Exercise 6.1

Jack Bartman

**Data Source:**

Two datasets were acquired for analysis and a third for annotation.

Trips\_by\_Distance.csv was sourced from Data.Transportation.Gov. It was prepared by the Bureau of Transportation Statistics by the Maryland Transportation Institute. It contains a historical record of people who took no trips, vs people who took trips, and the distances of those trips by miles traveled.

Motor\_Vehicle\_Registrations\_Dashboard\_data.csv was sourced from Data.Transportation.Gov. It was aggregated by the Federal Highway Administration, and maintained by the FHWA\_Datahub. It contains a historical record of vehicles registered by state.

State Covid Restrictions is a spreadsheet developed by the National Academy for State Health Policy. It documents which states had official travel restriction states of emergency during the Covid -19 pandemic, and when those restrictions started and ended.

**Data Collection:**

Trips\_by\_Distance.csv was produced from an anonymized national panel of mobile device data from multiple sources. All data sources used in the creation of the metrics contain no personal information. Data analysis is conducted at the aggregate national, state, and county levels. A weighting procedure expands the sample of millions of mobile devices, so the results are representative of the entire population in a nation, state, or county. To assure confidentiality and support data quality, no data are reported for a county if it has fewer than 50 devices in the sample on any given day.

Trips are defined as movements that include a stay of longer than 10 minutes at an anonymized location away from home. Home locations are imputed on a weekly basis. A movement with multiple stays of longer than 10 minutes before returning home is counted as multiple trips.

Motor\_Vehicle\_Registrations\_Dashboard\_data.csv was aggregated by the US Department of Transportation from reports submitted by each individual state and territory.

State Covid Restrictions were aggregated from ["See Which States Are Reopening and Which Are Still Shut Down"](https://www.nytimes.com/interactive/2020/us/states-reopen-map-coronavirus.html). *The New York Times*. [Archived](https://web.archive.org/web/20210130005858/https:/www.nytimes.com/interactive/2020/us/states-reopen-map-coronavirus.html) from the original on January 30, 2021and  ["This is where all 50 states stand on reopening"](https://www.cnn.com/interactive/2020/us/states-reopen-coronavirus-trnd/). *CNN*. [Archived](https://web.archive.org/web/20210210220832/https:/www.cnn.com/interactive/2020/us/states-reopen-coronavirus-trnd/) from the original on February 10, 2021.

**Data Contents:**

Trips by Distance

* Level
  + National or State
* State FIPs
  + State Federal Information Processing code identification number
* State Postal Code
  + State Abbreviation
* County FIPs
  + County Federal Information Processing code identification number
* County Name
* Population Staying at Home
* Population Not Staying at Home
* Number of Trips
  + Total number of Trips
* Number of Trips <1
  + Number of trips less than 1 mile
* Number of Trips 1-3
  + Number of trips 1 to 3 miles
* Number of Trips 3-5
  + Number of trips 3 to 5 miles
* Number of Trips 5-10
  + Number of trips 5 to 10 miles
* Number of Trips 10-25
  + Number of trips 10 to 25 miles
* Number of Trips 25-50
  + Number of trips 25 to 50 miles
* Number of Trips 50-100
  + Number of trips 50 to 100 miles
* Number of Trips 100-250
  + Number of trips 100 to 250 miles
* Number of Trips 250-500
  + Number of trips 250 to 500 miles
* Number of Trips >=500
  + Number of trips 500 or more miles
* Row ID
  + Unique row identifier
* Week
  + Numeric week of the month
* Month
  + Numeric Month of the year

Motor Vehicle Registrations Dashboard Data

* Year
* State
* Auto
  + Number of autos registered
* Truck
  + Number of trucks registered
* Motorcycles
  + Number of motorcycles registered

State Covid Restrictions

* State
  + Location of restriction
* Date\_enacted
  + Beginning of travel related state of emergency
* Date\_listed
  + Ending of travel related state of emergency

**Why this data:**

This data fulfilled the academic requirements of having a time and geographic variable set. I provided an opportunity to explore transportation-related data with the tool explored in CareerFoundry.

**Data Cleaning:**

Trips by Distance

Dropped Columns: State FIPS and County FIPS

Consistency Check:

A table with numbers and symbols

Description automatically generated

‘Population Not Staying at Home’ had a minimum value less then zero, which is not possible. 10 rows had a negative number in this column, as this was considered to be unreliable data, these rows were deleted.

Missing values: Missing values were found for State Postal Code and County Name in the National level rows, as expected.

Null values: 20 rows were found with no data, they were deleted.

Duplicates: None Found

Mixed types: State Postal Code and County Name due to the NaN values

Data frame broken in to a 2 data frames for analysis, a National level and a State level.

Motor Vehicle Registrations Dashboard Data

Consistency Check: No anomalies

Missing Values: None Found

Duplicates: None Found

Mixed Types: None Found

State Covid Restrictions

This dataframe is small enough to be reviewed visually in Excel, there was no Panda analysis.

Consistency Check: No anomalies

Missing Values: None Found

Duplicates: None Found

Mixed Types: None Found

**Data Limitations:**

Trips by Distance

This data is from 2019 to 2022. So, information before and after that is unknown

Motor Vehicle Registrations Dashboard Data

This data is from 1900 to 2020. So, won’t be able to capture what has happened in recent years

State Covid Restrictions

This dataframe only captures the imposing and lifting of emergency declarations, not details of the travel restrictions.

**Ethical Considerations:**

This information was gathered from Data.Gov and is publicly available for study. It does not carry any personal identifiers and does not have any ethical concerns regarding privacy, transparency, or fairness.

**Questions to Explore:**

* What are the differences in trips vs. people staying home
* What have the changes been over time?
* Have trips gotten longer or shorter?
* How much effect did Covid restrictions have on trips?
* Was there a lasting change after Covid restrictions?
* How has the number of trips compared to the number of vehicles registered?