# Assignment_Week_1&2_Venkidusamy_KesavAdithya

Kesav Adithya Venkidusamy

2022/06/11

```
knitr::opts_chunk$set(echo = TRUE)

library(readxl)
library(ggplot2)
library(tidyverse)
```

```
## Warning: package 'tidyverse' was built under R version 4.1.3
```

```
## -- Attaching packages --------------------------------------- tidyverse 1.3.1 --
```

```
## v tibble  3.1.4     v dplyr   1.0.7
## v tidyr   1.1.4     v stringr 1.4.0
## v readr   2.0.2     v forcats 0.5.1
## v purrr   0.3.4
```

```
## -- Conflicts ------------------------------------------ tidyverse_conflicts() --
## x dplyr::filter() masks stats::filter()
## x dplyr::lag()    masks stats::lag()
```

## Data Loading

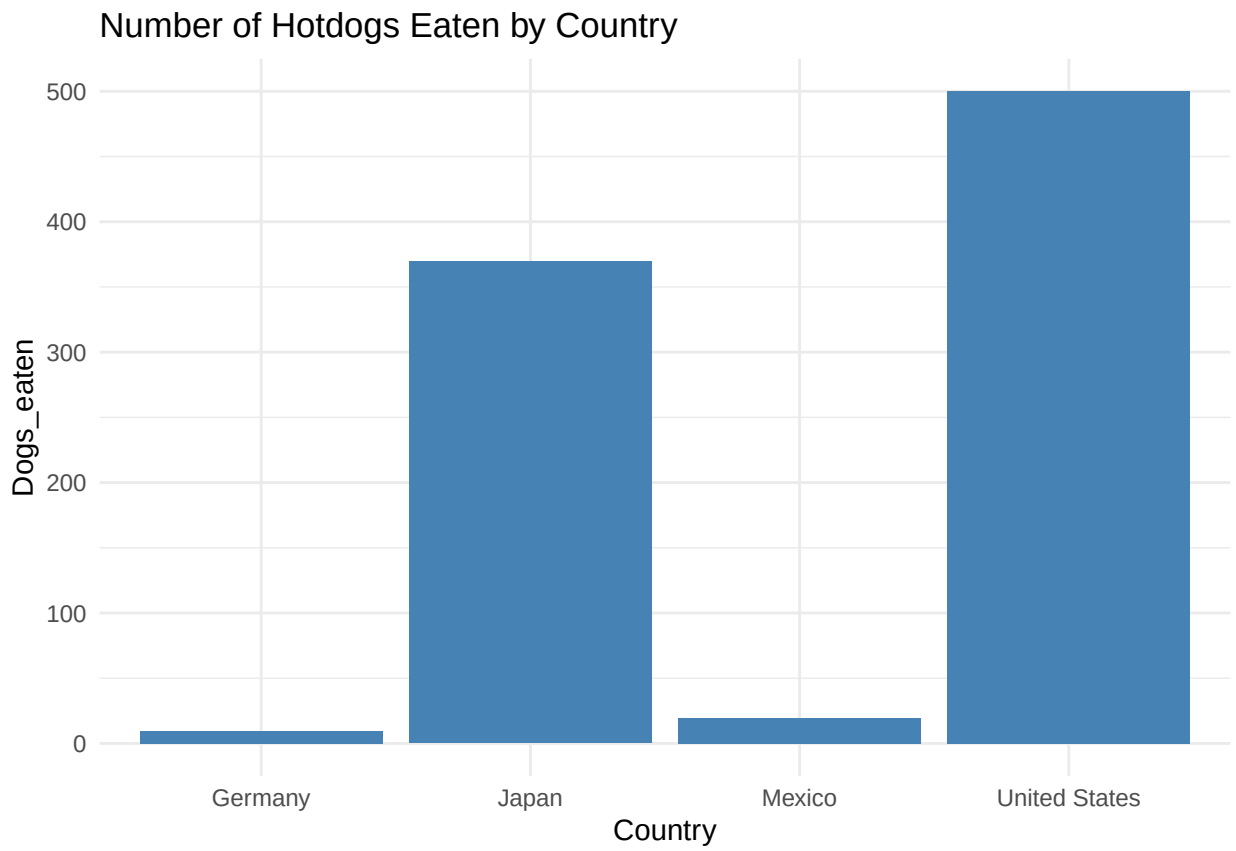
```
hotdog_df <- read_excel("E:/Personal/Bellevue University/Course/github/dsc640/Week 1&2/hotdog-contest-w
head(hotdog_df)
```

```
## # A tibble: 6 x 5
##    Year Winner                      Dogs_eaten Country       `New record`
##   <dbl> <chr>                            <dbl> <chr>                <dbl>
## 1  1980 Paul Siederman & Joe Baldini       9.1 United States            0
## 2  1981 Thomas DeBerry                    11   United States            0
## 3  1982 Steven Abrams                     11   United States            0
## 4  1983 Luis Llamas                       19.5 Mexico                   0
## 5  1984 Birgit Felden                      9.5 Germany                  0
## 6  1985 Oscar Rodriguez                   11.8 United States            0
```

```
# Total number of records present in the data set
nrow(hotdog_df)
```

```
## [1] 31
```

```
## Create Bar Chart

ggplot(hotdog_df, aes(x=Country, y=Dogs_eaten)) +
  geom_bar(stat = "identity", fill = 'steelblue')+
  ggtitle("Number of Hotdogs Eaten by Country")+
  theme_minimal()
```

