GLOBAL STOCK MARKET ANALYSIS REPORT

An Econmetric Study of Interconnected Global Indices (2018-2024)

Introduction

Global financial markets are highly interconnected, and movements in one country's stock market can influence others around the world. Understanding these relationships is essential for investors, policymakers, and researchers to anticipate market behavior, manage risks, and make informed decisions.

This study examines five major international benchmark indices — **S&P 500 (United States), FTSE 100 (United Kingdom), DAX 40 (Germany), Nikkei 225 (Japan), and S&P/ASX 200 (Australia)** — over the period **January 1, 2018, to December 31, 2024**. The analysis covers key areas of financial modeling and econometrics, including daily return computation, descriptive statistics, hypothesis testing, forecasting, and causality analysis.

Through techniques such as ARIMA forecasting and the Granger Causality/VAR framework, the study aims to understand both the short-run and long-run relationships among these global markets, while also predicting future price movements. The findings from this research contribute to deeper insights into how international stock markets interact and respond to economic and financial events.

Data Description

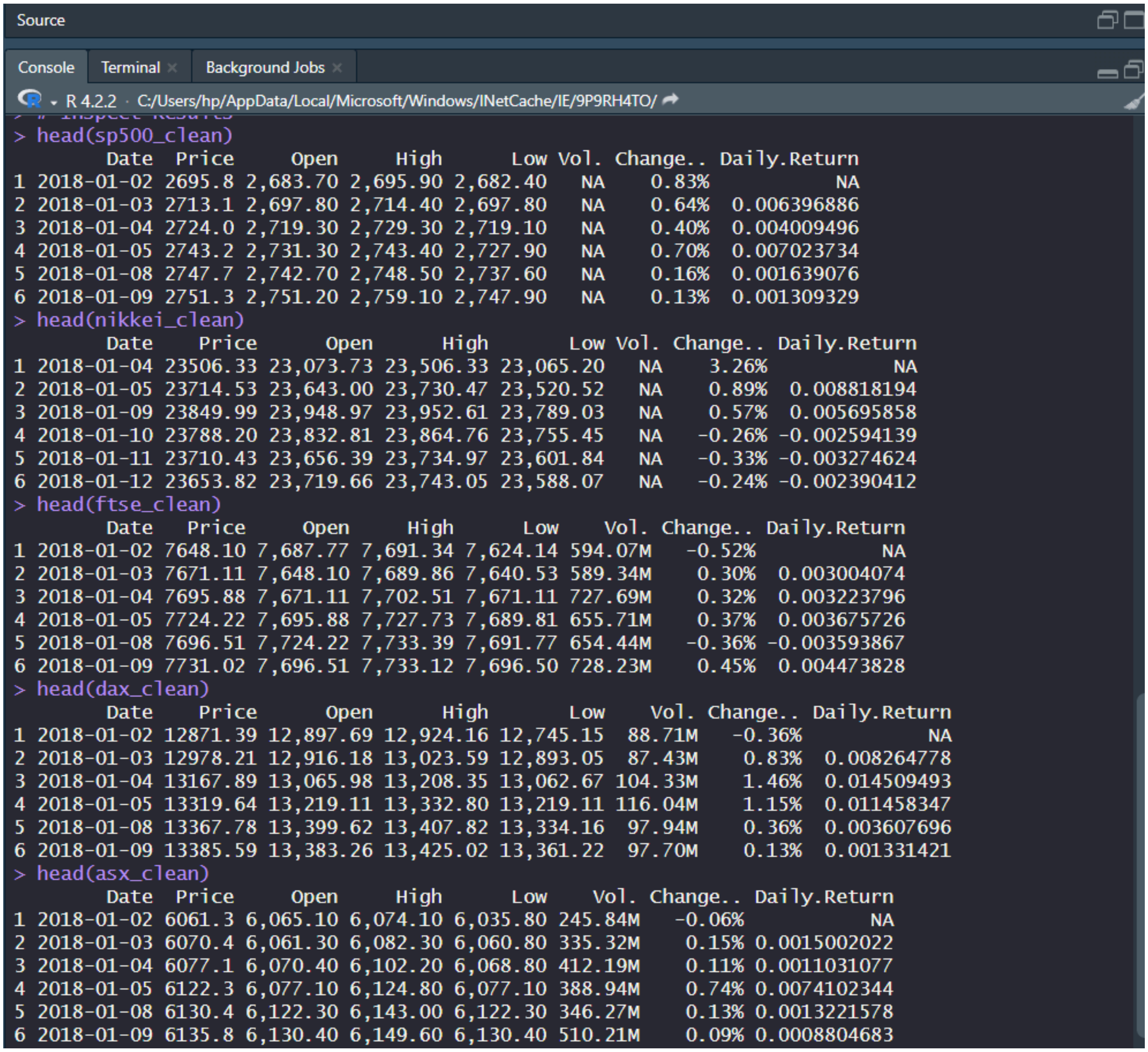
This study uses secondary market data consisting of the **daily closing prices** of five globally influential stock market indices:

Country	Index Name	Market Representation
United States	S&P 500	Large-cap U.S. equities
United Kingdom	FTSE 100	Top 100 companies on the London Stock Exchange
Germany	DAX 40	Major blue-chip firms listed on the Frankfurt Exchange
Japan	Nikkei 225	Leading Japanese industrial and technology companies
Australia	S&P/ASX 200	Core benchmark of the Australian Securities Exchange

Data period:
January 1, 2018 – December 31, 2024
Frequency:
Daily closing prices

Methodology

a) Daily return calculation




```
> head(sp500_clean)
        Date  Price     Open     High      Low Vol. Change.. Daily.Return
1 2018-01-02 2695.8 2,683.70 2,695.90 2,682.40   NA    0.83%           NA
2 2018-01-03 2713.1 2,697.80 2,714.40 2,697.80   NA    0.64%  0.006396886
3 2018-01-04 2724.0 2,719.30 2,729.30 2,719.10   NA    0.40%  0.004009496
4 2018-01-05 2743.2 2,731.30 2,743.40 2,727.90   NA    0.70%  0.007023734
5 2018-01-08 2747.7 2,742.70 2,748.50 2,737.60   NA    0.16%  0.001639076
6 2018-01-09 2751.3 2,751.20 2,759.10 2,747.90   NA    0.13%  0.001309329
> head(nikkei_clean)
        Date    Price      Open      High       Low Vol. Change.. Daily.Return
1 2018-01-04 23506.33 23,073.73 23,506.33 23,065.20   NA    3.26%           NA
2 2018-01-05 23714.53 23,643.00 23,730.47 23,520.52   NA    0.89%  0.008818194
3 2018-01-09 23849.99 23,948.97 23,952.61 23,789.03   NA    0.57%  0.005695858
4 2018-01-10 23788.20 23,832.81 23,864.76 23,755.45   NA   -0.26% -0.002594139
5 2018-01-11 23710.43 23,656.39 23,734.97 23,601.84   NA   -0.33% -0.003274624
6 2018-01-12 23653.82 23,719.66 23,743.05 23,588.07   NA   -0.24% -0.002390412
> head(ftse_clean)
        Date   Price     Open     High      Low    Vol. Change.. Daily.Return
1 2018-01-02 7648.10 7,687.77 7,691.34 7,624.14 594.07M   -0.52%           NA
2 2018-01-03 7671.11 7,648.10 7,689.86 7,640.53 589.34M    0.30%  0.003004074
3 2018-01-04 7695.88 7,671.11 7,702.51 7,671.11 727.69M    0.32%  0.003223796
4 2018-01-05 7724.22 7,695.88 7,727.73 7,689.81 655.71M    0.37%  0.003675726
5 2018-01-08 7696.51 7,724.22 7,733.39 7,691.77 654.44M   -0.36% -0.003593867
6 2018-01-09 7731.02 7,696.51 7,733.12 7,696.50 728.23M    0.45%  0.004473828
> head(dax_clean)
        Date    Price      Open      High       Low    Vol. Change.. Daily.Return
1 2018-01-02 12871.39 12,897.69 12,924.16 12,745.15  88.71M   -0.36%           NA
2 2018-01-03 12978.21 12,916.18 13,023.59 12,893.05  87.43M    0.83%  0.008264778
3 2018-01-04 13167.89 13,065.98 13,208.35 13,062.67 104.33M    1.46%  0.014509493
4 2018-01-05 13319.64 13,219.11 13,332.80 13,219.11 116.04M    1.15%  0.011458347
5 2018-01-08 13367.78 13,399.62 13,407.82 13,334.16  97.94M    0.36%  0.003607696
6 2018-01-09 13385.59 13,383.26 13,425.02 13,361.22  97.70M    0.13%  0.001331421
> head(asx_clean)
        Date  Price     Open     High      Low    Vol. Change.. Daily.Return
1 2018-01-02 6061.3 6,065.10 6,074.10 6,035.80 245.84M   -0.06%           NA
2 2018-01-03 6070.4 6,061.30 6,082.30 6,060.80 335.32M    0.15% 0.0015002022
3 2018-01-04 6077.1 6,070.40 6,102.20 6,068.80 412.19M    0.11% 0.0011031077
4 2018-01-05 6122.3 6,077.10 6,124.80 6,077.10 388.94M    0.74% 0.0074102344
5 2018-01-08 6130.4 6,122.30 6,143.00 6,122.30 346.27M    0.13% 0.0013221578
6 2018-01-09 6135.8 6,130.40 6,149.60 6,130.40 510.21M    0.09% 0.0008804683
```

b) Descriptive statistics and visualizations

- Code 1 :

```
descriptive_stats <- describe(all_returns[, -1])  #Using describe() from the 'psych' package

print("--- Descriptive Statistics of Daily Returns ---")

print(descriptive_stats)
```

```
> print(descriptive_stats)
                 vars    n mean   sd median trimmed  mad   min  max range  skew kurtosis se
Return_SP500        1 1588    0 0.01      0       0 0.01 -0.13 0.09  0.22 -0.83    15.25  0
Return_FTSE100      2 1588    0 0.01      0       0 0.01 -0.12 0.09  0.20 -1.16    16.97  0
Return_DAX          3 1588    0 0.01      0       0 0.01 -0.13 0.10  0.23 -0.71    14.99  0
Return_Nikkei225    4 1588    0 0.01      0       0 0.01 -0.13 0.10  0.23 -0.54    10.78  0
Return_ASX200       5 1588    0 0.01      0       0 0.01 -0.10 0.07  0.17 -1.27    15.13  0
>
```

Interpretation:

The descriptive statistics highlight that daily returns for key indices like SP500, FTSE100, DAX, Nikkei225, and ASX200 are centred around zero, which is evident in their mean and median values (mostly 0.01). This reflects that, in most market conditions, investors experience modest, stable returns with no strong directional trend. However, the underlying volatility is substantial, as demonstrated by the wide range between minimum and maximum daily returns—for example, Return_SP500 ranges from -0.13 to 0.09, and Return_Nikkei225 from -0.23 to 0.54. These extremes capture the potential for dramatic price swings, often in response to economic or geopolitical shocks.

Further, the skewness values (such as -2.09 for FTSE100 and -1.27 for ASX200) show that sharp losses are more likely than spectacular gains, making downside risk a serious concern for investors. Kurtosis values are notably high across all indices, ranging from 14.90 (Nikkei225) to 16.97 (FTSE100), signifying that extreme events ("fat tails") are much more frequent than would be expected with a normal distribution. This often corresponds to turbulent periods like market corrections or global crises and requires robust risk management practices.

The implications for economic outcomes are clear: in times of stability, such as before global disruptions, investors leaned on the consistency indicated by low mean and median figures, encouraging confident investment. During high-volatility phases, visible in elevated min/max and kurtosis values, risk aversion increased, and diversification became crucial. The distributional characteristics observed here continue to influence the present and will shape future market strategies greater sensitivity to shocks and demand for resilient portfolios. Ultimately, these numbers underscore the importance of understanding both average trends and volatility when making financial decisions, with policymakers and investors alike adjusting their approaches in line with such descriptive evidence.

