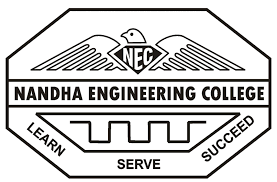
**NANDHA ENGINEERING COLLEGE**

(An Autonomous Institution, Affiliated to Anna University, Chennai)

# ERODE–638052



## A Project Report

***Submitted by***

*k.NISHANTHINI*

*In partial fulfillment for the award of the degree*

*of*

# BACHELOR OF TECHNOLOGY

# IN

# ARTIFICIAL INTELLIGENCE AND DATA SCIENCE

**DEPARTMENT OF ARTIFICIAL INTELLIGENCE AND**

**DATA SCIENCE**

What is Tableau?

Tableau is a powerful and easy-to-use data visualization tool that helps people understand data.  
It allows users to create beautiful charts, dashboards, and reports without needing to write complex code.  
With Tableau, we can turn raw data into clear, interactive visual stories to find trends, patterns, and insights.  
It is widely used in businesses, research, and education to make better decisions by seeing data visually.

Project: ALL MOVIES DATA

In this project, we are using Tableau to study and understand car sales data.  
The main aim is to turn the raw car sales information into beautiful and interactive dashboards that show:

* **When the movie came out**
* **Person who directed the movie**
* **People who wrote the script**
* Top-popular movies

Trending movies

Steps in the Project:

1. Connect Tableau to the movies data (from an Excel file or database).
2. Clean the data if needed (remove errors or missing values).
3. Create visualizations like bar charts, pie charts, line graphs, and maps.
4. Build dashboards that combine different charts together.
5. Analyze the dashboards to find patterns, trends, and useful insights.

Importance of This Project:

* **Analyzing movie ratings, votes, or revenue helps you understand **audience preferences** and **market dynamics**.**
* **Movies datasets are **perfect** for building **recommendation engines** like Netflix or YouTube does.**
* Shows sales trends across different months and years.
* Supports better business planning and marketing strategies.

## TITLE vs POPULARITY:

## C:\Users\pcsma\Pictures\Screenshots\Screenshot 2025-04-27 115359.pngScreenshot 2025-04-27 115359

## What This Chart Shows:

· ****Y-Axis (Vertical):** Movie titles (e.g., Avatar, Stranger Things, Mission Impossible, etc.)**

**· **X-Axis (Horizontal):** Popularity score (how popular each movie is, probably based on some metric like searches, votes, or views)**

## **Colour: ➡️ Each bar is colored differently for each movie, and the legend on the right side matches colors to movie name**

**Comparing Popularity:**

We can easily see **which movies are more popular** and which are less.

· **Top Popular Movies**

**"Mission: Impossible - ..."** has the highest popularity (about 45)

**"Avatar: The Way of Water"** and **"Avengers: Age of Ultron"** are also very popular (around 38–40).

## TITLE VS REVENUE:

## C:\Users\pcsma\Pictures\Screenshots\Screenshot 2025-04-27 115426.pngScreenshot 2025-04-27 115426

## What This Chart Shows:

## **No traditional X or Y axis — instead, the chart shows **portions of total revenue** for each movie as slices of the pie**

