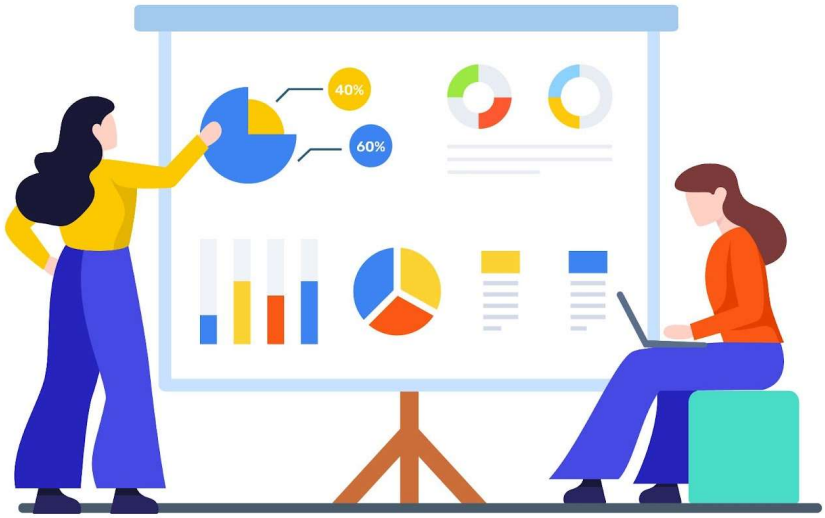


Exploratory Data Analysis Project Report

Arzoo Jangra (06)
Anitesh Minj (04)
Gaurav Kumar (60)

Submitted to :- Prof. Vasudha Bhatnagar

The Dataset

District Wise Crimes committed against women during 2001-2012

The dataset that we used in the project is taken from the government of India website : https://data.gov.in/resources/district-wise-crimes-committed-against-women-during-2001-2012

It was shared under the National Data Sharing and Accessibility Policy (NDSAP) and the major contributors were the Ministry of Home Affairs, Department of States, National Crime Records Bureau (NCRB).

DSTRCAW Dataset

STATE/UT	DISTRICT	Year	Rape	Kidnapping and Abduction	Dowry Deaths	Assault on women with intent to outrage her modesty	Insult to modesty of Women	Cruelty by Husband or his Relatives	Importation of Girls
ANDHRA PRADESH	ADILABAD	2001	50	30	16	149	34	175	0
ANDHRA PRADESH	ANANTAPUR	2001	23	30	7	118	24	154	0
ANDHRA PRADESH	CHITTOOR	2001	27	34	14	112	83	186	0
ANDHRA PRADESH	CUDDAPAH	2001	20	20	17	126	38	57	0
ANDHRA PRADESH	EAST GODAVARI	2001	23	26	12	109	58	247	0
ANDHRA PRADESH	GUNTAKAL RLY.	2001	0	0	0	1	0	0	0
ANDHRA PRADESH	GUNTUR	2001	54	51	7	139	129	378	0
ANDHRA PRADESH	HYDERABAD CITY	2001	37	39	24	118	27	746	0
ANDHRA PRADESH	KARIMNAGAR	2001	56	49	62	414	81	224	0
ANDHRA PRADESH	KHAMMAM	2001	47	30	17	180	336	172	0
ANDHRA PRADESH	KRISHNA	2001	37	21	10	208	72	265	0
ANDHRA PRADESH	KURNOOL	2001	29	47	13	141	107	92	0
ANDHRA PRADESH	MAHABOOBNAGAR	2001	59	27	14	176	41	69	0
ANDHRA PRADESH	MEDAK	2001	35	20	26	100	25	192	0
ANDHRA PRADESH	NALGONDA	2001	35	19	31	188	59	214	0
ANDHRA PRADESH	NELLORE	2001	46	80	10	207	228	287	0
ANDHRA PRADESH	NIZAMABAD	2001	21	21	19	55	15	228	0
ANDHRA PRADESH	PRAKASHAM	2001	19	12	5	140	100	119	0
ANDHRA PRADESH	RANGA REDDY	2001	72	83	37	113	55	421	7
ANDHRA PRADESH	SECUNDERABAD RLY	2001	0	0	1	0	1	0	0

Introduction to the Dataset

The dataset provides rich information about the crimes against women committed during the years 2001-2012.

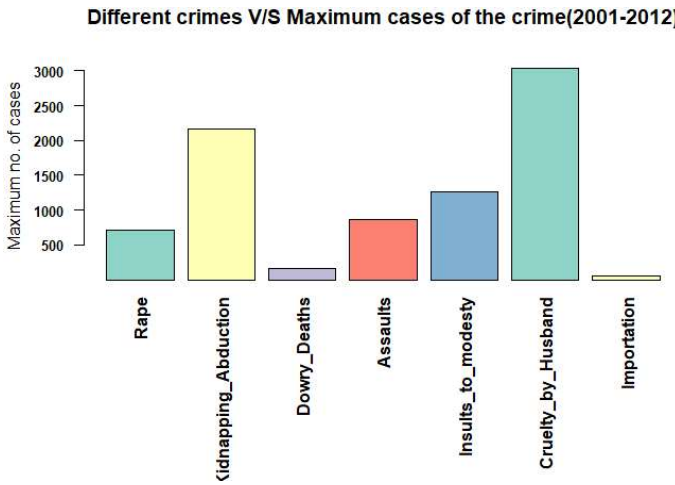
The data contains district wise details on crimes against women during the years 2001-2012 . The nature of such crimes includes **Rape**, **Kidnapping** and **Abduction**, **Dowry Death**, **Assault** on women with intent to outrage her modesty, **Insult to modesty** of Women, **Cruelty by Husband** or his Relatives and **Importation** of Girls. The districts are police districts and also include special police units. Therefore, these may be different from revenue districts.

We have first computed a few basic metrics for the data to understand the situation of crimes against women for the given years.

Then we set up some questions that striked us while observing this dataset and we tried to solve them. We have tried to analyze the data to the best of our knowledge and curiosity.

Bar graph for maximum of each crime in 2001-2012

```
63  #plotting maximum of each crime
64  par(mar=c(11,4,4,4))
65  barplot(height = y, name = crimes_names, col=coul3, ylab="Maximum no. of cases",
66          main="Different crimes V/S Maximum cases of the crime(2001-2012)",
67          las=2, cex.axis=0.8, cex.names=1,font.axis=2, ylim = c(50,3200))
```



Questions

The main questions that we framed and tried to solve are as follows :

Q1. Which year admitted the maximum number of crimes?

Q2 Compare the crimes year-wise and show which crime happened maximum in each year.

Q3 Examine which state witnessed the maximum no. of Dowry-Deaths in every particular year.

Q4. What can you say about the crime rate of the following 5 states/UT:

a)Delhi

b)Andhra Pradesh

c)West Bengal

d)Madhya Pradesh

e)Kerala

Did it increase or decrease over the period of time?

Q5. Give a 5 number summary for the different crimes in Bihar for the year it had the maximum total no. of crimes!

Q6. Compare the rates of Rapes and Kidnapping/Abduction in Rajasthan(district-wise) for the two years(take one year for which the total number of crimes was the maximum and other year for which the total number of crimes was minimum).

About the Project

This is an overview and a basic idea to our whole group project :

Platform Used :

We have implemented R Language and it's IDE called RStudio to analyze the data. It supports a variety of packages that we might need in data analysis while also being platform independent. We have also used it's various plotting techniques to get a visual view of the data and understand it better.

Loading the Dataset :

We loaded the dataset into the RStudio for analysis with the help of read.csv() function to a new dataframe. We then changed the names of some columns to make the dataframe look consistent. We also removed the extra "Total" row that was present after the rows of each state denoting the total of each crime in that state, because it might have caused hindrance to other queries.

Initial Steps :

After loading and organizing the data frame, a few basic metrics were computed such as the number of rows and columns, minimum and maximum of each individual crime in any state or year for the whole data, etc. We then proceeded to solve the previously mentioned questions.

Q1. Which year admitted the maximum number of crimes?

Ans:

Approach : To get the year in which most crimes were observed, we first calculated the total number of crimes(of all 7 types) for each year individually and then plotted a line graph to get our answer.

Variables Used : gtotal_2001 - Total number of crimes in the year 2001

gtotal_2002 - Total number of crimes in the year 2002

...and so on

years - A vector containing all years

totals - A vector containing the year-wise totals of all the crimes

Code Snippets :

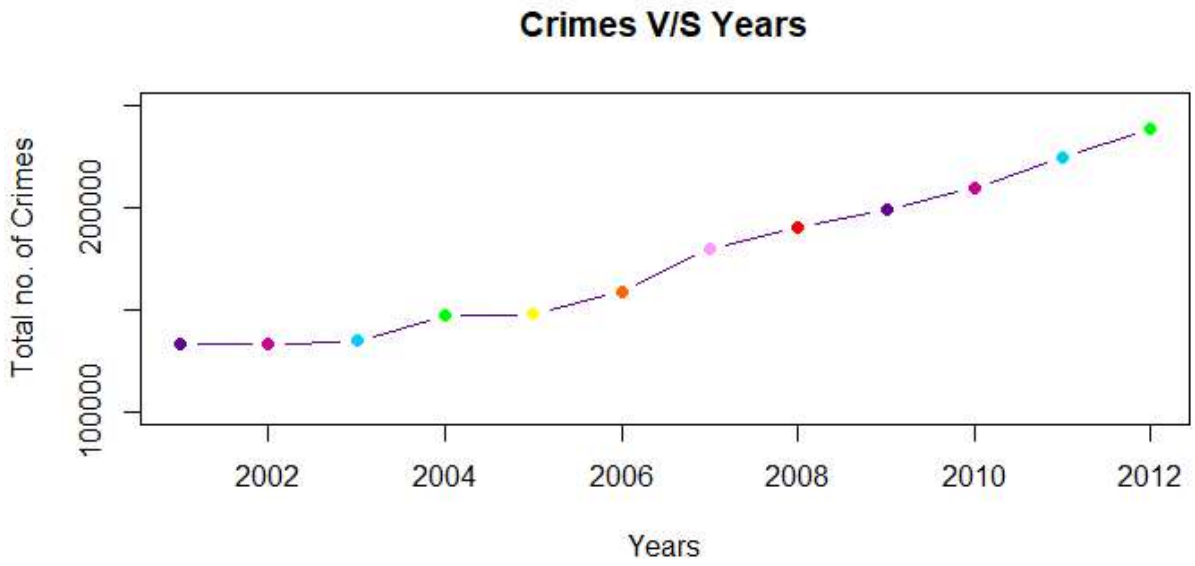
Calculating total number of crimes for each year

```
#finding the total number of crimes year-wise
gtotal_2001 <- sapply(crimes_2001[, 4:10], sum)
gtotal_2001
gtotal_2001 <- sum(gtotal_2001, na.rm = FALSE)
gtotal_2001    #total number of crimes recorded in 2001 : 130725

gtotal_2002 <- sapply(crimes_2002[, 4:10], sum)
gtotal_2002 <- sum(gtotal_2002, na.rm = FALSE)
gtotal_2002    #total number of crimes recorded in 2002 : 131112

gtotal_2003 <- sapply(crimes_2003[, 4:10], sum)
gtotal_2003 <- sum(gtotal_2003, na.rm = FALSE)
gtotal_2003    #total number of crimes recorded in 2003 : 131364
```

Plotting the line graph

```
#visualizing total crimes in each year
years <- c(2001, 2002, 2003, 2004, 2005, 2006, 2007, 2008, 2009, 2010, 2011, 2012)
years
totals <- c(gtotal_2001, gtotal_2002, gtotal_2003, gtotal_2004, gtotal_2005, gtotal_2006,
            gtotal_2007, gtotal_2008, gtotal_2009, gtotal_2010, gtotal_2011, gtotal_2012)
totals
plot(years, totals, type = "b", pch = 19, col = coul1, xlab = "Years",
     ylab = "Total no. of Crimes",main = "Crimes V/S Years", ylim = c(100000,250000), )
#Ques1 finished!! Ans: Year 2012 recorded the maximum number of crimes: 232528
```

The Plot :



Inference : The graph shows that the crimes against women are increasing slightly in the first few years but later, the crimes increased very frequently. And the graph concludes that the year 2012, which is the latest, witnessed the maximum number of crimes. This reflects that over time the crimes against women are increasing and that should spark a thought of concern in all of us.

