# R Notebook

1. Choose the response variable (Y) and one covariate (X) & Put thought for your response and covariate variable selection.

```r
BikeSharingInDay <- read.csv(file = "C:/Users/hugo1/Documents/MA575/Proj/day.csv")
# import dataset day.csv

casual <- as.numeric(unlist(BikeSharingInDay['casual']))
# choose counts of casual users as response variable (Y)

atemp <- as.numeric(unlist(BikeSharingInDay['atemp']))
# choose normalized temperature in Celsius as covariate (X)
```

---

2. Plot Y vx X (i.e. a scatterplot) from the data.

```r
plot(atemp, casual, xlab = 'Feeling Temperature', ylab = 'Counts of Casual Users')
```
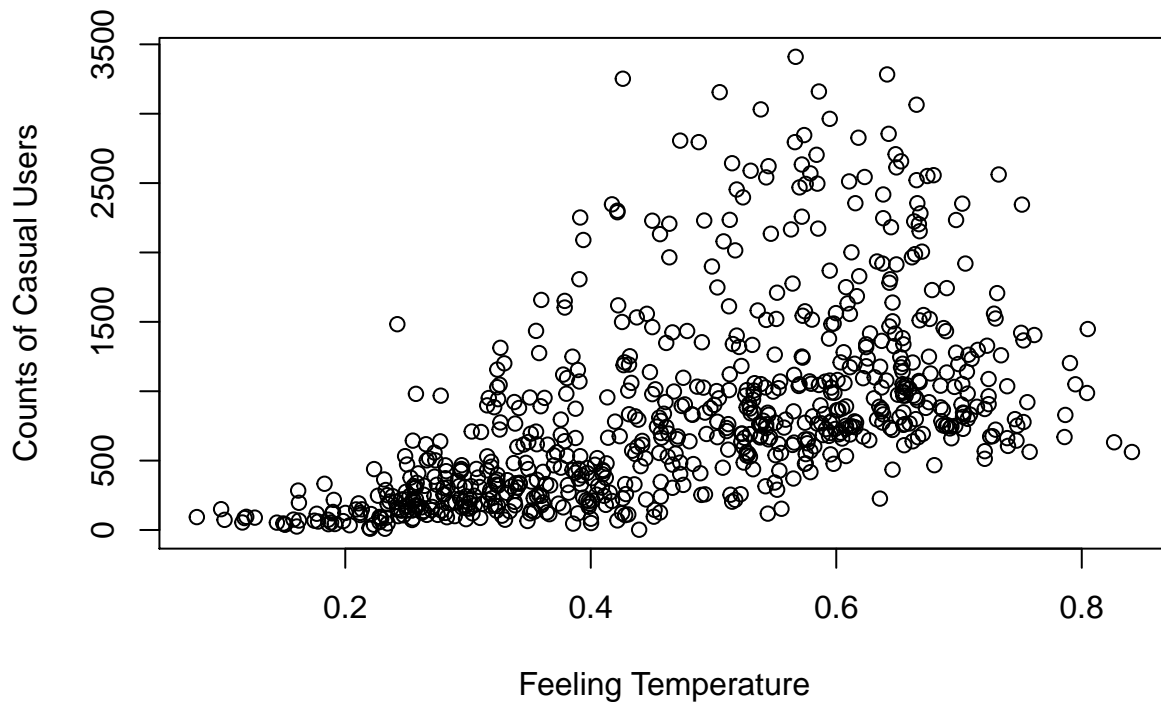


```r
# scatterplot of casual vs atemp
```

3&4. Perform OLS using R on your response and covariate & Submit the output from R of OLS (coefficient estimate, t values, p values) and interpret the results.

```
OLS_Casual_Atemp <- lm(casual~atemp)
# OLS on casual vs atemp

summary(OLS_Casual_Atemp)

##
## Call:
## lm(formula = casual ~ atemp)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1126.1  -343.9  -142.9   148.3  2514.3
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  -238.82      65.68  -3.636 0.000296 ***
## atemp        2291.52     130.95  17.499  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 576.6 on 729 degrees of freedom
## Multiple R-squared:  0.2958, Adjusted R-squared:  0.2948
## F-statistic: 306.2 on 1 and 729 DF,  p-value: < 2.2e-16
# summary of this OLS
```
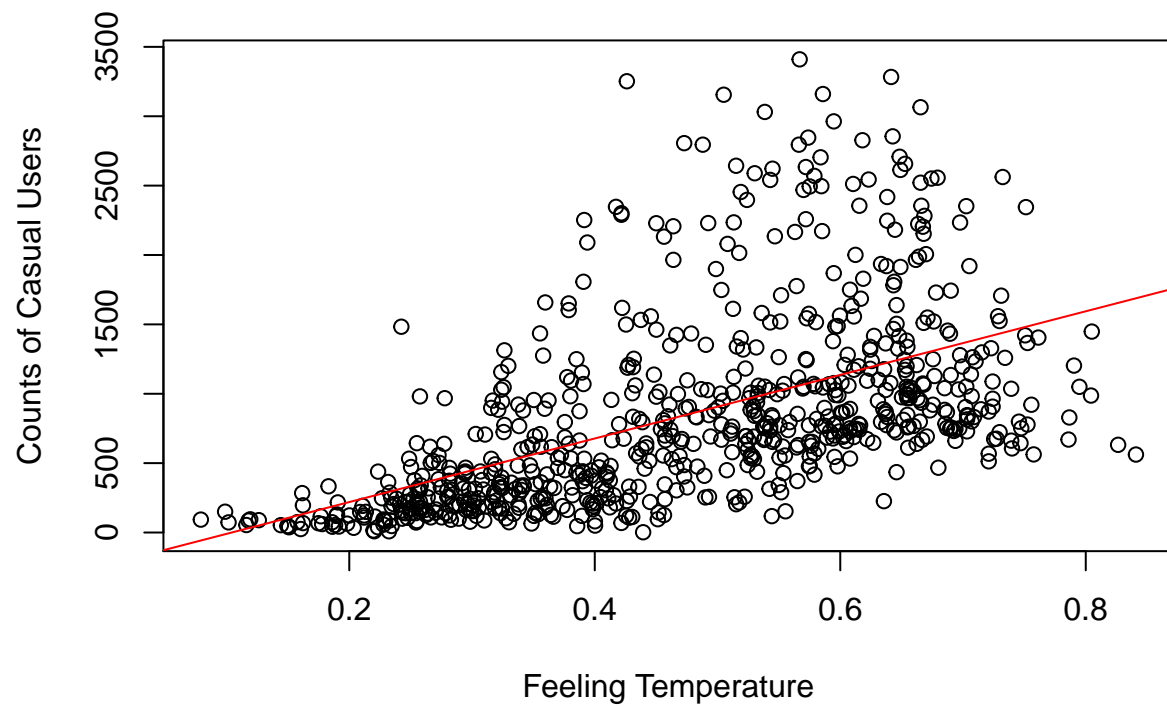
---

5. Plot also Y vs X for your choice of data set and overlay on your plot the linear regression fit obtained from R.

```
plot(atemp, casual , xlab = 'Feeling Temperature', ylab = 'Counts of Casual Users')
# scatterplot

abline(lsfit(atemp, casual), col = 'red')
```

```
# linear regression fit obtained from R
```