

Model Behavior Datafest Write Up

Zoe Chen (zoechen17@gmail.com), Rene Honarchian (rhonarchian@gmail.com), Sean Lashkari (slash5324@hotmail.com), Megan Li (meg.li494@gmail.com), Zach Loran (zachloran@gmail.com)

During this difficult time where many people around the world are stuck at home in quarantine, certain hobbies are taken up. One particular activity that has been shown to be popular among adults is drinking. Drinking has been proven to increase activity in the dopamine neurons, as well as opioid cells that release endorphins. Our team figured that alcohol sales may have increased during this time because we have observed this behavior first hand with our friends and loved ones. Instead of looking at the country as a whole, we decided to look into one particular state since it would be easier to analyze. We chose Iowa as our state.

On March 9th of 2020, the people of Iowa received the news to stay home and that non-essential businesses would close. We were interested in analyzing data for alcohol sales in Iowa over the period of dates from March 9-April 30. These dates correspond to the period right before the beginning of state-wide restrictions and range until the most current data available. To view trends in alcohol sales over the quarantine period, we also looked at data from this range of dates over the past four years as well. We used a dataset on alcohol sales in Iowa found on <https://data.iowa.gov/Sales-Distribution/Iowa-Liquor-Sales/m3tr-qhgy/data>. In the upper left plot below, we look at the average daily sales (in dollars) of alcohol purchases in Iowa in March and April from 2016-2020. We can see an overall upward trend in alcohol sales; particularly we see a larger increase in sales from previous years to this current year.

We did ANOVA tests to determine if the average daily alcohol sales differed per year. We found year to have a significant impact on sales with a p-value of $2.2e-16$.

```

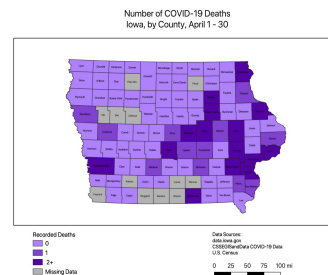
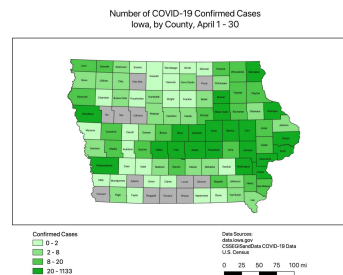
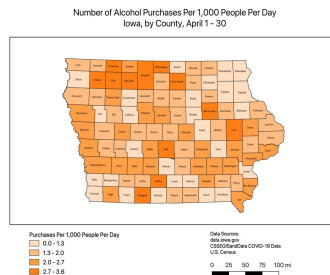
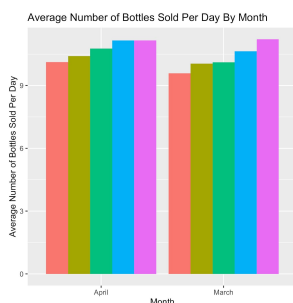
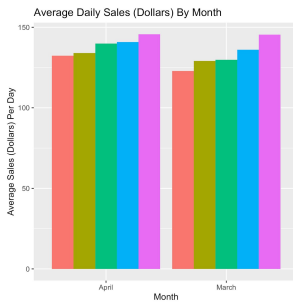
Response: sales
          Df    Sum Sq Mean Sq F value    Pr(>F)
year       4 6.2812e+07 15703022  81.556 < 2.2e-16 ***
Residuals 1655532 3.1876e+11  192543
    
```

We also did an ANOVA test to determine if the average number of bottles sold per day differed by year and again we found year to have a significant impact.

```

Response: bottles_sold
          Df    Sum Sq Mean Sq F value    Pr(>F)
year       4  374489   93622  127.21 < 2.2e-16 ***
Residuals 1655532 1218445247   736
    
```

These tests merely tell us that metrics from at least two of the years differed, not which years. The following bar graphs and line plot make it clear that 2020 induced the greatest average sales and bottles sold for the months of March and April (when COVID-19 mandated restrictions). We also obtained the data of daily coronavirus cases across the world and filtered to map the data by county. ([CSSEGISandData/COVID-19-GitHub](https://cseegisanddata.com/COVID-19-GitHub/)).



We also examined the spatial relationship between alcohol purchases and COVID deaths and cases, across different Iowa counties. From the maps (April maps shown above), it looks as though there is no spatial relationship, with a cluster of high-alcohol purchasing counties in the northwest, while most of the high COVID-19 death and cases counties are in the Southeast of the state. While this is the opposite of what we hypothesized, it shows only that there is no **spatial** relationship between alcohol and COVID -- not that there is no relationship at all.