Project Group - 11

Members: Mika Dawud, Luke Tros, Florine Vermeer, Brigitte Nauta and Lisette de Langen

Student numbers: 5379849, 5605946, 5351944, 5356148, 5476402

Introduction

In recent years, fuel prices have been a frequent topic of discussion. Fuel prices rose sharply at the end of 2021 and the beginning of 2022, partly due to the war in Ukraine, prompting the Dutch government to intervene. This was done by introducing an excise duty reduction on gasoline, diesel, and LPG starting from April 1, 2022 (ANWB, n.d.). As of July 1, 2023, half of this 'discount' expired, and by the end of this year, on December 31, 2024, this measure will be fully phased out (FiscaalRendement, 2023).
Despite rising fuel prices, the total number of cars in the Netherlands has not decreased. In fact, it has continued to grow (Compendium voor de Leefomgeving, 2022). This suggests that many Dutch people still consider the car an important mode of transport, regardless of the increasing fuel costs.
In a study at the University of Groningen, it was found that fuel price changes led to different transportation responses among individuals. For some, fuel costs were price-elastic, meaning they adjusted their driving habits and reduced kilometers driven in response to price increases. However, for others, fuel costs were price-inelastic, indicating that their driving behavior remained relatively unchanged despite fuel price fluctuations (Schrier, 2008).

In a study conducted in Japan, researchers estimated the price elasticity of fuel consumption for the entire population. The results indicated a price elasticity of 0.37, meaning that a 1% increase in fuel prices leads to a 0.37% reduction in kilometers driven. This demonstrates moderate responsiveness of fuel consumption to price changes (Knittel & Tanaka, 2019).

The impact of rising fuel prices on the total number of kilometers traveled by the Dutch population, as well as the ways in which individuals might adjust their mode of transportation in response, has not yet been thoroughly studied. This paper aims to address these gaps in the research.

Research Objective

Main research question: How does pump price correlate with travel distance of different modes between 2010 and 2017?

Subquestions:

- How has the pump prices changed between 2010 and 2017?
- How has the travel distance for car and train changed between 2010 and 2017?
- What is the difference in the correlations between provinces with least and the most yearly travel distance?

Contribution Statement

Mika Dawud: Figures and data pipeline for subquestion 2 and adding names for data mobility
Luke Tros: Figures and data pipeline for subquestion 1. Figure and data pipeline for main research question.
Florine Vermeer: Introduction, conclusion, explaining of the graphs of subquestions 1 and 2
Brigitte Nauta: background research, figures and data pipeline subquestion 3 (excl. correlation)
Lisette de Langen: Correlation part of subquestion 3

Data Used

Both of the data files that are used, are from Centraal Bureau voor de Statistiek (CBS).
Data File 1 `'fuel_prices.csv'` gives the pump prices per day from 2006 to 2024. (CBS, 2024.1)
Data File 2 `'mobility_info.csv'` and `'mobility_data.csv'` gives the total travel distances and per mode. (CBS, 2024.2)
Data File 3 `'total_kilometers.csv'` gives the total kilometers travelled per year from 2010 to 2017. (CBS, 2024.3)

Data Pipeline

1. Understand the data and know what it consist of.
2. Data processing: Filtering the data to the data that is needed. Example: change the data to pump prices per month instead of per day.
3. Vizualise: Presenting the data and the results in graphs.

How has the pump price changed between 2010 and 2017?

Formatting data

1) The data is imported using pandas.

2) The index column is set to 'Perioden', the column with the days.

3) The data is filtered, only the years 2010 to 2017 are relevant.

Now that the dataset is imported and the irrelevant years are filtered out, the data is further processed.

Processing data

1) Resample the DataFrame to years and months. We are interested in yearly and monthly average fuel prices.

2) Seperate the fuel types Benzine and Diesel. Since LPG is hardly used anymore, this fuel type is not analyzed further. We now obtained a monthly and yearly set of the average fuel price for both Benzine and Diesel.

3) Use the Benzine and Diesel prices to calculate an average fuel priced, which is used for this analyses and compared to both Benzine and Diesel.

These steps are taken in the file `'Subquestion 1 fuel prices.ipynb'`.

How has the travel distance for car and train changed between 2010 and 2017?
The dataset of the total kilometers travelled first needed to be filtered per transportmode and region. New datasets were created for the total yearly kilometres for all transport modes, car users and train users. The data for car drivers and car passenger are summed for the car users, because it is given in seperate sets and the total kilometers driven per person need to be known for this comparison. After the filtering there are three new dataframes. In these dataframes the data is ordened by year and represent the total kilometres travelled per year. The last step is to change the index names so it refers to the years. The total kilometres can then be plotted agianst the year for all transport modes, car users and train users. The first two rows have names to order the data per province and per transport modes, so for the plots it is neccesary to skip the first two rows by only select the data points out of the set.

This was done in file: `'Total_kilometers_travelled.ipynb'`.

For the average kilometers driven the dataset obtained had code names in the table to reduce the column width. There was an added file with the codenames and the explanantions. This file was used to create a dataframe with the corresponding names instead of the codes. This is done to work easier with the data while filtering.

This is done in file: `'‘Mobility_data_addnames.ipynb’`

To plot the data of average kilometers against time a similar method was used as described in the next paragraph for research question 3. The data for the Netherlands was available the same way it was for the provinces described there. The graphs produced for the total average kilometers for the Netherlands are created in the file `'Mobility_for_Netherlands.ipynb'`

What is the difference in the correlations between provinces with least and the most yearly travel distance?
First, the dataset was loaded using the Pandas library. Subsequently, it was filtered to

retain only the necessary data. In this process, it was decided to include data for the total population rather than only for residents aged twelve and older.

