ge-saka / **Phase-ONE-Project**

Type / to search

<> **Code**   Issues   Pull requests   Actions   Projects   Wiki   Security   Insights   Settings

Phase-ONE-Project Public

main   1 Branch   0 Tags   Go to file   Go to file   +   About   Code   ···

ge-saka Update README.md   bf797ff · 1 minute ago   2 Commits

README.md   Update README.md   1 minute ago

About

Analyzing Box office movies using three datasets to help Microsoft open a New movie Studio

📖 Readme

Activity

☆ 0 stars

👁 1 watching

⑂ 0 forks

README

Phase-ONE-Project

Analyzing Box office movies using three datasets to help Microsoft open a New movie Studio

Microsoft has decided to start creating video content by starting a new movie studio. In this project we explore what types of films are currently doing the best in the box office and translate the findings into actionable insights.

I use three datasets. The imdb database,tn dataset and bom.movie_gross.csv file. I import the libraries required to data clean these data sets.

Releases

No releases published
Create a new release

Packages

No packages published
Publish your first package

for data manipulation

import pandas as pd

for numeric computation

import numpy as np

for ploting our data and creating visualization

import matplotlib.pyplot as plt %matplotlib inline import seaborn as sns

To ignore warnings

import warnings

Connect to the database

read csv files and merge using python code.

we clean these data sets by checking for missing values and dropping them.

checking for missing values

data.isna().sum()

dropping missing values

data.dropna(inplace=True) also can replace with mean mode or median

exploratory data analysis

plot graphs and plots,histogram

Conclusion

given from thevisualizations

Recommendations

Derived from your findings