

```
library(readxl)
library(tidyverse)
library(kableExtra)
USArrests6.1973 <- as.data.frame(
  read_csv("~/git/SDS348/Project/Datasets/USArrests6.1973.csv"))
glimpse(USArrests6.1973)

## Rows: 50
## Columns: 5
## $ X1
##> chr> "Alabama", "Alaska", "Arizona", "Arkansas", "California", "Colo.
## $ Murder
##> dbl> 13.2, 10.0, 8.1, 8.8, 9.0, 7.9, 3.3, 5.9, 15.4, 17.4, 5.3, 2._
## $ Assault
##> dbl> 236, 263, 284, 190, 276, 284, 110, 238, 335, 211, 46, 129, 24._
## $ UrbanPop
##> dbl> 58, 48, 88, 50, 81, 78, 77, 72, 88, 68, 83, 54, 82, 65, 57, 6._
## $ Rape
##> dbl> 21.2, 44.5, 31.0, 19.5, 40.6, 38.7, 11.1, 15.8, 31.9, 25.8, 2._

allArrests <- as.data.frame(
  read_csv("~/git/SDS348/Project/Datasets/state_crime_all.csv"))
glimpse(allArrests)

## Rows: 2,751
## Columns: 751
## $ State
## $ Year
## $ Data.Population
## $ Data.Rates.Property.All
## $ Data.Rates.Property.Burglary
## $ Data.Rates.Property.Larceny
## $ Data.Rates.Property.Motor
## $ Data.Rates.Violent.All
## $ Data.Rates.Violent.Assault
## $ Data.Rates.Violent.Murder
## $ Data.Rates.Violent.Rape
## $ Data.Rates.Violent.Robbery
## $ Data.Totals.Property.All
## $ Data.Totals.Property.Burglary
## $ Data.Totals.Property.Larceny
## $ Data.Totals.Property.Motor
## $ Data.Totals.Violent.All
## $ Data.Totals.Violent.Assault
## $ Data.Totals.Violent.Murder
## $ Data.Totals.Violent.Rape
## $ Data.Totals.Violent.Robbery

colnames(USArrests6.1973[1]) <- "State"
allArrests1973 <- allArrests %>% filter(Year == 1973 & State != "United States")
USArrestscombined <- USArrests6.1973 %>% left_join(allArrests1973, by = c("State")) %>%
  filter(!is.na(State))
ME<- c("Connecticut", "Maine", "Massachusetts", "New Hampshire",
      "Rhode Island", "Vermont", "New Jersey", "New York",
      "Pennsylvania")
ME<- c("Indiana", "Illinois", "Michigan", "Ohio", "Wisconsin",
      "Iowa", "Kansas", "Minnesota", "Missouri", "Nebraska",
      "North Dakota", "South Dakota")
SE<- c("Delaware", "District of Columbia", "Florida", "Georgia",
      "Maryland", "North Carolina", "South Carolina", "Virginia",
      "West Virginia", "Alabama", "Kentucky", "Mississippi",
      "Tennessee", "Arkansas", "Louisiana", "Oklahoma", "Texas")
WE<- c("Arizona", "Colorado", "Idaho", "New Mexico", "Montana",
      "Utah", "Nevada", "Wyoming", "Alaska", "California",
      "Hawaii", "Oregon", "Washington")
USArrestscombined <- USArrestscombined %>%
  mutate(Region = case_when(State %in% ME ~ "Midwest",
                             State %in% W ~ "West",
                             State %in% NE ~ "Northeast",
                             State %in% S ~ "South")) %>% arrange(desc(Region))
USArrestscombinedImportant <- USArrestscombined %>% select(-c(Data.Rates.Violent.Murder, Data.Rates.Violent.Assault,
  Data.Rates.Violent.Robbery))

USArrestscombinedImportant %>%
  summarise(mean Rape = mean(Rape, na.rm = TRUE), mean Assault = mean(Assault, na.rm = TRUE),
    mean Murder = mean(Murder, na.rm = TRUE),
    mean Rape = mean(Rape, na.rm = TRUE),
    mean Burglary = mean(Data.Rates.Property.Burglary, na.rm = TRUE),
    mean Larceny = mean(Data.Rates.Property.Larceny, na.rm = TRUE),
    mean Motor = mean(Data.Rates.Property.Motor, na.rm = TRUE)) %>%
  kbl(caption = "Recreating booktabs style table") %>%
  kable_classic(full_width = F, html_font = "Cambria") %>% kable_material_dark()
```