****Labels:****

**Each slice has a **revenue value** next to it (e.g., 2,320,250,281 dollars).**

**Percentages like **7.5%**, **9.7%**, **6.6%** are also shown (these seem to represent share of total revenue).**

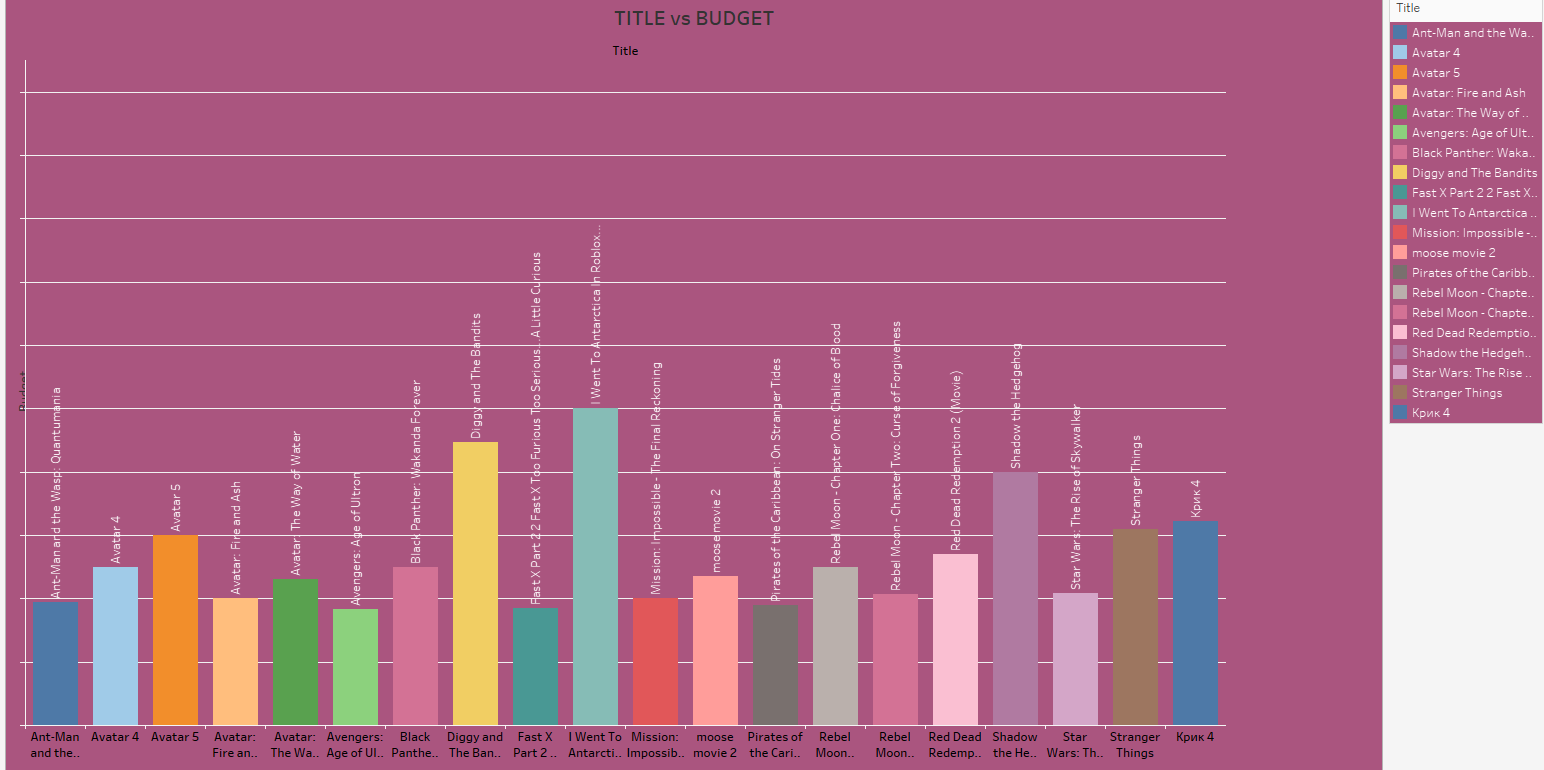
**Revenue Distribution:**

It shows **how much revenue each movie generated** compared to the total revenue of all movies

· **"Avatar: The Way of Water"** made the biggest amount (~2.3 billion dollars) — the largest slice of the pie.

· Other movies like **"Avatar 4"**, **"Avengers: Age of Ultron"**, and **"Black Panther: Wakanda Forever"** also made very high revenues (over 1 billion dollars).

