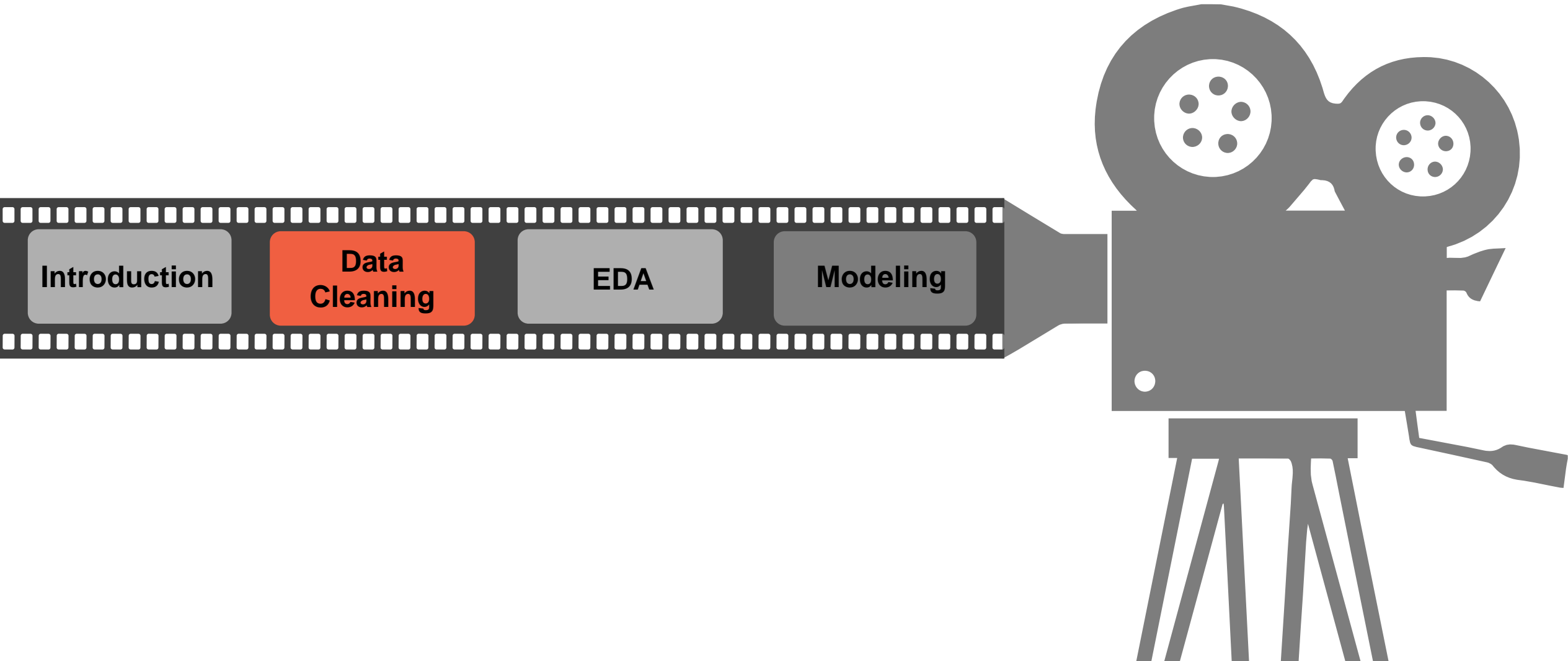# CINEMA TICKETS PREDICTION

Prepared by :
Mohammed Saeed Al shehri
Abdulelah Al dossari
Munira Alzhrani

Overview

Introduction

Data Cleaning

EDA

Modeling

Introduction

> **Cinema industry is not excluded from getting advantage of predictive modeling. Like other industry e.g., retail, banking, and restaurants, sale forecasts can help cinemas for cost reduction and better ROI. By forecasting sales, screening in different locations could be optimized as well as effective market targeting and pricing.**