Additionally, the modes of transportation—driver car, passenger car, train, and total—were selected, as only these will be utilized in subsequent analyses.

Comparison of total kilometers per province per year
To compare the data per province per year, the dataset was initially filtered for the transportation mode 'total' and subsequently sorted by provinces. Following this, a date was assigned to the data based on the available year in the dataset. This step ensures that the data can later be plotted chronologically.

Subsequently, separate datasets were created for each province, displaying the total number of kilometers per year. These datasets were then merged into a single dataset in which the data from different provinces are presented horizontally adjacent to one another. This combined dataset was utilized to plot the average number of kilometers per person per province per day.

Kilometers per Year per Mode of Transportation per Province
In a similar manner, the data was filtered by province, after which a date was added using the same methodology. A loop was then executed through all the distinct values in the 'RegioS' column, which represents the various provinces, as filtering had already been applied. Within this loop, a dataset was created for each province concerning the individual modes of transportation. By calling specific segments of the dataset based on province, the data can be visualized for each province. For the purposes of this analysis, this was performed for the provinces of Flevoland and Limburg, as these represent the provinces with the highest and lowest kilometers per person per year, respectively.

All these steps can be found in the notebook `'Mobility_per_province(subquestion3).ipynb'`.

Main research question: How does pump price correlate with travel distance of different modes between 2010 and 2017?

The notebooks `'Subquestion 1 fuel prices.ipynb'` and `'Total_kilometers_travelled.ipynb'` are combined and linear regression is performed. This is done in the file `'Research question correlation.ipynb'`. Also the dataset in the Appendix is used here, to find the correlation with the (extended) fuel prices.

Results

How has the pump prices changed between 2010 and 2017?

First, the development of fuel prices between 2010 and 2017 was studied. For this, data from CBS was used, which provides a monthly updated file of average fuel prices for LNG, benzine, and diesel in the Netherlands. The share of LNG users (3%) was very small,

so it was neglected in our analysis. The share of benzine users was 80% of the total and diesel 17%. The prices for benzine and diesel are plotted in two graphs. Additionally, the average price of benzine and diesel combined is also plotted.
To get a global view of the price trends, the average price per year was plotted, see Figure 1.

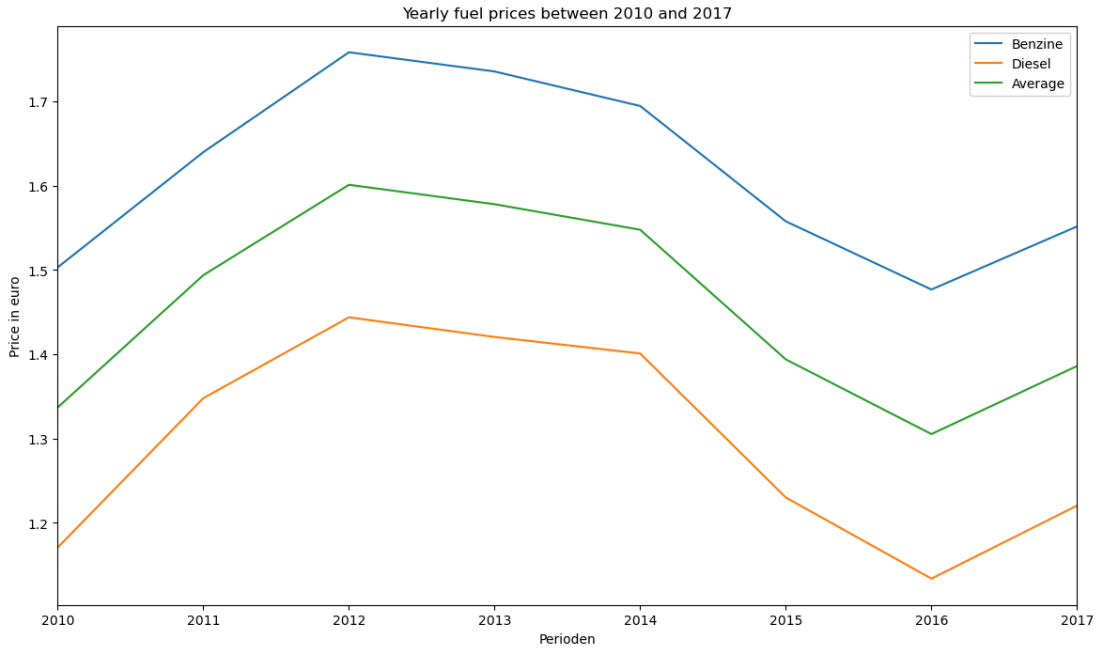

Figure 1: The average fuel prices per year between 2010 and 2017

From the graph, it can be seen that both diesel and benzine follow the same trend, with benzine being slightly more expensive on average than diesel. Prices rose from 2010 to 2012, then started declining after 2014, with an even steeper drop until 2016. From 2016 onwards, prices began rising again.
For a more detailed view of the price development, the price per month was plotted in the Figure 2 below.

