D_607_assign_2

Kevin Havis

2024-08-31

```r
movies <- read_csv("movie_survey_responses.csv")
```

```
## Rows: 6 Columns: 11
## -- Column specification ------------------------------------------------------
## Delimiter: ","
## chr (11): Timestamp, Alien: Romulus, The Barbie Movie, Dune: Part Two, Oppen...
##
## 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.
```

```r
glimpse(movies)
```

```
## Rows: 6
## Columns: 11
## $ Timestamp                              <chr> "8/31/2024 21:20:17", "8/31~
## $ `Alien: Romulus`                       <chr> "Have not seen", "Have not ~
## $ `The Barbie Movie`                     <chr> "Have not seen", "Have not ~
## $ `Dune: Part Two`                       <chr> "Have not seen", "Have not ~
## $ Oppenheimer                            <chr> "Have not seen", "Great", "~
## $ `Poor Things`                          <chr> "Have not seen", "Have not ~
## $ `Killers of the Flower Moon`           <chr> "Have not seen", "Have not ~
## $ `Asteroid City`                        <chr> "Have not seen", "Have not ~
## $ `Deadpool & Wolverine`                 <chr> "Have not seen", "Have not ~
## $ `First Name`                           <chr> "Corey", "Keith", "Corey", ~
## $ `What is your favorite movie of all time?` <chr> NA, "Lord of the Rings", "A~
```

```r
# Drop timestamp column and clean column names
movies <- movies |>
  select(everything(), -Timestamp) |>
  rename('reviewer_name' = 'First Name', 'favorite_movie' = 'What is your favorite movie of all time?')
```

```r
# Extract movies into a new tibble
movie_table <- tibble(names(movies)) |>
  rename('film' = 'names(movies)') |>
  filter(film != 'reviewer_name' & film != 'favorite_movie')

glimpse(movie_table)
```

```
## Rows: 8
## Columns: 1
## $ film <chr> "Alien: Romulus", "The Barbie Movie", "Dune: Part Two", "Oppenhei~
```

```r
# Extract favorite movies and append to move_table
fav_movies <- movies |>
  select(favorite_movie) |>
  filter(!is.na(favorite_movie)) |>
  rename('film' = 'favorite_movie') # allows us to join with movie_table

movie_table <- bind_rows(movie_table, fav_movies)

glimpse(movie_table)
```

```
## Rows: 13
## Columns: 1
## $ film <chr> "Alien: Romulus", "The Barbie Movie", "Dune: Part Two", "Oppenhei~
```

```r
# Reformat the reviews table
reviews_table <- movies |>
  pivot_longer(cols =(0:8), names_to = 'film', values_to = 'rating') |>
  select(everything(), -'favorite_movie')

# Add numerical rating field
reviews_table <- reviews_table |>
  mutate(num_rating = case_when(
    rating == 'Great' ~ 5,
    rating == 'Good' ~ 4,
    rating == 'Okay' ~ 3,
    rating == 'Poor' ~ 2,
    rating == 'Bad' ~ 1,
    TRUE ~ NA
  ))
```

```r
# Get audience member names
reviewer_table <- movies |>
  select(reviewer_name, favorite_movie)
```

```r
con <- DBI::dbConnect(duckdb::duckdb()) #, dbdir = "limp_lettuce")
```

```r
dbWriteTable(con, "movies", movie_table, overwrite = TRUE)

dbWriteTable(con, "reviews", reviews_table, overwrite = TRUE)

dbWriteTable(con, "reviewers", reviewer_table, overwrite = TRUE)
```

```r
dbListTables(con)
```

```
## [1] "movies"    "reviewers" "reviews"
```