Qualitative Visualization

Qualitative Research Questions

1. How does genre affect voting average?
2. How does genre affect runtime?
3. How does genre affect revenue?
4. What are the ratings, revenues, and runtimes across 3 popular movie franchises?

Load Packages

```
Attaching package: 'janitor'

The following objects are masked from 'package:stats':

    chisq.test, fisher.test


Attaching package: 'dplyr'

The following object is masked from 'package:kableExtra':

    group_rows

The following objects are masked from 'package:stats':

    filter, lag

The following objects are masked from 'package:base':

    intersect, setdiff, setequal, union
```

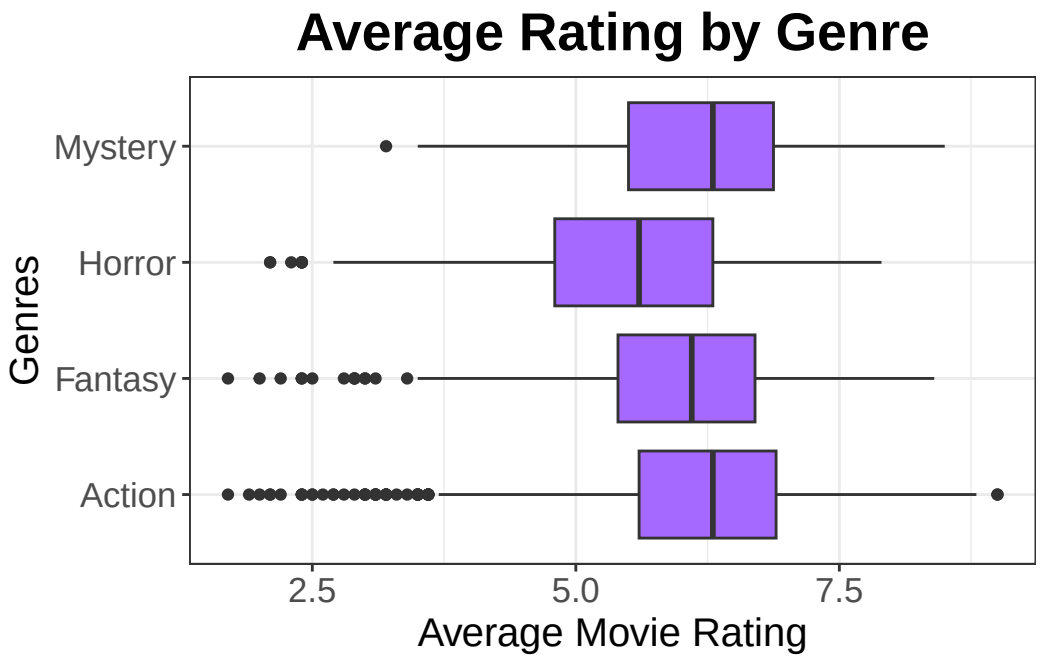
```
-- Attaching core tidyverse packages ------------------------ tidyverse 2.0.0 --
v forcats   1.0.0     v stringr   1.5.1
v lubridate 1.9.3     v tibble    3.2.1
v purrr     1.0.2     v tidyr     1.3.1
v readr     2.1.5
-- Conflicts ---------------------------------------- tidyverse_conflicts() --
x dplyr::filter()     masks stats::filter()
x dplyr::group_rows() masks kableExtra::group_rows()
x dplyr::lag()        masks stats::lag()
i Use the conflicted package (<http://conflicted.r-lib.org/>) to force all conflicts to becom

Attaching package: 'rvest'


The following object is masked from 'package:readr':

    guess_encoding



Attaching package: 'plotly'


The following object is masked from 'package:ggplot2':

    last_plot


The following object is masked from 'package:stats':

    filter


The following object is masked from 'package:graphics':

    layout
```

