IMPACT OF COVID-19 ON NORTHERN VIRGINIA TRAFFIC ACCIDENTS
ZACK OSGOOD
FEBRUARY 22ND, 2021

1. INTRODUCTION

1.1. BACKGROUND

In the early months of 2020, COVID-19 had taken ahold of American society and changed the way we functioned on a day-to-day basis. Multiple safety mandates and recommendations were quickly put in place. One of the most noticeable for all in the Northern VA region was the quick shift telework. With the Northern VA region having some of the worst commuting conditions in the country, many have enjoyed the ability to telecommute, yet how have the changes in conditions affected accidents?

1.2. PROBLEM

The following report explores the impact COVID has had on Northern VA traffic by comparing traffic accident data of 2019 (pre-COVID conditions) to traffic accident data of 2020 (COVID conditions). As vaccines are issued corporate & government leaders will need to plan for employees to begin returning to working on site. The information gathered in this report can help those leaders better anticipate the effect this return will have on traffic accidents in the region.

1.3. INTEREST

The target audience for this report is corporate and government leaders in the Northern VA region, yet anyone who lives in the region could find this information useful to better prepare for their own return to working onsite.

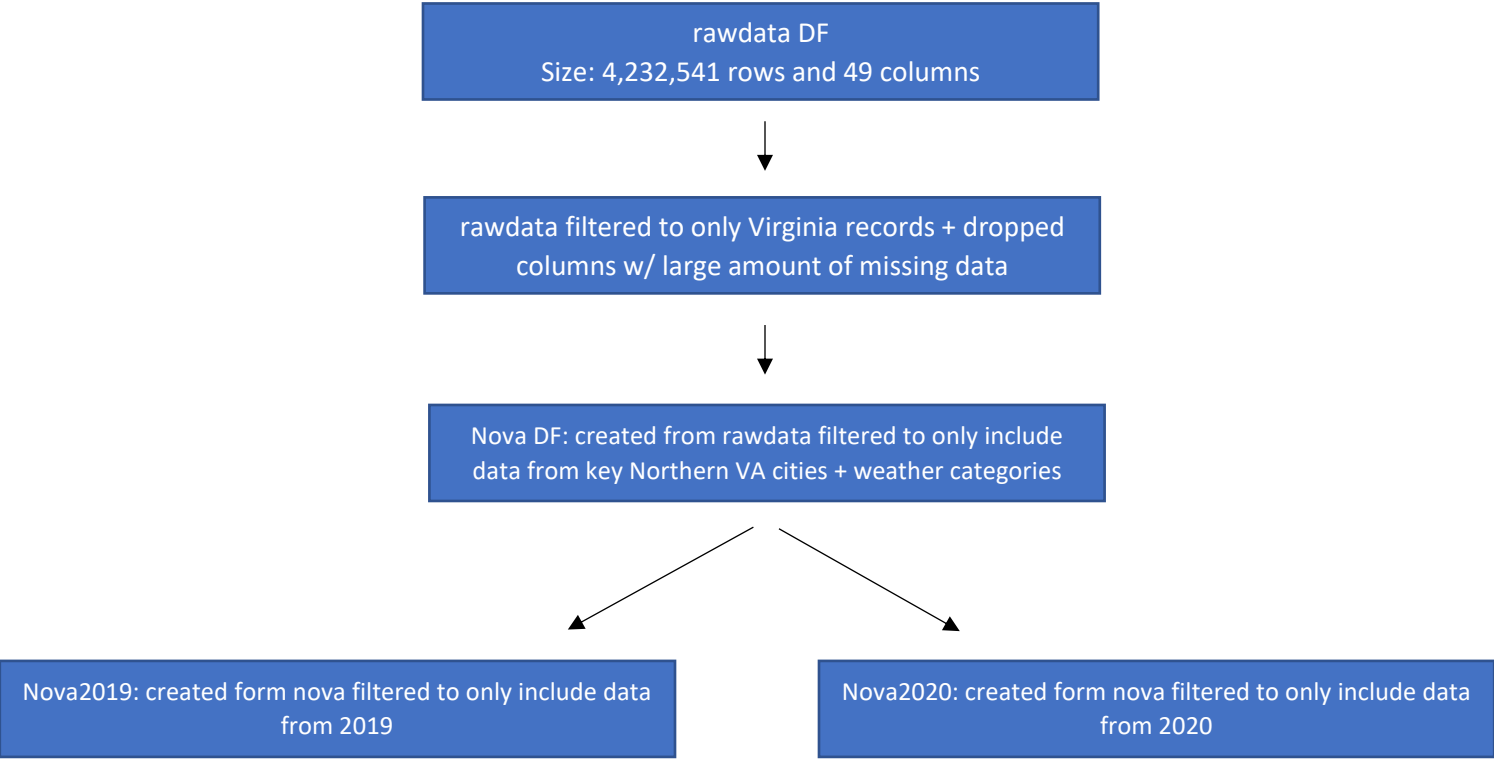## 2. DATA ACQUISTION & CLEANING

2.1. DATA SOURCES

To solve the question of COVID's impact on traffic accidents, I'll be using a dataset collected by Lyft Research Scientist, Sobhan Moosavi. This data has been continuously collected since February 2016, using several data providers, including two APIs which provide streaming traffic event data. The APIs broadcast traffic events captured by a variety of entities, such as the US and state departments of transportation, law enforcement agencies, traffic cameras, and traffic sensors within the road-networks. The dataset consist of roughly 4.2 million accidents recorded across 49 states.

