

7COM1079-0901-2024 - Team Research and Development Project

**Is there a correlation between the road tax and the mileage of used Toyota cars sold in the UK between the year 2010 and 2020?**

Group ID: A134

Dataset Number: DS138

Prepared by: 23068822 Laban Mahihu  
 23026708 Isaac Ngulat

23099101 Martin Thuranira

23080370 Emmanuel Chinedu

23004860 Onatemowo Oluwadamilare Temidayo

University of Hertfordshire

Hatfield

# Table of Contents

Table of Contents 2

1. Introduction 3

1.1. Problem Statement and Research Motivation 3

1.2. The Dataset 3

1.3. The Research Question 3

1.4 Null Hypothesis and Alternative Hypothesis 4

2. Background Research 5

2.1 Research Papers 5

2.2 Why the Research Question is of Interest 5

3. Visualization 6

3.1 Plots 6

3.2 Additional Information Relating to Understanding the Data 7

3.3 Useful Information for Data Understanding 7

4. Analysis 8

4.1 Statistical Tests Used 8

4.2 Null Hypothesis Interpretation 8

5. Evaluation 9

5.1 What went well 9

5.2 Points for Improvement 9

5.3 Group Time Management 10

5.4 Project’s Overall Judgement 10

5.5 Comment on Github Log Output 11

6. Conclusion 12

6.1. Results Explained. 12

6.2. Interpretation of the Results 12

6.3 Reasons and/or Implications for Future Work, Limitations of Our Study 12

References 13

# 

# Introduction

## 1.1. Problem Statement and Research Motivation

The relationship between road tax and mileage of used Toyota cars in the UK (2010–2020) remains underexplored. This study investigates whether a significant correlation exists between tax paid and mileage is driven, offering insights into consumer behaviour, vehicle valuation, and taxation in the automotive market. Although potential links between taxation systems and abandoned cars, tax avoidance, and scrappage schemes were considered, no clear connection was found (European Commission, 2021). A pay-by-mile tax system could align incentives, reduce emissions, and simplify cost calculations (Miles, 2021). Understanding this relationship aids stakeholders in vehicle pricing, purchase decisions, and taxation policy design.

## 1.2. The Dataset

Our dataset (DS138 - Toyota.csv) contains data on 6,738 Toyota vehicles sold in the UK between 2010 and 2020. It stores information such as the car's model, year, price, transmission type, mileage, fuel type, road tax, miles per gallon (mpg), and engine size. This comprehensive dataset provides a foundation for investigating the relationship between mileage (independent variable) and road tax (dependent variable), testing hypotheses to see if a correlation exists.

## 1.3. The Research Question

The research question aims to find out whether there is a correlation between road tax and mileage for used Toyota cars in the UK between the years 2010 and 2020. We aim to use statistical methods such as correlation tests and visualisations to answer the question and also find out the strength and direction of the relationship between the two variables.

## 1.4 Null Hypothesis and Alternative Hypothesis

The null hypothesis (H₀) states that there is no correlation between road tax and mileage for used Toyota cars sold in the UK between 2010 and 2020. This indicates that any observable relationship between the two variables is the result of random chance.

The alternative hypothesis (H₁) implies that there is a significant relation between road tax and mileage. Testing these hypotheses entails using statistical analysis, such as Pearson's or Spearman’s correlation, to evaluate whether the null hypothesis can be rejected, hence establishing the alternative hypothesis and revealing a potential link between road tax and mileage.

# 

# 2. Background Research

## 2.1 Research Papers

Road tax is officially known as Vehicle Excise Duty (VED) and is a charge that should be paid by any vehicle that is registered in the UK and is driven or kept on a public road (HM Revenue & Customs, 2024). The rates payable on each vehicle depend on the type of the vehicle and the date it was first registered. Since 2010, the UK government has adjusted the rates in accordance with inflation (HM Revenue & Customs, 2024). Vehicles that were first registered after 1st April 2017, are taxed against three categories; zero, standard or premium. On the other hand, vehicles registered between 1st March 2001 to 31st March 2017 are taxed against CO2 bands. The bands are classified to have lower rates with vehicles containing lower tailpipe emissions (HM Revenue & Customs, 2024).

