**Fuel consumption and CO2 emissions from passenger cars: Statistics and Analysis**

**Pravallika Avula**

**G01388664**

**Volgenau School of Engineering,**

**George Mason University.**

**Abstract**

This paper aims at investigating the types of passenger cars and their impacts on pollution. The results are determined based on the car specifications and carbon emissions. Modes of transportation have grown significantly in the last ten years. Cars are the most used vehicles worldwide. Transportation contributes the most to pollution. However, a variety of new car models that can reduce pollution are now available. I chose this UK data set that focuses on car models, fuel efficiency, and emissions. From 2000 to 2013, this study used data analytics, statistical analysis, and visualizations to uncover insights into how and which cars were causing the most pollution. These results were obtained using R, Python, and SQL. According to the findings, petrol is the source of the most pollution. As a result, appropriate steps must be taken to reduce the use of gasoline and find alternatives. In this research paper, we look further into the meaningful insights discovered through the analysis, discuss the study's limitations, and outline the important findings.

Keywords: Fuel efficiency, Vehicles, Carbon emissions, Energy

**Introduction**

Particularly in developing countries where the use of vehicles is growing, carbon emissions are rising. By 2050, there could be 2.5 to 3 billion vehicles on the planet, up from the current 850 million (Global Fuel Economy Initiative (GFEI), 2016). Ninety percent of this growth is anticipated to occur in developing and post-industrial nations (Global Fuel Efficiency initiative -State-of-the-World-report, 2016). The global fleet's greenhouse gas emissions are expected to triple because of these countries' stagnant average vehicle fuel economy. The growth of CO2 emissions in the transportation sector is the highest of any sector, according to the International Energy Agency (IEA) of 2013. Around the world, transportation accounts for more than 50% of oil consumption, and it accounts for almost all the recent and anticipated growth in oil consumption. By 2050, there will be three times as many vehicles on the planet as there are now, and the transport sector is growing more quickly than any other sector globally. Controlling the fuel energy demand and greenhouse gas (GHG) emissions has become a global concern due to the rapid growth in vehicle population, particularly private passenger vehicles. Thus, a major component of the global development agenda is promoting fuel efficiency and lowering carbon dioxide (CO2) emissions. Automobiles also emit methane (CH4), nitrous oxide (N2O), and hydrofluorocarbons from leaking air conditioners in addition to carbon dioxide (CO2). These gases emit fewer emissions than CO2, but because they have a higher global warming potential (GWP) than CO2, they can still have a significant impact. Although, some recent studies suggest that electric vehicles may be the answer to emission issues, Fernández (2018) discovered that the production of electric cars at the power plant level and the production of the electricity used to power their batteries pollute the environment. However, using renewable energy sources to generate electricity, such as wind, solar, and geothermal energy, as well as altering social customs, may reduce transportation emissions.

Since then, the EU has been working to further reduce CO2 emissions and fuel use from passenger cars, with mandatory, fleet-average, and mass-dependent targets of 95 g/km by 2021. A manufacturer can be fined between €5 and €95 per gram of excess CO2 per vehicle sold if they don't follow the rules. Over the past ten years, this policy has significantly altered the average official CO2 emissions and the key characteristics of European passenger cars. This is a direct result of engine downsizing for both diesel and gasoline engines and has been accompanied by a decrease in average engine capacity despite the apparent increase in engine power. Despite its importance for vehicle energy consumption, mass has consistently remained between 1300 and 1400 kg. However, there is debate over the veracity of these CO2 figures and whether they are representative of actual CO2 and fuel consumption. The maintenance of a clean environment has grown in importance in society. Vehicle and motorcycle emissions are some of the pertinent environmental issues that need to be addressed. As a result, reducing motor vehicle pollution is crucial for preventing climate change. It is believed that environmental emissions are proportional to the amount of fuel used by a vehicle. Fuel consumption, which can be defined as the amount of fuel used per unit distance, is the reciprocal of fuel efficiency. The vehicle is said to be more economical the lower the fuel consumption value. Although there is a lot of research on the environmental emissions produced by transportation systems, it is confirmed that variables like driving habits, vehicle design, and traffic conditions have a big impact. When driving in the real world, overlooked factors like side winds, rain, and road grade may significantly affect fuel consumption. A vehicle's fuel efficiency is influenced by several variables, including its design, the surrounding environment, traffic patterns, and user and driver behavior.

