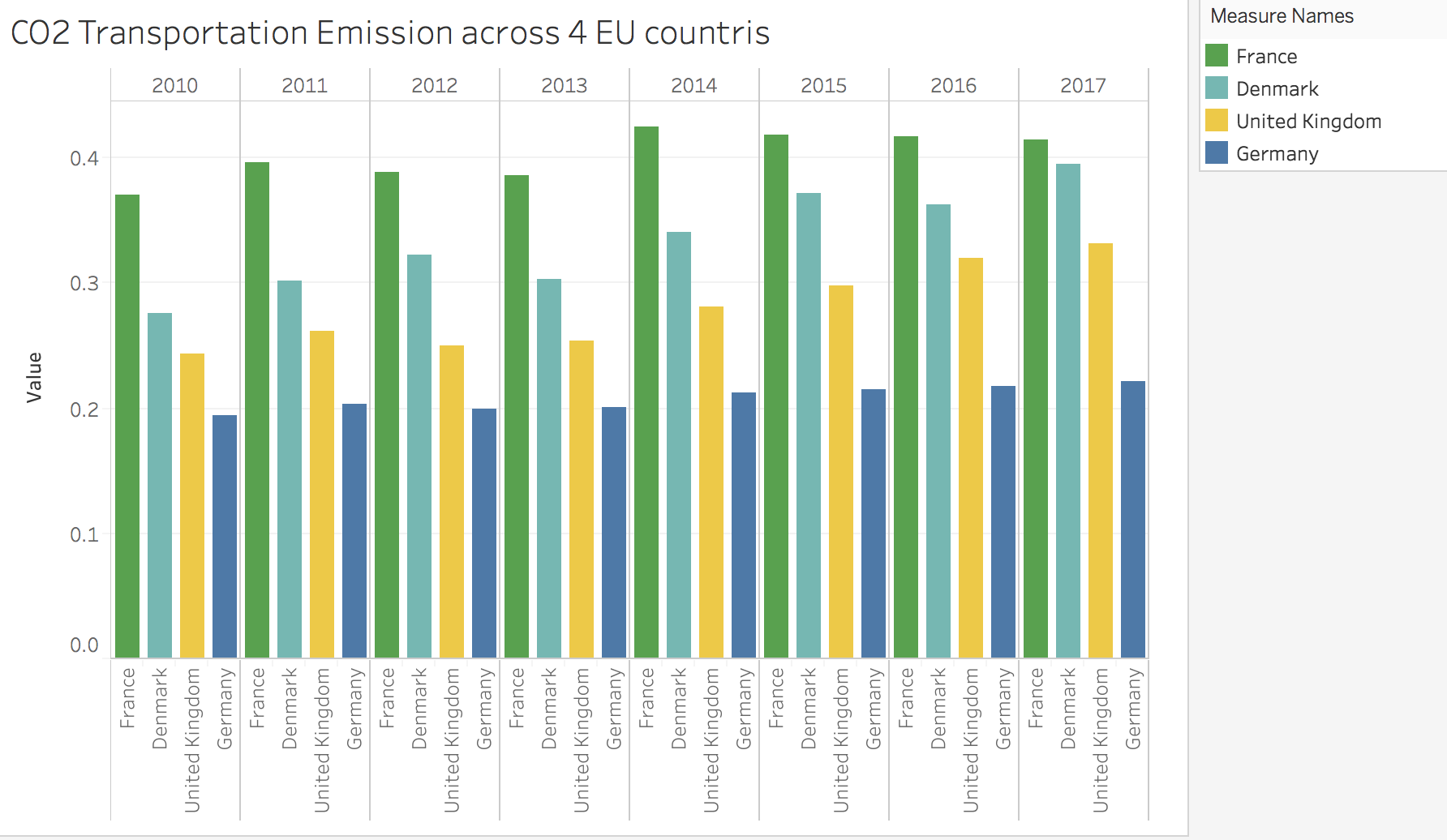


Figure 9 illustrates that FR has the largest CO2 transportation emission in recent years, next is Denmark, UK and Germany. From 2010-2017, it looks like CO2 transportation emission in Germany stays at the same level while other countries increase.

### **2.3 Statistical Methods**

From the visualization, some trends in dataset can be analysis with some statistical methods. As the dataset are mostly numeric values and follows a time series change and are not random variables. T test, permutation test, proportion test and mostly linear regression are decided to be the testing methods for the team.

2.3.1 T-test

The two sample T-test is chosen to compare the difference between different countries on the same industries. The similarity of industries can be explored through this method. From the basic observation, the team decides to focus on some small countries’ specific industries such as Finland and Slovakia. The other industry, team to test, is the transport industry between large population countries, such as France and Britain. The results of the method will be discussed in the results section.

The team picks Finland and Slovakia because these two countries have similar population at about 5.5 million with similar longitude and different latitude. By looking at the summary of the difference between these two countries, they have small difference. The following hypothesis was made:

H0: There is no difference in Manufacturing Industries and Construction CO2 emission between Finland and Slovakia.

Ha: There is a difference in Manufacturing Industries and Construction CO2 emission between Finland and Slovakia.

However, it is very unreasonable to compare the absolute value of consumption between countries. Because the countries are so different. Hence, the team decided to compare the countries’ industries with its proportion of the whole consumption and compare it with the EU as a whole union’s consumption. It may show that whether those large countries decide the trend of the whole union or all the countries contribute their parts.