Q2 Compare the crimes year-wise and show which crime happened maximum in each year.

Ans.

Approach : To find which crime happened the most in each year, we take the data frame which has the totals of each crime year -wise and state-wise and then plot a pie chart for it.

Variables Used :

total_2001 - Total number of crimes in the year 2001

total_2002 - Total number of crimes in the year 2002

… and so on

crime_total - A data frame containing the total number of crimes for different states for each year

STATE.UT	DISTRICT	Year	Rape	Kidnapping/Abduction	Dowry_Deaths	Assaults	Insults_to_modesty	Cruelty_by_Husband	Importation
WEST BENGAL	TOTAL	2001	709	695	265	954	48	3859	3
A & N ISLANDS	TOTAL	2001	3	2	0	19	1	9	0
CHANDIGARH	TOTAL	2001	18	50	3	24	15	36	0
D & N HAVELI	TOTAL	2001	6	2	0	7	0	4	0
DAMAN & DIU	TOTAL	2001	0	3	0	0	0	4	0
LAKSHADWEEP	TOTAL	2001	0	0	0	0	0	0	0
PUDUCHERRY	TOTAL	2001	9	3	1	35	27	3	0
ANDHRA PRADESH	TOTAL	2002	1002	854	449	3799	2024	7018	0
ARUNACHAL PRADESH	TOTAL	2002	38	38	0	68	2	13	0
ASSAM	TOTAL	2002	970	1276	70	984	7	1694	0
BIHAR	TOTAL	2002	1040	744	927	621	6	1577	38

total - stores the total cases of each crime in every particular year

max_crime - stores the maximum value of **total**

Max_crime_name - the crime whose total value is maximum in the **total**

Code Snippets :

Creating the data frame crime_total

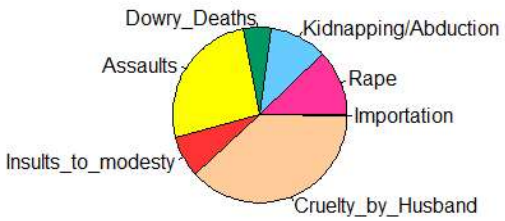
```
27  #storing state-wise total crimes in a variable
28  crime_total<-crimes_data[(crimes_data$DISTRICT=="TOTAL"),]
29  View(crime_total)
```

Storing total of each crime for a particular year in a variable and then plotting the pie chart

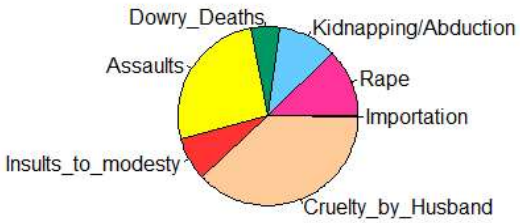
```
175 #storing the total of each crime for particular year in a variable
176 total_2001<-sapply(crime_total[crime_total$Year=='2001',][,4:10],sum)
177 pie(total_2001, main = "Crime pie chart 2001",col = coul5)
178 #Cruelty_by_Husband is the crime which was recorded for the maximum no. of times!
179 print(total_2001)
180
181 total_2002<-sapply(crime_total[crime_total$Year=='2002',][,4:10],sum)
182 pie(total_2002, main = "Crime pie chart 2002",col = coul5)
183 #Cruelty_by_Husband is the crime which was recorded for the maximum no. of times!
184 print(total_2002)
```

The Plots :

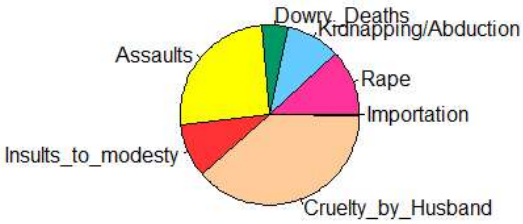
Crime pie chart 2001



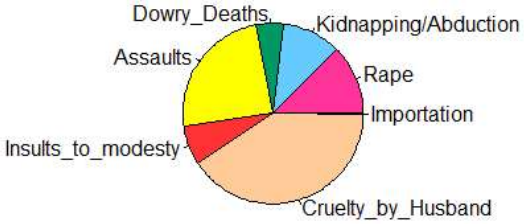
Crime pie chart 2002



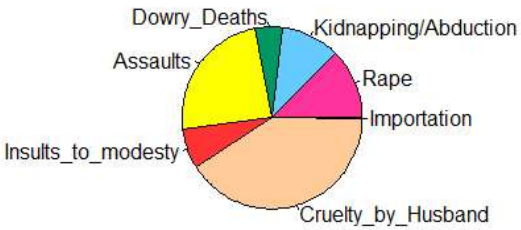
Crime pie chart 2003



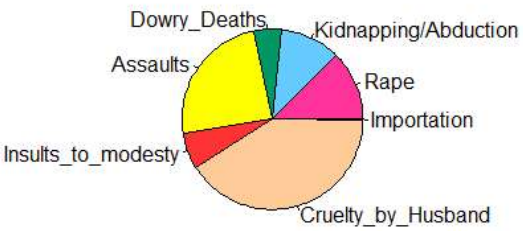
Crime pie chart 2004



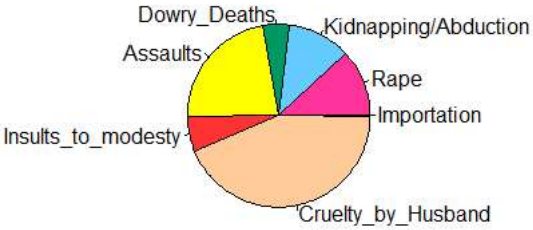
Crime pie chart 2005



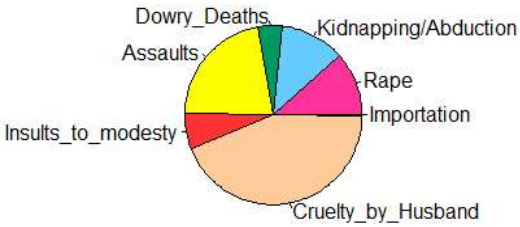
Crime pie chart 2006



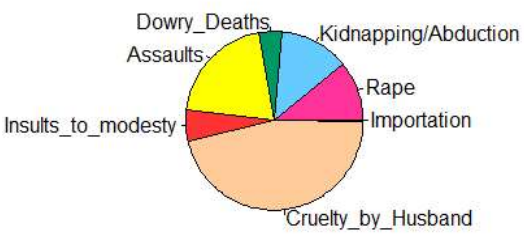
Crime pie chart 2007



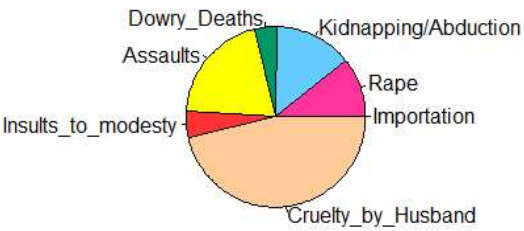
Crime pie chart 2008

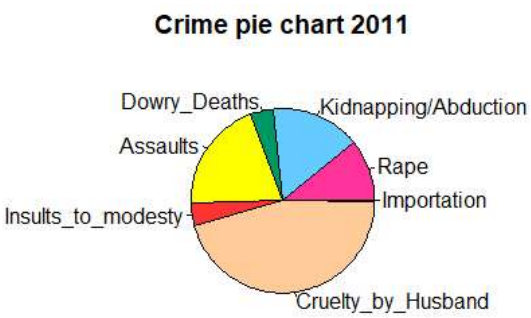


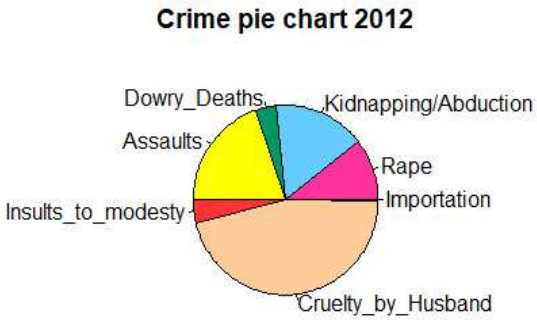
Crime pie chart 2009



Crime pie chart 2010







Inference : With the help of the pie charts we conclude that for all the years(2001-2012), the crime that occured the most is "Cruelty by Husband or his Relatives". This shows that in most cases of crimes related to women, women generally deal with cruelty after marriage.

Q3. Examine which state witnessed the maximum no. of Dowry-Deaths in every particular year.

Ans.

Approach : To solve this, we use the aggregate() function to compute the total number of Dowry related deaths for every individual year. We would get a list of states with the number of deaths related to dowry that occured there for a particular year. Then we plot a bar graph to actually observe which state had the most dowry-related deaths for each year.

