**Research Report**

Greenhouse Gas Emissions

Suhainsly Martha, student number 2602068

Tomas van Daalen, student number 2598668

Rob Schilder, student number 2572418  
  
Digitization: from Object to Data  
Vrije Universiteit Amsterdam

October 2018

**Introduction**

In 2018, it is impossible to ignore the reality of climate change; from droughts in Syria, to national weather changes. Even universities such as the Vrije University give people the option to study climate change with courses available in the Humanities. In political news, the European Union (EU) member states have set goals in various agreements to limit the increase of global temperatures. Countries globally agreed under the United Nations Framework Convention on Climate Change (UNFCCC) to limit 2°C compared to the average temperature in pre-industrial times to prevent the most severe impacts of climate change and possibly catastrophic changes in the global environment. To achieve this, the world must stop the growth in greenhouse gas emissions by 2020, and reduce it by 60% by 2050 compared with 2010. The new goals have been set in the Paris agreements in 2015. These were to decrease global average temperature to well below 2 °C above pre-industrial levels and to limit the increase to 1.5 °C, since this would substantially reduce the risks and effects of climate change. The rest of this paper will focus on this issue.

**The research**

This research is based on the dataset provided by the UNFCCC and the EU Greenhouse Gas Monitoring Mechanism. This dataset is a collection of the amounts of emissions of greenhouse gasses (GHG) per sector per country per year over the period 1990-2016. With the dataset, it is possible to look at which industry has the biggest growth in the emission of greenhouse gasses. The European Environment Agency (EEA) website where the dataset can be found, provides several graphs that displays which sector has the biggest growth in the emission of greenhouse gasses (figure 1). What becomes clear from this graph, is the fact that fuel combustion in the transportation sector makes it so that the industry has had the biggest growth in emissions in the period 1990-2016. However, the transportation sector is a big industry and it is not clear which specific transportation method has contribute more to the growth.

The goal of this research is to find out which transportation method in the transportation sector has the biggest percentage growth in the emissions of greenhouse gasses in the period 1990-2016. After finding out which transportation method has the biggest growth, further research will be done to explain this growth. The main question of this research is the following:

*Which transportation method had the biggest percentage growth in the emissions of greenhouse gasses between 1990-2016 and why?*

To find the answer to this question, the first objective is to find the transportation method that has the biggest percentage growth in the emission of greenhouse gasses. This will be done with the following sub-question:

*Which transportation method had the biggest percentage growth in the emissions of greenhouse gasses based on UNFCCC dataset in the period 1990-2016?*

This question will be answered by an analysis of the dataset conducted with programs like OpenRefine and Microsoft Excel. OpenRefine will be used to filter out the important information out of the big dataset. After the filtration with OpenRefine, Microsoft Excel will be used to make calculations, graphs and charts with the extracted dataset. The graphs and charts should help with finding out which transportation method has the biggest growth by visualizing the data.

After finding out which transportation method has the biggest growth in emissions, the next step will be to explain why this is. This will be done following our second sub-question:

*What are the reasons for the growth of emissions of greenhouse gasses in the biggest growing polluting industry?*

To answer this last sub-question, we will look at secondary literature about the subject.

**Relevance of the research**

The emission of greenhouse gasses, with global warming as a result, is a subject that affects everybody. The global sea level rising and the complete disappearance of vegetation are just some reasons as to why global warming should concern everybody. This research is focussed on the industry that has the biggest increase in emission of greenhouse gasses in the period of 1990-2016. The information that will be retrieved from this research can be used to tackle the problem of the emissions of greenhouse gasses and affect actual policy discussions. We will also look for explanations for the increase in the emission of greenhouse gasses. Because of this, it is possible to look at the problem from a new perspective, and with that, try to come up with new solutions to tackle a small part of the problem of the pollution of the earth.

