ACCIDENTALLY_LATE

J. KIRK, S. MCNAIR, N. FOSTER, K. HOLMBERG, M. ONEAL

SELECTED TOPIC

-Traffic accidents 2018-2022

-Impact of Covid: increase in accidents in Covid years vs. Pre-Covid years

-Impact of weather: temperature, visibility, wind speed, precipitation, weather condition, sunrise vs sunset, year, and date

REASON TOPIC SELECTED

This topic was chosen to see if a correlation exists between the severity of car accidents during the Covid years when compared to the severity of car accidents in pre-Covid years. In this context, “severity” refers to the delay a car accident causes in the traffic pattern. A longer delay is would mean a more substantial car accident.

Who/What Might Use the Data

-Weather Channels: utilize weather data to advise drivers of what to expect; delays, types of accidents, locations of increase of accidents

-Insurance Companies: raise rates in high severity/frequent areas

-Car rental agencies deciding price rates

-Companies debating working from home and start times for work

TOP 5 STATES WITH THE MOST ACCIDENT OCCURRENCES

ACCIDENTS UPDATED	PreCOVID accidents	COVID accidents
CALIFORNIA (611,261)	California (125,281)	California (480,030)
FLORIDA (355,328)	Oregon (35,137)	Florida (342,487)
TEXAS (111,194)	Minnesota (17,267)	Texas (103,287)
OREGON (100,907)	Utah (13,115)	Virginia (89,056)
VIRGINIA (93,357)	Florida (9,460)	Pennsylvania (85,880)

DESCRIPTION OF DATA SOURCE

-The data was pulled from US-Accidents: A Countrywide Traffic Accident Dataset - Sobhan Moosavi (smoosavi.org)

-Dataset has 47 columns of information; we included the following columns for the purpose of the project: severity, start time, state, temperature, visibility, wind speed, precipitation, weather condition, sunrise vs sunset, year, and date

QUESTIONS EXPECTED TO BE ANSWERED

What impact has Covid had on the severity of car accidents in the US?

Has weather played a role in the information?

Top 5 Weather Conditions with the most accident occurrences

	Accidents_updated	PreCOVID_accidents	COVID_accidents
Fair	1044151	115198	927250
Cloudy	335487	41604	293072
Mostly Cloudy	288935	32253	253466
Partly Cloudy	196112	31529	163688
Light Rain	121335	18454	89987

DATA EXPLORATION PHASE INFORMATION

According to the data, as of right now, the number of accidents went up during Covid years when compared to pre-Covid years

```
#11a. Get count of unique values in the 'Severity' column from 'accidents_upated' dataframe
print(accidents_updated['Severity'].value_counts())
```

```
2    2084426
3      71340
4      66076
1      23578
Name: Severity, dtype: int64
```

```
#11b. Get count of unique values for Severity column from 'PreCOVID_accidents'
print(PreCOVID_accidents['Severity'].value_counts())
```

```
2    228702
3     27134
4     21055
1       150
Name: Severity, dtype: int64
```

```
#11c. Get count of unique values for Severity Column from 'COVID_accidents'
print(COVID_accidents['Severity'].value_counts())
```

```
2    1828676
4      39448
3      36885
1      23422
Name: Severity, dtype: int64
```

ANALYSIS PHASE INFORMATION

```
#11. Create a PreCovid Dataframe
#accidents_updated = pd.PreCOVID_accidents(date)
start_date = '2018-06-01'
end_date = '2020-02-29'
# Select DataFrame rows between two dates
mask = (accidents_updated['date'] > start_date) & (accidents_updated['date'] <= end_date)
PreCOVID_accidents = accidents_updated.loc[mask]
PreCOVID_accidents.head()
```

	Severity	State	Temperature(F)	Visibility(mi)	Wind_Speed(mph)	Precipitation(in)	Weather_Condition	Sunrise_Sunset	year	date
582028	long_delay	NJ	89.0	10.0	5.0	0.0	clear_weather	Day	2019	2019-10-02
1295810	long_delay	LA	54.0	10.0	5.0	0.0	clear_weather	Day	2019	2019-11-01
1537770	long_delay	AZ	41.0	10.0	10.0	0.0	clear_weather	Night	2020	2020-02-16
1568027	long_delay	TX	79.0	10.0	25.0	0.0	bad_weather	Day	2019	2019-06-04
1756843	long_delay	VA	37.0	10.0	0.0	0.0	clear_weather	Night	2019	2019-10-19

```
#12. Create a COVID Dataframe
#accidents_updated = pd.COVID_accidents(date)
start_date = '2020-03-01'
end_date = '2021-12-31'
# Select DataFrame rows between two dates
mask = (accidents_updated['date'] > start_date) & (accidents_updated['date'] <= end_date)
COVID_accidents = accidents_updated.loc[mask]
COVID_accidents.head()
```

	Severity	State	Temperature(F)	Visibility(mi)	Wind_Speed(mph)	Precipitation(in)	Weather_Condition	Sunrise_Sunset	year	date
224945	short_delay	MA	42.0	10.0	12.0	0.0	clear_weather	Night	2021	2021-03-10
224946	short_delay	CA	54.0	2.0	6.0	0.0	bad_weather	Night	2021	2021-07-30
224947	short_delay	MD	79.0	10.0	9.0	0.0	clear_weather	Day	2021	2021-10-15
224948	short_delay	WA	38.0	10.0	0.0	0.0	clear_weather	Day	2021	2021-12-21
224949	short_delay	CA	52.0	10.0	9.0	0.0	clear_weather	Day	2021	2021-12-09

ANALYSIS CON'T

As noted in the visual, the pre-Covid dates are 6/1/18-2/29/20. And the data has been analyzed.

The Covid dates are 3/1/20-12/31/21 and those have been analyzed as well.

Both sets of dates have had the data pulled for the information listed on the "Description of Data" slide