Variables Used :

dowry_2001 - A list of state-wise dowry related deaths in the year 2001

dowry_2002 - A list of state-wise dowry related deaths in the year 2002

… and so on

states - Contains the names of all the states to be plotted on bar graph

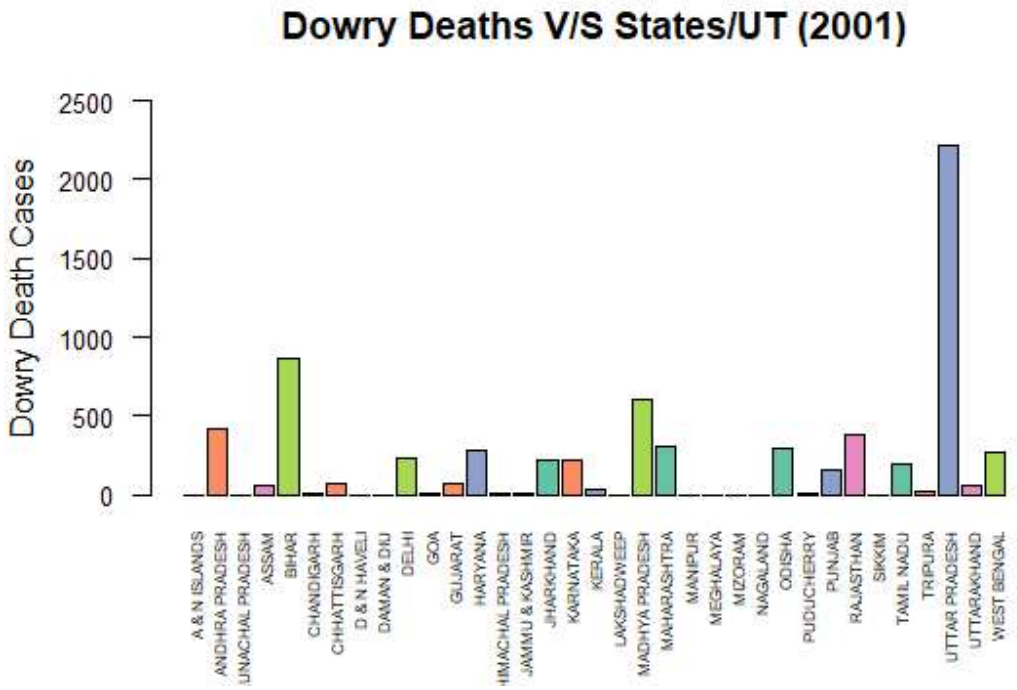
Code Snippets :

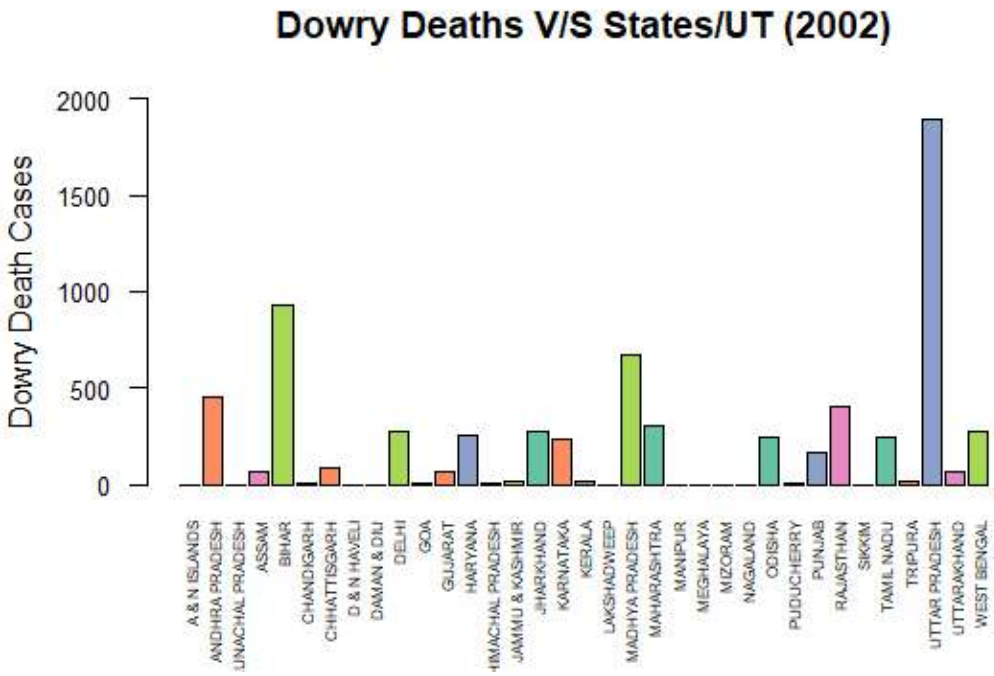
Calculating the sum of dowry related deaths for different states, one year at a time

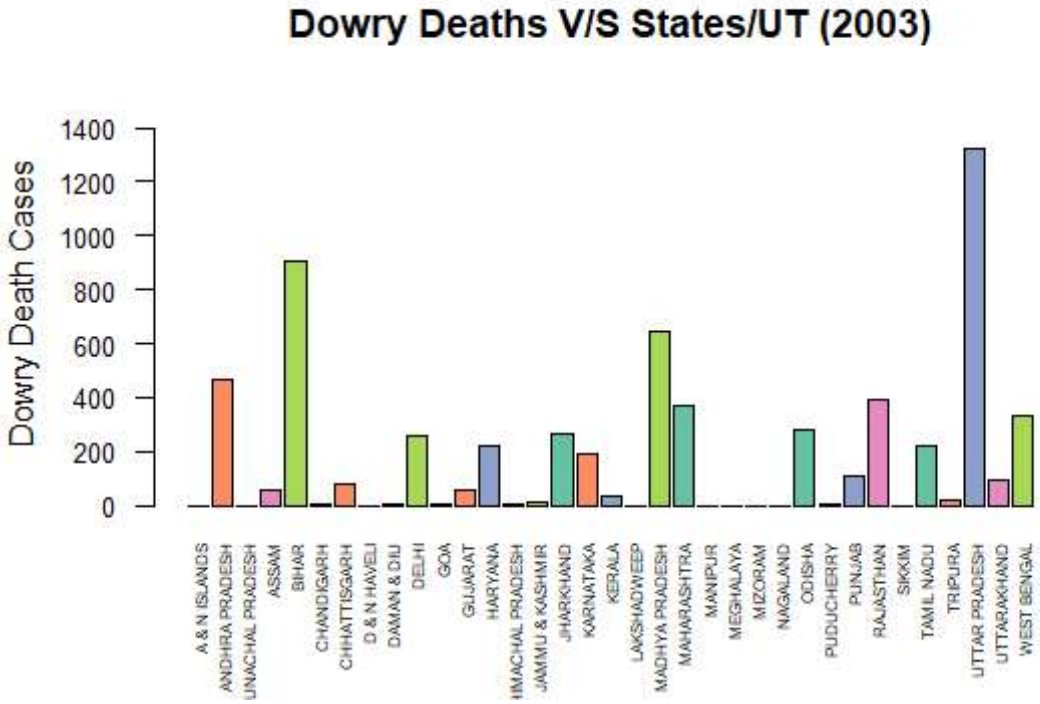
And then plotting the bar graphs

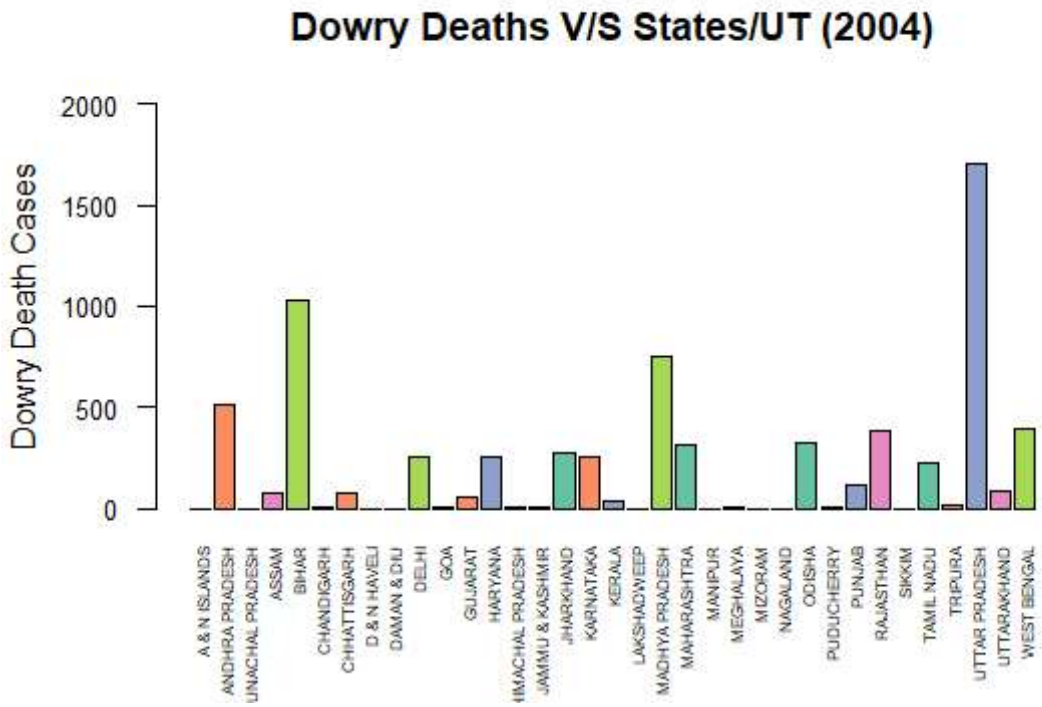
```
#For 2001
dowry_2001<-aggregate(x= crimes_2001$Dowry_Deaths,
                      by= list(crimes_2001$STATE.UT),
                      FUN=sum)
dowry_2001
states <- dowry_2001$Group.1
states
barplot(height = dowry_2001$x, name = states, col=coul4, ylab="Dowry Death Cases",
        main="Dowry Deaths V/S States/UT (2001)", las=2, cex.axis=0.8,
        cex.names=0.5, ylim = c(0,2500))


#For 2002
dowry_2002<-aggregate(x= crimes_2002$Dowry_Deaths,
                      by= list(crimes_2002$STATE.UT),
                      FUN=sum)
dowry_2002
states <- dowry_2002$Group.1
barplot(height = dowry_2002$x, name = states, col=coul4, ylab="Dowry Death Cases",
        main="Dowry Deaths V/S States/UT (2002)", las=2, cex.axis=0.8,
        cex.names=0.5, ylim = c(0,2000))
```

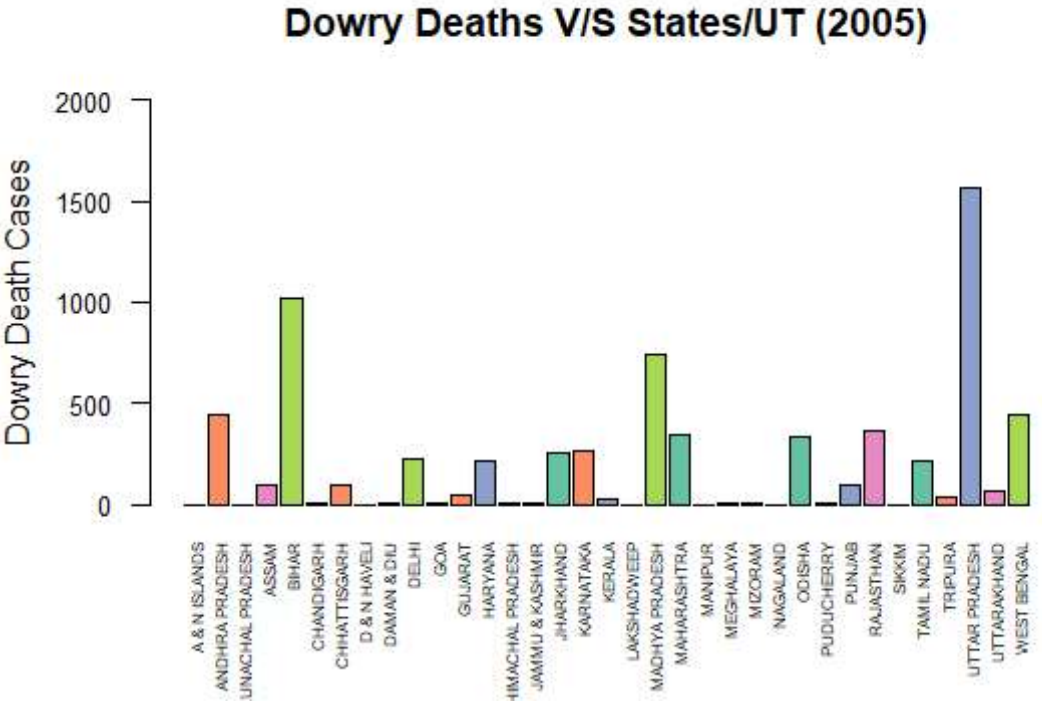
The Plots :

