Car Challenge

The impact of the pandemic in the automotive industry in Mexico

Team 6

Member's background

Diana Baltazar
Secretaría de Gobernación

Marco Cid
Public Accounting and Financial Strategy

Perla Martinez
GM Financial

Freddy Reyes
Electrical Engineer

Tania Rosas
Oil & Gas

Javier Sada
Start up: Robots

Motivation & Summary

Base on the data analysis, we want to demonstrate the effect of COVID-19 in the Mexican car industry.

Questions

Does the pandemic reduce the number of car sales in Mexico?

Did people buy cars during the pandemic to avoid public transportation?

What are the car brands that sell more?

What was the actual car sales in 2020?

What was the sale forecast for 2020 without Covid?

Does the stock price of the top brands increase or decrease?

Datasets

- **INEGI 2005-2021**

  Registro administrativo de la industria automotriz de vehículos ligeros (nuevos y seminuevos)

- **Forecast on INEGI dataset**

- **ALPHA VANTAGE**

  Free Stock Market API

Does the pandemic reduce the number of car sales in Mexico?

Did people buy cars during the pandemic to avoid public transportation?

What are the car brands that sell more?

What was the actual car sales in 2020?

Data Cleanup & Exploration

- **INEGI 2009-2021**
  1. Select range 2009-2019
  2. Excel export from INEGI
  3. Consolidation into pandas

1. Cleaning and Understanding the Data

Reading Data
Reading the dictionaries

In [3]:
```
import pandas as pd
import numpy as np
```

In [4]:
```
#Reading the Data
data_2009 = pd.read_csv('../raiavl_venta_mensual_2005_2021/conjunto_de_datos/raiavl_venta_mensual_tr_cifra_2009.csv')
data_2010 = pd.read_csv('../raiavl_venta_mensual_2005_2021/conjunto_de_datos/raiavl_venta_mensual_tr_cifra_2010.csv')
data_2011 = pd.read_csv('../raiavl_venta_mensual_2005_2021/conjunto_de_datos/raiavl_venta_mensual_tr_cifra_2011.csv')
data_2012 = pd.read_csv('../raiavl_venta_mensual_2005_2021/conjunto_de_datos/raiavl_venta_mensual_tr_cifra_2012.csv')
data_2013 = pd.read_csv('../raiavl_venta_mensual_2005_2021/conjunto_de_datos/raiavl_venta_mensual_tr_cifra_2013.csv')
data_2014 = pd.read_csv('../raiavl_venta_mensual_2005_2021/conjunto_de_datos/raiavl_venta_mensual_tr_cifra_2014.csv')
data_2015 = pd.read_csv('../raiavl_venta_mensual_2005_2021/conjunto_de_datos/raiavl_venta_mensual_tr_cifra_2015.csv')
data_2016 = pd.read_csv('../raiavl_venta_mensual_2005_2021/conjunto_de_datos/raiavl_venta_mensual_tr_cifra_2016.csv')
data_2017 = pd.read_csv('../raiavl_venta_mensual_2005_2021/conjunto_de_datos/raiavl_venta_mensual_tr_cifra_2017.csv')
data_2018 = pd.read_csv('../raiavl_venta_mensual_2005_2021/conjunto_de_datos/raiavl_venta_mensual_tr_cifra_2018.csv')
data_2019 = pd.read_csv('../raiavl_venta_mensual_2005_2021/conjunto_de_datos/raiavl_venta_mensual_tr_cifra_2019.csv')
```

In [8]:
```
#Create de Dataframe
sales_df = pd.concat([data_2009,data_2010,data_2011,data_2012,data_2013,data_2014,data_2015,data_2016,
                      data_2017,data_2018,data_2019], ignore_index=True)
sales_df.head()
```

Out[8]:

	PROD_EST	COBERTURA	ANIO	ID_MES	MARCA	MODELO	TIPO	SEGMENTO	ORIGEN	ID_PAIS_ORIGEN	UNI_VEH
0	Registro Administrativo de la Industria Automo...	Nacional	2009	1	Acura	Mdx	Camiones ligeros	SUV's	IMPORTADO	40	29
1	Registro Administrativo de la Industria Automo...	Nacional	2009	1	Acura	Rdx	Camiones ligeros	SUV's	IMPORTADO	66	34