**Data analysis**

The dataset provided by the UNFCCC is massive and is structured and given in csv format, although the initial download is of a zip file on the EEA website. The dataset and subsequently the data is copyrighted by this agency. However it is open, linked and available for the public to view as long as the user complies with the EEA standard reuse policy.[[1]](#footnote-0) Further metadata is found on the Eionet website which is a governmental open big data source. With Eionet EEA is trying to promote and further environmental research and policies via big data. This data is gathered through continuous monitoring and reporting. On this website the viewer can find an extensive overview of the metadata. Metadata is important for the interpretation and searching of particular data especially if the data comes from various sources (Janssen, Charalabidis & Zuiderwijk, 2012).

Diving into the dataset itself, there are several things that are apparent. First of all the level of standardization. For this the Intergovernmental Panel on Climate Change (IPCC) Common Reporting Format has been used. The data collected for this dataset has strict regulations for every country. The regulations are available on Eionet in the form of EU policy papers at which in the end there are approximately 50 tables that should be electronically submitted by every country annually (see Table 1). The tables represent the specific sectors and parameters set. This reporting process is also a massive undertaking consisting of several teams for every country collecting the data, curating the data and submitting the data. However, both the extensiveness of the metadata and the openness of the data mentioned previously have their limitations. In the case of this particular dataset, the original submitted reports with tables of the EU countries are not available for the public. This leads us to the issues of data collection and provenance. Because the original data is not available for the public there is no say into how the data was changed to fit the final dataset. Moreover, the country’s collection processes are not made public.

The dataset also brings up questions of quality. The data set is annotated, where there no values for the emission, there are annotations “C”, “IE”, “NA”, “NO” and a few more. What these notations mean are not immediately clear in the dataset. However, in the IPCC reporting guidelines the notations are defined. For example, for the notation “NA”, not estimated, the countries have to report why the estimation has not been made. This is however not made clear to the public because of the provenance issue. Perhaps the countries have made it clear in the required tables but this information did not make it back to the data set. So there is dumbing down of information once you look deeper. Given that there is no original data available, we researched further into the obligations that the countries’ have to stick to while collecting and writing down the data. One of these obligations is to not only make clear in the given tables that there is a level of uncertainty with values, but they have to make this clear in the form of completely different tables from the required 50 (see Table 2). It is clear, that not all data gathered can be said to be 100% conclusive. However, this shred of uncertainty is not made clear anywhere in the dataset. We can assume that the given information per country is correct in the dataset used, what the regular public can not assume is that the data reported is completely factual. So the data is public in name only but is still private in practice, which makes curating the dataset nearly impossible (Janssen, Charalabidis & Zuiderwijk, 2012). Nonetheless, curation is not necessary for this research. How the dataset will be manipulated is discussed in the following paragraph.

**Methodology**

The data in the dataset is complete enough to fully answer the research question. There are hardly any gaps in the dataset, which meant that no steps of enriching the data were necessary before analyzing the dataset. Since the dataset was in the .csv format, there were no problems with implementing the file in OpenRefine. The only problem with the dataset was the size if the file. With a size of 86,2MB, most laptops got stuck on processing the large data in OpenRefine. However, this problem was solved by simply using a computer with a better processor. In order to get the wanted results from the dataset, some data needed to be extracted. A short description of the methodology that is used to extract the wanted data and making the graphs will follow. A more extended description of the methodology can be found in the documentation of the dataset file.

In order to make graphs, data of four different years are needed, which are 1990, 2000, 2010 and 2016. These four years were chosen because this seemed like a nice division of the period 1990-2016. These four years with their corresponding data, can be filtered out of the dataset by making a text-facet of the year column. By doing this, it is possible to include the wanted years and save the filtered dataset as a new file. An additional benefit of making it into a new smaller file, is that modifying with OpenRefine becomes a lot faster.

For the four years, we needed information about the five transportation methods with their corresponding amount of greenhouse gasses. Same as for the years, this is done with text facets. After this filtration, the extracted dataset was almost ready to be used to make graphs. The only thing that needed to be done, was deleting the unnecessary columns. We deleted all the columns, except for year, amount of greenhouse gasses, country and transportation method.

