```r
library(ggplot2)
library(readr)
library(dplyr)
library(stringr)
library(plotrix)
library(gridExtra)
library(RColorBrewer)


#Reading Data
sc_data <- read.csv("D:/MCA/MCA 1/R P/Analysis-of-Suicides-in-India-DataSet-
master/DataSet/Suicides in India 2001-2012.csv")
head(sc_data,10)
#check for missing values in the dataset
is_na <- is.na(sc_data)
summary(is_na)

#Data Summary
str(sc_data)
summary(sc_data)

#_____#
#Start Analysis

##################################################################

#Which gender makes more suicide attempts?
gper<-sc_data %>% select(Gender,Total)%>% group_by(Gender)%>%
summarise(total_all=sum(Total))%>%mutate(rs=sum(total_all),
percent=round((total_all/rs)*100))
View(gper)
label <-  c( paste(gper$Gender[1],gper$percent[1],'%',sep=''),
paste(gper$Gender[2],gper$percent[2],'%',sep=''))
colr<-c("dodgerblue4","mediumpurple3","goldenrod3","orangered4",
        "lightsalmon4","mistyrose4","palevioletred3","slateblue4","slateblue",
"slategray4","tan")
#plot a bar graph
ggplot(gper,aes(x=Gender,y=percent,fill=Gender))+geom_bar(stat="identity")+sca
le_fill_manual(values=c("dodgerblue4","goldenrod3"))+geom_text(aes(label=perce
nt))
#plot a pie chart
pie3D(gper$percent,labels=label,labelcex=0.9,explode=0.15,col=colr)

##################################################################

#Which Age group makes more suicide attempts?
sc_data %>% select(Gender,Age_group,Total)%>% filter(!Age_group=="0-100+")%>%
group_by(Gender,Age_group)%>% summarise(atot=sum(Total))%>%
```

```r
ggplot(aes(x=Age_group,y=atot,fill=Gender))+geom_bar(stat="identity",position=
"dodge")+
  scale_fill_manual(values=c("dodgerblue4","goldenrod3"))+labs(y="Count")

pdata<-sc_data %>% select(Age_group,Total)%>% filter(!Age_group=="0-
100+")%>%group_by(Age_group)%>% summarise(atot=sum(Total))
pie3D(pdata$atot,labels=pdata$Age_group, labelcex=0.9, explode=0.1,col=colr,
main="Suicide and Age Groups ")

sc_data %>% select(Age_group,Total)%>% filter(!Age_group=="0-
100+")%>%group_by(Age_group)%>% summarise(atot=sum(Total))%>%
ggplot(aes(x=Age_group,y=atot,fill=Age_group))+geom_bar(stat="identity")+
  scale_fill_manual(values=colr)+labs(y="Count")

################################################################

#Education level and suicide
sc_type<-sc_data %>% filter(Type_code =="Education_Status")%>%
select(Gender,Total,Type)%>% group_by(Gender,Type)%>%
summarise(ttotal=sum(Total))
sc_type %>%
ggplot(aes(x=str_sub(Type,1,15),y=ttotal,fill=Type))+geom_boxplot()+scale_fill
_manual(values=colr)+theme(legend.position = "bottom",axis.text.x =
element_text(angle=90))+labs(x="Education Level",y="count")

################################################################

#Social Status of suicide Victims
ss_type<-sc_data %>% filter(Type_code =="Social_Status")%>%
select(Gender,Total,Type,Age_group)%>% group_by(Gender,Type,Age_group)%>%
summarise(ttotal=sum(Total))
ss_type%>%ggplot(aes(x=Type,y=ttotal,fill=Type))+geom_boxplot()+scale_fill_man
ual(values=colr)+theme(legend.position = "bottom",axis.text.x =
element_text(angle=90))+labs(x="Social Status" ,y="count")

################################################################

#Professional Status of suicide Victims

pp_type<-sc_data %>% filter(Type_code =="Professional_Profile")%>%
select(Gender,Total,Type,Age_group)%>% group_by(Gender,Type,Age_group)%>%
summarise(ttotal=sum(Total))
pp_type %>%
ggplot(aes(x=str_sub(Type,1,20),y=ttotal,fill=Gender))+geom_bar(stat="identity
",position="fill")+scale_fill_manual(values=colr)+theme(legend.position =
"bottom",axis.text.x = element_text(angle=90))+labs(x="Social
Status",y="Percent")
```

```r
pp_type%>%ggplot(aes(x=Type,y=ttotal,fill=Type))+geom_bar(stat="identity")+sca
le_fill_manual(values=colr)+
  theme(legend.position = "none",axis.text.x =
element_text(angle=90))+labs(x="Professional Profile" ,y="count")


###############################################################

##Means Adopted

ma_type<-sc_data %>% filter(Type_code =="Means_adopted")
%>%group_by(Type,Gender,Age_group)%>%summarize(mtot=sum(Total))
ma_type%>%ggplot(aes(x=Type,y=mtot,fill=Gender))+geom_bar(stat="identity",posi
tion="dodge")+scale_fill_manual(values=colr)+theme(legend.position =
"bottom",axis.text.x = element_text(angle=90))+labs(x="Means
Adopted",y="Count")
ma_type%>%ggplot(aes(x=Type,y=mtot,fill=Age_group))+geom_bar(stat="identity",
                                                      position="dodge")
+scale_fill_manual(values=colr)+theme(legend.position = "bottom",axis.text.x =
element_text(angle=90))+labs(x="Means Adopted",y="Count")