In [5]:
```
#Export to a new csv
sales_df.to_csv('../Excel/Sales Database.csv', index=False)
```

Data Cleanup & Exploration

- **INEGI 2009-2021**
  1. Select range 2009-2019
  2. Excel export from INEGI
  3. Consolidation into pandas
  4. Cleaning process
  5. Preparing data

2. Preparing the Data

Change the column's names
Check for missing data
Change the fields with the matching values in the dictionaries
Create the final base

In [53]:
```
#Rename Columns to match the names of the data dictionary
sales_df=sales_df.rename(columns={"PROD_EST":"Proyecto","COBERTURA":"Area_Geografica","ANIO":"Año_Referencia",
                                  "ID_MES":"Mes","MARCA":"Marca","MODELO":"Modelo","TIPO":"Clasificación",
                                  "SEGMENTO":"Segmento","ORIGEN":"Origen","ID_PAIS_ORIGEN":"Pais_Origen",
                                  "UNI_VEH":"Numero_Unidades"})
sales_df.head()
```

Out[53]:

	Proyecto	Area_Geografica	Año_Referencia	Mes	Marca	Modelo	Clasificación	Segmento	Origen	Pais_Origen	Numero_Unidades
0	Registro Administrativo de la Industria Automo...	Nacional	2009	1	Acura	Mdx	Camiones ligeros	SUV's	IMPORTADO	40	29
1	Registro Administrativo de la Industria Automo...	Nacional	2009	1	Acura	Rdx	Camiones ligeros	SUV's	IMPORTADO	66	34
2	Registro Administrativo de la Industria Automo...	Nacional	2009	1	Acura	Rl	Automóviles	De Lujo	IMPORTADO	126	2
3	Registro Administrativo de la Industria Automo...	Nacional	2009	1	Acura	Tl 3.5l	Automóviles	De Lujo	IMPORTADO	66	35
4	Registro Administrativo de la Industria Automo...	Nacional	2009	1	Acura	TSX 2.4L	Automóviles	De Lujo	IMPORTADO	40	49

Data Cleanup & Exploration

- **INEGI 2009-2021**
  1. Select range 2009-2019
  2. Excel export from INEGI
  3. Consolidation into pandas
  4. Cleaning process
  5. Preparing data

```
In [63]: #Merge the dictionary with the Main DataFrame to complete the missing values
         sales_final_df=pd.merge(sales_df,pais_df, on="Pais_Origen", how="left")

In [64]: sales_final_df.head()
```

Out[64]:

	Proyecto	Area_Geografica	Año_Referencia	Mes	Marca	Modelo	Clasificación	Segmento	Origen	Pais_Origen	Numero_Unidades	DESC
0	Registro Administrativo de la Industria Automo...	Nacional	2009	1	Acura	Mdx	Camiones ligeros	SUV's	IMPORTADO	40	29	
1	Registro Administrativo de la Industria Automo...	Nacional	2009	1	Acura	Rdx	Camiones ligeros	SUV's	IMPORTADO	66	34	
2	Registro Administrativo de la Industria Automo...	Nacional	2009	1	Acura	Rl	Automóviles	De Lujo	IMPORTADO	126	2	
3	Registro Administrativo de la Industria Automo...	Nacional	2009	1	Acura	Tl 3.5l	Automóviles	De Lujo	IMPORTADO	66	35	

...

```
In [78]: Final_Clean_Sales.head()
```

Out[78]:

	Area_Geografica	Año_Referencia	Marca	Modelo	Clasificación	Segmento	Origen	Numero_Unidades	Pais_Origen	Mes
0	Nacional	2009	Acura	Mdx	Camiones ligeros	SUV's	IMPORTADO	29	Canadá	Enero
1	Nacional	2009	Acura	Rdx	Camiones ligeros	SUV's	IMPORTADO	34	Estados Unidos	Enero
2	Nacional	2009	Acura	Rl	Automóviles	De Lujo	IMPORTADO	2	Japón	Enero
3	Nacional	2009	Acura	Tl 3.5l	Automóviles	De Lujo	IMPORTADO	35	Estados Unidos	Enero
4	Nacional	2009	Acura	TSX 2.4L	Automóviles	De Lujo	IMPORTADO	49	Canadá	Enero

...

