

# BA - Assignment 2 [Updated]

Prakkas

23 septembre 2015

```
#Statement 1
citidata <- read.csv("~/BA/CitiBike Data 1h.csv", header=TRUE, stringsAsFactors=FALSE)

dim(citidata)
```

```
## [1] 1158 15
```

```
mendata<-subset(citidata, citidata$gender==1)
pop<-citidata$tripduration
men<-mendata$tripduration

summary(pop)
```

```
##      Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  60.00   75.07   101.10   160.10   167.70   717.00
```

```
sd(pop)
```

```
## [1] 144.47
```

```
summary(men)
```

```
##      Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  60.00   77.66   110.80   182.30   202.00   717.00
```

```
sd(men)
```

```
## [1] 164.6802
```

*#Null hypothesis:men incur in overtimes fees which means men's trip duration is different than other users*

```
ttest<-(182.30-160.10)/(sd(men)/(sqrt(615)))
ttest
```

```
## [1] 3.343098
```

```
#t test value: 3.34 >1.96 then we can reject our null hypothesis
# at alpha=0.05 which means men dont incur in more overtimes fees with a confidence of 95%
```

```
#Statement 2
```

```
#Null hypothesis: subscribers incur in overtime fees
```

```
subscriberdata<-subset(citidata, citidata$usertype=="Subscriber")
subscriber<-subscriberdata$tripduration
dim(subscriberdata)
```

```
## [1] 784 15
```

```
summary(subscriber)
```

```
##      Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   60.00   77.27  109.70  180.10  201.10  717.00
```

```
sd(subscriber)
```

```
## [1] 161.0684
```

```
ttest2<-(mean(subscriber)-mean(pop))/(sd(subscriber)/sqrt(784))
ttest2
```

```
## [1] 3.466785
```

```
#ttest value:3.46>1.96 therefore we can reject our null hypothesis
# at alpha=0.05 which means subscribers dont incur in more overtimes fees with a confidence of 95%
```

```
#Statement 3
```

```
#Null hypothesis: Management should not be concerned about changes in variance
```

```
ttest3<-((200-160.1)/(185/sqrt(200)))
ttest3
```

```
## [1] 3.050115
```

```
#T test value=3.05> 1.96 then we can reject our hypothesis.
```