



# IMDB Movies Gross Prediction

---

# Table of contents

**01**

**Introduction**

**02**

**Methodology**

**03**

**Data preprocessing  
and Analysis**

**04**

**Conclusion**

---



01


# Introduction

# Introduction

Objective:

- IMDB Movies gross prediction
- Using multiple regression models.

Finding the best regression model.



---



02

## Methodology

# Methodology

## Exploratory Data Analysis

- Cleaning (nulls, duplicates, spacing)
- Data Analysis

## Regression

Linear  
K-fold  
Polynomial  
Ridge

01 ——— 02 ——— 03 ——— 04

## Data Extraction

Web scraping using  
Beautiful Soup.  
1000 records and 8  
features

## Data Preprocessing

---

# Tools

Jupyter Notebook

Pandas

Matplotlib

BeautifulSoup


NumPy

Seaborn

---

# Data Preprocessing

Steps:

- Feature Selection.
  - Data Splitting.
  - Regression models ( K-fold , Ridge, Linear , Polynomial).
- 



# Data Splitting

20% — 20% — 60%

Validation

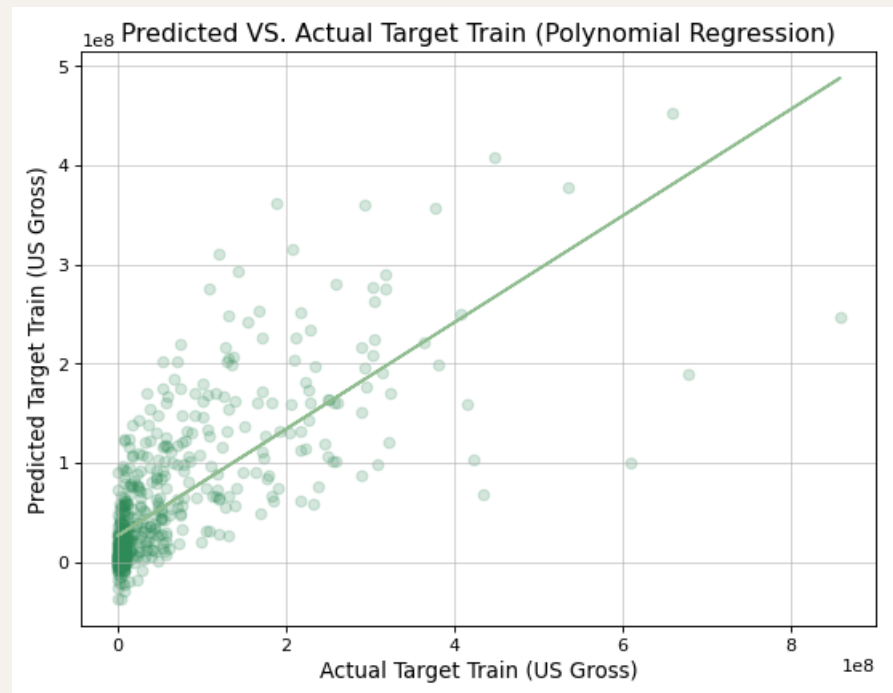
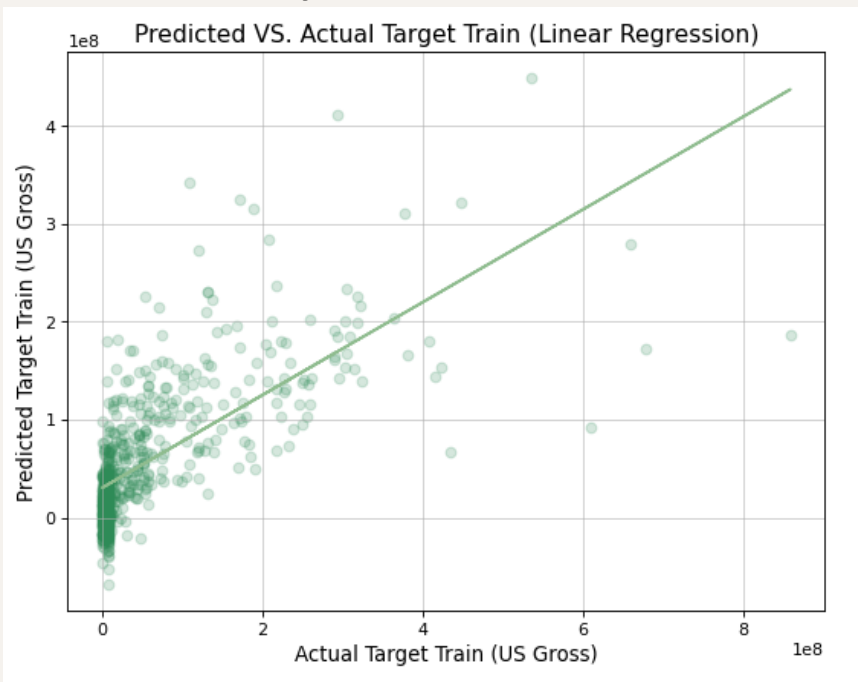
Test

Train

# Regression Models

Regression Models	Training Score	Validation Score
Normal LR	0.474546437	0.338569574
K-fold LR	0.474546438	0.459359399
Polynomial Regression	0.53764559	0.44003267
Ridge Regression	0.47454317	0.33885845
RR Cross-Validation	0.474492806	0.459387872

# Analysis



Testing Score:0.58248222

---

# Conclusion

## Polynomial Linear Regression

- After testing multiple models.
- Polynomial Regression is the best model for IMDB movies gross prediction

# Thanks

**CREDITS:** This presentation template was created by **Slidesgo**, including icons by **Flaticon**, and infographics & images by **Freepik**

Please keep this slide for attribution