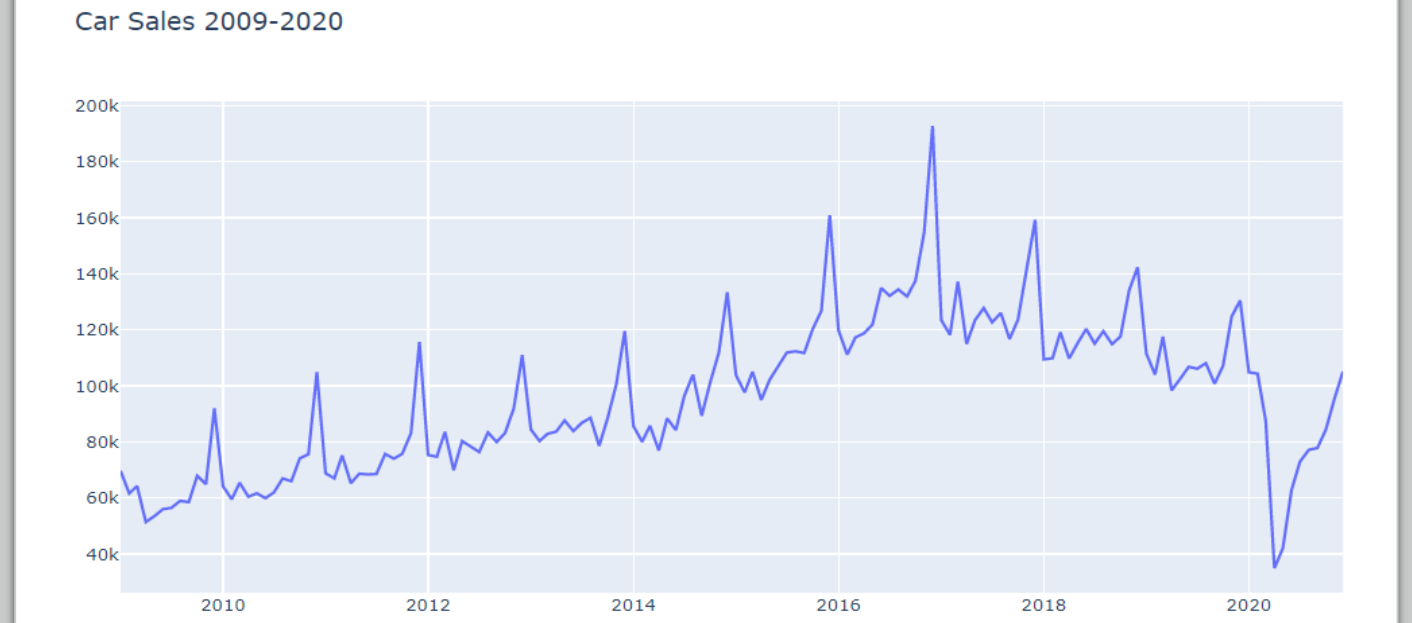
```
In [81]: Final_Clean_Sales['Marca'].astype('category').value_counts()

Out[81]: General Motors    6296
         Ford Motor        3897
         Nissan            3852
         Chrysler          3687
         Volkswagen        2838
         Mercedes Benz     2633
         Toyota            2340
         BMW               2311
         Audi              2073
         Honda             1718
```

Data Cleanup & Exploration

- **INEGI 2009-2021**
  1. Select range 2009-2019
  2. Excel export from INEGI
  3. Consolidation into pandas
  4. Cleaning process
  5. Preparing data
  6. Data exploration