Dataset

- 142524 Observations

- 14 Features

From

kaggle

Data Cleaning

1

Drop all values contain Nan values

2

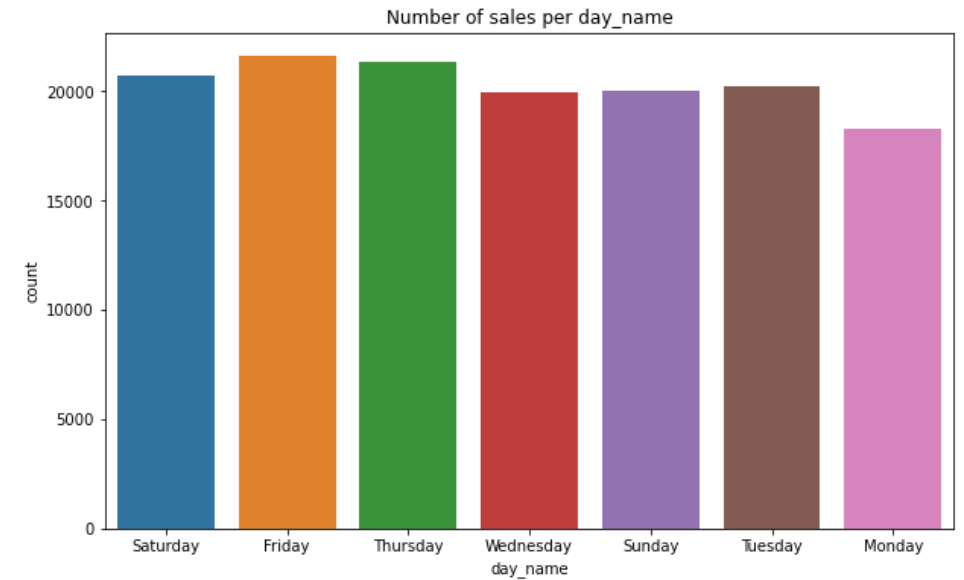Drop rows that contain irrational values

3
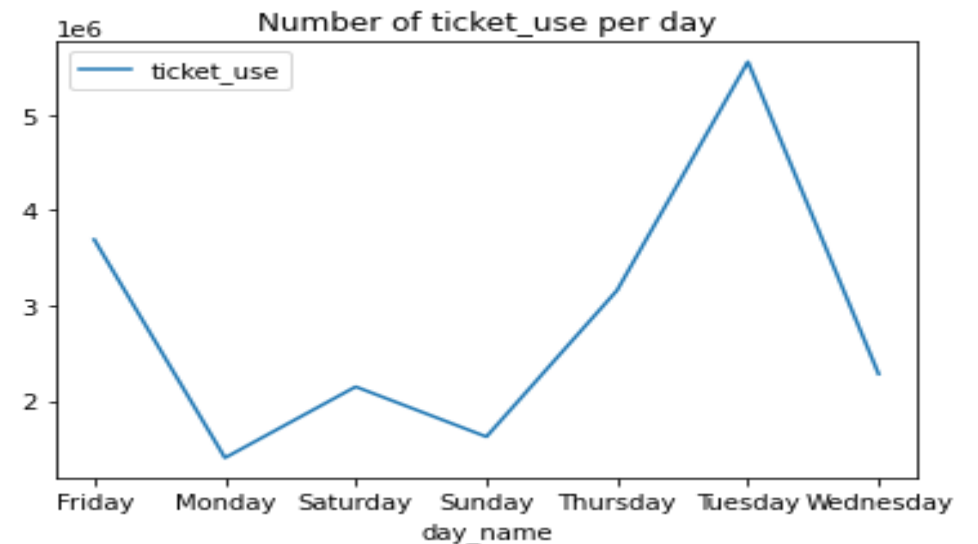
Change the column names to clearer names