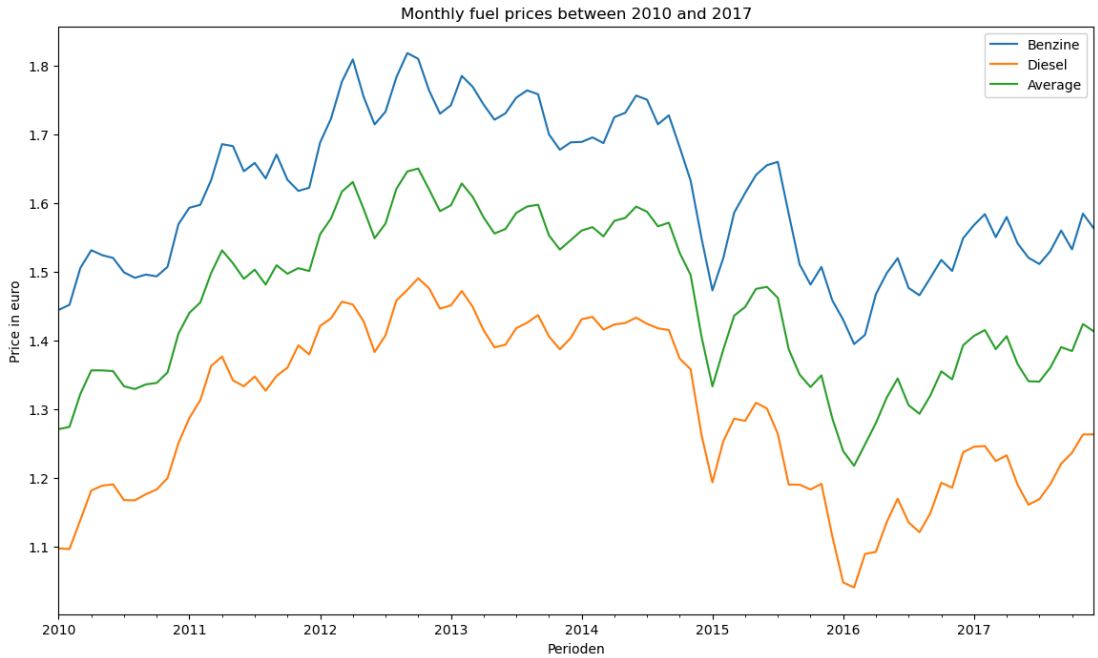

Figure 2: The average fuel prices per month between 2010 and 2017

The highest price for benzine was around €1.80 per liter, and for diesel, around €1.50, in mid-2012. The lowest price for benzine was €1.40, while for diesel, it was €1.50, both at the beginning of 2016. From the monthly graph, it can be seen that the prices follow a similar trend, but not exactly the same. For example, the price of diesel began rising slightly earlier than benzine at the end of 2011.

How has the travel distance for car and train changed between 2010 and 2017?

After studying fuel prices, the development of the average and total kilometers traveled by car and train was examined for the same period (2010-2017). The kilometers traveled by car are divided into two categories: kilometers driven by the car driver and kilometers traveled by passengers. The development is shown in Figure 3 below.

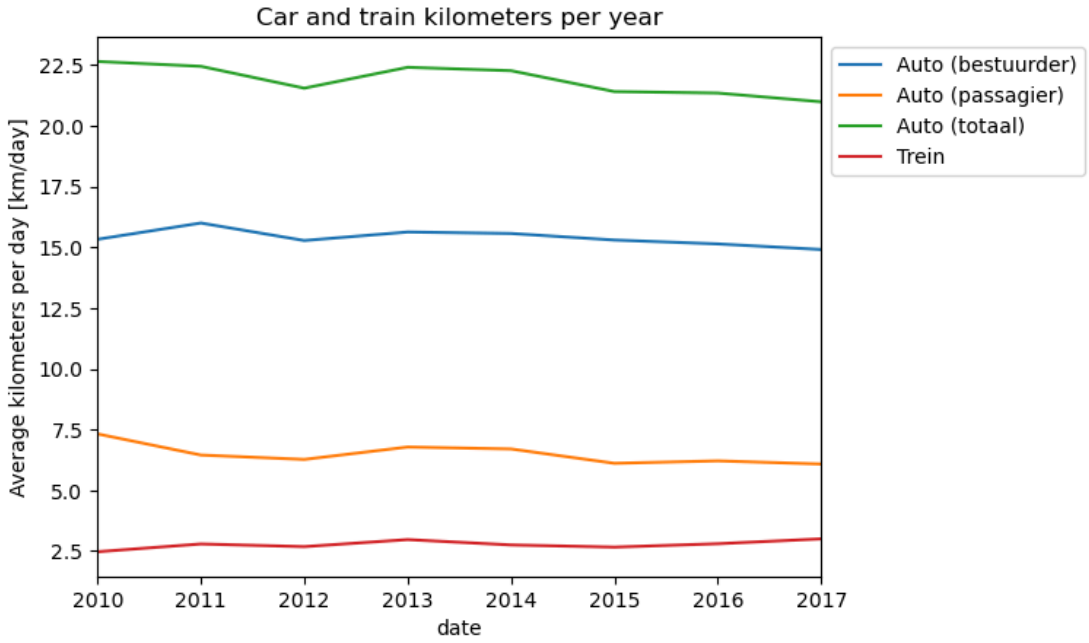

Figure 3: The amount of car and train kilometers per year in kilometers per day

From the graph, it can be seen that, on average, much more distance is covered by car than by train — about 20 km more per day per inhabitant of the Netherlands. The graph also shows that many cars are used only by the driver, as the number of kilometers driven by the driver is much higher than that by passengers. Over the years, the number of average kilometers traveled per mode appears relatively stable. A slight decline can be seen in the total kilometers traveled by car, and a small increase in train kilometers. The yearly graph might suggest stability, but when looking at the graph (see Figure 4) with monthly time steps, it becomes clear that the amount of kilometers traveled fluctuates.