- Trend shows sales increase every year with a seasonal drop each January
- 2017 drop down sales due to T-MEC

```
In [67]: fig=go.Figure(data = [go.Scatter(x=time_series_df.index, y=time_series_df['Units'])],layout_title_text="Car Sales 2009-2020")

         fig.show()
         fig.write_image("Graphs/CarSales09-20.jpeg")
```



Data Cleanup & Exploration

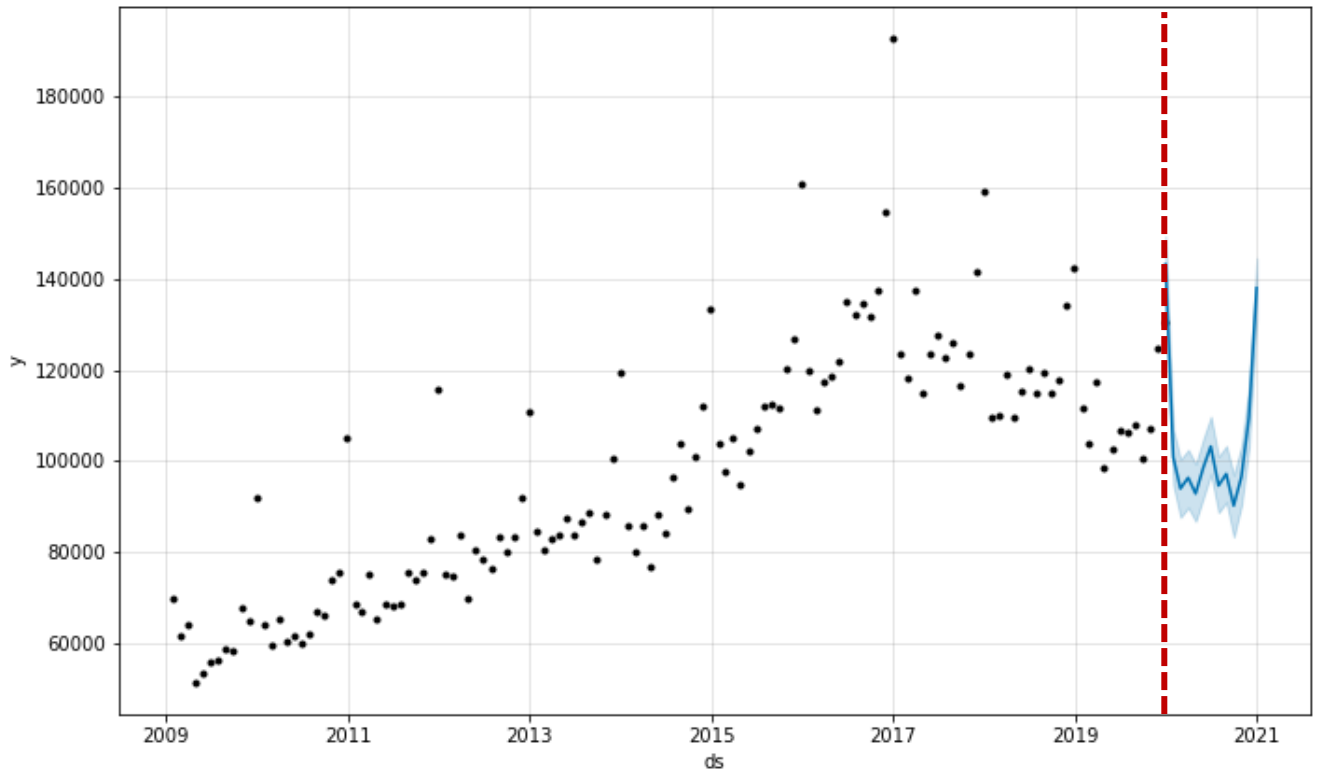
- **INEGI 2009-2019**

1. Forecast INEGI dataset for 2020 not considering actual 2020 sales

PROPHET

2. Prophet is an open-source library developed by Facebook and designed for automatic forecasting of univariate time series data.

```
In [5]: from fbprophet import Prophet

        mgm = Prophet()
        mgm.fit(gm)

        future_s = []

        for i in range(1,13):
            date = '2019-%02d' % i
            future_s.append(date)
        future_s = pd.DataFrame(future_s)
        future_s.columns = ['ds']
        future_s['ds'] = pd.to_datetime(future_s['ds'])

        fcstgm = mgm.predict(future_s)

        mgm.plot(fcstgm).savefig('../Graphs/07 GM car sales sample fcst.png')
        plt.show()
```



