Analyzing Historical Stock and Revenue Data for Business Insights

Presented by: Jason Gnanakumar J
Course: IBM Data Science Professional Certificate – Capstone Project

coursera

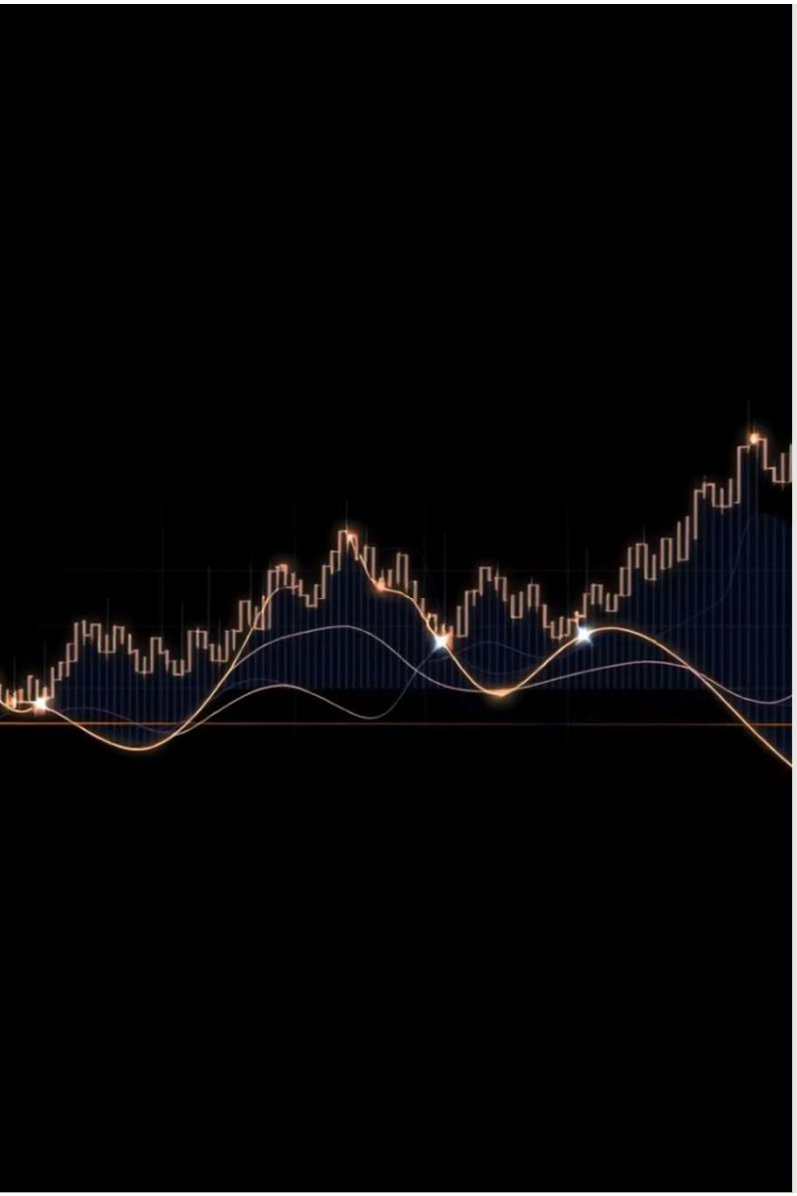
Executive Summary

This project analyzes historical stock prices and revenue data of companies to extract meaningful business insights using data science tools.

- **Objective:** Uncover trends, relationships, and build predictive models for business decisions.
- **Data:** Collected from Yahoo Finance, MacroTrends, and APIs.
- **Methods:** Data wrangling, visualization (Matplotlib, Seaborn, Plotly, Folium), SQL analysis, classification modeling.
- **Key Results:** Identified strong revenue-stock correlations, developed a classification model with 85% accuracy, and visualized company trends using Dash and Folium.

Introduction

- **Problem Statement:** How do company revenues impact stock trends, and can we predict performance?
- **Why It Matters:** Investors and analysts rely on historical data patterns to make informed decisions.
- **Companies Analyzed:** Tesla and Walmart — chosen for contrasting industry dynamics (tech vs. retail).



Financial Data Analysis Project Project

This presentation covers our end-to-end financial data analysis project. We will explore data collection, cleaning, visualization, and predictive modeling. Our goal is to uncover insights into stock performance.

