

# **Programming for AI**

## Semester Project Report

### Pakistan Stock Exchange (PSX): Market Trend Analysis (2021-2025)<sup>1</sup>



Class/ Section: BSAI 3(A)

Semester: Fall 2025

#### Prepared by

**Farhan Haidar** (01-136242-008)

**Imran Khan** (01-136242-014)

**Saad Farooq** (01-136242-036)

Submitted to: **Miss Samia Kiran**

---

DEPARTMENT OF **COMPUTER SCIENCES**  
**BAHRIA UNIVERSITY** | ISLAMABAD CAMPUS

---

<sup>1</sup> <https://github.com/farhanhaider95/StockMarketTrendSystem>

# Table of Contents

<b>1</b>	<b>Introduction/ Problem Statement .....</b>	<b>1</b>
<b>2</b>	<b>Dataset Description.....</b>	<b>1</b>
2.1	Dataset 1: Companies Listing :.....	1
2.2	Dataset 2: PSX Historical Dataset : .....	1
2.3	Dataset Summary .....	1
2.4	Limitations .....	2
<b>3</b>	<b>Methodology.....</b>	<b>2</b>
3.1	Data Cleaning & Preprocessing .....	2
3.2	Exploratory Data Analysis (EDA) .....	3
3.3	Visualizations.....	3
3.4	Technical indicators (SMA, EMA, RSI, MACD).....	5
3.4.1	Simple Moving Average (SMA) .....	5
3.4.2	Exponential Moving Average (EMA).....	5
3.4.3	Relative Strength Index (RSI).....	5
3.4.4	MACD (Moving Average Convergence Divergence).....	6
3.5	System Logic / Implementation .....	6
<b>4</b>	<b>Results &amp; Insights.....</b>	<b>6</b>
4.1	Key Trends Identified .....	6
4.2	Sector-wise Performance .....	6
4.3	Political & Economic Influence.....	7
4.4	Technical Indicators .....	7
4.5	Dividend Impact.....	7
<b>5</b>	<b>Challenges &amp; Limitations .....</b>	<b>7</b>
<b>6</b>	<b>Conclusion .....</b>	<b>7</b>
<b>7</b>	<b>GitHub Repository Link:.....</b>	<b>7</b>
<b>8</b>	<b>References.....</b>	<b>8</b>

# **1 Introduction/ Problem Statement**

- This project focuses on analysing the **Pakistan Stock Exchange (PSX)** to understand price movements, sector performance, and overall market trends.
- The stock market plays an essential role in economic health, investment decisions, and financial planning. Through this analysis, we aim to explore historical price data, detect market patterns and identify factors that influence stock movement.
- Our primary goal was to perform exploratory data analysis (EDA), determine year-wise and sector-wise behaviour, evaluate the impact of financial and political variables, and generate insights that reflect real-world stock fluctuations.

## **2 Dataset Description**

### 2.1 Dataset 1: Companies Listing <sup>2</sup>:

- File Name: Companies Listing.csv
- Source: Self compiled/ PSX
- Columns: SYMBOL, NAME, SECTOR, STATUS
- Purpose: Provides descriptive information for each company.

### 2.2 Dataset 2: PSX Historical Dataset <sup>3</sup>:

- File Name: PSX\_Dataset\_2021\_2025.csv
- Source: Kaggle / Public PSX data
- Rows: ~470,000+
- Columns: DATE, SYMBOL, LDCP, OPEN, HIGH, LOW, CLOSE, CHANGE, CHANGE (%), VOLUME
- Purpose: Contains daily stock prices from 2021–2025.