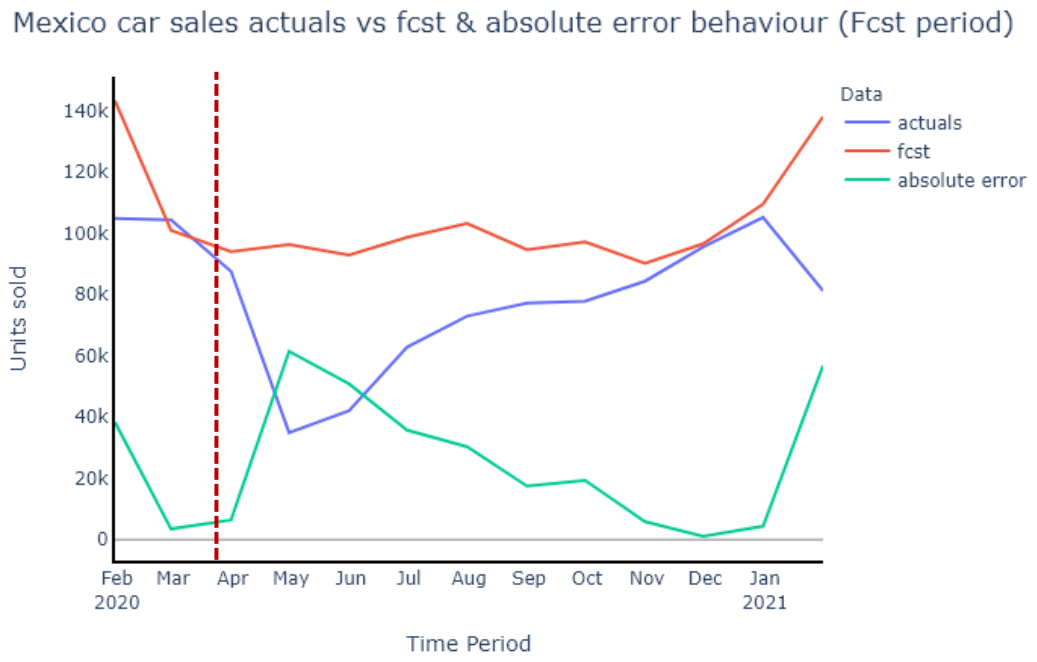
Data Cleanup & Exploration

- **Forecast 2020**
  1. Load & summarize dataset
  2. Load & plot dataset
  3. Forecast car sales 2020 with prophet
     - Fit Model
  4. Make in-sample forecast
  5. Make out-of-sample forecast

Insights: The biggest difference is shown April, after lockdown

Problems: Due to the limited number of variables used to train the model, this forecast shows an error of 25,000 units

In [7]:
```
from sklearn.metrics import mean_absolute_error

fcst = forecast[['ds','yhat']].resample('M', on='ds').sum().reset_index().merge(actuals, on='ds', how='left')
fcst = fcst[['ds','y','yhat']].rename(columns={'y':'actuals','yhat':'fcst'})
fcst.loc[:,'absolute error'] = abs(fcst.loc[:,'actuals'] - fcst.loc[:,'fcst'])

y_true = fcst['actuals'].values
y_fcst = fcst['fcst'].values

mae = mean_absolute_error(y_true,y_fcst)

print('The mean average error is of %.0f units' % mae)

avf = px.line(fcst, x='ds', y=fcst.columns, title='Mexico car sales actuals vs fcst & absolute error behaviour (Fcst period)',
              labels={
                  'variable':'Data',
                  'ds':'Time Period',
                  'value':'Units sold'
              })
avf.update_layout(
    xaxis=dict(dtick='M1', tickformat='%b\n%Y', showline=True, linewidth=2, linecolor='black'),
    yaxis=dict(showline=True, linewidth=2, linecolor='black', zeroline=True, zerolinewidth=1, zerolinecolor='grey'),
    plot_bgcolor='white'
)
avf.write_image('../Graphs/04 Mexico car sales v fcst in fcstp.png')
avf.write_html('../Graphs/04 Mexico car sales v fcst in fcstp.html')
avf.show()
```



