The Effect of Car Weight and Volume on CO2 Emissions

Sarah Bareh

Western Governors University

Table of Contents

A. Project Overview

A1. Research Question or Organizational Need

The goal of this project is to understand how the weight and volume of a car can impact its CO2 emissions, which will provide valuable insights to Future Cars, a car manufacturer aiming to have a positive impact on the environment. By analyzing this relationship, Future Cars will be able to make informed decisions about designing and manufacturing environmentally friendly cars.

A2. Context and Background

Environmental awareness has increased over the last decade and many manufacturers have chosen to explore ways to improve their impact on the environment. Future Cars is a car manufacturer committed to reducing the environmental impact of their vehicles and wants to understand the factors influencing CO2 emissions. Car weight and volume are important variables that could potentially contribute to higher or lower CO2 emissions. By studying the relationship between these variables, Future Cars can develop strategies to minimize emissions and improve their overall environmental impact.

A3. Summary of Published Works

- Less heavy vehicles cut CO2 emissions - international transport forum. (n.d.). https://www.itf-oecd.org/sites/default/files/docs/less-heavy-vehicles-cut-co2-emissions.pdf

  The International Transport Forum, an intergovernmental organization with 59 member countries, published a case-specific policy analysis stating that over the past forty years, the average weight of passenger cars in the European Union has risen by approximately 40%. In 2015, the average vehicle weighed 1,400 kg, compared to under 1,000 kg in 1975. The additional weight leads to higher energy consumption and increased CO2 emissions, making a reduction in vehicle weight crucial for achieving emissions reduction targets. Using a vehicle stock model, this study establishes two scenarios: a baseline scenario that assumes current policy trends continue, and a mass reduction scenario that assumes the average weight of new vehicles declines to levels seen four decades ago. In the vehicle mass reduction scenario, a gradual decrease in vehicle weight leads to a significant

reduction in CO2 emissions. Compared to 1990 levels, CO2 emissions decrease by 39% or 210 Megatonnes by 2050, with passenger cars contributing to around 85% of the reductions.

- Serrenho, A. C., Norman, J. B., & Allwood, J. M. (2017, June 13). *The impact of reducing car weight on global emissions: The Future Fleet in great britain*. Philosophical transactions. Series A, Mathematical, physical, and engineering sciences. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5415645/

  This study explores the impact of weight reduction of vehicles on CO2 emissions. The findings suggest that reducing vehicle weight can result in greater cumulative emissions savings by 2050. Lighter cars lead to less energy requirements and emissions during its use, which results in significantly lower net emissions.

- Frederik, J. (2020, February 24). *Your car has a weight problem and we need to regulate it*. The Correspondent. https://thecorrespondent.com/310/your-car-has-a-weight-problem-and-we-need-to-regulate-it

  This article discusses the relationship between car weight and CO2 emissions. It specifically addresses the issue of increased manufacturing of Sports Utility Vehicles (SUVs) which is having a significant impact on the environment. SUVs are traditionally heavier cars, weighing well over 8600 pounds. The article states that SUVs account for a significant portion of new car sales globally, that they have become the second largest source of CO2 emissions after electrical power generation, surpassing emissions from agriculture and aviation.

A3a. Relation of Published Works to Project

The published works provide valuable insights that directly relate to the project's objective of understanding the relationship between car weight and volume, and CO2 emissions. They offer insights into the potential benefits of weight reduction strategies, the importance of policy interventions, and the environmental impact of heavy vehicles. Integrating the findings and

recommendations from these works into the project's planning can enhance its comprehensiveness and effectiveness in addressing the challenges associated with car weight and volume, and emissions.

A4. Summary of Data Analytics Solution

This project will be investigating the relationship between car weight and volume and CO2 emissions through a python notebook. This notebook will analyze the relationship by using statistical calculations such as correlation coefficient, p-value, r-squared, mean, and median. These calculations will all be based on a preexisting dataset that includes information about car model, volume, weight and CO2 emissions. The input will be the data points from the dataset representing the variables of interest. The output will be scatterplots and multilinear regression table describing the relationship between car weight, volume and CO2 emissions.

A5. Benefit to Organization and Decision-Making Process

The project's stakeholders include Future Cars, other automotive manufacturers, governmental organizations and officials, and environmental groups and activists. The project's goal is to find out how the weight and volume of a car, is related to its CO2 emissions. The project will benefit stakeholders by analyzing and explaining the relationship between CO2 emissions and the car weight and volume. Future Cars and other manufacturers can use these insights to design more environmentally friendly cars, while government agencies and policymakers can make informed decisions regarding regulations and standards. Environmental groups and activists can use the project's findings to promote eco-friendly transportation methods.

