TSAF2024

TSAF 2024

TSAF2024

Forecasting stock prices and stock returns
Holt-Winters exponential smoothing model
ARIMA
Efficient Market Hypothesis
Random Walk Hypothesis
Exploratory data analysis
Alpha and Beta
Distributions and correlations of stock returns
Modern portfolio theory
Mean-Variance Optimization
Efficient frontier, Sharpe ratio, Tangency portfolio
APT and CAPM (Capital Asset Pricing Model)
Factor models
Volatility modeling and forecasting
Machine learning & Algo trading

Required
Decent Python coding skills
Numpy, Matplotlib, Pandas, and Scipy
Probability & Statistics

Recomended References

1. Chris Brooks (2019) Introductory Econometrics for Finance, Fourth Edition, Cambridge University Press and Tao, Ran and Tao, Ran and Brooks, Chris, Python Guide to Accompany Introductory Econometrics for Finance (October 25, 2019). Available at SSRN: https://ssrn.com/abstract=3475303 or http://dx.doi.org/10.2139/ssrn.3475303

Code and data sets are available at https://www.cambridge.org/gb/academic/subjects/economics/finance/introductory-econometrics-finance-4th-edition?format=PB&isbn=9781108422536

* (3rd ed) https://sites.google.com/view/davidgabauer/use-rpython

2. Philip Hans Franses, Dick van Dijk and Anne Opschoor (2014). *Time Series Models for Business and Economic Forecasting*, 2nd ed. Cambridge University Press.

3. Marco Peixeiro (2022). *Time Series Forecasting in Python*. Manning Publications.

4. Rob J Hyndman and George Athanasopoulos (2021). *Forecasting: Principles and Practice* (3rd ed) OTexts. https://otexts.com/fpp3/

5. tsa4.pdf (ucla.edu)

6. ⓘ ts 531 (Winter 2016) Analysis of Time Series (ionides.github.io)

TSAF2024

- John Y. Campbell, Andrew W. Lo, A. Craig MacKinlay () The Econometrics of Financial Markets, Princeton University Press
- Greg N. Gregoriou, Razvan Pascalau (2011) Financial Econometrics Modeling: Market Microstructure, Factor Models and Financial Risk Measures, Palgrave Macmillan
- Stan Hurn, Vance L. Martin, Jun Yu, Peter C. B. Phillips (2020) Financial Econometric Modeling, Oxford University Press

YouTube

- Shmueli <https://www.youtube.com/watch?v=y5HG8ZQ_mfE&list=PLoK4oIB1jeK0LHLbZW3DTT05e4srDYxFq>

Assessment

Cat I (15%) + DataCamp + Empirical paper (20%) + Attendance

Final Exam (60%)

DAY ONE

- Revision of Statistics

DAY TWO (OHCD data)

<data> <lecture>

We will cover:

- Returns (simple vs compound vs cumulative vs log returns)
- Adjusted Prices/Volume
- Some basic trading strategies (by excel) - MACD

  Here are some practice questions:

  - Why use Adjusted Price for backtesting? Why not use the actual prices, such as the closing price?
  - Why is just adding simple daily returns (i.e., the cumulative returns) not an appropriate measure of investment returns over time? What is the correct way to fine the cumulative returns?
  - Suggest three reasons for using logarithms? If you have the log (say, log price), how can you find the original value (price)?

TSAF2024

<excel> <notes> <lecture> <lecture>

- What is a white noise?
- What do we mean by a stationary series?
- How do you use the autocorrelation function (ACF) to tell whether a series is stationary of non-stationary?
- Suggest four kinds of non-stationarity series.
- Given a series, use Excel to find the ACF.

DAY FOUR (Smoothing)

<notes> <excel> <lecture>

- How does the simple exponential smoother (SES) get its name?
- In what situations do we apply the MA, SES, Holts and Holts-Winter methods?
- How do we measure which of the methods fit the data best? What are MSE, RMSE, MAP, MAPE, etc.?
- We learnt Solver in Excel in class. Can you (a) think of a how we can use Solver to find the max/min of a quadratic equation? (b) how to solve the classical least squares problem using Solver?

