NETFLIX DATA ANALYSIS

Kashifa Shah

Internship ID: STU6828a35915ed01747493721

PROBLEM STATEMENT

“Netflix needs to understand how its content library—Movies vs. TV Shows, genres, and country contributions—has evolved over the years to make data-driven strategic decisions for content acquisition and production.”

Key Points:

- Netflix’s library is vast and diverse, spanning multiple countries and genres.
- Competitors like Amazon Prime, Disney+, and regional OTT platforms are growing.
- Netflix must identify content trends, audience preferences, and underrepresented categories.
- The goal is to provide actionable insights to optimize content strategy, improve audience engagement, and strengthen global

Project Description

The project will involve analyzing a Netflix dataset (7,789 records, 11 columns) to uncover content trends and patterns. The analysis will cover:

1. **Distribution Analysis:** How Movies vs. TV Shows have evolved over the years.
2. **Genre Analysis:** Identify most popular genres and track their popularity over time.
3. **Country-wise Contributions:** Examine content contributions from different countries.
4. **Strategic Insights:** Provide recommendations based on trends for content production, acquisition, and global expansion.

The project combines **data cleaning, exploratory data analysis (EDA), visualization, and insights generation** to produce actionable recommendations.

WHO ARE THE END USERS?

1. **Netflix Decision Makers:** Content strategists, production planners,and market analysts.

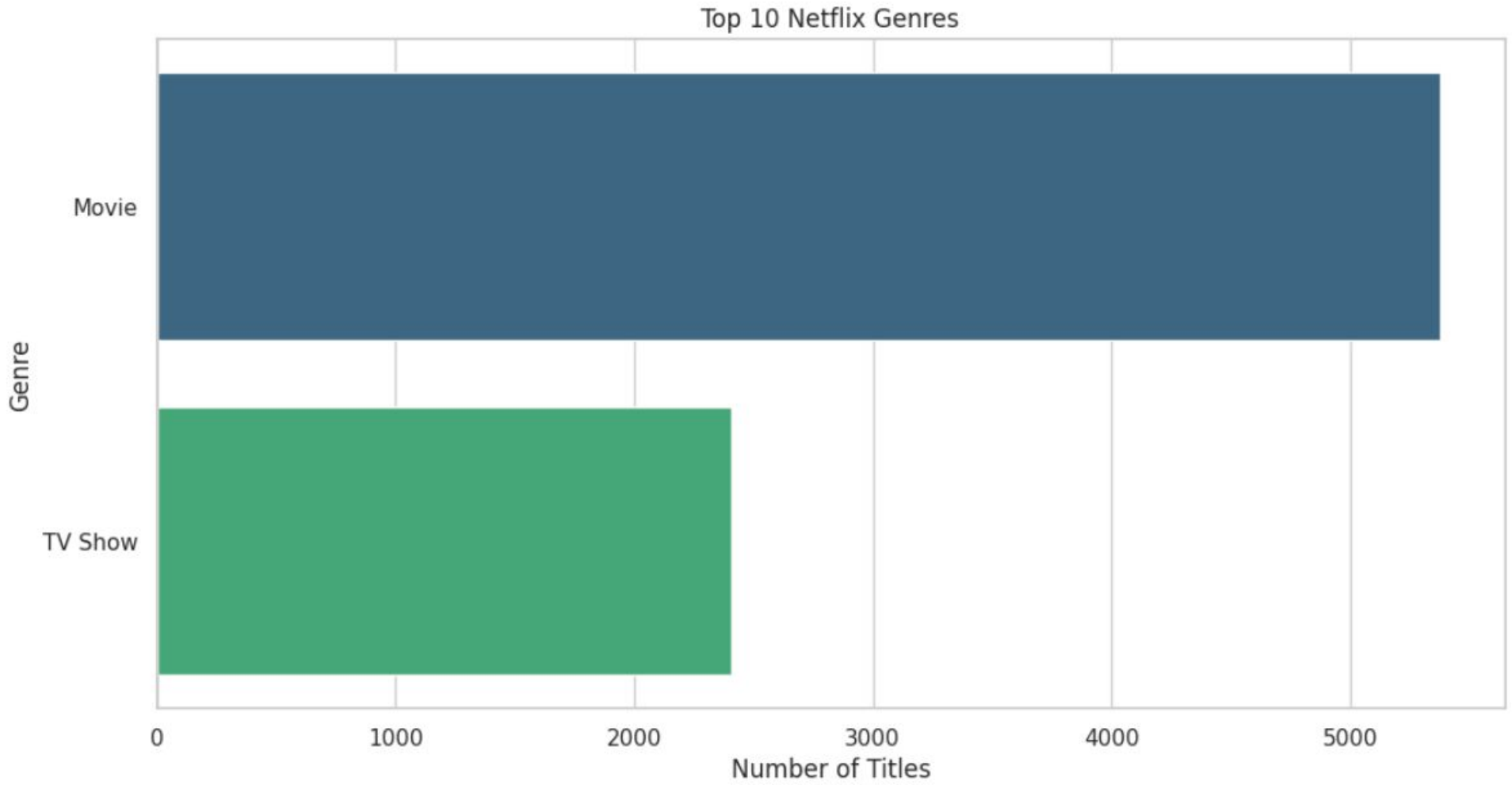
2. **OTT Industry Analysts:** Researchers who track streaming content trends and market dynamics.

