R Notebook

Installing Libraries

```
library("pacman")
```

```
## Warning: package 'pacman' was built under R version 4.0.4
```

```
pacman::p_load(lubridate, ggplot2, zoo, dplyr, caTools, Hmisc, caret)
```

Reading Data Files

```
solarpred <- read.csv('SolarPrediction.csv')
str(solarpred, vec.len = 1)
```

```
## 'data.frame':    32686 obs. of  11 variables:
##  $ UNIXTime              : int  1475229326 1475229023 ...
##  $ Data                  : chr  "9/29/2016 12:00:00 AM" ...
##  $ Time                  : chr  "23:55:26" ...
##  $ Radiation             : num  1.21 1.21 ...
##  $ Temperature           : int  48 48 ...
##  $ Pressure              : num  30.5 ...
##  $ Humidity              : int  59 58 ...
##  $ WindDirection.Degrees.: num  177 ...
##  $ Speed                 : num  5.62 3.37 ...
##  $ TimeSunRise           : chr  "06:13:00" ...
##  $ TimeSunSet            : chr  "18:13:00" ...
```

```
head(solarpred)
```

```
##     UNIXTime                  Data     Time Radiation Temperature Pressure
## 1 1475229326 9/29/2016 12:00:00 AM 23:55:26      1.21          48    30.46
## 2 1475229023 9/29/2016 12:00:00 AM 23:50:23      1.21          48    30.46
## 3 1475228726 9/29/2016 12:00:00 AM 23:45:26      1.23          48    30.46
## 4 1475228421 9/29/2016 12:00:00 AM 23:40:21      1.21          48    30.46
## 5 1475228124 9/29/2016 12:00:00 AM 23:35:24      1.17          48    30.46
## 6 1475227824 9/29/2016 12:00:00 AM 23:30:24      1.21          48    30.46
##   Humidity WindDirection.Degrees. Speed TimeSunRise TimeSunSet
## 1       59                177.39  5.62    06:13:00   18:13:00
## 2       58                176.78  3.37    06:13:00   18:13:00
## 3       57                158.75  3.37    06:13:00   18:13:00
## 4       60                137.71  3.37    06:13:00   18:13:00
## 5       62                104.95  5.62    06:13:00   18:13:00
## 6       64                120.20  5.62    06:13:00   18:13:00
```

Creating Month Variable

```
solarpred$Month = month(solarpred$Data)
solarpred$Month = as.factor(solarpred$Month)
solarpred$Month = factor(solarpred$Month ,labels = c("September", "October", "November", "December"))
```

Removing Unwanted Data Columns

```
solarpred$UNIXTime = NULL
solarpred$TimeSunRise = NULL
solarpred$TimeSunSet = NULL
solarpred$Pressure = NULL
```

Basic Visualisations

```
with(solarpred, tapply(Radiation,Month, mean))
```

```
## September   October  November  December
##  404.7526  391.0416  384.8184  240.0325
```

```
with(solarpred, tapply(Temperature,Month, mean))
```

```
## September   October  November  December
##  56.91637  54.99827  53.78501  49.44946
```

```
with(solarpred, round(tapply(Radiation, Time, mean ),2))
```

```
##      5      6      7      8      9     10     11     12     13     14     15
##   1.36   9.81 128.77 370.39 550.66 681.41 713.70 727.47 649.52 515.22 370.72
##     16     17     18
## 208.74  54.61   2.97
```

Creating Lag term

```
solarpred$RadiationLag = Lag(solarpred$Radiation,-10)
```

Create Model

```
model1 <- lm(Radiation ~ Temperature + RadiationLag, data = train)
summary(model1)
```

```
##
## Call:
## lm(formula = Radiation ~ Temperature + RadiationLag, data = train)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -979.38  -99.49  -14.33  118.67 1033.86
##
```

```
## Coefficients:
##                Estimate Std. Error t value Pr(>|t|)
## (Intercept)  -4.622e+02  1.815e+01  -25.47   <2e-16 ***
## Temperature   1.062e+01  3.676e-01   28.89   <2e-16 ***
## RadiationLag  6.935e-01  6.831e-03  101.52   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 183.8 on 12687 degrees of freedom
##   (7 observations deleted due to missingness)
## Multiple R-squared:  0.7167, Adjusted R-squared:  0.7167
## F-statistic: 1.605e+04 on 2 and 12687 DF,  p-value: < 2.2e-16
```

Predicting Results !

```
prediction <- predict(model1, newdata = test)
```

```
sse <- sum((fitted(model1) - solarpred$Radiation)^2)
```

```
## Warning in fitted(model1) - solarpred$Radiation: longer object length is not a
## multiple of shorter object length
```