**Project01- Team 3**

**“Gas Prices per States”**

**Hypothesis: Location affects gas prices in USA states.**

We know gasoline prices varies widely across all different US states. But we want to know, what influences this?

We will look into the relationship between gasoline prices across the country and the different variables that we believe may affect it such as locations, GDP per state, number of vehicles, historical prices and distance from refineries. We will be focusing on getting insights and discover patterns that might be relevant towards determining gasoline prices across the USA.

For our data exploration & cleanup we follow the next steps:

1. Obtain Data:
   1. Requests for Gas Price API and EIA API.
   2. Download csv formats of GDP per state, total cars per state and list of oil refineries.
   3. Google Geolocation API for latitude & longitude per state.
   4. Google Geolocation API to retrieve latitude & longitude of refineries.
2. Treat Data:
   1. Reduce EIA API gas price time series to 52 weeks.
   2. Calculate distance between refinery and the center of state.
   3. Create data frames for each source of information.
   4. Remove data NaN information and other irrelevant information.
   5. Remove outliers in the analysis.
3. Merge Data:
   1. Merge all different data sources into one single data frame.

As we evaluate the gas prices per state we were focus mainly on the following questions:

* **Does the price change between locations (North/South/East/West)?**

A screen shot of a computer

Description automatically generated

On our first analysis using scatter plots we saw 4 outliners states: Hawaii, Alaska, Washington, California, where prices were higher than average. After this first analysis we divide into east and west graphs and come upon a relation between state longitude and prices where states on the east part have a lower average price on gas than states on west.

A picture containing indoor, table, computer, desk

Description automatically generatedA picture containing indoor, table, computer, photo

Description automatically generated

* **Does the price of gas change in relation to the proximity of a refinery?**

When you considered all the refineries the average distance directly affect the price of gas; as the average distance increase the price is higher. In the min distance graph you cannot get any conclusions as they are some states like Texas that will have more refineries and the distance will be lower.

A close up of a map

Description automatically generated

A close up of a computer

Description automatically generatedA picture containing table, computer, photo, laptop

Description automatically generatedr=0.4747 r=0.0176

* **Is the price of gas on each state influenced by its GDP?**

Our regression analysis shows that our average R-square value in all gas types is closer to 0 so there´s a negative relation between these 2 variables.

r=0.1347 r=0.1230

A picture containing photo, table, computer, desk

Description automatically generatedA picture containing table, photo, computer, desk

Description automatically generated

r=0.1250 r=0.1638

A picture containing table, photo, computer, desk

Description automatically generatedA picture containing photo, table, computer, white

Description automatically generated

* **Is the price of gas of each state influenced by how many cars each state has?**

The gas price is not influenced by the number of vehicles on each state. Our r-square value is close to 0.

A picture containing computer, table

Description automatically generated

* **How does gas price vary across 1 year in each state?**

For this analysis we used the prices per week for the last year.

The trend we see is that the gas prices decreased between April-May 2020. This effect is related to the drop of global demand due to COVID. The OPEC (Organization of the Petroleum Exporting Countries-Plus) held meetings in response to the drop of global demand and price oil expecting to cut in oil productions but we can see the trend on lower prices continues for several more weeks.

A close up of a map

Description automatically generated

Conclusions:

~~For our analysis we conclude the location of the states may affect the price of the gas. As a result on the scatter plot graphs , east states have a lower average price than the states on west. There is also a relation between the distance from refineries and the price of the gas per state.~~

~~On the different scenarios we run with these variables we saw mainly 4 states as outliers: Hawai, Alaska, California & DC. According to the NASC (National Advancing Convenience & Fuel Retailing) they are 3 main reasons on why the average price is higher in those states, this is mainly driven by: taxes, fuel blends, and margins. California requires a specific fuel blend that impact directly on their high prices. The state faces unique environmental conditions that have led state regulators to require a cleaner burning gasoline to improve air quality.~~

~~Las dos que influencias mas son la variación en longitud entre east y west y la distancia promedio entre el centro de los estados y la refinería, las demás variables no dieron una~~

~~Una continuación seria bajar hacia ciudades para ver cuales~~

Data Source:

* RapidApi Gas Price 🡪 <https://rapidapi.com/collectapi/api/gas-price?endpoint=apiendpoint_dd843904-8bb0-413b-8610-a5eb9ebba72e>.
* Google Maps API 🡪 <https://cloud.google.com/maps-platform?hl=es>
* EIA (US Energy Information Administration) 🡪<https://www.eia.gov/>.
* BEA (US Bureau of Economic Analysis 🡪 <https://www.bea.gov/data/gdp/gdp-state>.
* Office of Highway Policy Information🡪<https://www.fhwa.dot.gov/policyinformation/>
* <https://www.convenience.org/Topics/Fuels/Why-Do-Prices-Vary-Around-the-Country-or-Around-th>