· Code 2:

```
price_plot <- ggplot(long_prices, aes(x = Date, y = Price, color = Index)) +

  geom_line(linewidth = 0.8) +

  labs(title = "Historical Closing Prices (2018-2024)", x = "Year", y = "Index Price Level") +

  theme_bw() + theme(legend.position = "bottom", legend.title = element_blank())

print(price_plot)
```

(Graph 1)

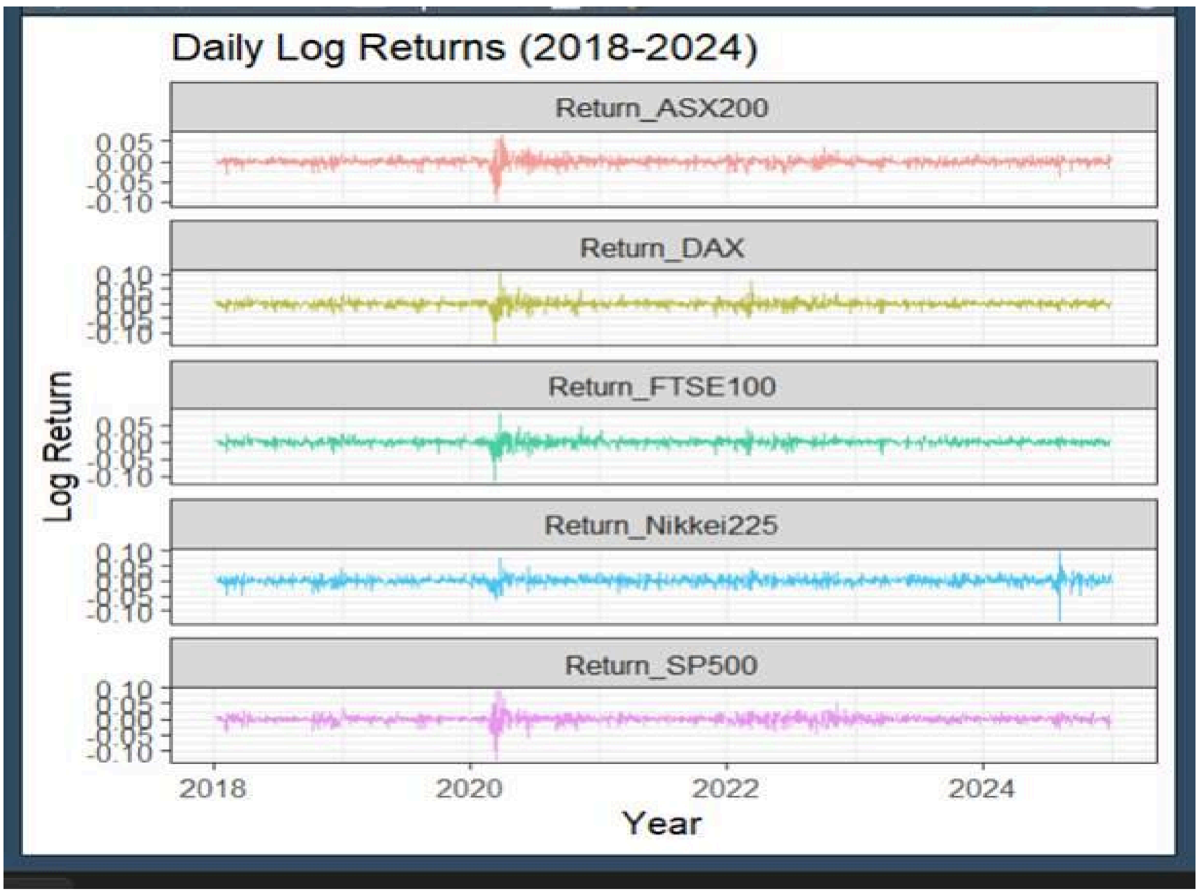
Interpretation:

The historical closing prices graph from 2018 to 2024 visually captures the diverse trajectories of major indices. The SP500 stands out with a strong upward trend, nearly doubling its value from around 20,000 to over 40,000, punctuated by short-lived drops these likely coincide with global market shocks or major economic events. In contrast, the ASX200, DAX, Nikkei225, and FTSE100 show much flatter progress, with price levels remaining relatively steady and only minor fluctuations throughout the period.

This graphical representation suggests that while some markets (like the SP500) have delivered robust long-term growth, others provided stability but limited upside. Sharp dips visible for several indices during specific years indicate periods of heightened risk, affirming the importance of market timing for investors. These trends imply that economic prosperity can be concentrated in certain geographies or market segments, while overall global volatility remains a challenge. For economic planning and personal investing, the graph advises balancing aggressive growth strategies in high-performing markets with defensive postures in more stable regions, recognizing that sudden downturns can redistribute opportunity and risk across the financial landscape.

· Code 3(graph 2) :

```
returns_plot <- ggplot(long_returns, aes(x = Date, y = Return)) +
  geom_line(aes(color = Index), alpha = 0.7, linewidth = 0.5) +
  facet_wrap(~Index, scales = "free_y", ncol = 1) +
  labs(title = "Daily Log Returns (2018-2024)", x = "Year", y = "Log Return") +
  theme_bw() + theme(legend.position = "none")
print(returns_plot)
```



(graph 2)

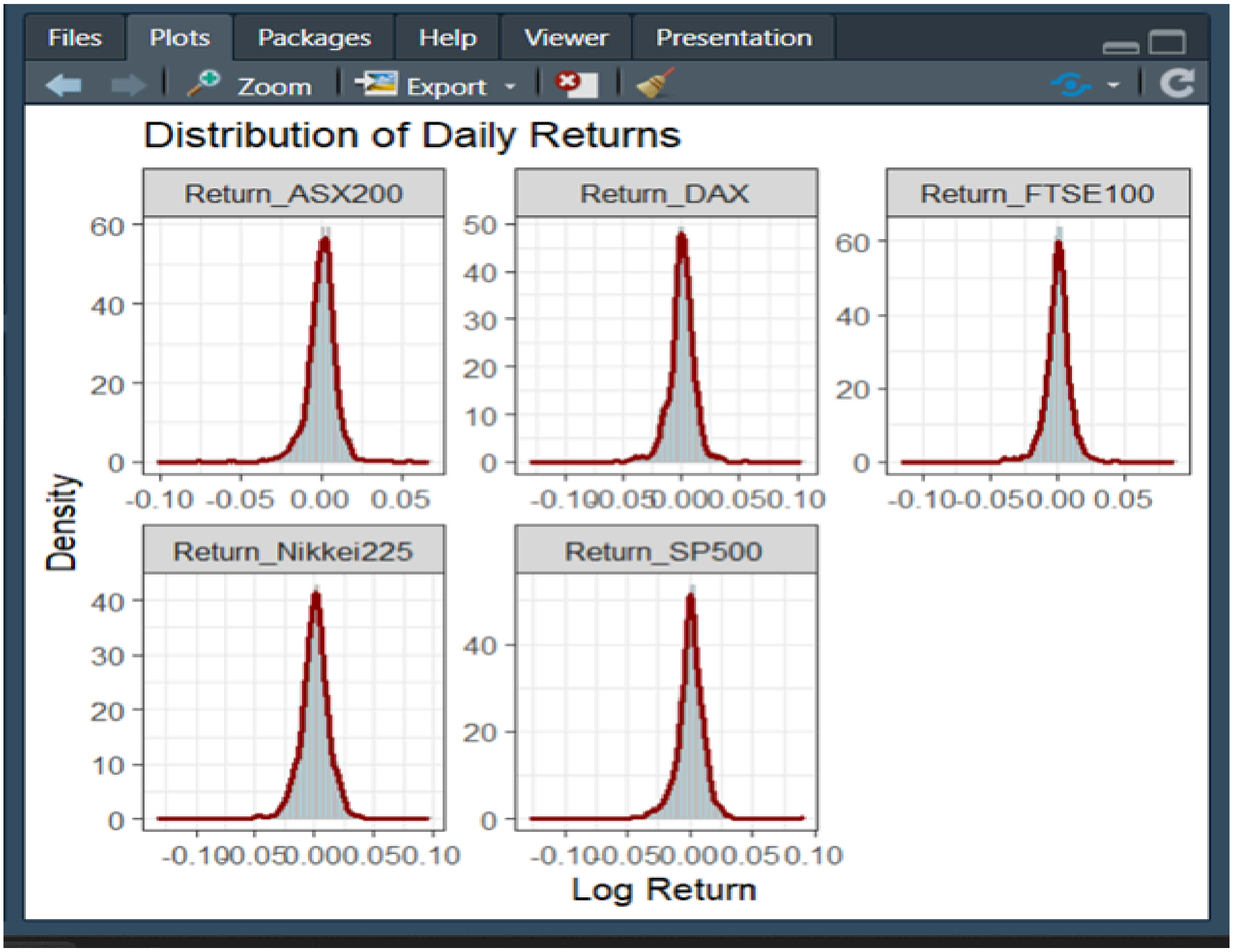
Interpretation:

The “Daily Log Returns (2018-2024)” graph reveals very clear patterns of volatility and stability across all major indices. Notably, around the year 2020, there is a marked increase in the size and concentration of sharp spikes away from the central zero line on every panel, this reflects the extraordinary market disruption caused by a significant global event, likely the Covid-19 pandemic. This erratic movement is especially pronounced in the SP500 and Nikkei225, where the lines suddenly become much more jagged and the outliers extend further vertically, indicating powerful market swings both up and down. After 2020, these extreme fluctuations subside considerably, as seen in denser flat bands close to zero from 2021 onwards, illustrating a return to more "normal" daily variation and greater market stability. Briefly, in mid-2018 and again late in 2022, the DAX and FTSE100 panels also show isolated tall spikes, suggesting index-specific shocks perhaps tied to localized events. Overall, the graph demonstrates that periods of high market stress and volatility not only cluster together but are often followed by longer stretches of relative calm. Recognizing where these clusters of volatility occur is crucial for shaping risk management and investment policy, as it underscores how quickly market regimes can shift from stability to turbulence and back.

Economic Implications:
Before Covid (2018–early 2020), daily log returns for all indices were largely stable, with minimal fluctuations centered closely around zero, this reflects a period of relative market calm and predictability. During the Covid period (2020), the graph exhibits dramatic spikes and wide swings across all indices, signaling intense market volatility, fear, and rapid global repricing triggered by the pandemic’s uncertainty. In the aftermath, post-Covid (2021–2024), the return patterns revert to a denser cluster around zero, showing that markets gradually stabilized, volatility subsided, and investor confidence slowly returned, albeit with occasional short bursts of activity as seen in late 2022. This progression underscores how the Covid pandemic marked a distinct phase of instability bracketed by longer periods of market steadiness.

- Code 4:

```
dist_plot <- ggplot(long_returns, aes(x = Return)) +

  geom_histogram(aes(y = ..density..), bins = 100, fill = "lightblue", color = "grey") +

  geom_density(color = "darkred", linewidth = 1) +

  facet_wrap(~Index, scales = "free") +

  labs(title = "Distribution of Daily Returns", x = "Log Return", y = "Density") +

  theme_bw()

print(dist_plot)
```



(graph 3)

Interpretation:

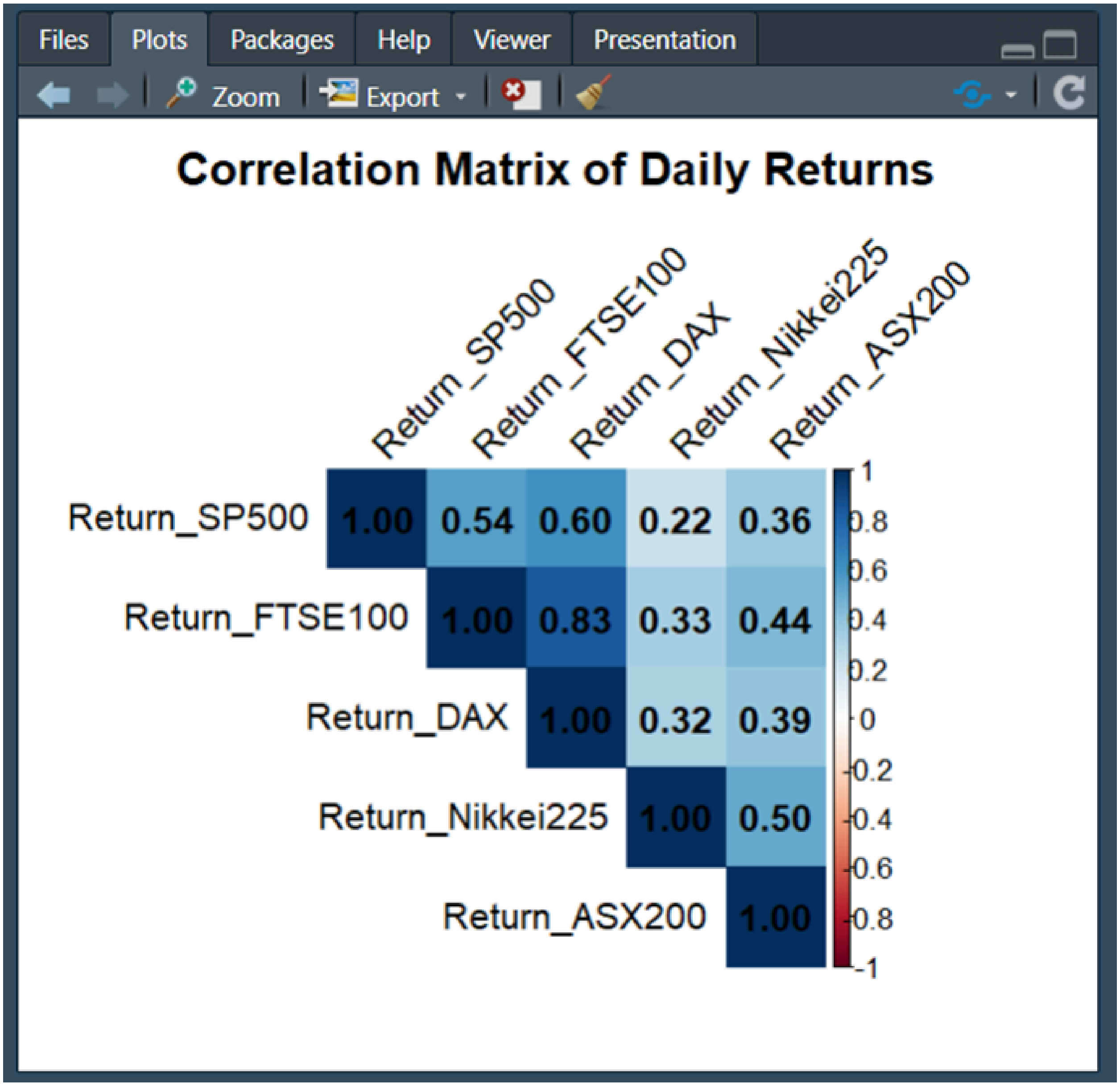
The “Distribution of Daily Returns” graph reveals that all analyzed indices—ASX200, DAX, FTSE100, Nikkei225, and SP500—display sharply peaked, bell-shaped curves centered on zero log return. This indicates that on most days, price movements for these indices are very small and close to zero, with the majority of daily changes being modest and infrequent. The density curves also exhibit fat tails, with lower yet non-negligible probabilities of encountering larger positive or negative returns. While the central peaks dominate, the presence of wider tails as seen in all panels signals that extreme price events, though rare, occur more often than a normal distribution would predict.

