**\*Loading importing libraries**\*

import pandas as pd

import matplotlib.pyplot as plt

import seaborn as sns

import numpy as np

**\* Importing csv file into panda datframe\***

pd.read\_csv('https://raw.githubusercontent.com/Waderonan/UCDPA\_RonanWade/main/IMDB\_movie\_reviews\_details.csv')

movies = pd.read\_csv('https://raw.githubusercontent.com/Waderonan/UCDPA\_RonanWade/main/IMDB\_movie\_reviews\_details.csv')

print(movies.head())

**\*Cleaning up data and Analysing\***

movies.isnull().sum()

movies.gross.fillna("unknown",inplace= True)

**\*Replace missing values\***

movies["year"].value\_counts()

movies.gross.fillna("unknown",inplace= True)

**\*sorting, indexing, grouping\***

print(movies.info())

print(movies.index)

print(movies.set\_index("name"))

print(movies.describe())

print(movies.dtypes)

movies.columns

movies.dtypes

**\*Slicing\***

movies[0:10]

movies.iloc[0:10, 1:7]

movies.iloc[0:20, 1:7]

pd.unique(movies['genre'])

movies['year'].count()

**\*iterrows\***

movies.iterrows

next(movies.iterrows())

row = next(movies.iterrows())[1]

movies.head(n=2).iterrows()

index = movies.head(n=2).iterrows()

print(index, row)

print(index, row['year'], row['runtime'], row['genre'])

\*Numpy\*

movies = pd.read\_csv('https://raw.githubusercontent.com/Waderonan/UCDPA\_RonanWade/main/IMDB\_movie\_reviews\_details.csv')

year1 = np.array([movies['year']])

print(year1.dtype)

print(year1.shape)

**\*Matplotlib\***

x=movies.genre.value\_counts().head(20)

plt.figure(figsize=(10,10))

ax=sns.barplot(x.values,x.index)

ax.set\_xlabel("Number of Movies ")

ax.set\_ylabel("Genre")

plt.title('Ranking of Movies by Genre'

x=movies.year.value\_counts().head(20)

plt.figure(figsize=(10,10))

ax=sns.barplot(x.values,x.index)

ax.set\_xlabel("Number of Movies Made")

ax.set\_ylabel("Year of Movie")

plt.title('Ranking of Movies by Year')

x=movies.runtime.value\_counts().head(10)

plt.figure(figsize=(16,10))

plt.xlabel("Runtime in Minutes")

plt.ylabel("Number of Movies")

sns.barplot(x=x.index ,y= x.values)

plt.title('Top 10 Runtimes')

**\*Seaborn\***

**sns.histplot(movies['runtime'])**

**plt.show**

**sns.stripplot(x="runtime",y="genre",data=movies)**

**plt.ylim(0,10)**

**plt.xlim(0,200)**

**plt.show**