**BUDGET REPRESENTATION:**



What This Chart Shows:

**This chart shows **how much budget** each movie had for production.**

**AXES:**

****X-axis:** Movie Titles (each bar represents one movie)**

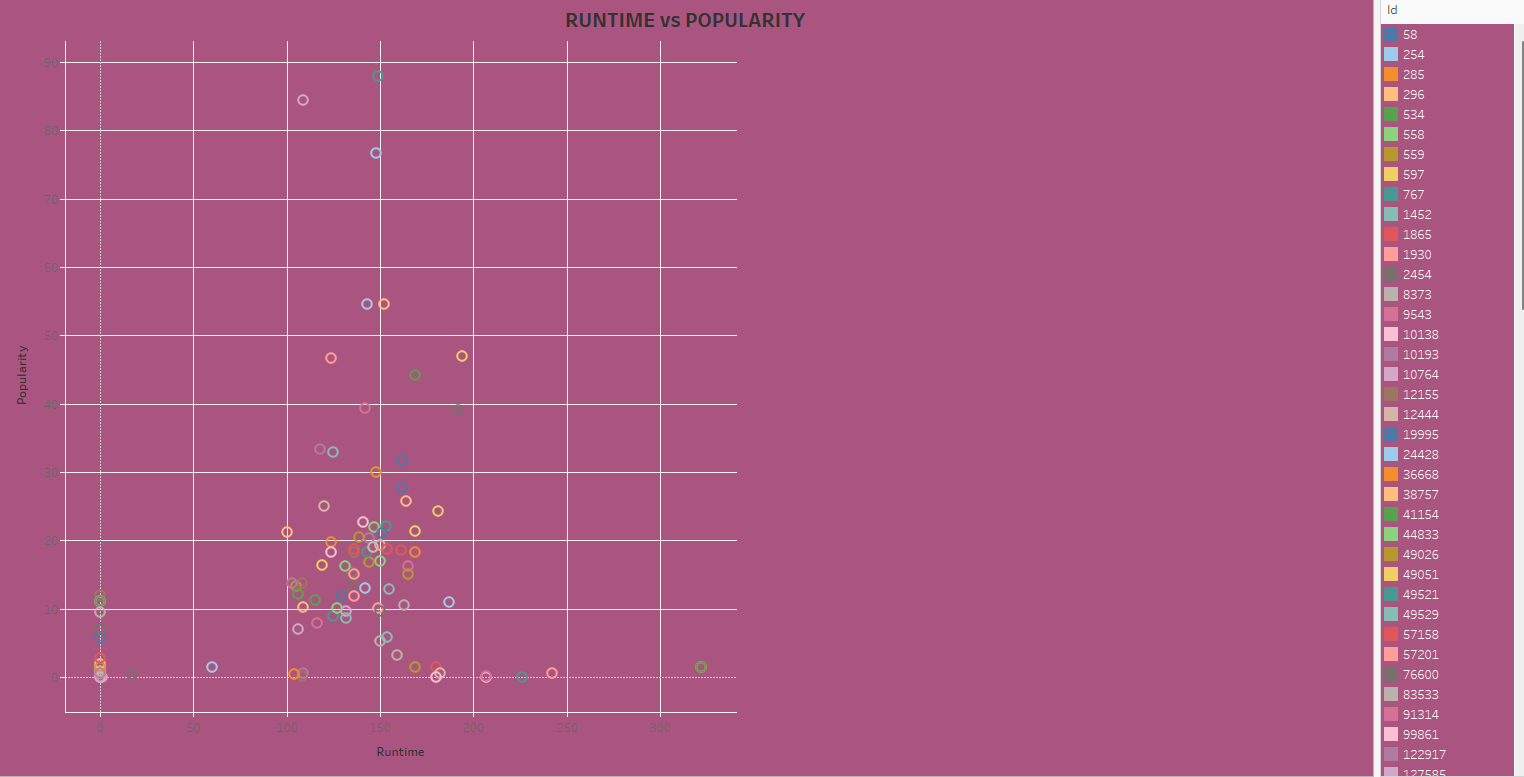
****Y-axis:** Budget (amount of money spent to make the movie)**

**Some movies like **"I Went To Antarctica"** and **"Diggy and The Bandits"** show very high budgets compared to others**

**Some titles that don't sound like major Hollywood productions (like "Diggy and The Bandits") surprisingly have **high budgets**, which might point to either a reporting error, an indie project with unusual funding, or a special project.**

