📊 Netflix Movies & TV Shows Data Analysis Project

# 1. Overview

This project analyzes the Netflix Movies & TV Shows dataset to uncover insights such as:  
- Number of movies vs TV shows  
- Most common genres  
- Content release trends over time  
- Top directors  
- Most common content ratings  
  
Dataset Source: Netflix Movies & TV Shows - Kaggle (https://www.kaggle.com/datasets/shivamb/netflix-shows)

# 2. Project Folder Structure

Organize your project in the following way:  
  
netflix-data-analysis/  
├── data/  
│ └── netflix\_titles.csv # dataset (or link in README if too large)  
├── notebooks/  
│ └── Netflix\_Analysis.ipynb # your Jupyter notebook/code  
├── results/  
│ └── plots.png # (optional: save some graphs as images)  
├── requirements.txt # dependencies  
└── README.md # documentation

# 3. Requirements.txt

Run this in your project folder to create requirements.txt:  
  
pip freeze > requirements.txt

# 4. README.md Template

# 🎬 Netflix Movies & TV Shows Analysis  
  
## 📌 Overview  
This project analyzes the Netflix Movies & TV Shows dataset.  
  
## 🚀 Features  
- Data cleaning & preprocessing  
- Yearly trend of Netflix content added  
- Top 10 most common genres  
- Top 10 directors  
- Ratings distribution  
  
## 🛠️ Tools & Libraries  
- Python, pandas, numpy  
- matplotlib, seaborn  
- collections (Counter)  
- wordcloud (optional)  
  
## 📊 Example Insights  
- Netflix has more movies than TV shows  
- Dramas and Comedies dominate the platform  
- 2018–2020 saw a huge spike in content  
- TV-MA and TV-14 are most frequent ratings  
  
## ⚙️ Installation  
1. Clone the repo  
2. Install dependencies  
3. Run Jupyter Notebook  
  
## 📌 Future Improvements  
- Interactive dashboard with Plotly/Power BI  
- Deploy as Streamlit app  
  
## 👩‍💻 Author  
- Rashmitha Thatikonda

# 5. GitHub Upload Steps

1. Initialize git repo:  
 git init  
  
2. Add files:  
 git add .  
  
3. Commit:  
 git commit -m "Netflix Data Analysis Project"  
  
4. Add remote:  
 git remote add origin https://github.com/yourusername/netflix-data-analysis.git  
  
5. Push:  
 git push -u origin main

# 6. LinkedIn Post Draft

🎬 Just completed my Netflix Movies & TV Shows Data Analysis Project 📊  
  
🔍 Explored 8,800+ Netflix titles to analyze:  
✅ Movies vs TV Shows distribution  
✅ Top genres and directors  
✅ Yearly trends in content release  
✅ Content ratings distribution  
  
🛠️ Tools: Python, pandas, matplotlib, seaborn  
  
📈 Key Insights:  
- Netflix has significantly more movies than TV shows  
- Dramas and Comedies dominate the platform  
- 2018–2020 was the peak for content additions  
- TV-MA is the most frequent maturity rating  
  
🔗 GitHub Repository: https://github.com/yourusername/netflix-data-analysis  
  
Excited to share more Data Analytics projects in the coming weeks 🚀