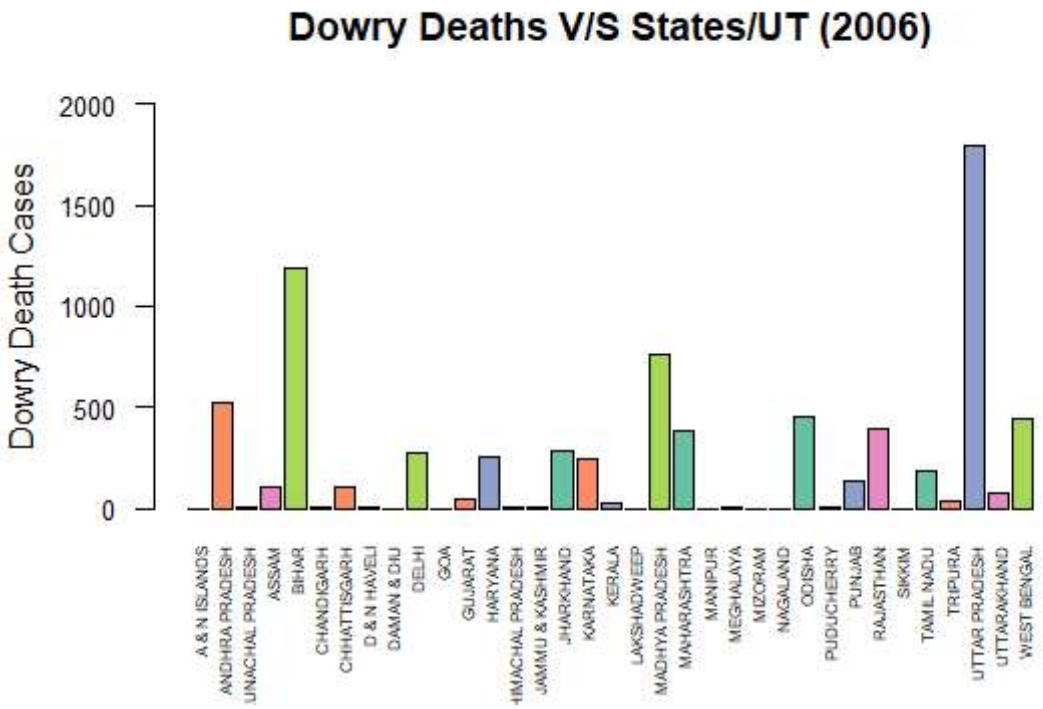


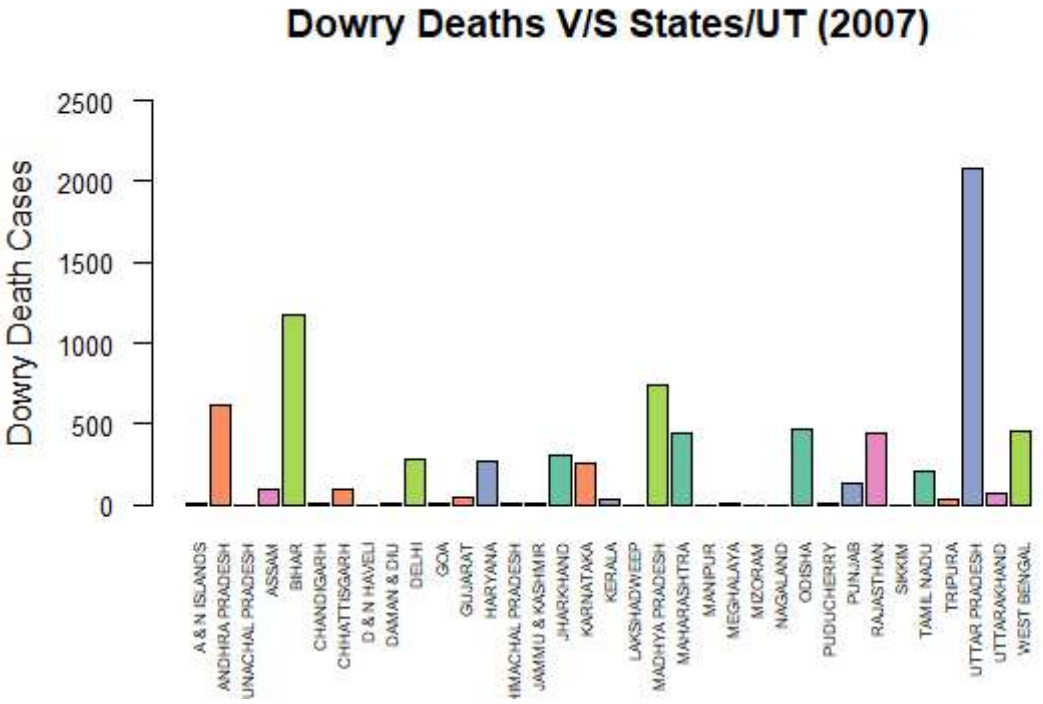

Dowry Deaths V/S States/UT (2002)
Dowry Death Cases
2000
1500
1000
500
0
A & N ISLANDS
ANDHRA PRADESH
ARUNACHAL PRADESH
ASSAM
BIHAR
CHANDIGARH
CHHATTISGARH
D & N HAVELI
DAMAN & DIU
DELHI
GOA
GUJARAT
HARYANA
HIMACHAL PRADESH
JAMMU & KASHMIR
JHARKHAND
KARNATAKA
KERALA
LAKSHADWEEP
MADHYA PRADESH
MAHARASHTRA
MANIPUR
MEGHALAYA
MIZORAM
NAGALAND
ODISHA
PUDUCHERRY
PUNJAB
RAJASTHAN
SIKKIM
TAMIL NADU
TRIPURA
UTTAR PRADESH
UTTARAKHAND
WEST BENGAL

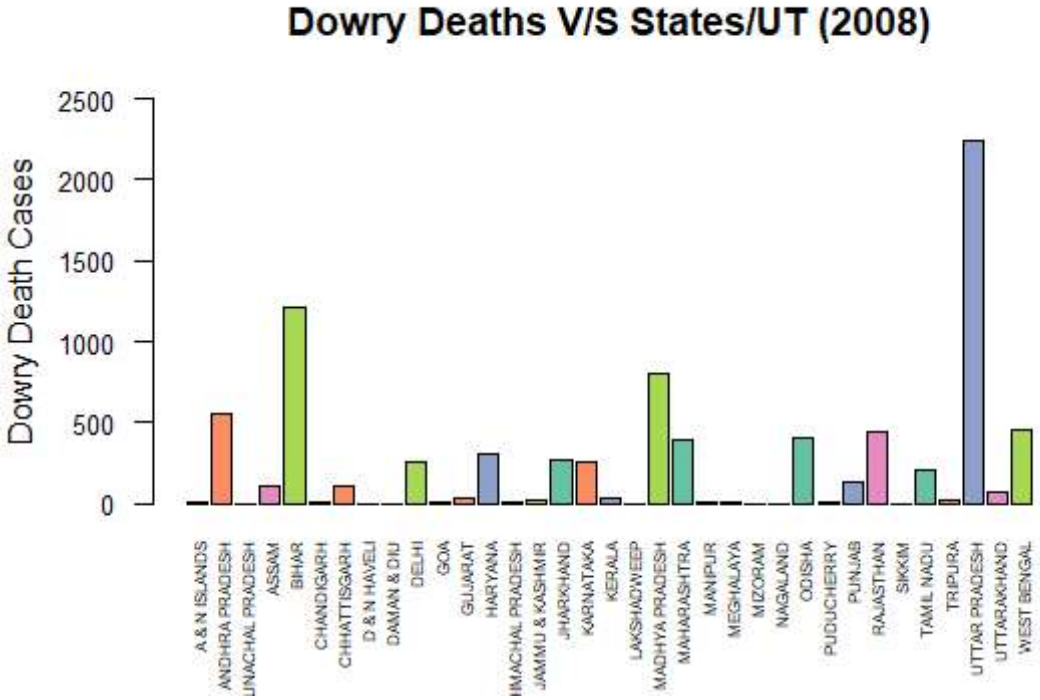

Dowry Deaths V/S States/UT (2003)
Dowry Death Cases
1400
1200
1000
800
600
400
200
0
A & N ISLANDS
ANDHRA PRADESH
ARUNACHAL PRADESH
ASSAM
BIHAR
CHANDIGARH
CHHATTISGARH
D & N HAVELI
DAMAN & DIU
DELHI
GOA
GUJARAT
HARYANA
HIMACHAL PRADESH
JAMMU & KASHMIR
JHARKHAND
KARNATAKA
KERALA
LAKSHADWEEP
MADHYA PRADESH
MAHARASHTRA
MANIPUR
MEGHALAYA
MIZORAM
NAGALAND
ODISHA
PUDUCHERRY
PUNJAB
RAJASTHAN
SIKKIM
TAMIL NADU
TRIPURA
UTTAR PRADESH
UTTARAKHAND
WEST BENGAL

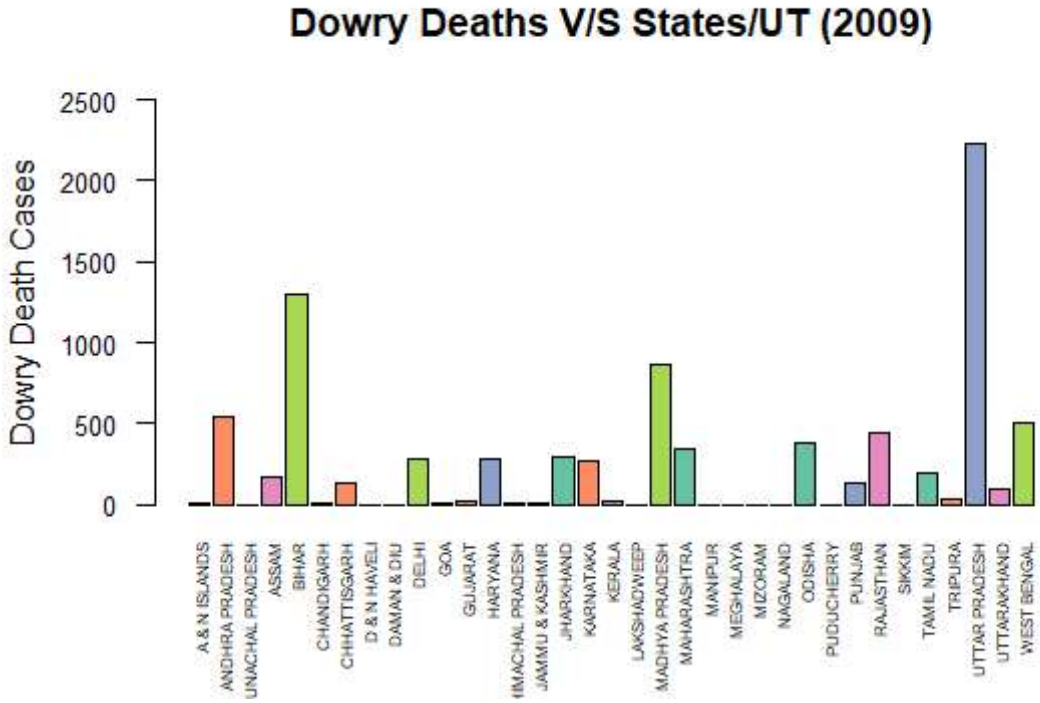

Dowry Deaths V/S States/UT (2004)
Dowry Death Cases
2000
1500
1000
500
0
A & N ISLANDS
ANDHRA PRADESH
UNACHAL PRADESH
ASSAM
BIHAR
CHANDIGARH
CHHATTISGARH
D & N HAVELI
DAMAN & DIU
DELHI
GOA
GUJARAT
HARYANA
HIMACHAL PRADESH
JAMMU & KASHMIR
JHARKHAND
KARNATAKA
KERALA
LAKSHADWEEP
MADHYA PRADESH
MAHARASHTRA
MANIPUR
MEGHALAYA
MIZORAM
NAGALAND
ODISHA
PUDUCHERRY
PUNJAB
RAJASTHAN
SIKKIM
TAMIL NADU
TRIPURA
UTTAR PRADESH
UTTARAKHAND
WEST BENGAL

