IST-687 Chapter Notes Template: After Completing Please Submit as a PDF.
Originality Assertion: By submitting this file you affirm that this writing is your own.

Name: Yicun Deng
Class: M003
Date: 9/29/20
Chapter Number: #12
Title of Chapter: Pictures Versus Numbers

Module 6 Lectures - Part 1

This class is mainly focusing on making the data visualize, for example, hist()
Unlike information visualization, science visualization has more physical help and cleared labeled data. Visualization can improve the data's human readable a lot.

Module 6 - Roundtable 1

Bar Chart-Oil Chart Oil chart is better.
Bar Chart is more readable but Oil chart contains more information. And sometimes, visualization can be misleading by its racial and size.
As a data scientist, we must make sure that our visualized data is not misleading. Different people would have different perspect of viewing the visualized information. Don't expect people to feel the same for one visualization.
10 principles:
1. Simplicity
2. Encoding
3. Patterns vs Details(show different perspect of the information)
4. Rnages
5. Transformation
6. Show density
7. Connection when missing data
8. Aggregates
9. Comparison
10. Colors

Module 6 - Roundtable 2

Its all about what you want to extract from the data.
I feel the man chart is hard to read too, feels like the focus is mimic the data into composition of human.
A good visualization always contain information while remaining simple to read.

IST-687 Chapter Notes Template: After Completing Please Submit as a PDF.
Originality Assertion: By submitting this file you affirm that this writing is your own.

Module 6 Lectures - Part 2 - GGplot

Module 6 Lectures - Part 2 - GGplot

package GGPLOT2

Plot = Data + Aesthetics + Geometry

```
ggplot(data, aes(x=, y=))+geom_histogram(binwidth=,color=”black”,
fill=”white”)
geom_plot()
```

aes(x=, group=, color=)

Module 6 - Histograms and Line Graphs with GGplot

```
mtc<-mtcars
hist(mt$mpg, breaks=4)
g<- ggplot(mtc, aes(x=mpg))+geom_histogram(bins=5, color =
“black”,fill=”white”)
g+ggtitle(“mpg buckets”)
g<- ggplot(mtc, aes(x=mpg))+geom_histogram(binwidth=10, color =
“black”,fill=”white”)
ggplot(df, aes(x=day, y=timeToNYC,group=1))+geom_line(color=”red”,
linetype=”dashed”,size=1.5)
g+geom_point(color=”blue”,size=4)+ylab(“time to NYC(in hours)”)
geom_point()set point geom_line() set line
ggplot(df, aes(x=dayOfWeek, group=week,
color=week))+geom_line(aes(y=time))+ylab()+ggtitle
```

Module 6 - Bar and Scatter Plots in GGplot

Module 6 - Bar and Scatter Plots in GGplot

```
ggplot(mtcars, aes(x=factor(0),mpg))+geom_boxplot()
ggplot(mtcars, aes(group=cyl, mpg))+geom_boxplot()+coord_flip()

ggplot(mtcars,aes(x=cyl))+geom_bar()

car.names <-rownames(mtc)
ggplot(mtc,
aes(x=car.names,y=wt))+geom_bar(stat=”identity”)+theme(axis.text,x=element_t
ext(angle=90,hjust=1))
```

ggplot(mtc, aes(x=cyl, fill=factor(gear))) + geo_bar(position=”dodge”)

ggplot(mtc,aes(x=mpg,
y=wt))+geom_point(aes(size=qsec),color=qsec)+geom_text(aes(label=car.name),size=3)

ggplot(mtc,aes(x=mpg,y=reorder(car.names,mpg)))+geom_point(size=3)

IST-687 Chapter Notes Template: After Completing Please Submit as a PDF.
Originality Assertion: By submitting this file you affirm that this writing is your own.

Exercise Review
Step1. Add the region into dfStates
Step2. Install ggplot
Step3. Use ggplot()+geom_boxplot() which would show the mean

Question for Class
When is the group project start and when would we know which group we are in? Thank you!