2.2. DATA CLEANSING

The first action performed in the data cleansing process was uploading the dataset into a Pandas dataframe titled 'rawdata'. The rawdata dataframe consisted of 4,232,541 rows and 49 columns. From here I performed the following actions

1) Dropped columns that had large amounts of missing information
2) Filtered rawdata to only include data that was recorded in the state of Virginia, new size of rawdata now equals 127,949 rows and 30 columns
3) Created 'nova' dataframe from rawdata that only consisted of data collected from the following Northern VA cities : Arlington, Reston, Alexandria, Fairfax, Falls Church, Springfield, Ashburn, Chantilly, Fort Belvoir, Herndon, McLean, Vienna. Nova dataframe consist of 17,561 rows and 30 columns.
4) Categorize weather conditions into 6 main categories (overcast, fair, rain, storm, fog, winter conditions).
5) Created nova2019 dataframe which only consist of data recorded in 2019
6) Crreated nova2020 dataframe which only consisted of data recorded in 2020

2.3. DATA CLEANSING WORKFLOW VISUAL

rawdata DF
Size: 4,232,541 rows and 49 columns

↓

rawdata filtered to only Virginia records + dropped columns w/ large amount of missing data

↓

Nova DF: created from rawdata filtered to only include data from key Northern VA cities + weather categories

Nova2019: created form nova filtered to only include data from 2019

Nova2020: created form nova filtered to only include data from 2020

3. EPLORARTORY DATA ANALYSIS

3.1. OVERALL 2019 VS 2020 COMPARISON

In 2019 there were a total of 2,901 accidents in Northern VA compared to 5,236 accidents recorded in 2020. Given these numbers we can see there has been an 80.49% increase in accidents from Pre-COVID conditions to COVID conditions in the Northern VA region.
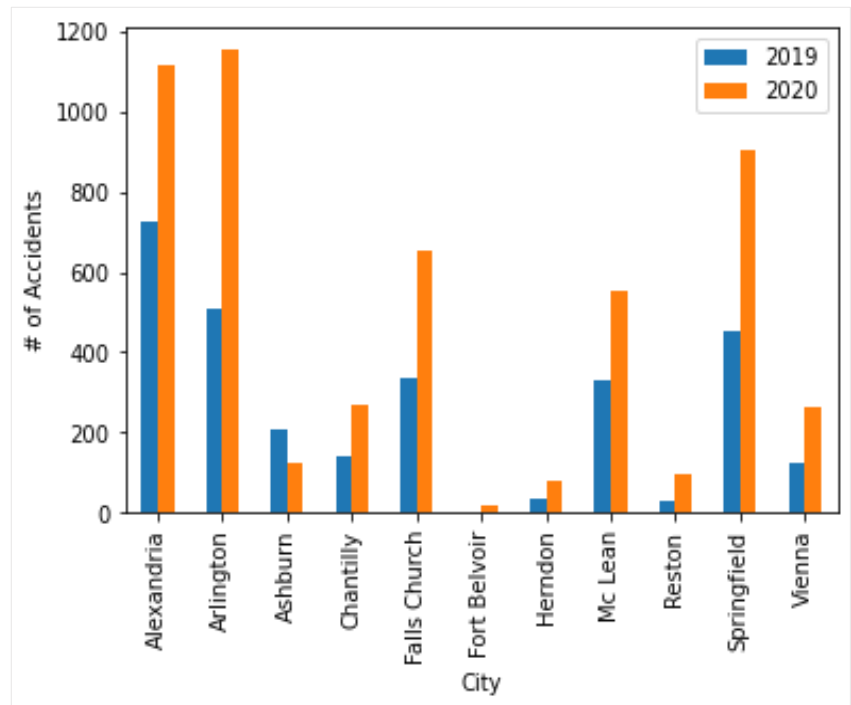
2019 vs 2020: Total Accidents
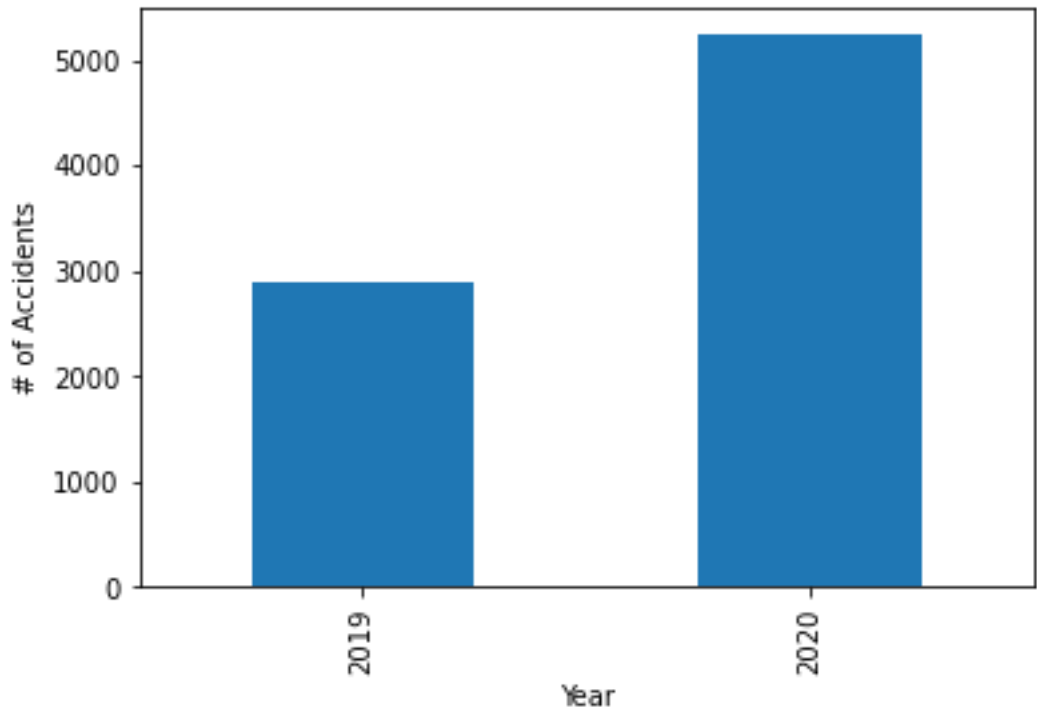
3.2 2019 vs 2020: COMPARISON BY CITY