The one sample t-test is focused on the transport industry proportion of the countries. Because, from the basic visualization, the team observed that the percentage of the transport in large countries are very stable and no large fluctuation happens. So the team wants to explore if the proportion of a large developed country is somehow static to the overall GHG consumption use. France and Britain are chosen to conduct the test and test results are in the results section.

2.3.2 Permutation Test

The permutation test is also made to test for the mean difference between the CO2 emission in Manufacturing Industries and Construction. The data of the CO2 emission of the two countries was combined and resampled for the permutation test. The team want to use the permutation test result to compare with the result of t test.

2.3.3 Linear Regression

The dataset contains data from 1990 to 2017, year by year. Also from the basic visualization, there are trends in some industries and overall consumption being observed. For example, the per capital consumption of GHG looks likely follow a linear trend and decreasing. However, the international aviation is observed to be linearly increasing. The linear regression model is conducted to those values and industries and values to observe the relationship and predict the GHG consumption for future.

# 3. Results

**3.1 T-test Results**

Two sample T-test results

Table3.1 T-test of FR and UK Transport



Table3.1 shows the result of the t-test for the difference between FR and UK’s Transport CO2. The t value is -7.26 and p-value is close to 0. So it is very unlikely that FR and UK have the same volume of CO2 consumption. From the summaries and visualizations, all large countries in the EU have very different consumption on major industries. The transport is the only one that seemingly close. However, from the test, it shows that from the absolute value, it is not close in statistical concern.

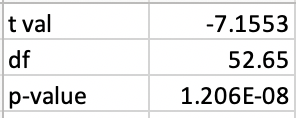
Table3.2 T-test of FR and EU Transport Percentage



Table3.3 T-test of UK and EU Transport Percentage



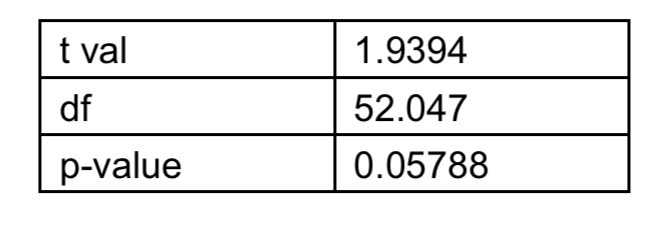
Table3.4 T-test of DE and EU Transport Percentage



From table3.2 to 3.4 are the two-way t-test results between the highest 3 GDP countries and the whole EU. Only the UK has similar proportion to EU. It is unlikely that the EU follows only the largest countries.

Two sample T-test for FI and SK:

Table3.5 T-test of FI and SK Manufacturing Industry and Constructions



The above graph is the two-sample test of difference between the mean of the CO2 emission of the Manufacturing Industry and Constructions between Finland and Slovakia. The p-value is 0.05788, which is larger than 0.05. It also tells that 0 falls into the 95 percent confidence interval [-0.04119663, 2.41970267]. Therefore, we do not reject the null hypothesis. We conclude that there is no difference between the mean of the CO2 emission of the Manufacturing Industryand Constructions in these two countries.

One sample t-test Confidence Interval

Germany Transport Percentage has a 95% CI from (0.1948478, 0.2049552)

France Transport Percentage has a 95% CI from (0.3586600, 0.3792641)

UK Transport Percentage has a 95% CI from (0.2335988, 0.2570855)

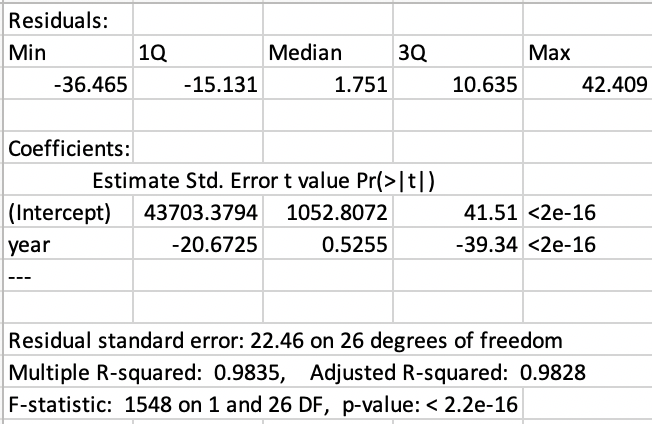
From the test results, the team concludes that the transport has a very stable percentage in a country total CO2 consumption.

**3.2 Permutation Test Results**

The p-value for the permutation tests for the mean is 0.056, we do not reject the null hypothesis at 0.05 significance level. We conclude that there is no difference between the mean of the CO2 emission of the Manufacturing Industry and Constructions in these two countries. Base on the two-sample result and the permutation test result, it is possible that the industrial scale of these two countries might be similar.

