Week 1 Assignment Question 1

December 19, 2016

Question 1

Make a plot that answers the question: what is the relationship between mean covered charges (Average.Covered.Charges) and mean total payments (Average.Total.Payments) in New York?

1. Load Data

```
mypath <- "./data/week1_payments.csv"
mydf <- read.csv(mypath)
dim(mydf)
```

```
## [1] 6401    12
```

```
names(mydf)
```

```
##  [1] "DRG.Definition"
##  [2] "Provider.Id"
##  [3] "Provider.Name"
##  [4] "Provider.Street.Address"
##  [5] "Provider.City"
##  [6] "Provider.State"
##  [7] "Provider.Zip.Code"
##  [8] "Hospital.Referral.Region.Description"
##  [9] "Total.Discharges"
## [10] "Average.Covered.Charges"
## [11] "Average.Total.Payments"
## [12] "Average.Medicare.Payments"
```

2. Subset Data

```
summary(mydf$Provider.State)
```

```
##   CA   FL   IL   NY   PA   TX
## 1514  972  743  933  808 1431
```

```
mydf2 <- mydf[mydf$Provider.State == "NY", ]
dim(mydf2)
```

```
## [1] 933   12
```

3. Plot

```
plot(mydf2$Average.Covered.Charges, mydf2$Average.Total.Payments,
     pch = 1, cex = 0.5,
     main = "Plot 1 Relationship between Mean Covered Charges
             and Mean Total Payments in New York",
```

```
    xlab = "Mean Covered Charges",
    ylab = "Mean Total Payments")
abline(lm(data = mydf2,
          formula = Average.Total.Payments~Average.Covered.Charges),
       col = "green")
```

**Plot 1 Relationship between Mean Covered Charges
and Mean Total Payments in New York**