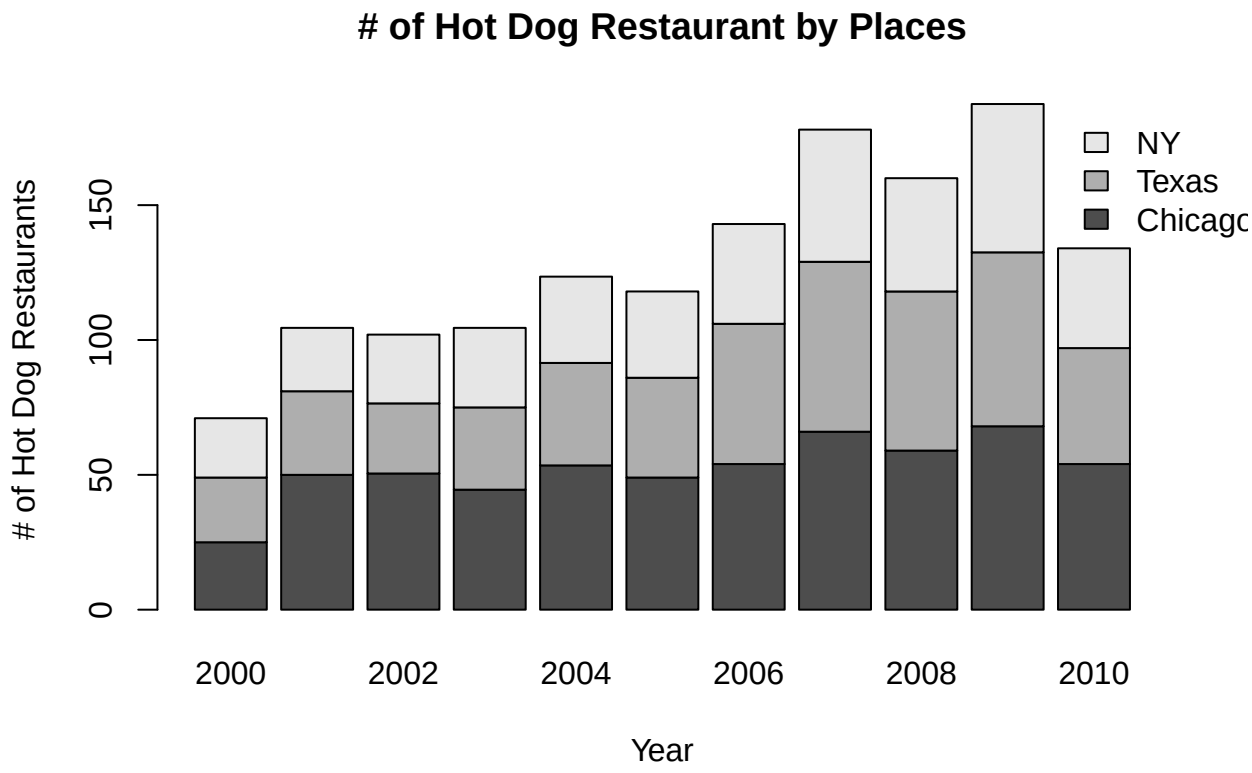



```
# Loadind the 2nd dataset for stacked bar
hdp_df <- read_excel("E:/Personal/Bellevue University/Course/github/dsc640/Week 1&2/hotdog-places.xlsm")

# Showing sample records
head(hdp_df)
```

```
## # A tibble: 3 x 11
##   `2000` `2001` `2002` `2003` `2004` `2005` `2006` `2007` `2008` `2009` `2010`
##    <dbl>  <dbl>  <dbl>  <dbl>  <dbl>  <dbl>  <dbl>  <dbl>  <dbl>  <dbl>  <dbl>
## 1     25   50     50.5   44.5   53.5     49     54     66     59   68       54
## 2     24   31     26     30.5   38       37     52     63     59   64.5     43
## 3     22   23.5   25.5   29.5   32       32     37     49     42   55       37
```

```
## Plotting stacked bar plot
barplot(as.matrix(hdp_df),  main="# of Hot Dog Restaurant by Places",
        xlab="Year",
```

```
        ylab="# of Hot Dog Restaurants",
        legend.text =  c('Chicago','Texas','NY'),
        args.legend = list(x = "topright",bty='n', inset=c(-0.1,0))
        )
```