In September 2017, the Worldwide Harmonized Light Vehicle Test Procedure (WLTP) took the place of the New European Driving Cycle (NEDC) test procedure for measuring the official fuel consumption and CO2 emissions of new cars. By September 2018, all new cars (powered by an internal combustion engine) were required to comply with the WLTP test procedure. By providing model-specific values at the point of sale, the new testing regime seeks to provide a closer representation of "real-world" fuel consumption and CO2 figures. Even though all tests will still be carried out in a lab, allowing for meaningful comparisons between cars, it differs from the previous procedure in several ways.

To have a better understanding of cars and the pollution they cause, this paper attempts to address key questions of concern. The following are the research questions-

1. *Which car manufacturer produces the most air pollution and noise pollution?*
2. *Which car model produces the most air pollution and noise pollution?*
3. *How are fuel costs and emissions related?*
4. *What are the characteristics of the car with low emissions?*

By finding a solution to these questions, we can raise awareness and discontinue the production of those models. By studying the dataset, one will gain an understanding of the characteristics of the car models and will be able to take the necessary steps to change the car's characteristics.

**Literature review**

Pollution is now a major concern worldwide. Although, there are many reasons for this, vehicles contribute a major part in creating pollution. The amount of CO2 a car emits is directly related to the amount of fuel it consumes. A typical passenger vehicle emits about 4.6 metric tons of carbon dioxide per year. Road transport contributes about one-fifth of the European Union's (EU) total emissions of carbon dioxide (CO2), the main Greenhouse Gas (GHG), 75% of which originates from passenger cars. Considering, the fuel consumption and CO2 emissions of UK alone, published by the Vehicle Certification Agency (VCA), an Executive Agency of the United Kingdom Department for Transport has given information of all the vehicles, including new and old cars, and their properties like fuel consumption, CO2 emissions, model, etc.

Several factors have been identified as contributing to the growing disparity between official fuel consumption and real-world driving conditions. Due to the variety of operating conditions, driver behavior, vehicle usage, and other external factors, it is unlikely that any test protocol, no matter how carefully designed, will be able to accurately capture the real-world performance of vehicles. As a result, there will always be a need to identify which factors influence emissions under real-world driving conditions and which are captured by vehicle certification tests to assess their impact on real-world fuel consumption.

In the research paper [1], Georgios Fontaras, has compared official fuel consumption conditions and real-world driving conditions. Through a qualitative perspective, he performed a Reality vs Certification gap estimation for an average 2015 passenger car. He explained the reasons and the assumptions considered for the official fuel conditions, that lead to a gap between them. Analyzation of potential influence of different factors on CO2 emissions over real-world conditions compared to the official test value is done. Average CO2 emissions based on various factors of the car are determined. The NEDC-based procedure for CO2 and fuel consumption measurement is considered.

Juliet Namukasa in [2] aimed at establishing the relationship between age, engine capacity, fuel type and fuel efficiency and carbon emissions. The finding showed a relation between a vehicle's age, engine capacity, vehicle category (Light Duty, Medium Duty, and Heavy Duty), and fuel efficiency and carbon emissions. Petrol vehicles had lower fuel consumption and emissions, whereas diesel vehicles had higher carbon emissions and fuel consumption due to increased engine capacity. Most of the fuel efficiency and carbon emissions research relies on survey data rather than panel data.

In [3], estimated GHG emissions from example journeys across the UK, comparing different modes of transport, and presents statistics on GHG emissions and air pollution on transport. The author considered three cases namely, indirect emissions, direct emissions, indirect effects. Map plots indicating **CO2 emissions from transport by local authority, 2019; Percentage change in CO2 emissions from transport, 2005 to 2019 in UK are analyzed and plotted. ‘Emissions and mileage for cars, vans, HGVs and buses in 2019’ and change in emissions are produced as a table plot to give insights on the amount of CO2 emitted by cars.**

**Materials and Methods**

This publicly accessible online dataset was obtained from the gov.uk, transport department of UK. This dataset contains information about the description of cars and fuel efficiencies like manufacturer, model, fuel type, noise levels(dB), carbon emissions(mg/km), engine power in Kw. The structure of the data obtained from the new test is one of the immediately noticeable differences. The terms "urban," "extra-urban," and "combined," which were previously used to describe test cycle driving phases for petrol/diesel cars, are no longer acceptable. The data is being expressed as "Low," "Medium," "High," "Extra-high," and "Combined" in their place.   This dataset could be loaded into many tools more easily because it was obtained as a CSV file. Dataset consists of 5678 rows. This dataset contains data from 2000 to 2013 in UK alone. All information about all new cars currently on sale in the UK and used cars that were first registered on or after 1 March 2001.