Recreating booktabs style table

mean Rape	mean Assault	mean Murder	mean Burglary	mean Larceny	mean Motor
21.232	170.76	7.788	1096.822	2041.032	369.798

```
USArrestscombinedImportant %>%
  summarise(min Rape = min(Rape, na.rm = TRUE), min Assault = min(Assault, na.rm = TRUE),
    min Murder = min(Murder, na.rm = TRUE),
    min Burglary = min(Data.Rates.Property.Burglary, na.rm = TRUE),
    min Larceny = min(Data.Rates.Property.Larceny, na.rm = TRUE),
    min Motor = min(Data.Rates.Property.Motor, na.rm = TRUE)) %>%
  kbl(caption = "Recreating booktabs style table") %>%
  kable_classic(full_width = F, html_font = "Cambria") %>% kable_material_dark()
```

Recreating booktabs style table

min Rape	min Assault	min Murder	min Burglary	min Larceny	min Motor
7.3	45	0.8	383.4	824.9	107.1

```
USArrestscombinedImportant %>%
  summarise(max Rape = max(Rape, na.rm = TRUE), max Assault = max(Assault, na.rm = TRUE),
    max Murder = max(Murder, na.rm = TRUE),
    max Burglary = max(Data.Rates.Property.Burglary, na.rm = TRUE),
    max Larceny = max(Data.Rates.Property.Larceny, na.rm = TRUE),
    max Motor = max(Data.Rates.Property.Motor, na.rm = TRUE)) %>%
  kbl(caption = "Recreating booktabs style table") %>%
  kable_classic(full_width = F, html_font = "Cambria") %>% kable_material_dark()
```

Recreating booktabs style table

max Rape	max Assault	max Murder	max Burglary	max Larceny	max Motor
46	337	17.4	2149.8	3720.1	1109.6

```
USArrestscombinedImportant %>%
  summarise(sd Rape = sd(Rape, na.rm = TRUE), sd Assault = sd(Assault, na.rm = TRUE),
    sd Murder = sd(Murder, na.rm = TRUE),
    sd Burglary = sd(Data.Rates.Property.Burglary, na.rm = TRUE),
    sd Larceny = sd(Data.Rates.Property.Larceny, na.rm = TRUE),
    sd Motor = sd(Data.Rates.Property.Motor, na.rm = TRUE)) %>%
  kbl(caption = "Recreating booktabs style table") %>%
  kable_classic(full_width = F, html_font = "Cambria") %>% kable_material_dark()
```

Recreating booktabs style table

sd Rape	sd Assault	sd Murder	sd Burglary	sd Larceny	sd Motor
9.366385	83.33766	4.35551	405.4954	671.5986	200.4927

```
USArrestscombinedImportant %>%
  summarise(variation Rape = var(Rape, na.rm = TRUE), variation Assault = var(Assault, na.rm = TRUE),
    variation Murder = var(Murder, na.rm = TRUE),
    variation Burglary = var(Data.Rates.Property.Burglary, na.rm = TRUE),
    variation Larceny = var(Data.Rates.Property.Larceny, na.rm = TRUE),
    variation Motor = var(Data.Rates.Property.Motor, na.rm = TRUE)) %>%
  kbl(caption = "Recreating booktabs style table") %>%
  kable_classic(full_width = F, html_font = "Cambria") %>% kable_material_dark()
```

Recreating booktabs style table

variation Rape	variation Assault	variation Murder	variation Burglary	variation Larceny	variation Motor
87.72916	6945.166	18.97047	164426.5	451044.7	40197.32

```
USArrestscombinedImportant %>%
  summarise(distinct Values Rape = n_distinct(Rape, na.rm = TRUE), distinct Values Assault = n_distinct(Assault, n
    a.rm = TRUE), distinct Values Murder = n_distinct(Murder, na.rm = TRUE), distinct Values Burglary = n_distinct(Dat
    a.Rates.Property.Burglary, na.rm = TRUE), distinct Values Larceny = n_distinct(Data.Rates.Property.Larceny, na.rm
    = TRUE), distinct Values Motor = n_distinct(Data.Rates.Property.Motor, na.rm = TRUE)) %>%
  kbl(caption = "Recreating booktabs style table") %>%
  kable_classic(full_width = F, html_font = "Cambria") %>% kable_material_dark()
```