**This chart tells us **how much money filmmakers spent** on making each movie, and **helps spot the most expensive productions** easily**

**RUNTIME VS POPULARITY:**



What This Chart Shows:

· ****X-axis (horizontal):** Runtime (movie length in minutes)**

**· **Y-axis (vertical):** Popularity (some numerical measure of how popular the movie is — like search volume, ratings buzz, etc.)**

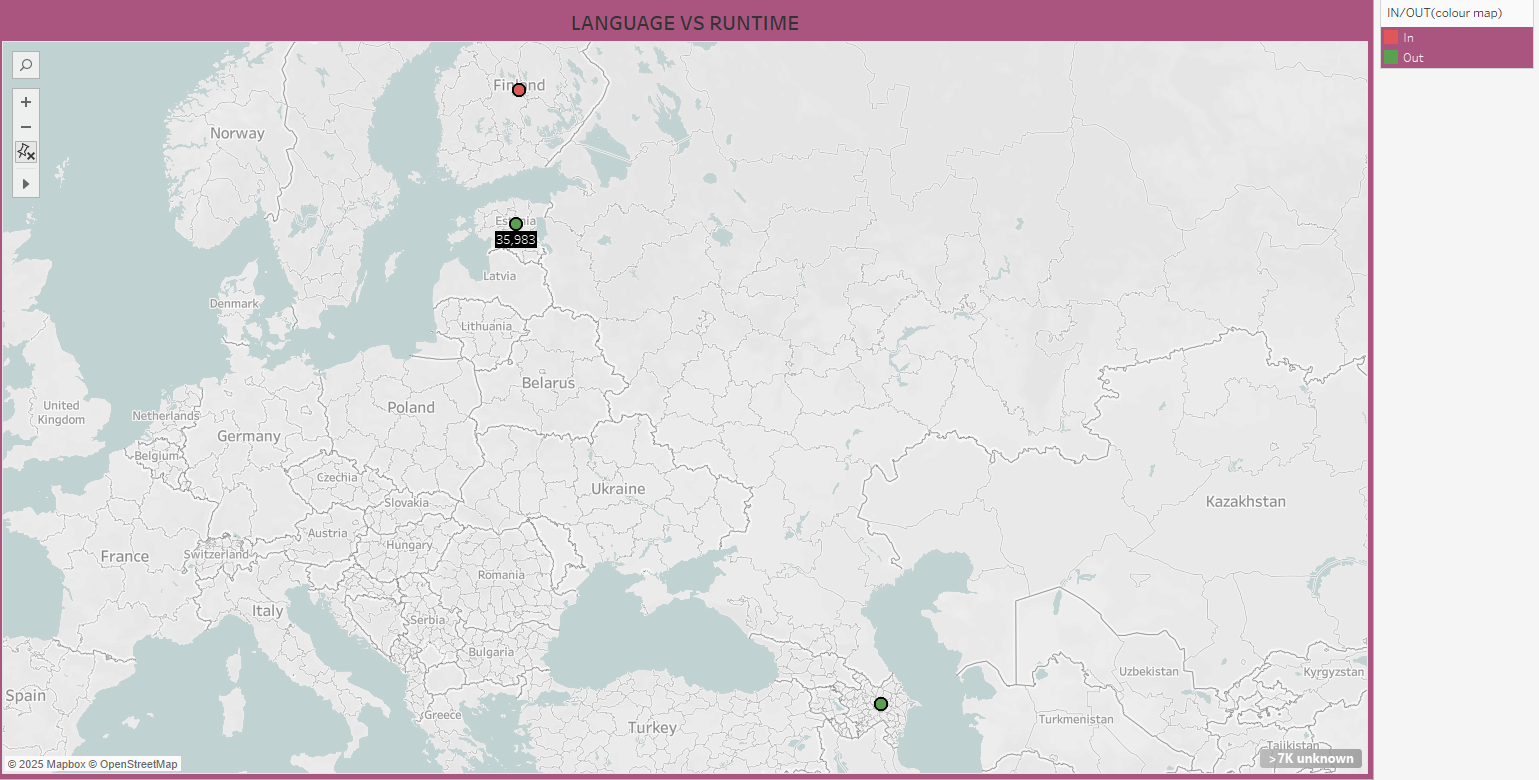
**Runtime vs Popularity Relationship:**

* **Most movies have runtimes between **100 and 180 minutes**.**
* ****Popularity is scattered** — there's no very strong correlation between runtime and popularity.**