By viewing this data, we can see that 10 out of 11 cities had an increase in accidents (Ashburn was the only city that decreased). Overall, there was an average 124% increase in accidents across all cities , with the largest change occurring at Fort Belvoir (425% increase). We conclude that the increase in accidents was region wide and not specific to any one city.

	2019	2020	Percent_Change
Alexandria	724	1118	54.419890
Arlington	511	1153	125.636008
Ashburn	210	126	-40.000000
Chantilly	141	270	91.489362
Falls Church	338	651	92.603550
Fort Belvoir	4	21	425.000000
Herndon	35	77	120.000000
Mc Lean	332	554	66.867470
Reston	31	99	219.354839
Springfield	451	905	100.665188
Vienna	124	262	111.290323

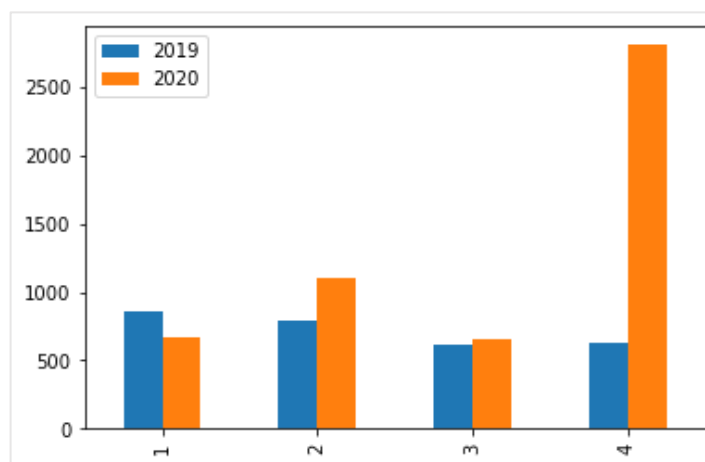## 3.3 2019 vs 2020: TIMELINE COMPARISON

 When analyzing the number of accidents that occurred between 2019 and 2020 three timeline comparisons were made;

1. Total number of accidents per month
2. Total number of accidents per quarter
3. Percent of total accidents per quarter