With the extracted dataset, it became possible to make graphs and charts of the data. We made a pie chart for every year in the dataset, in order to display the share of each transportation method in the total transportation sector. The pie charts are made by adding up all the emission values of each country per transportation method in Excel. With the total values of each transportation method, it became possible to make pie charts and graphs of the shares of each transportation method in the total transportation sector. These charts and graphs can be used to find out which transportation method has the biggest share every year and for this research specifically, it can help with the question which transportation method has the biggest percentage growth over the period 1990-2016.

**Findings and interpretations**

To get a better insight in to the growth, we first look at the total emissions of each transport sector in 1990, 2000, 2010 and 2016. Using the EEA (2018) database, we calculated that the greenhouse gas emissions originating from the transport sector were 76.984.408.453 Gg CO2-eq in 1990. Ultimately, the emissions increased by approximately 29 percent to 99.466.482.644 Gg CO2-eq in 2015. An increase of 29% may not seem like a lot, but in the context of climate change it is very much significant. The new dataset made several things clear, the transportation methods for instance. The ones the EEA checks for are *aviation*, *road transportation*, *railway*, *navigation* and the rest of the gasses are contributed by what are categorized as *others*. Based on this dataset several chieparts were made which play a big part in the coming conclusions (see Graphs 1 to 4).

From the pie charts it is clear that the the transportation method *road transportation* is the bigger polluter. This is not a surprise, since there are more vehicles than airplanes, trains etc. throughout the whole of the European Union, around 290 million vehicles according to the European Automobile Manufacturers Association[[2]](#footnote-1). This including passenger cars, light and heavy commercial vehicles, busses and total commercial vehicles (think of cargo transport trucks). However growth wise, this industry has remained steady from 1990 to 2010, around 92% of the total GHG in the transportation sector (see Graph 1 to 3). In 2016 it had a slight increase of 2%, making it around 94% of the GHG emissions. However this 2% increase is not sufficient to explain the 29% growth in the whole of the sector. So despite the growth in the transportation sector, the astronomical amount of road transportation gasses does not play a significant a role, which is important when talking about reduction of GHG in this sector.

Furthermore, there are several other industries in this sector that aren’t steady as the road transportation industry but are actually decreasing. The railway industry has been steadily dropping in their GHG emissions since 1990. It went from being around 1,9% of the GHG emissions in the transportation sector to being around 0,7% (see Graph 1 and 4). The same counts for navigation industry that decreased from around 4,1% to 2,2% from 1990 to 2016 (see Graph 1 and 4). So it seems that none of these industries really are the biggest cause for the 29% growth in GHG emissions in the transportation sector.

This leaves out the aviation industry. This industry contributed around 1,3% of the GHG emissions in 1990, however it almost tripled the amount of emission by 2000 (see Graph 1 and 2). By 2000 this industry contributed 3% of the GHG, which again compared to road transportation that contributed 92% that year is not a lot (see Graph 2). However this *growth* is significant when looking at the total increase of 29% from 1990 to 2016. By 2016 this industry did decrease 0.4% but has never reached the amount of GHG emissions that it had in 1990 (see Graph 5). Additionally, according to EUROCONTROL the skytraffic reached an all-time record of 10,190,903 flights with a daily average traffic of 27,844 flights in 2016, surpassing the former record set in 2008. So the previously mentioned decrease of 0.4% could actually not be relevant anymore and the next set of data gathered from UNFCCC could again show a major increase in GHG emissions from this industry. Which will definitely lead to another increase in the transportation sector. The big increase of the emissions of greenhouse gasses in the aviation industry is a direct result of three factors according to Grassl (2003). First of all the low pricing of flights, the wide availability of flights and finally the large demand for flights. These three factors are closely connected, because low pricing causes a large demand, a large demand causes more availability and more availability causes lower prices. It is clear that the factors are very interdependent on each other, which might be a clear indication why the aviation industry is growing this hard.

