



EDA Assignment

Netflix Customer Data - EDA Assignment

Objective:

Explore and analyze the Netflix customer dataset to derive insights into user behavior, subscription patterns, and churn tendencies.

[Dataset](#)

1. Basic Information

- Total records: 5000
 - Columns: 14
 - Types of Data:
 - **Numerical:** `age`, `watch_hours`, `last_login_days`, `monthly_fee`, `number_of_profiles`, `avg_watch_time_per_day`, `churned`
 - **Categorical:** `gender`, `subscription_type`, `region`, `device`, `payment_method`, `favorite_genre`
-

2. Assignment Tasks

Task 1: Data Overview

- Display the first 5 rows of the dataset.
- Check for missing/null values in the dataset.
- Check for unique values in columns like `gender`, `subscription_type`, `region`, etc.

✓ Task 2: Univariate Analysis

- Plot the distribution of:
 - age
 - watch_hours
 - monthly_fee
 - churned (0 = No churn, 1 = Churn)
- Count plots for categorical variables:
 - subscription_type
 - gender
 - region
 - device
 - payment_method
 - favorite_genre

✓ Task 3: Bivariate Analysis

- Compare average watch_hours and monthly_fee across:
 - subscription_type

- `region`
 - `device`
- Plot average `avg_watch_time_per_day` by `favorite_genre`.
- Analyze churn rate based on:
 - `gender`
 - `region`
 - `subscription_type`
 - `payment_method`

✓ Task 4: Correlation Analysis

- Create a heatmap of correlation between numerical variables.
- Look for potential predictors of churn, e.g., `watch_hours`, `last_login_days`, `monthly_fee`.

✓ Task 5: Insights & Recommendations

Write 5–7 insights based on your analysis. For example:

- Are customers with lower watch hours more likely to churn?
- Do Premium users churn more or less than Basic users?
- Which region has the highest average `watch_hours`?