Border Patrol Analysis

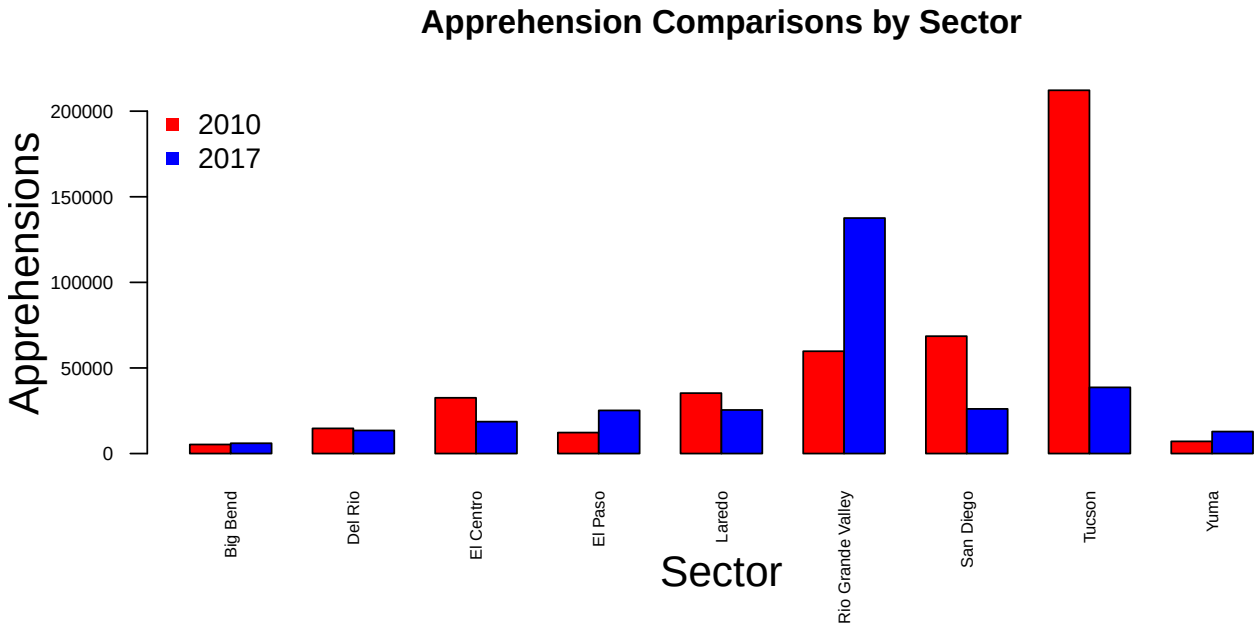
Udochi Maduakor, Ozair Meghani ,Junpei Xiao

