examine the impact of:

gender of a passanger on survival

embarkation point on survival

cabin on survival

fare level on survival

do both visual exploration and hypothesis testing and submit your results on IVLE

Task-2

take any one dataset of your choice

form 5-8 questions and get the answers through visual (or non-visual) data exploration

one of the many resources on internet:

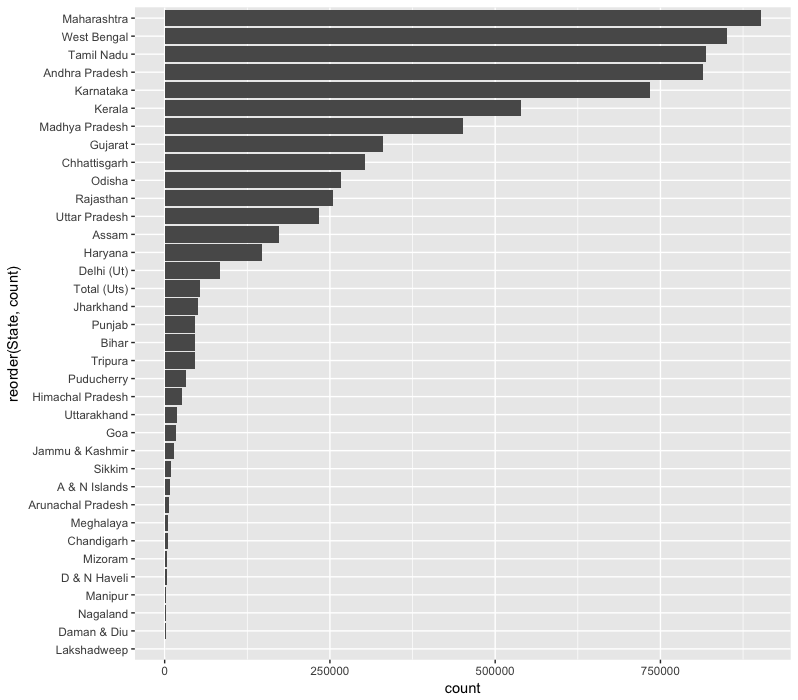
* 1. Which state has highest suicide rate?

data %>% group\_by(State) %>%

summarize(count=sum(Total)) %>%

ggplot(aes(x=reorder(State,count),y=count))+

geom\_bar(stat='identity') + coord\_flip()



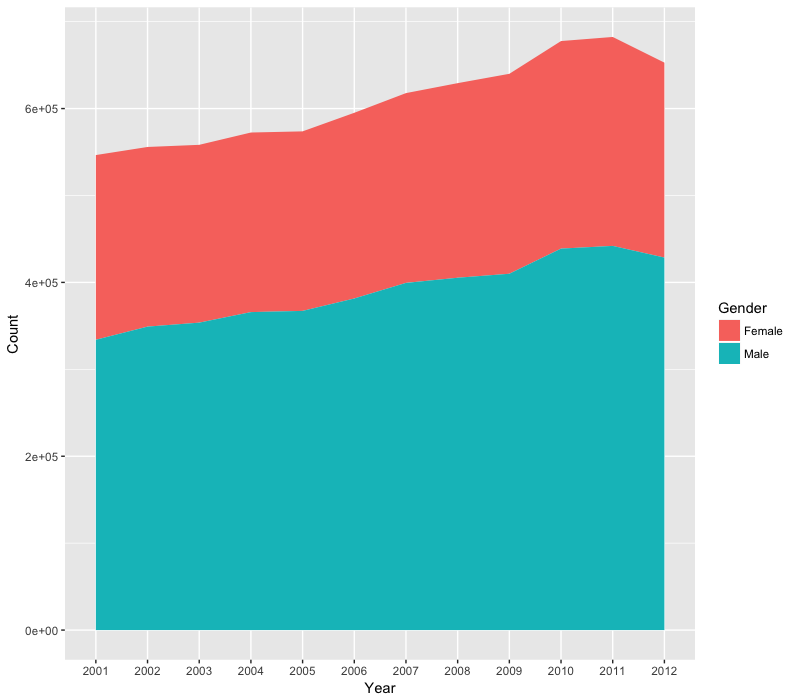
* 1. Is there any impact of Gender on suicide?

Code:

gper<-data %>% select(Gender,Total)%>% group\_by(Gender)%>% summarise(total\_all=sum(Total))%>%

transform(percent=round(total\_all/sum(total\_all)\*100))

ggplot(gper,aes(x=Gender,y=percent,fill=Gender))+geom\_bar(stat="identity")+geom\_text(aes(label=percent))



* 1. Most significant cause or reason of suicide for the year in which maximum suicides happened?

data %>% filter(Type\_code=="Causes" & !(Type=="Causes Not known") & !(Type=='Other Causes (Please Specity)'))%>% select(Year,Total,Type)%>%