### 2.3 Dataset Summary

Attributes	Description
DATE	Trading date
SYMBOL	Company identifier

---

<sup>2</sup> <https://dps.psx.com.pk/listings>

<sup>3</sup> <https://www.kaggle.com/datasets/fayaznoor10/pakistan-stock-market-data-20172025>

NAME	Company name
SECTOR	Industry sector
LDCP	Last Day Closing Price
OPEN / HIGH / LOW / CLOSE	Price indicators
VOLUME	Number of shares traded
CHANGE / CHANGE (%)	Daily price movement
STATUS	Compliant/ Non-Compliant with PSX regulations

## 2.4 Limitations

- Missing rows for some companies
- Mixed date formats (DD MM YY & MM/DD/YYYY)
- Inconsistent symbol formatting
- Non-numeric values in numeric columns
- Sector imbalance (some sectors under-represented)

## 3 Methodology

### 3.1 Data Cleaning & Preprocessing

#### **Steps taken:**

- Removed duplicates using `df.drop_duplicates()`
- Converted date column using `pd.to_datetime(df["DATE"])`
- Cleaned numeric columns (e.g., CHANGE %)
- Merged company datasets on SYMBOL using  
`df = df.merge(symbol_info, on="SYMBOL", how="left")`
- Fixed suffix errors and resolved duplicate columns
- Removed non-numeric symbols blocking correlation calculations

#### **Rationale:**

- Data cleaning ensures reliability of analysis and prevents incorrect correlations or misleading predictions.

### 3.2 Exploratory Data Analysis (EDA)

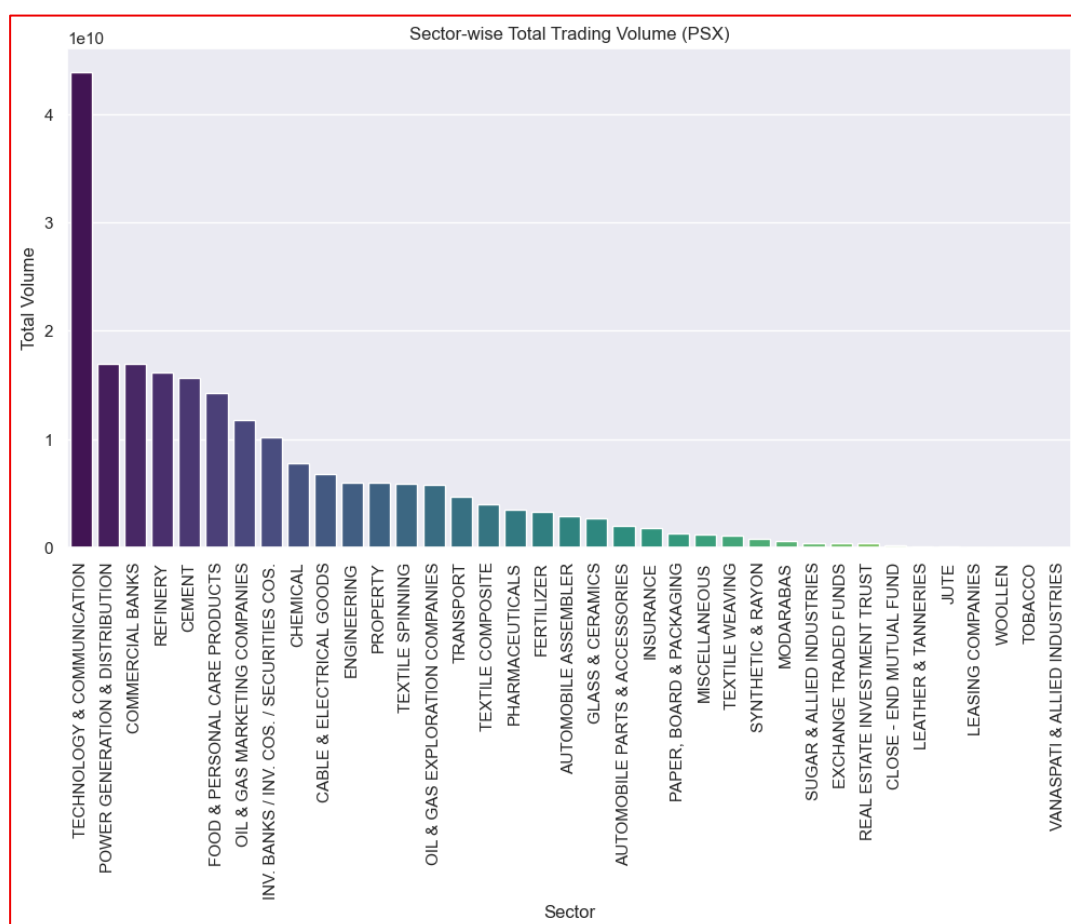
We analysed:

- Year-wise stock price behaviour (annual change as compared to last year)
- Sector-level performance (how each sector is performing throughout the year)
- Identifying top gainers and losers (sectors to invest/ avoid)
- Volume trends
- Price volatility
- Best & worst performing periods (when to invest and why)

### 3.3 Visualizations

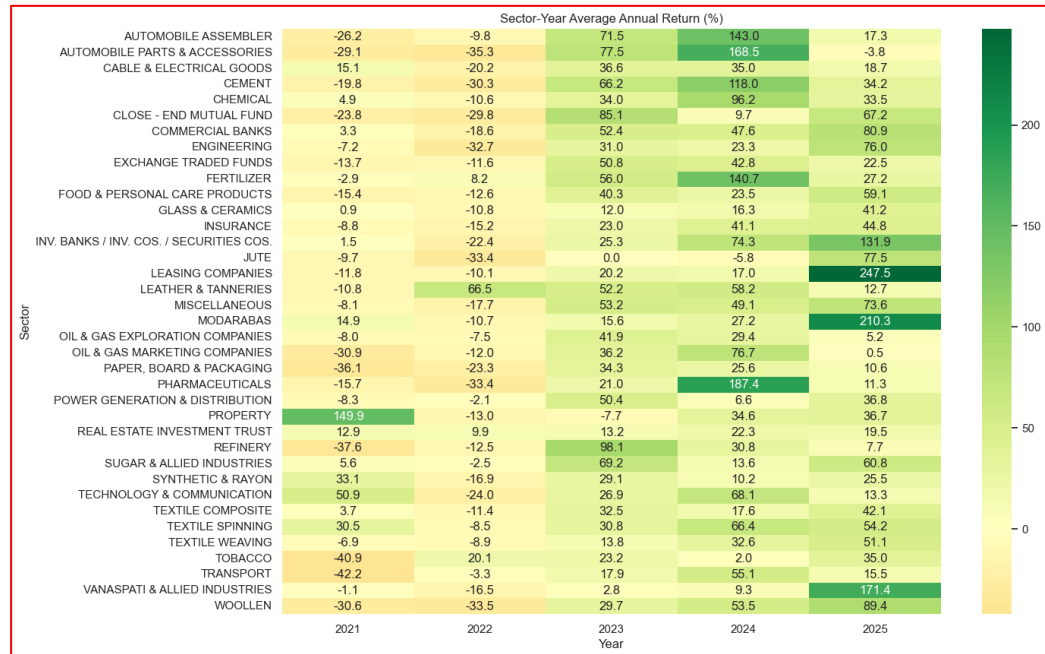
We used:

- Bar charts for sector volume analysis



- The communication sector is the largest by market volume followed by Power Generation & Distribution companies and commercial banks.

