ITD105–FM254

**Members:**

Cañete, Alenne Jasper

Halinon, Aldrin

Gallano, Cyrus Jade

**Beyond the Screen:**

The Exploration with Cinemalytics to Unveil Box Office Big Data

**Project Description**

The box office analytics project, Cinemalytics, unveils a meticulously developed box office web application alongside the backend prowess of Django, with an unwavering focus on Big Data approaches and SQLite3 as the database foundation. Dive into the core of this project, wherein the primary goal transcends conventional boundaries, which aims to construct a comprehensive box office analytics project that goes further with the integration of Big Data processing techniques to build a scalable box office prediction model and implement visualization charts for thorough data insights.

**Implemented Features**

1. **Big Data Processing**

Utilizes the Big Data processing techniques for an efficient management of large–scale box office information.

1. **Data Cleansing**

Implements thorough cleansing of box office data in order to leverage the capabilities of Big Data processing.

1. **Descriptive Analytics**

Incorporates stellar functions for quick and efficient retrieval of summary statistics, key metrics, and data distribution.

1. **Predictive Analytics**

Employs a multiple linear regression model to predict box office performance based on extensive box office features.

1. **Data Visualization**

Utilizes the bell curve, line graph, histogram, radar chart, scatter plot, and correlation matrix based on box office features to provide comprehensive insights.

**Table of Contents**

1. Data and Database
2. Big Data–Driven Box Office Prediction
3. Visualization
4. Deployment Instructions

**Data and Database**

The box office analytics project, Cinemalytics, sources the box office information from The TMDb (The Movie Database), a comprehensive film database that provides information concerning films, which includes details such as titles, ratings, release dates, revenue, genres, and so on, to be employed for Big Data processing techniques. The sourced database contains more than 963,000 various films which were released as far as from the year 1899 to 2023. By means of this, those details such as genres, adult, vote count, popularity, and vote average offer a nuanced understanding of audience preferences, film reception, and marketing appeal; poster path helps to visualize the nature and impression within the film; while the budget contributes as the core for the box office status based on 2.5x break–even multiplier and revenue prediction.

**Big Data–driven Box Office Prediction**

Within the box office analytics project, Cinemalytics, a multiple linear regression model, was integrated into the developed web application that draws upon an extensive array of box office features with an R–squared value of at least 0.60. Accordingly, the web application embraces user–centric design to let box office enthusiasts and industry professionals contribute to such predictive prowess. Users can seamlessly input data across various attributes – budget, vote count, vote average, popularity, and genres – which fosters an interactive and intuitive experience that transcends the conventional boundaries of revenue prediction in the cinematic realm.

**Data Visualization**

The analytics journey of this box office analytics project, Cinemalytics, extends beyond mere prediction; as such employs a diverse arsenal of visualization techniques. From the nuanced depiction of value versus residual distributions through a bell curve and histogram to the dynamic portrayal of value trends via line graph and scatter plot, the platform transforms raw data into visual narratives. Accordingly, the insights are further enriched by the integration of a radar chart and correlation matrix, which offers a holistic perspective on the intricate relationships within box office features.

**Deployment Instructions**

1. **Clone the Repository**

Begin by cloning the Cinemalytics repository to your local machine using the following command: *git clone https://github.com/alennejasper/Cinemalytics.git*

1. **Navigate to the Cinemalytics Directory**

Move into the Cinemalytics directory: *cd Cinemalytics*

1. **Install Python**

Ensure Python 3.10.2 is installed on your system. If not, you can download and install it from *python.org*.

1. **Install Pipenv**

Ensure Pipenv is installed on your system. If not, you can install it using: *pip install pipenv*

1. **Install Django**

Install Django using the following command: *pip install django*

1. **Install Dependencies**

Install dependencies from the *requirements.txt* file using Pipenv by running this in the Cinemalytics directory: *pip install -r requirements.txt*

1. **Activate the Virtual Environment**

Activate the virtual environment using Pipenv: *pipenv shell*

1. **Apply Migrations**

Apply any necessary database migrations: *python manage.py migrate*

1. **Run the Development Server**

Start the development server using the following command: *python manage.py runserver*

1. **Access the Application**

Once the server is running, open your web browser and then visit the local host address below: *http://127.0.0.1:8000/*

1. **Explore and Contribute**

The application is now up and running! Feel free to explore the box office features, and do not hesitate to delve into the codebase. When you are done, use *Ctrl+C* in the terminal to stop the development server, and you can deactivate the virtual environment with the *exit* command.