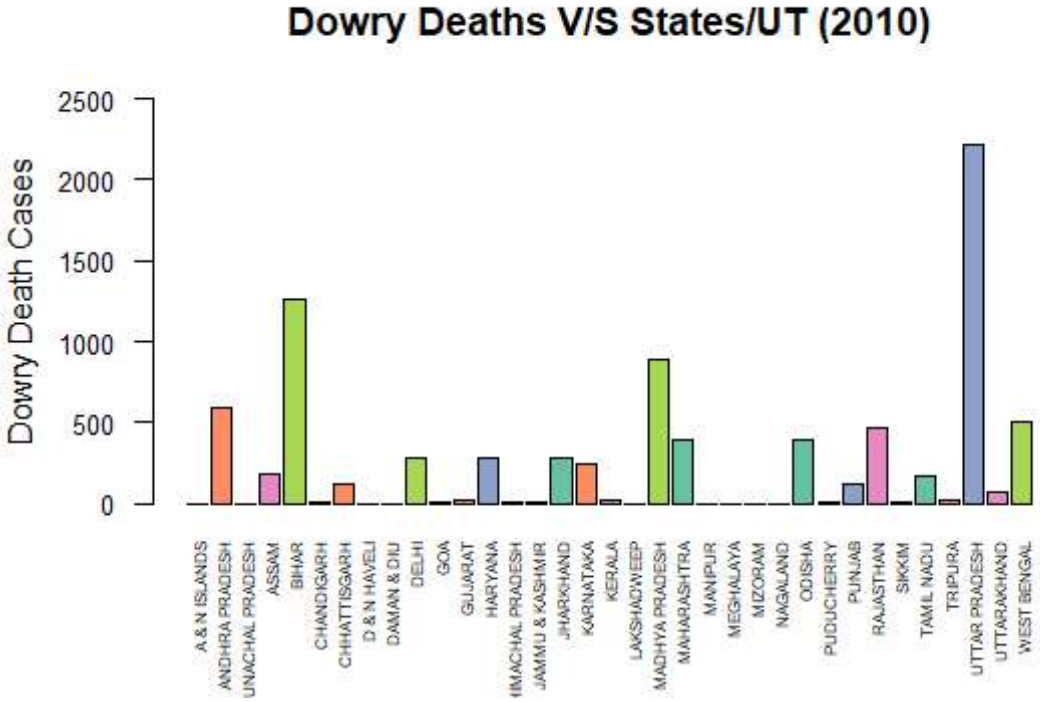

Dowry Deaths V/S States/UT (2005)
Dowry Death Cases
2000
1500
1000
500
0
A & N ISLANDS
ANDHRA PRADESH
UNACHAL PRADESH
ASSAM
BIHAR
CHANDIGARH
CHHATTISGARH
D & N HAVELI
DAMAN & DIU
DELHI
GOA
GUJARAT
HARYANA
HIMACHAL PRADESH
JAMMU & KASHMIR
JHARKHAND
KARNATAKA
KERALA
LAKSHADWEEP
MADHYA PRADESH
MAHARASHTRA
MANIPUR
MEGHALAYA
MIZORAM
NAGALAND
ODISHA
PUDUCHERRY
PUNJAB
RAJASTHAN
SIKKIM
TAMIL NADU
TRIPURA
UTTAR PRADESH
UTTARAKHAND
WEST BENGAL

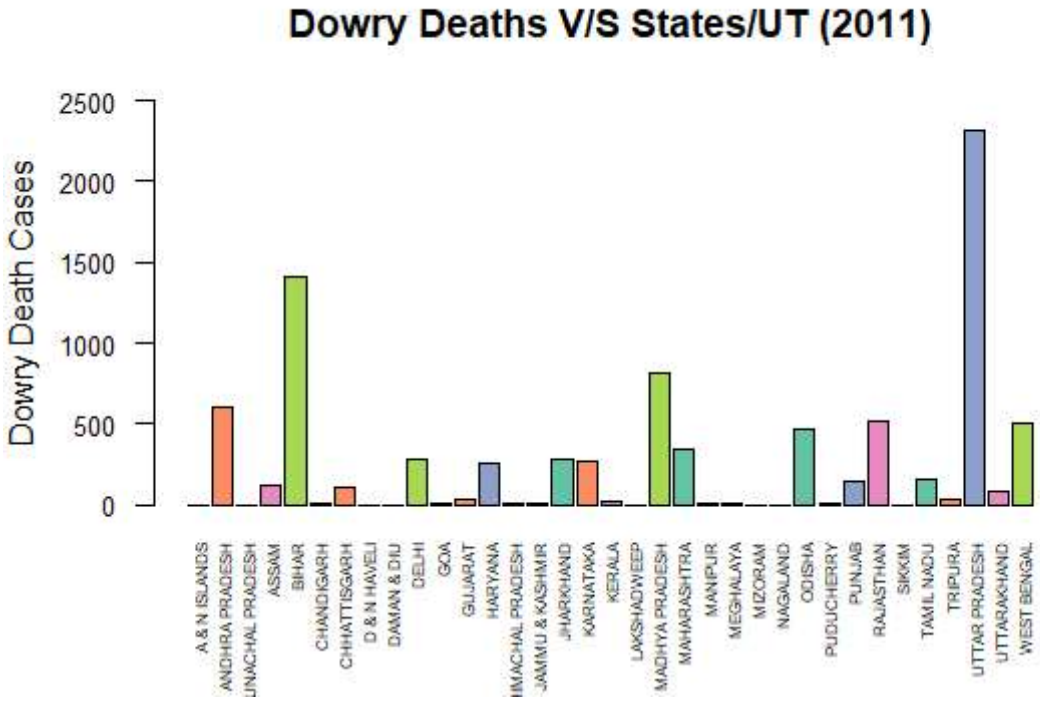

Dowry Deaths V/S States/UT (2006)
Dowry Death Cases
2000
1500
1000
500
0
A & N ISLANDS
ANDHRA PRADESH
ARUNACHAL PRADESH
ASSAM
BIHAR
CHANDIGARH
CHHATTISGARH
D & N HAVELI
DAMAN & DIU
DELHI
GOA
GUJARAT
HARYANA
HIMACHAL PRADESH
JAMMU & KASHMIR
JHARKHAND
KARNATAKA
KERALA
LAKSHADWEEP
MADHYA PRADESH
MAHARASHTRA
MANIPUR
MEGHALAYA
MIZORAM
NAGALAND
ODISHA
PUDUCHERRY
PUNJAB
RAJASTHAN
SIKKIM
TAMIL NADU
TRIPURA
UTTAR PRADESH
UTTARAKHAND
WEST BENGAL

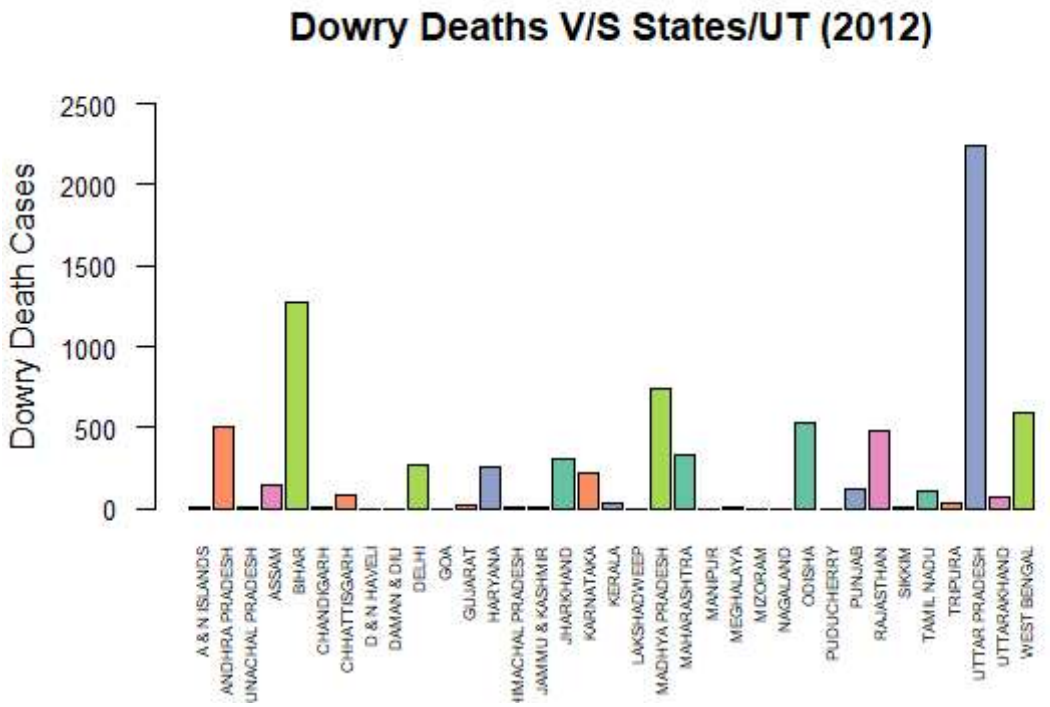

Dowry Deaths V/S States/UT (2007)
Dowry Death Cases
2500
2000
1500
1000
500
0
A & N ISLANDS
ANDHRA PRADESH
ARUNACHAL PRADESH
ASSAM
BIHAR
CHANDIGARH
CHHATTISGARH
D & N HAVELI
DAMAN & DIU
DELHI
GOA
GUJARAT
HARYANA
HIMACHAL PRADESH
JAMMU & KASHMIR
JHARKHAND
KARNATAKA
KERALA
LAKSHADWEEP
MADHYA PRADESH
MAHARASHTRA
MANIPUR
MEGHALAYA
MIZORAM
NAGALAND
ODISHA
PUDUCHERRY
PUNJAB
RAJASTHAN
SIKKIM
TAMIL NADU
TRIPURA
UTTAR PRADESH
UTTARAKHAND
WEST BENGAL


Dowry Deaths V/S States/UT (2008)
Dowry Death Cases
2500
2000
1500
1000
500
0
A & N ISLANDS
ANDHRA PRADESH
UNACHAL PRADESH
ASSAM
BIHAR
CHANDIGARH
CHHATTISGARH
D & N HAVELI
DAMAN & DIU
DELHI
GOA
GUJARAT
HARYANA
IMACHAL PRADESH
JAMMU & KASHMIR
JHARKHAND
KARNATAKA
KERALA
LAKSHADWEEP
MADHYA PRADESH
MAHARASHTRA
MANIPUR
MEGHALAYA
MIZORAM
NAGALAND
ODISHA
PUDUCHERRY
PUNJAB
RAJASTHAN
SIKKIM
TAMIL NADU
TRIPURA
UTTAR PRADESH
UTTARAKHAND
WEST BENGAL