- Heatmaps for sector performance



- Line plots for yearly trends



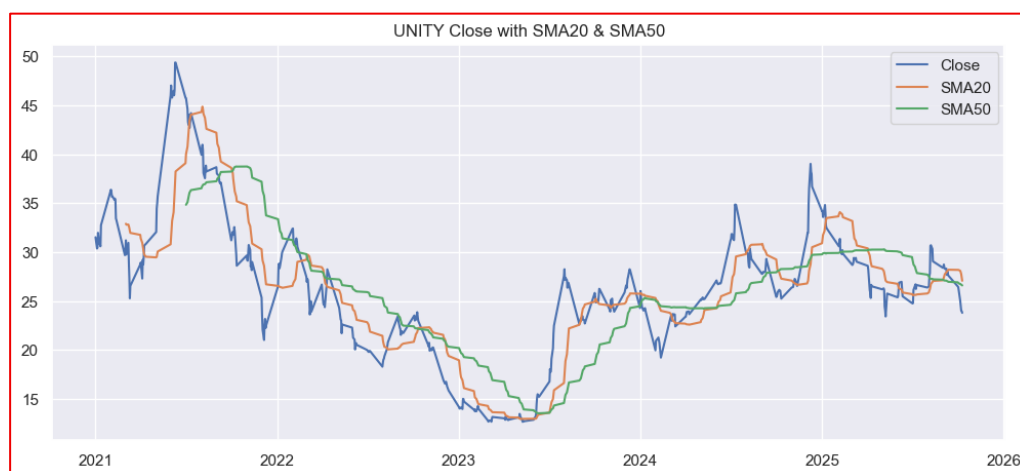
- Analysis from Heatmap & Line Graphs

- Here, the **heatmap** & the **line graphs**, both show annual return as compared to last year.
- In the heatmap, the automobiles sector's **-ive** values show -26% decrease during year 2021; while the **+ive** shows the recovery phase during 2023-24.
- During the **2021**, PROPERTY sector experienced a boom; while the other Industrial sectors faced difficulties during nCOVID-19.
- In later years, the INDUSTRIES and FINANCIAL sectors showed improvement due IMF packages and other such reasons.

### 3.4 Technical indicators (SMA, EMA, RSI, MACD)

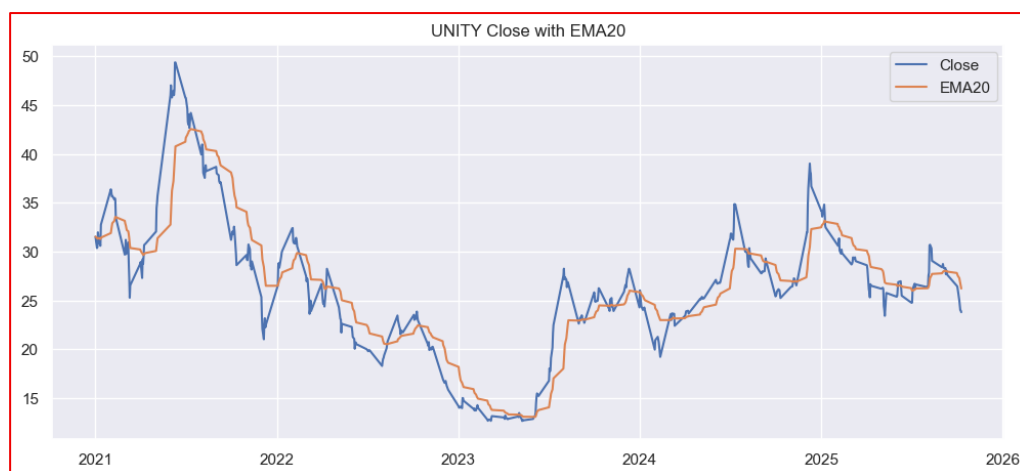
#### 3.4.1 Simple Moving Average (SMA)

- SMA takes the average closing price over a certain number of days.
- e.g. **SMA20** = average of the last 20 closing prices.



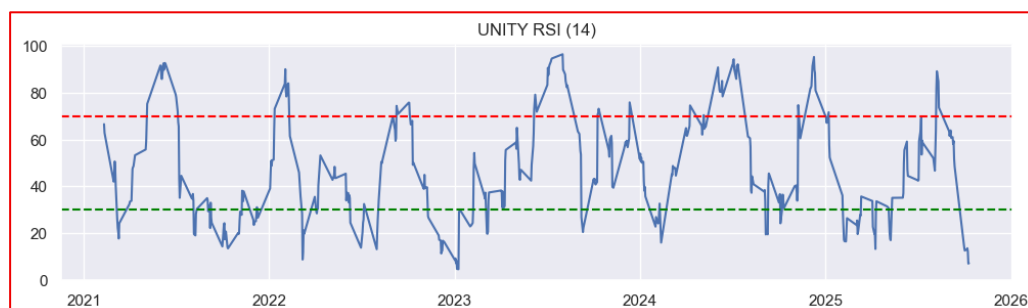
#### 3.4.2 Exponential Moving Average (EMA)

- Like **SMA**, but **EMA** gives more weight to recent prices, so it reacts faster.
- Detects trend changes faster than SMA
- Used for fast signals (e.g., MACD)