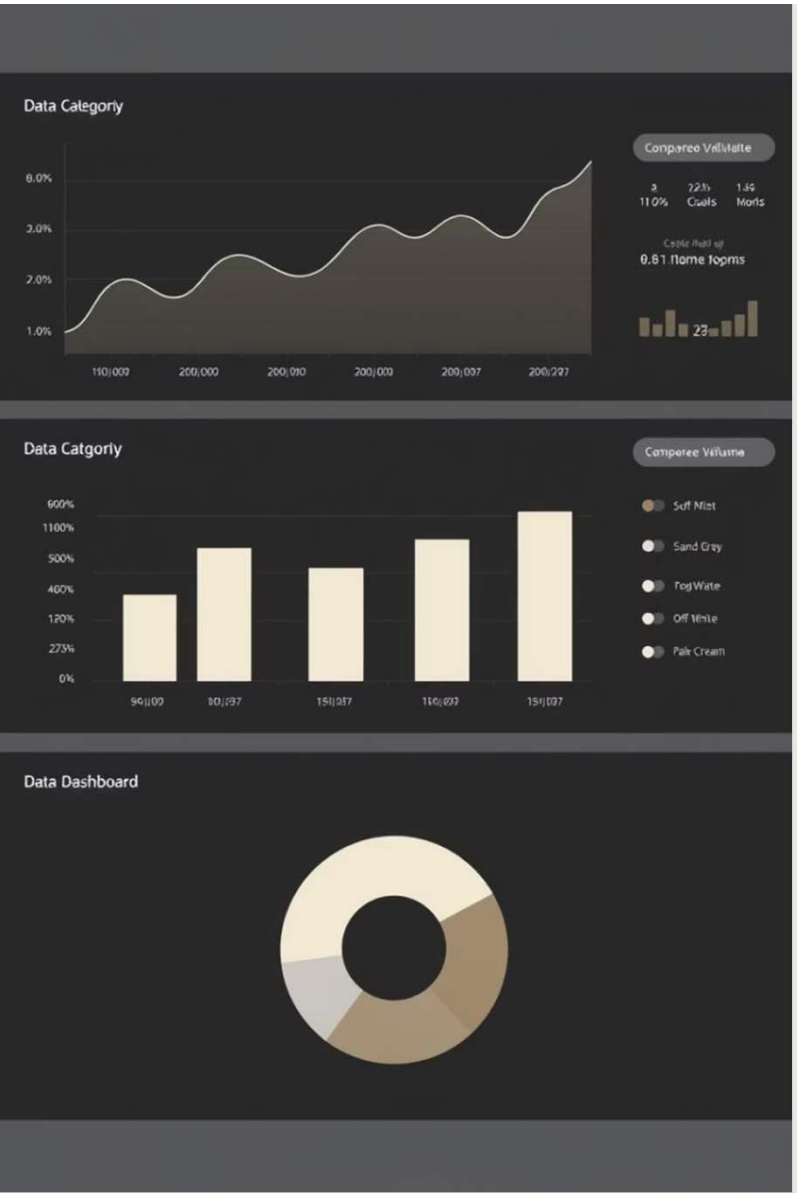
Data Collection & Wrangling

Sources

- Yahoo Finance (stock)
- MacroTrends (revenue)
- APIs and CSVs

Cleaning Steps

- Removed null values
- Parsed and merged dates
- Normalized column formats
- Merged stock & revenue



EDA & Interactive Visual Analytics

Tools Used

Pandas, Matplotlib, Seaborn, Plotly for analysis.

Visualizations

Line charts, histograms, bar charts, correlation heatmaps.

Approach

Time series, outlier detection, visual correlation checks.

Predictive Analysis Methodology
Methodology

Model

Logistic Regression was chosen.