Before March 2021, the UK government imposed road tax based on the size of the engine. This meant that cars with larger engine sizes paid more taxes than the latter (Cerruti, Alberini and Linn, 2017). However, After March 2021, vehicles were placed on a CO2 emissions band that determines tax in g/km. According to (gov.uk, 2011) vehicles on tax band A with CO2 emissions of up to 100g/km pay £0 which is the lowest amount while vehicles on tax band M with CO2 emissions of upto 255g/km pay up to £771.75.

In this research, we have taken a general assumption that cars with generally older registration dates have a higher likelihood of having a higher mileage. This is attributed to the fact that the government does not hold a public database of vehicles taxed against its mileage.

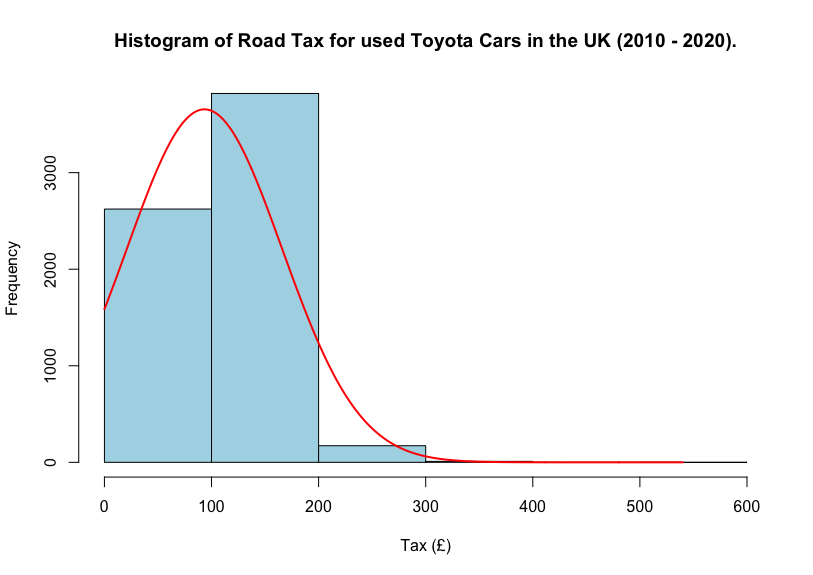
## 2.2 Why the Research Question is of Interest

Tax and mileage is surprisingly a low research topic with very little research conducted to understand the relationship between the two. When a first time car buyer is on the market for a new car, they tend to overlook how mileage may affect the amount of tax due on their vehicle, specifically, Toyota cars. It is for this reason that we have delved into this research topic to provide more insights on how taxes are related to mileage on Toyota vehicles.

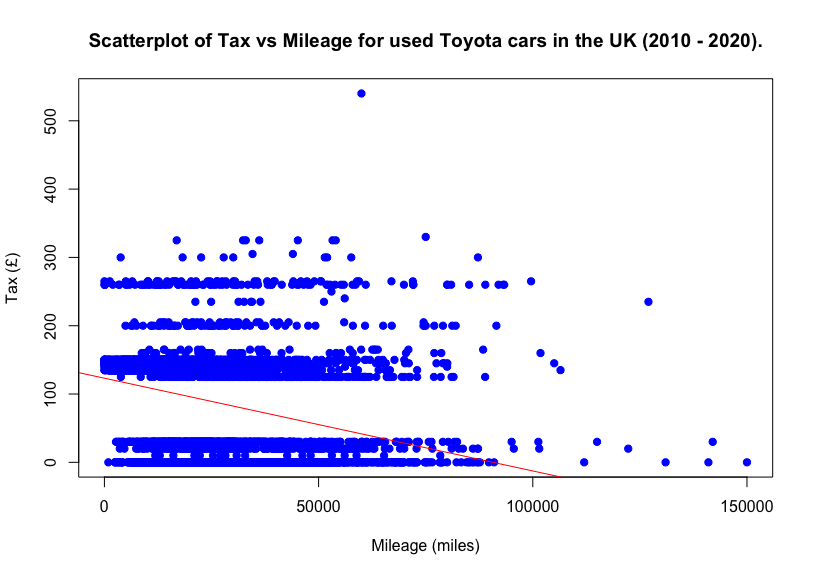
# 3. Visualization

## 3.1 Plots

Since our research question is a correlation question, we will be implementing histogram and scatter plot diagrams.