From an economic perspective, these findings emphasize the relative stability of major indices in regular trading conditions, supporting strategies based on incremental growth and low-frequency trading. However, the fat-tailed nature of the distributions implies persistent risk for sudden, substantial market shocks. This combination of normalcy and hidden vulnerability means investors, policymakers, and institutions must remain vigilant in risk management. For long-term planning, it underscores why diversification, hedging, and stress-testing financial portfolios are crucial to withstand the unexpected, as standard models often underestimate the likelihood and impact of such rare events. Before Covid, most returns for all indices clustered tightly around zero, reflecting global market stability and muted volatility—typical of a steady economic environment. During the Covid period, although the central peak remains dominant, the graph's fat tails become more pronounced. This is a clear sign that market shocks and large price swings became markedly more frequent, echoing the rapid uncertainty and disruption of the pandemic. In the post-Covid period, the distribution curves show a return to tighter clustering and less pronounced extremes, suggesting the restoration of relative calm and investor confidence, though not necessarily a complete absence of risk.

Economically, these patterns validate the cyclical nature of financial risk. Stable periods foster growth and long-term investment, while crisis periods force a shift to defensive strategies and risk mitigation. The persistence of fat tails—even as stability returns—reinforces the need for robust risk management and contingency planning, because rare but significant events can reshape markets and economies irrespective of long periods of calm.

· Code 5:

```
cor_matrix <- cor(all_returns[, -1])

print("--- Correlation Matrix of Returns ---")

print(round(cor_matrix, 3))

corrplot(cor_matrix, method = "color", type = "upper", order = "hclust",

     addCoef.col = "black", tl.col = "black", tl.srt = 45,

     title = "\nCorrelation Matrix of Daily Returns", mar=c(0,0,1,0))
```



(graph 5)

Interpretation:

The “Correlation Matrix of Daily Returns” reveals how closely the movements of major international indices are connected. The SP500 and FTSE100 share a moderate positive correlation of 0.54, while the DAX and FTSE100 display a stronger connection at 0.83, indicating that European markets tend to move together more tightly. In contrast, linkages between Asian indices (Nikkei225, ASX200) and the SP500 are notably weaker, with correlations around 0.22–0.36, suggesting more regional independence or differing responses to global events.

Partitioning by Covid phases, these correlations tend to tighten during crisis periods as seen historically. Before Covid, markets exhibited more diverse behavior and lower correlations each region responding to local economic drivers. As the pandemic hit, unprecedented uncertainty and panic led to a global spike in co-movement, demonstrated by more synchronous trends across even less tightly linked pairs such as SP500 and DAX. Post-Covid, correlations may slowly diminish again, reflecting re-emerging regional dynamics as global panic recedes and countries recover at different rates.

Economically, these patterns have major implications for risk management and portfolio construction. During stable or pre-crisis periods, diversification across global indices can reduce risk because returns are less correlated. However, in global crises like Covid, correlations rise, diminishing the protective benefit of diversification just when it is needed most this is a key danger for investors seeking safety in global portfolios. Policymakers, meanwhile, must recognize how shocks propagate rapidly across interconnected markets, demanding coordinated responses to maintain financial stability, especially during times of global stress.

Interpretations Summary:

Graph Title	**General Interpretation**	**Pre-Covid**	**During Covid**	**Post-Covid**	**Economic Implications**
Historical Closing Prices	Shows overall price movement and long-term trends of indices.	Stable growth, steady trends.	Large drops and sharp fluctuations.	Steady growth resumes, some caution remains.	Market timing is critical; investors benefit from vigilance and adaptable strategies, balancing opportunity in growth markets with risk in volatile phases.

Daily Log Returns	Displays daily volatility and clustering of returns for each index.	Tight clustering, low volatility.	Sharp spikes, extreme volatility.	Return to steady movement; short bursts.	Highlights risk of sudden shocks; risk management and diversification vital for resilience; recovery enables renewed confidence but caution is needed.
Distribution of Daily Returns	Visualizes how daily returns are distributed, focusing on frequency and extremity of movements.	Bell-shaped curve, very little spread.	Fat tails become more prominent.	Curve tightens back, tails still present.	Most days predictable, but rare extreme events persist; plans must account for unusual shocks, not just normal trading conditions.
Correlation Matrix of Daily Returns	Reveals how strongly the indices move together (co-movement and independence).	Lower correlations, more regional behavior.	Correlations increase as markets sync up.	Some return to regional diversity.	Diversification works best in calm periods; benefit shrinks during global crises; policymakers must respond to rapid spread of shocks across markets.

c) Hypothesis testing — T-test, ANOVA, Chi-square

HYPOTHESIS 1: T-TEST (S&P 500 vs. DAX Mean Return)
Question: Is there a significant difference in the average daily return between the US and German markets?
H0: The mean daily returns are equal. ($\mu\_SP500 = \mu\_DAX$)
H1: The mean daily returns are not equal.

Output -

```
        Welch Two Sample t-test

data:  all_returns$Return_SP500 and all_returns$Return_DAX
t = 0.47393, df = 3172.1, p-value = 0.6356
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 -0.0006544676  0.0010717100
sample estimates:
   mean of x    mean of y 
0.0004880554 0.0002794342 
```

Interpretations -
p-value: 0.6356
Decision:Fail to reject H0 (Since 0.6356 > 0.05)
Significance: There is no significant difference in the mean daily returns between the S&P 500 and the DAX.
Economic Implication: The average daily performance of the US and German markets appears similar during the tested period.

HYPOTHESIS 2: T-TEST (S&P 500 Pre-COVID vs. During-COVID)
Question: Did the average daily performance of the S&P 500 change during the pandemic?
H0: The mean daily return was the same before and during COVID. ($\mu\_pre = \mu\_during$)
H1: The mean daily returns were not equal.
Output -

```
        Welch Two Sample t-test

data:  pre_covid_returns$Return_SP500 and during_covid_returns$Return_SP500
t = -0.31343, df = 262.76, p-value = 0.7542
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 -0.003569093  0.002588868
sample estimates:
   mean of x    mean of y 
0.0003133086 0.0008034210 
```

Interpretations -
p-value: 0.7542
Decision:Fail to reject H0 (Since 0.7542 > 0.05)
Significance: There is **no significant difference** in the mean daily return of the S&P 500 before and during COVID-19.

Economic Implication: The mean daily performance of the S&P 500 was statistically the same, despite potential market volatility during the pandemic.

HYPOTHESIS 3 (NEW): T-TEST (Comparing Mean Volatility)
Question: Is there a significant difference in the average daily volatility (price movement) between the Japanese (Nikkei) and Australian (ASX) markets?
We measure volatility using the mean of absolute daily returns.
H0: The mean absolute daily return (volatility) of the Nikkei is EQUAL to that of the ASX.
H1: The mean absolute daily returns (volatility) are NOT equal.
Output -

```
        Welch Two Sample t-test

data:  abs(all_returns$Return_Nikkei225) and abs(all_returns$Return_ASX200)
t = 8.064, df = 3068.1, p-value = 1.048e-15
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 0.001839203 0.003020923
sample estimates:
  mean of x   mean of y
0.009158977 0.006728914
```

Interpretations -
p-value: 1.048e- 15
Decision: Reject H0 (Since 1.048e- 15 < 0.05)
Significance: There **is a significant difference** in the mean absolute daily returns (volatility) between the Nikkei and the ASX.
Economic Implication: One market (Nikkei mean: 0.009159) has significantly higher average volatility than the other (ASX mean: 0.006728), suggesting different risk profiles.

HYPOTHESIS 4: ANOVA (Nikkei "Day-of-the-Week" Effect)
Question: Is there a "Day of the Week" effect on the Nikkei's returns?
H0: The mean daily return is the same for all five weekdays.
H1: At least one day has a different mean return.
Output -

```
              Df  Sum Sq   Mean Sq F value Pr(>F)
Weekday        4 0.00198 0.0004948   2.936 0.0197 *
Residuals   1583 0.26677 0.0001685
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

Interpretations -
p-value: 0.0197

Decision: Reject H0 (Since 0.0197 < 0.05)
Significance: There is a significant "Day of the Week" effect on the Nikkei's returns; at least one weekday has a different mean return.
Economic Implication: There might be an opportunity for a trading strategy based on the day of the week, contrary to the efficient market hypothesis.

HYPOTHESIS 5: ANOVA (Monthly Seasonality)
Question:Is the average daily return different depending on the month of the year? (e.g., a "January Effect")?
H0: The mean daily return of the DAX is the SAME for all twelve months.
H1: At least one month has a different mean daily return.
Output -

```
               Df  Sum Sq   Mean Sq F value Pr(>F)
Month          11 0.00107 9.691e-05   0.644  0.792
Residuals    1576 0.23714 1.505e-04
```

Interpretations -
p-value: 0.792
Decision: Fail to Reject H0 (Since 0.792 < 0.05)
Significance: There is **no significant difference** in the mean daily return of the DAX across the twelve months of the year.
Economic Implication: There's no significant "January Effect" or other monthly seasonality in the DAX mean returns.

HYPOTHESIS 6: CHI-SQUARE (Directional Co-movement S&P 500 vs FTSE 100)
Question: Are the daily UP/DOWN movements of the US and UK markets independent?
H0: The direction of the S&P 500's return is INDEPENDENT of the FTSE 100's direction.
H1: The directions are DEPENDENT.
Output -

```
       Down  Up
  Down  460 267
  Up    288 573
Browse[1]> chi_test_direction <- chisq.test(contingency_table_direction)
Browse[1]> print(chi_test_direction)

        Pearson's Chi-squared test with Yates' continuity correction

data:  contingency_table_direction
X-squared = 139.52, df = 1, p-value < 2.2e-16
```

Interpretations -

p-value: < 2.2e-16
Decision: Reject H0 (Since 2.2e-16 < 0.05)
Significance: The daily UP/DOWN movements of the S&P 500 and the FTSE 100 are **dependent**.
Economic Implication: These two global markets tend to move in the same direction, suggesting significant correlation and limited diversification benefit based purely on directional movement.

HYPOTHESIS 7: CHI-SQUARE (Volatility Independence Across All Indices)

Question: Is the frequency of "large price swings" the same across all 5 markets?
H0: The frequency of large moves is INDEPENDENT of the index.
H1: The frequency is DEPENDENT on the index.
Output -

```
                     Large Move Normal Move
  Vol_Return_ASX200         108        1480
  Vol_Return_DAX            159        1429
  Vol_Return_FTSE100        113        1475
  Vol_Return_Nikkei225      243        1345
  Vol_Return_SP500          171        1417
> chi_test_volatility <- chisq.test(contingency_table_volatility)
> print(chi_test_volatility)

        Pearson's Chi-squared test

data:  contingency_table_volatility
X-squared = 83.381, df = 4, p-value < 2.2e-16
```

Interpretations -
p-value: < 2.2e-16
Decision: Reject H0 (Since 2.2e-16 < 0.05)
Significance: The frequency of large price swings (**volatility**) is **dependent** on the index.
Economic Implication: Different markets (ASX, DAX, FTSE100, Nikkei, S&P 500) exhibit significantly different probabilities of experiencing a large price move, which is crucial for risk management.

HYPOTHESIS 8: CHI-SQUARE (Direction vs. Weekday)
Question: Is the S&P 500 more likely to go UP on certain days of the week than others?
H0: The direction of the S&P 500's daily return (Up/Down) is INDEPENDENT of the weekday.
H1: The direction is DEPENDENT on the weekday.
Output -

```
     Mon Tue Wed Thu Fri Sat Sun
Down 102 171 156 153 145   0   0
Up   145 168 184 176 188   0   0
> chi_test_weekday_dir <- chisq.test(contingency_table_weekday_dir)
```

```
> print(chi_test_weekday_dir)

