

A blue parallelogram and a light green parallelogram are positioned on the left side of the slide, overlapping each other and the dark background. The blue shape is on the left, and the green shape is to its right, partially overlapping it.

MICROSOFT MOVIE ANALYSIS

RECOMMENDER SYSTEM



Business understanding

Here the role expected of a data scientist is to explore through the datasets and be able to come up with the solution to question problem

- In the dataset we should execute data that is relevant to come up with the best movies

Based on the foreign gross,domestic gross and ratings of a specific movie produced per year



DATA UNDERSTANDING

- The two datasets in csv format have similarities in some movies produced eg the title names of the movies together with the times they were released
- Analysis is to be done on selected columns such as the year released, the vote averages referring to ratings and the gross values both domestic and foreign



DATA ANALYSIS

- Data analysis was using the EDA EXPLORATORY ANALYSIS

- It contains two methods :Univariate analysis-used to analyse data considering one variable

- Bivariate analysis-analysis based on the relationship between two variables

- From the results yielded „the metric for success was absolutely obtained hence a Null hypothesis

- The findings were visualized in graphs and scatter plots for easy understanding



RECOMMENDATIONS

-Based on the findings from the movie analysis :

>I highly recommend creation of movies based on their ratings per series produced

>Also consider the domestic and foreign gross earned by the movie

*In short, go for highly rated movies such as 'Toy Story 3' movies and those that relate closely with it



NEXT STEPS

_In the future i would love to encode more features analysis using the univariate and bivariate methods and also explore other ways to perform the analysis to its best.



THANKYOU

BY:GLORIA CHEPNGENO