Project 2 Report (First Draft)

JessLMurray

10/21/2019

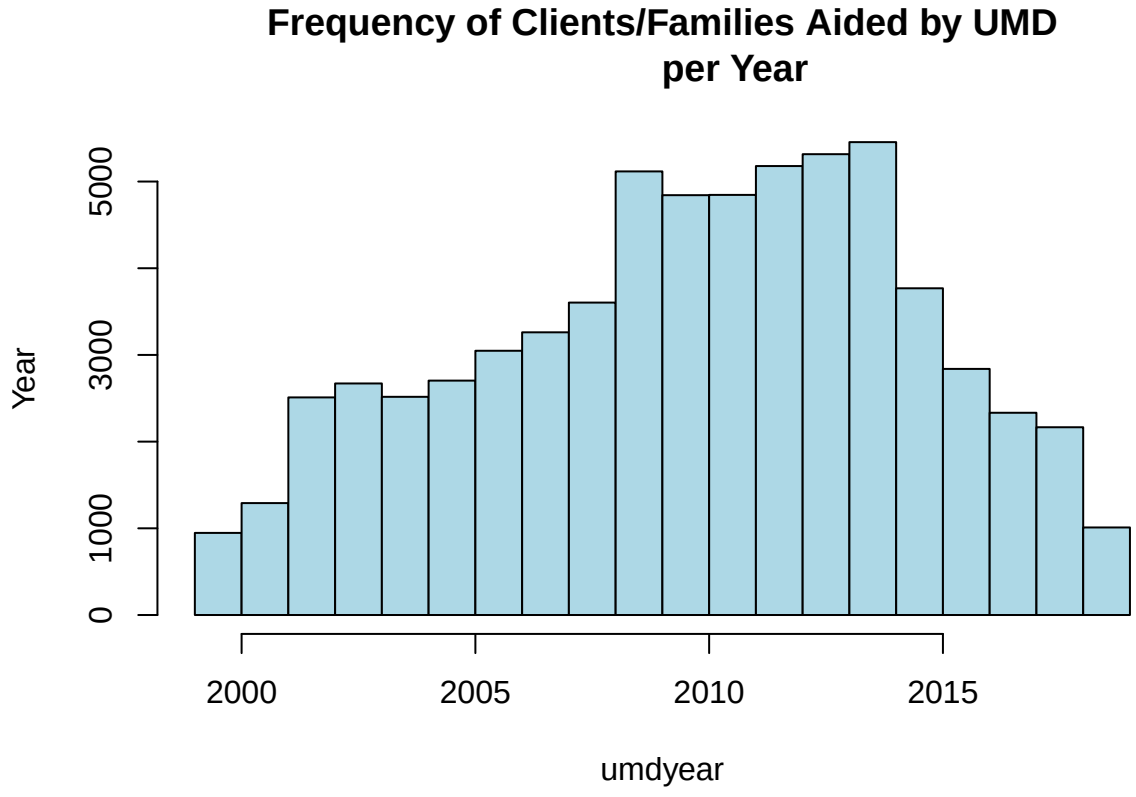
Urban Ministries of Durham Data Analysis & Shiny Dashboard

The goal of this project is to create a shiny app that will predict the number of instances service will be provided to UMD clients/families based on the national unemployment rate (data provided by the Bureau of Labor Statistics).

UMD Data Cleaning & Analysis

Note: This code was partially taken from the analysis performed in project 1.

```
#histogram of the frequency of clients/families aided per year
umdyear <- umd20$year
histo <- hist(umdyear, main="Frequency of Clients/Families Aided by UMD
              per Year", ylab="Year", col="light blue")
```