3. **Academics/Students:** Those studying media analytics, data-driven decision-making, or business intelligence.

Technology Used

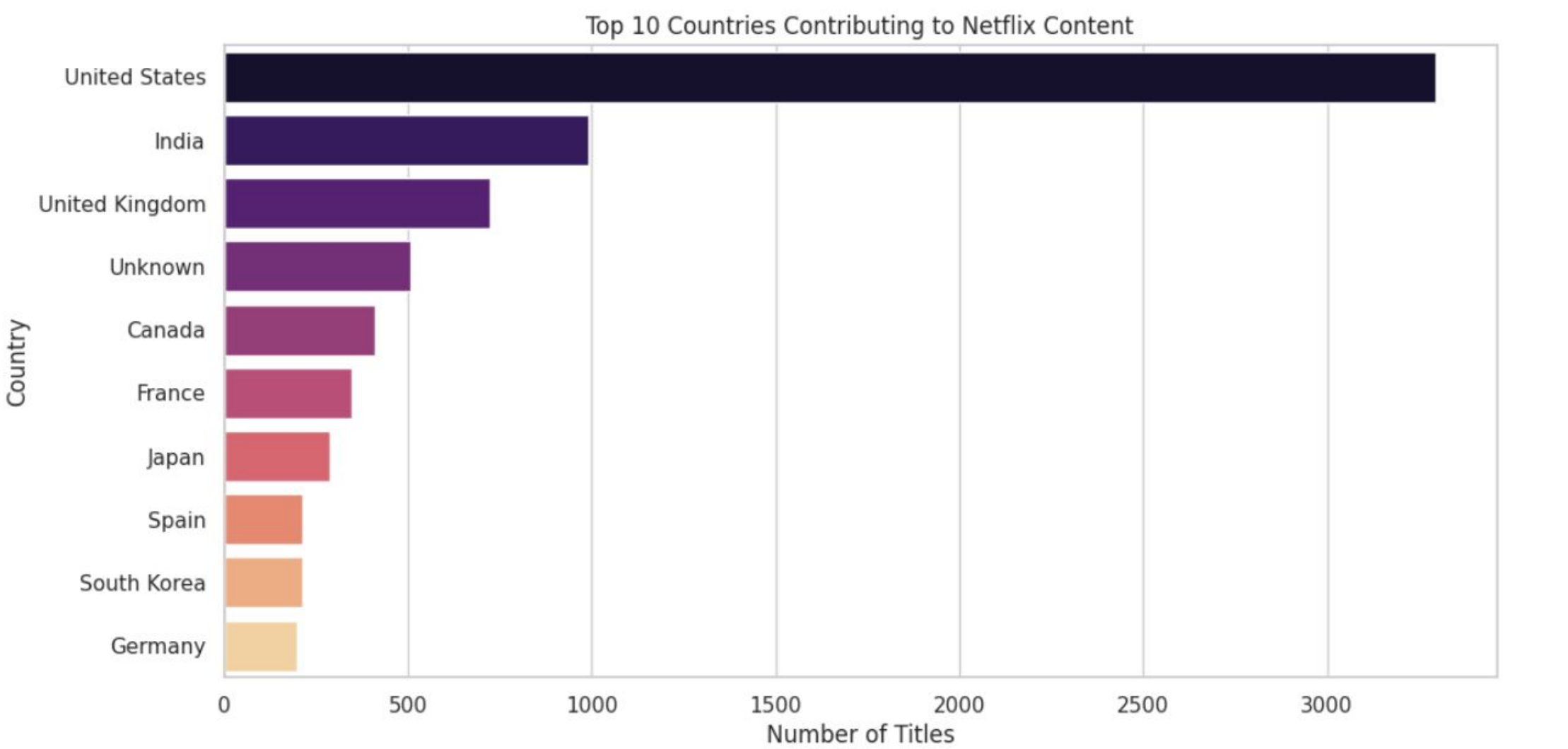
- **Programming Language:** Python
- **Libraries:**
  - `pandas` for data manipulation
  - `numpy` for numerical operations
  - `matplotlib` & `seaborn` for data visualization
  - `plotly` or `dash` for interactive dashboards (optional)
- **IDE/Environment:** Jupyter Notebook or VS Code
- **Data Storage:** CSV or Excel file
- **Optional:** Tableau/Power BI for enhanced dashboards