The histogram is used to see the distribution of the dependent variable, tax, which in turn helps us to find out whether the data is normally distributed or not. The scatter plot is then used to visualise the relationship between our two variables, tax and mileage.



## 3.2 Additional Information Relating to Understanding the Data

From the histogram, it shows that our dependent variable is skewed to the right, with most values ranging from £0 and £200. Our scatter plot then reveals a slight negative correlation between tax and mileage, with notable clustering patterns and outliers, particularly in tax values between £150 and £300.

## 3.3 Useful Information for Data Understanding

The data displays a clear horizontal banding in tax values, indicating standardized tax brackets. Tax ranges from £0 to £500 and mileage ranges from 0 to 150,000 miles. Some high-mileage Toyota cars have very low tax rates, while others have higher tax rates regardless of their mileage.

# 4. Analysis

## 4.1 Statistical Tests Used

For the correlation test, we used the Kendall method because of various reasons. First, it does not assume a normal distribution, which is important because the histogram of our dependent variable, tax, is not normally distributed. Second, it effectively handles well, which is important considering the scatterplot’s distinct tax brackets. Third, the test is resistant to outliers, which are present in our dataset. Finally, it assesses linear relationships, making it useful for detecting broad trends in tax and mileage.

## 4.2 Null Hypothesis Interpretation

For our research question, we reject the null hypothesis because we obtained a small p-value (< 2.2e-16), which is way lower than the standard significance level of 0.05. Kendall's tau value of -0.2780 implies a weak to moderate negative correlation between road tax and mileage. This shows that when mileage increases, there is a slight trend for road tax to decrease, however, the link is not substantial. The negative correlation is statistically significant, but it may not be practically relevant due to the low tau value and evident clustering in the scatterplot.

# 5. Evaluation

## 5.1 What went well

**Clear** **Focus** **on** **Objectives**: The analysis explored the potential correlation between road tax and mileage, ensuring that all efforts were aligned with the goal.

**Data Relevance**: The study focused on a specific and relevant dataset—Toyota cars sold in the UK from 2010 to 2020—allowing for meaningful conclusions within a defined scope.

**Statistical Accuracy** : Using correlation and regression analyses provided quantitative backing for the conclusions, ensuring the findings were data-driven and objective.

**Actionable Suggestions**: The evaluation provided clear recommendations for further research, ensuring that the study’s insights could be built upon.

## 5.2 Points for Improvement

**Data Completeness**: The analysis could benefit from a more comprehensive dataset that includes additional variables such as CO₂ emissions, engine size, fuel type, and vehicle age to provide a fuller picture of the factors influencing road tax.

**Sample Size and Diversity**: Expanding the sample size to include a wider variety of Toyota models and other car brands could make the findings more generalizable across the used car market.

**Accounting** **for** **Policy** **Changes**: Given the significant shifts in UK road tax policies from 2010 to 2020, incorporating a timeline of these changes would better contextualize the results and account for their impact.

## 5.3 Group Time Management

**Adherence to Deadlines**: The group successfully met key deadlines for various project stages, such as data collection, analysis, and report preparation, ensuring the project remained on track.

**Effective Scheduling**: Regular meetings were scheduled to discuss progress and address challenges, allowing for consistent communication and alignment among team members.

**Task** **Prioritization**: Tasks were effectively prioritized. Critical activities like data cleaning and analysis were completed early, leaving ample time for review and refinement.

**Responsiveness to Challenges**: The group demonstrated flexibility in reallocating time to address unforeseen issues, such as refining the dataset or revisiting analysis methods.

## 5.4 Project’s Overall Judgement

The project achieved its primary objective of analyzing the potential correlation between road tax and mileage of used Toyota cars sold in the UK from 2010 to 2020. The findings provided valuable insights into how road tax policies are influenced more by emissions than mileage while highlighting indirect relationships between mileage, vehicle age, and taxation.