Read in, Clean, and Wrangle Data

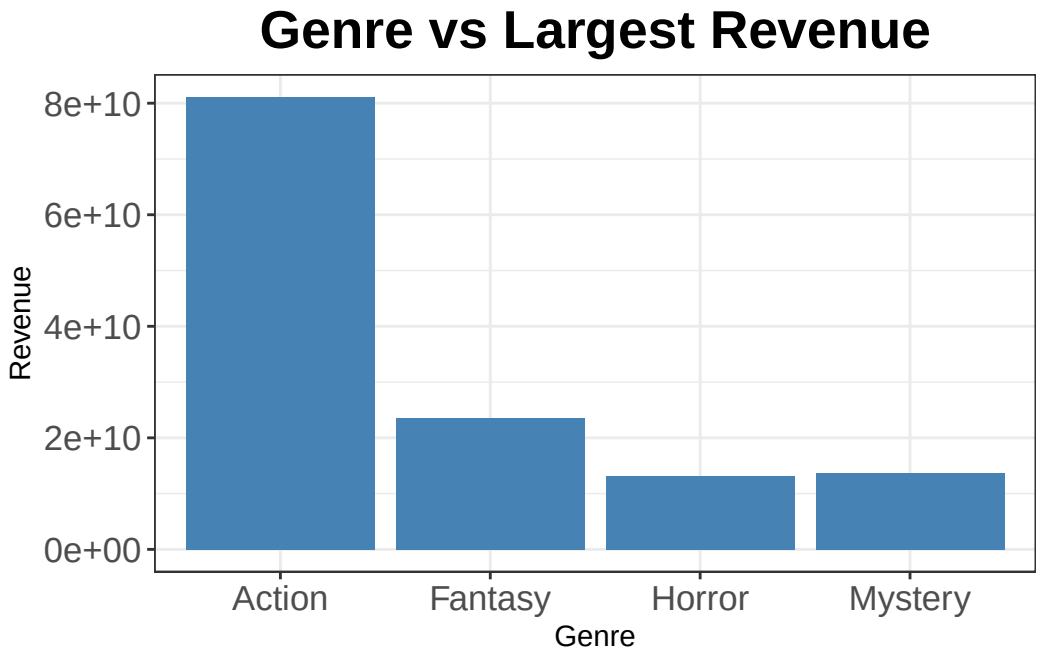
```
Rows: 17163 Columns: 14
-- Column specification --------------------------------------------------------
Delimiter: ","
```

```
chr (11): movie_id, movie_name, year, certificate, runtime, genre, descripti...
dbl  (3): rating, votes, gross(in $)

i Use `spec()` to retrieve the full column specification for this data.
i Specify the column types or set `show_col_types = FALSE` to quiet this message.
Rows: 52452 Columns: 14
-- Column specification --------------------------------------------------------
Delimiter: ","
chr (11): movie_id, movie_name, year, certificate, runtime, genre, descripti...
dbl  (3): rating, votes, gross(in $)

i Use `spec()` to retrieve the full column specification for this data.
i Specify the column types or set `show_col_types = FALSE` to quiet this message.
Rows: 36682 Columns: 14
-- Column specification --------------------------------------------------------
Delimiter: ","
chr (11): movie_id, movie_name, year, certificate, runtime, genre, descripti...
dbl  (3): rating, votes, gross(in $)

i Use `spec()` to retrieve the full column specification for this data.
i Specify the column types or set `show_col_types = FALSE` to quiet this message.
Rows: 18960 Columns: 14
-- Column specification --------------------------------------------------------
Delimiter: ","
chr (11): movie_id, movie_name, year, certificate, runtime, genre, descripti...
dbl  (3): rating, votes, gross(in $)

i Use `spec()` to retrieve the full column specification for this data.
i Specify the column types or set `show_col_types = FALSE` to quiet this message.
Joining with `by = join_by(movie_id, movie_name, year, certificate, runtime, genre, rating, c
Joining with `by = join_by(movie_id, movie_name, year, certificate, runtime, genre, rating, c
Joining with `by = join_by(movie_id, movie_name, year, certificate, runtime, genre, rating, c
```