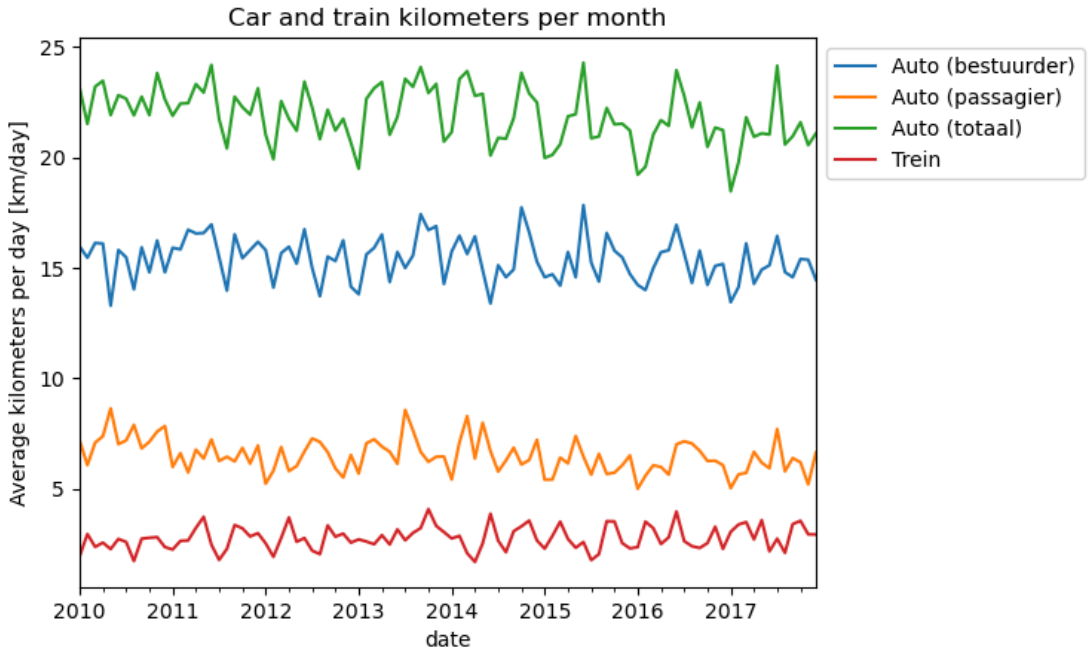

Figure 4: The amount of car and train kilometers per month in kilometers per day

These monthly fluctuations can be explained by holiday periods when people travel less or by seasonal effects, with people switching between car and train as the weather gets colder. However, as previously concluded, the total average kilometers traveled on a yearly basis remain relatively stable.

For the total kilometers driven it is also concluded that a much higher percentage of people travel by car. In Figure 5 it can be seen that the percentage of car users in this dataset is about 80 %. The biggest drops in total kilometers driven is observed in 2012 and 2015. Furthermore, if the percentage of car user drops it looks like the train user is following the opposite pattern, although it has to be said that the margins on these changes differ massively.

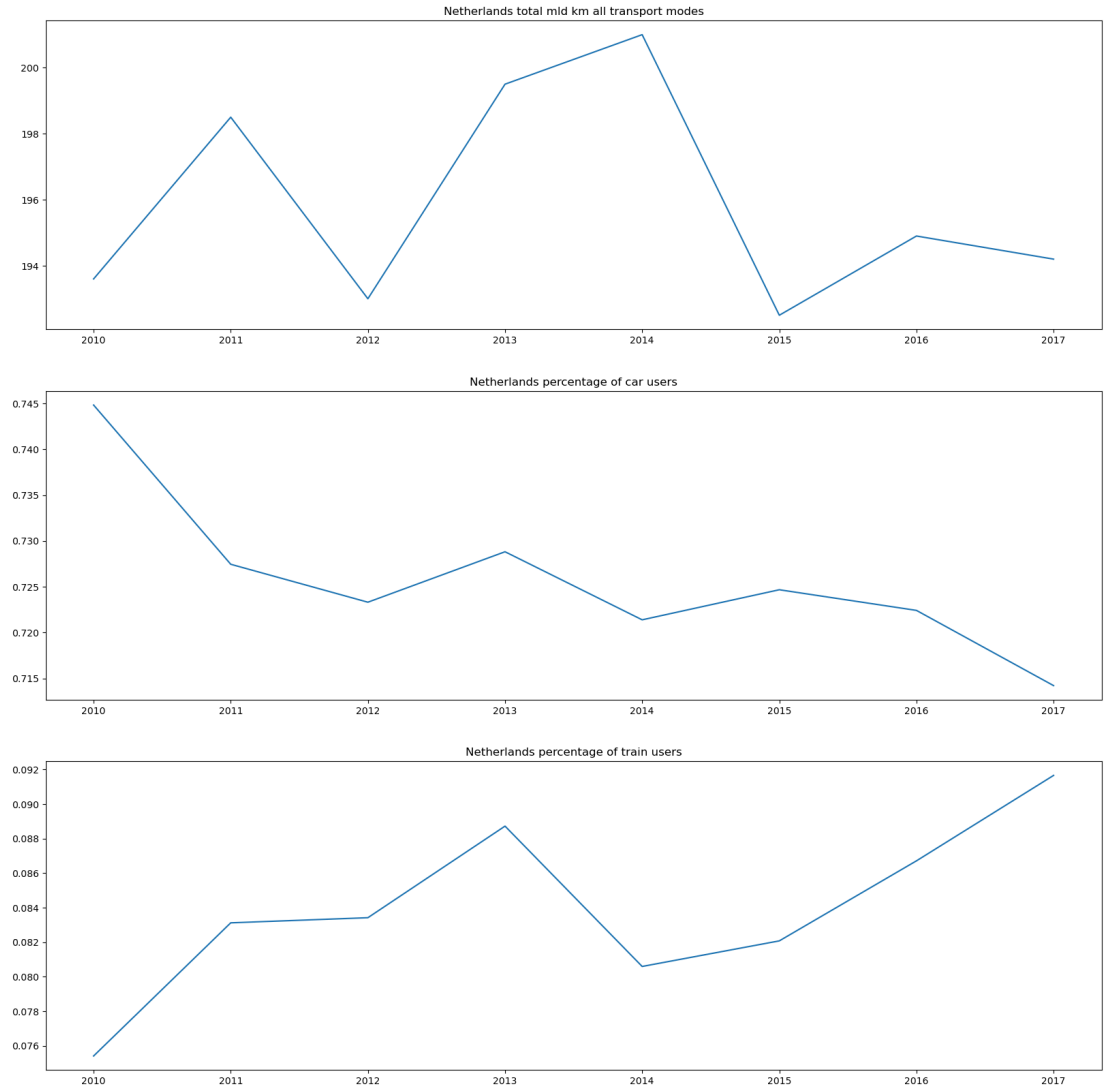

Figure 5: Total kilometers for all transport modes in the Nederlands and the percentage of train and car users between 2010 and 2017

