**Netflix Dataset Insights & Metrics**

**1. Dataset Overview**

* **Initial shape:** df\_read.shape (you can include exact numbers from your run).
* **Columns with missing values:** director, cast, country, date\_added, rating, duration.

**2. Data Cleaning Steps (Using Pandas)**

* **Handled Missing Values:**
  + director → "Unknown"
  + cast → "Not Available"
  + country → "Unknown"
  + date\_added → "Not Available"
  + rating → "Not Rated"
  + duration → "Unknown"
* **Dropped rows** missing critical fields (title, type).
* **Removed duplicates** to ensure unique entries.
* **Standardized text fields** for consistency:
  + duration → stripped whitespace and capitalized text
  + cast → stripped whitespace

**3. Dataset after Cleaning**

* **Shape after null handling:** df\_read.shape
* **Shape after dropping duplicates:** df\_read.shape
* **Sample data:** Display first 5 rows for reference.

**4. Metrics & Insights (from your analysis, can be added once you share)**

* Top directors by number of titles
* Distribution of movies vs. TV shows
* Countries with most content
* Most common ratings