## Executive Summary

This project implements a comprehensive data warehouse solution to analyze the relationships between major economic indicators and market performance. Through dimensional modeling and advanced analytics, we explore the intricate connections between GDP, S&P 500, housing prices, and other key economic metrics.

This analysis provides valuable insights for investment strategies, risk management, and policy considerations, demonstrating the power of structured data analysis in understanding complex economic relationships.

## Problem Statement

In an increasingly complex financial landscape, investors and policymakers face the challenge of understanding the intricate relationships between macroeconomic indicators (GDP, housing prices) and financial market performance (S&P 500). The ability to analyze these relationships is crucial for making informed investment decisions and policy recommendations, particularly given the historical patterns of market volatility and economic cycles from 1974-2013. This project addresses the critical need to understand how different economic sectors interact and influence each other over time, with a specific focus on:

1. The correlation between GDP and S&P 500 performance, which impacts investment strategies and economic forecasting
2. The housing market's relationship with financial performance, particularly relevant given the significant market events like the 2008 financial crisis visible in your data
3. The interplay between multiple economic indicators (GDP, CPI, unemployment rate) and their combined effect on market performance

### Significance:

* Investment Decision Making: The analysis provides crucial insights for portfolio management and risk assessment by revealing historical patterns and correlations between economic indicators and market performance.
* Economic Policy: Understanding these relationships helps policymakers anticipate the potential impact of economic changes on different sectors.
* Risk Management: The historical analysis spanning multiple decades captures various economic cycles and crises, offering valuable insights for future risk assessment.
* Market Efficiency: The project contributes to understanding how quickly and effectively markets integrate macroeconomic information, which is essential for both institutional and retail investors.

Through dimensional modeling and OLAP analysis, this project creates a comprehensive framework for analyzing these complex relationships, enabling more informed decision-making in both financial markets and policy development.

## Data Collection and Preparation

The data for this analysis was sourced from the Federal Reserve Economic Data (FRED) database, The data had varying granularities across different economic indicators.

##### Source Data Overview:

* S&P 500 Index: Daily frequency trading data including open, high, low, close, volume, and adjusted close prices
* GDP: Quarterly economic output measurements
* House Price Index: Quarterly measurements of housing market trends
* CPI (Consumer Price Index): Monthly inflation indicator
* Unemployment Rate: Monthly labor market statistics

##### Granularity Harmonization:

* To create a cohesive analytical framework, data granularity was standardized through:
* Aggregation of daily S&P 500 data to monthly and quarterly levels using averaging techniques when needed.
* Alignment of monthly CPI and unemployment data with quarterly GDP and housing data through appropriate temporal aggregation.
* Creation of a unified date dimension (DW110.JULIAN\_DAYS) to support various time-based analyses.

##### Julian day Integration:

* Julian day was added to all fact tables and the data was updated through Transformed existing date values into corresponding Julian day numbers.

## Database Design

### OLAP Implementation

The data warehouse follows a star schema design optimized for OLAP operations, with one shared time dimension and three fact tables supporting different analytical perspectives.

1. **Dimensional Model Architecture**

Time Dimension (DW110.JULIAN\_DAYS):

* Serves as the central dimension table connecting all fact tables
* Implements a comprehensive date hierarchy with multiple grain levels:
  + Day level
  + Month level
  + Quarter level
  + Year Level
* Supports drill-down and roll-up operations across different time granularities
* Uses surrogate keys (JULIAN\_DAY) as PK to optimize join performance

Fact Tables:

1. SP500\_INDEX\_EOD\_FACTS:
   * Grain: Daily stock market data
   * Measures: OPEN, HIGH, LOW, CLOSE, VOLUME, ADJUSTED\_CLOSE
   * Foreign Key: JULIAN\_DAY links to dimension table
2. FACT\_QUARTERLY\_INDICATORS:
   * Grain: Quarterly economic metrics
   * Measures: GDP, HOUSE\_PRICE\_INDEX
   * Foreign Key: JULIAN\_DAY links to dimension table
3. FACT\_MONTHLY\_INDICATORS:
   * Grain: Monthly economic metrics
   * Measures: CPI, UNRATE (Unemployment Rate)
   * Foreign Key: JULIAN\_DAY links to dimension table
4. **OLAP Cube Implementation**

The design enables multiple analytical perspectives through a virtual OLAP cube:

* Dimensions: Time (Day, Month, Quarter, Year)
* Measures: Stock market metrics, Economic indicators
* Supported operations:
  + Slice: Analysis by specific time periods
  + Dice: Combining multiple indicators
  + Drill-down: From yearly to daily granularity
  + Roll-up: Aggregating daily data to higher time levels

Query Optimization Techniques

1. Temporary Tables:
   * Used for staging intermediate aggregations
   * Reduces repeated computations in complex analyses
   * Improves query performance for multi-step calculations
2. Common Table Expressions (CTEs):
   * Implements hierarchical time-based calculations
   * Enables window functions for trend analysis
   * Simplifies complex joins between fact tables at different granularities

Design Benefits

* Analytical Flexibility:
  + Easy integration of indicators at different granularities
  + Efficient historical trend analysis
  + Simplified correlation studies between economic indicators
* Performance Optimization:
  + Minimized join complexity through star schema
  + Efficient aggregation through proper indexing
  + Reduced storage through normalized time dimension
* Maintainability:
  + Clear separation of facts and dimensions
  + Easy addition of new economic indicators
  + Consistent handling of temporal relationships

This design effectively supports complex analytical queries while maintaining performance and flexibility for various types of financial and economic analyses.

A diagram of a number

Description automatically generated

Figure :Dimensional model (OLAP)

### Transactional Model (OLTP)

**A transactional model would involve:**

* **HOUSING\_TRANSACTIONS**: Captures raw housing transaction details.
* **LABOR\_SURVEY\_RESPONSES**: Logs individual labor survey results.
* **STOCK\_TRADES**: Tracks stock trades at the transactional level.

**A screenshot of a computer program

Description automatically generated**

The diagram illustrates how the OLAP would feed from the OLTP Tables. Each table record data on a 1 row per record level which would then be aggregated to give quarterly, monthly or summarized tables.

### Analysis Data Relationships and Trends

This section provides some analytical queries along with outputs which have been truncated in some cases (due to volume of data). These queries will show the relations and trends between various economic indicators and the financial data. Later these have been used for visualizations.

#### **Analytical Queries:**

##### **Analysis 1: Market Volatility Study**

Query Purpose:

* Calculates monthly price volatility of S&P 500
* Measures trading volume patterns
* Identifies historically volatile periods

Business Insights:

* Identifies historical stress periods in the market
* Shows clustering of volatile periods (e.g., 2008 crisis)
* Demonstrates how market behavior has evolved over time
* Helps in understanding risk patterns across different market cycles

This analysis provides a foundation for understanding market behavior during different economic conditions, which is crucial for risk assessment and investment strategy development.

A computer code with numbers and symbols

Description automatically generated

A screenshot of a data

Description automatically generated

##### **Analysis 2: Cross-Metric Economic Overview**

Query Purpose:

* Creates a comprehensive quarterly view of all key economic indicators
* Aligns metrics of different granularities (daily, monthly, quarterly)
* Enables multi-factor economic analysis

Value of Analysis:

* Shows interrelationships between economic indicators
* Captures economic cycle progression
* Enables identification of leading/lagging indicators
* Provides context for understanding market behavior during economic transitions

This comprehensive view helps understand how different economic metrics interact during various phases of the economic cycle, particularly useful for both historical analysis and future pattern recognition.

A screenshot of a computer code

Description automatically generated

A screenshot of a graph

Description automatically generated

##### **Analysis 3: Day-of-Week Trading Patterns**

Query Purpose:

* Analyzes trading behavior patterns across different days of the week
* Examines both volume and price movement characteristics
* Identifies potential day-of-week effects in market activity

Business Implications:

* Trading activity concentrates mid-week
* Day-of-week effects visible in both volume and price movements
* Could inform optimal trading day decisions
* Helps in understanding market liquidity patterns

A screenshot of a computer code

Description automatically generated

A screenshot of a computer

Description automatically generated

##### **Analysis 4: Quarterly Growth Comparison (S&P 500 vs GDP)**

Query Purpose:

* Calculates and compares quarterly growth rates for both S&P 500 and GDP
* Uses lag functions to compute period-over-period changes
* Enables direct comparison of market and economic growth patterns

A screenshot of a computer code

Description automatically generated

A screenshot of a computer

Description automatically generated

##### **Analysis 5: S&P 500 and GDP Correlation Study**

Query Purpose:

* Calculates the statistical correlation between S&P 500 prices and lagged GDP values
* Uses sophisticated windowing functions to align time periods
* Measures the strength of relationship between market and economic growth

**A screenshot of a computer code

Description automatically generated**

A screenshot of a computer

Description automatically generated

Key Finding:

* Strong Positive Correlation: 0.923
  + Indicates very strong relationship between market performance and GDP
  + Correlation coefficient close to 1 shows nearly linear relationship
  + Suggests market largely moves in tandem with economic growth

This analysis provides statistical validation of the relationship between economic growth and market performance, supporting evidence-based decision making in investment and policy contexts.

##### **Analysis 6: S&P 500 and Housing Price Index Correlation Study**

Query Purpose:

* Examines relationship between stock market and housing market
* Similar methodology to GDP correlation, but focuses on real estate sector
* Uses lagged values to account for potential delayed effects

A screenshot of a computer code

Description automatically generated

A computer screen shot of a computer error

Description automatically generated

Key Finding:

* Strong Positive Correlation: 0.901
  + Very strong relationship between stock and housing markets
  + Slightly lower than GDP correlation (0.923)
  + Indicates strong wealth effect between markets

This analysis shows how two major asset classes (stocks and real estate) move together, which is crucial for investment strategy and economic policy considerations.

##### **Analysis 7: Housing Market Annual Growth Trends**

Query Purpose:

* Analyzes year-over-year changes in Housing Price Index (HPI)
* Calculates average annual HPI values
* Computes percentage growth rates
* Identifies long-term housing market trends

Business Value:

* Identifies housing market cycles
* Shows periods of rapid appreciation
* Helps understand real estate market dynamics

A text on a white background

Description automatically generated

A screenshot of a graph

Description automatically generated

##### **Analysis 8: Quarterly Housing Market Patterns**

Query Purpose:

* Examines seasonal patterns in housing market
* Averages house prices by quarter across years
* Calculates typical quarterly growth rates
* Identifies seasonal trends in housing market

A computer code with text

Description automatically generated with medium confidence

A screenshot of a computer

Description automatically generated

Key Insights:

* Seasonal Pattern Identified:
  + Strongest growth in Q2 (Spring season)
  + Weakest growth in Q4 (Winter season)
  + Spring prices typically highest
  + More moderate activity in Q3 and Q1

Business Applications:

* Optimal timing for real estate transactions
* Seasonal adjustment of housing market analysis
* Real estate investment strategy planning
* Understanding cyclical market behavior

This analysis helps market participants understand and anticipate regular seasonal patterns in the housing market, useful for both timing decisions and price expectations.

The analysis demonstrates the power of dimensional modeling and OLAP capabilities in understanding complex economic relationships, supporting both investment decisions and policy considerations. The findings provide a solid foundation for understanding market behavior, economic relationships, and timing considerations across different asset classes.

## Data Visualization and Storytelling

##### **The Three Acts of America's Economic Story (1978-2013)**

A graph of different colored lines

Description automatically generated

This visualization presents a synchronized view of four key economic indicators - GDP, Unemployment Rate, Housing Prices, and S&P 500 Adjusted Closing Price. The data naturally divides into three distinct periods, color-coded for clarity:

1. Steady Growth Era (Orange, 1975-1993):

* Stable GDP growth
* Gradually declining unemployment
* Modest housing appreciation
* Conservative market growth

1. Boom Period (Blue, 1994-2007):

* Accelerated economic expansion
* Low unemployment
* Housing market surge
* Strong stock performance

1. Crisis & Recovery (Red, 2008-2013):

* Economic volatility
* High unemployment
* Housing market correction
* Market turbulence

The parallel arrangement of these metrics reveals how they interact during different economic cycles, clearly showing the transition from stability to boom and subsequent correction. There is a powerful illustration of the 2008 financial crisis impact.

##### **Market Returns vs Economic Growth: A Tale of Volatility and Recovery (1974-2012)**

A graph with red and blue lines

Description automatically generated

This dual-axis time series visualization compares quarterly GDP growth (red line) with quarterly S&P 500 returns (blue line), revealing several key insights:

1. Pattern Recognition:

* GDP growth shows gradual decline from high volatility (3.5% in late 1970s) to more stable rates
* S&P 500 returns demonstrate higher volatility, especially during crisis periods
* Most dramatic divergence during 2008 crisis (S&P 500 -10% return)

1. Key Periods:

* 1970s: High but declining GDP growth, moderate market returns
* 1980s-1990s: Stabilizing GDP growth with fluctuating market returns
* 2008: Severe market decline followed by sharp recovery

The visualization effectively shows how market returns can be more volatile than underlying economic growth, with the S&P 500 often experiencing more extreme movements than GDP, particularly during economic stress periods.