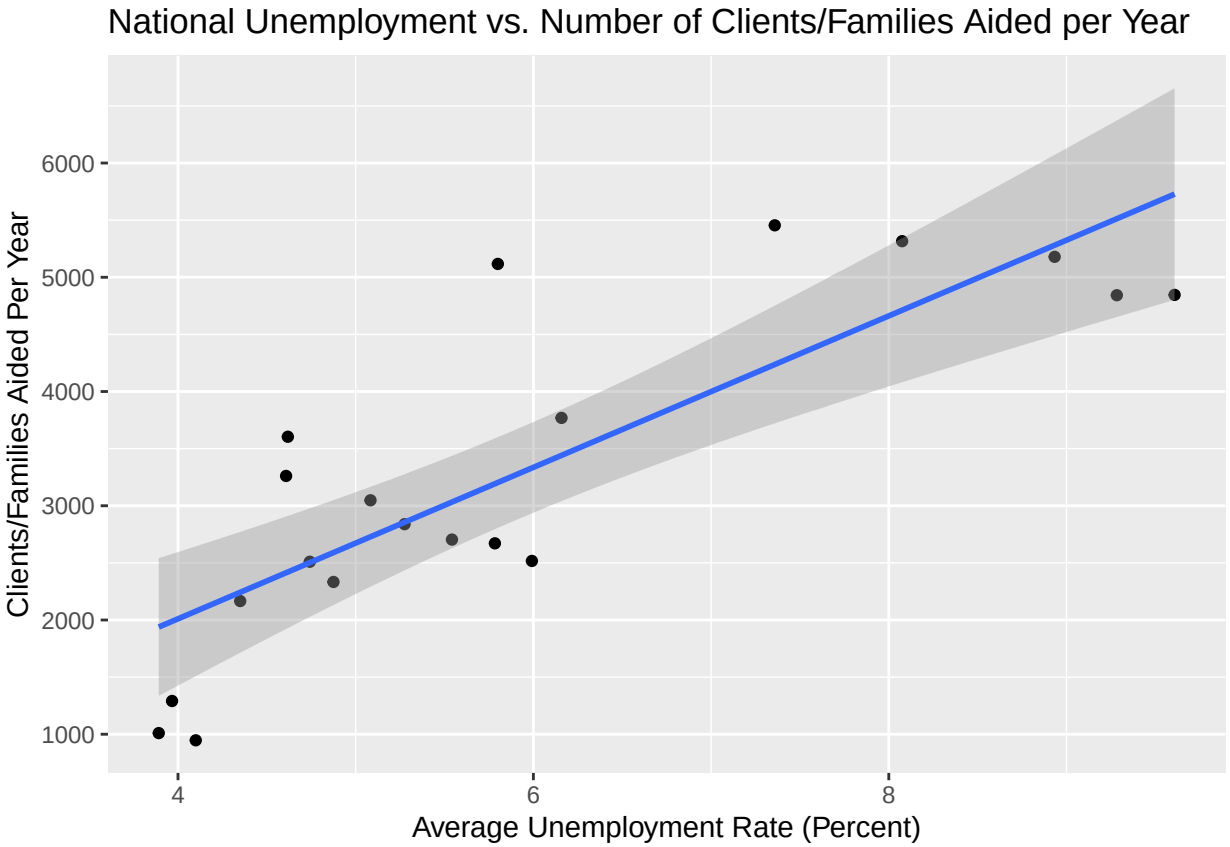



```
#correlation between the unemployment rate and the number of clients/families aided per year
blscorr <- data.frame("Clients/Families Aided Per Year" = counts,
```

```
                      "Unemployment Rate (Percent)" = bls_avg$`avg_unemployment`)
blscorr
```

```
##    Clients.Families.Aided.Per.Year Unemployment.Rate..Percent.
## 1                              947                    4.100000
## 2                             1291                    3.966667
## 3                             2510                    4.741667
## 4                             2671                    5.783333
## 5                             2517                    5.991667
## 6                             2704                    5.541667
## 7                             3048                    5.083333
## 8                             3261                    4.608333
## 9                             3604                    4.618182
## 10                            5117                    5.800000
## 11                            4843                    9.283333
## 12                            4846                    9.608333
## 13                            5179                    8.933333
## 14                            5316                    8.075000
## 15                            5455                    7.358333
## 16                            3769                    6.158333
## 17                            2839                    5.275000
## 18                            2333                    4.875000
## 19                            2166                    4.350000
## 20                            1010                    3.891667
```

```
ggplot(blscorr, aes(x=Unemployment.Rate..Percent., y=Clients.Families.Aided.Per.Year)) +
  geom_point() +
  geom_smooth(method = 'lm', se=TRUE, formula=y~x) +
  labs(x="Average Unemployment Rate (Percent)", y="Clients/Families Aided Per Year",
       title = "National Unemployment vs. Number of Clients/Families Aided per Year")
```



There appears to be a positive correlation between the national unemployment rate and the number of clients/families helped per year by UMD. Correlation analysis reveals that this association is statistically significant:

```
##
## Call:
## lm(formula = blscorr$Clients.Families.Aided.Per.Year ~ blscorr$Unemployment.Rate..Percent.)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1129.66  -676.57   -89.02   397.00  1913.42
##
## Coefficients:
##                                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)                          -641.2      661.7  -0.969    0.345
## blscorr$Unemployment.Rate..Percent.    662.9      107.5   6.169 7.99e-06
##
## (Intercept)
## blscorr$Unemployment.Rate..Percent. ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 843.8 on 18 degrees of freedom
## Multiple R-squared:  0.6789, Adjusted R-squared:  0.6611
## F-statistic: 38.06 on 1 and 18 DF,  p-value: 7.989e-06
```