Recreating booktabs style table

Distinct Values Rape	Distinct Values Assault	Distinct Values Murder	Distinct Values Burglary	Distinct Values Larceny	Distinct Values Motor
48	45	43	50	50	50

```
USArrestscombinedImportant %>%
  summarise(Quantile Rape = quantile(Rape, na.rm = TRUE), Quantile Assault = quantile(Assault, na.rm = TRUE), Quant
    ile Murder = quantile(Murder, na.rm = TRUE), Quantile Burglary = quantile(Data.Rates.Property.Burglary, na.rm = TR
    UE), Quantile Larceny = quantile(Data.Rates.Property.Larceny, na.rm = TRUE), Quantile Motor = quantile(Data.Rates
    .Property.Motor, na.rm = TRUE)) %>% kbl(caption = "Recreating booktabs style table") %>%
  kable_classic(full_width = F, html_font = "Cambria") %>% kable_material_dark() %>% footnote(general = "Here are
    general comments of the table.",
    number = c("First value for column is at 0%; ", "Second value for column is at 25%;"),
    alphabet = c("Third value for column is at 50%; ", "Fourth value for column is at 75%; "),
    symbol = c("Fifth value for column is at 100% ")
  )
```

Recreating booktabs style table

Quantile Rape	Quantile Assault	Quantile Murder	Quantile Burglary	Quantile Larceny	Quantile Motor
7.300	45	0.800	383.400	824.900	107.100
15.075	109	4.075	813.325	1494.775	217.350
20.100	159	7.250	1027.200	2004.500	340.800
26.175	249	11.250	1274.050	2509.275	502.475
46.000	337	17.400	2149.800	3720.100	1109.600

Note:

Here are general comments of the table.

1 First value for column is at 0%;

2 Second value for column is at 25%;

3 Third value for column is at 50%;

4 Fourth value for column is at 75%;

5 Fifth value for column is at 100%;

```
USArrestscombinedImportant %>%
  group_by(Region) %>%
  summarise(mean Rape = mean(Rape, na.rm = TRUE), mean Assault = mean(Assault, na.rm = TRUE), mean Murder = mean(Mu
    rder, na.rm = TRUE), mean Burglary = mean(Data.Rates.Property.Burglary, na.rm = TRUE), mean Larceny = mean(Data.Ra
    tes.Property.Larceny, na.rm = TRUE), mean Motor = mean(Data.Rates.Property.Motor, na.rm = TRUE)) %>%
  kbl(caption = "Recreating booktabs style table") %>%
  kable_classic(full_width = F, html_font = "Cambria") %>% kable_material_dark()
```

Recreating booktabs style table

Region	mean Rape	mean Assault	mean Murder	mean Burglary	mean Larceny	mean Motor
MidWest	18.44167	120.3333	5.700000	894.525	1972.400	317.7750
NorthEast	13.77778	126.6667	4.700000	1041.656	1677.667	508.7333
South	21.16250	220.0000	11.706250	1025.006	1657.412	293.0000
West	29.05385	187.2308	7.030769	1410.138	2828.092	416.1538

```
USArrestscombinedImportant %>%
  group_by(Region) %>%
  summarise(min Rape = min(Rape, na.rm = TRUE), min Assault = min(Assault, na.rm = TRUE), min Murder = min(Murder, n
    a.rm = TRUE), min Burglary = min(Data.Rates.Property.Burglary, na.rm = TRUE), min Larceny = min(Data.Rates.Prope
    rty.Larceny, na.rm = TRUE), max Motor = max(Data.Rates.Property.Motor, na.rm = TRUE)) %>%
  kbl(caption = "Recreating booktabs style table") %>%
  kable_classic(full_width = F, html_font = "Cambria") %>% kable_material_dark()
```

Recreating booktabs style table

Region	min Rape	min Assault	min Murder	min Burglary	min Larceny	min Motor
MidWest	7.3	45	0.8	383.4	1406.0	131.4
NorthEast	7.8	48	2.1	685.0	1090.7	136.2
South	9.3	81	5.7	415.8	824.9	107.1
West	14.2	46	2.6	699.7	2237.8	207.0

```
USArrestscombinedImportant %>%
  group_by(Region) %>%
  summarise(max Rape = max(Rape, na.rm = TRUE), max Assault = max(Assault, na.rm = TRUE), max Murder = max(Murder,
    na.rm = TRUE), max Burglary = max(Data.Rates.Property.Burglary, na.rm = TRUE), max Larceny = max(Data.Rates.Prope
    rty.Larceny, na.rm = TRUE), max Motor = max(Data.Rates.Property.Motor, na.rm = TRUE)) %>%
  kbl(caption = "Recreating booktabs style table") %>%
  kable_classic(full_width = F, html_font = "Cambria") %>% kable_material_dark()
```