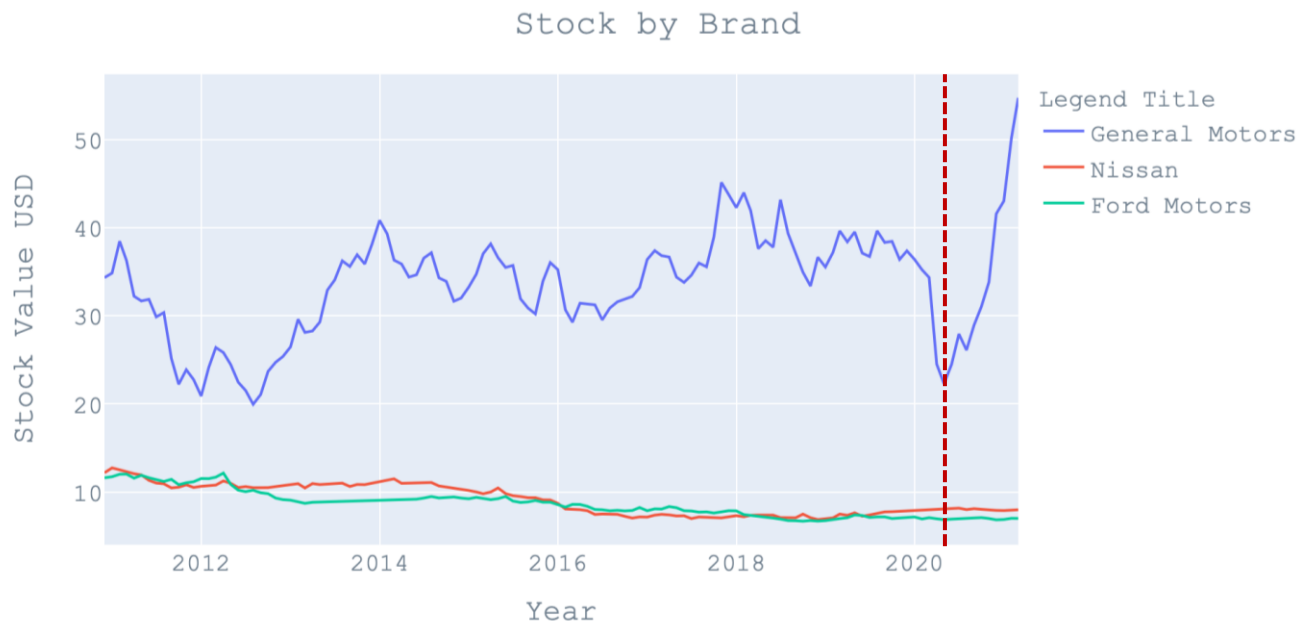
Data Cleanup & Exploration

- **ALPHA VANTAGE**
  Free Stock Market API

The ticker names were searched manually to be consulted using the API

Ticker	Company Name
GM	General Motors Company
Ford	Ford Motor Company
NSANY	Nissan Motor Co., Ltd. (ADR)

```
In [22]: fig = go.Figure([go.Scatter(x=gm_df["date"], y = gm_df["high"])])
fig.update_layout(
    title="Time VS Stock",
    xaxis_title="Year",
    yaxis_title="Stock Value USD",
    legend_title="Legend Title",
    font=dict(
        family="Courier New, monospace",
        size=18,
        color="slategray"
    )
)
fig.update_layout(
    title={
        'text': "Time VS Stock",
        'y':0.9,
        'x':0.5,
        'xanchor': 'center',
        'yanchor': 'top'})

fig.show()
```