**3.3 Linear Regression**

Table3.6 Summary of Linear Regression Model EU



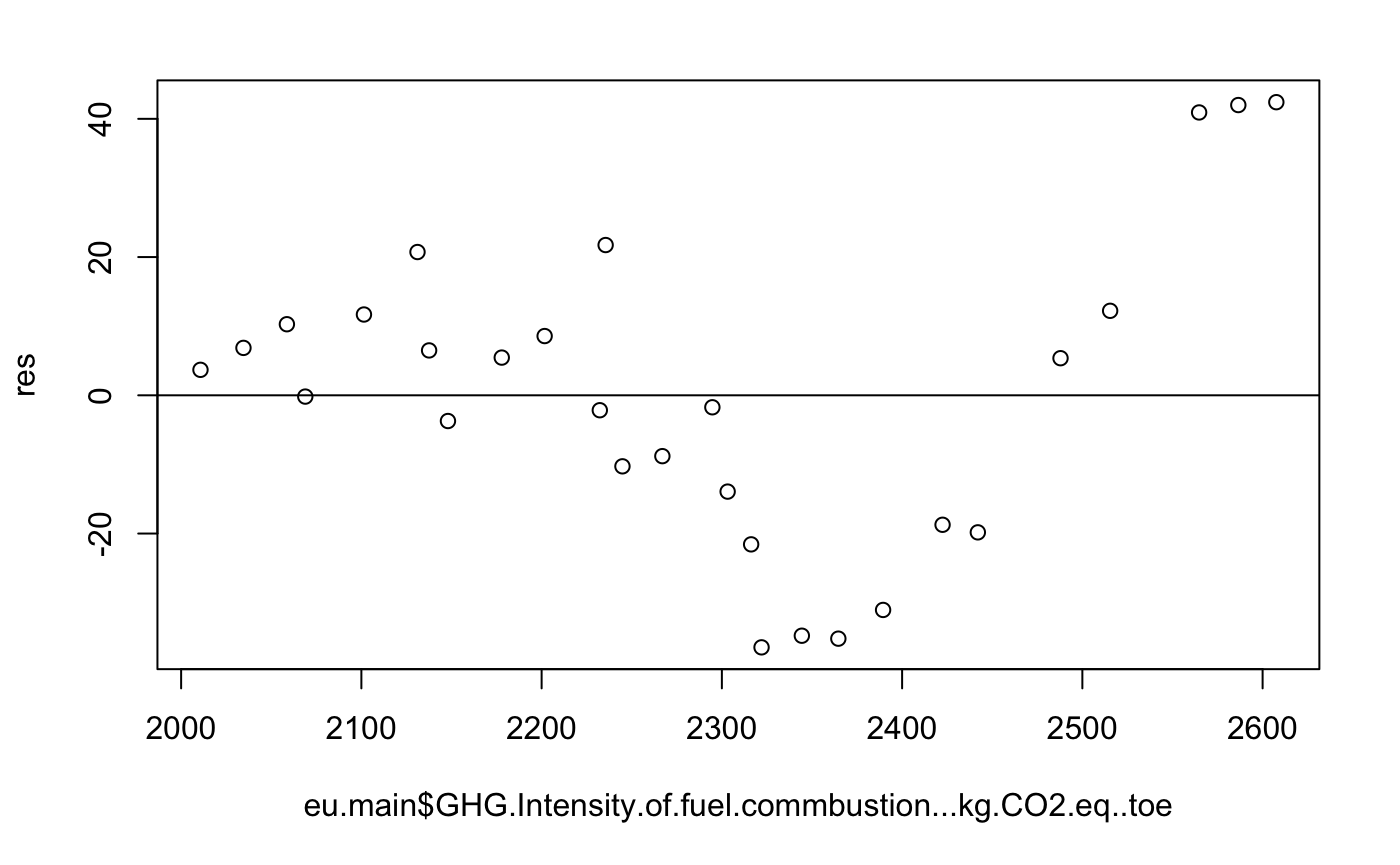
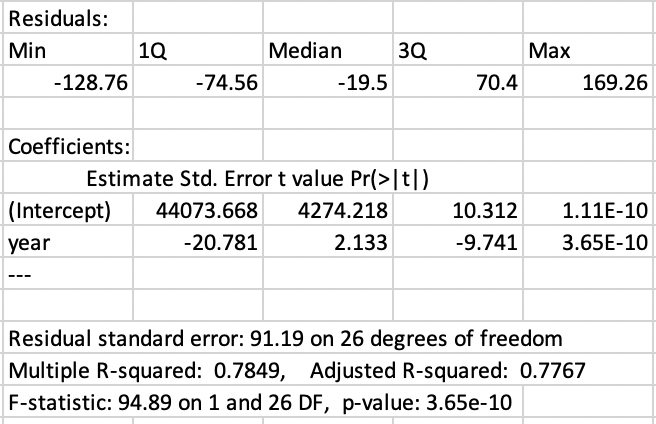


Fig3.1 Residual Plot EU Consumption Model

Table 3.6 shows the summary of the linear regression model of EU GHG intensity and year. R2 is 0.98. It shows a very good fit of the model. The coefficients are not very important in this model. As the year of start can be defined differently. In some studies, the first year or some other year can be taken as a standard year or starts from 1.

Fig3.1 is the residual plot of the EU model. There are only 28 points in the dataset. So it is hard to clarify whether a linear line model is enough or appropriate.