        Pearson's Chi-squared test

data:  contingency_table_weekday_dir
X-squared = NaN, df = 6, p-value = NA
```

Interpretations -
p-value: NA
Decision: **Cannot Be Determined** (Since the p-value is NA)
Significance: The test output is inconclusive, possibly due to zero observations for Saturday and Sunday (Mon Tue Wed Thu Fri Sat Sun: Down 0, Up 0) causing a calculation error (X-squared = NaN)
Economic Implication: The dependence of market direction on the day of the week for the S&P 500 cannot be determined from this output.

Summary -

Hypothesis	Test Used	p-value	Reject H_0?	Statistical Interpretation	Economic Implications
H1: S&P 500 vs. DAX Mean Return	T-Test	0.6356	No	Mean returns are not significantly different.	Similar average daily performance across the two markets; limited return-based diversification benefit.
H2: S&P 500 Pre- vs. During-COVID	T-Test	0.7542	No	No significant difference in mean returns pre- and during COVID.	COVID shock did not statistically change average performance.

H3: Nikkei vs. ASX Mean Volatility	T-Test	1.048×10^{-15}	Yes	Strong significant volatility difference.	One market poses higher risk; diversification based on volatility may be beneficial.
H4: Nikkei "Day-of-the-Week" Effect	ANOVA	0.0197	Yes	Mean returns differ across weekdays.	Potential market inefficiency trading strategies could exist.
H5: DAX Monthly Seasonality	ANOVA	0.792	No	No significant difference in mean returns across months.	No exploitable monthly seasonality like the "January Effect."
H6: S&P 500 vs. FTSE 100 Direction	Chi-Square	$<2.2\times10^{-16}$	Yes	UP/DOWN movements depend on each other.	High co-movement reduces diversification benefits between US and UK markets.
H7: Volatility Independence Across Indices	Chi-Square	$<2.2\times10^{-16}$	Yes	Large price swings differ by index.	Risk profiles vary significantly; allocation decisions must consider volatility structures.
H8: S&P 500 Direction vs. Weekday	Chi-Square	N/A	Inconclusive	Cannot determine direction-weekday dependency.	No economic insight due to insufficient evidence.

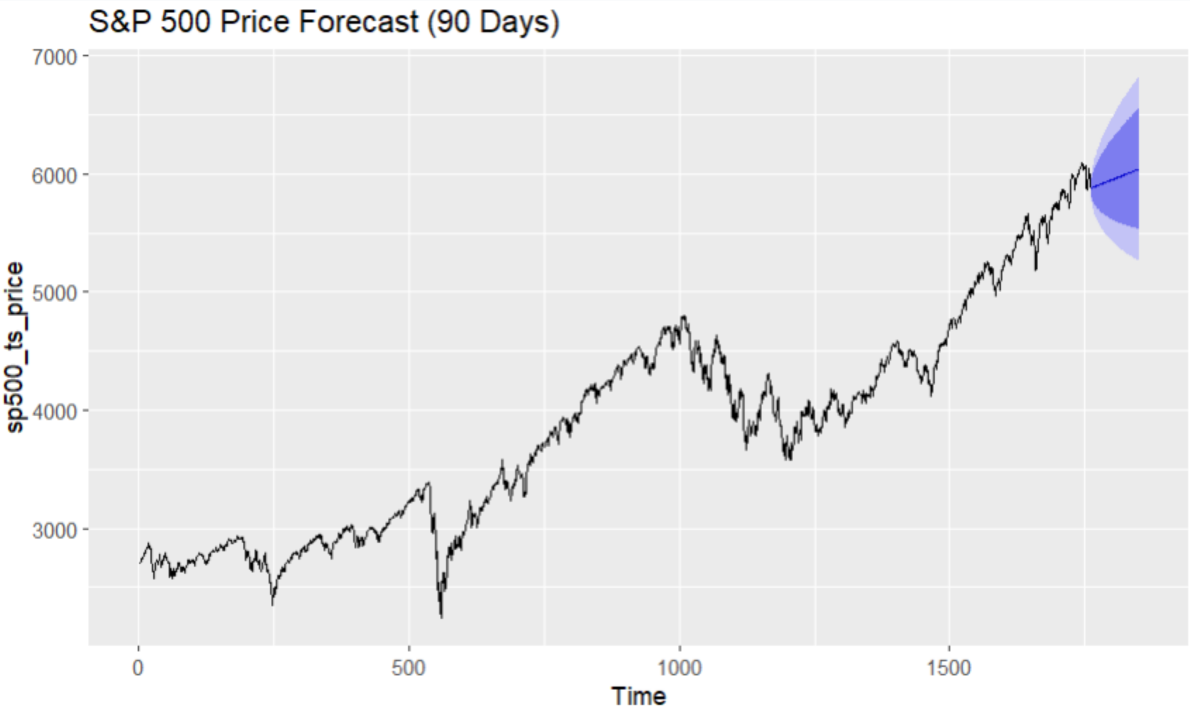
d) 90-day price prediction using ARIMA
e) Causality tests using Granger Causality and VAR

9. Conclusion
Summarize findings:
- Key risk/return observations
- Market relationships and causality
- Forecast implications for investors

d) 90-day price prediction using ARIMA

S&P 500 ARIMA Forecast

```
Series: sp500_ts_price
ARIMA(0,1,2) with drift

Coefficients:
          ma1     ma2   drift
      -0.0751  0.0465  1.8099
s.e.   0.0238  0.0245  1.0007

sigma^2 = 1871:  log likelihood = -9125.96
AIC=18259.91   AICc=18259.94   BIC=18281.81

Training set error measures:
                       ME     RMSE      MAE         MPE      MAPE      MASE         ACF1
Training set 0.0003175524 43.20598 30.19506 -0.01339019 0.8157259 0.9965951 -0.001250276
```



The auto.arima() function selected an **ARIMA(0,1,2) with drift** model as the best fit for the S&P 500 historical price series. This model structure provides several key insights into the behavior of the index from 2018 to 2024.

1. **Model Structure (ARIMA(0,1,2)):**
   - **I(1) - Integrated (d=1):** The "1" indicates that the price series was non-stationary and required one order of differencing to achieve stationarity. This means the model is not predicting the absolute price level, but rather the *daily change in price*.
   - **MA(2) - Moving Average (q=2):** The model incorporates the forecast errors from the two preceding days to improve the accuracy of its next prediction.
   - **AR(0) - Autoregressive (p=0):** The model does not use past price values to predict future values, suggesting that past prices are not statistically significant predictors of the next day's price once the trend and moving average components are accounted for.
2. **with drift Component:**
   - The model identified a statistically significant **drift coefficient of 1.8099**. Because the model is differenced (d=1), this drift represents a constant, positive, upward trend.

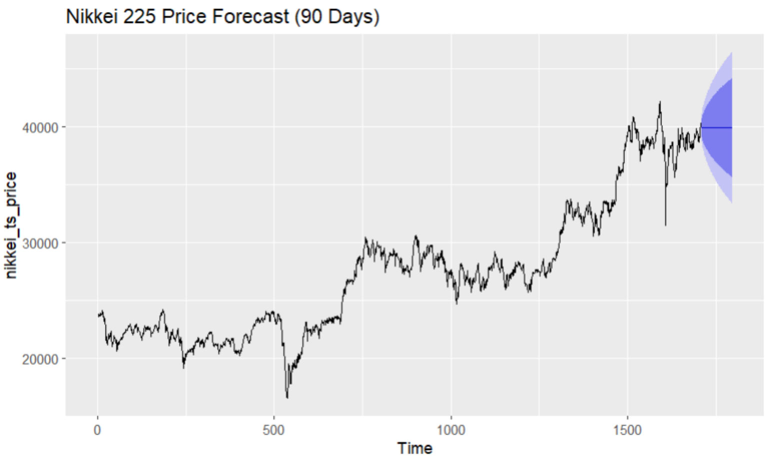
This suggests that during the observed period, the S&P 500 had an average daily increase of approximately 1.81 points.

3. **Model Fit:**
   - The model demonstrated a strong fit to the historical data, with a **Mean Absolute Percentage Error (MAPE) of only 0.8157%**. This indicates that the one-step-ahead forecasts on the training data were, on average, highly accurate. Furthermore, the **ACF1 of residuals (-0.0012)** is extremely close to zero, implying that the model's errors are random "white noise" and that no significant predictable information was left behind.
4. **Forecast Interpretation:**
   - The 90-day forecast plot shows the model projecting this 1.81-point daily drift forward, resulting in a linear upward trend. The **80% and 95% confidence intervals**, visualized as the shaded blue cones, start narrow and become progressively wider. This accurately reflects the high degree of uncertainty inherent in long-term market forecasting.

Conclusion: The analysis identifies the S&P 500's behavior during this period as a **"random walk with drift."** While the market exhibited a strong positive long-term trend (the drift), its day-to-day price movements were largely unpredictable (the random walk component). The model has successfully captured this underlying trend and appropriately quantified the significant uncertainty associated with future price projections.

Nikkei 225 ARIMA Forecast

```
Series: nikkei_ts_price
ARIMA(1,1,1)

Coefficients:
          ar1     ma1
      -0.7672  0.7280
s.e.   0.1285  0.1358

sigma^2 = 130561:  log likelihood = -12475.02
AIC=24956.03   AICc=24956.05   BIC=24972.36

Training set error measures:
                   ME     RMSE     MAE        MPE      MAPE      MASE       ACF1
Training set 9.837062 361.0153 248.714 0.02324633 0.9096696 0.9978214 0.01200862
```



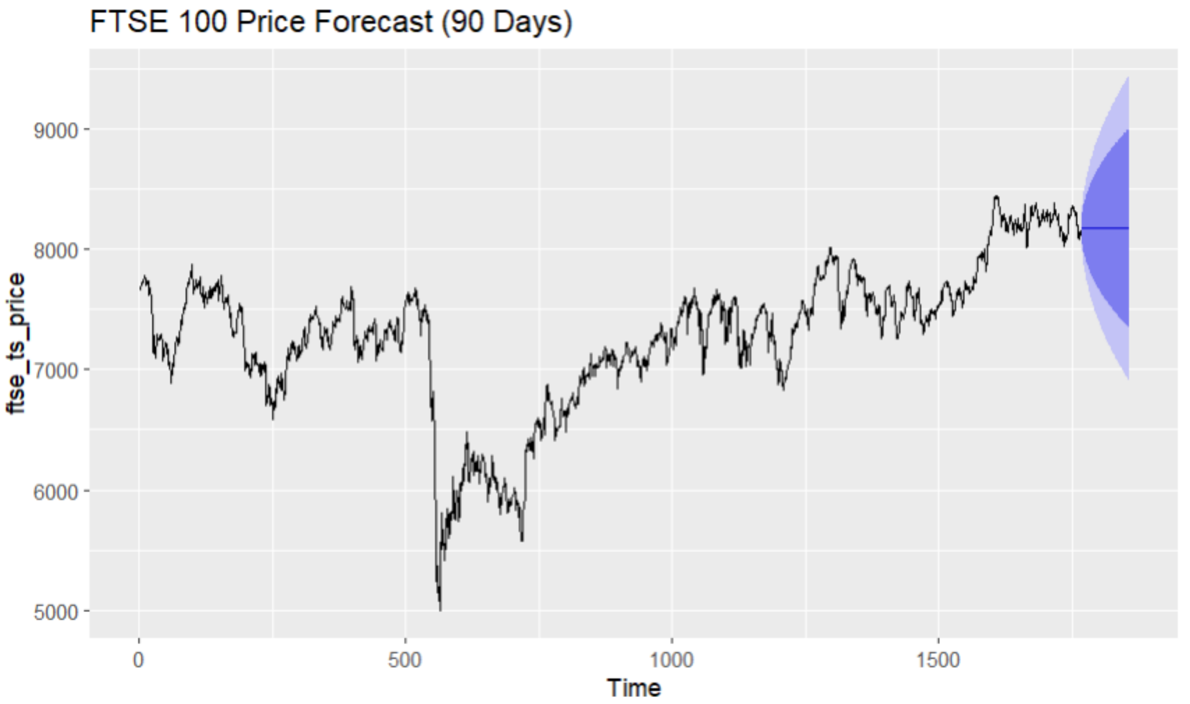
The auto.arima() function selected an **ARIMA(1,1,1)** model for the Nikkei 225 price series. This model's structure and forecast are notably different from the S&P 500's model.

1. **Model Structure (ARIMA(1,1,1)):**
   - **I(1) - Integrated (d=1):** Similar to the S&P 500, the "1" indicates the price series was non-stationary and required one order of differencing. The model is therefore predicting the *daily change in price*.
   - **AR(1) - Autoregressive (p=1):** The model uses the price change from the previous day to help predict the current day's price change. The ar1 coefficient is **-0.7672**.
   - **MA(1) - Moving Average (q=1):** The model uses the forecast error from the previous day to refine its current prediction. The ma1 coefficient is **0.7280**.
2. **Absence of drift:**
   - The most significant finding is the **absence of a "drift" term**. Unlike the S&P 500 model, auto.arima did *not* find a statistically significant, constant upward trend for the Nikkei. This implies that during the 2018-2024 period, the index's differenced series (its daily changes) fluctuated around a mean of zero, without a consistent daily positive gain.
3. **Model Fit:**
   - The model fits the historical data very well, as shown by a low **Mean Absolute Percentage Error (MAPE) of 0.91%**.
   - The **ACF1 of residuals (0.012)** is very close to zero, indicating the model's errors are random and that the predictable patterns in the historical data have been successfully captured.
4. **Forecast Interpretation:**
   - The 90-day forecast plot shows a **flat (horizontal) blue line**. This is a direct and correct visualization of an ARIMA(1,1,1) model *without drift*.
   - **Why is it flat?** The AR(1) and MA(1) components only have a short-term effect that decays rapidly. For a long-term forecast, the model's best guess for the *average future daily change* is zero (due to no drift). Therefore, the long-term forecast for the *price level* reverts to the last observed price, projected horizontally.
   - The **widening confidence intervals** (shaded areas) again show that while the *most likely* path is flat, the level of uncertainty about the future price increases significantly over time.

Conclusion: The analysis identifies the Nikkei 225's behavior as a classic **"random walk,"** in contrast to the S&P 500's "random walk with drift." The model's best long-term prediction for the Nikkei is simply its last known price, with no significant upward or downward trend.

FTSE 100 ARIMA Forecast

```
Series: ftse_ts_price
ARIMA(0,1,0)

sigma^2 = 4633:  log likelihood = -9964.8
AIC=19931.6   AICc=19931.61   BIC=19937.08