February 11, 2018

Data Analysis for United States Border of Apprehensions between 2010 and 2017

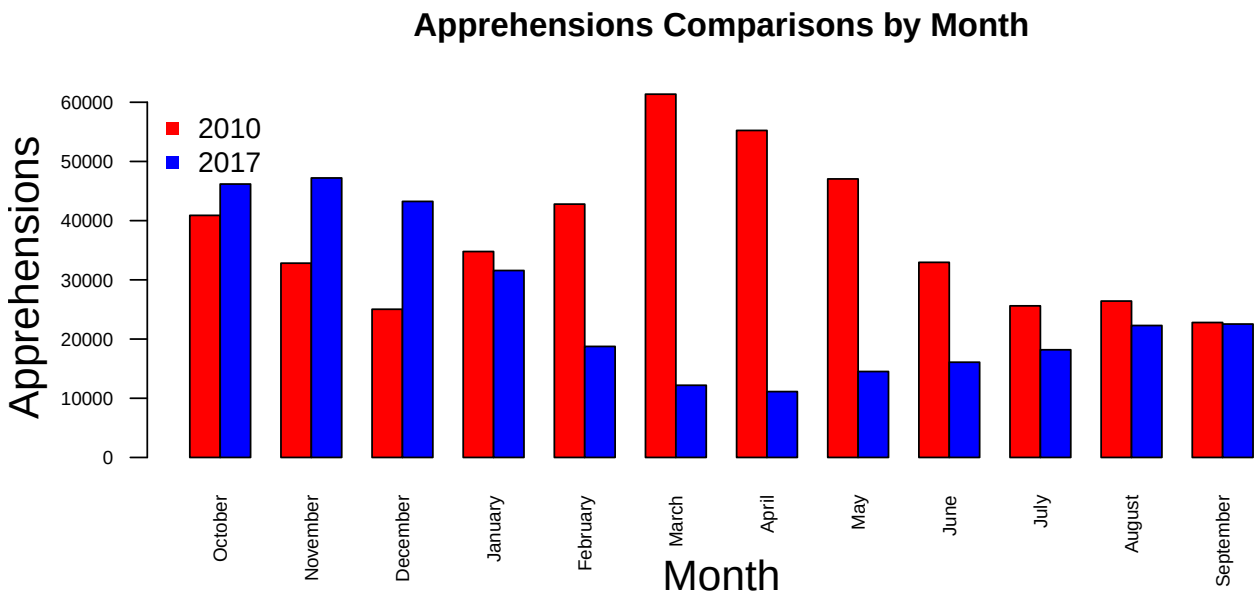
According to CNN Border Patrol Apprehensions reached an all time low in the month of April 2017. They claimed to have seen record breaking lows in apprehensions, that may be influenced by Donald Trump's rhetoric and strong public views on immigrants and immigration laws. We have decided to take a look at the data ourselves to determine if these claims are valid. Our observations and results are reported below.

Bar Plot comparison between the sectors in 2010 and 2017



From the bar plot comparisons for sectors shown above there is not a consistent trend in the change in Border Patrol apprehensions between 2010 and 2017 among all of the sectors as a whole. In some sectors such as El Paso and Rio Grande Valley the number of apprehensions increased from 2010 to 2017, while in other sectors it decreased. One major observation that can be seen is the notable decrease in apprehensions in Tucson over the 7 year time period.

Bar Plot Comparison between the Months in 2010 and 2017



Once again there is no visible trend noticable in changes in Border Patrol apprehensions from 2010 to 2017 based on the monthly comparison bar plot above. Some attention worthy changes to note are the significant decreases in apprehensions in the months of March, April, and May. There are also some unexpected increases in apprehensions during the months of October, November, and December.

T-Tests comparing Max Apprehensions by Sector in 2010 and 2017

```
##
##  F test to compare two variances
##
## data:  maxsect2010 and maxsect2017
## F = 0.93205, num df = 11, denom df = 11, p-value = 0.9092
## alternative hypothesis: true ratio of variances is not equal to 1
## 95 percent confidence interval:
##  0.2683166 3.2376648
## sample estimates:
## ratio of variances
##           0.932051
```

```
##
##  Two Sample t-test
##
## data:  maxsect2010 and maxsect2017
## t = 1.9547, df = 22, p-value = 0.06344
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##   -379.1203 12819.1203
## sample estimates:
## mean of x mean of y
##   17683.5   11463.5
```

The next portion of our analysis involved conducting a T-test between the sector with the most apprehensions in 2010 with the sector that had the most apprehensions in 2017. We wanted to determine if there was a statistically significant decrease in between sectors with maximum apprehensions. Based on observing the numbers alone, we can see that Tucson had the maximum apprehensions in 2010 with 212,202, while Rio Grande Valley had the maximum apprehensions in 2017 with 137,562, which is a decrease of roughly 75,000. First we conducted a Variance F-Test to see if these two sectors had equal variance, which they did. Then, we conducted an equal variance T-test to determine if there really was a significant decrease in mean apprehensions between these sectors. Based on our results shown above, (p-value of about 0.06) we were not able to reject the null hypothese, and thus concluded that there was not enough evidence to support a statistically significant difference in means. So although there was a visible decrease, the difference in means was not significant between the max apprehensions in 2010 vs. 2017.

