Project Description: Exploratory Data Analysis for Microsoft's New Movie Studio

The aim of this project is to help Microsoft's new movie studio determine the types of movies that are currently performing well at the box office. This will be achieved by conducting an exploratory data analysis (EDA) on two provided datasets: "tmdb.movies.csv" and "tn.movie\_budgets.csv". The project will involve three deliverables, which are a non-technical presentation, a Jupyter Notebook, and a GitHub repository.

The first step in the project will be to understand the business problem and the data provided. Microsoft is interested in creating original video content, and the goal of the project is to identify what types of films are currently performing well at the box office. The two datasets provided contain information on movie budgets, revenue, genres, and production studios, among others.

Next, the project will involve conducting an EDA on the datasets to identify trends and patterns in the data. This will involve cleaning and preprocessing the data, performing exploratory visualizations, and conducting statistical analyses to identify relationships and correlations in the data.

The analysis will focus on three main areas: budget, movie runtime, and production studio recommendation. The budget will be analyzed to identify how it affects a movie's revenue and profitability. Movie runtime will be analyzed to identify trends in the optimal length of movies. Production studio recommendation will be based on an analysis of the most successful production studios based on revenue and profitability.

The Jupyter Notebook will contain the code used to clean and preprocess the data, perform visualizations, and conduct statistical analyses. The Notebook will also contain explanations of the code, interpretations of the results, and recommendations for the new movie studio based on the insights generated from the analysis.

The non-technical presentation will provide an overview of the project, the data used, and the insights generated from the analysis. The presentation will be tailored towards business stakeholders and will provide actionable recommendations based on the analysis.

Finally, the GitHub repository will contain the Jupyter Notebook and all associated files used in the project. The repository will also contain a README file that provides an overview of the project, instructions for running the code, and acknowledgments.

Overall, this project aims to provide actionable insights for Microsoft's new movie studio based on an EDA of two provided datasets. The insights generated from the analysis will help the studio make data-driven decisions on the types of movies to create based on factors such as budget, movie runtime, and production studio recommendation.