4

Insert a column called`day_name`

Exploratory Data Analysis (EDA) ▶▶
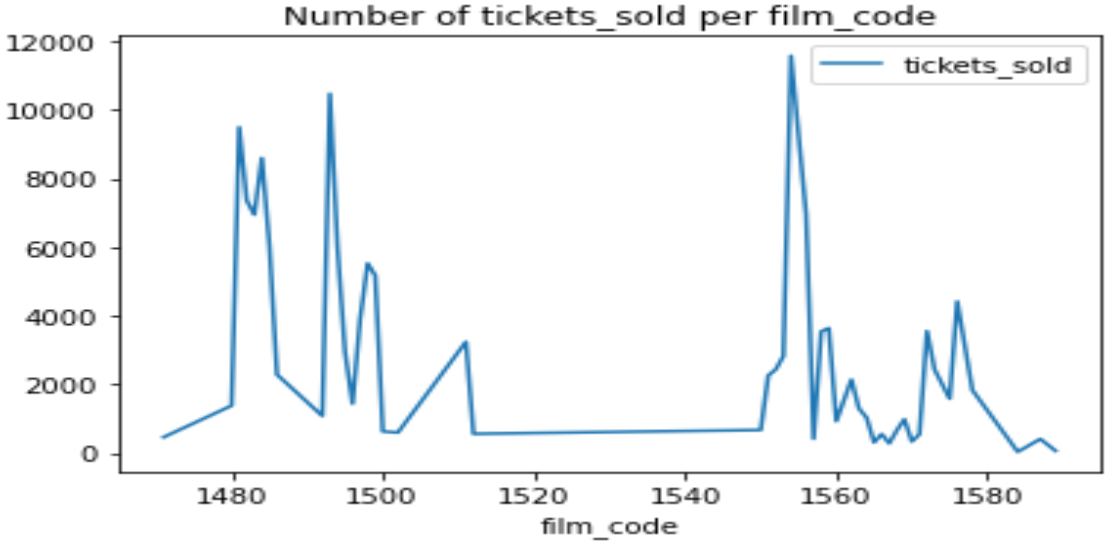
Q1. Which `day_name` has the most number of sales?
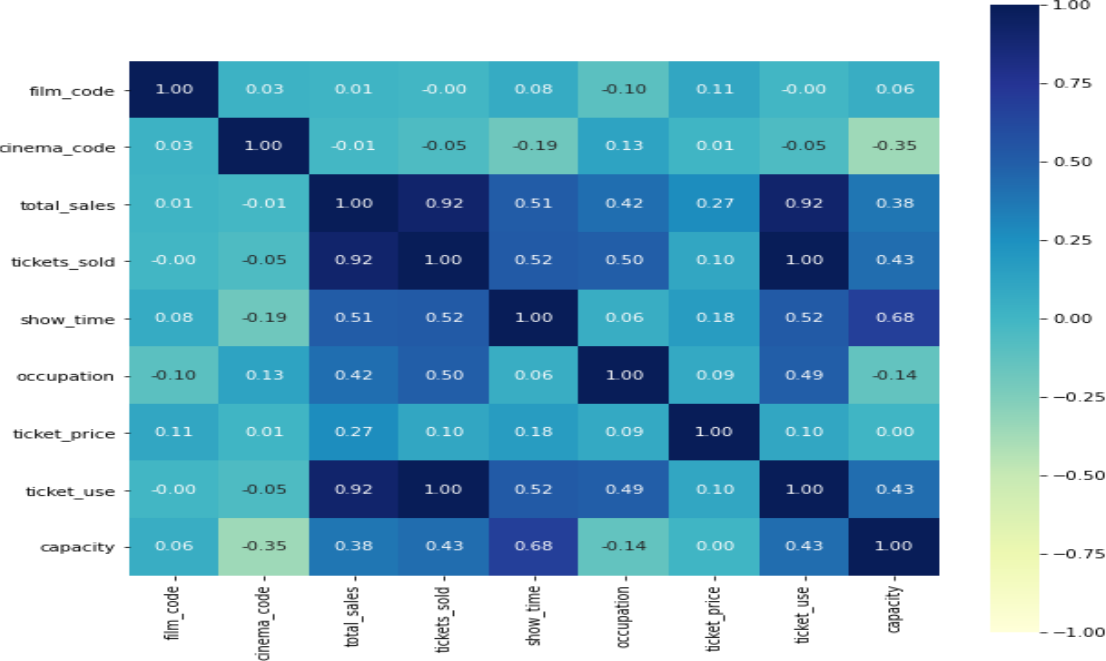

Q2. How many `ticket_use` based on `day_name`?