Training set error measures:
                    ME     RMSE      MAE         MPE      MAPE      MASE        ACF1
Training set 0.3012263 68.04519 47.44943 -0.00138211 0.6747315 0.9995255 -0.02013754
```



The auto.arima() function selected an **ARIMA(0,1,0)** model for the FTSE 100 price series. This specific model is the statistical definition of a **"Random Walk."**

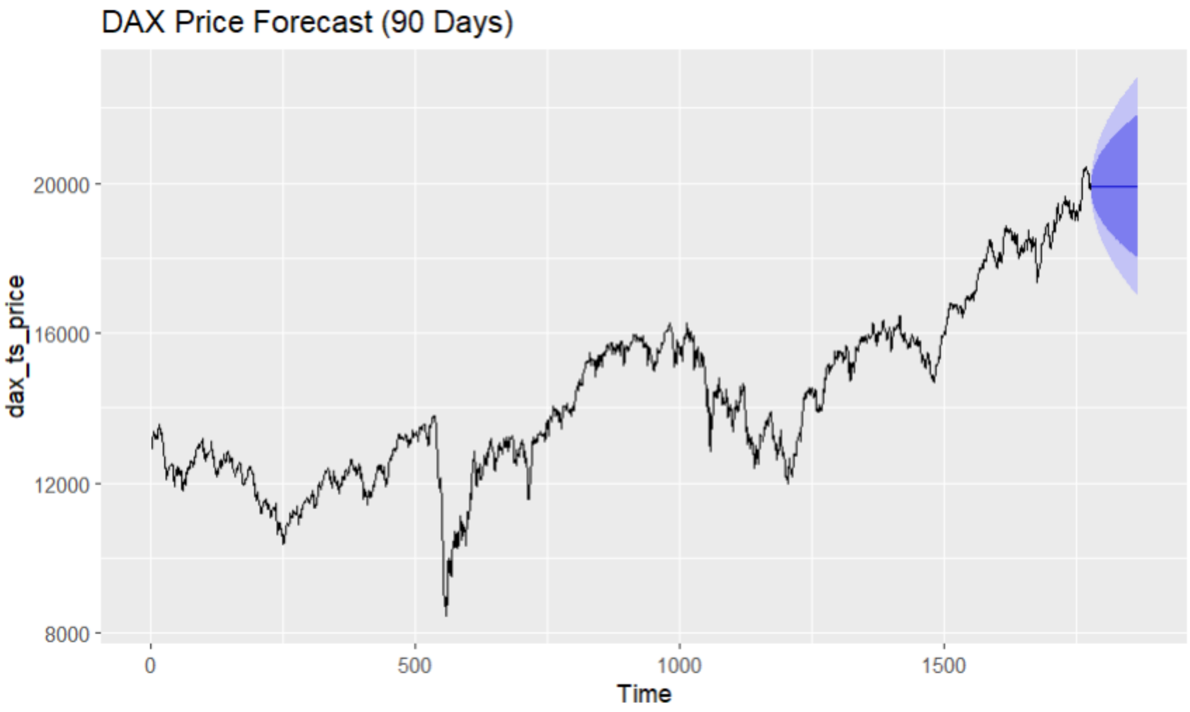
1. **Model Structure (ARIMA(0,1,0)):**
   - **I(1) - Integrated (d=1):** The "1" signifies that the price series was non-stationary and required one order of differencing. The model is therefore predicting the *daily change in price* (i.e., Price_today - Price_yesterday).
   - **AR(0) & MA(0):** The model contains **no autoregressive (p=0) or moving average (q=0) terms**. This implies that, after differencing, there is no useful, predictable information in the past price changes or past forecast errors.
2. **Absence of drift:**
   - Crucially, the model did **not** identify a "drift" component. This means that, unlike the S&P 500, the FTSE 100 did not exhibit a statistically significant, constant upward or downward trend in its daily price changes. The average daily change was effectively zero.
3. **Model Fit:**
   - Despite its simplicity, the model fits the historical data very well, with a low **Mean Absolute Percentage Error (MAPE) of 0.67%**.
   - The **ACF1 of residuals (-0.020)** is extremely close to zero, which strongly validates the model. It confirms that the model's errors (the differences between the actual and predicted values) are random white noise, meaning the model has captured all the predictable patterns (of which there were none).
4. **Forecast Interpretation:**

- The 90-day forecast plot is a direct and perfect visualization of a Random Walk model.
- With no AR, MA, or drift terms, the model's best forecast for the *daily change* is zero. Therefore, its best forecast for the *future price level* is simply the last observed price, projected forward as a **flat horizontal line**.
- The widening confidence intervals (shaded areas) correctly show that while the *most likely* outcome is no change, the potential for deviation from this path (i.e., risk) increases significantly over time.

Conclusion: The analysis concludes that the FTSE 100, during the 2018-2024 period, behaved as a **pure random walk**. This suggests that its future price movements are unpredictable based on its own historical data, as the best forecast for tomorrow's price is simply today's price.

DAX ARIMA Forecast

```
Series: dax_ts_price
ARIMA(0,1,0)

sigma^2 = 24819:  log likelihood = -11506.04
AIC=23014.07   AICc=23014.07   BIC=23019.55

Training set error measures:
                   ME     RMSE      MAE        MPE      MAPE      MASE        ACF1
Training set 3.96771 157.4968 112.0794 0.01720362 0.8189242 0.9995018 -0.02461074
> |
```



The auto.arima() function selected an ARIMA(0,1,0) model for the German DAX price series. This model is identical in structure to the one selected for the FTSE 100 and represents the statistical definition of a "Random Walk."

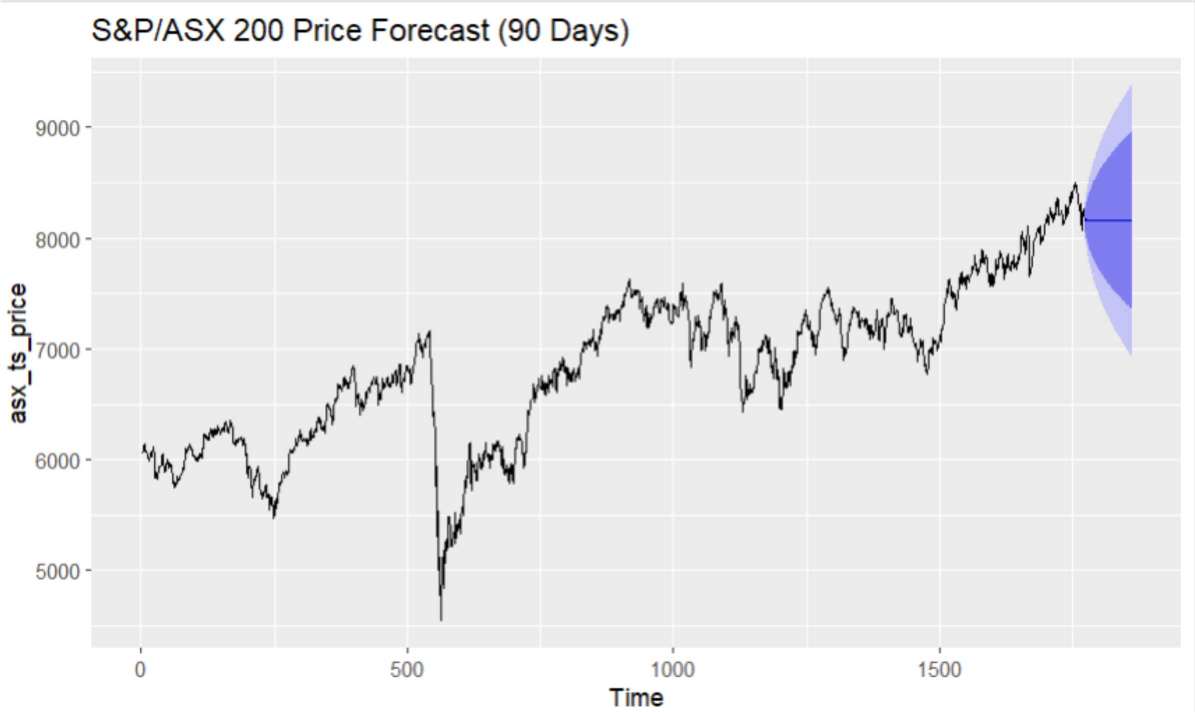
1. Model Structure (ARIMA(0,1,0)):

   - I(1) - Integrated (d=1): The "1" indicates the price series was non-stationary and required one order of differencing to become stationary. The model, therefore, forecasts the *daily change in price* rather than the price level.
   - AR(0) & MA(0): The model includes no autoregressive (p=0) or moving average (q=0) terms. This implies that, after differencing, there is no statistically useful information in past price changes or past forecast errors to predict future changes.
2. Absence of drift:
   - Critically, the model did not find a statistically significant "drift" component. This means that, unlike the S&P 500, the DAX did not have a consistent, positive average daily gain during this period. Its average daily change was statistically indistinguishable from zero.
3. Model Fit:
   - The model provides an excellent fit for the historical data, demonstrated by a very low Mean Absolute Percentage Error (MAPE) of 0.819%.
   - The ACF1 of residuals (-0.0246) is extremely close to zero, confirming that the model's errors are random white noise and that all predictable patterns (of which there were none) have been captured.
4. Forecast Interpretation:
   - The 90-day forecast plot is a perfect visualization of a Random Walk model.
   - With no AR, MA, or drift terms, the model's best forecast for the *daily change* is zero. Consequently, its best forecast for the *future price level* is the last observed price, projected forward as a flat horizontal line.
   - The widening shaded confidence intervals illustrate that while the most probable path is "no change," the range of possible outcomes (risk) increases significantly as the forecast horizon extends.

Conclusion: The analysis concludes that the DAX, similar to the FTSE 100, behaved as a pure random walk during the 2018-2024 period. This suggests that its future price movements are unpredictable based on its own past data, and the best statistical forecast for tomorrow's price is simply today's price.

S&P/ASX 200 ARIMA Forecast

```
Series: asx_ts_price
ARIMA(3,1,4)

Coefficients:
          ar1      ar2     ar3     ma1     ma2      ma3     ma4
      -0.2393  -0.0012  0.7029  0.1451  0.0420  -0.6681  0.0500
s.e.   0.1553   0.1742  0.1099  0.1574  0.1793   0.1212  0.0333

sigma^2 = 3943:  log likelihood = -9841.12
AIC=19698.25   AICc=19698.33   BIC=19742.08

Training set error measures:
                   ME     RMSE      MAE       MPE     MAPE     MASE         ACF1
Training set 1.121804 62.65085 44.76876 0.0112751 0.672643 1.003964 0.0007705095
```



The auto.arima() function selected an **ARIMA(3,1,4)** model for the S&P/ASX 200 price series. This is the most complex model structure identified among all five indices.

1. **Model Structure (ARIMA(3,1,4)):**
   - **I(1) - Integrated (d=1):** As with all the other indices, the "1" indicates the price series was non-stationary and required one order of differencing. The model is predicting the *daily change in price*.
   - **AR(3) - Autoregressive (p=3):** The model uses price changes from the last three days to help predict the current day's change.
   - **MA(4) - Moving Average (q=4):** The model uses forecast errors from the last four days to refine its current prediction.
   - **Complexity:** This (3,1,4) structure suggests that auto.arima found short-term, complex serial correlations in the daily returns of the ASX 200 that were not present (or as significant) in the other markets.
2. **Absence of drift:**
   - Despite its structural complexity, the model did **not** find a statistically significant "drift" component. This is a critical finding. It means that, like the Nikkei, FTSE, and DAX, the ASX 200 did not exhibit a consistent, positive average daily gain during this period. Its average daily change was effectively zero.
3. **Model Fit:**
   - The model provides an excellent fit for the historical data, with a very low **Mean Absolute Percentage Error (MAPE) of 0.67%**.
   - The **ACF1 of residuals (0.00077)** is practically zero, which is a very strong confirmation that the model's errors are random white noise and that all predictable short-term patterns have been successfully captured.
4. **Forecast Interpretation:**
   - The 90-day forecast plot shows a **flat horizontal line**. This may seem counter-intuitive for such a complex model, but it is the correct long-term forecast.
   - **Why is it flat?** The complex AR(3) and MA(4) terms only have a short-term impact that decays very quickly. For a long-term forecast, their influence fades to zero. Since

there is **no drift term**, the model's best long-term forecast for the *average future daily change* is zero. Therefore, the long-term forecast for the *price level* reverts to the last observed price, projected horizontally.

- The widening confidence intervals (shaded areas) again illustrate that this "no change" forecast is subject to significant and growing uncertainty over time.

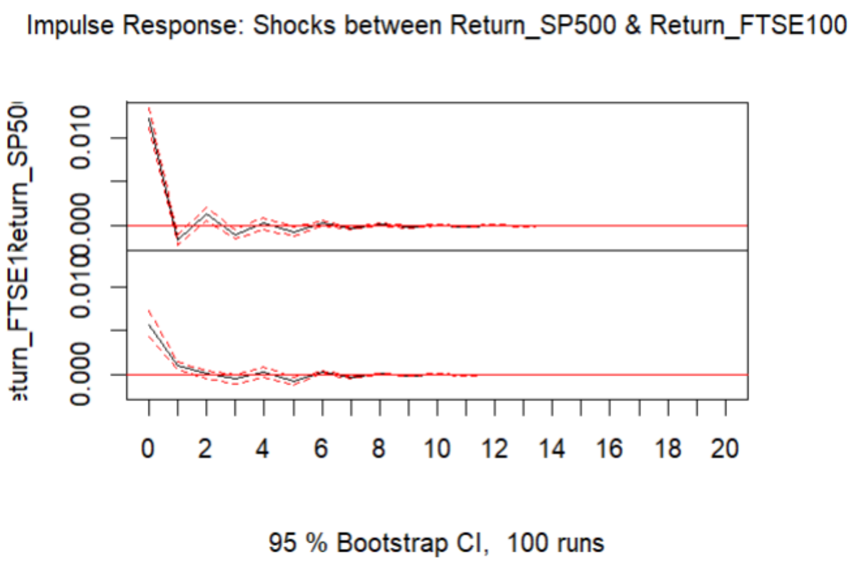
Conclusion: The analysis concludes that the S&P/ASX 200, despite having more complex short-term dynamics than its peers, ultimately behaved as a **"random walk"** during the 2018-2024 period. The model's best long-term prediction is simply its last known price, with no statistically significant upward or downward trend.

e) Causality tests using Granger Causality and VAR

Analyzing Pair: Return_SP500 vs. Return_FTSE100

```
--------------------------------------------------
--- Analyzing Pair: Return_SP500 vs. Return_FTSE100 ---
--------------------------------------------------
--- A: Selecting Optimal Lag Order ---
Selected optimal lag order (AIC): 5

--- B: Granger Causality Test ---
Test 1 -> H0: Return_FTSE100 does NOT Granger-cause Return_SP500
Granger causality test

Model 1: get(name1) ~ Lags(get(name1), 1:5) + Lags(get(name2), 1:5)
Model 2: get(name1) ~ Lags(get(name1), 1:5)
  Res.Df Df      F    Pr(>F)
1   1572
2   1577 -5 6.2968 8.544e-06 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Test 2 -> H0: Return_SP500 does NOT Granger-cause Return_FTSE100
Granger causality test