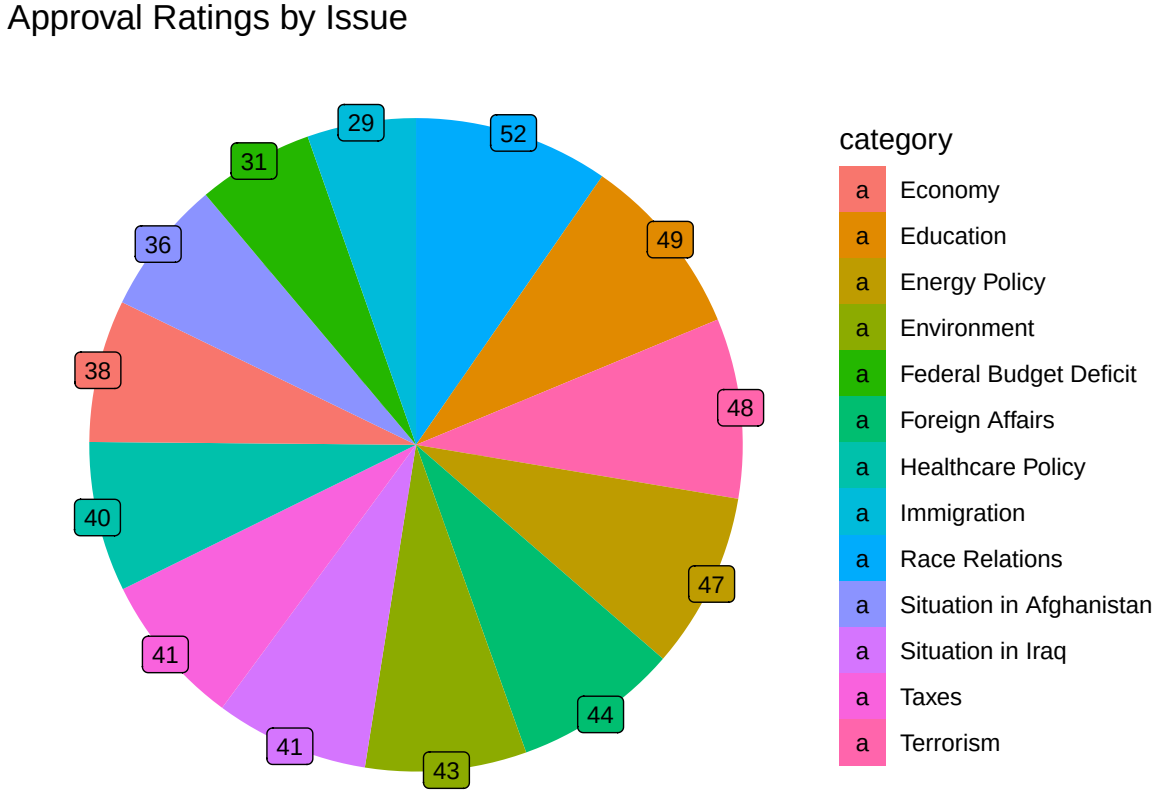



```
#Read in third file for the pie and donut charts
obama <- read_excel('E:/Personal/Bellevue University/Course/github/dsc640/Week 1&2/obama-approval-ratin

## Displaying few records
head(obama)
```

```
## # A tibble: 6 x 4
##   Issue           Approve Disapprove  None
##   <chr>             <dbl>      <dbl> <dbl>
## 1 Race Relations       52         38    10
## 2 Education            49         40    11
## 3 Terrorism            48         45     7
## 4 Energy Policy        47         42    11
## 5 Foreign Affairs      44         48     8
## 6 Environment          43         51     6
```