Table3.7 Summary of Linear Regression Model UK



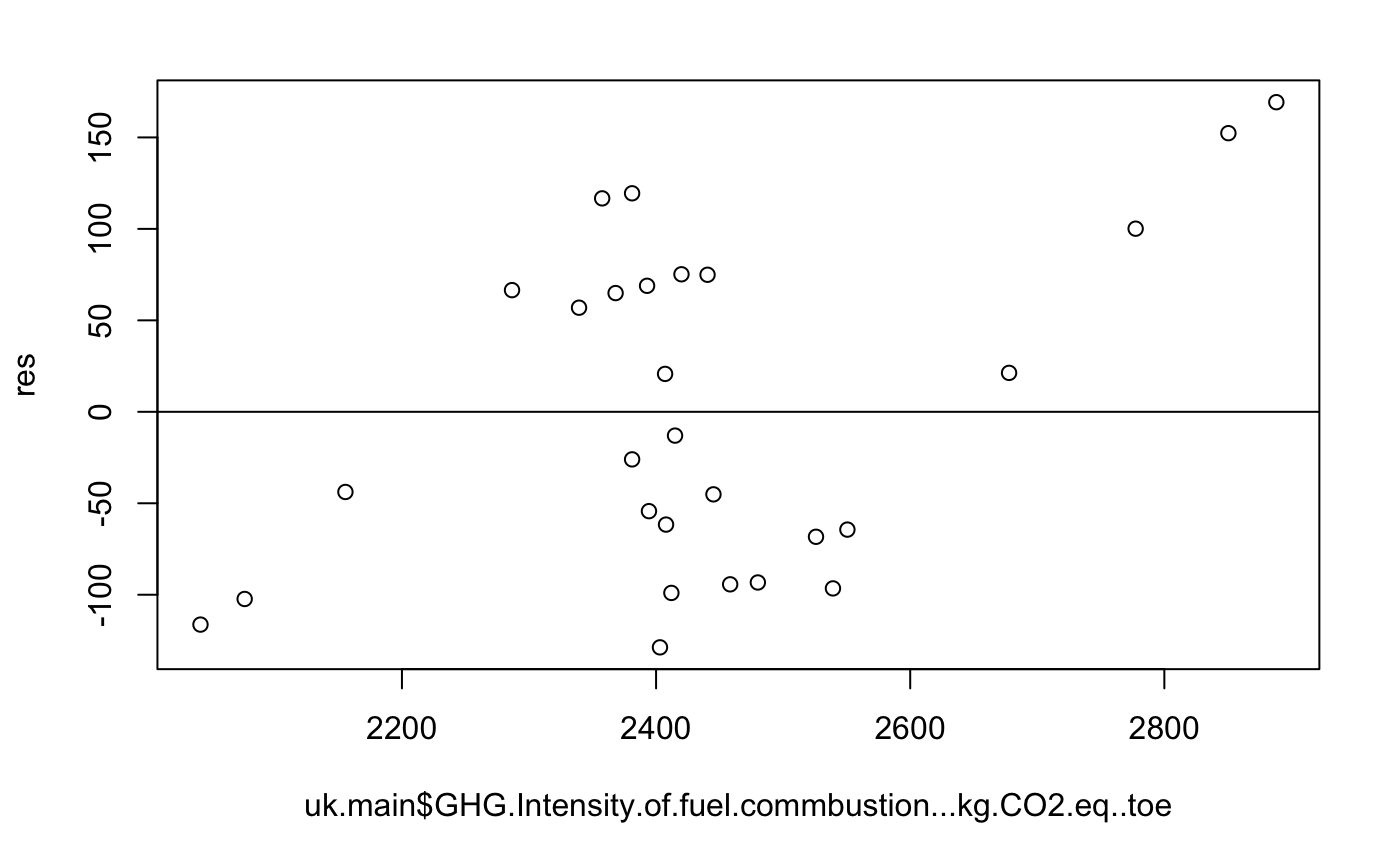
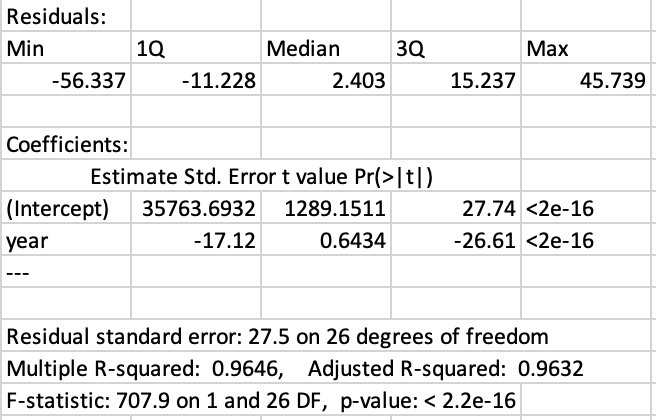


Fig3.2 Residual Plot UK Consumption Model

Table 3.7 shows the summary of the linear regression model of UK GHG intensity and year. R2 is 0.78. It shows the fitness of the model. The coefficients are not very important in this model. As the year of start can be defined differently. In some studies, the first year or some other year can be taken as a standard year or starts from 1.

Fig3.1 is the residual plot of the UK model. There are only 28 points in the dataset. So it is hard to clarify whether a linear line model is enough or appropriate.