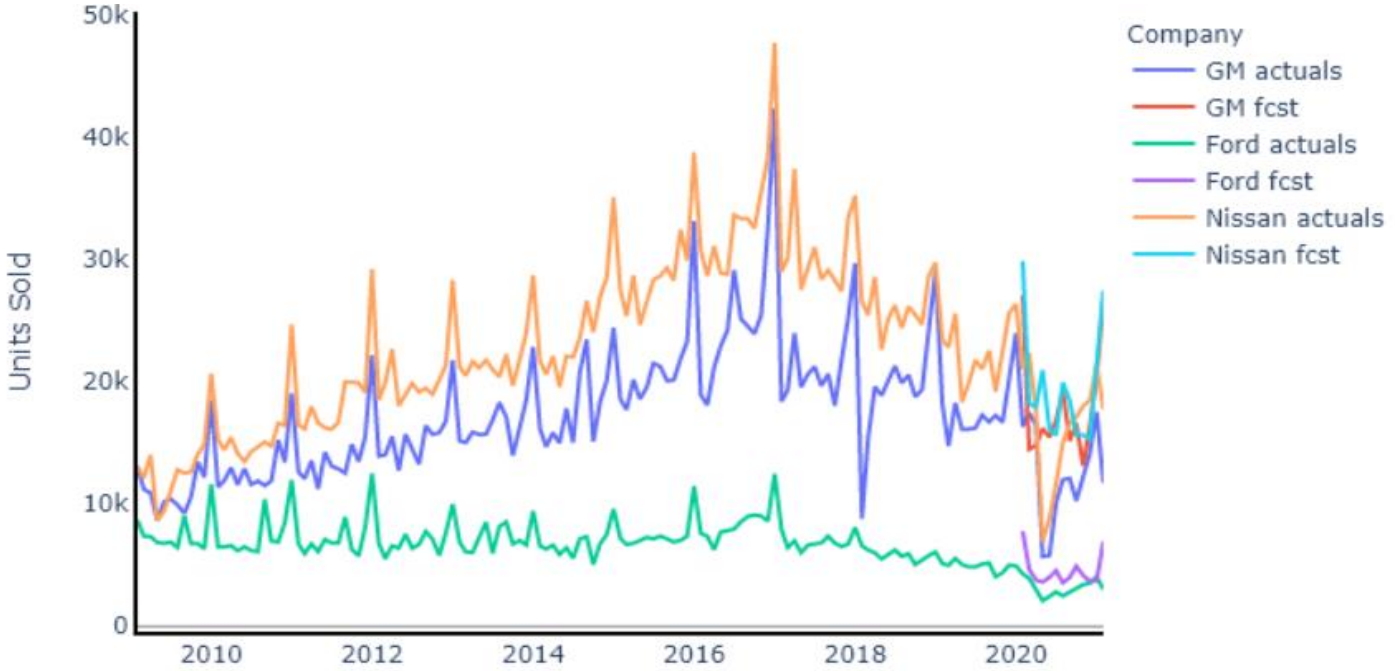
Data Analysis

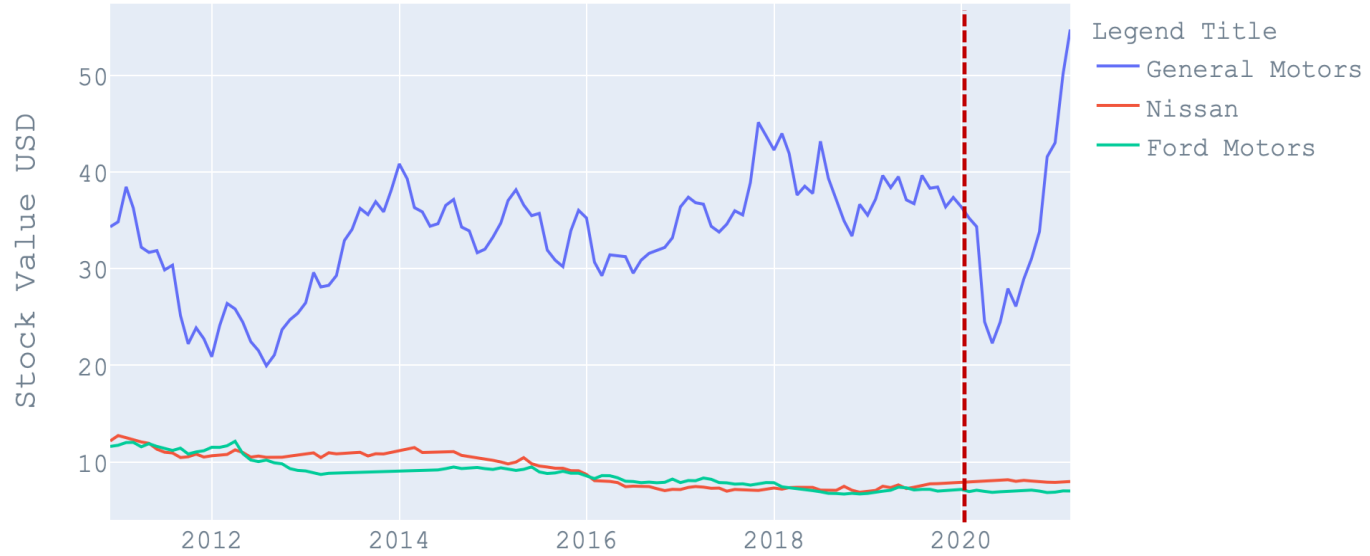
Does the pandemic reduce the number of car sales in Mexico?

Did people buy cars during the pandemic to avoid public transportation?

What are the car brands that sell more?

What was the actual car sales in 2020?





Discussion

- Car Sales Drop

  Macroeconomic uncertainty fueled by COVID-19's global spread and supply chain disruptions resulted in a significant sales drop with a mean absolute error of 25,000 units in 2020.