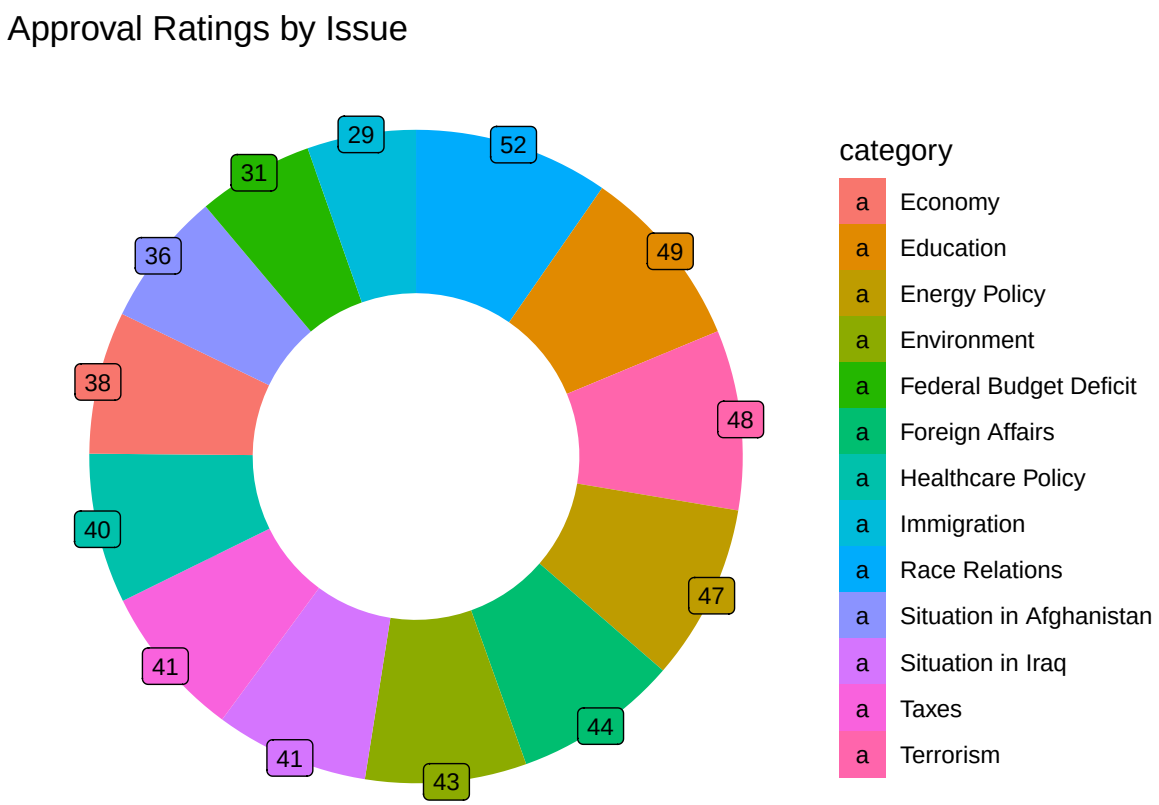
```
#Set up for pie and donut charts
#Creating smaller dataframe from original dataset
dat <- data.frame(count=c(obama$Approve), category=c(obama$Issue))
```

```
#Compute percentages
dat$fraction = dat$count / sum(dat$count)

#Compute the cumulative percentages
dat$ymax = cumsum(dat$fraction)

#Compute the bottom of each rectangle
dat$ymin = c(0, head(dat$ymax, n=-1))

#Compute label position
dat$labelPosition <- (dat$ymax + dat$ymin) / 2

#Compute a good label
dat$label <- paste0(dat$count)

#Pie chart
ggplot(dat, aes(ymax=ymax, ymin=ymin, xmax=4, xmin=3, fill=category)) +
  geom_rect() +
  geom_label( x=4, aes(y=labelPosition, label=label), size=3) +
  ggtitle("Approval Ratings by Issue") +
  coord_polar(theta="y") +
  theme_void()
```