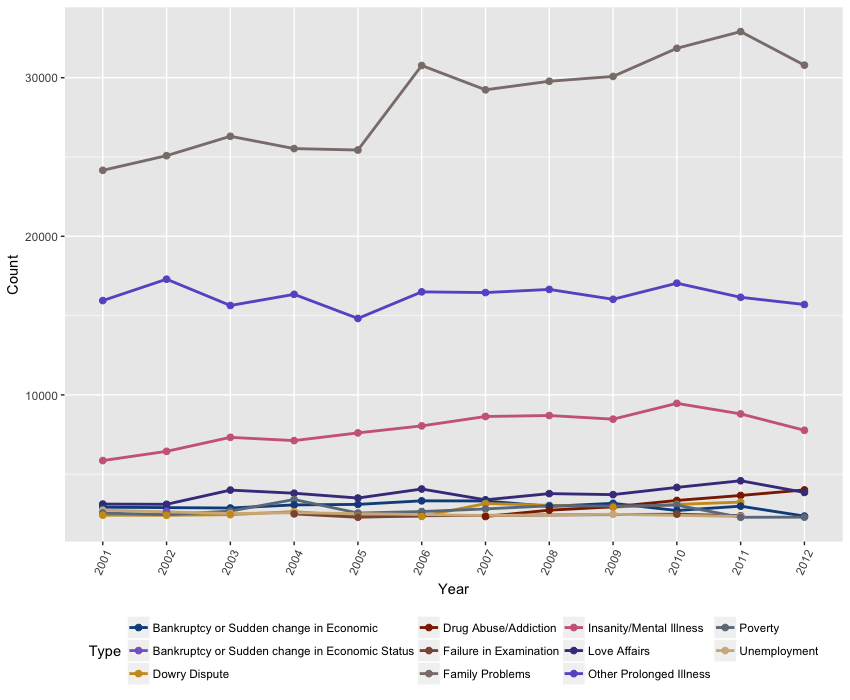
group\_by(Year,Type)%>%summarise(total=sum(Total))%>%arrange(desc(total))%>%head(100)%>%

ggplot(aes(x=factor(Year),y=total,color=Type,group=Type))+

geom\_line(size=1)+scale\_color\_manual(values=colr)+

theme(legend.position = "bottom",axis.text.x = element\_text(angle=65,vjust=0.5))+

labs(x="Year",y="Count")+geom\_point(size=2)



* 1. Impact of age on suicide?

**1st**

data %>%

filter(!Age\_group=="0-100+")%>%

group\_by(Age\_group)%>%

summarise(total =sum(Total))%>%

ggplot(aes(x=Age\_group,y=total,fill=Age\_group))+geom\_bar(stat="identity")

**2nd**

data%>%

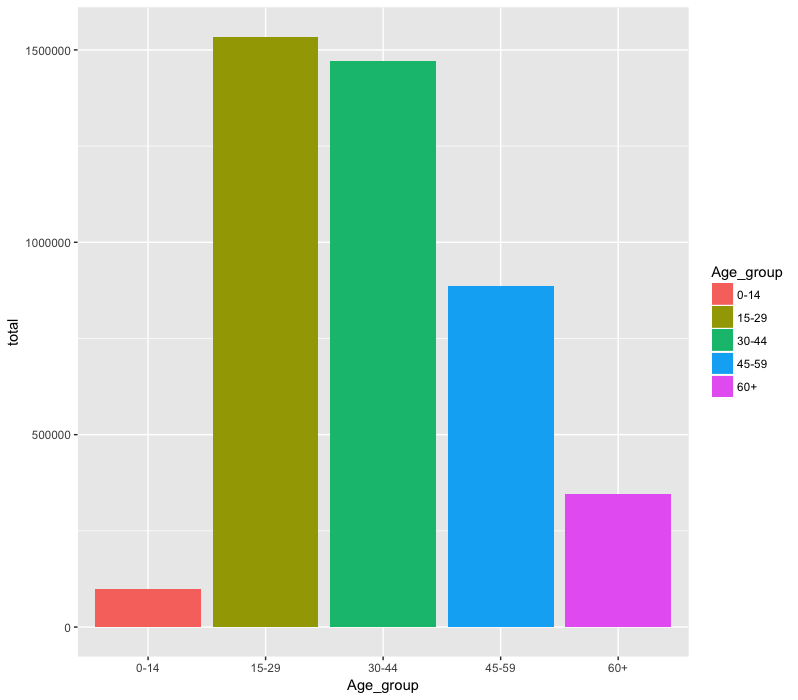
filter(!Age\_group=="0-100+")%>%

group\_by(Age\_group,Year)%>%

summarise(ptot=sum(Total))%>%

ggplot(aes(x=factor(Year),y=ptot,group=Age\_group,fill=Age\_group))+

geom\_area()+labs(x="Year",y="Count",fill="Age Group")



* 1. Common Means Adopted ?

set.seed(1234)

total<- as.data.frame( data %>% filter(Type\_code =="Means\_adopted")%>%

group\_by(Type)%>%

summarise(tot=sum(Total)))

wordcloud(words=total$Type, freq=total$tot, min.freq=25,max.freq=200)



6. Impact of Education level on Suicides.

data %>% filter(Type\_code =="Education\_Status")%>%

group\_by(Gender,Type)%>%

summarise(total=sum(Total))%>%

arrange(desc(total)) %>%

ggplot(aes(x=Type,y=total,fill=Type))+

geom\_bar(stat="identity")+scale\_fill\_manual(values=colr)+coord\_flip()

