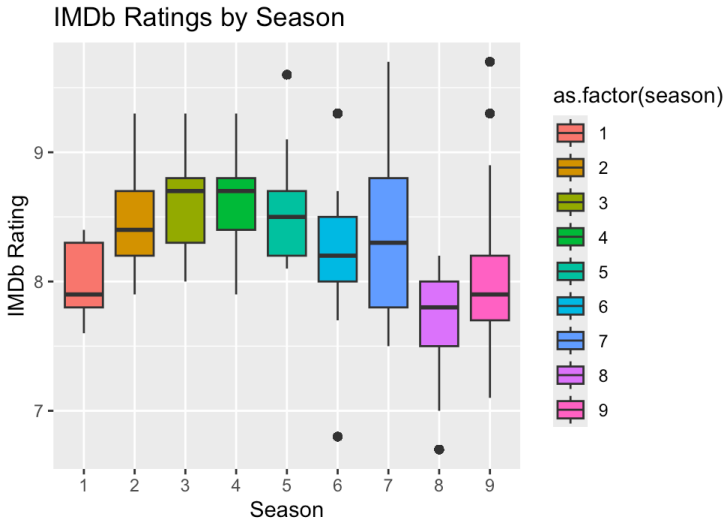
Utilizing R studio to analyze the office dataset. I utilized the following libraries to aid in the analysis: ggplot2, dplyr, gganimate, transformr

I created boxplots for episode IMDB ratings to analyze which season had the least and the most variability in ratings.

Code: # side-by-side boxplots for episode IMDB rating

```
boxplot_ratings <- theoffice |>
ggplot(aes( x= as.factor (season), y = imdb_rating, fill = as.factor(season))) +
  geom_boxplot() +
  labs(x = "Season", y = "IMDb Rating", title = "IMDb Ratings by Season") +
  scale_color_manual(values=c('skyblue4','orange','pink2', 'purple', 'blue','green','yellow','red'))
#Display of the Chart
boxplot_ratings
```



Season 4 has the least variability in ratings and season 7 has the most variability in ratings.

I created a bar chart showing the average number of lines per episode Dwight has per season.

```
#Filter the dataset for all Dwight's lines
dwight_lines <- theoffice |>
  filter(character == "Dwight") |>
```

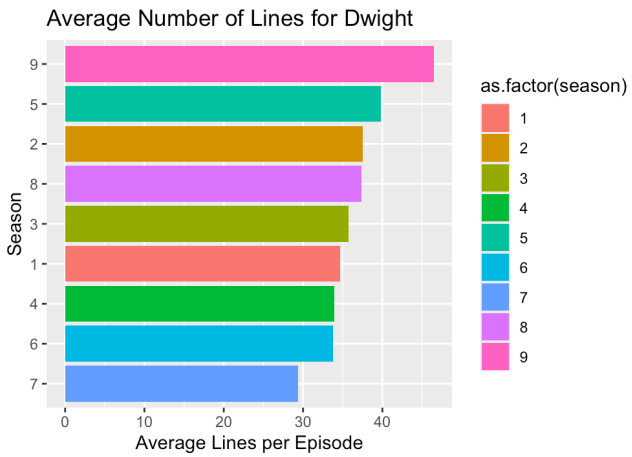
```
  group_by(season, episode) |>
  summarize(lines = n()) |>
  group_by(season) |>
  summarize(avg_lines = mean(lines, na.rm = TRUE))
```

#Step 2 Create the bar chart

```
barplot_dwight <- dwight_lines |>
  ggplot(aes(x = reorder(as.factor(season), avg_lines), y = avg_lines, fill = as.factor(season)))+
  geom_bar(stat = "identity") +
  coord_flip() +
  labs(x = "Season", y = "Average Lines per Episode", title = "Average Number of Lines for Dwight") +
  scale_color_manual(values=c('skyblue4','orange','pink2', 'purple', 'blue','green','yellow','red'))
```

#Display chart

```
barplot_dwight
```



I created a bar chart to display the maximum IMDb rating per season.

Question 3. #Step 1 Summarizing the ratings data

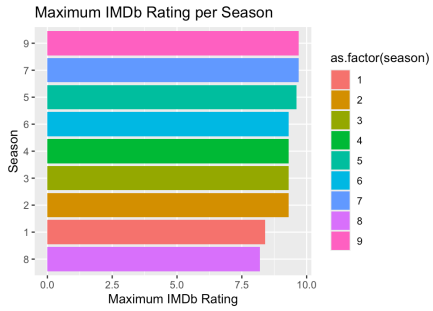
```
max_ratings <- theoffice |>
  group_by(season) |>
  summarize(max_rating = max(imdb_rating, na.rm = TRUE))
```

#Step 2 Plotting the chart

```
barplot_max_rating <- max_ratings |>
  ggplot(aes(x = reorder(as.factor(season), max_rating),
             y = max_rating, fill = as.factor(season))) +
  geom_bar(stat = "identity") +
  coord_flip() +
  labs(x = "Season", y = "Maximum IMDb Rating", title = "Maximum IMDb Rating per Season")+
  scale_color_manual(values=c('skyblue4','orange','pink2', 'purple', 'blue','green','yellow','red'))
```

#Displaying the chart

```
barplot_max_rating
```



Lastly, I created an animated bar chart to show how the average number of lines per episode for three minor characters Darryl, Oscar, and Toby changes over the 9 seasons.

#Step 1 Filtering the data

```
minor_characters <- theoffice |>
  filter(character %in% c("Darryl", "Oscar", "Toby")) |>
```

```
  group_by(season, episode, character) |>
  summarize(lines = n()) |>
  group_by(season, character) |>
  summarize(avg_lines = mean(lines, na.rm = TRUE))
```

#Step 2 Creating the chart

```
animated_plot <- minor_characters |>
    ggplot(aes(x = reorder(character, avg_lines), y = avg_lines, fill = character)) +
    geom_bar(stat = "identity") +
    coord_flip() +
    labs(x = "Character", y = "Average Lines per Episode", title = "Average Number of Lines per Episode by Character: Season {frame_time}") +
    theme_minimal() +
    transition_states(season, transition_length = 2, state_length = 1) +
    ease_aes()
```

#Step 3 Creating the animated gif

```
animate(animated_plot, nframes = 150, fps = 10, width = 600, height = 400, renderer = gifski_renderer("minor_characters_lines.gif"))
```