Correlation between the pump price and travel distance of different modes between 2010 and 2017

To study the influence of fuel prices on the number of kilometers traveled, a graph was created showing the data points for the kilometers traveled per year alongside the fuel price for each year (see Figure 6). The result is shown in the graph, where the data points are scattered around the regression line. This suggests that there is little to no significant correlation between fuel price and kilometers traveled, likely due to the fact that, for most people, the car is an irreplaceable mode of transportation. When fuel prices rise, people continue to pay because demand for fuel is price inelastic.

Interestingly, the regression line indicates that as fuel prices increase, the number of kilometers traveled also increases. This can be attributed to the general upward trend in kilometers traveled each year, alongside rising fuel prices over the studied period.

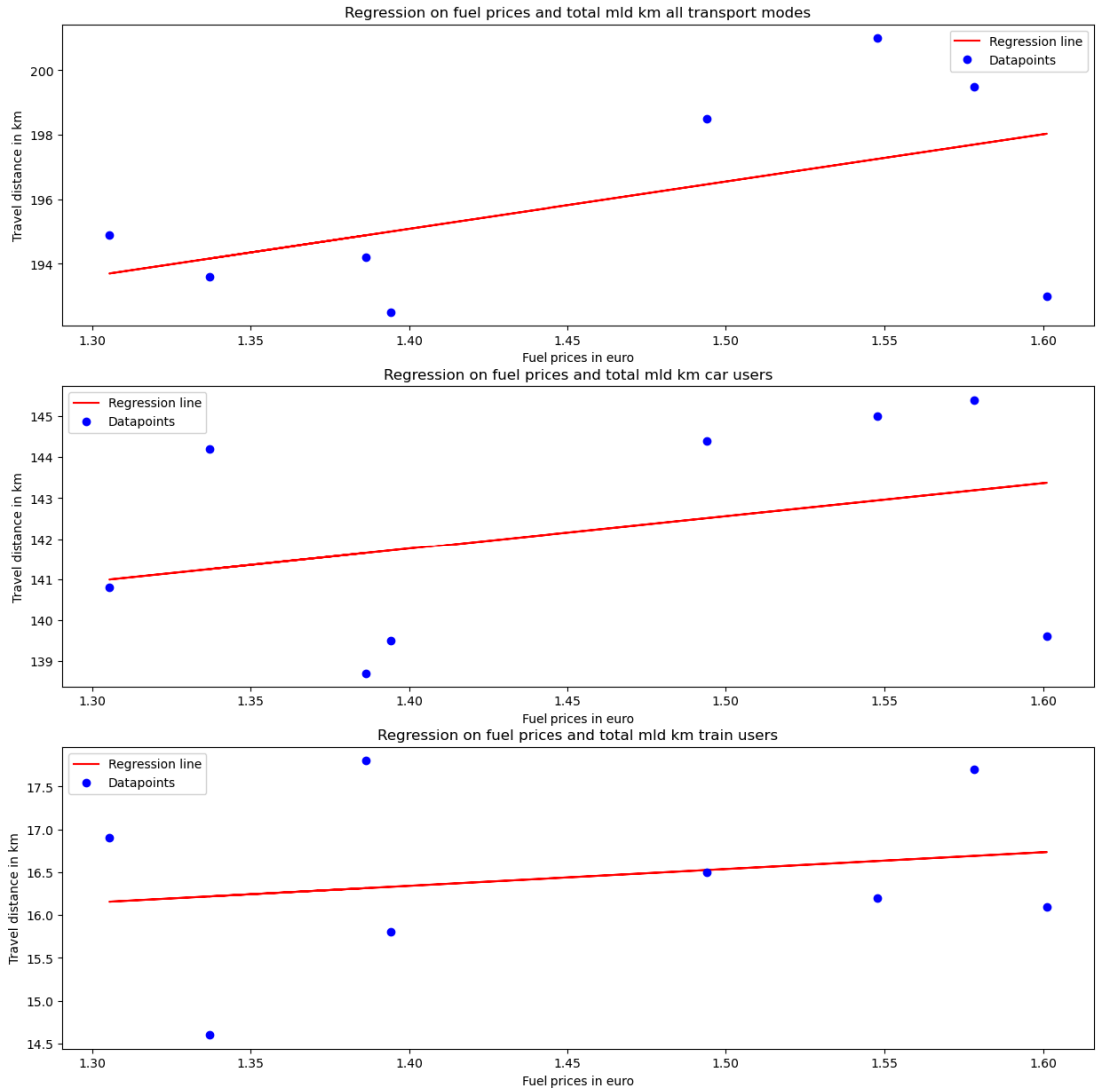

Figure 6: Correlation between pump price and different travel modes

What is the difference in the correlations between provinces with least and the most yearly travel distance?

To address this sub-question, it is first necessary to identify which province exhibits the highest and lowest average kilometers per person. For this purpose, the total annual kilometers per person for each province have been plotted in a graph, see Figure 7.

This visualization indicates that Limburg has the lowest average kilometers per person per year, while Flevoland has the highest average kilometers per person per year.