B. Data Analytics Plan

B1. Goals, Objectives, and Deliverables

This project will use python programming language to analyze the relationship between CO2 emissions and the car weight and volume, with the aim of helping Future Cars to improve their environmental impact and reduce emissions. The objectives for this goal are:

- Data Collection and Preparation:

    - Deliverables are:
        - Cleaned dataset that contain no missing values, outliers, or formatting issues.
        - Exploratory data analysis to gain insights into the dataset and identify any patterns or relationships.
- Statistical analysis by applying multilinear regression to investigate the relationship between car weight, volume, and CO2 emissions.
    - Deliverable: A table containing calculations of relevant metrics such as correlation coefficient, p-value, and R-squared to assess the statistical significance and strength of the relationship.
- Interpretation of the results to provide meaningful insights into how car weight and volume impact CO2 emissions.
    - Deliverables are:
        - A comprehensive data analysis report that summarizes the findings of the statistical analysis with clear explanations and interpretations of the statistical metrics and their implications.
        - Appropriate visualizations, such as scatterplots and tables to communicate the relationship between car weight, volume, and CO2 emissions.

B2. Scope of Project

The scope of this project will include an analysis of the relationship between car weight, volume, and CO2 emissions, based on a preexisting dataset. It will include the development of an analytical report that will explain how CO2 emissions are based on car weight and volume. The scope of the project will not include physical modification or alteration of vehicles to reduce weight or improve environmental performance. It will not include the collection of real-time or live data on CO2 emissions from vehicles or any regulatory or policy changes regarding vehicle weight or emissions standards. It does not involve direct implementation or enforcement of measures to reduce emissions.

B3. Standard Methodology

The project will use the Waterfall methodology, which involves a sequential approach to project phases. The first phase is requirements gathering, where the specific requirements of the project will be identified and documented. The project scope, stakeholder expectations, and the resources required to finish the project will be defined.

The project will move into the design phase, during which a detailed plan will be created for developing the analytical report. The design will involve determining which programming language and environment will be used.

The next phase is the implementation phase, where the designed project will be implemented using appropriate programming languages and tools to achieve the project requirements. After the development phase, the project will proceed to the testing phase, where the functionality and accuracy of the developed code will be assessed.

Finally, the project will enter the maintenance phase, where the code will be monitored for performance. Any issues or bugs that arise will be addressed promptly. Feedback from users and stakeholders will be collected to identify potential improvements or enhancements. The code will be updated as needed to incorporate new data or address changing requirements.

B4. Timeline and Milestones

The following table indicates each milestone its projected start and end dates, and its projected duration:

Milestone	Projected Start Date	Projected End Date	Duration (days/hours)
Requirement gathering	06-15-2023	06-17-2023	2 days
Design	06-17-2023	06-20-2023	3 days
Implementation	06-20-2023	06-21-2023	1 days
Testing	06-21-2023	06-22-2023	1 days
Maintenance	06-21-2023	Ongoing	N/A

B5. Resources and Costs

Category	Cost Description	Amount (USD)
Personnel	Project Owner/Developer	0
Technology	Personal Computer	0
Infrastructure	Existing Internet Connection	0
Total Cost		0

In this table, the costs are listed as zero since there are no financial expenses involved. The project will be managed, developed, and owned solely by me. The technology cost includes the use of my personal computer for the project, which is already available. The infrastructure cost is also zero as it utilizes my preexisting internet connection for accessing necessary tools and resources.

B6. Criteria for Success

The success of this project will be assessed based on the following metrics:

Criterion/Metric	Required Data	Cut Score for Success
Strong Correlation	The dataset and statistical calculations.	A high correlation coefficient between car wight, volume and CO2 emissions indicates a strong relationship.
Statistically Significant Results	Statistical calculations	p-value below a predetermined significance level indicate a statistically significant relationship between car wight, volume and CO2 emissions.
High Predictive Power	Multilinear regression model	A high r-squared indicates that the model can effectively predict CO2 emissions based on car weight and volume.

C. Design of Data Analytics Solution

C1. Hypothesis

There is a strong positive relationship between car weight and volume and the amount of CO2 emissions produced by it.

C2. Analytical Method

The analytical method that will be used is multilinear regression. Multilinear regression is both a descriptive and a predictive analytical method. It involves creating a statistical model that would describe and possibly predict the relationship between the variable involved in this project.

C2a. Justification of Analytical Method

Multilinear regression provides a clear and interpretable relationship between the independent variables (car weight and volume) and the dependent variable (CO2 emissions). We can also determine the statistical significance of the variables through the regression coefficients which indicate the magnitude and direction of the impact of each independent variable on the dependent variable, allowing for easy understanding and explanation of the relationship.