**Conclusion**

In the future, the economic growth has to be more sustainable if we want to continue to live in ease on this planet. For that we have to lower the amount of greenhouse gasses with a rapid speed. A lot of measures have been agreed on by international conferences, yet from our research we can conclude that the aviation industry (and for a part the road transportation industry) is still growing in its pollution.

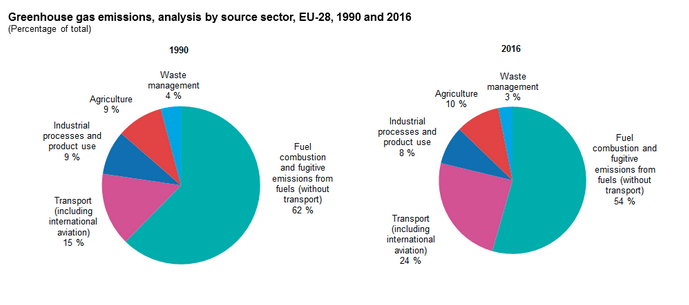
This research looked at which transportation method had the biggest growth in the emissions of greenhouse gasses in the period of 1990-2016. This research is conducted with a dataset that is provided by the United Nations Framework Convention on Climate Change (UNFCCC). From this we see that all industries in the transport sector are growing, with the aviation industry growing the fastest. Taking a closer look at the dataset gives us the conclusion that there has been a significant growth from 1990 to 2016. Although the growth has come to a stop around 2006, the growth does continue again in 2016. Hence, we can conclude that the problems of the growth of the emissions in the transport sector will mainly come from the aviation industry.

The importance of such a research is glaring, as it shows the issue in Europe’s greenhouse gas emissions. It also makes it clear that maybe the policies implemented aren’t working as they should. Especially in a time where flying starts to become more popular, measures should be taken to find alternative ways of travelling to avoid the growth in the emissions. This research forms a basis for future research towards alternative, non-polluting ways of transport.