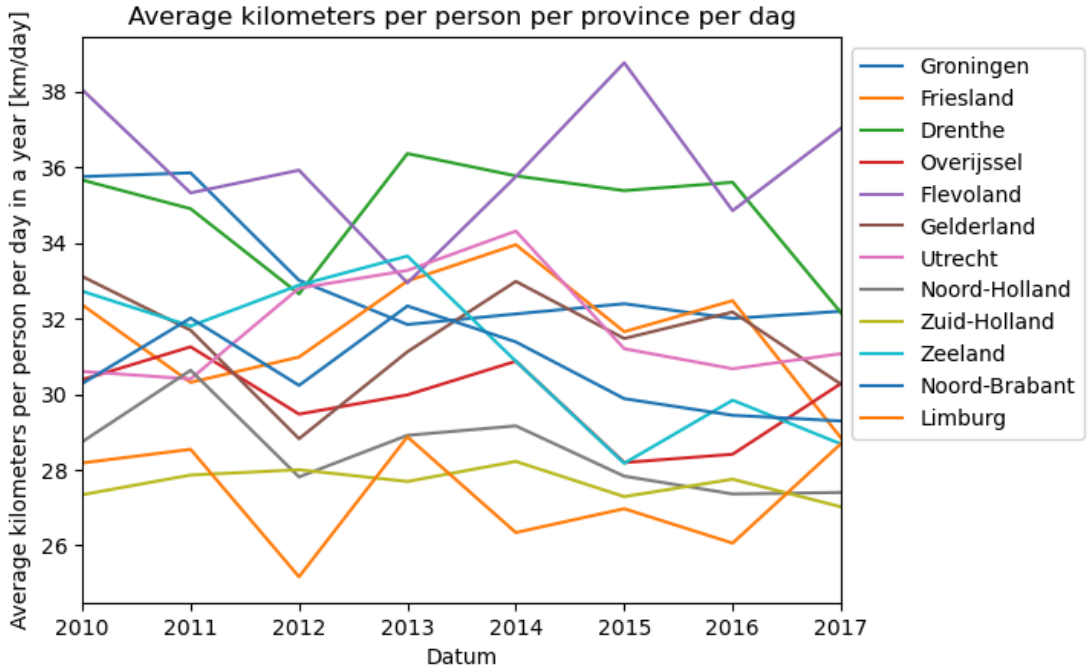

Figure 7: Average kilometers per person per province

Subsequently, the various modes of transportation—train, passenger car, and driver car—were examined. Unfortunately, the dataset was found to be incomplete, leading to the omission of several data points in the graph, see Figure 8.

When analyzing these two graphs, a slight decrease in the number of kilometers driven by car (driver) and a slight increase in train kilometers are observed in Flevoland. In contrast, Limburg exhibits a slight increase in the number of kilometers driven by car (driver), while train kilometers remain relatively stable. Moreover, there is little discernible difference in the number of kilometers traveled by car (passenger) between the two provinces.

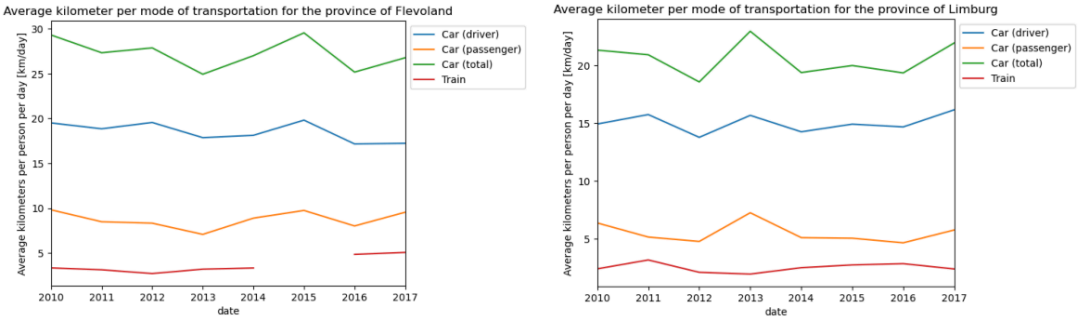

Figure 8: Kilometer per mode of transportation per person per day each year for the province of Flevoland (with the most travel distance per year) and Limburg (with the least travel distance per year)

Next, the correlations between fuel price and usage of car and train can be evaluated by performing linear regression and calculating the Pearson's r coefficient. The results are plotted in Figure 9.