RESULTS1



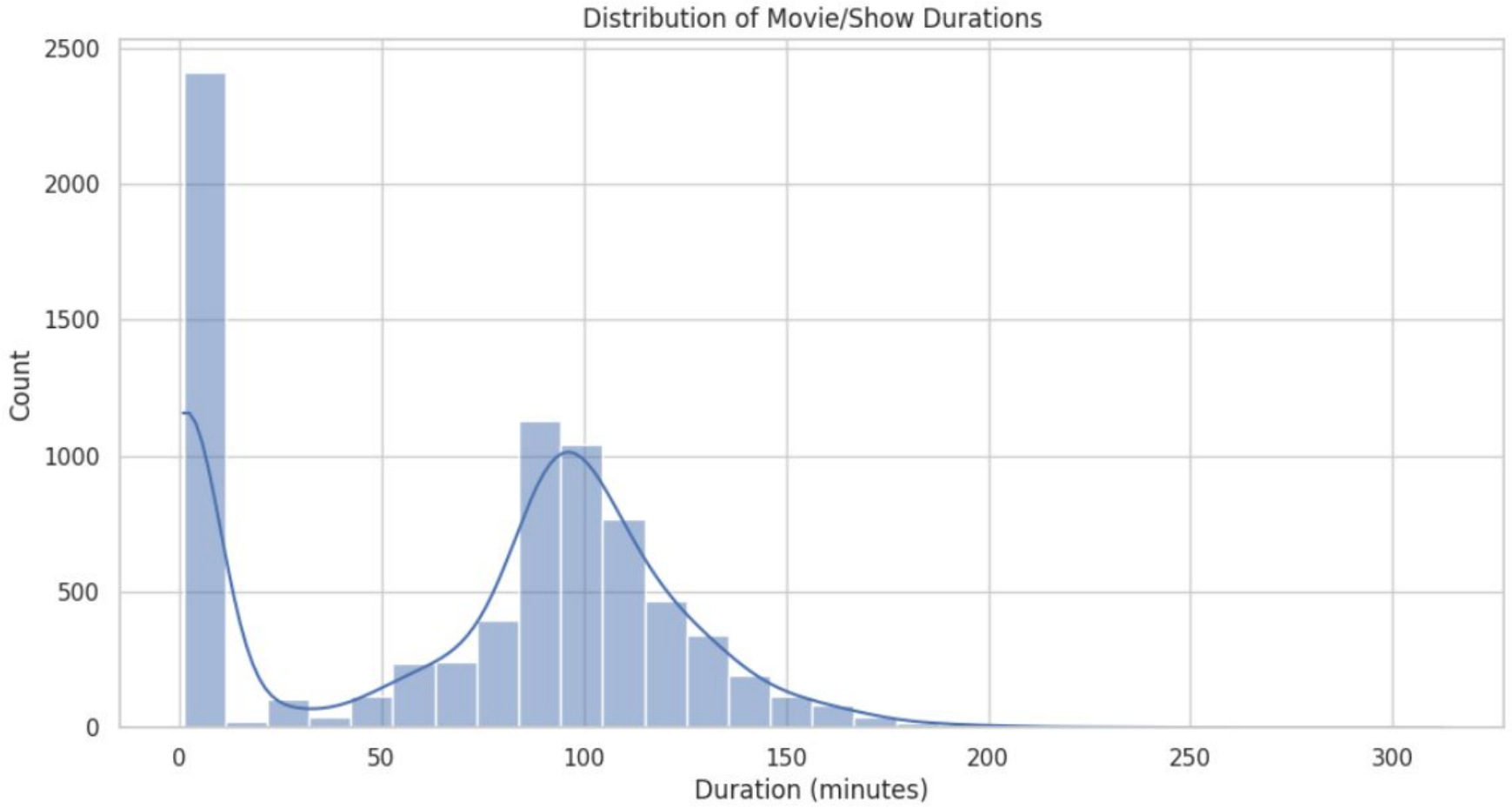
[Demo Link]