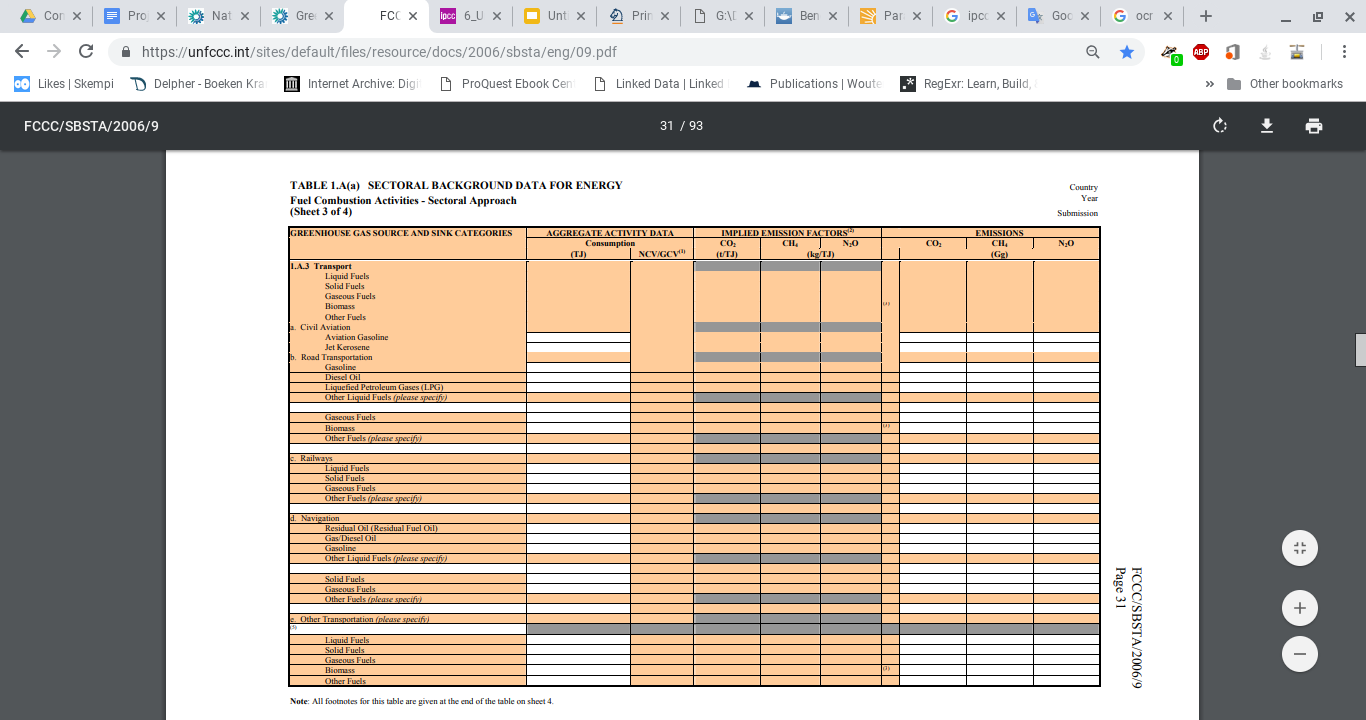
**Possibilities for future research**

The result of the research of the dataset made clear that aviation has the biggest growth in the emissions of greenhouse gasses in the period of 1990-2016. This research only focussed on why there was such a big growth in the emissions of greenhouse gasses in the aviation industry. For further research, it might be interesting to come up with solutions for this problem. These solutions can be based on this research, but it is also possible to come up with another kind of research. It might be possible that the pollution of the aviation industry is vastly reduced in the future, because of improved planes or a more sustainable type of fuel. In this case, it might be possible to do further research on new threats for the environment with the same method as this research. This will probably give new results, which can be investigated. Finally, as made clear by the data analysis, it would be prudent to research the issues of provenance and data collection of these submitted values of greenhouse gas emissions.

