

IMDB THE SQL

STARRING: ADAM, ANDREW, MELINA, AND DERRICK

Project Intention

To research and discuss winners of Various Academy Awards (Oscars) and show the representation (or lack thereof) in categories like Best Actor, Best Actress, and Best Picture.

To do this, we would find data sets about the Academy Awards, read them into Pandas, create a new CSV file with all the data, then upload this into a SQL database

Coding/Libraries Used

Python/Pandas

PySpark/API

SQL/SQLite



Our Process

- 1) Find usable datasets
- 2) Create dataframes in Pandas
- 3) Rename columns, merge dataframes, delete redundancies/irrelevant columns, then export this as a new CSV
- 4) Create SQL Database/Tables
- 5) Import CSV files
- 6) Create API

Datasets/Python/Pandas

Step 1: Find Data!

Retrieved from Kaggle

Dataset Titles:

Oscar Winners - Director.csv

Oscars.csv

the_oscar_award.csv



Step 2: Read into Pandas

Once settled, import these CSVs into Pandas

-Delete "Oscar Winners- Director" (redundant)

-rename columns to have like names

Next Step: Data Problems/Merging DFs

Representation- Black or White

Go through CSV files and edit them one line at a time to reflect actual race of the people involved

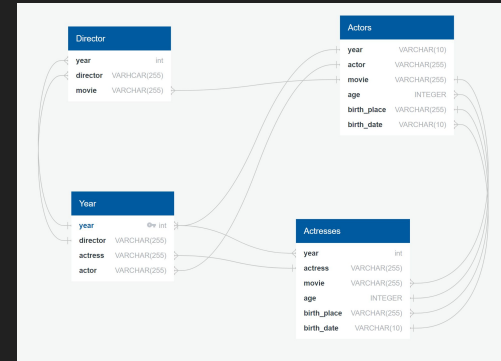
Once that was solved, merge the Data Frames together and output a new CSV file (“full_oscar.csv”) to be used for the SQL Database



Melina talks SQL

- 1.) Choose Valuable Data
- 2.) Create ERD
- 3.) Create Tables
- 4.) Import Data
- 5.) Finish Tables

- 1.) Assess Data
- 2.) What is it telling us?
- 3.) Is age a factor?
- 4.) Is race a factor?



Your Friendly Neighborhood API

#Prompt/Interface for User

```
def userInterface():
```

```
    print("Greetings user and welcome to the IMDB The Sequel API!");
```

```
    print("");
```

```
    print("This database is a compilation of oscar winning films from 1928 until 2023.");
```

```
    print("To query the database, you may chose to enter the following commands 1) year span
```

```
    userResponse = input("Please enter your search field: ");
```

```
    if userResponse == "year span":
```

```
        #grab each year and store it
```

```
        startYr = input("Please enter the starting year: ")
```

```
        endYr = input("Please enter the end year: ")
```

```
        readtoCSV(userInterfaceYear(startYr, endYr))
```

```
    elif userResponse == "nominee name":
```

```
        name = input("Please enter the nominee name to search for: ")
```

```
        userInterfaceName(name)
```

```
    elif userResponse == "director name":
```

```
        name = ("Please enter the director name to search for: ")
```

```
        userInterfaceDir(name);
```

```
else:
    1 oscar_no,oscar_yr,award,name,movie,age,birth_pl,birth_date,birth_mo,birth_d,birth_y,index,Gender,Race,Director(s),Film,Nomination/Winner
```

```
    print("-----2 1,1929,Best actress,Janet Gaynor,7th Heaven,22,Pennsylvania,1906-10-06,10,6,1906,5,Male,White,Frank Lloyd,The Divine Lady,Winner
```

```
    print("Input was found to3 1,1929,Best actress,Janet Gaynor,7th Heaven,22,Pennsylvania,1906-10-06,10,6,1906,6,Male,White,Lionel Barrymore,Madame X,Nomination
```

```
    print("-----4 1,1929,Best actress,Janet Gaynor,7th Heaven,22,Pennsylvania,1906-10-06,10,6,1906,7,Male,White,Harry Beaumont,The Broadway Melody,Nominati
```

```
5 1,1929,Best actress,Janet Gaynor,7th Heaven,22,Pennsylvania,1906-10-06,10,6,1906,8,Male,White,Irving Cummings,In Old Arizona,Nomination
```

```
6 1,1929,Best actress,Janet Gaynor,7th Heaven,22,Pennsylvania,1906-10-06,10,6,1906,9,Male,White,Frank Lloyd,Drag,Nomination
```

```
7 1,1929,Best actress,Janet Gaynor,7th Heaven,22,Pennsylvania,1906-10-06,10,6,1906,10,Male,White,Frank Lloyd,Weary River,Nomination
```

```
8 1,1929,Best actress,Janet Gaynor,7th Heaven,22,Pennsylvania,1906-10-06,10,6,1906,11,Male,White,Ernst Lubitsch,The Patriot,Nomination
```

```
9 1,1929,Best actor,Emil Jannings,The Way of All Flesh & The Las,44,Switzerland,1884-07-23,7,23,1884,5,Male,White,Frank Lloyd,The Divine La
```

```
0 1,1929,Best actor,Emil Jannings,The Way of All Flesh & The Las,44,Switzerland,1884-07-23,7,23,1884,6,Male,White,Lionel Barrymore,Madame X
```

```
1 1,1929,Best actor,Emil Jannings,The Way of All Flesh & The Las,44,Switzerland,1884-07-23,7,23,1884,7,Male,White,Harry Beaumont,The Broadw
```

✓ 0.0s



Conclusions/observations

1. Summary of findings:
2. Representation in best Actor category:
3. Representation in best Actress category:
4. Representation in best Picture category:
5. Intersectionality and diversity:
6. Limitations:
7. Conclusions: