# PROJECT REPORT

**Project Name: S&P 500 Stock Trends – Market Performance & Sector Analysis**

**Project Type:** Individual, Company-Based Analytics Project
**Candidate: Pran Wasnik**
**Role:** Data Analyst Intern
**Tools:** Python (Pandas, NumPy, Matplotlib, Seaborn), Google Collab, PowerPoint, MS Word
**Dataset Source:** Kaggle (Dummy / Practice Dataset)

---

## 1. Executive Summary

This project focuses on analyzing historical stock price trends of S&P 500 companies to understand **overall market performance, sector-level behavior, risk patterns, and portfolio outcomes**. The objective is to simulate how financial data analysts support **investment decisions, strategic planning, and risk management** using data-driven insights.

The S&P 500 index is widely considered a benchmark for the U.S. equity market. By analyzing its historical data, businesses and investors can identify **growth opportunities, risk exposure, and diversification benefits** across different market sectors.

This project delivers insights into:

- Market trend behavior over time
- Sector-wise risk and return comparison
- Correlation between sectors
- Portfolio performance under different allocation strategies

All analysis was performed using a **dummy/practice dataset sourced from Kaggle**, strictly for learning and skill demonstration.

## 2. Business Problem & Use Case

### Business Problem

Financial markets are volatile and influenced by multiple economic factors. Organizations and investors often face challenges in:

- Understanding which sectors drive market growth
- Balancing risk and return
- Designing portfolios that are resilient to market fluctuations

Without data-backed analysis, investment decisions can become speculative and risky.

### Use Case

This analysis helps answer critical business questions such as:

- How has the S&P 500 market behaved over time?
- Which sectors generate higher returns and which carry higher risk?
- How correlated are different sectors?
- How do portfolio strategies impact long-term performance?

**Business Users**

- Investment and portfolio management teams
- Corporate finance and strategy teams
- Financial analysts and decision-makers

**3. Data Overview**

The dataset used for this project was sourced from **Kaggle**.

**Dataset Characteristics:**

- **Type:** Historical daily stock price data
- **Companies:** S&P 500 listed companies
- **Key Fields:**
    - Date
    - Open
    - High
    - Low
    - Close
    - Volume
- **Time Coverage:** Multiple years of historical data

**Important Note:**
The dataset is a **dummy/practice dataset** used strictly for educational and analytical simulation. It does not represent real-time or live market data.

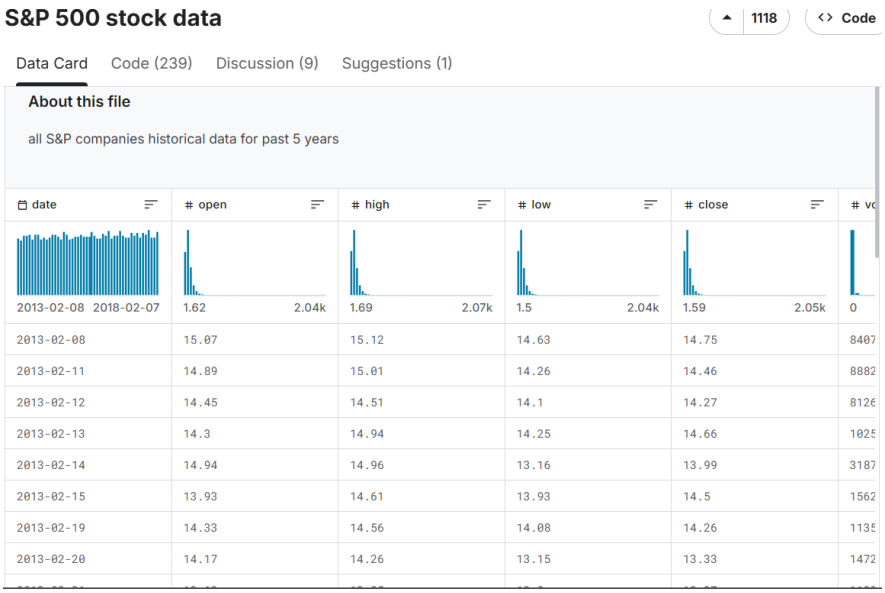


**Figure 3.1: S&P 500 Dataset Overview from Kaggle**

## 4. Data Preparation & Cleaning

Before analysis, the dataset was processed to ensure accuracy and reliability.