RESULTS2



[Demo Link]

RESULTS3



[Demo Link]

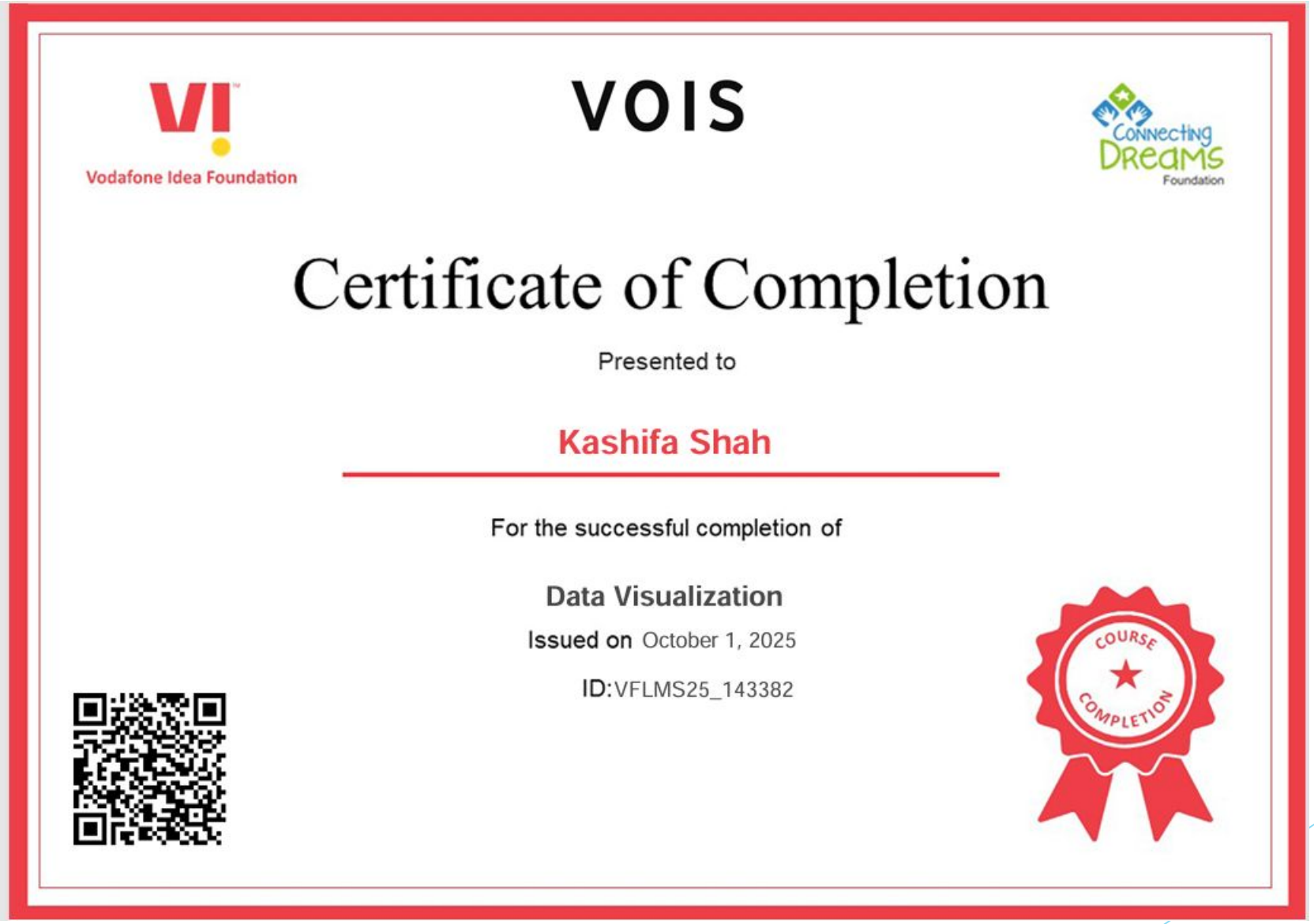
GitHub repository

https://github.com/kashifashah1/VOIS_AICTE_Oct2025_MajorProject_KashifaShah.git

Demo Link

Getting started with Basics of Python Certificate

Data Visualization Certificate



Thank you