The conclusions were credible and rigorous because of the use of statistical methods, including correlation and regression analysis. Additionally, contextual knowledge about UK road tax policies was incorporated, adding depth to the evaluation. Segmenting data by fuel type and vehicle category further enriched the analysis, making the results more actionable and meaningful.

Overall, the project demonstrates a strong understanding of the topic, excellent teamwork, and solid analytical execution, making it a commendable effort by the group.

## 5.5 Comment on Github Log Output

Commit 1: “Update: Data cleaning, Statistical test, visualisation plots”.

In this commit the following operations were carried out:

* Importing the dataset
* Removing the outliers using Inter Quartile Range Method (IQR)
* Filtering data with less than 10 observations
* Calculating the mean Prices for each model
* Performing Kruskal test on the data
* Generating a box plot and histogram
* Checking for skewness

Commit 2: “Visualisation using histogram and scatterplot of the research question

In this commit the following operations were carried out:

* Use of the hist R function to generate a histogram for the dataset
* Use of the plot R function to generate a scatterplot for the dataset

Commit 3: “Correlation code”

In this commit the following operations were carried out:

* Preparing the dataset
* Preparing the tax histogram
* Calculating the bell curve
* Correlation test between tax and mileage
* Scatter Plot of tax vs Mileage

# 6. Conclusion

## 6.1. Results Explained.

From the analysis, there is a weak negative correlation between road tax and mileage for used Toyota cars sold in the UK between 2010 and 2020. This indicates that vehicles with higher mileage tend to have slightly lower tax values. The Kendall correlation test illustrated this with a significant p-value of 2.2e-16, prompting us to reject the null hypothesis. While this is statistically significant, the correlation is weak, suggesting that mileage alone does not heavily influence road tax rates.

## 6.2. Interpretation of the Results

The results imply that road tax is more likely influenced by factors such as engine size and fuel efficiency(mpg) rather than mileage alone. This finding is essential for both buyers and sellers in the used car market, as it highlights that high-mileage vehicles may offer lower road tax, potentially making them more appealing to budget-conscious buyers. However, given the weak correlation, other factors should be considered when assessing vehicle costs.

## 6.3 Reasons and/or Implications for Future Work, Limitations of Our Study

Future research should address the impact of outliers, consider recent tax policy changes, and explore additional variables such as fuel efficiency. The limitation of our study is that the dataset only covers vehicles up to 2020, potentially excluding critical recent trends. A broader approach examining engine size and fuel efficiency could yield more comprehensive insights.

# 

# References

European Commission (2021) *The taxation framework in the EU automotive market*. Available at:<https://taxation-customs.ec.europa.eu/document/download/924dce3e-ce48-4c70-8561-b30d5623310e_en> (Accessed: 5 January 2025).

Miles, EVP0073 (2021) *Evidence submitted to the UK Parliament*. Available at:<https://committees.parliament.uk/writtenevidence/22813> (Accessed: 5 January 2025).

HM Revenue & Customs (2024). *Vehicle Excise Duty rates for cars, vans and motorcycles — from 1 April 2025*. [online] GOV.UK. Available at: https://www.gov.uk/government/publications/vehicle-excise-duty-rates-for-cars-vans-and-motorcycles-from-1-april-2025/c2addbb5-258e-4636-971e-afa4d78f8672#:~:text=Vehicle%20Excise%20Duty%20is%20a,subject%20to%20Vehicle%20Excise%20Duty. [Accessed 5 Jan. 2025].

‌

Cerruti, D., Alberini, A. and Linn, J. (2017). Charging Drivers by the Pound: The Effects of the UK Vehicle Tax System. *SSRN Electronic Journal*. [online] doi:<https://doi.org/10.2139/ssrn.2967312>.

GOV.UK (2011). *Vehicle tax rates*. [online] GOV.UK. Available at: https://www.gov.uk/vehicle-tax-rate-tables/rates-for-cars-registered-on-or-after-1-march-2001 [Accessed 5 Jan. 2025].

‌