C3. Tools and Environments of Solution

Python is selected for data extraction in this project, as it will provide powerful tools and functions to manipulate the dataset. Jupyter notebooks will be used as an environment to execute the python code. Python is a versatile and widely used programming language for data analysis. It includes many powerful libraries as Pandas, NumPy, Scikit and Matplotlib. Pandas provide convenient functions for reading data from various files and allows for efficient data filtering, transformation, and aggregation. NumPy and Matplotlib provide functions for data analysis and visualization. The environment that will be used is Jupyter notebooks as it will allow me to execute my code in line and debug it instantly. It will also allow for visualizations to be displayed in line. It is a free tool that can be used locally on my machine.

C4. Methods and Metrics to Evaluate Statistical Significance

The statistical significance will be evaluated by creating a null hypothesis and an alternative hypothesis. The null hypothesis will be that there is no relationship between car weight and volume and CO2 emissions. My goal will be to reject the null hypothesis by using metrics of statistical significance such as z-score, p-value, alpha score. If p-value is lower than alpha score, then we will be able to reject the null hypothesis and accept the alternative hypothesis that there is a positive relationship between car weight and volume and CO2 emissions.

C4a. Justification of Methods and Metrics

These metrics and methods provide objective measures to evaluate the significance of the relationship between variables. As a result, we can use these findings to make evidence-based conclusions that will enhance the credibility and reliability of the project findings.

C5. Practical Significance

The findings of this project will have a practical significance by affecting the decision-making process involved in manufacturing environmentally friendly cars. As an example, Future car and other car manufacturers can use these findings to design and develop cars by choosing lightweight materials, aerodynamic designs and efficient engine technologies that can reduce emissions. Policymakers can use the results to write new laws and standards to regulate the production of eco-friendly cars.

C6. Visual Communication

the findings of this project will be communicated by using scatter plots and tables. Scatter plots can display the relationship between the variables of interest. It can be easily interpreted to indicate a positive or a negative relationship. I will also use a table to display the results of the multilinear regression analysis. The table will include important metrics such as p-value, correlation coefficient and r-squared.

D. Description of Datasets

D1. Source of Data

The dataset that will be used is preexisting on Kaggle.com under the title CO2 emission of cars dataset.

L, M. D. (2022, July 18). *CO2 Emission of Cars Dataset*. Kaggle.

https://www.kaggle.com/datasets/midhundasl/co2-emission-of-cars-dataset

D2. Appropriateness of Dataset

This dataset was chosen as it closely correlates with purpose of this project. It includes 5 variables, car make, model, volume, weight, and CO2 emissions, which directly relate to the research question. It is available under a public domain license so it can be used without any legal ramifications.

D3. Data Collection Methods

The dataset was obtained from Kaggle, a popular platform for hosting and sharing datasets. It is a publicly available dataset and was selected for this project due to its relevance to the research question. It was not collected specifically for this project but was compiled by the creator and uploaded to Kaggle.com. The specific steps taken to ensure accuracy and quality during the data collection process are not explicitly mentioned in the dataset documentation.

The advantage of using a preexisting dataset is the convenience and time-saving aspect. Collecting data from scratch can be a time-consuming process, requiring significant effort in data collection, cleaning and validation. By utilizing an existing dataset, these steps are already completed which allows for a more efficient analysis. However, the disadvantage to using preexisting datasets is the lack of control over the data collection process. The dataset may contain errors, missing values, or inconsistencies that can affect the quality and reliability of the analysis.

D4. Data Quality

To ensure data accuracy and quality, I will be checking the data for inconsistencies, outliers, missing or duplicate values. The dataset will be cleaned and appropriately prepared for the analysis. Any columns, or rows with null, missing or inconsistent values will be deleted to ensure consistency and accuracy.

D5. Data Governance, Privacy and Security, Ethical, Legal, and Regulatory Compliance

Data governance, privacy and security, ethical considerations, and legal and regulatory compliance are critical aspects to address when working with any dataset. In this project, data governance is the responsibility of the data creator, who uploaded it on Kaggle.com. The dataset does not contain any sensitive information. It is available under a public domain license so there will be no ethical or legal implications for using it.

D5a. Precautions

For data preparation, the dataset will be cleaned and transformed to ensure data quality before analysis. During analysis, appropriate statistical methods will be used and documented for

reproducibility. The findings and the code used during the project will be backed up regularly to protect against data loss. When disseminating data, it will be presented in an easily understandable format for the intended audience.

E. Sources

- Serrenho, A. C., Norman, J. B., & Allwood, J. M. (2017, June 13). *The impact of reducing car weight on global emissions: The Future Fleet in great britain*. Philosophical transactions. Series A, Mathematical, physical, and engineering sciences. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5415645/
- Less heavy vehicles cut CO2 emissions - international transport forum. (n.d.). https://www.itf-oecd.org/sites/default/files/docs/less-heavy-vehicles-cut-co2-emissions.pdf
- Frederik, J. (2020, February 24). *Your car has a weight problem and we need to regulate it*. The Correspondent. https://thecorrespondent.com/310/your-car-has-a-weight-problem-and-we-need-to-regulate-it