**Key Steps Performed:**

- Identification and handling of missing or inconsistent values
- Conversion of date fields into proper datetime format
- Sorting data chronologically for time-series analysis
- Calculation of **daily stock returns** from closing prices
- Mapping companies to their respective market sectors
- Validation checks for duplicates and anomalies

These steps ensured that the data was **clean, structured, and analysis-ready**.

---

all_stocks_5yr.csv

```
df = pd.read_csv("/content/all_stocks_5yr.csv")
df.head()
df.info()
# Convert date column to datetime
df['date'] = pd.to_datetime(df['date'])

# Sort properly for time-series calculations
df = df.sort_values(['Name', 'date'])

# Check missing values
df.isnull().sum()
df['daily_return'] = df.groupby('Name')['close'].pct_change()
total_return = df.groupby('Name').agg(
    start_price=('close', 'first'),
    end_price=('close', 'last')
)
```

**Figure 4.1: Data Loading and Cleaning in Python**

```
total_return['total_return_%'] = (
    (total_return['end_price'] / total_return['start_price'] - 1) * 100
)
top_10 = total_return.sort_values(
    'total_return_%', ascending=False
).head(10)

top_10
plt.bar(top_10.index, top_10['total_return_%'])
plt.title("Top 10 S&P 500 Companies by 5-Year Total Return")
plt.ylabel("Total Return (%)")
plt.xticks(rotation=45)
plt.show()
```

**Figure 4.2: Daily Return Calculation**

### 5. Exploratory Data Analysis (EDA)

Exploratory Data Analysis was conducted to understand overall market behavior and stock performance patterns.

**EDA Insights:**

- Identified long-term market growth trends
- Observed periods of high and low volatility
- Analyzed distribution of daily returns
- Highlighted top-performing stocks based on historical returns

EDA helped uncover meaningful patterns before deeper sector and portfolio analysis.

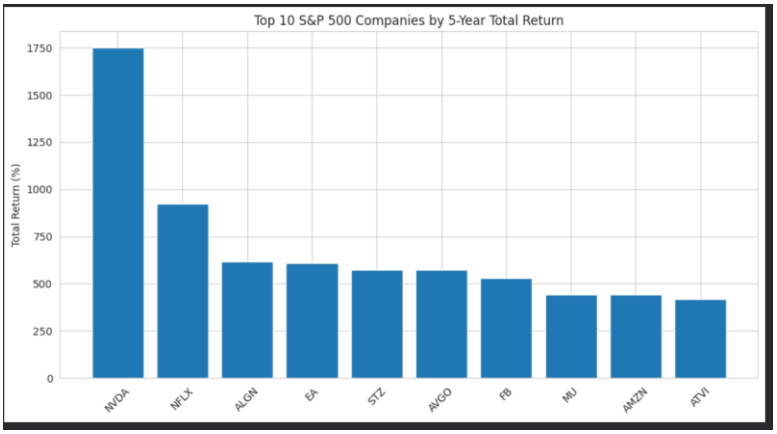


**Figure 5.1: Overall Market Trend Visualization**

### 6. Sector-Level Analysis

Sector-level analysis was performed to evaluate **risk and return behavior across industries**.

**Key Metrics:**

- Average sector-wise returns
- Sector-wise volatility (risk measurement)

- Risk vs return comparison

**Interpretation:**

- Growth-oriented sectors showed higher returns with increased volatility
- Defensive sectors exhibited stable returns with lower risk
- Sector behavior aligned with real-world economic cycles

This analysis helps businesses align investment strategies with risk tolerance.


**Figure 6.1: Sector-Wise Average Returns**

```
Top 10 Companies by 5-Year Return:
      start_price  end_price  total_return_%
Name
NVDA      12.3700     228.80     1749.636217
NFLX      25.8528     264.56      923.332096
ALGN      32.7300     234.33      615.948671
EA        17.3700     123.05      608.405296
STZ       31.8500     214.15      572.370487
AVGO      35.3200     237.38      572.083805
FB        28.5450     180.18      531.213873
MU         7.7500      42.01      442.064516
AMZN     261.9500    1416.78      440.858943
ATVI      13.4100      69.46      417.971663

Sector Risk-Return Summary:
            avg_return  volatility
Sector
Energy        0.000144    0.015202
Financials    0.000631    0.014060
Technology    0.001175    0.015898

Sector Correlation Matrix:
Sector         Energy  Financials  Technology
Sector
Energy       1.000000    0.534014    0.280018
Financials   0.534014    1.000000    0.381177
Technology   0.280018    0.381177    1.000000
```

**Figure 6.2: Top 10 Companies by 5 Year Return**

## 7. Correlation Analysis

Correlation analysis was conducted to understand relationships between major market sectors such as Technology, Financials, Energy, and Healthcare.