Even from a quick glance it was obvious that majority of the increase in accidents occurred in quarter 4 of 2020. On average each quarter increased by 92%, while quarter 4 increased by 348% percent in 2020. Overall, 53% of all accidents in 2020 occurred in quarter 4, while only 21.5% of total accidents in 2019 occurred in quarter 4. From this data we can conclude that quarter 4 was the only quarter in the year that saw a significant increase in accidents.
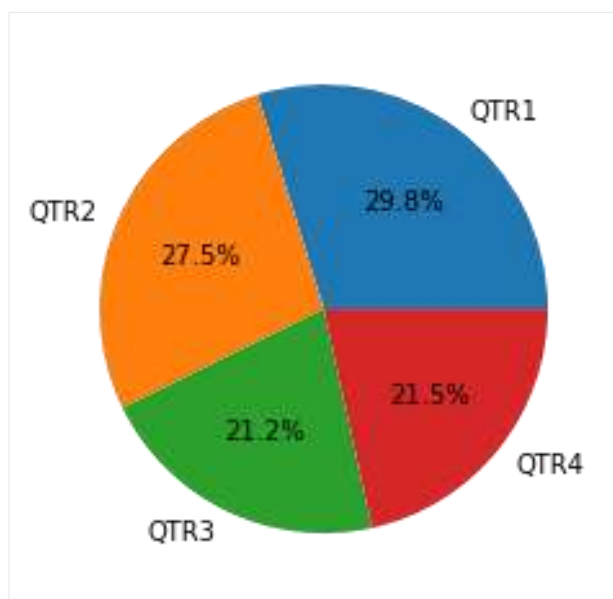
2019 vs 2020: ACCIDENTS BY QUARTER

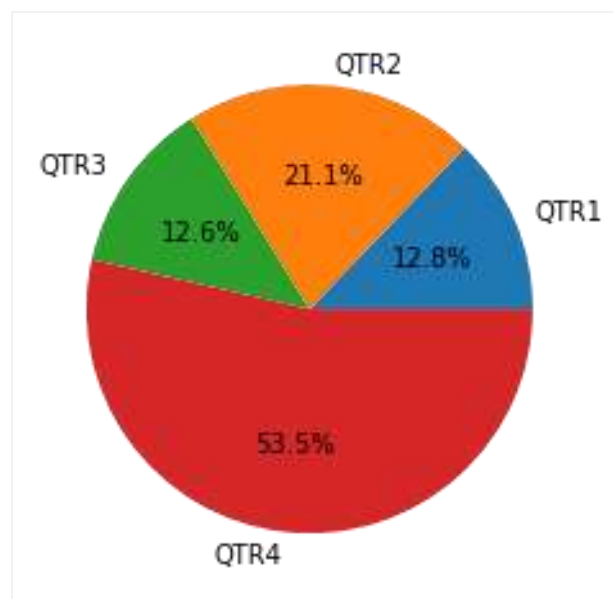|	2019	2020	Percent_Change
1	864	670	-22.453704
2	797	1103	38.393977
3	615	660	7.317073
4	625	2803	348.480000

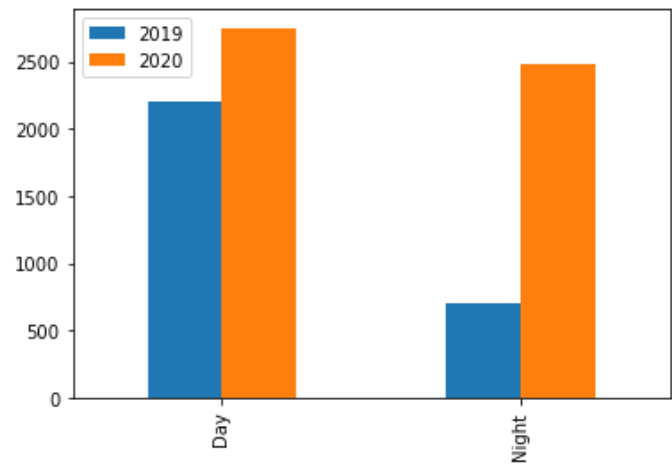# 2019 vs 2020: ACCIDENTS BY MONTH
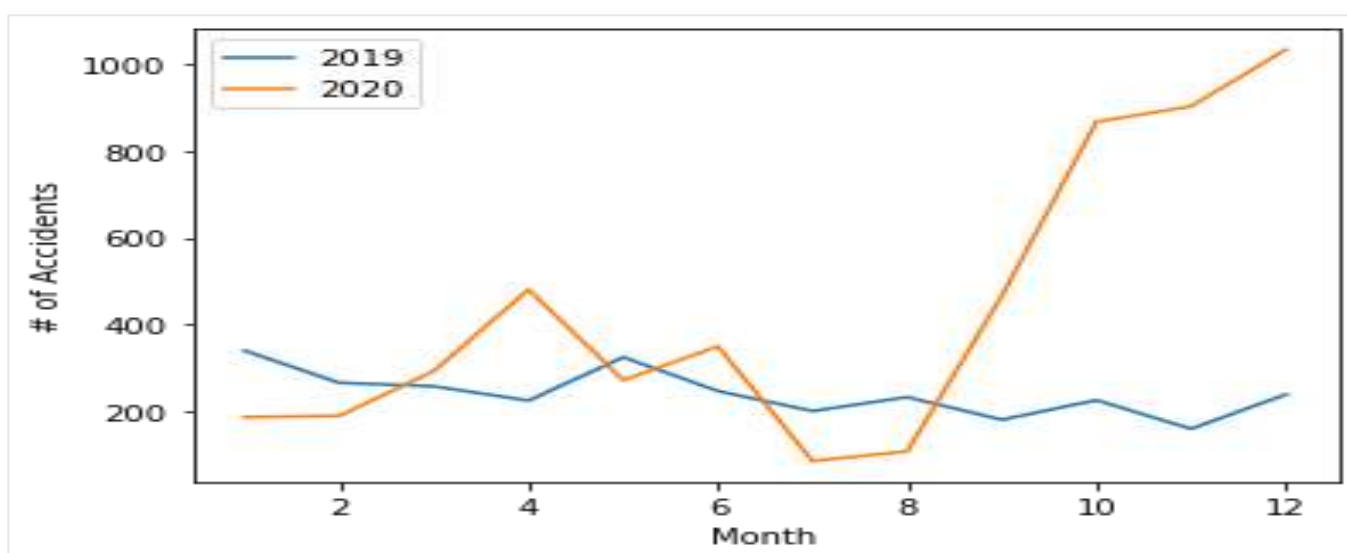
# 2019: ACCIDENTS BY QUARTER