Table3.8 Summary of Linear Regression Model FR



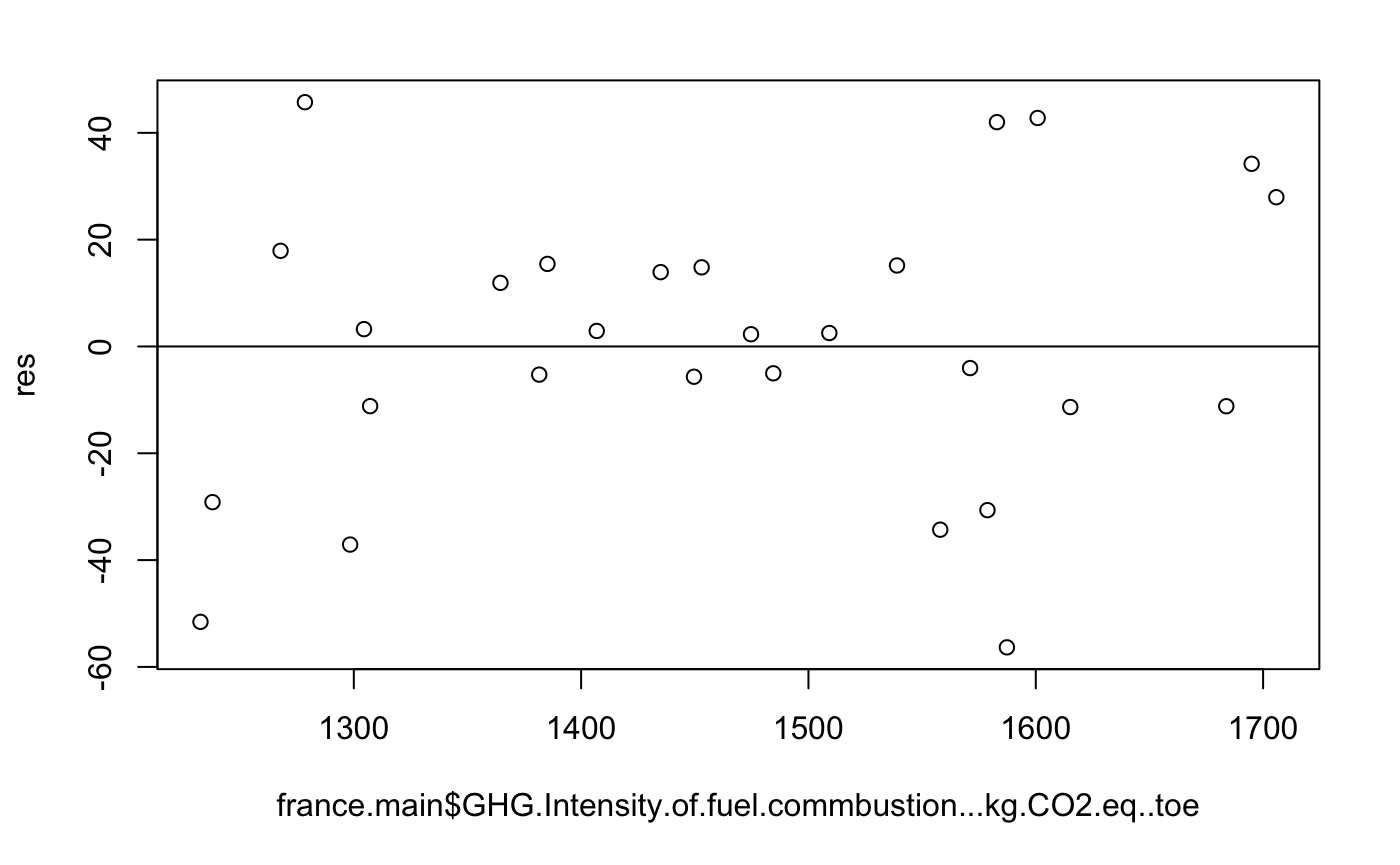
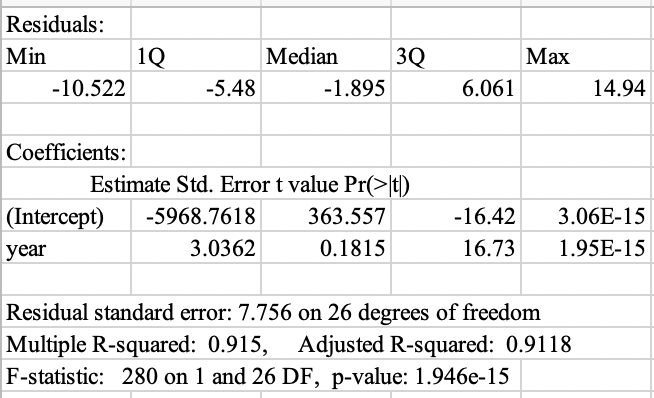


Fig3.2 Residual Plot FR Consumption Model

Table 3.8 shows the summary of the linear regression model of UK GHG intensity and year. R2 is 0.96. It shows the fitness of the model, better than UK and little lower than EU. The coefficients are not very important in this model. As the year of start can be defined differently. In some studies, the first year or some other year can be taken as a standard year or starts from 1.

Fig3.1 is the residual plot of the FR model. There are only 28 points in the dataset. So it is hard to clarify whether a linear line model is enough or appropriate. But this one looks like no trend.