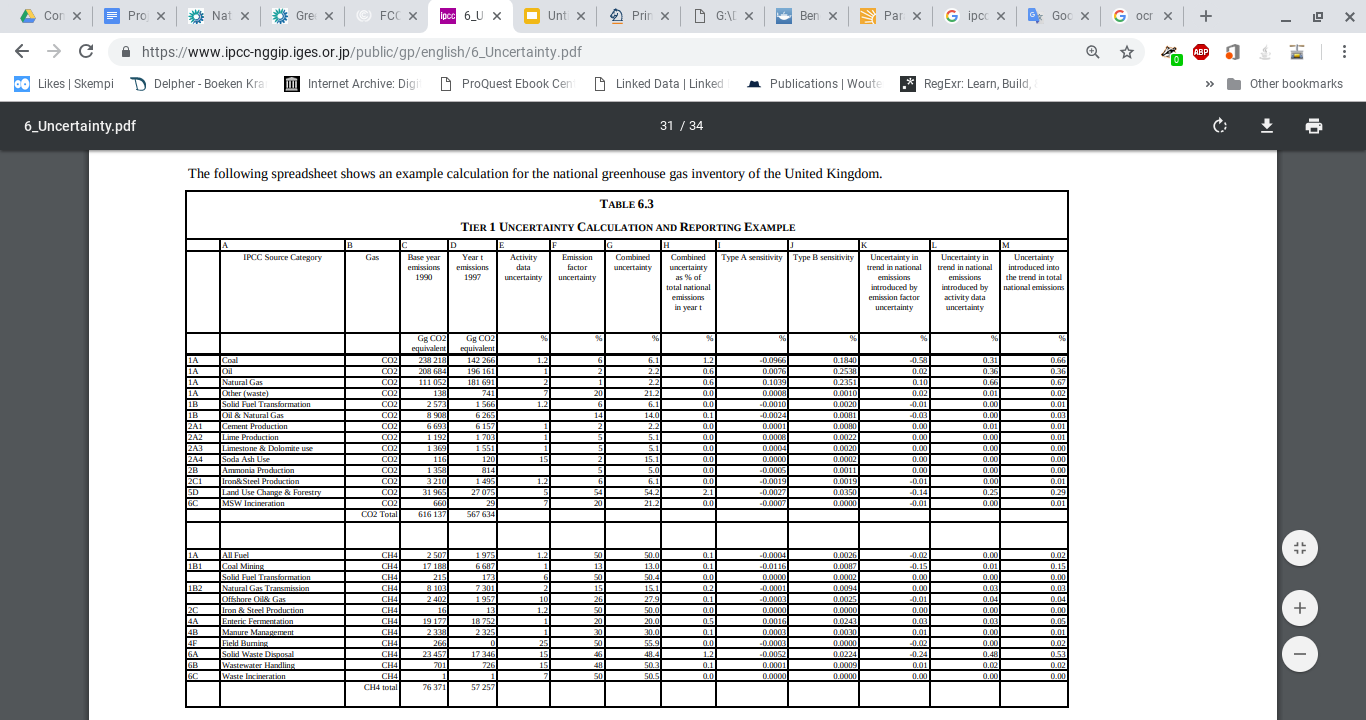
**Appendix**

**Figure 1: Greenhouse gas emissions, analysis by source sector, EU-28, 1990 and 2016**

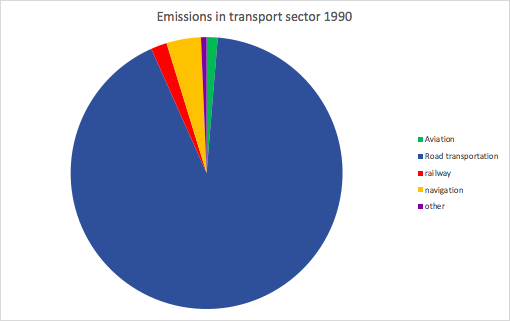
**Table 1: Example Emission Reporting Table**

****

**Table 2: Example IPCC Uncertainty Reporting Table**

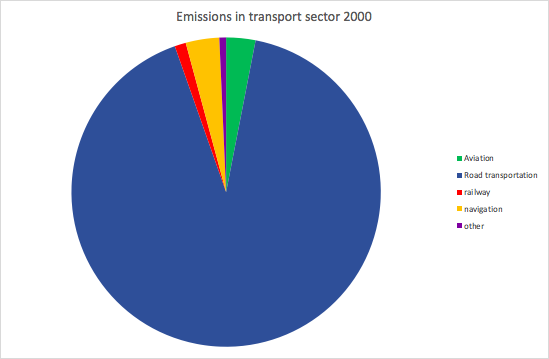
****

**Graph 1: Emissions in transport sector 1990**

****

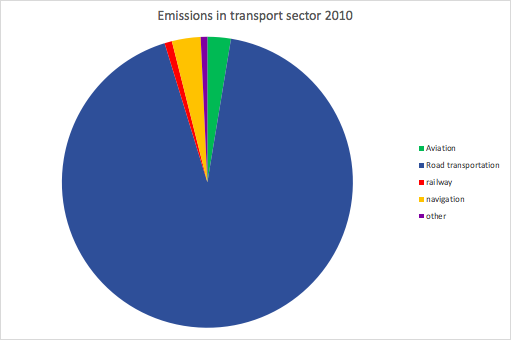
|  |  |
| --- | --- |
| Total | 76.984.408.453 |
| Aviation | 1.012.948.785 |
| Road transportation | 70.840.624.648 |
| railway | 1.461.711.039 |
| navigation | 3.157.065.060 |
| other | 512.058.921 |

