#Problem Set 1: Solutions and Analysis using R  
##R edit by RJ Neel  
##Original Author: Ben Lambert

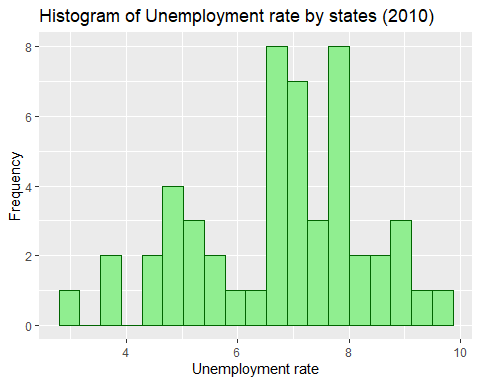
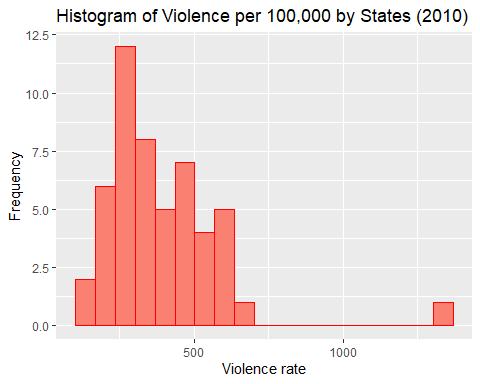
###Getting started with R  
R is a free programming language used for statistical analysis (and much more). Other softwares used for statistical analysis are Gretl, IBM SPSS, MATLAB, Python and many more. Python is the leading data analysis language. R is popular among academics.

The following link covers all the necessary information regarding R and R studio installation Link:<https://www.youtube.com/watch?v=9-RrkJQQYqY>

###The Crime-Unemployment data set

**#Questions 1 to 2: Histograms**  
Let’s begin by Visualizing the data.We do this by plotting Histograms.

## State Unemployment Violence  
## 1 Alabama 7.2 383.7  
## 2 Alaska 7.0 635.3  
## 3 Arizona 7.7 413.6  
## 4 Arkansas 7.2 503.5  
## 5 California 8.9 439.6  
## 6 Colorado 6.9 323.7



**#Questions 3: Maximum Violence state**  
The State with Maximum violence is

##   
## Attaching package: 'dplyr'

## The following objects are masked from 'package:stats':  
##   
## filter, lag

## The following objects are masked from 'package:base':  
##   
## intersect, setdiff, setequal, union

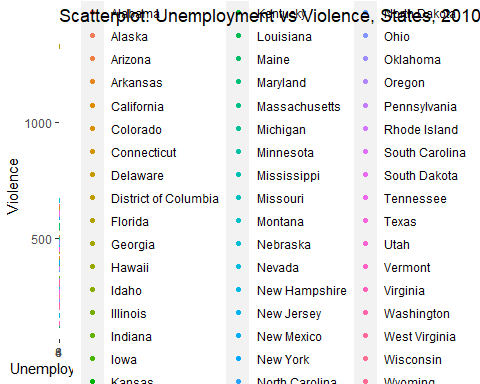
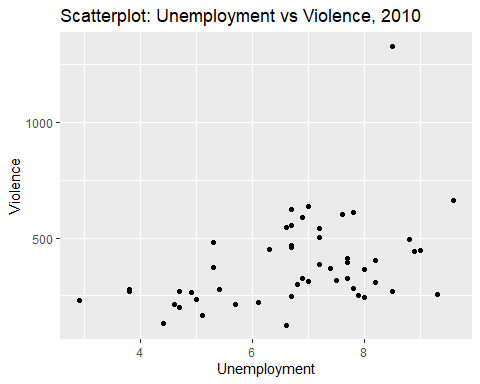
## State Unemployment Violence  
## 9 District of Columbia 8.5 1326.8

District of Columbia

**#Questions 4: Summary of Data**  
Summary Statistics

summary(data)

## State Unemployment Violence   
## Length:51 Min. :2.900 Min. : 122.1   
## Class :character 1st Qu.:5.550 1st Qu.: 260.9   
## Mode :character Median :6.900 Median : 323.7   
## Mean :6.765 Mean : 385.3   
## 3rd Qu.:7.800 3rd Qu.: 475.1   
## Max. :9.600 Max. :1326.8

**#Questions 5: Relationship between Violence and Unemployment**  
Let us try to check this using scatterplot  


*It appears to have a* ***positive*** *relation*

**#Questions 6: Correlation**  
The Correlation between Violence and unemployment is 0.42

**#Questions 7 to 8: Regression**

##   
## Call:  
## lm(formula = data$Violence ~ data$Unemployment)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -263.13 -110.84 -40.78 87.49 848.95   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 24.40 113.90 0.214 0.8313   
## data$Unemployment 53.35 16.43 3.248 0.0021 \*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 178.8 on 49 degrees of freedom  
## Multiple R-squared: 0.1771, Adjusted R-squared: 0.1603   
## F-statistic: 10.55 on 1 and 49 DF, p-value: 0.002102

## (Intercept) data$Unemployment   
## 24.39789 53.34785

**#Questions 9: Interpretation of Beta**  
For every 1 percentage change in unemployment we can expect an increase of 53 crimes per 100,000.  
**#Questions 10: Regression of Unemployment (Dependent) on Violence (Independent)**

##   
## Call:  
## lm(formula = data$Unemployment ~ data$Violence)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -3.3474 -0.9712 0.0427 1.0554 2.9599   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 5.485383 0.440645 12.449 <2e-16 \*\*\*  
## data$Violence 0.003320 0.001022 3.248 0.0021 \*\*   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 1.411 on 49 degrees of freedom  
## Multiple R-squared: 0.1771, Adjusted R-squared: 0.1603   
## F-statistic: 10.55 on 1 and 49 DF, p-value: 0.002102

## (Intercept) data$Violence   
## 5.485382917 0.003320499

**#Questions 11: Interpretation of Beta**  
For every 1 unit increase in violence per 100,00 we can expect an increase of 0.003 percentage unit in Unemployment.

**#Questions 11: Interpretation of Beta**  
Regression does not indicate causality but on average we can expect higher unemployment to be associated with high violence. Also, Higher violence may itself lead to unemployment.

#Other Analysis

## Warning: Use of `data$Unemployment` is discouraged. Use `Unemployment` instead.

## Warning: Use of `data$Violence` is discouraged. Use `Violence` instead.

## Warning: Continuous x aesthetic -- did you forget aes(group=...)?