2020: ACCIDENTS BY QUARTER

3.4 2019 vs 2020: TIME OF DAY COMPARISON

The main finding observed in the time-of-day data was that majority of the increase in accidents in 2020 occurred at night. Overall daytime accidents increased by 25% in 2020, while nighttime accidents occurred by a staggering 255%. From this data we could easily conclude that more accidents occurred at night during COVID compared to a more even distribution with daytime pre-COVID.
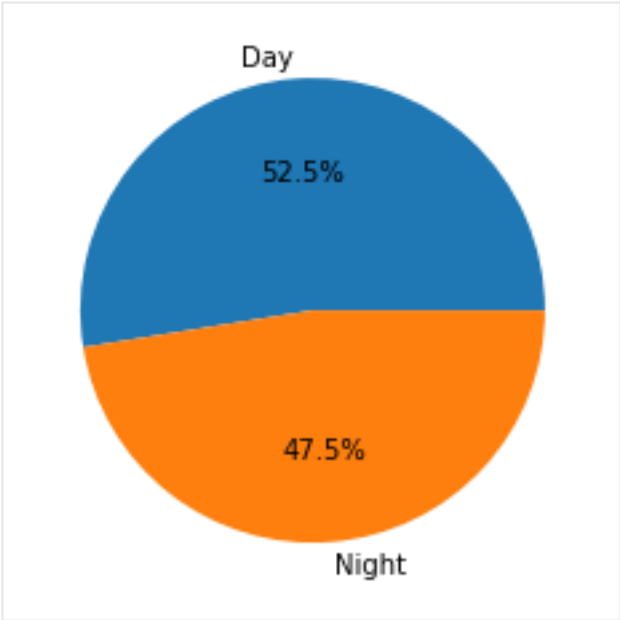
	2019	2020	Percent_Change
Day	2201	2751	24.988642
Night	700	2485	255.000000

2019 ACCIDENTS: DAY vs NIGHT
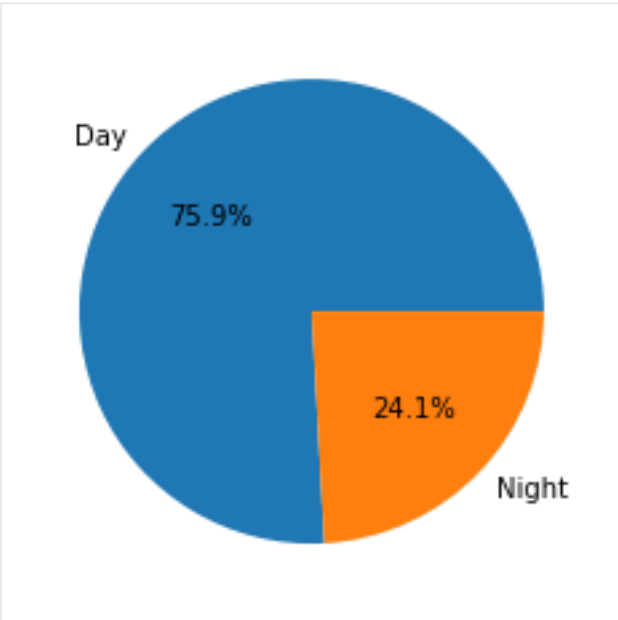
**2020 ACCIDENTS: DAY vs NIGHT**

3.5 2019 vs 2020: SEVERITY COMPARISON

On average each severity category increased by 60% from 2019 to 2020. The largest increased occurred in category 2 where a 174% in accidents occurred. Analyzing the data, it can be concluded that overall, the distribution of data across the different severity categories stayed consistent except category 2.