Table3.9 Summary of Linear Regression Model International Aviation EU



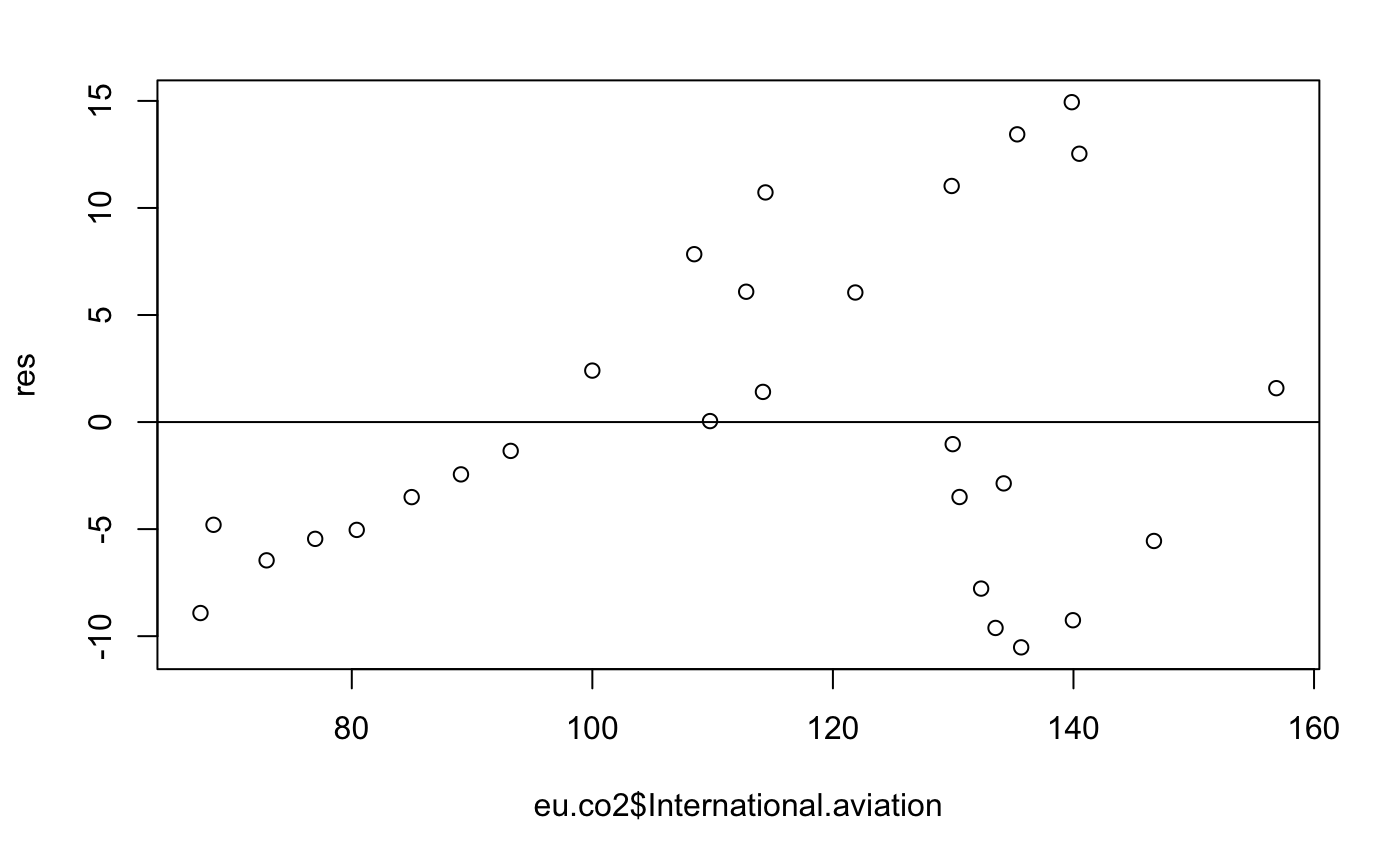
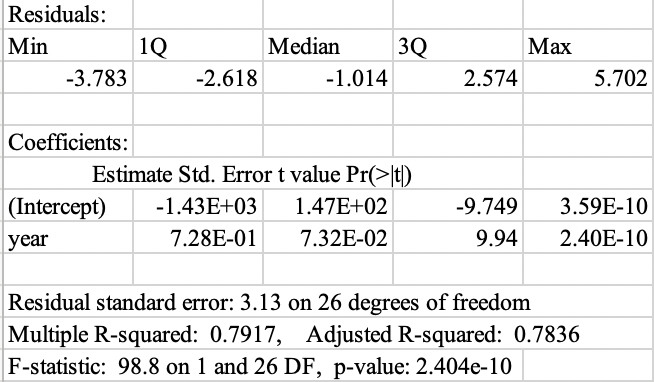


Fig3.4 Residual Plot EU International Aviation Model

Table 3.9 shows the summary of the linear regression model of EU International Aviation and year. R2 is 0.91. It shows a very good fit of the model. The coefficient of year is positive. It implies that the CO2 consumption increases with the time which is different to the total trend.

Fig3.4 is the residual plot of the EU International Aviation model. There are only 28 points in the dataset. So it is hard to clarify whether a linear line model is enough or appropriate.