DAY FIVE (Forecasting)

<excel> <python> <python> <lecture>

- What is the difference between static and dynamic forecasting?
- Get some time series data and do a forecast after splitting the series into -in-sample and out-of-sample periods using a forecasting model you have learnt. Compare the forecasting performance with the random walk (or what is known as the naive model). Here are some examples of time series data <link> from Kaggle.

DAY SIX (Random Walk model)

<excel> <lecture>

DAY SEVEN (More trading strategies + portfolio risk and returns CAPM)

<excel> <lecture>

We cover VWAP, Rolling VWAP and RSI (For more see DataCamp).

We also looked at CAMP, or expected returns and expected SD of portfolios.

TSAF2024

Dec 3 (Tuesday)

>> Start Online 6.00 p.m. <<

DAY SEVEN (ARIMA modeling)

<excel> <lecture>

This is the workhorse of time series econometrics. A good introduction can be found in Shmueli's YouTube VDOs. The DataCamp courses on Time Series and ARIMA should also be useful!

- Note the mathematical form of AR and MA models (see my notes posted above on Day 3).
- Make sure you know how to derive the expected value or mean and variance of a AR(1) process and show that it is stationary.
- Also derive the expected value or mean and variance of a MA(1) process.
- Combining gives he ARMA(p, q) model.
- Note that ARMA models require the series to be stationary.

  Here are some questions to think about:

  - Show that MA(1) is stationary.
  - Under what conditions is the AR(1) model stationary? How about the AR(2) model?

DAY EIGHT (ARIMA modelling con'td)

<python> <lecture>

- Fitting an ARIMA model to data
- Using ACF and PACF to determine the order of ARIMA
- Relevance of information criteria (AIC, BIC, etc.)
- In-sample and out-sample.
- Evaluating "fit" - RMSE again!

MERRY CHRISTMAS EVERYONE & A HAPPY NEW YEAR

See you Jan 7, 2025

DAY NINE (ARIMA modelling con'td)

<python> <lecture>

S ⓘ for the delay. Had some trechnical issues accessing google services.

TSAF2024

<lecture>

Ch12-13 POE <link>

Pairs Trading <VDO> <python> <EWC> <EWA>

DAY ELEVEN

VAR + VECM

JEP - VAR <link>

notes <link>

Lecture <link>

ARCH-GARCH

Ch14 POE <link>

JEP - GARCH <link>

Excel <link>

Lecture <link>

Exercises/revision <link>

Link to submit your term empirical paper/project <link>

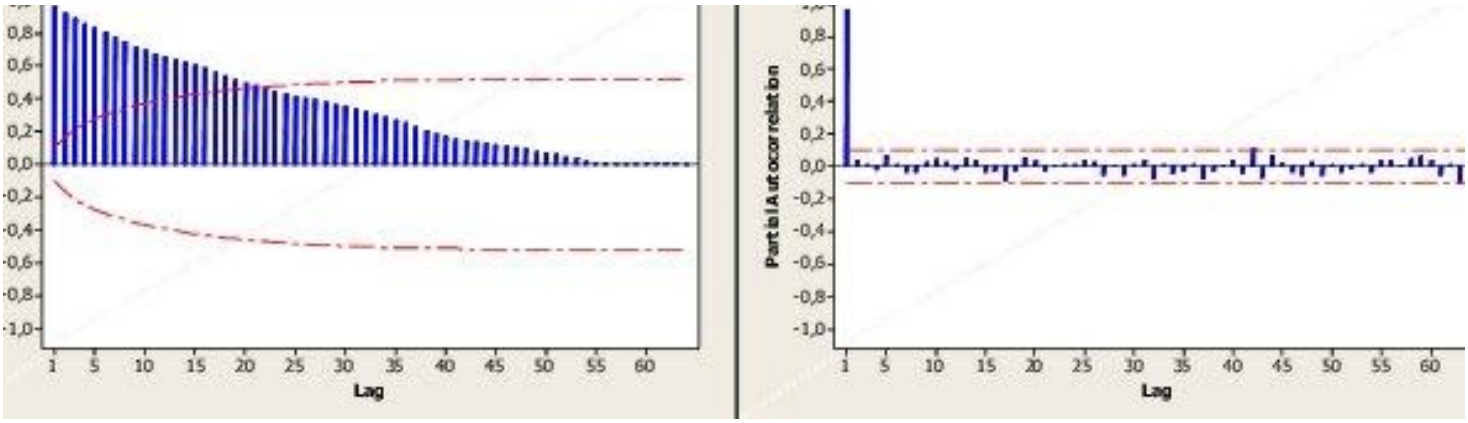
DEADLINE Feb 10 2025 (on final exam day!!)

