library(ggplot2)

library(forecast)

library(treemap)

library(ggthemes)

library(ggfortify)

theme\_set(theme\_classic())

train<-read.csv(file="C:/Users/Pooja/Documents/SPL\_traindata.csv")

train

dim(train)

table(is.na(train))

colSums(is.na(train))

mean(train$Item\_Weight, na.rm=TRUE)

mean<-mean(train$Item\_Weight, na.rm=TRUE)

mean

train$Item\_Weight.mean<-ifelse(is.na(train$Item\_Weight),mean,train$Item\_Weight)

dim(train)

table(is.na(train))

colSums(is.na(train))

#ggplot(data=train)+geom\_bar(mapping=aes(x=Outlet\_Type,fill=Outlet\_Location\_Type),alpha=0.2,position="dodge")

#ggplot(data=train)+geom\_point(mapping=aes(x=Outlet\_Type,y=Item\_Outlet\_Sales,shape=Outlet\_Location\_Type))

#ggplot(data=train)+geom\_jitter(mapping=aes(x=Outlet\_Type,fill=Outlet\_Location\_Type),alpha=0.2,position="dodge")

#TreeMap

treemap(train,index=c('Outlet\_Location\_Type','Outlet\_Type'),border.col=c('black','white'),vSize='Item\_Outlet\_Sales',

align.labels = list(c('center','center'),c('left','top')),

overlap.labels = 0.5,

inflate.lables=F,

vColor='Outlet\_Location\_Type',

palette='Set3',

title='Sales Distribution in Different Tiers')

#Stacked\_Bar

ggplot(data=train)+

geom\_bar(mapping=aes(x=Item\_Type,fill=Outlet\_Type))+

theme(axis.text.x=element\_text(angle=90,vjust=0.7))+

labs(title="Item distribution in Different Outlets")+

scale\_fill\_brewer(palette="Accent")

#Pyramid\_Plot

ggplot(train,mapping=aes(x=Item\_Type,fill=Item\_Fat\_Content,y=ifelse(test = Item\_Fat\_Content=='Regular',yes=-Item\_Outlet\_Sales,no=Item\_Outlet\_Sales)))+

geom\_bar(stat="identity")+

scale\_y\_continuous(labels=abs,limits=max(train$Item\_Outlet\_Sales)\*c(-1,1))+

labs(y="Item\_Sales",title="Fat Contents in Different Items")+coord\_flip()+scale\_fill\_brewer(palette="Dark2")