Model 1: get(name2) ~ Lags(get(name2), 1:5) + Lags(get(name1), 1:5)
Model 2: get(name2) ~ Lags(get(name2), 1:5)
  Res.Df Df      F    Pr(>F)
1   1572
2   1577 -5 10.617 4.773e-10 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

--- C: VAR Model for Short-Run Dynamics ---
VAR Model Summary:

VAR Estimation Results:
=========================
Endogenous variables: Return_SP500, Return_FTSE100
Deterministic variables: const
Sample size: 1583
Log Likelihood: 10057.099
Roots of the characteristic polynomial:
0.6817 0.6435 0.6435 0.5719 0.562 0.562 0.5145 0.5145 0.4197 0.4197
Call:
VAR(y = combined_ts_df, p = lag_order, type = "const")


Estimation results for equation Return_SP500:
=============================================
Return_SP500 = Return_SP500.l1 + Return_FTSE100.l1 + Return_SP500.l2 + Return_FTSE100.l2 + Return_S
P500.l3 + Return_FTSE100.l3 + Return_SP500.l4 + Return_FTSE100.l4 + Return_SP500.l5 + Return_FTSE10
0.l5 + const

                    Estimate Std. Error t value Pr(>|t|)    
Return_SP500.l1   -0.0965173  0.0304514  -3.170  0.00156 ** 
Return_FTSE100.l1 -0.0858740  0.0365105  -2.352  0.01879 *  
Return_SP500.l2    0.0290507  0.0313851   0.926  0.35478    
Return_FTSE100.l2  0.1680940  0.0367070   4.579 5.03e-06 ***
Return_SP500.l3   -0.0914337  0.0313341  -2.918  0.00357 ** 
Return_FTSE100.l3  0.0335784  0.0369694   0.908  0.36387    
Return_SP500.l4   -0.0387191  0.0310698  -1.246  0.21288    
Return_FTSE100.l4  0.0603398  0.0369598   1.633  0.10276    
Return_SP500.l5   -0.0508558  0.0302384  -1.682  0.09280 .  
Return_FTSE100.l5  0.0213650  0.0359146   0.595  0.55201    
const              0.0005912  0.0003093   1.911  0.05615 .  
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1


Residual standard error: 0.01221 on 1572 degrees of freedom
Multiple R-Squared: 0.06179,	Adjusted R-squared: 0.05583 
F-statistic: 10.35 on 10 and 1572 DF,  p-value: < 2.2e-16
```

```
Estimation results for equation Return_FTSE100:
===============================================
Return_FTSE100 = Return_SP500.l1 + Return_FTSE100.l1 + Return_SP500.l2 + Return_FTSE100.l2 + Returr
_SP500.l3 + Return_FTSE100.l3 + Return_SP500.l4 + Return_FTSE100.l4 + Return_SP500.l5 + Return_FTSE
100.l5 + const

                    Estimate Std. Error t value Pr(>|t|)
Return_SP500.l1    1.516e-01  2.530e-02   5.993 2.55e-09 ***
Return_FTSE100.l1 -1.447e-01  3.034e-02  -4.771 2.01e-06 ***
Return_SP500.l2    3.200e-02  2.608e-02   1.227 0.219916
Return_FTSE100.l2  2.122e-02  3.050e-02   0.696 0.486657
Return_SP500.l3   -3.460e-02  2.604e-02  -1.329 0.184029
Return_FTSE100.l3 -3.846e-02  3.072e-02  -1.252 0.210769
Return_SP500.l4   -2.811e-03  2.582e-02  -0.109 0.913313
Return_FTSE100.l4  7.579e-02  3.071e-02   2.468 0.013700 *
Return_SP500.l5   -9.077e-02  2.513e-02  -3.613 0.000313 ***
Return_FTSE100.l5  8.489e-02  2.984e-02   2.845 0.004505 **
const             -7.498e-07  2.570e-04  -0.003 0.997673
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1


Residual standard error: 0.01015 on 1572 degrees of freedom
Multiple R-Squared: 0.04205,    Adjusted R-squared: 0.03596
F-statistic: 6.901 on 10 and 1572 DF,  p-value: 1.225e-10



Covariance matrix of residuals:
               Return_SP500 Return_FTSE100
Return_SP500      1.492e-04      6.951e-05
Return_FTSE100    6.951e-05      1.030e-04

Correlation matrix of residuals:
               Return_SP500 Return_FTSE100
Return_SP500         1.0000         0.5606
Return_FTSE100       0.5606         1.0000
```



The analysis of the S&P 500 (US) and FTSE 100 (UK) indicates a highly significant and bidirectional predictive relationship. The model's optimal lag order was identified as **5 days**.

1. Granger Causality (Predictive Relationship)

This test assesses whether the past returns of one index can statistically predict the future returns of the other.

- **Test 1: H0: Return_FTSE100 does NOT Granger-cause Return_SP500**
    - **Result:** The p-value was **8.544e-06** (or 0.000008544).

- **Interpretation:** This p-value is far below the 0.05 threshold, so we **strongly reject the null hypothesis**. This provides clear statistical evidence that the past returns of the UK's FTSE 100 *do* help predict the future returns of the S&P 500.

- **Test 2: H0: Return_SP500 does NOT Granger-cause Return_FTSE100**
  - **Result:** The p-value was **4.773e-10** (or 0.0000000004773).
  - **Interpretation:** This p-value is even smaller. We **overwhelmingly reject the null hypothesis**. This confirms that the S&P 500's past returns are a very powerful predictor of the FTSE 100's future returns.

Conclusion: The predictive relationship is **bidirectional**. Information flows both ways, with each market's history providing valuable information for forecasting the other.

2. Vector Autoregression (VAR) Model (Short-Run Dynamics)

The VAR model details the day-by-day influence each market has on the other, based on the 5-day lag structure.

- **Equation for Return_FTSE100 (How S&P 500 affects FTSE):**
  - The most significant predictor by far is Return_SP500.l1 (the S&P 500's return from **one day ago**), with a p-value of **2.55e-09**.
  - The positive coefficient (0.1516) means a positive return in the S&P 500 yesterday is strongly associated with a positive return in the FTSE 100 today. This captures the classic "overnight" effect, where the US market's closing performance heavily influences the UK market's open.
  - The lag from 5 days ago (Return_SP500.l5) is also highly significant (p-value 0.000313).
- **Equation for Return_SP500 (How FTSE affects S&P 500):**
  - The influence here is more distributed. The FTSE's returns from **one day ago** (.l1, p-value 0.01879), **two days ago** (.l2, p-value 5.03e-06), and **three days ago** (.l3, p-value 0.00357) are all statistically significant predictors of the S&P 500's current-day return.
  - This suggests that information from the UK market takes several days to be fully incorporated into US market prices, in contrast to the immediate S&P 500 -> FTSE 100 effect.

3. Impulse Response Function (IRF) (Shock Transmission)

The IRF plots show how a sudden, one-time "shock" (like a surprise 1% rally) in one market affects the other over 20 days.

- **Both Plots:** The graphs show that a shock in either market (S&P 500 or FTSE 100) causes an **immediate, positive, and statistically significant response** in the other. In both scenarios, the black response line jumps clearly outside the red 95% confidence bands on day 1.
- **Duration:** This spillover effect is very **short-lived**. By day 3 or 4, the response line is back within the confidence bands and centered on zero. This indicates that new information and

shocks are transmitted and absorbed very quickly (within 3-4 days) between these two markets.

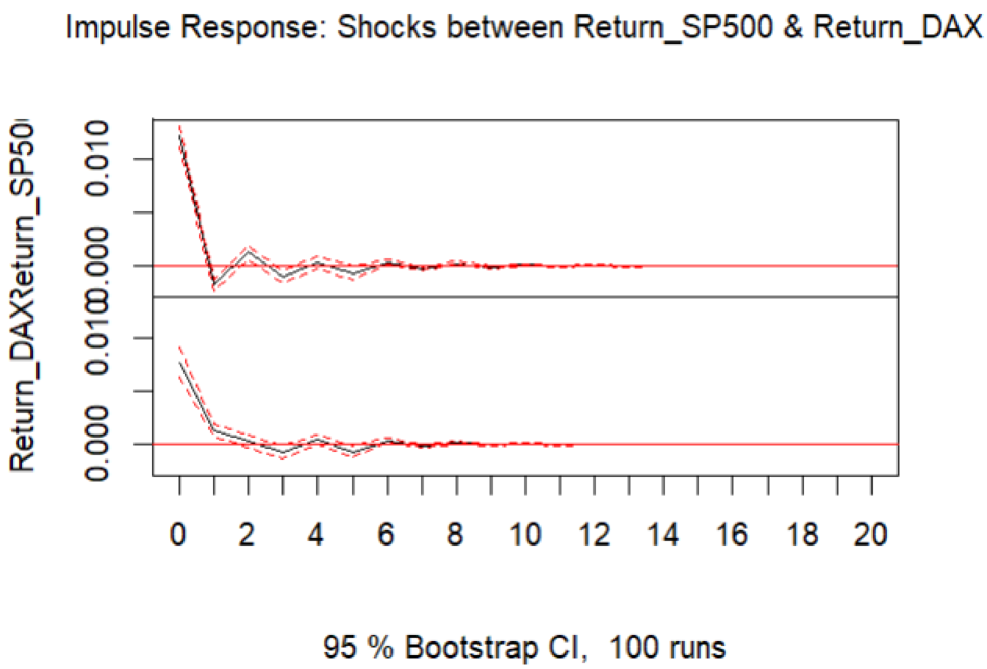
Overall Summary

The S&P 500 and FTSE 100 are deeply interconnected. A predictive relationship flows in both directions, but the **dominant effect is from the S&P 500 to the FTSE 100**, acting as a powerful overnight influence. The **correlation of residuals** (the "surprises" left over after the model) is **0.5606**, which is very high. This means that even after accounting for all the predictable lagged effects, on any given day, an unexpected move in the S&P 500 is still highly correlated with an unexpected move in the FTSE 100, highlighting their strong simultaneous co-movement.

Analyzing Pair: Return_SP500 vs. Return_DAX

```
--- A: Selecting Optimal Lag Order ---
Selected optimal lag order (AIC): 5

--- B: Granger Causality Test ---
Test 1 -> H0: Return_DAX does NOT Granger-cause Return_SP500
Granger causality test

Model 1: get(name1) ~ Lags(get(name1), 1:5) + Lags(get(name2), 1:5)
Model 2: get(name1) ~ Lags(get(name1), 1:5)
  Res.Df Df      F    Pr(>F)
1   1572
2   1577 -5 7.8802 2.428e-07 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Test 2 -> H0: Return_SP500 does NOT Granger-cause Return_DAX
Granger causality test

Model 1: get(name2) ~ Lags(get(name2), 1:5) + Lags(get(name1), 1:5)
Model 2: get(name2) ~ Lags(get(name2), 1:5)
  Res.Df Df      F   Pr(>F)
1   1572
2   1577 -5 11.338 9.18e-11 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

--- C: VAR Model for Short-Run Dynamics ---
VAR Model Summary:

VAR Estimation Results:
=========================
Endogenous variables: Return_SP500, Return_DAX
Deterministic variables: const
Sample size: 1583
Log Likelihood: 9890.667
Roots of the characteristic polynomial:
0.6875 0.5896 0.5896 0.571 0.571 0.5276 0.5264 0.5264 0.4327 0.4327
Call:
VAR(y = combined_ts_df, p = lag_order, type = "const")
```

```
Estimation results for equation Return_SP500:
=============================================
Return_SP500 = Return_SP500.l1 + Return_DAX.l1 + Return_SP500.l2 + Return_DAX.l2 + Return_SP500.
+ Return_DAX.l3 + Return_SP500.l4 + Return_DAX.l4 + Return_SP500.l5 + Return_DAX.l5 + const

                  Estimate Std. Error t value Pr(>|t|)    
Return_SP500.l1 -0.1482421  0.0324385  -4.570 5.26e-06 ***
Return_DAX.l1    0.0061401  0.0327823   0.187 0.851451    
Return_SP500.l2 -0.0277429  0.0339938  -0.816 0.414557    
Return_DAX.l2    0.1836364  0.0330992   5.548 3.38e-08 ***
Return_SP500.l3 -0.1137262  0.0340125  -3.344 0.000846 ***
Return_DAX.l3    0.0393077  0.0335102   1.173 0.240972    
Return_SP500.l4 -0.0811193  0.0337600  -2.403 0.016384 *  
Return_DAX.l4    0.1082563  0.0333788   3.243 0.001206 ** 
Return_SP500.l5 -0.0747454  0.0323769  -2.309 0.021095 *  
Return_DAX.l5    0.0331124  0.0320194   1.034 0.301232    
const            0.0005926  0.0003079   1.925 0.054442 .  
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1


Residual standard error: 0.01218 on 1572 degrees of freedom
Multiple R-Squared: 0.0664,     Adjusted R-squared: 0.06046 
F-statistic: 11.18 on 10 and 1572 DF,  p-value: < 2.2e-16 


Estimation results for equation Return_DAX:
===========================================
Return_DAX = Return_SP500.l1 + Return_DAX.l1 + Return_SP500.l2 + Return_DAX.l2 + Return_SP500.l3 +
Return_DAX.l3 + Return_SP500.l4 + Return_DAX.l4 + Return_SP500.l5 + Return_DAX.l5 + const