2019 vs 2020: ACCIDENT BY SEVERITY

|	2019	2020	Percent_Change
1	NaN	48	NaN
2	1403.0	3847	174.198147
3	1373.0	1190	-13.328478
4	125.0	151	20.800000
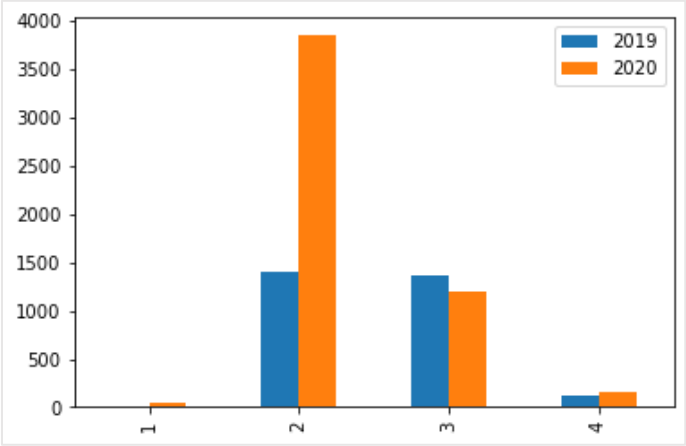
3.6 2019 vs 2020: WEATHER COMPARISON

On average each weather category saw a 67% increase between 2019 and 2020. Taking a deeper dive into the data it can be concluded that the Fair and Rain categories were the two areas that saw significant increase, 157% and 136% respectively. Proportionally the weather category accident distributions stayed the same with overcast weather have the strongest correlation with accidents.