Recreating booktabs style table

Region	max Rape	max Assault	max Murder	max Burglary	max Larceny	max Motor
MidWest	35.1	255	12.1	1584.6	2771.3	548.3
NorthEast	26.1	254	11.1	1348.2	2312.3	1109.6
South	31.9	337	17.4	1857.2	3048.6	547.8
West	46.0	294	12.2	2149.8	3720.1	635.9

```
USArrestscombinedImportant %>%
  group_by(Region) %>%
  summarise(sd Rape = sd(Rape, na.rm = TRUE), sd Assault = sd(Assault, na.rm = TRUE), sd Murder = sd(Murder, na.rm
    = TRUE), sd Burglary = sd(Data.Rates.Property.Burglary, na.rm = TRUE), sd Larceny = sd(Data.Rates.Property.Larcen
    y, na.rm = TRUE), sd Motor = sd(Data.Rates.Property.Motor, na.rm = TRUE)) %>% kbl(caption = "Recreating booktabs s
    tyle table") %>%
  kable_classic(full_width = F, html_font = "Cambria") %>% kable_material_dark()
```

Recreating booktabs style table

Region	sd Rape	sd Assault	sd Murder	sd Burglary	sd Larceny	sd Motor
MidWest	7.981736	71.53935	3.556345	337.1593	354.8975	139.8711
NorthEast	5.942806	64.85754	3.047950	250.4261	351.4370	334.4643
South	5.027536	74.20782	3.760934	332.2478	626.5774	129.6427
West	10.997774	80.32761	3.062511	480.6029	426.8271	152.4576

```
USArrestscombinedImportant %>%
  group_by(Region) %>%
  summarise(variation Rape = var(Rape, na.rm = TRUE), variation Assault = var(Assault, na.rm = TRUE), variation Mur
    der = var(Murder, na.rm = TRUE), variation Burglary = var(Data.Rates.Property.Burglary, na.rm = TRUE), variation L
    arceny = var(Data.Rates.Property.Larceny, na.rm = TRUE), variation Motor = var(Data.Rates.Property.Motor, na.rm =
    TRUE)) %>% kbl(caption = "Recreating booktabs style table") %>%
  kable_classic(full_width = F, html_font = "Cambria") %>% kable_material_dark()
```

Recreating booktabs style table

Region	variation Rape	variation Assault	variation Murder	variation Burglary	variation Larceny	variation Motor
MidWest	63.70811	5117.879	12.661818	113675.70	125952.3	19563.93
NorthEast	35.31694	4206.500	9.290000	62714.25	123508.0	11866.37
South	31.66917	5506.800	14.144625	110388.59	392599.3	16807.23
West	120.95103	6452.526	9.378974	230979.14	182181.4	22343.33

```
USArrestscombinedImportant %>%
  group_by(Region) %>%
  summarise(distinct Values Rape = n_distinct(Rape, na.rm = TRUE), distinct Values Assault = n_distinct(Assault, n
    a.rm = TRUE), distinct Values Murder = n_distinct(Murder, na.rm = TRUE), distinct Values Burglary = n_distinct(Dat
    a.Rates.Property.Burglary, na.rm = TRUE), distinct Values Larceny = n_distinct(Data.Rates.Property.Larceny, na.rm
    = TRUE), distinct Values Motor = n_distinct(Data.Rates.Property.Motor, na.rm = TRUE)) %>%
  kbl(caption = "Recreating booktabs style table") %>%
  kable_classic(full_width = F, html_font = "Cambria") %>% kable_material_dark()
```

Recreating booktabs style table

Region	Distinct Values Rape	Distinct Values Assault	Distinct Values Murder	Distinct Values Burglary	Distinct Values Larceny	Distinct Values Motor
MidWest	12	12	12	12	12	12
NorthEast	9	9	8	9	9	9
South	16	16	14	16	16	16
West	13	12	13	13	13	13

```
USArrestscombinedImportant %>%
  group_by(Region) %>%
  summarise(Quantile Rape = quantile(Rape, na.rm = TRUE), Quantile Assault = quantile(Assault, na.rm = TRUE), Quant
    ile Murder = quantile(Murder, na.rm = TRUE), Quantile Burglary = quantile(Data.Rates.Property.Burglary, na.rm = TR
    UE), Quantile Larceny = quantile(Data.Rates.Property.Larceny, na.rm = TRUE), Quantile Motor = quantile(Data.Rates
    .Property.Motor, na.rm = TRUE)) %>% kbl(caption = "Recreating booktabs style table") %>%
  kable_classic(full_width = F, html_font = "Cambria") %>% kable_material_dark() %>% footnote(general = "Here are
    general comments of the table.",
    number = c("First value of region for column is at 0%; ", "Second value of region for column is at
    25%;"),
    alphabet = c("Third value of region for column is at 50%; ", "Fourth value of region for column is at 75%; "),
    symbol = c("Fifth value of region for column is at 100% ")
  )
```