Q3. Which `film_code` has the most `tickets_sold`?

Q4. What is the correlation between the features?

Q5. Does `show_time` affect the `tickets_sold?

Number of tickets_sold per nember of show_time
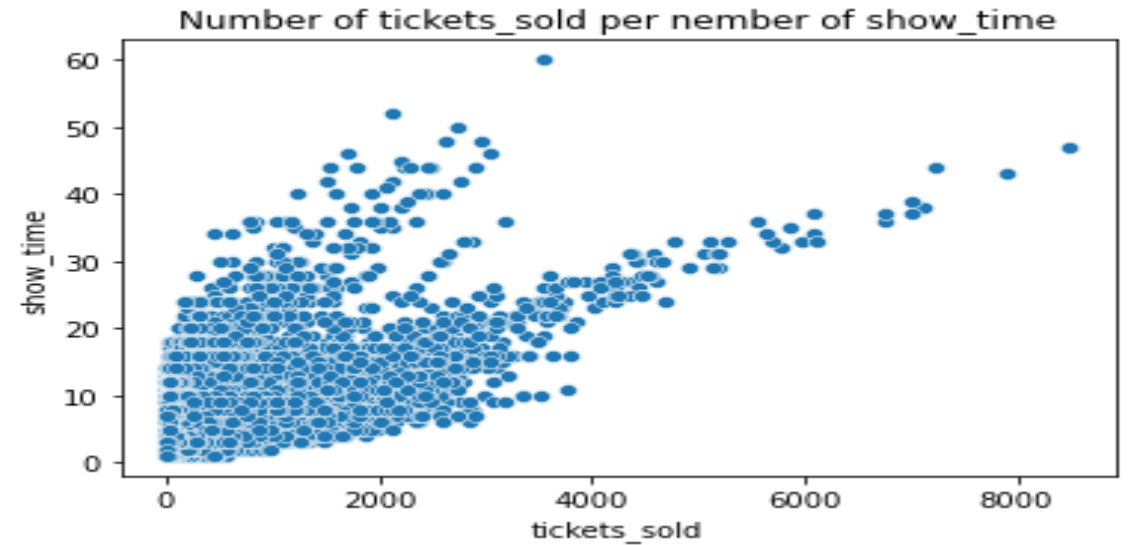
Q6. Which `film_code` is the most popular film?

Number of times film_code has been shown
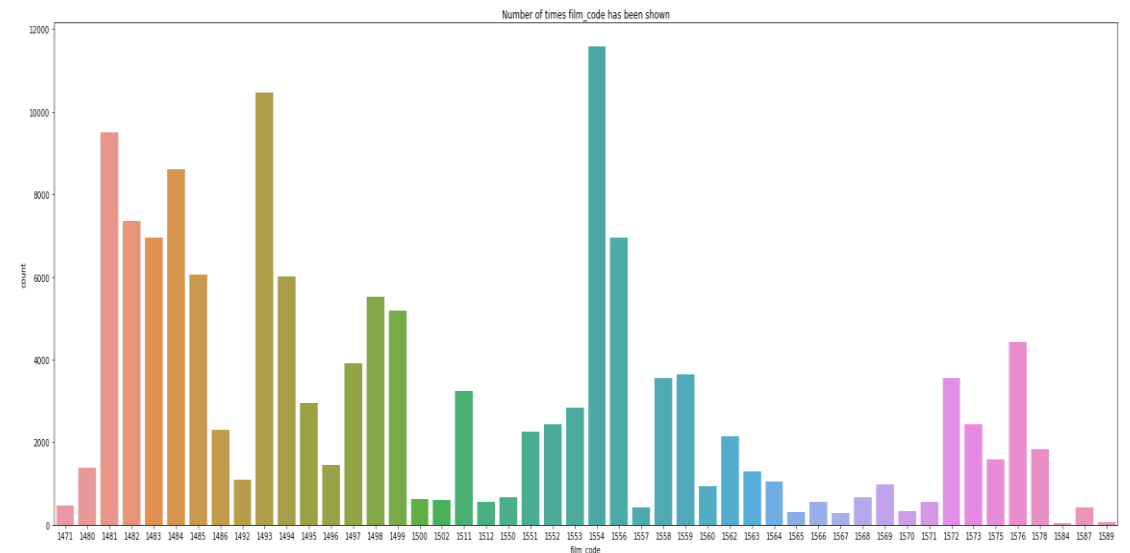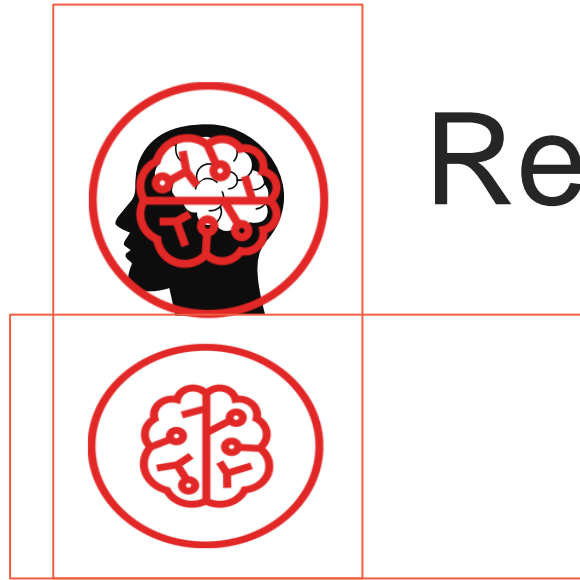
Regression