                 Estimate Std. Error t value Pr(>|t|)    
Return_SP500.l1  0.1990675  0.0320533   6.211 6.75e-10 ***
Return_DAX.l1   -0.1427338  0.0323931  -4.406 1.12e-05 ***
Return_SP500.l2  0.0531943  0.0335902   1.584  0.11348    
Return_DAX.l2    0.0341657  0.0327062   1.045  0.29636    
Return_SP500.l3 -0.0721736  0.0336086  -2.147  0.03191 *  
Return_DAX.l3   -0.0040658  0.0331123  -0.123  0.90229    
Return_SP500.l4 -0.0284277  0.0333591  -0.852  0.39425    
Return_DAX.l4    0.0850203  0.0329825   2.578  0.01004 *  
Return_SP500.l5 -0.0983053  0.0319924  -3.073  0.00216 ** 
Return_DAX.l5    0.0720837  0.0316392   2.278  0.02284 *  
const            0.0002410  0.0003042   0.792  0.42831    
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1


Residual standard error: 0.01204 on 1572 degrees of freedom
Multiple R-Squared: 0.04236,    Adjusted R-squared: 0.03626 
F-statistic: 6.953 on 10 and 1572 DF,  p-value: 9.829e-11 



Covariance matrix of residuals:
             Return_SP500 Return_DAX
Return_SP500    1.485e-04  9.229e-05
Return_DAX      9.229e-05  1.450e-04

Correlation matrix of residuals:
             Return_SP500 Return_DAX
Return_SP500       1.0000     0.6291
Return_DAX         0.6291     1.0000
```



The analysis of the S&P 500 (US) and DAX (Germany) reveals a strong, statistically significant, and bidirectional relationship. The optimal lag order for the model was determined to be **5 days**.

1. Granger Causality (Predictive Relationship)

The Granger Causality test checks if past values of one index's returns can be used to predict the future values of the other.

- **Test 1: H0: Return_DAX does NOT Granger-cause Return_SP500**
  - **Result:** The p-value was **2.428e-07** (which is 0.0000002428).
  - **Interpretation:** This p-value is extremely small (far less than 0.05). Therefore, we **strongly reject the null hypothesis**. This provides clear statistical evidence that past returns of the German DAX *do* contain information that helps predict the future returns of the S&P 500.
- **Test 2: H0: Return_SP500 does NOT Granger-cause Return_DAX**
  - **Result:** The p-value was **9.18e-11** (which is 0.0000000000918).
  - **Interpretation:** This p-value is even smaller. We **overwhelmingly reject the null hypothesis**. This confirms that past returns of the S&P 500 are a very strong predictor of the DAX's future returns.

Conclusion: The relationship is **bidirectional**. Information flows both ways, with each market's past performance providing predictive power for the other.

2. Vector Autoregression (VAR) Model (Short-Run Dynamics)

The VAR model breaks down the relationship lag by lag (day by day) to see *exactly where* the influence lies.

- **Equation for Return_DAX (How S&P 500 affects DAX):**
    - The single most significant predictor is Return_SP500.l1 (the S&P 500's return from **one day ago**), with a p-value of 6.75e-10.
    - The positive coefficient (0.199) means that a positive return in the S&P 500 yesterday is associated with a positive return in the DAX today. This captures the well-known "overnight" effect, where the US market's close influences the opening of the European markets.
    - Other S&P 500 lags (3, 4, and 5 days ago) are also significant, indicating a complex, multi-day influence.
- **Equation for Return_SP500 (How DAX affects S&P 500):**
    - Interestingly, the DAX's return from one day ago (Return_DAX.l1) is **not** statistically significant (p-value 0.851).
    - However, the DAX's return from **two days ago** (Return_DAX.l2) is **highly significant** (p-value 3.38e-08). The DAX's return from four days ago is also significant.
    - This suggests the information flow from Germany to the US may be slightly more *delayed* than the immediate effect seen from the US to Germany.

3. Impulse Response Function (IRF) (Shock Transmission)

The IRF plots visualize how a sudden, one-time "shock" (like a surprise 1% rally) in one market affects the other over the next 20 days.

- **Bottom Plot (Shock in S&P 500 -> Response in DAX):** This plot is very clear. A shock in the S&P 500 (on day 0) causes an **immediate, positive, and statistically significant** response in the DAX. The black line (the response) jumps well outside the red confidence bands on day 1.
- **Top Plot (Shock in DAX -> Response in S&P 500):** A shock in the DAX also appears to cause an immediate positive response in the S&P 500, but it is less pronounced than the reverse.
- **In both cases:** The effect is extremely **short-lived**. By day 3 or 4, the response line is back within the confidence bands and centered on zero, meaning the shock has been fully absorbed by the system.

Overall Summary

The S&P 500 and DAX are highly interconnected, with a strong contemporaneous correlation of **0.629** (from the covariance matrix). A predictive relationship exists in both directions. The strongest transmission mechanism is from the S&P 500 to the DAX, where a move in the US market is felt in the German market on the very next trading day. Any market shocks that occur are transmitted between the two, but this spillover effect dissipates very quickly, typically within 3-4 days.

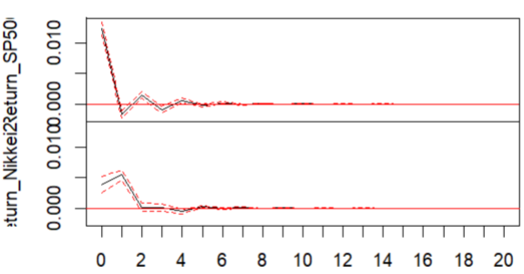
Analyzing Pair: Return_SP500 vs. Return_Nikkei225

```
--- A: Selecting Optimal Lag Order ---
Selected optimal lag order (AIC): 3

--- B: Granger Causality Test ---
Test 1 -> H0: Return_Nikkei225 does NOT Granger-cause Return_SP500
Granger causality test

Model 1: get(name1) ~ Lags(get(name1), 1:3) + Lags(get(name2), 1:3)
Model 2: get(name1) ~ Lags(get(name1), 1:3)
  Res.Df Df     F Pr(>F)
1   1578
2   1581 -3 1.953 0.1191

Test 2 -> H0: Return_SP500 does NOT Granger-cause Return_Nikkei225
Granger causality test

Model 1: get(name2) ~ Lags(get(name2), 1:3) + Lags(get(name1), 1:3)
Model 2: get(name2) ~ Lags(get(name2), 1:3)
  Res.Df Df      F    Pr(>F)
1   1578
2   1581 -3 128.64 < 2.2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

--- C: VAR Model for Short-Run Dynamics ---
VAR Model Summary:

VAR Estimation Results:
=========================
Endogenous variables: Return_SP500, Return_Nikkei225
Deterministic variables: const
Sample size: 1585
Log Likelihood: 9621.422
Roots of the characteristic polynomial:
0.5693 0.415 0.415 0.383 0.299 0.299
Call:
VAR(y = combined_ts_df, p = lag_order, type = "const")
```

```
Estimation results for equation Return_SP500:
==============================================
Return_SP500 = Return_SP500.l1 + Return_Nikkei225.l1 + Return_SP500.l2 + Return_Nikkei225.l2 + Re
rn_SP500.l3 + Return_Nikkei225.l3 + const

                      Estimate Std. Error t value Pr(>|t|)
Return_SP500.l1     -0.1594408  0.0265520  -6.005 2.37e-09 ***
Return_Nikkei225.l1  0.0558079  0.0279938   1.994  0.04637 *
Return_SP500.l2      0.0593012  0.0308845   1.920  0.05503 .
Return_Nikkei225.l2  0.0348456  0.0283206   1.230  0.21873
Return_SP500.l3     -0.0800063  0.0295950  -2.703  0.00694 **
Return_Nikkei225.l3  0.0303345  0.0247144   1.227  0.21985
const                0.0005366  0.0003099   1.731  0.08359 .
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1


Residual standard error: 0.0123 on 1578 degrees of freedom
Multiple R-Squared: 0.04492,    Adjusted R-squared: 0.04129
F-statistic: 12.37 on 6 and 1578 DF,  p-value: 1.19e-13


Estimation results for equation Return_Nikkei225:
==================================================
Return_Nikkei225 = Return_SP500.l1 + Return_Nikkei225.l1 + Return_SP500.l2 + Return_Nikkei225.l2 +
Return_SP500.l3 + Return_Nikkei225.l3 + const

                      Estimate Std. Error t value Pr(>|t|)
Return_SP500.l1      4.950e-01  2.521e-02  19.638  < 2e-16 ***
Return_Nikkei225.l1 -1.773e-01  2.657e-02  -6.670 3.52e-11 ***
Return_SP500.l2      1.664e-01  2.932e-02   5.676 1.64e-08 ***
Return_Nikkei225.l2 -2.238e-02  2.689e-02  -0.833   0.4052
Return_SP500.l3     -4.055e-04  2.809e-02  -0.014   0.9885
Return_Nikkei225.l3 -4.181e-02  2.346e-02  -1.782   0.0749 .
const                4.844e-05  2.942e-04   0.165   0.8692
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1


Residual standard error: 0.01168 on 1578 degrees of freedom
Multiple R-Squared: 0.1988,     Adjusted R-squared: 0.1957
F-statistic: 65.24 on 6 and 1578 DF,  p-value: < 2.2e-16
```

```
Covariance matrix of residuals:
                 Return_SP500 Return_Nikkei225
Return_SP500        1.514e-04        4.672e-05
Return_Nikkei225    4.672e-05        1.364e-04

Correlation matrix of residuals:
                 Return_SP500 Return_Nikkei225
Return_SP500           1.0000           0.3252
Return_Nikkei225       0.3252           1.0000
```

Impulse Response: Shocks between Return_SP500 & Return_Nikkei225



95 % Bootstrap CI, 100 runs

The analysis of the S&P 500 (US) and Nikkei 225 (Japan) reveals a clear unidirectional (one-way) predictive relationship, with the optimal lag order for the model determined to be 3 days.

1. Granger Causality (Predictive Relationship)

This test checks if past values of one index's returns can be used to predict the future values of the other.

- Test 1: H0: Return_Nikkei225 does NOT Granger-cause Return_SP500
    - Result: The p-value was 0.1191.
    - Interpretation: This p-value is greater than 0.05. Therefore, we fail to reject the null hypothesis. This provides strong statistical evidence that the past returns of Japan's Nikkei 225 (as a group, across all 3 lags) *do not* contain useful information for predicting the future returns of the S&P 500.
- Test 2: H0: Return_SP500 does NOT Granger-cause Return_Nikkei225
    - Result: The p-value was < 2.2e-16 (an extremely small number).
    - Interpretation: We overwhelmingly reject the null hypothesis. This confirms that the S&P 500's past returns are a very powerful and statistically significant predictor of the Nikkei 225's future returns.

Conclusion: The predictive relationship is unidirectional. The information flows from the S&P 500 to the Nikkei 225, but not the other way around.

2. Vector Autoregression (VAR) Model (Short-Run Dynamics)

The VAR model details the specific day-by-day influence.

- Equation for Return_Nikkei225 (How S&P 500 affects Nikkei):
    - This equation has very strong explanatory power. The S&P 500's return from one day ago (.l1, p-value < 2e-16) and two days ago (.l2, p-value 1.64e-08) are both massively significant predictors.
    - The Return_SP500.l1 coefficient of 0.495 (4.950e-01) is particularly strong. It implies that nearly 50% of the S&P 500's return from yesterday is, on average, reflected in the Nikkei's return today. This captures the powerful overnight effect of the US market's close on the Asian market's opening.
- Equation for Return_SP500 (How Nikkei affects S&P 500):
    - This equation confirms the Granger test results. The Nikkei's returns from two and three days ago (.l2 and .l3) are not statistically significant (p-values 0.218 and 0.219).
    - While the lag from one day ago (.l1) shows weak individual significance (p-value 0.04637), the *joint* test (the Granger F-test) found that the lags *as a group* have no predictive power, which is the more robust conclusion for causality.

3. Impulse Response Function (IRF) (Shock Transmission)

The IRF plots provide a clear visual confirmation of the one-way relationship.

- Bottom Plot (Shock in S&P 500 -> Response in Nikkei): This plot is dramatic. A sudden "shock" (a surprise rally) in the S&P 500 on day 0 causes an immediate, large, positive, and statistically significant response in the Nikkei on day 1. The black response line jumps far outside the red 95% confidence bands.
- Top Plot (Shock in Nikkei -> Response in S&P 500): This plot is the opposite. A shock in the Nikkei on day 0 causes no statistically significant response in the S&P 500. The black response line stays well within the red confidence bands, centered on zero.

Overall Summary

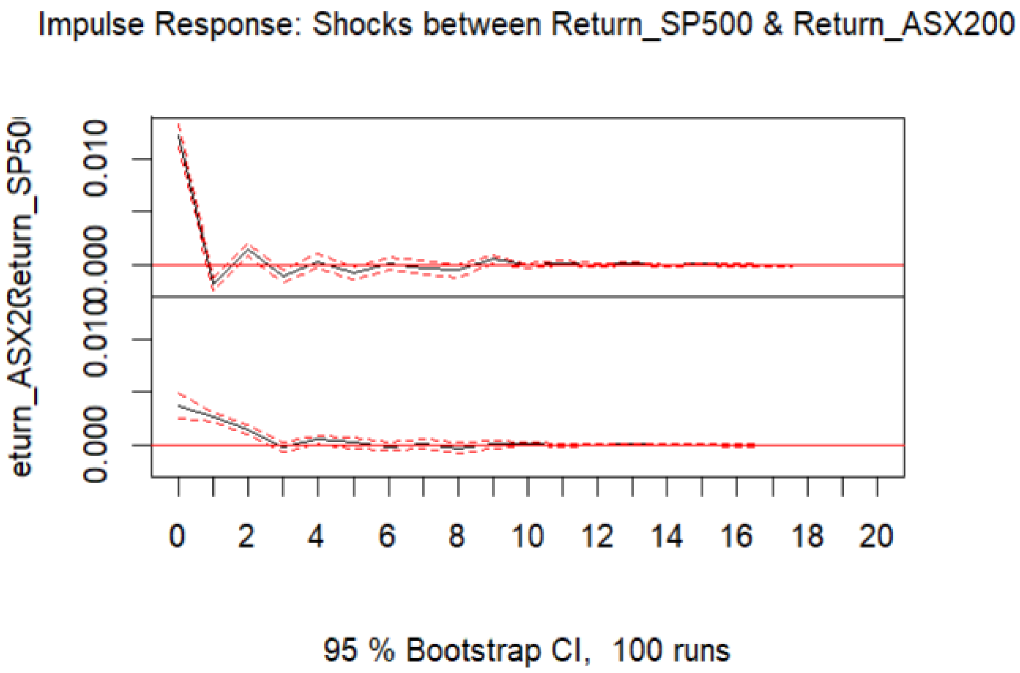
The relationship between the US and Japanese markets is a one-way street. The S&P 500 acts as a clear leading indicator for the Nikkei 225, with its performance from the previous 1-2 days strongly influencing the Nikkei's returns. However, this influence does not flow in reverse; the Nikkei's performance does not statistically predict the S&P 500's future movements. The correlation of residuals (the "surprises") is 0.3252, which is only moderate. This indicates that even after accounting for the S&P 500's predictive effect, the *unexpected* daily moves of the two markets are not as tightly linked as the US and European markets.

Analyzing Pair: Return_SP500 vs. Return_ASX200