Recreating booktabs style table

Region	Quantile Rape	Quantile Assault	Quantile Murder	Quantile Burglary	Quantile Larceny	Quantile Motor
MidWest	7.300	45.0	0.800	383.400	1406.000	131.400
MidWest	12.425	68.0	2.675	636.550	1834.650	205.550
MidWest	17.250	107.5	5.150	979.700	2024.000	319.700
MidWest	22.050	134.5	7.725	1031.825	2071.975	403.775
MidWest	35.100	255.0	12.100	1584.600	2771.300	548.300
NorthEast	7.800	48.0	2.100	685.000	1090.700	136.200
NorthEast	9.500	83.0	2.200	801.500	1489.900	189.600
NorthEast	11.200	110.0	3.400	1029.300	1729.200	516.700
NorthEast	16.300	159.0	6.300	1244.500	1873.400	615.000
NorthEast	26.100	254.0	11.100	1348.200	2312.300	1109.600
South	9.300	81.0	5.700	415.800	824.900	107.100
South	16.900	180.0	8.725	891.000	1265.950	213.100
South	20.950	223.5	12.850	985.950	1462.900	297.050
South	25.575	264.0	14.650	1200.575	2030.050	357.350
South	31.900	337.0	17.400	1857.200	3048.600	547.800
West	14.200	46.0	2.600	699.700	2237.800	207.000
West	20.200	120.0	4.900	969.300	2614.300	301.200
West	29.300	161.0	6.800	1535.500	2865.500	408.200
West	38.700	263.0	9.000	1607.700	2988.500	545.600
West	46.000	294.0	12.200	2149.800	3720.100	635.800

Note:

Here are general comments of the table.

1 First value of region for column is at 0%;

2 Second value of region for column is at 25%;

3 Third value of region for column is at 50%;

4 Fourth value of region for column is at 75%;

5 Fifth value of region for column is at 100%;

```
USArrestscombinedImportantCorrelation <- USArrestscombinedImportant %>% select_if(is.numeric) %>% select(contains
  ("Rates"))
USArrestscombinedImportantCorrelation$Murder <- USArrestscombinedImportant$Murder
USArrestscombinedImportantCorrelation$Assault <- USArrestscombinedImportant$Assault
USArrestscombinedImportantCorrelation$Rape <- USArrestscombinedImportant$Rape
USArrestscombinedImportantCorrelation <- USArrestscombinedImportantCorrelation %>% select(-c(Data.Rates.Property.A
  ll, Data.Rates.Violent.All)) %>% rename(
  Burglary = Data.Rates.Property.Burglary,
  Larceny = Data.Rates.Property.Larceny,
  Motor = Data.Rates.Property.Motor,
  Robbery = Data.Rates.Property.Robbery,
  )
```

Correlation Matrix of Different State Crimes in 1973

	Burglary	Larceny	Motor	Robbery	Murder	Assault	Rape
Burglary	1.0000000	0.7800244	0.6437819	0.6149926	0.3331508	0.5621404	0.7787774
Larceny	0.7800244	1.0000000	0.4897432	0.3459301	-0.0333869	0.3225882	0.6489864
Motor	0.6437819	0.4897432	1.0000000	0.6415893	0.0808595	0.3431231	0.4194696
Robbery	0.6149926	0.3459301	0.6415893	1.0000000	0.4921382	0.5804301	0.6056911
Murder	0.3331508	-0.0333869	0.0808595	0.4921382	1.0000000	0.8018733	0.5635788
Assault	0.5621404	0.3225882	0.3431231	0.5804301	0.8018733	1.0000000	0.6052412
Rape	0.7787774	0.6489864	0.4194696	0.6056911	0.5635788	0.6052412	1.0000000

```
cor(USArrestscombinedImportantCorrelation) %>%
  # Save as a data frame
  as.data.frame %>%
  # Convert row names to an explicit variable
  rownames_to_column %>%
  # Pivot so that all correlations appear in the same column
  pivot_longer(<-, names_to = "other_var", values_to = "correlation") %>%
  # Specify variables are displayed alphabetically from top to bottom
  ggplot(aes(rowname, factor(other_var, levels = rev(levels(factor(other_var))))), fill=correlation)) +
  # Heatmap with geom_tile
  geom_tile() +
  # Change the scale to make the middle appear neutral
  scale_fill_gradient2(low="red", mid="white", high="blue") +
  # Correl values
  geom_text(aes(label = round(correlation, 2)), color = "black", size = 4) +
  # Give title and labels
  labs(title = "Correlation matrix for State Crime Rates in 197
```