Project Process – OTT Platform & TV Show Analysis

1. Project Objective
This project focuses on analyzing Movies and TV Shows data using Excel to identify content trends based on type, release year, genre, runtime, and IMDb ratings. The goal is to clean, process, and visualize data to support data-driven decision-making through an interactive dashboard.

2. Dataset Overview
The dataset was sourced from Kaggle and includes information such as content type, release year, genres, runtime, IMDb scores, and votes. The raw dataset contained inconsistencies and missing values that required preprocessing.

3. Data Cleaning & Preparation
Duplicate records were removed to ensure data accuracy. Missing numerical values were handled using average imputation, while text fields such as genres and countries were standardized. This step ensured clean and reliable data for analysis.

4. Data Analysis & Visualization
Multiple Excel visualizations were used to extract insights:
• Pie charts to compare Movies vs TV Shows and age certifications.
• Bar charts to analyze IMDb scores by content type and release year trends.
• Histograms to understand runtime distribution and skewness.
• Line charts to study content growth and IMDb voting trends.
• Scatter plots to analyze correlation between runtime and IMDb ratings.
• Boxplots to identify outliers in IMDb scores.

5. Outlier Treatment & Statistical Analysis
Outliers were identified using boxplots and treated to improve distribution symmetry. Measures of central tendency and spread were analyzed, revealing that TV shows generally perform better in ratings, while movies show higher variability.

6. Dashboard & Conclusion
An interactive Excel dashboard was created to summarize insights visually. The project concludes that runtime has minimal impact on ratings and that TV shows tend to receive higher IMDb scores compared to movies.