Dowry Deaths V/S States/UT (2009)
Dowry Death Cases
2500
2000
1500
1000
500
0
A & N ISLANDS
ANDHRA PRADESH
UNACHAL PRADESH
ASSAM
BIHAR
CHANDIGARH
CHHATTISGARH
D & N HAVELI
DAMAN & DIU
DELHI
GOA
GUJARAT
HARYANA
IMACHAL PRADESH
JAMMU & KASHMIR
JHARKHAND
KARNATAKA
KERALA
LAKSHADWEEP
MADHYA PRADESH
MAHARASHTRA
MANIPUR
MEGHALAYA
MIZORAM
NAGALAND
ODISHA
PUDUCHERRY
PUNJAB
RAJASTHAN
SIKKIM
TAMIL NADU
TRIPURA
UTTAR PRADESH
UTTARAKHAND
WEST BENGAL


Dowry Deaths V/S States/UT (2010)
Dowry Death Cases
2500
2000
1500
1000
500
0
A & N ISLANDS
ANDHRA PRADESH
ARUNACHAL PRADESH
ASSAM
BIHAR
CHANDIGARH
CHHATTISGARH
D & N HAVELI
DAMAN & DIU
DELHI
GOA
GUJARAT
HARYANA
HIMACHAL PRADESH
JAMMU & KASHMIR
JHARKHAND
KARNATAKA
KERALA
LAKSHADWEEP
MADHYA PRADESH
MAHARASHTRA
MANIPUR
MEGHALAYA
MIZORAM
NAGALAND
ODISHA
PUDUCHERRY
PUNJAB
RAJASTHAN
SIKKIM
TAMIL NADU
TRIPURA
UTTAR PRADESH
UTTARAKHAND
WEST BENGAL


Dowry Deaths V/S States/UT (2011)
Dowry Death Cases
2500
2000
1500
1000
500
0
A & N ISLANDS
ANDHRA PRADESH
ARUNACHAL PRADESH
ASSAM
BIHAR
CHANDIGARH
CHHATTISGARH
D & N HAVELI
DAMAN & DIU
DELHI
GOA
GUJARAT
HARYANA
HIMACHAL PRADESH
JAMMU & KASHMIR
JHARKHAND
KARNATAKA
KERALA
LAKSHADWEEP
MADHYA PRADESH
MAHARASHTRA
MANIPUR
MEGHALAYA
MIZORAM
NAGALAND
ODISHA
PUDUCHERRY
PUNJAB
RAJASTHAN
SIKKIM
TAMIL NADU
TRIPURA
UTTAR PRADESH
UTTARAKHAND
WEST BENGAL



Inference : From all the above plotted bar graphs, it can be easily interpreted that the state of Uttar Pradesh faced the most number of dowry related deaths every year. This points towards the fact that suppression and torture of women for Dowry in Uttar Pradesh is far more as compared to other states.

Q4. What can you say about the crime rate of the following 5 states/UT:

a)Delhi

b)Andhra Pradesh

c)West Bengal

d)Madhya Pradesh

e)Kerala

Did it increase or decrease over the period of time?

Ans.

Approach : We would first filter out our data on the basis of State/UT name as per the question. Then for the chosen state, the total number of cases for every single year would be calculated. And then we would plot a line graph for our data to assess the increment or decrement of criminal cases over time.

Variables Used :

dlcrimes - Year wise total number of crimes of Delhi

apcrimes - Year wise total number of crimes of Andhra Pradesh

wbcrimes - Year wise total number of crimes of West Bengal

mpcrimes - Year wise total number of crimes of Madhya Pradesh

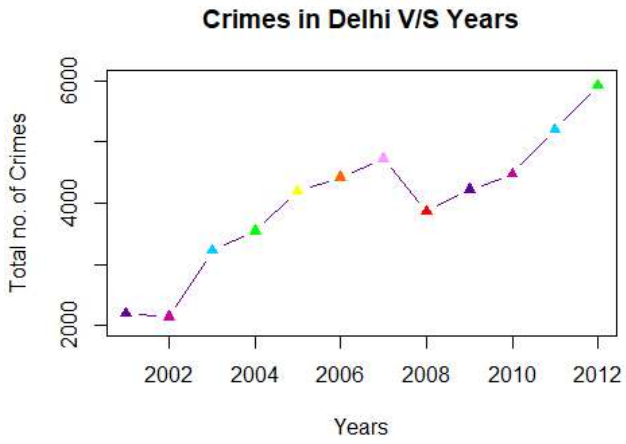
klcrimes - Year wise total number of crimes of Kerala

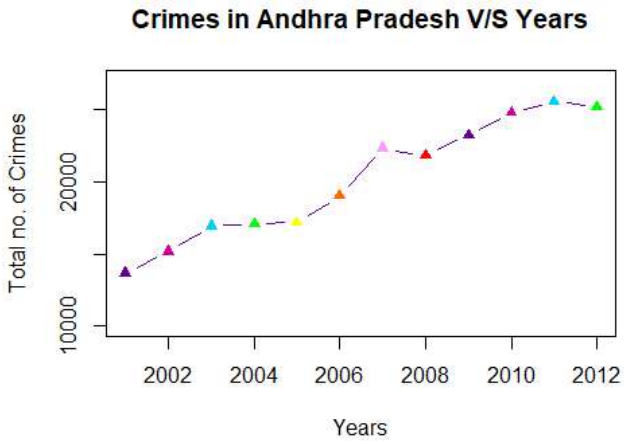
Code Snippets :

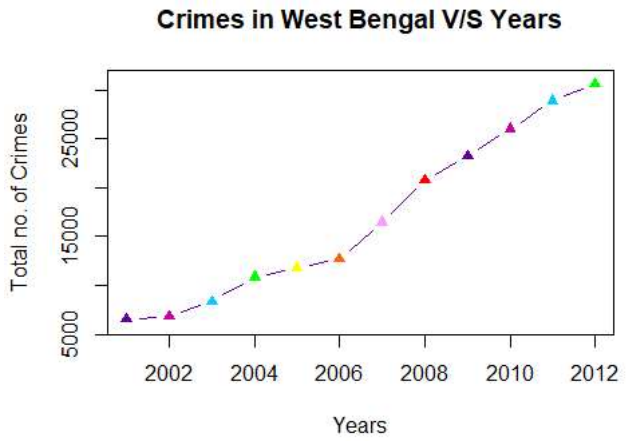
Storing year-wise total crimes for individual states and plotting the graph

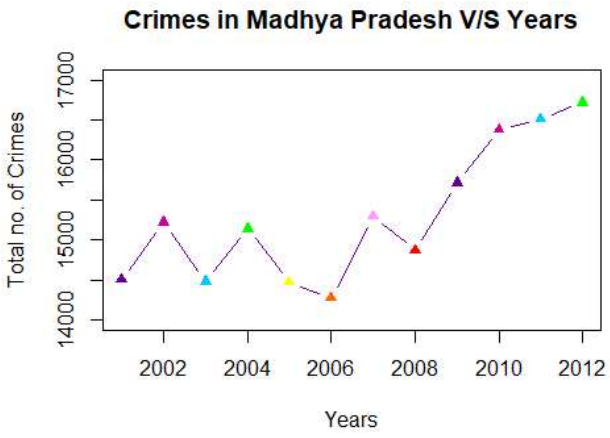
```
#State/UT 1: Delhi
dlcrimes <- crimes_data[(crimes_data$DISTRICT=="DELHI UT TOTAL"),]
dlcrimes$total = rowSums(dlcrimes[,c(4,5,6,7,8,9,10)])
dlcrimes
plot(dlcrimes$Year, dlcrimes$total, type = "b", pch = 17, col = coul1, xlab = "Years",
     ylab = "Total no. of Crimes",main = "Crimes in Delhi V/S Years", ylim = c(2000,6000), )
#Ans: Increasing over the period of time.

#State/UT 2: Andhra Pradesh
apcrimes <- crime_total[(crime_total$STATE.UT == "ANDHRA PRADESH"),]
apcrimes$total = rowSums(apcrimes[,c(4,5,6,7,8,9,10)])
apcrimes
plot(apcrimes$Year, apcrimes$total, type = "b", pch = 17, col = coul1, xlab = "Years",
     ylab = "Total no. of Crimes",main = "Crimes in Andhra Pradesh V/S Years", ylim = c(10000,27000), )
#Ans: Increasing over the period of time.
```

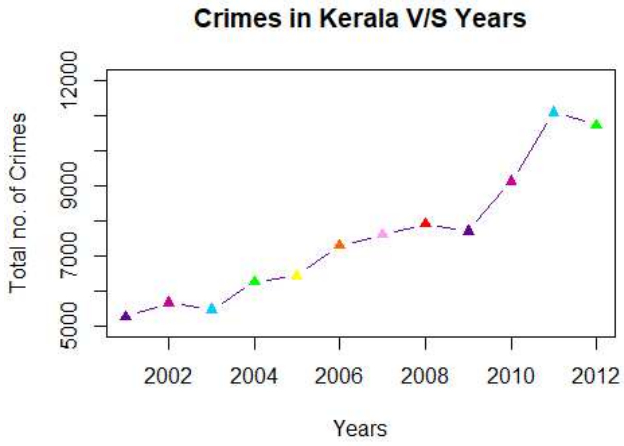
The Plots :











Inference : The general trend is that over the years crimes against women have increased drastically which is a serious concern. More specifically, in the case of :

Delhi - The crime rates took a dip in 2008 but have continued to grow from then on

Andhra Pradesh - The crime rates here too declined slightly but then continued to grow.

West Bengal - We can interpret from the graph that West Bengal has seen the most continuous growth in crime rates.

Madhya Pradesh - Here, we see fluctuations in crime rates before 2008, but after 2008 the crime rates have increased continuously.

Kerala - Here, we see moderate levels of crimes before 2009 but after that the crime rates grew before taking a dip in 2012

Q5. Give a 5 number summary for the different crimes in Bihar for the year it had the maximum total no. of crimes!

Ans.

Approach : For this, we would first compute the year which had the maximum number of crimes in Bihar (comes out to be 2012). Then, we would plot a boxplot to represent the 5 number summary for each individual crime type.