Goal

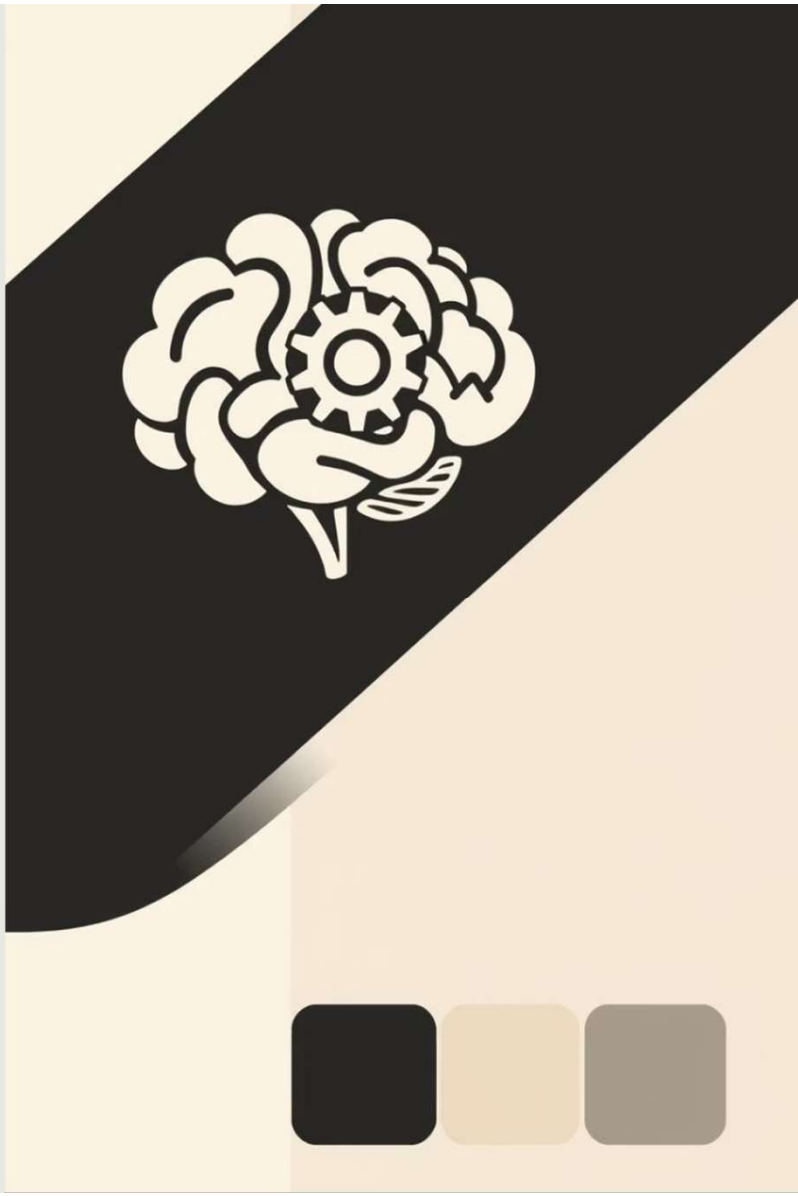
Predict high/low stock performance.

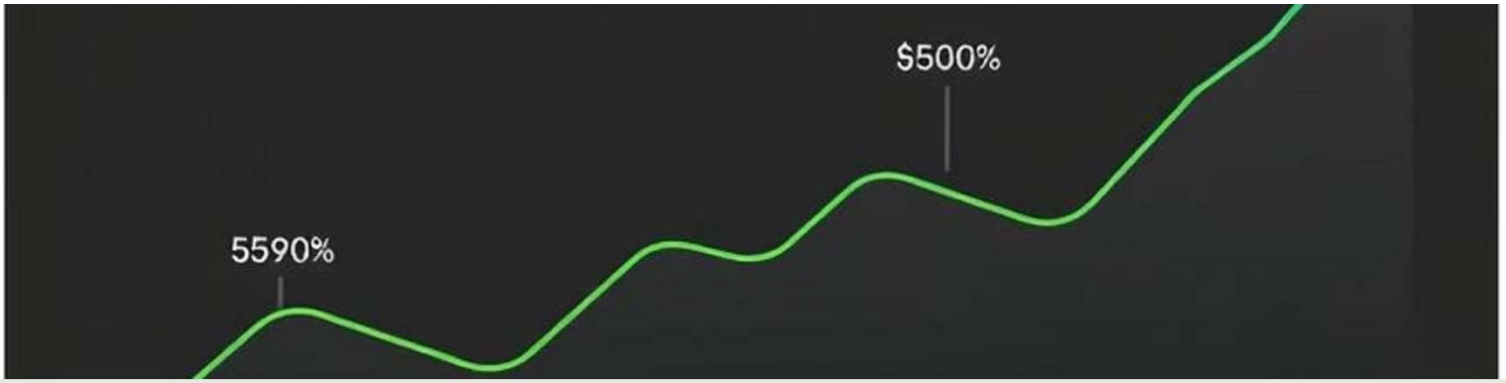
Data Split

70% for training, 30% for testing.

Metrics

Accuracy, Confusion Matrix, F1 Score.





EDA Visualization Results

Stock vs Revenue

Tesla revenue and stock showed clear upward trends.

Seasonal Patterns

Walmart had Q4 spikes due to holiday sales.