Regression Algorithms

Linear Regression

Ridge

polynomial

Random Forest

Decision Tree
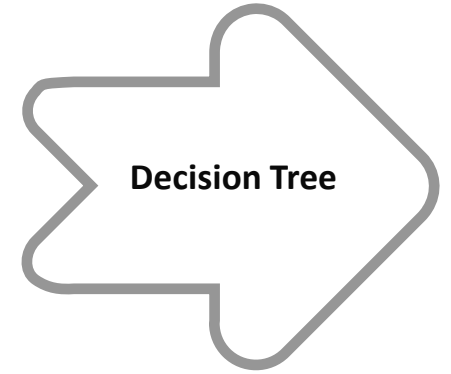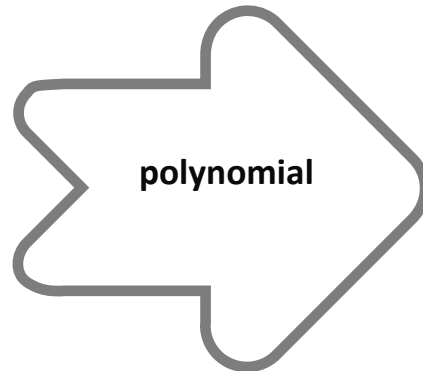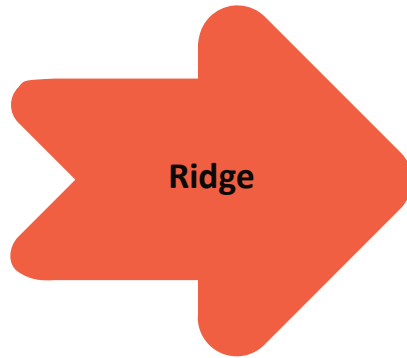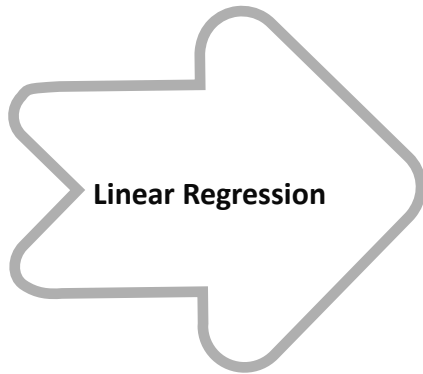
MODELS