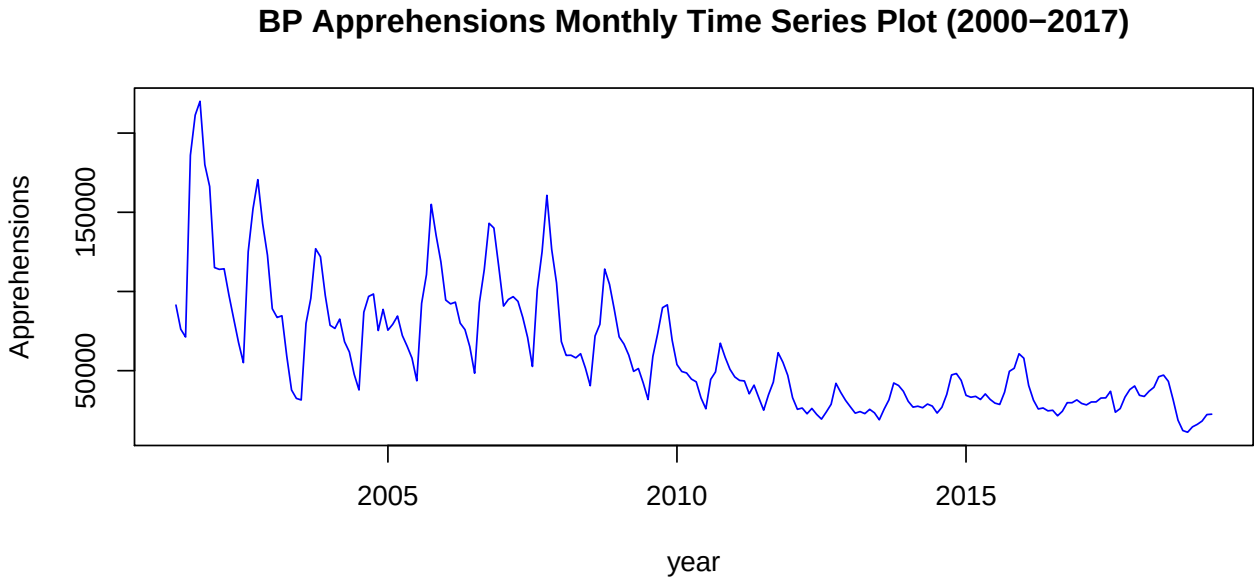
T-Tests comparing Max Apprehensions by 3-Month Period in 2010 and 2017

```
##
##  F test to compare two variances
##
## data:  max3mnth2010 and max3mnth2017
## F = 1.4469, num df = 26, denom df = 26, p-value = 0.3522
## alternative hypothesis: true ratio of variances is not equal to 1
## 95 percent confidence interval:
##  0.6594069 3.1750302
## sample estimates:
## ratio of variances
##            1.44694
```

```
##
##  Two Sample t-test
##
## data:  max3mnth2010 and max3mnth2017
## t = 0.48741, df = 52, p-value = 0.628
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -3116.650  5116.427
## sample estimates:
## mean of x mean of y
##  6060.852  5060.963
```

Our next test involves comparing the 3-month period with the most apprehensions in 2010 and 2017. For 2010 our 3-month period with the most apprehensions is March, April, and May with a total of 163,643 and a mean of 6060.852, which was is an average of the number of apprehensions at each sector during those 3 months. For 2017 our 3-month period with the most apprehensions is October, November, and December with a total of 136,646 and a mean of 5060.963. We once again conducted a Variance Test and based on the results of that an equal variance T-test. Our results (above) showed us that there is not a statistically significant difference in means between these 3-month periods.

Time Series Plot for the Monthly Summaries



The above plot is a time series plot of the monthly apprehensions from the year 2000 to 2017. As can be seen from the plot there has been a study decrease in Borer Patrol apprehensions over the years.