Sample Exam Questions

1. Below is the ACF and PACF for a series called "Low".

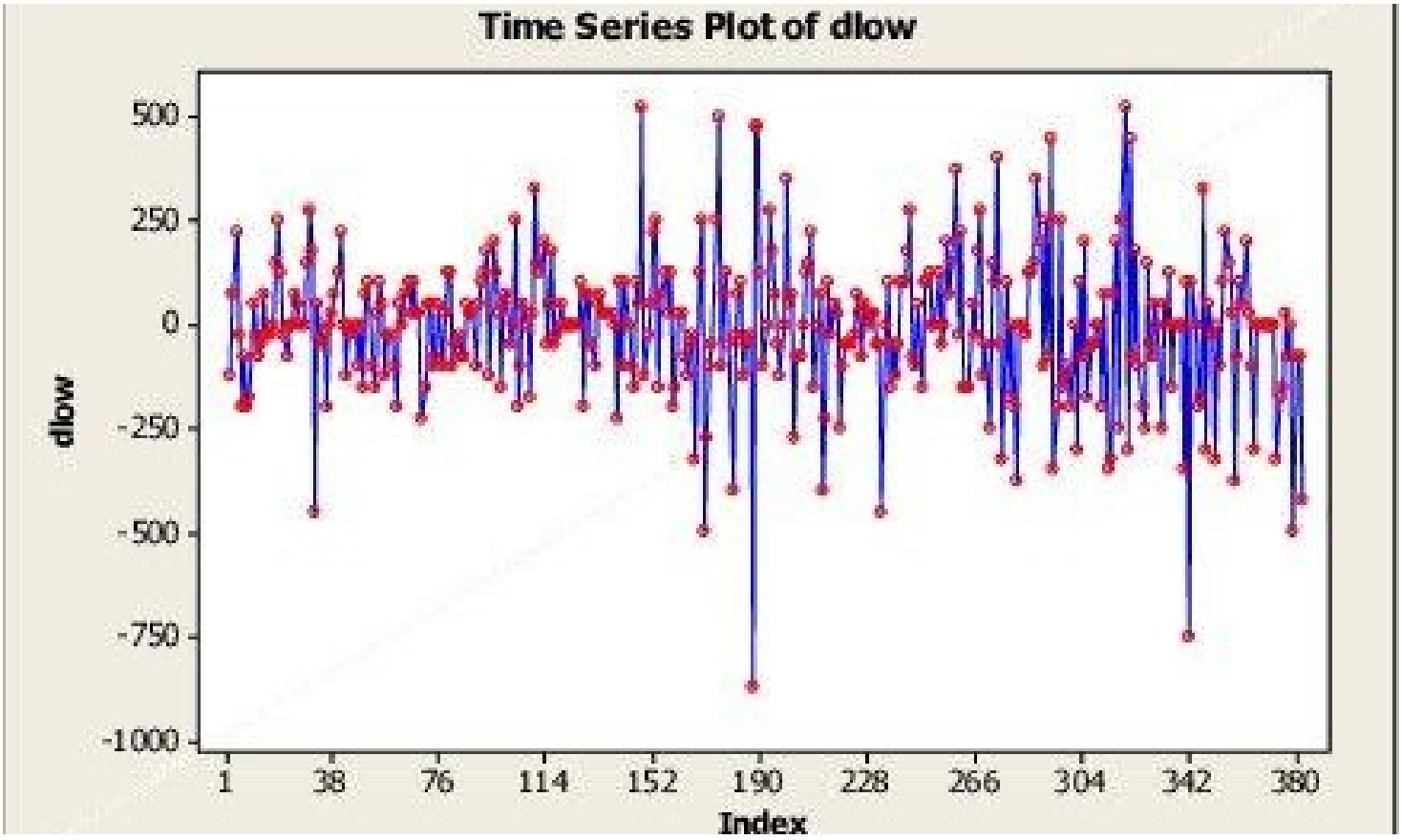
(a) Do you think the series is stationary or non-stationary?