- Stock Price

  Fiscal support was injected in the stock exchange as part of the crisis management response from the US government and we see the value of the tickers increased during 2020.

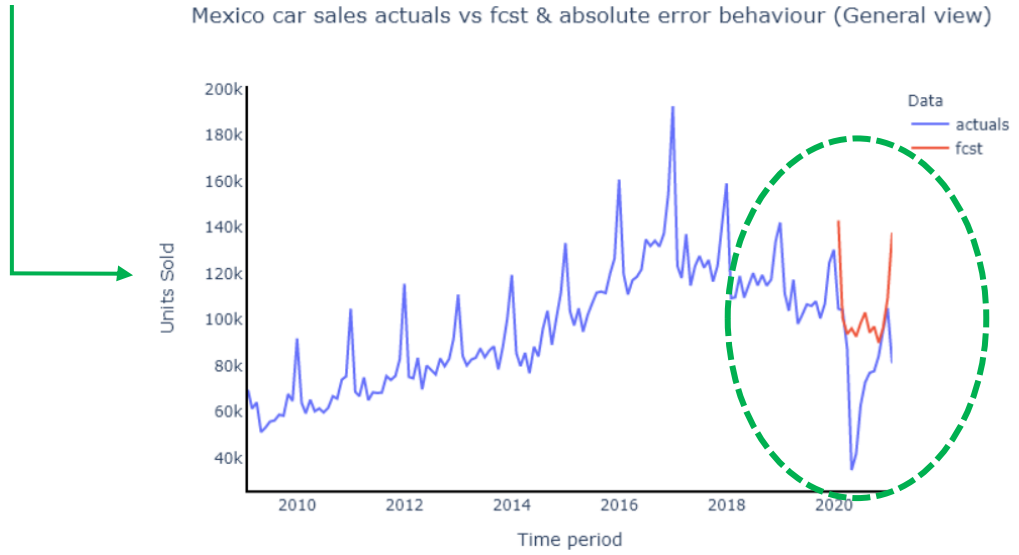
  Using the stock price was not a clear support of our hypothesis.

Postmortem

- Semáforo de movilidad (Mexico)

  https://datos.cdmx.gob.mx/dataset/afluencia-preliminar-en-transporte-publico/resource/5d33f9c7-e033-4676-a02d-9e2129017acf

  SEMOVI - API



Questions?