



Solace Online Project

Problem Context

Streaming media services are media service offered directly to viewers via the internet. They are typically accessed via websites on personal computers, apps on smartphones and tablets, or televisions with integrated Smart TV platforms.

The global streaming media market size was valued at \$121.61 billion in 2019 and is projected to reach \$1,039.03 billion by 2027, growing at a CAGR of 29.4% from 2020 to 2027. The shift from television to streaming services for entertainment is driven by benefits such as on-demand services, ease of access, and access to better networks and digital connectivity.

Presently, Streaming services are widely accepted as a trending technology across the globe, coupled with the increasing change in customers' social behavior, which is shifting from traditional subscriptions to broadcasting services and Streaming on-demand video and music subscriptions every year, Booming streaming market is expected to grow at a very fast pace.



Business Goal (Objective)

Solace Online is an Online Streaming service provider and Platform that offers a wide variety of content (movies, web shows, etc.) for its subscribers. The Streaming Company wants to determine the driver variables for first-day content viewership so that they can take necessary measures to improve the viewership of the content on their platform. Some of the reasons for the decline in viewership of content would be the decline in the number of people coming to the platform, decreased marketing spend, content timing clashes, weekends and holidays, etc. They have hired you as a Data Scientist, shared the data of the current content in their platform, and asked you to analyze the data to determine the driving factors for first-day viewership.

Best Practices for EDA:

The final notebook should be well-documented, with inline comments explaining the functionality of code and markdown cells containing comments on the observations and insights.



Here you are required to apply Data visualization techniques to carry out an exploration of the data set. You are expected to carry out Univariate, Bivariate and Multivariate Analysis to explore, figure out relationships and identify insights from the Data set.

■ Kindly go through the Data dictionary to see the variables involved, further details and a case study template has been provided to guide your analysis

Data Description

The data contains the different factors to analyze for the content. The detailed data dictionary is given below.

Data Dictionary

- visitors: Average number of visitors, in millions, to the platform in the past week
- ad_impressions: Number of ad impressions, in millions, across all ad campaigns for the content (running and completed)
- major_sports_event: Any major sports event on the day
- genre: Genre of the content
- dayofweek: Day of the release of the content
- season: Season of the release of the content
- views_trailer: Number of views, in millions, of the content trailer
- views_content: Number of first-day views, in millions, of the content



**Create a detailed
analysis and upload
on your Github.**

(Tag @10Alytics)