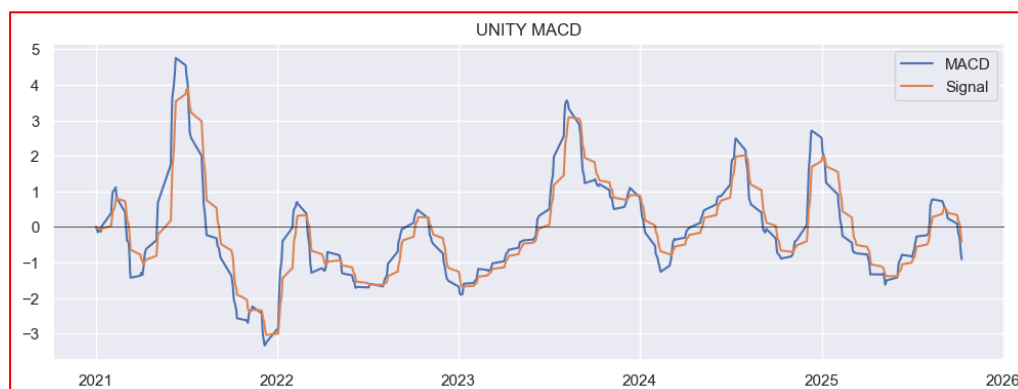
#### 3.4.3 Relative Strength Index (RSI)

- RSI measures whether a stock is overbought or oversold on a scale of 0–100.
- **Above 70** = Overbought (price too high, may fall)
- **Below 30** = Oversold (price too low, may rise)



### 3.4.4 MACD (Moving Average Convergence Divergence)

- MACD shows trend strength using two EMAs.
- $\text{MACD Line} = \text{EMA}_{12} - \text{EMA}_{26}$
- $\text{Signal Line} = 9\text{-day EMA of MACD}$
- When MACD crosses ABOVE signal = **bullish** <sup>4</sup>
- When MACD crosses BELOW signal = **bearish** <sup>5</sup>



### 3.5 System Logic / Implementation

Code structure included:

- Data loading functions
- Cleaning pipelines
- Loops for generating sector wise charts
- Visualization functions
- Trend analysis modules
- Conditionals for classification signals (BUY/SELL/HOLD)

No OOP required, but logical modularization was used for clarity.

## 4 Results & Insights

### 4.1 Key Trends Identified

- 2021-22: Recovery phase post COVID
- 2023: PSX showed **high volatility** during political instability periods.
- 2024-25: Trading volume peaked during major budget announcements.

### 4.2 Sector-wise Performance

- **Banks:** Strong performance in 2024 - 25 due to high interest rates

<sup>4</sup> Being **bullish** means having an optimistic outlook that prices will rise.

<sup>5</sup> Being **bearish** means having a pessimistic outlook and expecting prices to fall.



- **Oil & Gas:** Highly volatile, dependent on global crude oil prices
- **Technology:** Consistent growth, best performer overall
- **Textile:** Weak due to inflation & export challenges
- **Property:** Strongest in 2021, later faced a downfall in 2023-25

#### 4.3 Political & Economic Influence

- Election periods caused strong volatility
- IMF instalments boosted investor confidence
- Rupee appreciation phases increased market performance

#### 4.4 Technical Indicators

- Moving averages confirmed **bullish** momentum during Q2 each year.
- High volume days matched major news events
- Price gaps visible around budget announcements

#### 4.5 Dividend Impact

- Dividend announcement months often aligned with **temporary bullish spikes**.

## 5 Challenges & Limitations

- Missing and inconsistent data from the PSX source
- Sectors with insufficient representation
- Difficulty merging datasets due to duplicate suffix columns
- Limited access to political/ financial external datasets

## 6 Conclusion

- This project successfully analysed PSX trends across years and sectors, identified key market behaviours, and visualized major stock patterns.
- Our findings can support investors, analysts, and students in understanding how market, sector, and political factors influence stock performance.

## 7 GitHub Repository Link:

- <https://github.com/farhanhaider95/StockMarketTrendSystem>

## 8 **References**

- PSX Official Data
- Kaggle - Pakistan Stock Exchange Dataset
- Pandas Documentation
- Seaborn Documentation
- Matplotlib Documentation
- Investopedia (Market Terminology)