TSAF2024



(b) You decide to fit a ARIMA(2, 1, 0) model. Write down the model specification in full.

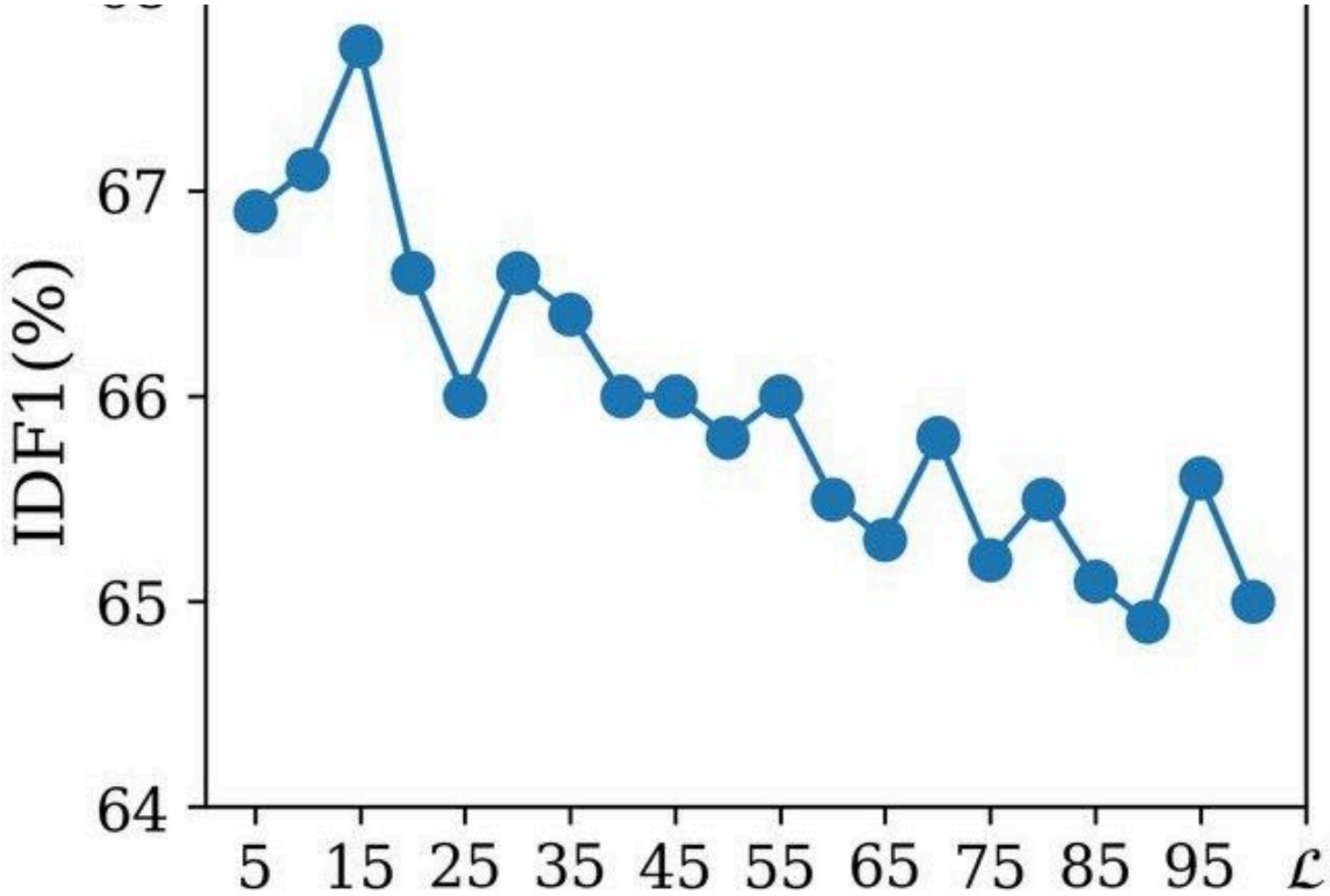
(c) The figure below plots the residuals for your ARIMA(2, 1, 0) model. Explain why you would like to check the residuals and whether you are okay with what you see.