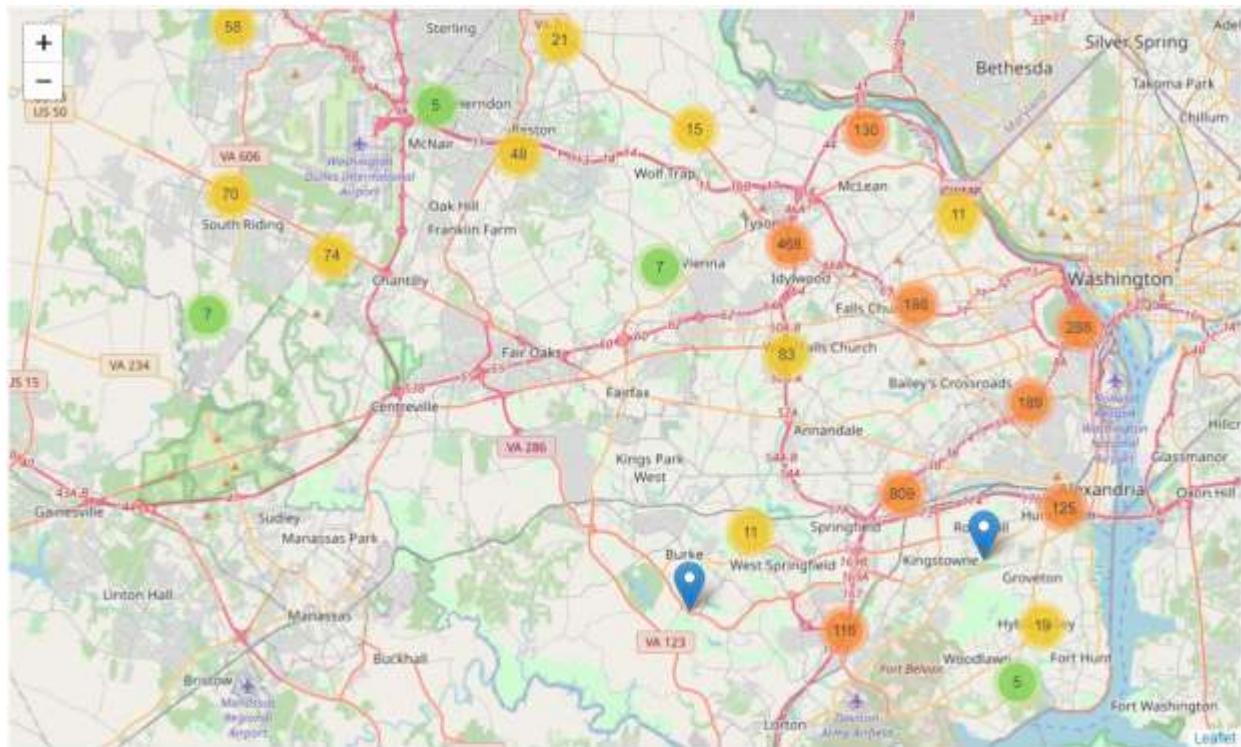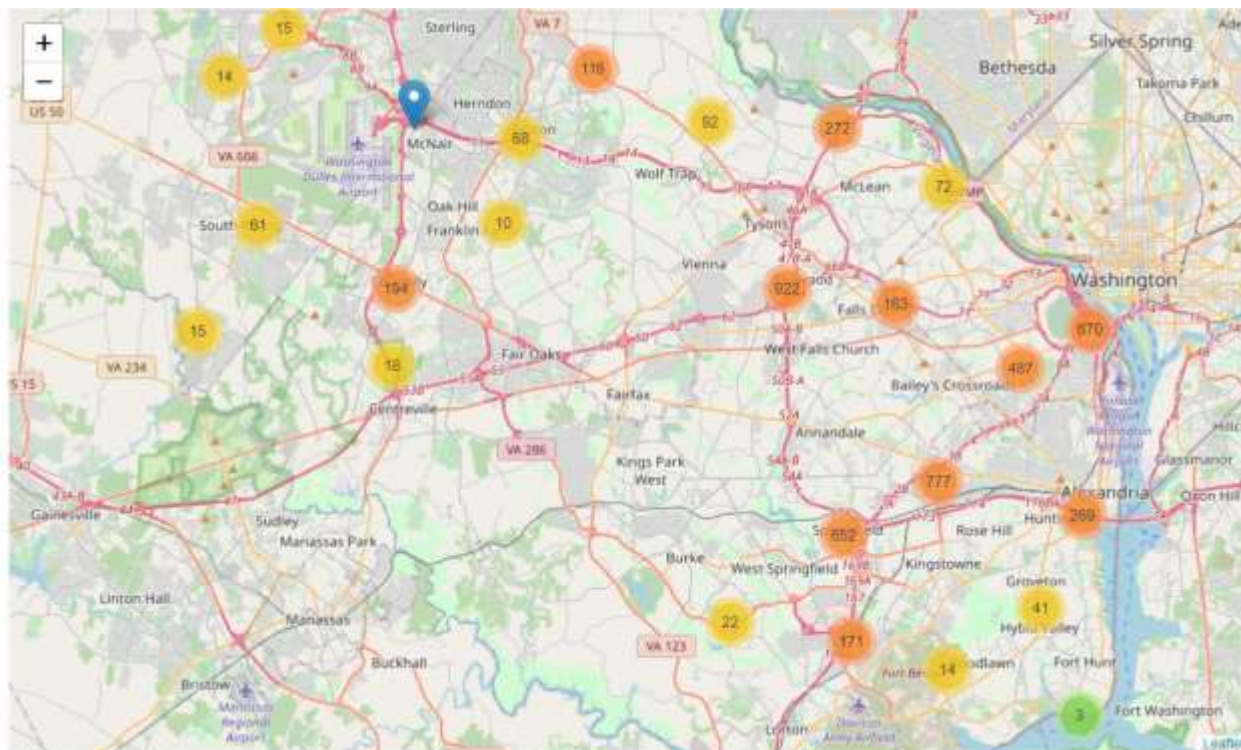|	2019	2020	Percent_Change
Fair	746	1758	135.656836
Fog	46	30	-34.782609
Overcast	1824	2776	52.192982
Rain	214	552	157.943925
Storm	38	49	28.947368
Winter Conditions	24	39	62.500000
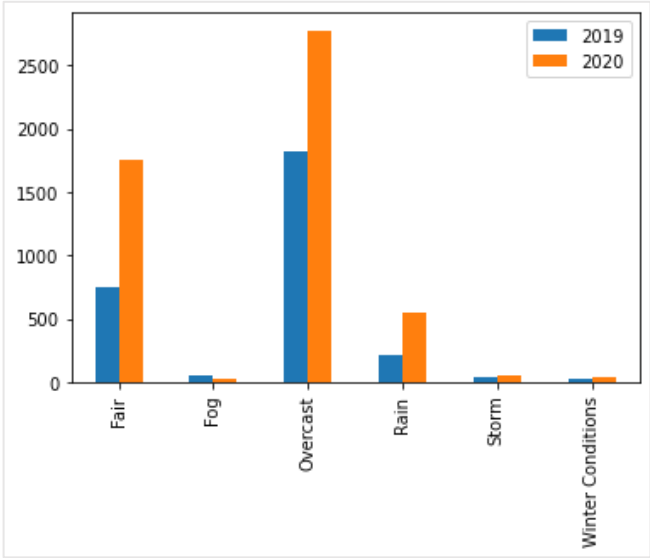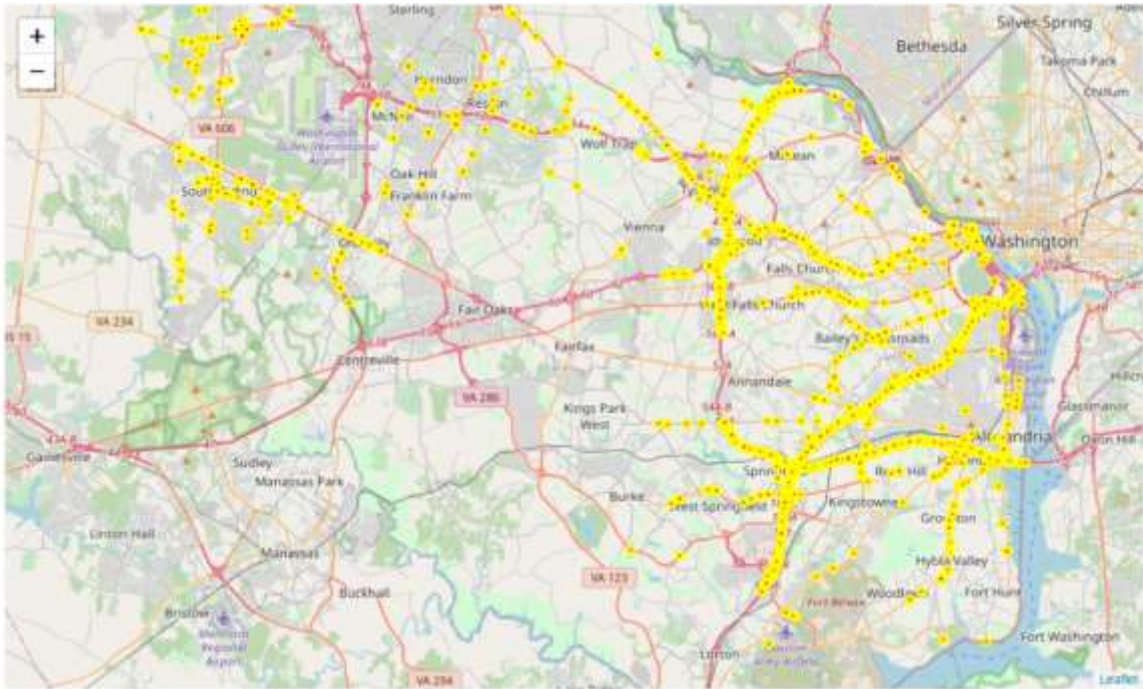
2019 ACCIDENTS
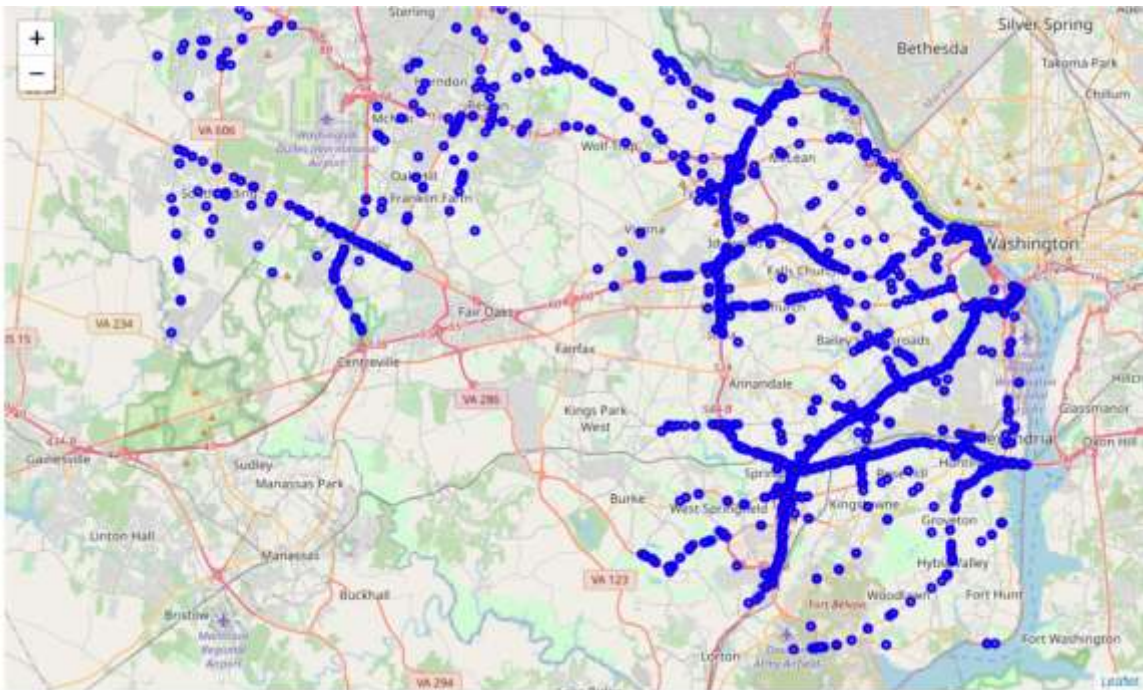
2020 ACCIDENTS

2019 ACCIDENTS

2020 ACCIDENTS

4. CONCLUSIONS

In this study, I analyzed the differences in traffic accidents data between 2019 and 202 to determine the impact COVID-19 has had on traffic accidents in the Northern VA region. I identified how location, time of day, weather, time of year, and severity influenced the number of accidents that occurred. The average percent change and maximum percent change was calculated for each factor to determine the level of significance in each factor and its relationship to the accidents that occurred. These calculations can help leaders in the area prepare for traffic accident changes as work conditions return to normal.