Table3.10 Summary of Linear Regression Model International Aviation UK



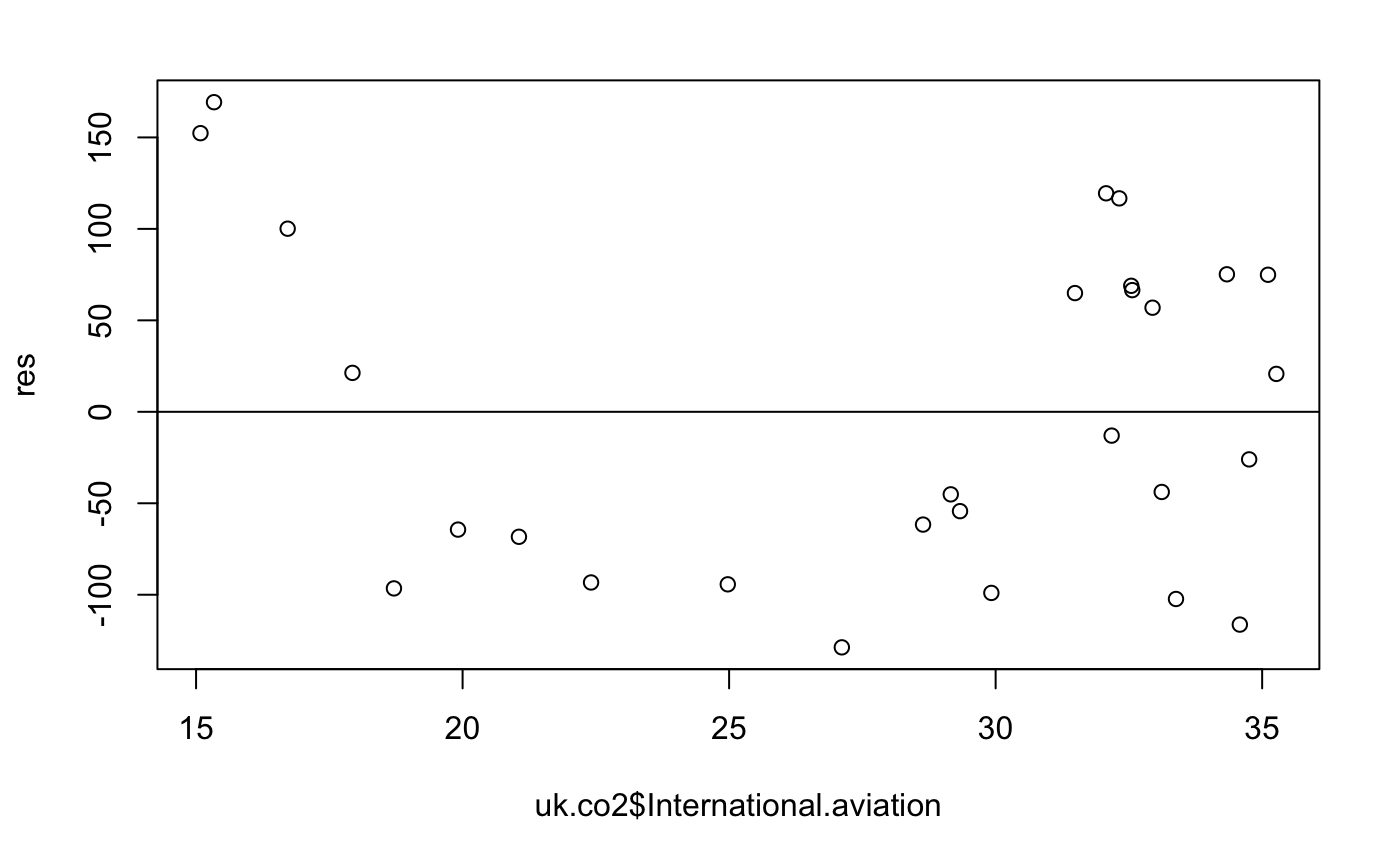
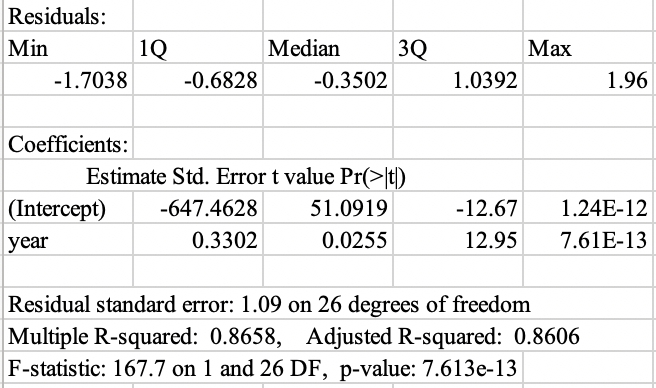


Fig3.5 Residual Plot UK International Aviation Model

Table 3.10 shows the summary of the linear regression model of UK International Aviation and year. R2 is 0.79. It shows the fitness of the model. The coefficient of year is positive. It implies that the CO2 consumption increases with the time which is different to the total trend of UK.

Fig3.4 is the residual plot of the UK International Aviation model. There are only 28 points in the dataset. So it is hard to clarify whether a linear line model is enough or appropriate.