Data cleaning was the first step, and Python was mainly used for this. The categorical values in the dataset contained a significant number of null values. To make the data easier to analyze and visualize, these data columns were removed. A few string values also needed to be changed due to misspellings and spaces between words in the variables. I was able to eliminate duplicate values by doing this. I began statistical analysis after the data had been cleaned. Python and RStudio both assisted me in creating analyses. To make visualizations and interpretations, I used R studio and Python. Oracle SQL Developer was additionally employed to analyze and extract some data from the dataset based on frequencies and percentages.

For the summary and visualizations of the dataset in Python, libraries like Matplotlib, Pandas, NumPy, etc. were used. Matplotlib for plotting the visualizations, Pandas is used to handle and analyze data, and NumPy for adding support for large, multi-dimensional arrays and matrices, along with a large collection of high-level mathematical functions to operate on these arrays.

Data analysis has been done using R programming and libraries like ggplot2, tidy verse, GGally, tinytex. ggplot2 for visualizations, GGally, a helper to ggplot2. I've created simple MySQL queries and shown the results.

**Results**

1. *Which car manufacturer produces the most air pollution and noise pollution?*

*Chart, bar chart, histogram

Description automatically generated*

Fig.1- car manufacturer producing the most air pollution

*Table

Description automatically generated*

Fig.2-Table viewing the bar chart information

The above visualization is used to display the car manufacturers with respect to carbon dioxide emissions. I used bar chart to create this visualization. We can interpret that Rolls Royce manufacturer is causing the most air pollution of 380 mg/km. It is indeed surprising to see this.

Chart, bar chart

Description automatically generated

Fig.3- Car manufacturer producing the most noise pollution

The above visualization is used to display the car manufacturers with respect to noise levels. I used bar chart to create this visualization. We can interpret that Kia manufacturer is causing the most noise pollution of 89.2 db.

Graphical user interface, application, table, Excel

Description automatically generated

Fig.4-Table viewing the bar chart information

1. *Which car model produces the most air pollution and noise pollution?*

Chart, bar chart

Description automatically generated

Fig.5- Car model producing the most air pollution

The above visualization is used to display the car models with respect to carbon emissions. I used bar chart to create this visualization. We can interpret that Cullinan model is causing the most air pollution.

Chart, bar chart

Description automatically generated

Fig.6- Car model producing the most noise pollution

The above visualization is used to display the car models with respect to noise levels. I used bar chart to create this visualization. We can interpret that GLE Estate model year 2023 model is causing the most noise pollution.

1. *How are fuel costs and emissions related?*

Chart, bar chart

Description automatically generated

Fig.7- Fuel type v/s WLTP Metric Low fuel consumption

The above graph is created to display fuel type and its fuel consumption - WLTP\_Metric\_Low. This means that the fuel consumptions in urban areas are plotted above. This graph depicts that diesel consumption is more in urban areas.

Chart, bar chart, histogram

Description automatically generated

Fig.8- Fuel type v/s WLTP Metric Medium fuel consumption

The above graph is created to display fuel type and its fuel consumption - WLTP\_Metric\_Medium. This means that the fuel consumptions in Semi-urban areas are plotted above. This graph depicts that Petrol consumption is more in Semi-urban areas.

Chart, bar chart, histogram

Description automatically generated

Fig.9- Fuel type v/s WLTP Metric High fuel consumption

The above graph is created to display fuel type and its fuel consumption - WLTP\_Metric\_High. This means that the fuel consumptions in Rural areas are plotted above. This graph depicts that Diesel consumption is more in Rural areas.

Chart, bar chart, histogram

Description automatically generated

Fig.10- Fuel type v/s WLTP Metric Medium fuel consumption

The above graph is created to display fuel type and its fuel consumption - WLTP\_Metric\_Extra\_High. This means that the fuel consumptions on highways are plotted above. This graph depicts that Diesel consumption is more in Rural areas.

From these above graphs we can understand that diesel consumption is more in all the areas.

Chart, bar chart

Description automatically generated

Fig.11- Fuel type v/s CO2 Emissions

The above graph displays fuel type and Co2 emissions as a bar chart. From this graph we can tell that Co2 emissions are more when fuel type is petrol. The next place goes to diesel. From the above graphs between fuel types and fuel consumptions, we can insist that Petrol is the major contributor of air pollution. As diesel is mostly used in many areas, Diesel is causing major pollution. As the diesel is the most common fuel type the diesel price must be low. Electricity and LPG are causing less pollution and the consumption of fuel is also less. Hence LPG must be provided to citizens at lower prices.