**Movies that are **very short** (close to 0 runtime) and **very long** (over 250 minutes) tend to have **low popularity**.**

**This chart shows that **moderate-length movies tend to be the most popular**, while **very short or very long movies don't do as well**.**

**LANGUAGE VS RUNTIME:**



**Points are plotted based on **location**, with colors representing **In** (pink/red) and **Out** (green)**.

****Locations:****

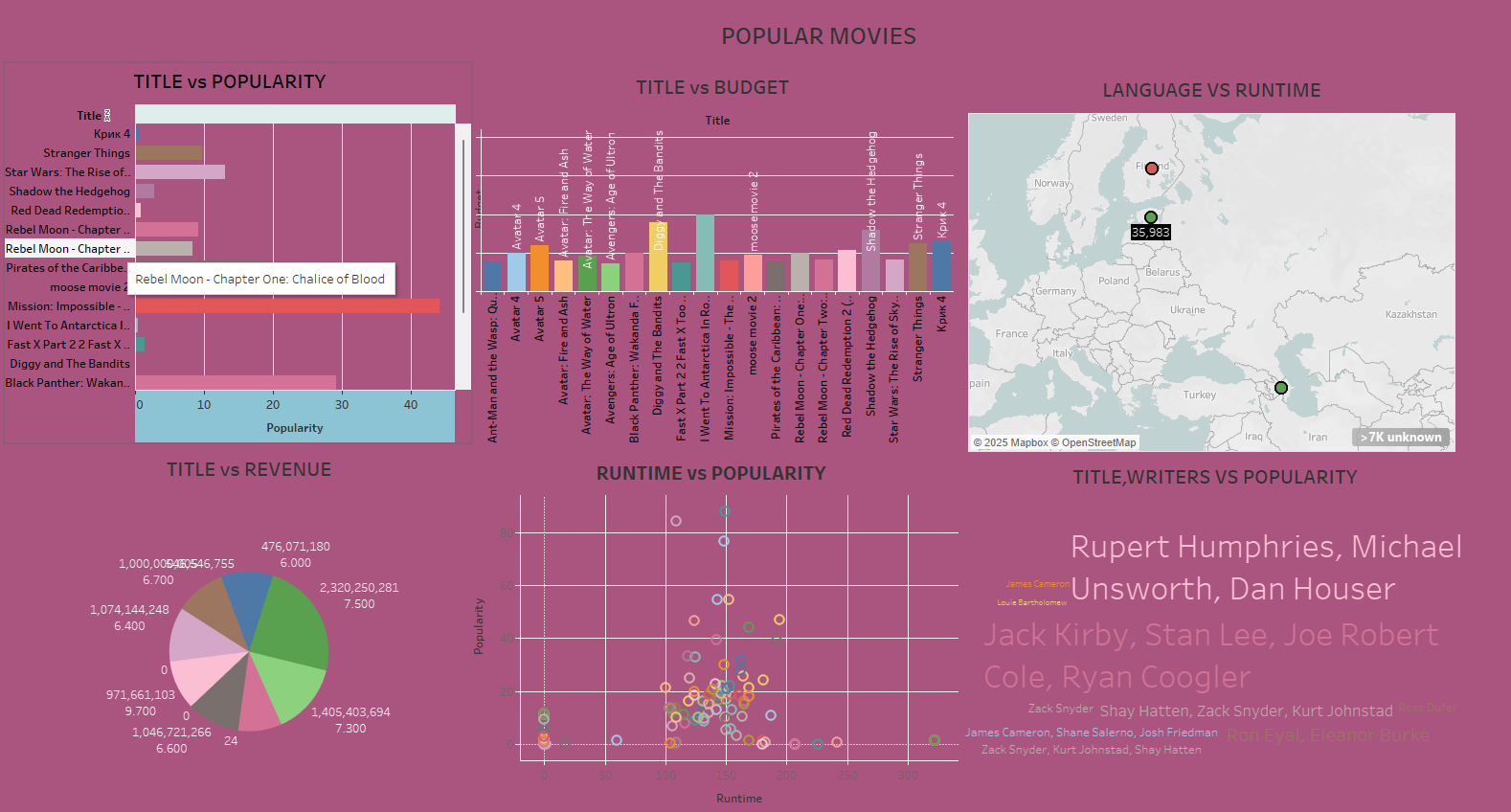
* **Most data points shown here are focused around **Eastern Europe** — especially near **Finland**, **Latvia**, and possibly **Georgia****
* **Language vs Runtime Meaning:**