(d) Alternatively you try fitting an ARIMA(0, 2, 1) model. Write down the model specification in full and explain why would you want to do this?

(e) How would you decide which is the better model for forecasting? ARIMA(2, 1, 0) or ARIMA(0, 2, 1)?

TSAF2024



(a) We carry out an ADF test and find:

```
ADF Statistic: -0.731503
p-value: 0.838420
Critical Values: 1%: -3.578 ; 5%: -2.925 ; 10%: -2.601
```

State the null and alternative and conclude whether we have a unit-root or not.

(b) Taking the difference of the series gives:

```
ADF Statistic: -3.047661
p-value: 0.030677
```

What do you conclude about the original series?

3. Below are forecasted and actual sales for 3 products as well as the **Mean absolute percentage error (MAPE)** calculated for each day.

TSAF2024

Product A		Mon	Tue	Wed	Thu	Fri	Sat	Sun	Average
	Sales	1	1	6	5	7	4	4	
	Forecast	2	2	2	2	2	2	2	
	MAPE	100%	100%	67%	60%	71%	50%	50%	71%

Product B		Mon	Tue	Wed	Thu	Fri	Sat	Sun	Average
	Sales	7	14	6	6	23	15	10	
	Forecast	12	12	12	14	19	25	18	
	MAPE	71%	14%	100%	133%	17%	67%	80%	69%

Product C		Mon	Tue	Wed	Thu	Fri	Sat	Sun	Average
	Sales	18	28	19	29	46	28	44	
	Forecast	16	18	21	22	37	42	40	
	MAPE	11%	36%	11%	24%	20%	50%	9%	23%

(a) Write down the formula for the MAPE, mean absolute deviation (MAD), mean square error (MSE), and root mean square forecasting error (RMSE) , explaining and highlighting the strength and weakness of each measure.

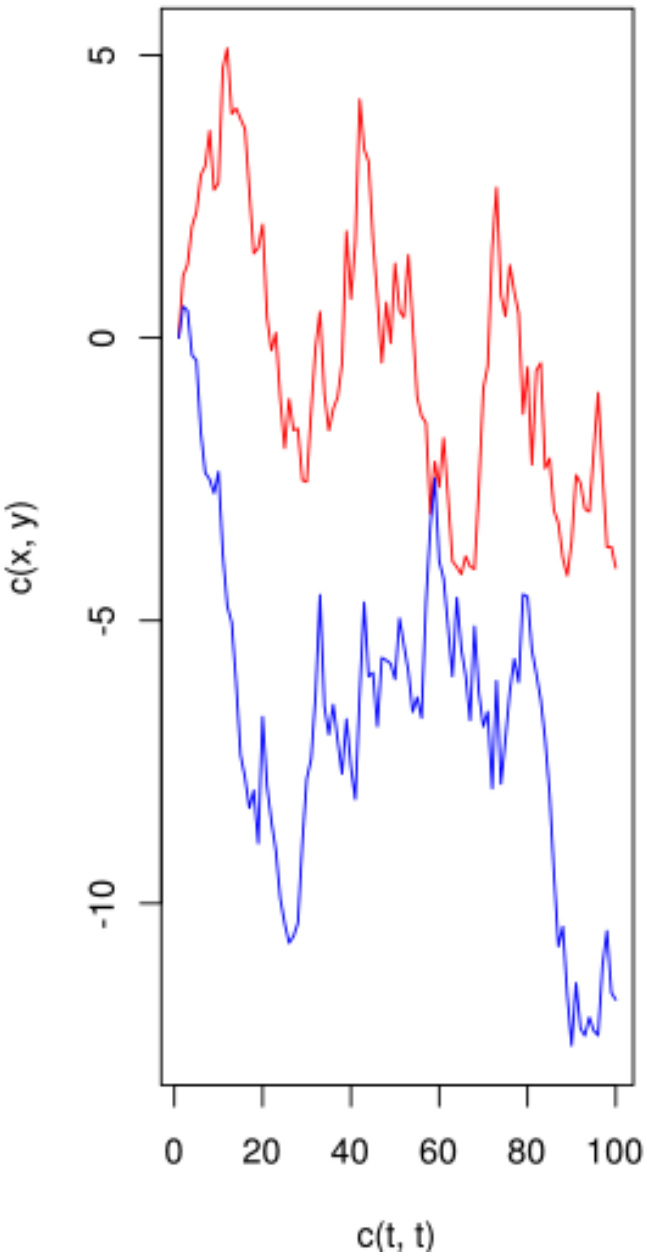
(b) Perform all the calculations for MAE, MSE, RMSE and decide for which product the forecasting method is best suited.