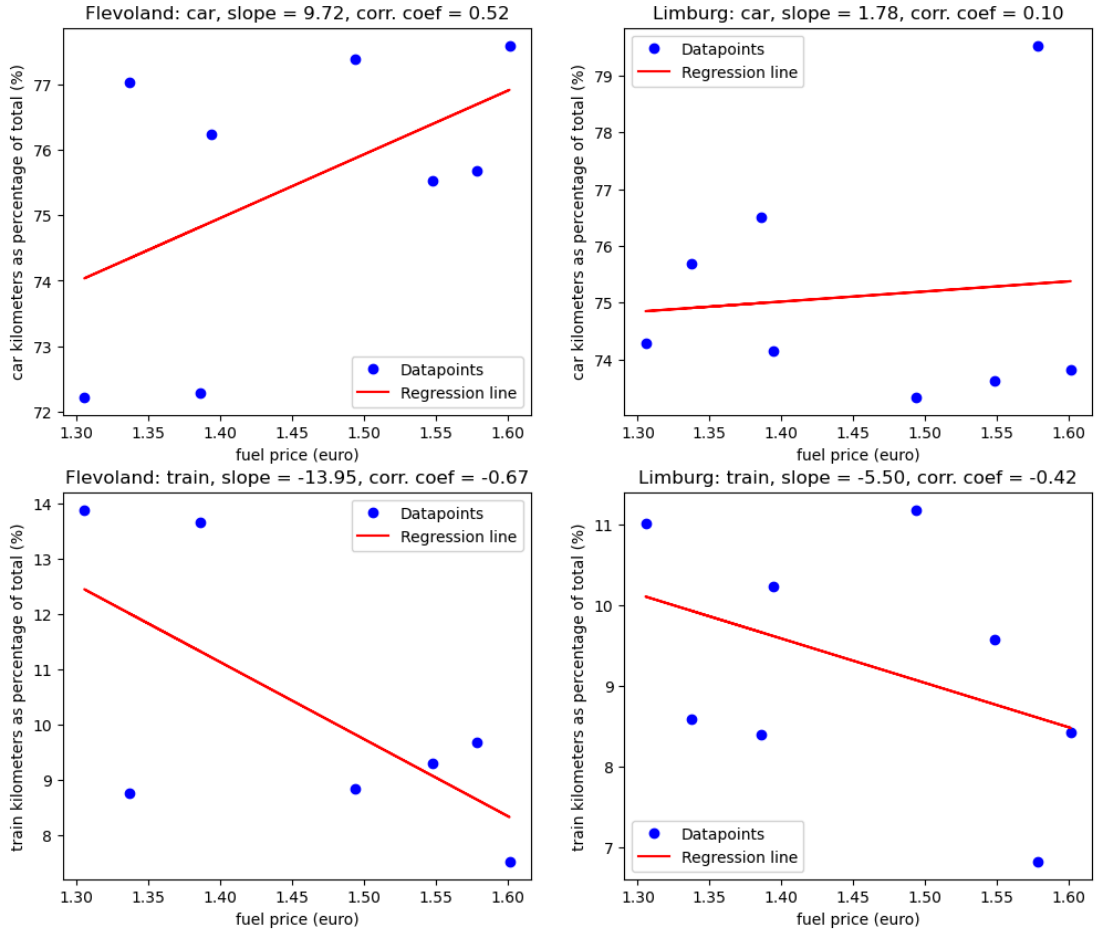

Figure 9: Correlation between fuel price and car or train for Flevoland and Limburg

In Figure 9 can be seen that for both provinces, fuel price correlates stronger with train usage than it does with car usage. Furthermore, the correlation coefficients for data in Flevoland are larger than those of Limburg. This indicates that a change of fuel price has a bigger effect in Flevoland, where the average travel distance is high, than in Limburg, where travel distance is low. Lastly, it is notable that train usage is negatively correlated, while car usage is positively correlated.

Conclusion

The prices for gasoline and diesel followed the same trend. During the studied time window (2010-2017), the highest prices were reached in 2012, while the lowest were in 2016. On average, people in the Netherlands traveled 20 kilometers by car per day per inhabitant, which is significantly more than the average distance traveled by train. Over the studied period, the number of kilometers traveled remained relatively stable; however, when looking at a more detailed graph, the kilometers vary more significantly due to seasonal factors such as holidays.

An analysis of the correlation between fuel prices and traveled distance reveals no significant correlation, suggesting that the demand for fuel is inelastic with respect to fuel price. When comparing the provinces with the highest and lowest number of traveled kilometers—Flevoland and Limburg, respectively—similar travel patterns emerge over the studied period. For both provinces, train usage has a negative correlation with

fuel price, while car usage has a positive correlation. The difference lies in the intensity of correlation, Flevoland saw a stronger correlation in both car usage and train usage compared to Limburg.

References

ANWB (n.d.). Accijnsverlaging brandstof deels teruggedraaid. https://www.anwb.nl/verkeer/nieuws/nederland/2023/juli/brandstofaccijns-omhoog-1-juli-2023

FiscaalRendement (2023). Accijnzen op benzine en diesel verder omhoog per 1 januari 2024. https://www.rendement.nl/reiskosten/nieuws/accijnzen-op-benzine-en-diesel-verder-omhoog-per-1-januari-2024.html

Compendium voor de Leefomgeving (2022). Aantal motorvoertuigen, 1990-2021. https://www.clo.nl/indicatoren/nl002624-aantal-motorvoertuigen-1990-2021

Knittel, C. R., & Tanaka, S. (2019). Driving behavior and the price of gasoline: Evidence from fueling-level micro data. Massachusetts Institute of Technology Center for Energy and Environmental Policy Research.

Schrier, W. (2008). The impact of rising fuel prices on automobilization (Master's thesis). University of Groningen, Netherlands. Retrieved from Rijksuniversiteit Groningen repository

CBS (2024.1). Pompprijze motorbrandstoffen; brandstofsoort, per dag. https://opendata.cbs.nl/statline/#/CBS/nl/dataset/80416ned/table?ts=1727428179572

CBS (2024.2). Personenmobiliteit; reiskenmerken en vervoerwijzen, regio's, 2010-2017 https://opendata.cbs.nl/#/CBS/nl/dataset/83498NED/table?ts=1727857444917

CBS (2024.3). Totale reizigerskilometers in Nederland per jaar; 2010-2014 https://opendata.cbs.nl/statline/#/CBS/nl/dataset/81126ned/table?ts=1727428018607

Appendix

extra answer for question 2: Another dataset was found with data from 1990 to 2018 with the total kilometers for motor vehicles. As this dataset only contains data for car usage it was not used to answer the main research question. However it was still considered relevant as this dataset has more datapoints. This dataset was plotted against time from 2006 to 2018 as this was the maximum time period that could be taken from the fuel prices dataset, see Figure 10. This dataset did not need a lot of data filtering as it was already ordened. The data was only selected for 'all motor vehicles' and for 'kilometers in the Netherlands'. Furthermore the data was used to plot the correlation between total travel distance of motor vehicles and the yearly fuel price, see figure 11.