##### **Parallel Paths: The Long-Term Dance of Markets and Economic Growth (1974-2012)**

A graph showing the growth of a stock market

Description automatically generated

This stacked visualization powerfully illustrates the relationship between S&P 500 prices (top) and GDP (bottom) over four decades:

1. Growth Patterns:

* GDP shows steady, consistent upward trajectory
* S&P 500 demonstrates more dramatic growth phases
  + Relatively flat until early 1990s
  + Steep acceleration in late 1990s
  + Notable peaks and valleys post-2000

1. Key Transitions:

* Pre-1990: Both metrics show modest, steady growth
* 1995-2000: Market acceleration outpaces GDP growth
* 2000-2012: Market volatility despite continued GDP growth
  + Dot-com bubble (2000-2002)
  + Financial crisis (2008-2009)
  + Post-crisis recovery

These visuals complement the previous growth rate comparison by showing absolute values, reinforcing our correlation findings (0.923) and highlighting how market values can swing dramatically even as GDP maintains steady growth.

##### **Twin Bubbles: Housing and Stock Market Trajectories (1975-2011)**

A graph showing a line

Description automatically generated

This parallel visualization reveals the intricate relationship between the S&P 500 and Housing Price Index (HPI), highlighting several critical periods:

1. Pre-1995: Stable Growth

* Both markets show steady but modest appreciation
* Housing prices demonstrate more consistent growth
* Stock market maintains relatively flat trajectory

1. 1995-2006: Dual Boom

* Stock market surges dramatically
* Housing prices begin accelerating
* Both markets show unprecedented growth

1. 2007-2011: Crisis and Recovery

* Sharp stock market corrections (2008)
* Housing prices plateau and decline
* Stock market shows faster recovery than housing

The visualization supports our high correlation finding (0.901) while revealing how both markets experienced similar boom-bust cycles, though with different recovery patterns. Housing prices show more "stickiness" during downturns compared to the more volatile stock market.

##### **Housing Market Growth vs S&P 500: A Tale of Two Metrics (1974-2012)**

A graph with a line going up

Description automatically generated

This overlay visualization combines:

* S&P 500 values (blue area chart)
* Housing market yearly growth rates (orange line)

Key Insights:

1. Early Volatility (1976-1980):

* Highest housing growth rates (~14%)
* Relatively modest S&P 500 values
* Period of high inflation

1. Mid-Period Stability (1985-2000):

* Moderate housing growth (4-6%)
* Steady S&P 500 appreciation
* More balanced growth period

1. Crisis Period (2006-2010):

* Dramatic housing decline (-5%)
* S&P 500 maintains value better
* First negative housing growth in decades

This visualization effectively shows how housing growth rates don't always mirror stock market performance, with housing showing more extreme reactions during crisis periods despite the overall high correlation (0.901) between the two markets.

# Conclusion

Our visual and analytical exploration of financial markets and economic indicators from 1974-2013 reveals several key insights:

1. Market-Economic Relationships

* Strong correlation between S&P 500 and GDP (0.923)
* High correlation between housing and stock markets (0.901)
* Markets show higher volatility than underlying economic metrics

1. Distinct Economic Periods

* 1975-1993: Steady growth with high but declining volatility
* 1994-2007: Accelerated growth across all metrics
* 2008-2013: Crisis and recovery phase

1. Critical Patterns

* Stock market demonstrates faster recovery from shocks
* Housing market shows more "sticky" behavior during downturns
* GDP maintains steadier growth despite market fluctuations
* Day-of-week effects in trading patterns

This analysis provides valuable insights for:

* Investment strategy development
* Risk assessment and management
* Understanding economic cycles
* Policy impact evaluation

The data demonstrates how different economic indicators interact and influence each other, highlighting the interconnected nature of financial markets and the broader economy.

##### **Project Conclusion:**

This project demonstrates the power of proper data warehousing and analytical techniques in understanding complex economic relationships, providing valuable insights for both investment decisions and policy considerations.