```
---
title: "Young_DSC 640_Assignment 1.2_R"
author: "Bret Young"
date: "9/9/2020"
output:
  html_document:
    df_print: paged
  word_document: default
  pdf_document: default
---
```

````
```{r setup, include=FALSE}
knitr::opts_chunk$set(
    echo = TRUE,
    message = FALSE,
    warning = FALSE,
    fig.align='center',
    fig.pos='H',
    fig.path = "plots/",
    dev = c("png"),
    dpi=500
)
```
````

````
```{r}
# load library needed for Excel files
library(readxl)

# load data
file = 'hotdog-contest-winners.xlsm'
data_1 = read_xlsx(file, col_names = TRUE)
```
````

````
```{r}
# Print dtaa
head(data_1)
```
````

````
```{r}
# Load library for plotting data
library(ggplot2)

# Create Plot with annotations and custom features
ggplot(data_1, aes(Year, `Dogs eaten`, fill = as.factor(`New record`))) +
    geom_bar(stat = "identity") +
    scale_fill_manual(values = c("light gray", " light blue")) +
    ggtitle("Nathan's Hot Dog Eating Contest 1980 - 2010",
        subtitle = "Nathan's hot dog eatting contest happens ever year on July 4th.\nContestants compete against
eachother to see who can eat the most hot\ndogs and buns. Highlighted bars indicate new World Record set.") +
    labs(caption = "Source: Data Collected By Nathan Yau From Wikipedia",
        y = "Hot Dogs &\nBuns Eaten") +
    theme_classic() +
    theme(legend.position = "none",
        plot.title = element_text(face = "bold", size = 18),
        plot.subtitle = element_text(color = "light gray"),
        plot.caption = element_text(color = "light gray"),
        axis.title.y = element_text(angle = 0),
        ) +
    scale_x_continuous(breaks=seq(1980, 2010, 2)) +
    scale_y_continuous(breaks=seq(0, 70, 5)) +
    annotate("text", x = 1989, y = 25, label = "21.5 hot dogs eaten by\nFrank Dellarosa",
        size=2.5, hjust = "right") +
    annotate("segment", x = 1989.1, xend = 1991, y = 25, yend = 21.5) +
```
````

```
    annotate("text", x = 1999, y = 55, label = "Takeru Kobayashi's more\nthan doubles the\nexisiting World Record!",
size=2.5, hjust = "right") +
    annotate("segment", x = 1999.1, xend = 2001, y = 55, yend = 50)
```


```{r}
# Load dataset 2
file_2 = "hotdog-places.xlsm"
data_2 = read_xlsx(file_2, col_names = TRUE)
```


```{r}
head(data_2)
```

```{r}
# Rename rows of dat frame
row.names(data_2) = c("First", "Second", "Third")
# add row names to data for us in melt
data_2$Row_Names = row.names(data_2)
```


```{r}
#load library
library(data.table)
# Convert data to long version
data_melt <- melt(data_2, id.vars = )
print(data_melt)
```


```{r}
# Create stacked bar plot with custom styles
ggplot(data_melt, aes(x = variable, y = value, fill = Row_Names)) +
    geom_col(position = position_stack(reverse = T)) +
    scale_fill_manual(values = c("blue", "dark gray", "light gray")) +
    ggtitle("Top Three Competitors (2000 - 2010)",
        subtitle = "The top competitors for the Nathan's Hot Dog Eating Contests show\nsimilar numbers for 2000,
but for the next 5 years, the leader produced a\nlarge gap. In 2006 and on, the numbers eaten are more\nequally
distributed again.") +
    labs(caption = "Source: Data Collected By Nathan Yau From Wikipedia",
        x = "Year",
        y = "Hot Dogs &\nBuns Eaten") +
    theme_classic() +
    theme(plot.title = element_text(face = "bold", size = 18),
        plot.subtitle = element_text(color = "light gray"),
        plot.caption = element_text(color = "light gray"),
        axis.title.y = element_text(angle = 0),
        legend.title = element_blank()) +
    scale_y_continuous(breaks=seq(0, 200, 20))

```


```{r}
# Load third data file
file_3 = "obama-approval-ratings.xls"
data_3 = read_xls(file_3)
head(data_3)
```


```{r}
# Convert data to long form with melt
data_3_melt = melt(data_3)
print(data_3_melt)
```

```{r fig.height = 7, fig.width = 7}
# Create pie chart for each issue in data with facet
ggplot(data_3_melt, aes(x = "", y =value, fill = variable)) +
    geom_bar(stat="identity", width=1) +
    coord_polar("y", start=0) +
    ggtitle("Obama Approval Ratings By Issue",
            subtitle = "Approval ratings are calculated on percent of total answers for political issues.
            ") +
    labs(caption = "Source: Data Collected By Nathan Yau From Wikipedia") +
    theme_void() +
    theme(plot.title = element_text(face = "bold", size = 18),
        plot.subtitle = element_text(color = "light gray"),
        plot.caption = element_text(color = "light gray"),
        axis.text = element_blank(),
        axis.ticks = element_blank(),
        panel.grid  = element_blank()) +
    scale_fill_manual(values=c("#E69F00", "#56B4E9", "#999999"),
                name="",
                labels=c("Disapprove", "Approve", "None")) +
    geom_text(aes(label = paste0(round(value), "%")),
            position = position_stack(vjust = 0.5), color = "white", size = 2) +
    facet_wrap(facets = .~ Issue, labeller = label_value)
```

```{r}
library(dplyr)

data_3_donut = as.data.frame(melt(data_3[2:4])) %>% group_by(variable) %>%
    summarize(Sum = sum(value, na.rm = TRUE)) %>% mutate(percent = Sum / sum(Sum)) %>% mutate(ymax =
cumsum(percent)) %>% mutate(ymin = c(0, head(ymax, n=-1))) %>% mutate(pos = (ymax + ymin)/2)
print(data_3_donut)
```

```{r}
ggplot(data_3_donut, aes(ymax = ymax, ymin = ymin, xmax = 4, xmin = 3, fill = variable)) +
    geom_rect() +
    geom_text(x = 2, aes(y = pos, label = paste0(round(percent * 100), "%")),
            size = 4,
            color = c("#E69F00", "#56B4E9", "#999999")) +
    coord_polar(theta = "y") +
    xlim(c(-1, 4)) +
    ggtitle("Overall Obama Approval Rating",
            subtitle = "Approval ratings are combined into an overall rating.
            ") +
    labs(caption = "Source: Data Collected By Nathan Yau From Wikipedia") +
    theme_void() +
    theme(plot.title = element_text(face = "bold", size = 18),
        plot.subtitle = element_text(color = "light gray"),
        plot.caption = element_text(color = "light gray"),
        axis.text = element_blank(),
        axis.ticks = element_blank(),
        panel.grid  = element_blank()) +
    scale_fill_manual(values=c("#E69F00", "#56B4E9", "#999999"),
                name="",
                labels=c("Disapprove", "Approve", "None"))

```