1. *What are the characteristics of the car with low emissions?*

Table

Description automatically generated

Fig.12- Table view of manufacturer and carbon emissions in ascending order

The above table is created by ordering the dataset according to carbon emissions in ascending order. From the above table it’s clear that ‘KIA’ manufacturer has the lowest emissions.

Table

Description automatically generated

Fig.13- Table view of model and carbon emissions in ascending order

From the above table it’s clear that ‘Sportage MY22’ model has the lowest emissions.

Similarly, “'GT-Line' 1.6 T-GDI 261 bhp PHEV” is the description, ‘AT6’ transmission, ‘Automatic’ gear, with ‘1598’ engine capacity, ‘Electricity/Petrol’ fuel type, ‘Plug-in Hybrid Electric Vehicle (PHEV)’ power train are the characteristics that a car should have for to release less emissions.

Table

Description automatically generated

Fig.14 – Table view for the SQL query

The above figure show the dataset after loading the data in SQL by importing the data and writing the SQL query. ‘select \* from car;’ , query is used to display the table loaded into the SQL.

Graphical user interface

Description automatically generated with medium confidence

Fig.15 – Table view for the SQL query displaying fuel type and engine power

‘select fuel type, engine power from car where WLTPCO2>100;’ query is used to display the fuel types and engine power for those where engine power is greater than 100 Kw.

Chart, scatter chart

Description automatically generated

Fig.16-Scatterplot displaying fuel type with respect to co2 emissions and noise levels

The above graph is a scatterplot displaying fuel types considering noise level and carbon emissions. The graph provides us the following insights:

1. Electric provides more noise pollution
2. Petrol causes more carbon emissions
3. The carbon emissions increase with noise level

Chart, scatter chart

Description automatically generated

Fig.17-Scatterplot displaying different emissions

The above graph is a scatter plot of all the four emissions. They are

**CO2 emissions [g/km]:**

The main byproducts of burning gasoline or diesel for energy are carbon dioxide and water vapor. The most significant greenhouse gas causing climate change is carbon dioxide. A car will produce about its own weight in CO2 after 6000 miles.

**CO Emissions [mg/km]:**

Because carbon monoxide decreases the blood's ability to carry oxygen, less oxygen may be available to vital organs.

**HC Emissions [mg/km]**

The creation of ozone is aided by hydrocarbons. Some HCs have the potential to cause cancer and act as indirect greenhouse gases.

**NOx emissions ( mg/km)**

Nitrogen dioxide (NO2) is created when nitrogen oxides react in the atmosphere, which can be harmful to health, especially for those who already have respiratory conditions. Additionally, NOx can harm vegetation and contribute to the formation of smog and acid rain.

The green points indicate carbon emissions whereas the black points indicate hydrocarbons and nitrogen dioxide emissions. The plot indicates that carbon dioxide and carbon monoxide emissions are more than hydrocarbons and nitrogen dioxide.

Chart, scatter chart

Description automatically generated

Fig.18-Scatterplot displaying different fuel consumptions

The above graph is a scatter plot of all the four Fuel consumptions. They are

**W LTP Metric Low:**

These are the fuel consumption values in urban areas according to WLTP.

**W LTP Metric Medium:**

These are the fuel consumption values in Semi-urban areas according to WLTP.

**W LTP Metric High:**

These are the fuel consumption values in Rural areas according to WLTP.

**W LTP Metric Extra High:**

These are the fuel consumption values in Highway areas according to WLTP.

The blue points indicate Urban and Semi urban areas whereas the red points indicate rural and highway areas. The above scatter plot indicates that there are more vehicles in urban and semi urban areas than in rural and highway areas.

**Limitations and further research needed**

I have faced many limitations while performing research. Firstly, the dataset is large and and consisted of errors. As the dataset is large cleaning the data was difficult. Secondly, different software’s gave different outputs. I had to create and sort tables to compare the visualizations for the right answer.

In the 4 research questions, I didn’t consider the noise pollution. Kia manufacturer has low carbon emissions but more noise pollution. Hence, one must conduct more research and analysis to determine the optimal car characteristics if they want to reduce the noise levels too.