Variables Used :

crime_bihar -contains total crime in bihar for each crime of each year from 2001 to 2012

crime_total _bihar -contains the sum of row of crime_bihar

FilteredData_bihar -Contains the crime data of bihar for year 2012

Code Snippets :

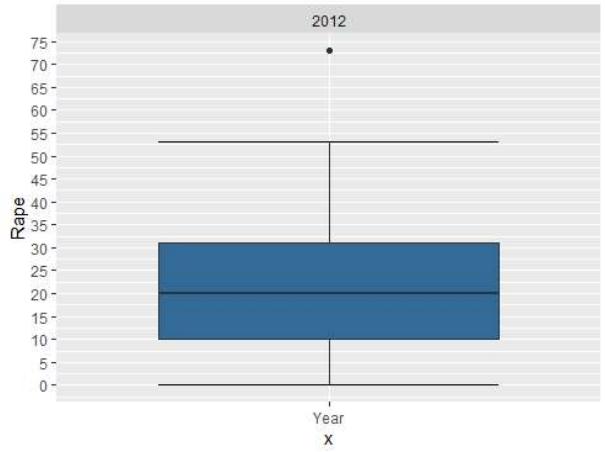
Checking in which year Bihar had the most number of crimes

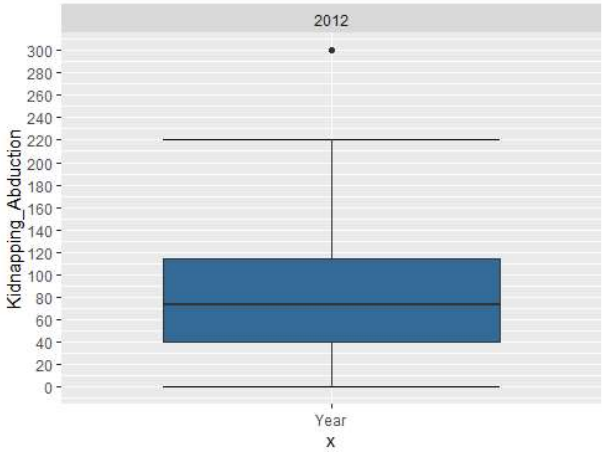
```
682  crime_bihar<-crime_total[crime_total$STATE.UT=='BIHAR',]
683  crime_total_bihar<-rowSums(crime_bihar[4:10])
684  print(crime_total_bihar)
685  print(max(crime_total_bihar))
686  View(crime_total_bihar)
687  #Bihar records highest crime in 2012
```

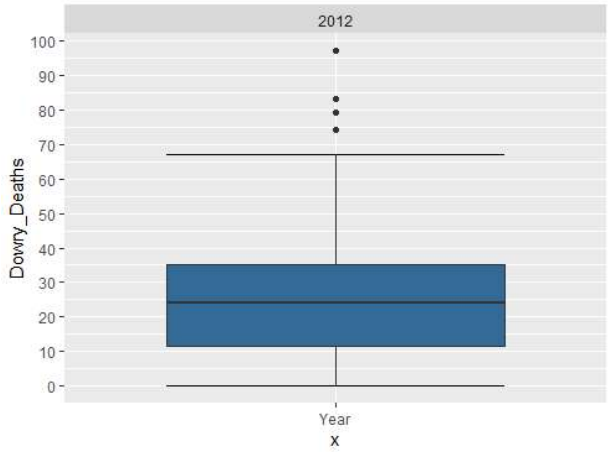
Plotting boxplots for every crime type for the year, in which Bihar had most number of crimes

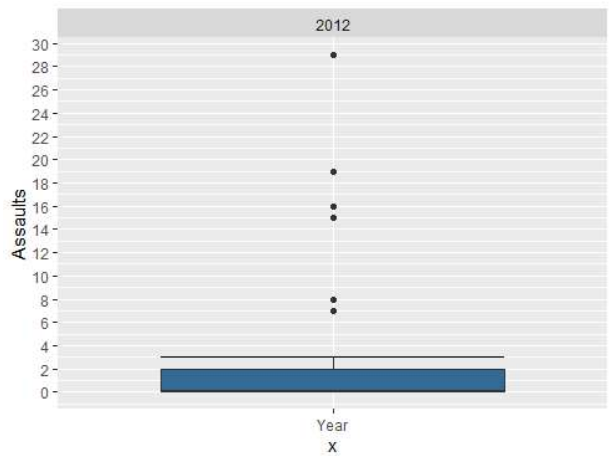
```
#so plotting box plot for 2012 crimes in bihar
crimes_bihar<-crimes_data[(crimes_data$STATE.UT=="BIHAR"),]

FilteredData_bihar = subset ( crimes_bihar,Year=='2012')

ggplot(FilteredData_bihar, aes(x = "Year", y = Rape, fill = Year)) +
  stat_boxplot(geom='errorbar')+
  geom_boxplot() +
  scale_y_continuous(breaks = scales::pretty_breaks(n = 14)) +
  facet_wrap( ~ Year, scales="free")+
  theme(legend.position = "none")

ggplot(FilteredData_bihar, aes(x = "Year", y = Kidnapping_Abduction, fill = Year)) +
  stat_boxplot(geom='errorbar')+
  geom_boxplot() +
  scale_y_continuous(breaks = scales::pretty_breaks(n = 14)) +
  facet_wrap( ~ Year, scales="free")+
  theme(legend.position = "none")
```

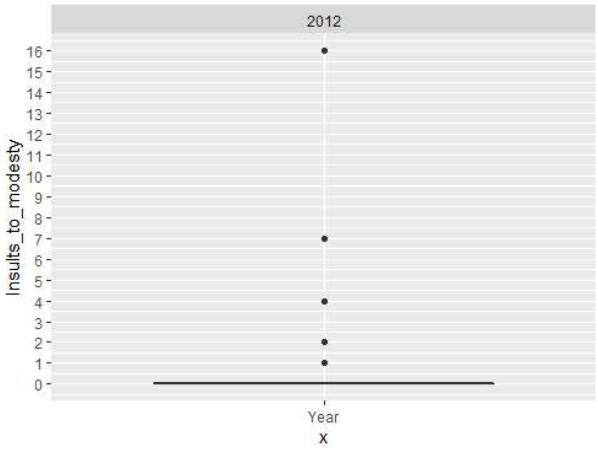
The Plots :

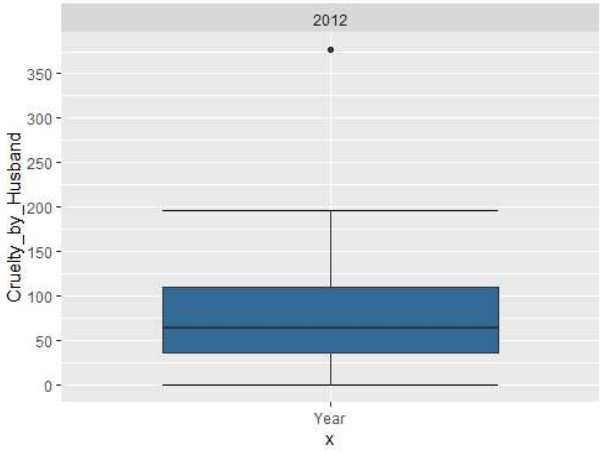


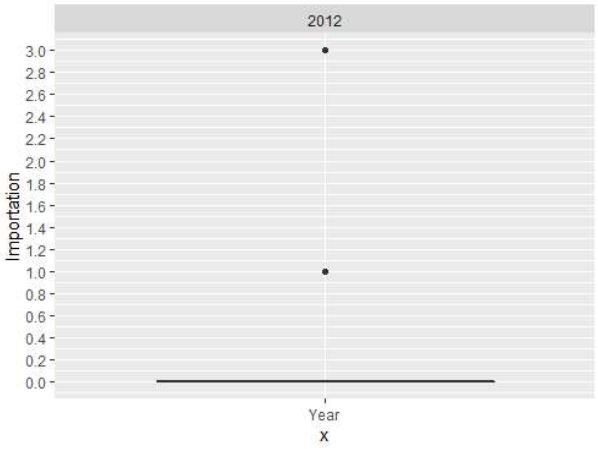












Inference : Box plots show the five-number summary of a set of data: including the minimum score, first (lower) quartile, median, third (upper) quartile, and maximum score.

Crimes	Min	1st Q	2nd Q	3rd Q	Max	Mean
Rape	0.0	10.0	20.0	31.0	73.0	21.07
Kidnapping/ Abduction	0.0	40.25	73.50	114.75	300.0	86.11
Dowry_Deaths	0.0	11.50	24.0	35.0	97.0	28.98

Assaults	0.0	0.0	0.0	2.0	29.0	2.682
Insults_to_modesty	0.0	0.0	0.0	0.0	16.0	0.8409
Cruelty_by_Husband	0.0	37.0	64.50	110.25	377.0	83.77
Importation	0.0	0.0	0.0	0.0	3.0	0.09091

Q6. Compare the rates of Rapes and Kidnapping/Abduction in Rajasthan(district-wise) for the two years(take one year for which the total number of crimes was the maximum and other year for which the total number of crimes was minimum).

Ans.

Approach : For the solution of this question, we have to first check which two year had the maximum and minimum total crimes. Then, for the year with minimum total crimes, we would first access the data for Rapes and plot a bar graph and then do the same for Kidnapping/Abduction. In the similar manner, we would do these operations for the year with the maximum number of crimes and thus we would get our answer.

Variables Used :

tcrimes - Contains year wise total number of crimes

rajasthan_2001 - Contains all the crimes of Rajasthan for 2001

rapes_rj_2001 - District wise no. of rapes in Rajasthan for 2001

kidnapping_rj_2001 - District wise no. of kidnapping/abduction in Rajasthan for 2001

rajasthan_2012 - Contains all the crimes of Rajasthan for 2012

rapes_rj_2012 - District wise no. of rapes in Rajasthan for 2012

kidnapping_rj_2012 - District wise no. of kidnapping/abduction in Rajasthan for 2012

Code Snippets :

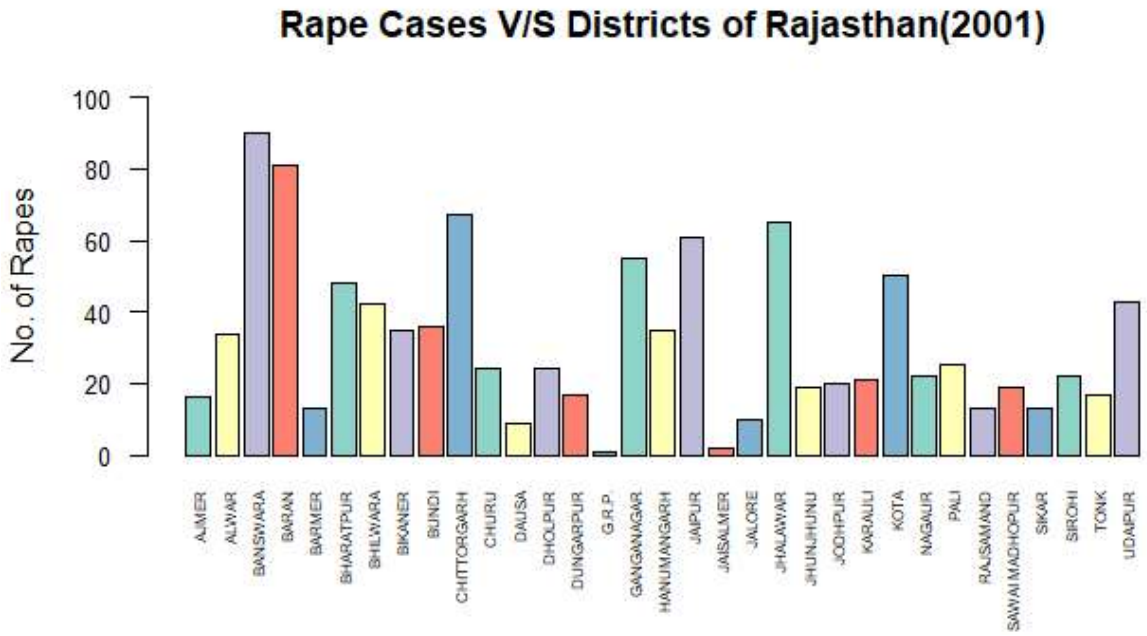
Picking out Rapes cases from Rajasthan for 2001 and plotting a bargraph