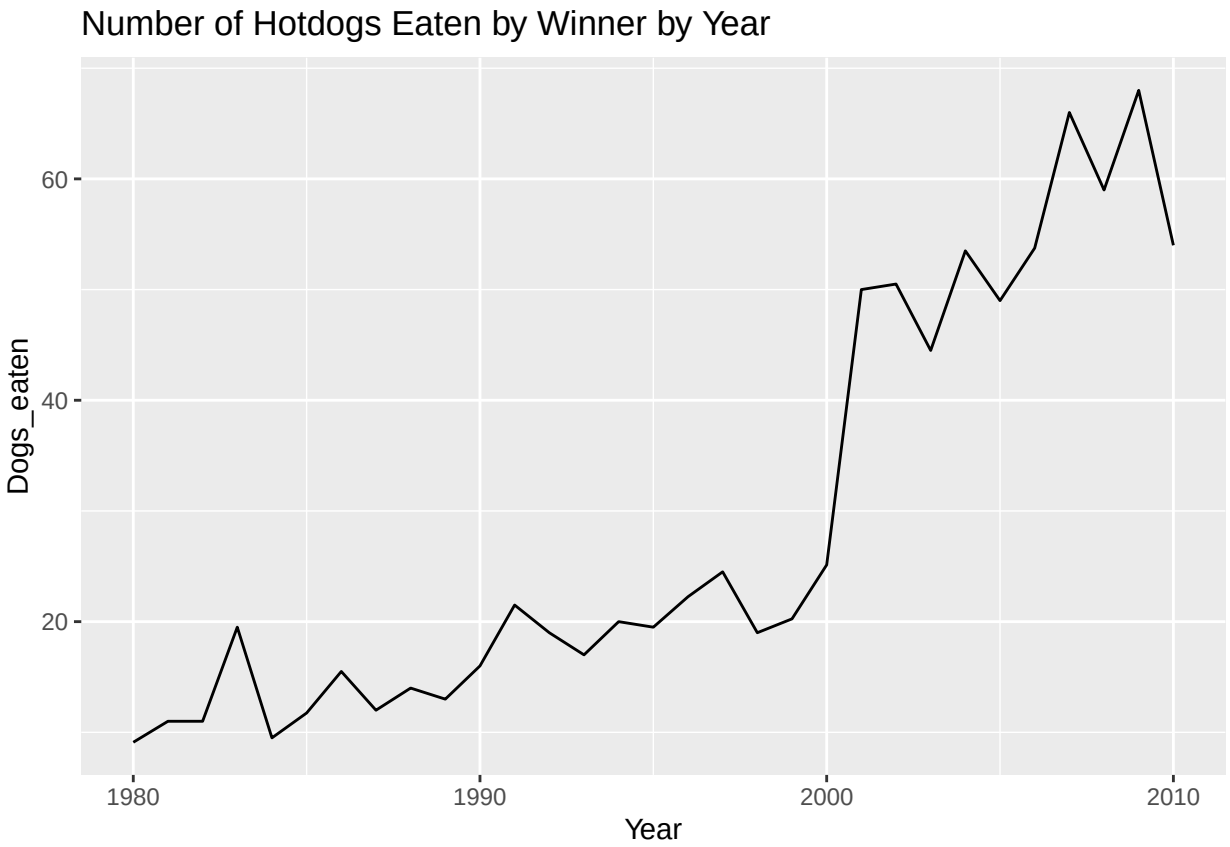



```
#Donut chart
ggplot(dat, aes(ymax=ymax, ymin=ymin, xmax=4, xmin=3, fill=category)) +
```

```
  geom_rect() +
  geom_label( x=4, aes(y=labelPosition, label=label), size=3) +
  ggtitle("Approval Ratings by Issue") +
  coord_polar(theta="y") +
  theme_void() +
  xlim(c(2, 4))
```




```
# We will use the same hotdog dataframe for this plot

ggplot(hotdog_df, aes(x=Year, y=Dogs_eaten)) +
  geom_line() +
  ggtitle("Number of Hotdogs Eaten by Winner by Year")
```


Number of Hotdogs Eaten by Winner by Year
Dogs_eaten
60
40
20
1980
1990
2000
2010
Year