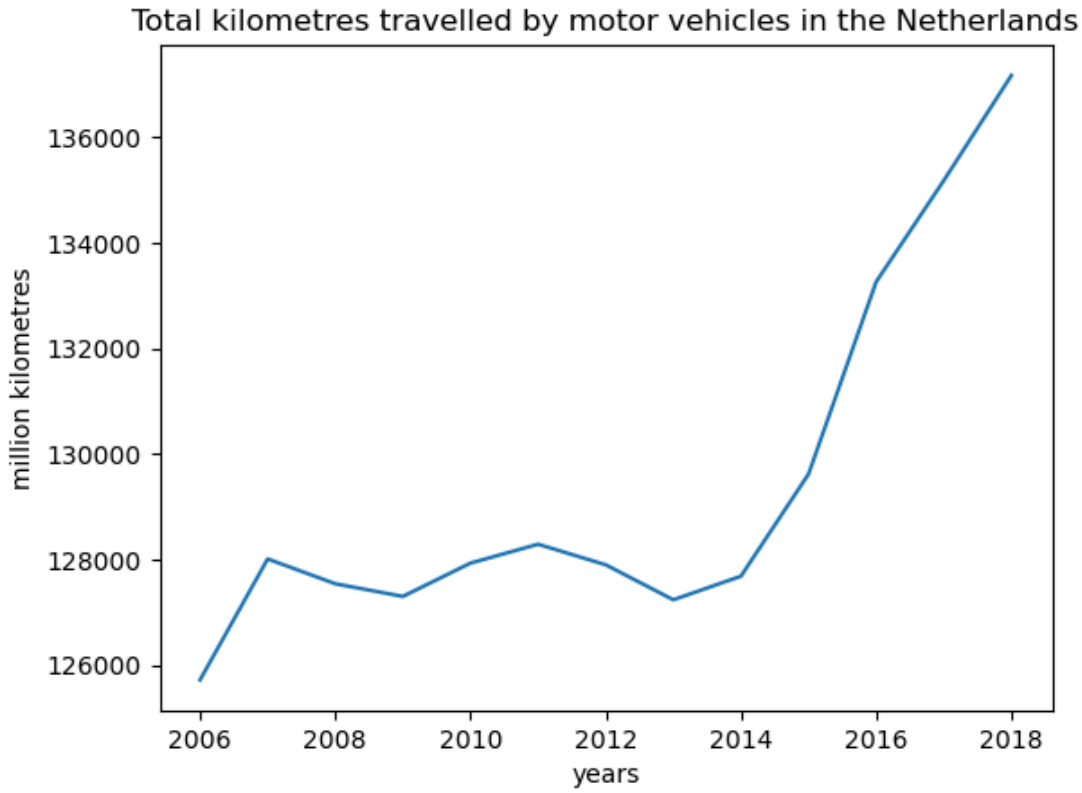

Figure 10: Total kilometers driven for motor vehicles against time

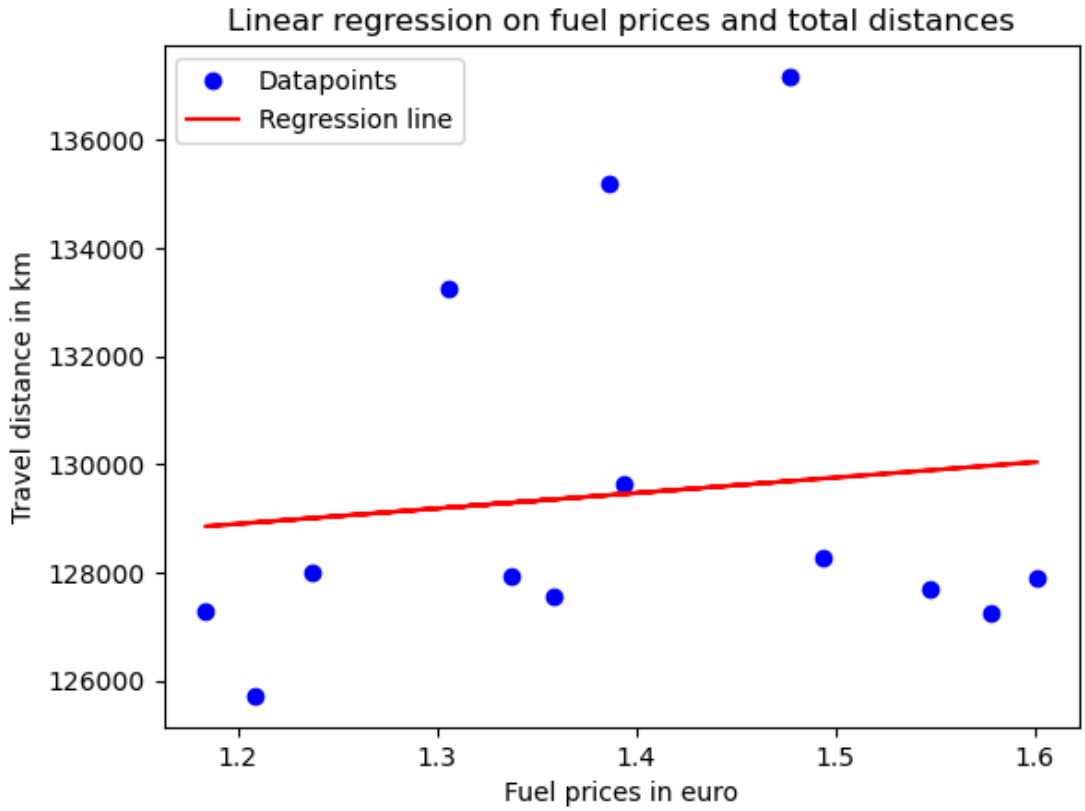

Figure 11: Linear regression of the fuel prices and the total distances