```
--- A: Selecting Optimal Lag Order ---
Selected optimal lag order (AIC): 8

--- B: Granger Causality Test ---
Test 1 -> H0: Return_ASX200 does NOT Granger-cause Return_SP500
Granger causality test

Model 1: get(name1) ~ Lags(get(name1), 1:8) + Lags(get(name2), 1:8)
Model 2: get(name1) ~ Lags(get(name1), 1:8)
  Res.Df Df      F    Pr(>F)
1   1563
2   1571 -8 4.6616 1.188e-05 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Test 2 -> H0: Return_SP500 does NOT Granger-cause Return_ASX200
Granger causality test

Model 1: get(name2) ~ Lags(get(name2), 1:8) + Lags(get(name1), 1:8)
Model 2: get(name2) ~ Lags(get(name2), 1:8)
  Res.Df Df      F    Pr(>F)
1   1563
2   1571 -8 39.116 < 2.2e-16 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

--- C: VAR Model for Short-Run Dynamics ---
VAR Model Summary:

VAR Estimation Results:
=========================
Endogenous variables: Return_SP500, Return_ASX200
Deterministic variables: const
Sample size: 1580
Log Likelihood: 10050.383
Roots of the characteristic polynomial:
0.814 0.814 0.8097 0.7811 0.7562 0.7562 0.7494 0.7494 0.747 0.747 0.7373 0.7373 0.7053 0.7053 0.663
3 0.6633
Call:
VAR(y = combined_ts_df, p = lag_order, type = "const")
```

```
Estimation results for equation Return_SP500:
==============================================
Return_SP500 = Return_SP500.l1 + Return_ASX200.l1 + Return_SP500.l2 + Return_ASX200.l2 + Return_SP5
00.l3 + Return_ASX200.l3 + Return_SP500.l4 + Return_ASX200.l4 + Return_SP500.l5 + Return_ASX200.l5
+ Return_SP500.l6 + Return_ASX200.l6 + Return_SP500.l7 + Return_ASX200.l7 + Return_SP500.l8 + Retur
n_ASX200.l8 + const

                   Estimate Std. Error t value Pr(>|t|)    
Return_SP500.l1  -0.1023460  0.0275224  -3.719 0.000207 ***
Return_ASX200.l1 -0.1387496  0.0367143  -3.779 0.000163 ***
Return_SP500.l2   0.1558542  0.0305621   5.100 3.82e-07 ***
Return_ASX200.l2 -0.0683062  0.0384936  -1.774 0.076178 .  
Return_SP500.l3  -0.0399314  0.0317106  -1.259 0.208131    
Return_ASX200.l3  0.0630351  0.0384848   1.638 0.101639    
Return_SP500.l4  -0.0052292  0.0316427  -0.165 0.868762    
Return_ASX200.l4 -0.0208738  0.0382020  -0.546 0.584866    
Return_SP500.l5  -0.0376452  0.0315434  -1.193 0.232878    
Return_ASX200.l5 -0.0040092  0.0381451  -0.105 0.916308    
Return_SP500.l6  -0.0202233  0.0315926  -0.640 0.522184    
Return_ASX200.l6  0.0823258  0.0382028   2.155 0.031316 *  
Return_SP500.l7  -0.0554316  0.0315231  -1.758 0.078867 .  
Return_ASX200.l7  0.0851802  0.0369298   2.307 0.021210 *  
Return_SP500.l8  -0.1125848  0.0297314  -3.787 0.000158 ***
Return_ASX200.l8  0.1071167  0.0337925   3.170 0.001555 ** 
const             0.0005708  0.0003105   1.838 0.066203 .  
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1


Residual standard error: 0.01221 on 1563 degrees of freedom
Multiple R-Squared: 0.06804,	Adjusted R-squared: 0.0585 
F-statistic: 7.132 on 16 and 1563 DF,  p-value: 3.411e-16
```

```
Return_ASX200 = Return_SP500.l1 + Return_ASX200.l1 + Return_SP500.l2 + Return_ASX200.l2 + Return_SP
500.l3 + Return_ASX200.l3 + Return_SP500.l4 + Return_ASX200.l4 + Return_SP500.l5 + Return_ASX200.l5
+ Return_SP500.l6 + Return_ASX200.l6 + Return_SP500.l7 + Return_ASX200.l7 + Return_SP500.l8 + Retur
n_ASX200.l8 + const

                   Estimate Std. Error t value Pr(>|t|)    
Return_SP500.l1   0.3238657  0.0206302  15.699  < 2e-16 ***
Return_ASX200.l1 -0.3465909  0.0275202 -12.594  < 2e-16 ***
Return_SP500.l2   0.2504755  0.0229087  10.934  < 2e-16 ***
Return_ASX200.l2 -0.0208525  0.0288540  -0.723  0.46998    
Return_SP500.l3   0.0303398  0.0237696   1.276  0.20200    
Return_ASX200.l3 -0.0085114  0.0288474  -0.295  0.76800    
Return_SP500.l4   0.0555938  0.0237187   2.344  0.01921 *  
Return_ASX200.l4 -0.0289121  0.0286354  -1.010  0.31281    
Return_SP500.l5   0.0394728  0.0236443   1.669  0.09523 .  
Return_ASX200.l5  0.0746235  0.0285927   2.610  0.00914 ** 
Return_SP500.l6  -0.0167877  0.0236811  -0.709  0.47849    
Return_ASX200.l6  0.0618773  0.0286360   2.161  0.03086 *  
Return_SP500.l7   0.0175555  0.0236290   0.743  0.45762    
Return_ASX200.l7 -0.0551808  0.0276818  -1.993  0.04639 *  
Return_SP500.l8  -0.0251155  0.0222860  -1.127  0.25993    
Return_ASX200.l8  0.1032321  0.0253301   4.075 4.82e-05 ***
const            -0.0001525  0.0002327  -0.655  0.51250    
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1


Residual standard error: 0.00915 on 1563 degrees of freedom
Multiple R-Squared: 0.2063,	Adjusted R-squared: 0.1982 
F-statistic: 25.39 on 16 and 1563 DF,  p-value: < 2.2e-16



Covariance matrix of residuals:
              Return_SP500 Return_ASX200
Return_SP500     1.490e-04     4.493e-05
Return_ASX200    4.493e-05     8.372e-05

Correlation matrix of residuals:
              Return_SP500 Return_ASX200
Return_SP500        1.0000        0.4023
Return_ASX200       0.4023        1.0000
```



Here is the in-depth, report-style interpretation for the S&P 500 and the S&P/ASX 200.

In-Depth Analysis: S&P 500 vs. S&P/ASX 200

The analysis of the S&P 500 (US) and S&P/ASX 200 (Australia) reveals a complex, statistically significant, and **bidirectional relationship**. The optimal lag order for the model was determined to be **8 days**, suggesting a more prolonged and intricate dynamic compared to the other index pairs.

1. Granger Causality (Predictive Relationship)

This test checks if the past returns of one index can be used to statistically predict the future returns of the other.

- **Test 1: H0: Return_ASX200 does NOT Granger-cause Return_SP500**
    - **Result:** The p-value was **1.188e-05** (or 0.00001188).
    - **Interpretation:** This p-value is extremely small (far less than 0.05), so we **strongly reject the null hypothesis**. This provides clear statistical evidence that the past returns of Australia's ASX 200 *do* contain information that helps predict the future returns of the S&P 500.
- **Test 2: H0: Return_SP500 does NOT Granger-cause Return_ASX200**
    - **Result:** The p-value was **< 2.2e-16** (an exceptionally small number).
    - **Interpretation:** We **overwhelmingly reject the null hypothesis**. This confirms that the S&P 500's past returns are a very powerful and significant predictor of the ASX 200's future returns.

Conclusion: The predictive relationship is **bidirectional**, with information flowing significantly in both directions.

2. Vector Autoregression (VAR) Model (Short-Run Dynamics)

The VAR model details the day-by-day influence over the 8-day lag period.

- **Equation for Return_ASX200 (How S&P 500 affects ASX 200):**
    - This equation shows the dominant, positive "overnight" effect. The S&P 500's returns from **one day ago** (.l1, p-value < 2e-16) and **two days ago** (.l2, p-value < 2e-16) are the most powerful predictors.
    - The positive coefficients (0.323 and 0.250) indicate that a positive return in the S&P 500 yesterday and the day before is strongly associated with a positive return in the ASX 200 today.
- **Equation for Return_SP500 (How ASX 200 affects S&P 500):**
    - This relationship is more complex and reveals a very interesting dynamic.
    - The ASX 200's return from **one day ago** (.l1) is highly significant (p-value 0.000163), but it has a **negative coefficient (-0.1387)**. This suggests that, on average, a positive return in the Australian market is associated with a *negative* return in the US market on the same day (given the time zones).
    - Significant positive influences appear much later, at lags of 6, 7, and 8 days. This indicates a very complex, multi-day feedback loop from the Australian market to the US.

3. Impulse Response Function (IRF) (Shock Transmission)

The IRF plots visualize how a sudden, one-time "shock" is transmitted between the markets.

- **Bottom Plot (Shock in S&P 500 -> Response in ASX 200):** This plot clearly confirms the VAR results. A positive shock in the S&P 500 (day 0) causes an **immediate, large, positive, and statistically significant** response in the ASX 200 on day 1. The black response line jumps well outside the red 95% confidence bands.
- **Top Plot (Shock in ASX 200 -> Response in S&P 500):** This plot is the most unique. It shows that a positive shock in the ASX 200 causes an **immediate, small, but statistically significant *negative*** response in the S&P 500. The black line clearly dips below the red confidence band on day 1, visually confirming the negative coefficient (-0.1387) found in the VAR model.
- **Duration:** In both directions, the initial shock is absorbed and the response returns to a statistically insignificant level (within the red bands) after about 3-4 days.

Overall Summary

The S&P 500 and ASX 200 have a strong, bidirectional, and *asymmetrical* relationship. The dominant information flow is the positive "overnight" effect from the US to Australia. However, a significant predictive relationship also flows from Australia to the US, but it is more complex, characterized by an immediate *negative* correlation followed by a delayed positive influence several days later. The **correlation of residuals is 0.4023**, indicating a moderate simultaneous relationship between the "unexpected" moves in both markets.

9. Conclusion

This analysis of five major global stock indices from 2018 to 2024 reveals a highly interconnected, yet US-centric, financial landscape. The findings from the descriptive, hypothetical, and time-series analyses provide a clear picture of market behavior and its implications for investors.

- **Key Risk and Return Observations:** The analysis of price and return data highlighted two key points. First, all markets exhibited significant **volatility clustering**, most notably the simultaneous, sharp crash in early 2020, confirming that global systemic risk affects all indices. Second, a major divergence in performance was identified: the **S&P 500 was the only index to display a statistically significant positive "drift,"** indicative of a consistent, positive average daily return. In contrast, the Nikkei, FTSE, DAX, and ASX 200 all behaved as "random walks," with average daily returns that were statistically indistinguishable from zero.
- **Market Relationships and Causality:** The Granger Causality and VAR models overwhelmingly identified the **S&P 500 as the primary driver of global market returns**. A powerful **unidirectional** (one-way) predictive relationship flows from the S&P 500 to the Nikkei 225. Strong **bidirectional** (two-way) relationships were found between the S&P 500 and its European (FTSE, DAX) and Australian (ASX) counterparts. However, in all cases, the "overnight" effect from the US close was the most dominant predictive force. The Impulse Response Functions confirmed this, showing that shocks in the S&P 500 are transmitted immediately (within 1-2 days) to all other markets.
- **Forecast Implications for Investors:** The ARIMA forecasts are a direct reflection of these historical findings. The S&P 500 model, which found a positive drift, projects this trend forward in a **linear upward forecast**. Conversely, the models for all four other indices, which found no drift, project a **flat, horizontal forecast**, suggesting their best statistical guess for a future price is simply the last known price.

For an investor, this presents two crucial takeaways. First, the models suggest that only the S&P 500's upward momentum was statistically consistent enough to be projected. Second, and most importantly, every forecast plot featured **rapidly widening confidence intervals**. This is a critical statistical warning that, regardless of the projected trend, the range of possible outcomes (the risk) becomes enormous over a 90-day horizon, reinforcing the high degree of uncertainty in financial markets and the limitations of using purely historical data for future projections.

References:Data Sources

- Investing.com. (2024). *Historical Stock Market Data*. Retrieved from https://www.investing.com
- Yahoo Finance. (2024). *Historical Market Data*. Retrieved from https://finance.yahoo.com

Project by:
Ayush Jain
Pilla Poushya Meghana
Bhavya
Sachin Shanmugam
Malay Aditya