

# Final Project First Draft

Curtis Leaver

## Import Packages

```
library(tidyverse)
library(ggplot2)
library(dplyr)
library(skimr)
library(ggthemes)
```

## Import Data

```
library(tidyverse)
library(skimr)
library(readr)
library(utf8)

file_encoding <- guess_encoding("week7_starwars.csv")$encoding
cat("File Encoding:", file_encoding, "\n")
```

File Encoding: windows-1252

```
starwars <- read_csv("week7_starwars.csv", locale = locale(encoding = file_encoding)) %>%
  mutate(across(where(is.character), utf8_encode))
```

## Data Cleaning

```
starwars <- starwars |>
  janitor::clean_names()
cleandata <- starwars |>
  select("have_you_seen_any_of_the_6_films_in_the_star_wars_franchise", "x22", "x6")
cleandata <- cleandata |>
  rename("darth_vader_opinion" = "x22") |>
  rename("seen_any" = "have_you_seen_any_of_the_6_films_in_the_star_wars_franchise") |>
  rename("seen_ep_3" = "x6")
cleandata <- cleandata |>
  mutate(seen_ep_3 = ifelse(is.na(seen_ep_3), "No", seen_ep_3)) |>
  mutate(seen_ep_3 = recode(seen_ep_3, "Star Wars: Episode III Revenge of the Sith" = "Yes"))
cleandata <- cleandata |>
  drop_na() |>
  select("darth_vader_opinion", "seen_ep_3") |>
  slice(-1)
favorability_order <- c(
  "Unfamiliar (N/A)",
  "Very unfavorably",
  "Somewhat unfavorably",
  "Neither favorably nor unfavorably (neutral)",
  "Somewhat favorably",
  "Very favorably"
)
cleandata$darth_vader_opinion <- factor(
  cleandata$darth_vader_opinion,
  levels = favorability_order
)
percentage_data <- cleandata |>
  group_by(seen_ep_3, darth_vader_opinion) %>%
  summarise(count = n()) %>%
  group_by(seen_ep_3) %>%
  mutate(percentage = count / sum(count) * 100)
```

## The Visualization

```
star_wars_theme <- theme_minimal() +
  theme(
    text = element_text(family = "Star Jedi", color = "white"),
    plot.title = element_text(size = 12),
    axis.title = element_text(size = 6),
    axis.text = element_text(size = 9, color = "white"),
    axis.line = element_line(color = "white"),
    panel.grid.major = element_blank(),
    panel.grid.minor = element_blank(),
    plot.background = element_rect(fill = "black"),
    panel.background = element_rect(fill = "black")
  )
ggplot(percentage_data, aes(x = seen_ep_3, y = percentage, fill = darth_vader_opinion)) +
  geom_bar(stat = "identity", position = "stack") +
  scale_fill_manual(values = c(
    "Very favorably" = "darkgreen",
    "Somewhat favorably" = "lightgreen",
    "Neither favorably nor unfavorably (neutral)" = "gray",
    "Somewhat unfavorably" = "lightcoral",
    "Very unfavorably" = "darkred",
    "Unfamiliar (N/A)" = "white"
  )) +
  #, labels = c("unfamiliar (N/A)", "very unfavorably", "somewhat unfavorably", "neither f
  labs(
    title = "Favorability Of Darth Vader Is Higher For Viewers of Star Wars: Episode 3",
    x = "Have you seen Star Wars: Episode 3 Revenge of the Sith?",
    y = "Percentage",
    fill = "Favorability of Darth Vader"
  ) +
  scale_y_continuous(labels = scales::percent_format(scale = 1)) +
  theme_minimal() +
  theme(
    panel.background = element_rect(fill = "black"),
    plot.background = element_rect(fill = "black"),
    text = element_text(color = "white", size = 10), # Adjust size as needed
    axis.title = element_text(size = 7), # Adjust size as needed
    axis.text = element_text(size = 6, color = "white"),
    plot.title = element_text(size = 10, face = "bold"),
    panel.grid.major = element_blank(),
```

```
    panel.grid.minor = element_blank()  
)  
  
Favorability Of Darth Vader Is Higher For Viewers of Star Wars: Episode 3
```