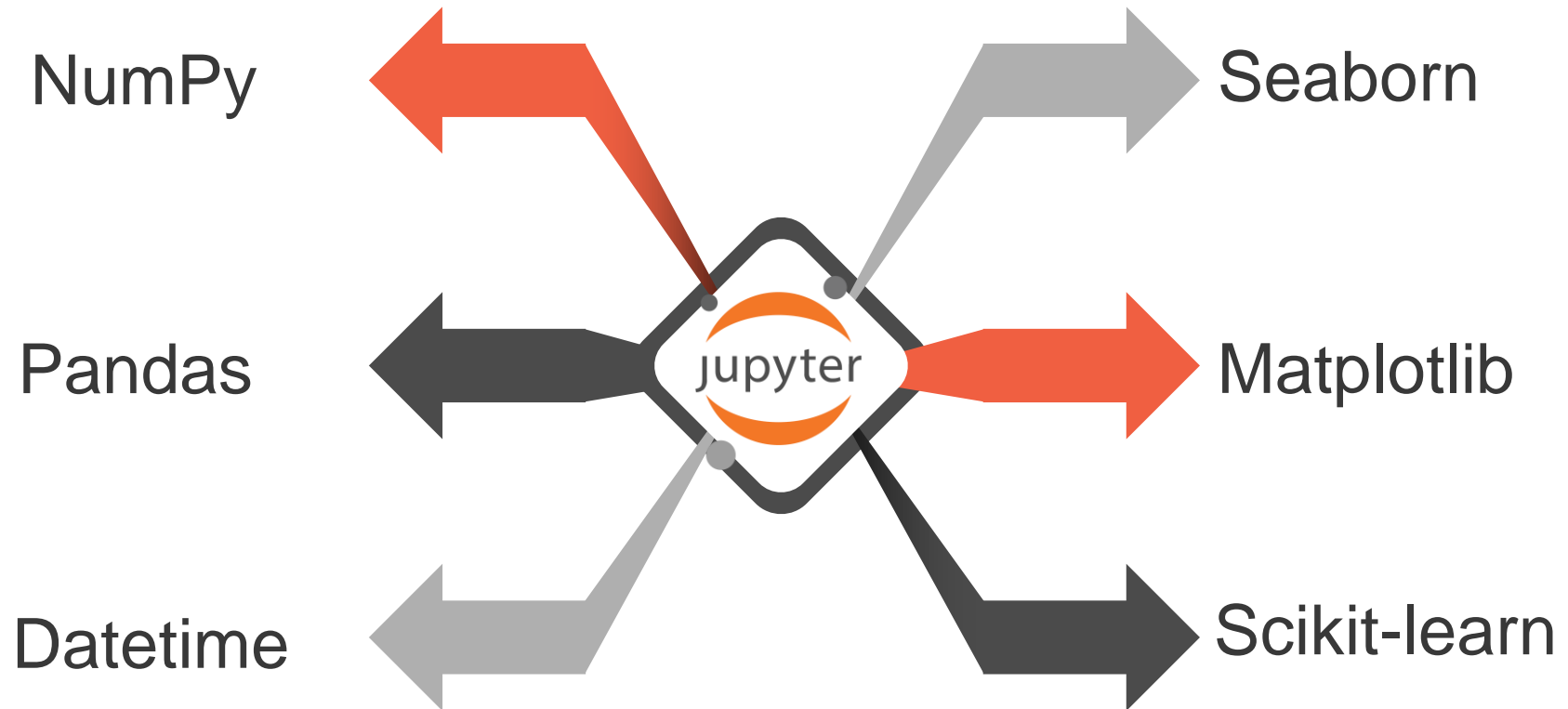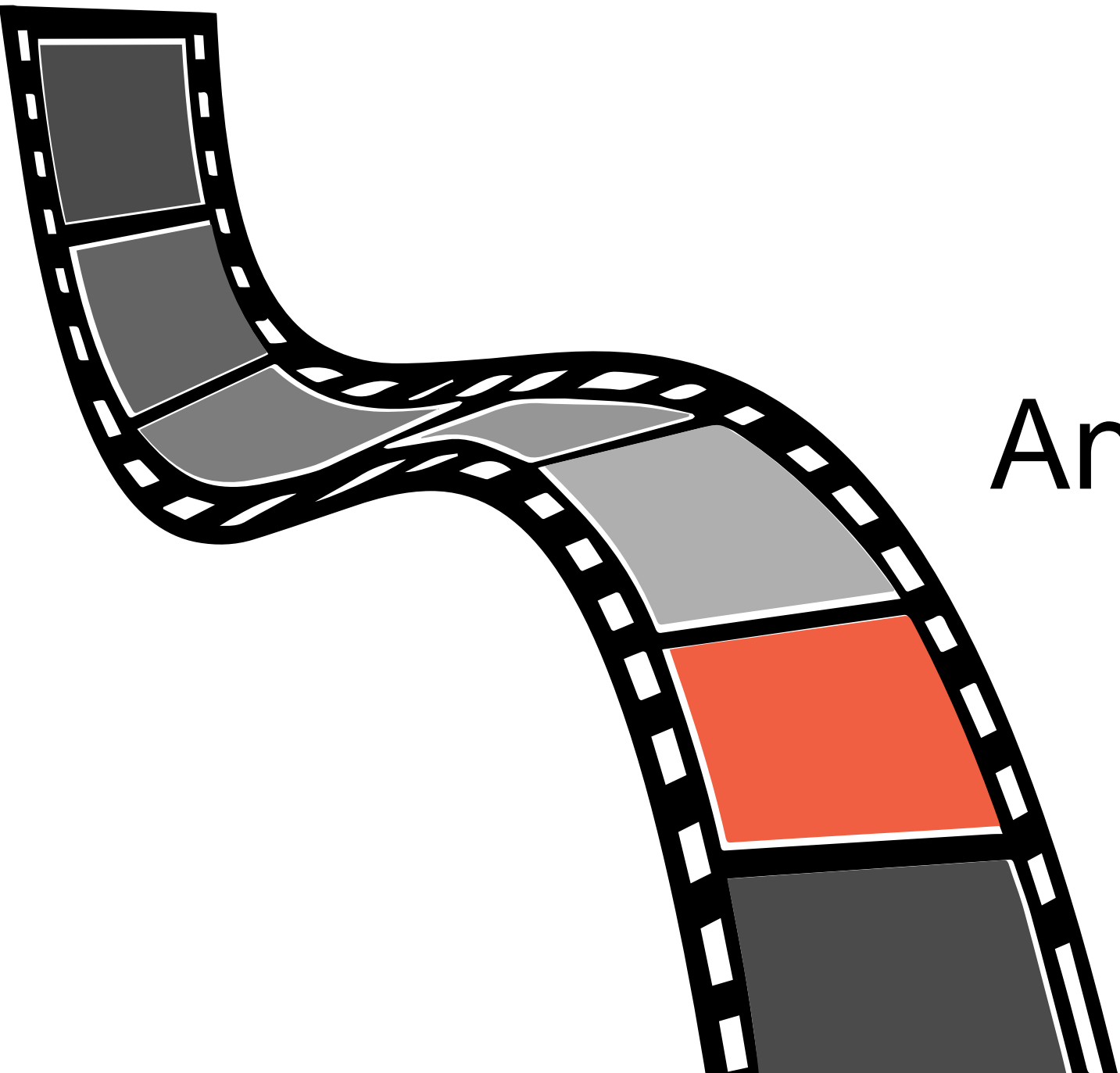| Algorithms	r^2 score
Linear Regression	The r^2 score value for train is 0.879 The r^2 score value for test is 0.880
Ridge	The r^2 score value for train is 0.879 The r^2 score value for test is 0.880
polynomial	The r^2 score value for train is 1.000 The r^2 score value for test is 1.000
Random Forest	The r^2 score value for train is 0.996 The r^2 score value for test is 0.997
Decision Tree	The r^2 score value for train is 0.993 The r^2 score value for test is 0.996

Conclusion

To sum up, to know the days when the gross income is low and trying to increase it by means of discounts for ticket and to develop it. Also to create offer on the ticket to invest more tickets in some days that people cannot go to it

Tools

Any Questions ?