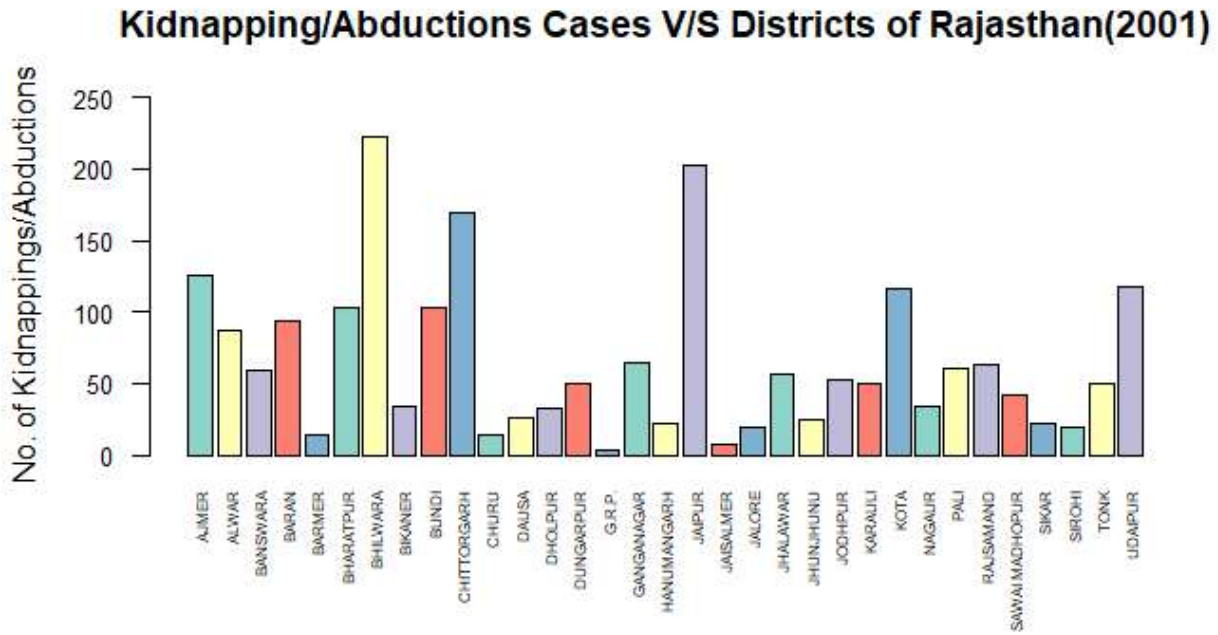
```
#for 2001(min no. of total crimes)
rajasthan_2001 <- crimes_2001[(crimes_2001$STATE.UT == "RAJASTHAN"),]
rajasthan_2001

#rapes in 2001
rapes_rj_2001 <- aggregate(x= rajasthan_2001$Rape,
                           by= list(rajasthan_2001$DISTRICT),
                           FUN=sum)
rapes_rj_2001   #district wise no. of rapes in rajasthan for 2001

#plotting the graph for rapes district-wise
barplot(height = rapes_rj_2001$x, name = rapes_rj_2001$Group.1, col=coul3,
        ylab="No. of Rapes",
        main="Rape Cases V/S Districts of Rajasthan(2001)",
        las=2, cex.axis=0.8, cex.names=0.5,font.axis=2, ylim = c(0,100))
```

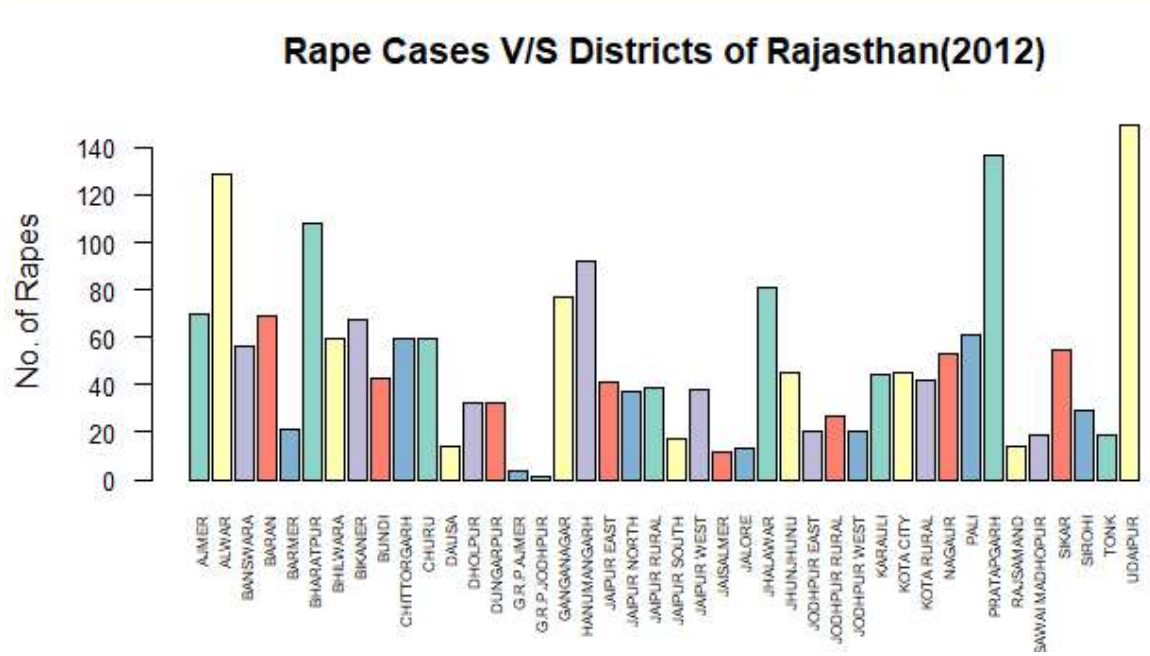
Picking out Kidnapping/Abduction cases from Rajasthan for 2001 and plotting a bar graph

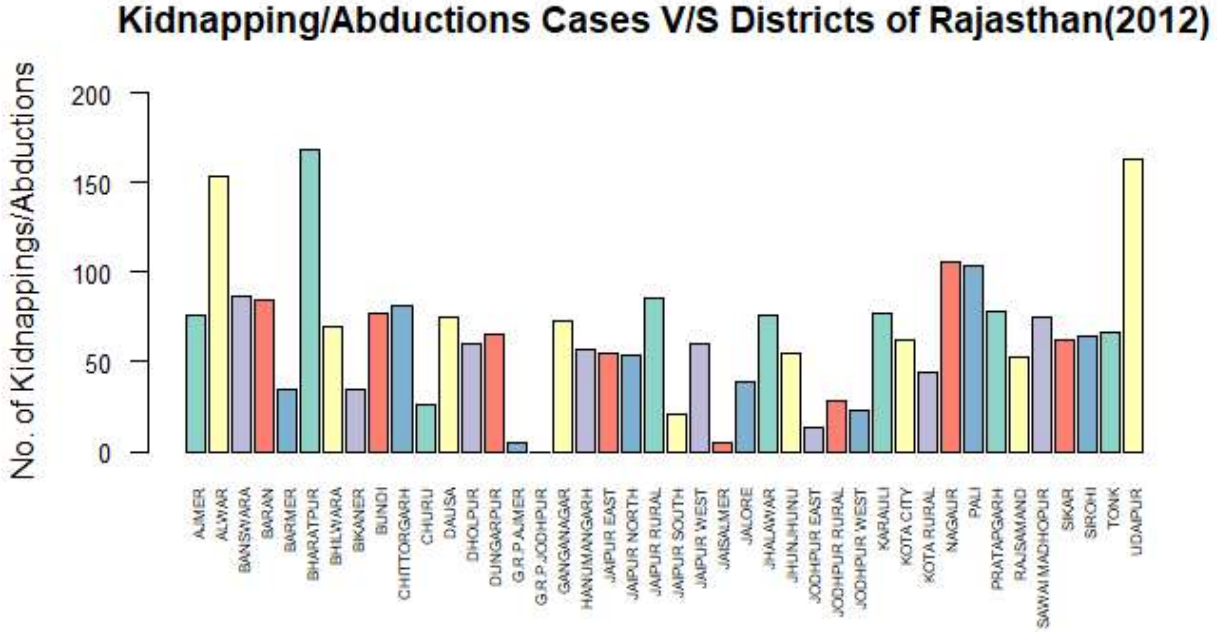
```
#kidnapping/abduction in 2001
kidnapping_rj_2001 <- aggregate(x= rajasthan_2001$`Kidnapping/Abduction`,
                                by= list(rajasthan_2001$DISTRICT),
                                FUN=sum)
kidnapping_rj_2001
#district wise no. of kidnapping/abduction in rajasthan for 2001

#plotting the graph for kidnapping district-wise
barplot(height = kidnapping_rj_2001$x, name = kidnapping_rj_2001$Group.1,
        col=coul3, ylab="No. of Kidnappings/Abductions",
        main="Kidnapping/Abductions Cases V/S Districts of Rajasthan(2001)",
        las=2, cex.axis=0.8, cex.names=0.5,font.axis=2, ylim = c(0,250))
```

The Plots :


Rape Cases V/S Districts of Rajasthan(2001)
No. of Rapes
0
20
40
60
80
100
AJMER
ALWAR
BANSWARA
BARAN
BARMER
BHARATPUR
BHILWARA
BIKANER
BUNDI
CHITTORGARH
CHURU
DAUSA
DHOLPUR
DUNGARPUR
G.R.P.
GANGANAGAR
HANUMANGARH
JAIPUR
JAISALMER
JALORE
JHALAWAR
JHUNJHUNU
JODHPUR
KARAULI
KOTA
NAGAUR
PALI
RAJSAMAND
SAWAI MADHOPUR
SIKAR
SIROHI
TONK
UDAIPUR


Kidnapping/Abductions Cases V/S Districts of Rajasthan(2001)
No. of Kidnappings/Abductions
0
50
100
150
200
250
AJMER
ALWAR
BANSWARA
BARAN
BARMER
BHARATPUR
BHILWARA
BIKANER
BUNDI
CHITTORGARH
CHURU
DAUSA
DHOLPUR
DUNGARPUR
G.R.P.
GANGANAGAR
HANUMANGARH
JAIPUR
JAISALMER
JALORE
JHALAWAR
JHUNJHUNU
JODHPUR
KARAULI
KOTA
NAGAUR
PALI
RAJSAMAND
SAWAI MADHOPUR
SIKAR
SIROHI
TONK
UDAIPUR





Inference : We can interpret from the plotted bar graphs that

For the year with minimum total number of crimes i.e. 2001, in Rajasthan :

District “BANSWARA” had maximum number of “Rape” cases

District “G.R.P.” had the minimum number of “Rape” cases

District “BHILWARA” had the maximum number of “Kidnapping/Abduction” cases

District “G.R.P.” had the minimum number of “Kidnapping/Abduction” cases

For the year with maximum total number of crimes i.e. 2012, in Rajasthan :

District “UDAIPUR” had maximum number of “Rape” cases

District “G.R.P. JODHPUR” had the minimum number of “Rape” cases

District “BHARATPUR” had the maximum number of “Kidnapping/Abduction” cases

District “G.R.P. JODHPUR” had the minimum number of “Kidnapping/Abduction” cases

----------END OF PROJECT—----------