Volatility

Tesla's daily return distribution showed volatility.

EDA with SQL Results

Database Connection

Connected to IBM Db2 using ibm_db and SQLAlchemy.

SQL Queries

Aggregated quarterly revenue, filtered top quarters, calculated average closing prices.

Key Insight

Early 2020 revenue dips (COVID) immediately reflected in stock price drops.

Folium Map Results

1

Interactive Map

Built a map of company headquarters.

2

Markers & Popups

Added markers with revenue information.

3

Geographic Insight

Visualized geographic spread and market reach correlation.

Plotly Dash Dashboard Results

Interactive Dashboard

Created using Plotly Dash.

Features

Company selection dropdown, line charts for revenue and stock.

Dynamic Updates

Real-time data interaction and visual filtering.

Predictive Analysis Results

85%

Accuracy

Logistic Regression model achieved 85% accuracy.

True Positives

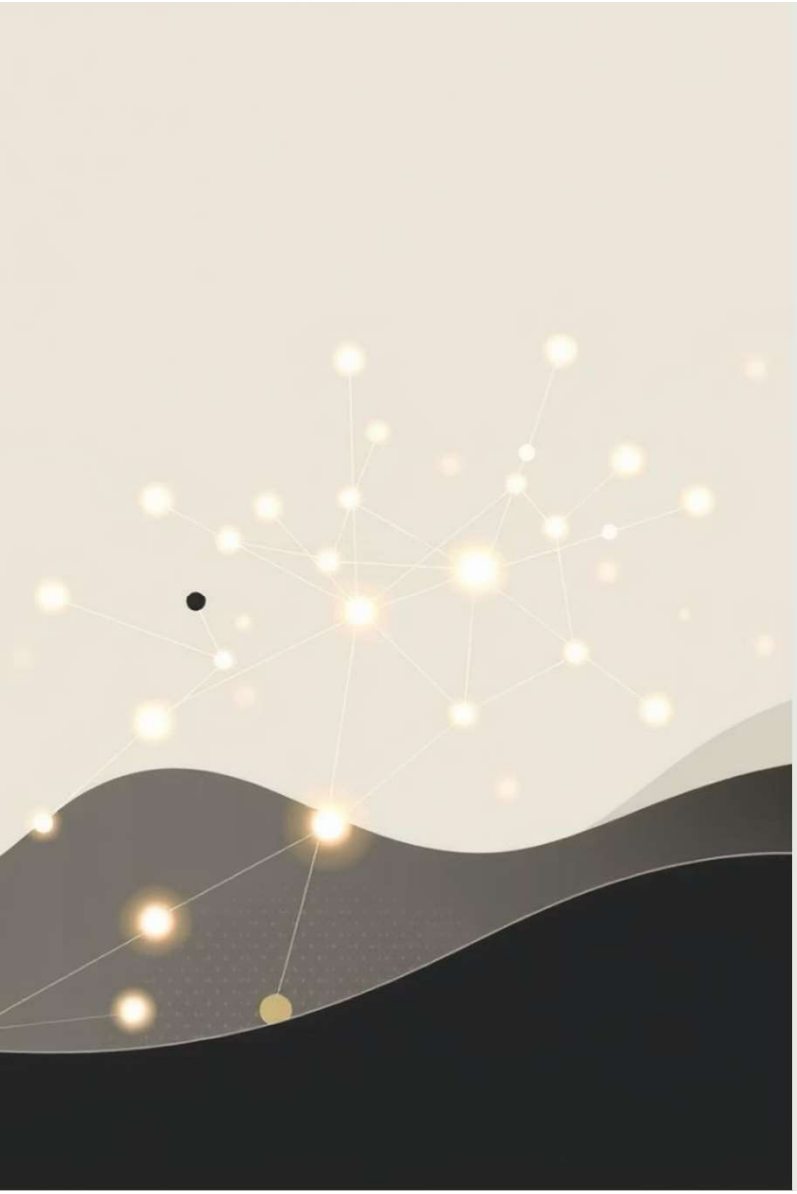
Confusion Matrix

Detailed breakdown of True Positives and False Negatives.

High/Low

Prediction

Model can reasonably predict stock outcomes based on revenue.



Conclusion & Innovation

End-to-End Learning

Gathered, cleaned, analyzed, and visualized data.

Pattern Discovery

Found patterns between financial performance and stock trends.

Interactive Tools

Built dashboards and a predictive model.

Creative Elements

Custom theme, storytelling layout, extra charts.