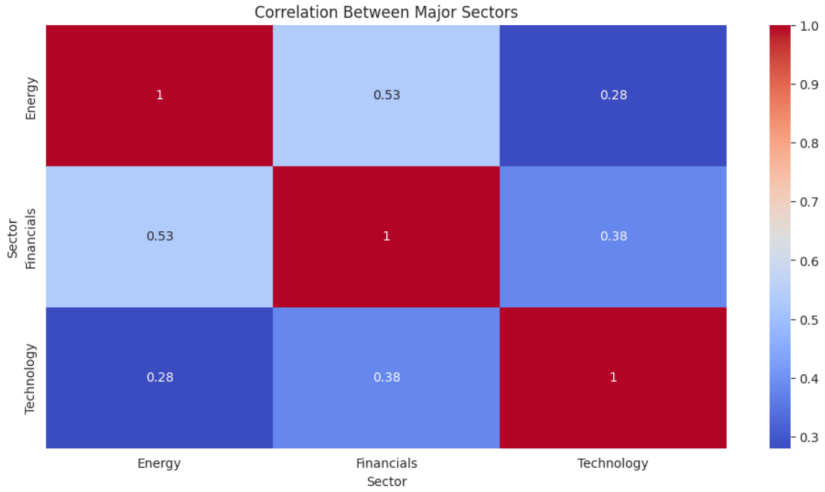


**Figure 7.1: Correlation Between Major Sectors**

**Business Interpretation:**

- High correlation indicates similar movement patterns
- Low or negative correlation supports diversification
- Sector correlation insights help reduce portfolio risk

## 8. Portfolio Analysis

Two portfolio strategies were analyzed to compare performance and risk exposure.

**Portfolio Types:**

1. **Equal-Weight Portfolio**
    - Equal allocation across all stocks
    - Broad market exposure
    - Reduced concentration risk
2. **Sector-Weighted Portfolio**
    - Allocation based on sector importance
    - Strategic focus on selected sectors
    - Higher potential returns with increased risk

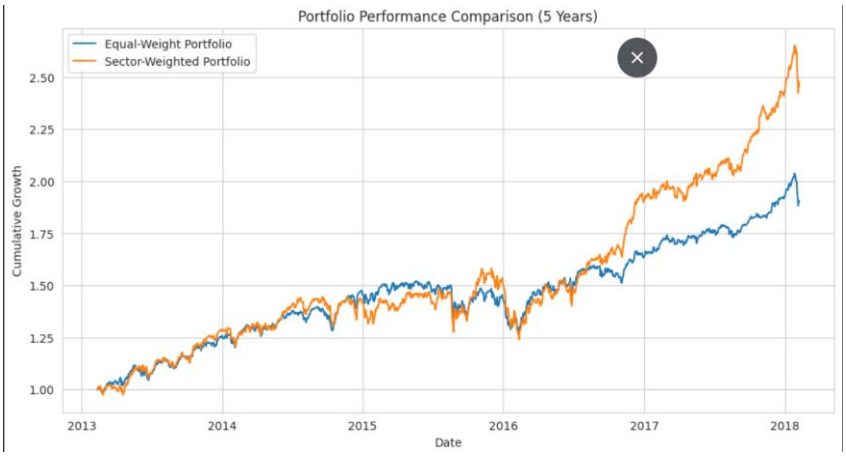


**Figure 8.1: Portfolio Performance Comparison**

**Outcome:**

Portfolio performance varied significantly based on allocation strategy.

**9. Key Business Insights**

- Sector diversification improves portfolio stability
- Higher returns are associated with higher volatility
- Correlation analysis supports better risk management
- Portfolio strategy impacts performance more than individual stock selection
- Data-driven analysis enables informed financial decisions

**10. Limitations & Assumptions**

- Dataset is dummy/practice-based
- Sector classification may be simplified
- Real-time market factors are not included
- Historical performance does not guarantee future results

---

**11. Conclusion**

This project successfully demonstrates a **real-world financial analytics workflow** from data sourcing to executive-level reporting. It highlights how data analysts:

- Clean and validate financial data
- Analyze market and sector behavior
- Evaluate portfolio strategies
- Translate analysis into business insights

- Overall, the project reflects **practical responsibilities of a Data Analyst working in finance and investment analytics**.