**It seems to imply that **movies from these regions are being flagged** based on their inclusion or exclusion in the dataset**

**Runtime may relate to clustering — e.g., languages with shorter or longer average movie runtimes**

**This map shows **a very small number of movies with known language and runtime information** plotted by location**

**ALL MOVIES DASHBOARD:**



**Key Insights from Such Charts:**

****Most Popular Genres**: Action movies might dominate the dataset..**

****Language Trends**: English movies may be more prominent, but other languages can still make a noticeable percentage.**

Key Insights

* ****Rating Distribution**: Most movies likely hover around mid-to-high ratings (e.g., 6-8).**
* ****Budget and Earnings**: Higher-budget movies generally bring in more revenue, but there are exceptions**
* ****Global Production**: Hollywood might be the leader in movie production, but countries like India (Bollywood) or France may also have a significant presence**

### ****Bar Chart - Most Popular Genres****

**This could display the number of movies in each genre**

****X-axis**: Genre (e.g., Action, Comedy, Drama)**

****Y-axis**: Count of Movie**

****Purpose**: Shows which genres are most common in the dataset.**

**Example:**

****Action** = 100 movies**

****Drama** = 80 movies**

****Comedy** = 50 movieS**

### **2. **Pie Chart - Language Distribution****

**This shows how the movies are distributed by **language****

****Purpose**: To visualize the proportion of movies in different languages (e.g., English, Spanish, French).**

****Example:****

****English**: 75%**

****Spanish**: 15%**

****French**: 10%**

### **3. **Histogram - Ratings Distribution****

**A histogram can show the distribution of movie ratings.**

****X-axis**: Rating value (e.g., from 0 to 10)**

****Y-axis**: Number of Movies**

****Purpose**: To see how movies are rated (e.g., Most movies might have ratings between 7-8).**

**EXAMPLE:**

**Movies with **8-9 rating**: 200 movies**

### **4. **Scatter Plot - Budget vs. Box Office****

**This shows the relationship between the **budget** and **box office earnings** for each movie.**

****X-axis**: Budget (e.g., $1 million, $100 million)**

****Y-axis**: Box Office Revenue (e.g., $10 million, $500 million)**

****Purpose**: To identify if high-budget movies tend to perform better.**

**Example:**

**Movies with **large budgets** tend to have **higher box office earnings**, but there could be some surprises where a low-budget movie earned significantly more.**

### **5. **World Map - Movie Production by Country****

**This chart can map the countries where movies were produced.**

****Purpose**: To see how many movies came from each country.**

****Example**:**

****USA**: 500 movies**

****India**: 300 movies**

****UK**: 100 movies**

**Conclusion**

****All Movies Dataset** is a powerful way to explore and understand the film industry through data. It allows for creative and analytical problem-solving while also giving you a chance to sharpen your skills in **data processing, visualization, and machine learning**. Whether you’re interested in entertainment, marketing, or simply want to work with an interesting dataset, the insights derived from this dataset can be both fun and insightful.**

**The movies dataset is typically represented in tabular format, where each row corresponds to a movie, and each column represents a specific attribute**

**The data can reveal insights into the **commercial side** of movies, such as the relationship between budget and earnings or how ratings correlate with movie popularity.**