**Conclusion**

Overall, this dataset gives us insights about the car specifications and requirements. It also mentions the emissions and noise levels from various models and manufacturers. The research questions are answered and gives us result that are beneficial for the society.

1. *Which car manufacturer produces the most air pollution and noise pollution?*

From the visualizations we can tell that ‘Rolls Royce’ and ‘Kia’ manufacturer is causing more pollution hence he should consider the factors that are leading to the more emissions and try to avoid them.

1. *Which car model produces the most air pollution and noise pollution?*

From the visualizations and interpretations, we can tell that ‘Cullinan’ ab ‘GLE Estate model year 2023’ models are causing more air and noise pollution. Proper measures can be taken for the models by trying to change the design or engine.

1. *How are fuel costs and emissions related?*

Diesel is the most common fuel used in many areas including rural areas. Hence, diesel is contributing to the most of carbon emissions. Although Diesel is mostly used, Petrol causes more pollution. LPG, Electricity led to less emissions. Hence, public must be made aware of this information and government measures must be taken to use the fuels that create less pollution.

1. *What are the characteristics of the car with low emissions?*

The characteristics, a car must have so that it produces less emissions are ‘Kia’ is the manufacturer. ‘Sportage MY22’ model has the lowest emissions, “'GT-Line' 1.6 T-GDI 261 bhp PHEV” is the description, ‘AT6’ transmission, ‘Automatic’ gear, with ‘1598’ engine capacity, ‘Electricity/Petrol’ fuel type, ‘Plug-in Hybrid Electric Vehicle (PHEV)’ power train are the characteristics that a car should have for to release less emissions.

**References**

Bishop, J. D. K., Martin, N. P. D., & Boies, A. M. (2014). Cost-effectiveness of alternative powertrains for reduced energy use and CO2 emissions in passenger vehicles. Applied Energy, 124, 44-61. <https://doi.org/10.1016/j.apenergy.2014.02.019>

Environmental Protection Agency. (n.d.). EPA. Retrieved June 30, 2022, from https://www.epa.gov/greenvehicles/greenhouse-gas-emissions-typical-passenger-vehicle#:~:text=typical%20passenger%20vehicle%3F-,A%20typical%20passenger%20vehicle%20emits%20about%204.6%20metric%20tons%20of,8%2C887%20grams%20of%20CO2.

Fontaras, G., Zacharof, N. G., & Ciuffo, B. (2017). Fuel consumption and CO2 emissions from passenger cars in Europe–Laboratory versus real-world emissions. Progress in Energy and Combustion Science, 60, 97-131. <https://doi.org/10.1016/j.pecs.2016.12.004>

*The Worldwide Harmonised Light Vehicle Test procedure (WLTP)*. Vehicle Certification Agency. (n.d.). Retrieved December 5, 2022, from https://www.vehicle-certification-agency.gov.uk/fuel-consumption-co2/the-worldwide-harmonised-light-vehicle-test-procedure/

Fontaras, G., Zacharof, N.-G., & Ciuffo, B. (2017, February 21). *Fuel consumption and CO2 emissions from passenger cars in Europe – laboratory versus real-world emissions*. Progress in Energy and Combustion Science. Retrieved October 30, 2022, from <https://www.sciencedirect.com/science/article/pii/S0360128516300442>

Namukasa, J., Namagembe, S., & Nakayima, F. (2020, January). *Fuel efficiency vehicle adoption and carbon emissions in a country context*. Fuel Efficiency Vehicle Adoption and Carbon Emissions in a Country Context. Retrieved October 30, 2022, from https://www.researchgate.net/publication/338537706\_Fuel\_Efficiency\_Vehicle\_Adoption\_and\_Carbon\_Emissions\_in\_a\_Country\_Context/fulltext/5e1a7e0a4585159aa4c8bd6e/Fuel-Efficiency-Vehicle-Adoption-and-Carbon-Emissions-in-a-Country-Context.pdf

GOV.UK. (2021, October 19). Transport and environment statistics: Autumn 2021. Retrieved October 30, 2022, from https://www.gov.uk/government/statistics/transport-and-environment-statistics-autumn-2021/transport-and-environment-statistics-autumn-2021#mileage-and-fuel-use

*Car Fuel Data, CO2 and vehicle tax tools*. Select a search : Directgov - Car fuel data, CO2 and vehicle tax tools. (n.d.). Retrieved October 17, 2022, from https://carfueldata.vehicle-certification-agency.gov.uk/downloads/download.aspx?rg=latest