**Graph 2: Emissions in transport sector 2000**

****

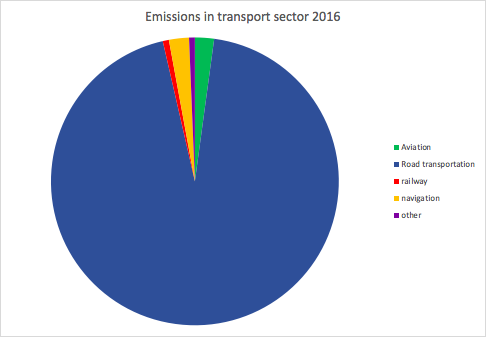
|  |  |
| --- | --- |
| Totalen: | 79.190.778.590 |
| Aviation | 2423325375 |
| Road transportation | 72.501.249.650 |
| railway | 932.612.925 |
| navigation | 2.774.395.214 |
| other | 559.195.426 |

**Graph 3: Emissions in transport sector 2010**

****

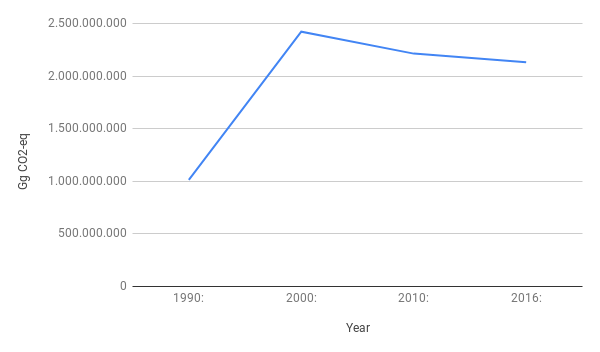
|  |  |
| --- | --- |
| Total | 85.161.429.750 |
| Aviation | 2.214.251.441 |
| Road transportation | 78.896.603.917 |
| railway | 704.321.107 |
| navigation | 2.717.640.696 |
| other | 628.612.589 |

**Graph 4: Emissions in transport sector 2016**

****

|  |  |
| --- | --- |
| Total | 99.466.482.644 |
| Aviation | 2.132.318.101 |
| Road transportation | 93.763.671.775 |
| railway | 687.535.860 |
| navigation | 2.237.826.237 |
| other | 645.489.640 |

**Graph 5: Emissions of greenhouse gasses in the aviation industry**

****

**Bibliography**

Aero, A. ‘79 new airlines begin life in 2017; 25 go out of business – Europe has the highest number of start-ups, but also the most failures’ [Airline Network News and Analysis]. (18 october 2017). Retrieved from: <<https://www.anna.aero/2017/10/18/78-new-airlines-begin-life-2017-25-go-out-business/>>

Grassl, H. (2003). Aviation and Atmosphere In: Meyers R.A. (eds) Encyclopedia of Sustainability Science and Technology. Springer, New York, NY. <<https://link-springer-com.vu-nl.idm.oclc.org/referenceworkentry/10.1007/978-1-4419-0851-3_555>>

Janssen, M., Charalabidis, Y. & Zuiderwijk, A. (2012) Benefits, Adoption Barriers and Myths of Open Data and Open Government. *Information Systems Management, 29*(4), 258-268. DOI: 10.1080/10580530.2012.716740

De Koning, B. (15 july 2017). Vliegen is veel te goedkoop en dat is onhoudbaar. *De Volkskrant,* Retrieved from: <<https://www.volkskrant.nl/columns-opinie/-vliegen-is-veel-te-goedkoop-en-dat-is-onhoudbaar-~b17d4836/>>

Noack, R. (12 october 2017). Why are flights so much cheaper in Europe than in the U.S.?. *The Washington Post*, Retrieved from: <<https://www.washingtonpost.com/news/worldviews/wp/2017/10/12/why-are-flights-so-much-cheaper-in-europe-than-in-the-u-s/?utm_term=.578fd332ddd5>>

1. <https://www.eea.europa.eu/legal/copyright> [↑](#footnote-ref-0)
2. <https://www.acea.be/statistics/article/vehicles-in-use-europe-2017> [↑](#footnote-ref-1)