Table3.11 Summary of Linear Regression Model International Aviation FR



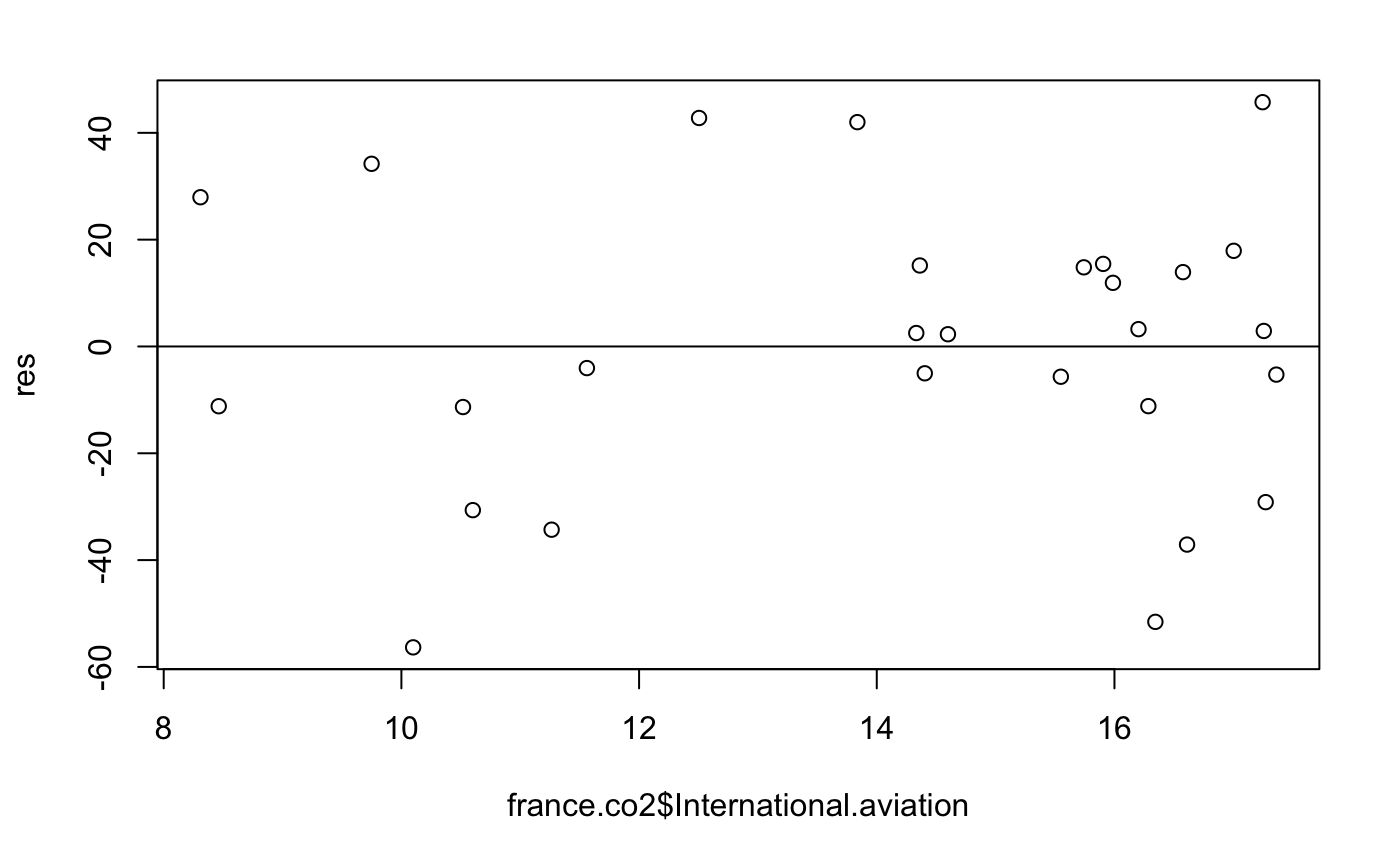


Fig3.6 Residual Plot FR International Aviation Model

Table 3.11 shows the summary of the linear regression model of FR International Aviation Model and year. R2 is 0.86. It shows the fitness of the model. The coefficient of year is positive. It implies that the CO2 consumption increases with the time which is different to the total trend of UK.

Fig3.6 is the residual plot of the FR International Aviation model. There are only 28 points in the dataset. So it is hard to clarify whether a linear line model is enough or appropriate.

# 4.Conclusion

The dataset has a lot of information that can be worked. However, due to the time limitation, only the obvious trends and features are looked into. There are several things that are discovered in project.

Overall, the Europe Union is doing a very good job on reducing CO2 and greenhouse consumption in general. Though there are some fluctuations in some small industries and countries, the total amount of the consumption is definitely reducing. For example, in the project, team did a lot with transport industry and transport industry maintains a stable percentage in those large countries. However, there are still few industries have a different trend such as international aviation.

After 1990, people pay more attention on environmental issues. Europe union, a union of most developed countries, has made impressive progress on reducing those consumption from our result. From the result’s section, the reduction of emission follows a linear curve and has a good fitness; predicting the emission on 2050, the reduction will be half from the 1990. However, it may still be not enough. For the specific industries, the main industries, though with fluctuations, are as the same trend as the total consumption. However, those industries are different across countries. It is possible that because as a union, countries can develop different aspects in an industry. For example, energy industry, France is famous for its nuclear energy and Germany is the solid fuels and United Kingdom is the renewable energy especially wind resource. So do other industries have this situation. A lot of further studies can be done with more data and other source information to provide more insights on those issues.

However, there are still some industries’ consumption are similar among small size countries. Especially, when those industries are mainly served only for their own citizens. So, the team concludes that it is the cooperation between countries making the industries among large countries different. That also explains why international aviation and international maritime transport are increasing. Right now, those two are still the main methods of transport between countries. Although there is slowing trend contemporary, those consumptions are still expected to be increasing.

Europe union, as a union of developed countries, actually provided a possible road for other countries. There is not a lot of super countries such as United States and China. More small countries should try to develop their own advanced industries and trade with others. At least this way of reducing GHG consumption is proved by the Europe Union.

Future work on the environmental issues, team considers that collecting data more specifically and narrow down the topics. It is possible to provide specific suggestions on some industries.

# Reference:

1. Global Warming Effects.” *Global Warming and Climate Change Effects: Information and Facts*, 4 Feb. 2019, <https://www.nationalgeographic.com/environment/global-warming/global-warming-effects/>.
2. “The Causes of Climate Change.” *NASA*, NASA, 30 Sept. 2019, <https://climate.nasa.gov/causes/>.

# Appendix

The R rmd file is submitted together. See the attached rmd file for details.

### 