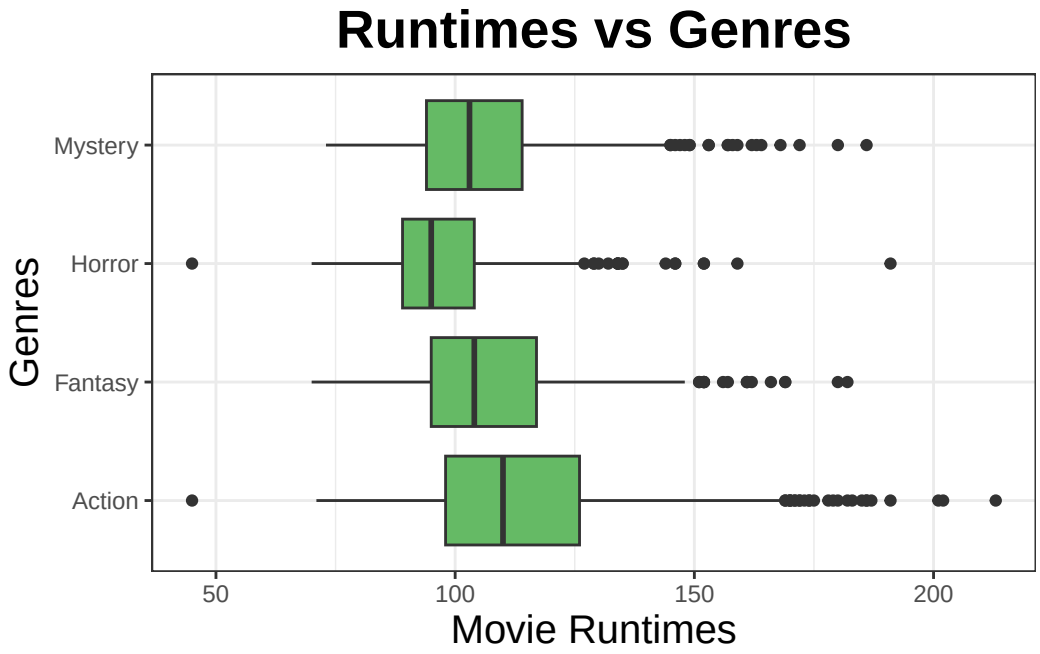
Genre and Rating



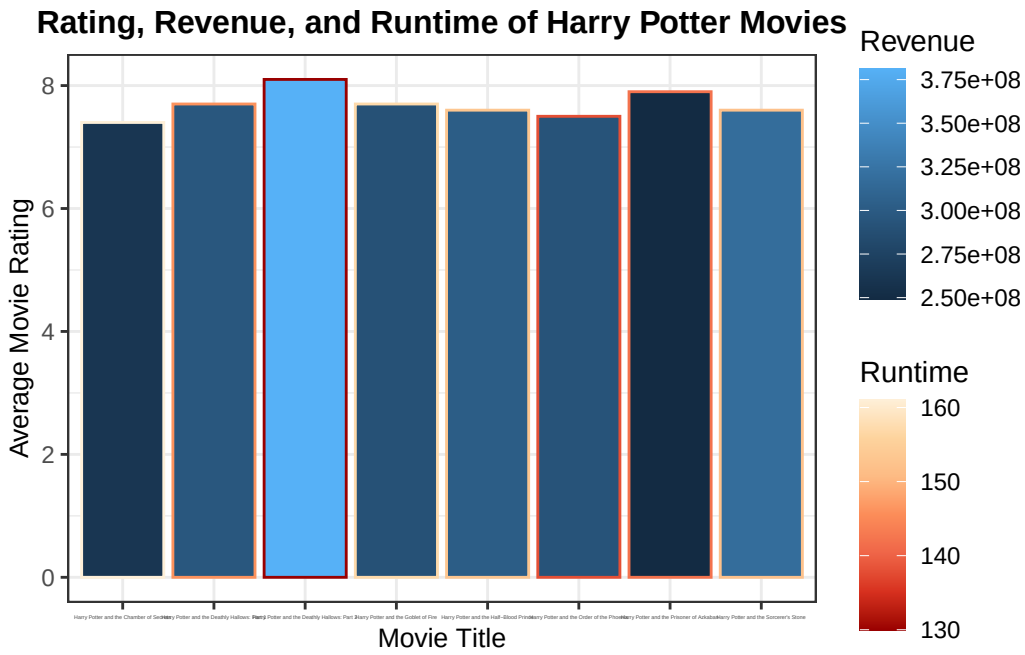
Genre and Revenue



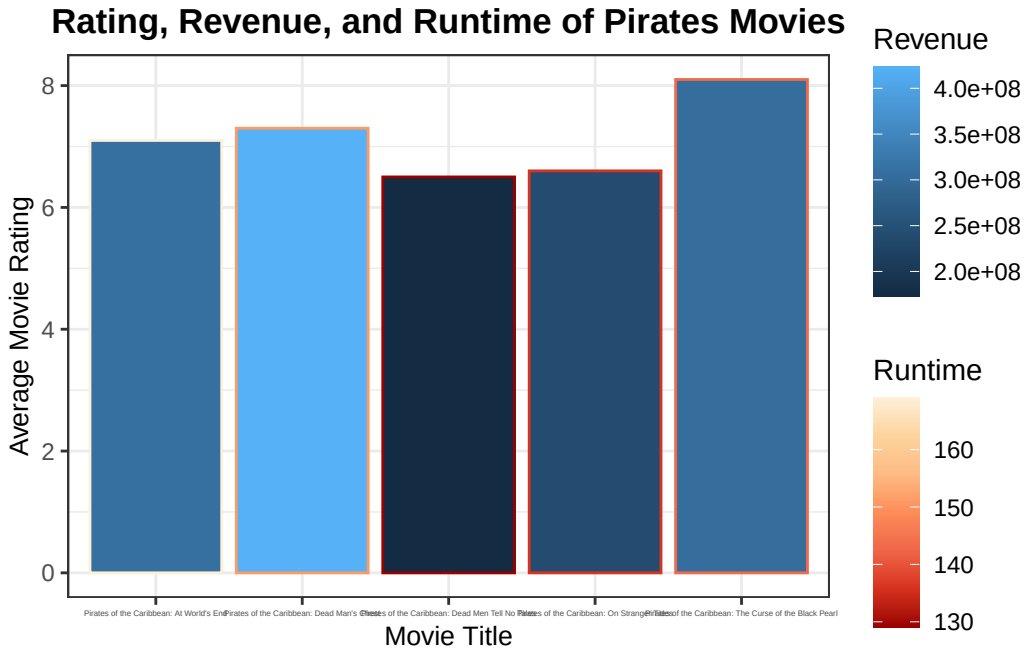
Genre vs Runtime



Harry Potter Movies



Pirates of the Caribbean Movies



Spiderman Movies