###############################################################

##Major Causes

ca_type<-sc_data %>% filter(Type_code =="Causes")%>%
group_by(Type,Gender,Age_group)%>%summarise(ctot=sum(Total))
ca_type%>%ggplot(aes(x=Type,y=ctot,color=Gender))+geom_point()+scale_color_man
ual(values=colr)+theme(legend.position = "bottom",axis.text.x =
element_text(angle=90,vjust=0.5))+labs(x="Causes",y="Count")


options(repr.plot.width=6, repr.plot.height=4)
p1<-ca_type%>% filter(Age_group =="0-14" & ctot!=0)%>%
ggplot(aes(x=Type,y=ctot,fill=Type))+geom_bar(stat="identity")+theme(legend.po
sition = "none",axis.text.x = element_text(angle=90,vjust=0.5))+labs(x="Major
causes-Age 0-14",y="Count")
p2<-ca_type%>% filter(Age_group =="15-29" & ctot!=0)%>%
ggplot(aes(x=Type,y=ctot,fill=Type))+geom_bar(stat="identity")+theme(legend.po
sition = "none",axis.text.x = element_text(angle=90,vjust=0.5))+labs(x="Major
causes-Age 15-29",y="Count")
p3<-ca_type%>% filter(Age_group =="30-44" & ctot!=0)%>%
ggplot(aes(x=Type,y=ctot,fill=Type))+geom_bar(stat="identity")+theme(legend.po
sition = "none",axis.text.x = element_text(angle=90,vjust=0.5))+labs(x="Major
causes-Age 30-44",y="Count")
p4<-ca_type%>% filter(Age_group =="45-59" & ctot!=0)%>%
ggplot(aes(x=Type,y=ctot,fill=Type))+geom_bar(stat="identity")+theme(legend.po
sition = "none",axis.text.x = element_text(angle=90,vjust=0.5))+labs(x="Major
causes-Age 45-59",y="Count")
```

```r
p5<-ca_type%>% filter(Age_group =="60+" & ctot!=0)%>%
ggplot(aes(x=Type,y=ctot,fill=Type))+geom_bar(stat="identity")+theme(legend.po
sition = "none",axis.text.x = element_text(angle=90,vjust=0.5))+labs(x="Major
causes-Age 60+",y="Count")
p1
p2
p3
p4
p5


##############################################################

#Analysing the trend for the major causes for the suicides
sc_data %>% filter(Type_code=="Causes" & Type %in% c("Failure in
Examination","Family Problems","Other Prolonged Illness","Unemployment","Dowry
Dispute","Poverty","Insanity/Mental Illness"))%>%select(Year,Total,Type)%>%
group_by(Year,Type)%>%summarise(ytot=sum(Total))%>%  ggplot(aes(x=factor(Year)
,y=ytot,color=Type,group=Type))+geom_line(size=1)+scale_color_manual(values=co
lr)+
  theme(legend.position = "bottom",axis.text.x =
element_text(angle=65,vjust=0.5))+labs(x="Year",y="Count")+geom_point(size=2)


##############################################################

###Sucicides per Year

sc_data$Year<-as.factor(sc_data$Year)
sc_data%>%select(Year,Total,Age_group) %>% group_by(Year,Age_group) %>%
summarise(total=sum(Total))%>%ggplot(aes(x=Year,y=total,group=Age_group,fill=A
ge_group))+
  geom_area()+scale_fill_manual(values=colr)


##############################################################

#Which state in India has got highest number of suicides.
sc_data%>%select(State,Year,Total) %>% group_by(State,Year)%>%
summarise(tot=sum(Total)) %>%
ggplot(aes(x=State,y=tot,fill=Year))+geom_bar(stat="identity")+
  theme(legend.position="bottom",axis.text.x=element_text(angle=90))


##############################################################

sc_data %>%filter(!State %in% c("Total (All India)","Total (States)","Total
(Uts)"))%>%ggplot(aes(x=factor(Year),y=State,fill=Total))+geom_tile()+facet_wr
ap(~Gender)+scale_fill_gradientn(colors=brewer.pal(5,"Reds"))+theme(axis.text.
x = element_text(angle=90))+labs(x="Year")


##############################################################
```

```r
#Finding the top 10 and bottom 10 states in india based on the number of
suicides.

options(repr.plot.width=6, repr.plot.height=4)
sc1<-sc_data%>%filter(!State %in% c("Total (All India)","Total
(States)","Total (Uts)"))%>% select(State,Year,Total) %>% group_by(State)%>%
summarise(tot=sum(Total)) %>%
arrange(desc(tot))%>%head(10)%>%  ggplot(aes(x=factor(State,level=State),y=tot
,color=State))+geom_point(size=4)+geom_segment(aes(xend=State,y=0,yend=tot),si
ze=2)+theme(legend.position="none",axis.text.x=element_text(angle=90))+scale_c
olor_manual(values=colr)+geom_text(aes(label=tot),vjust=0.3)+labs(x="State -
More Suicides")
sc2<-sc_data%>%filter(!State %in% c("Total (All India)","Total
(States)","Total (Uts)"))%>% select(State,Year,Total) %>% group_by(State)%>%
summarise(tot=sum(Total)) %>%
arrange(desc(tot))%>%tail(10)%>%   ggplot(aes(x=factor(State,level=State),y=to
t,color=State))+geom_point(size=4)+geom_segment(aes(xend=State,y=0,yend=tot),s
ize=2)+theme(legend.position="none",axis.text.x=element_text(angle=90))+scale_
color_manual(values=colr)+geom_text(aes(label=tot),vjust=0.3)+labs(x="State -
Less Suicides")
grid.arrange(sc1,sc2,nrow=2)
```