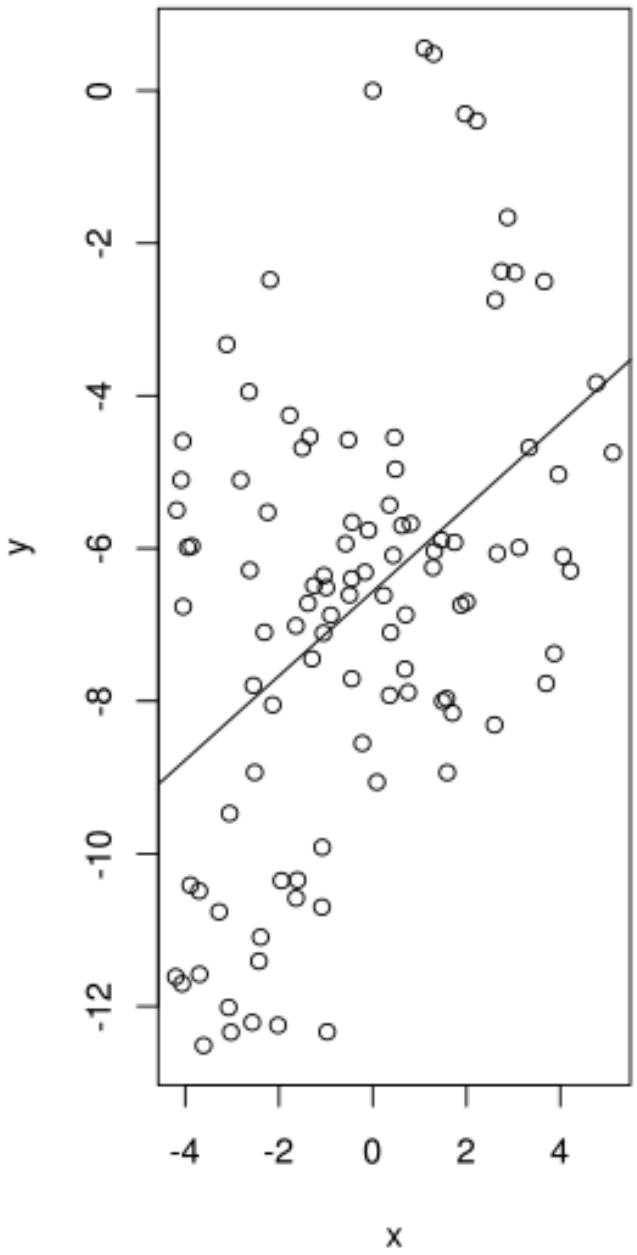
(c) What is the difference between forecasting bias and forecasting efficiency/variance? Given the bias-variance tradeoff, in which situation would you prefer the other in forecasting?

(d) What do we mean by over-fitting and under-fitting and why are these concepts important in the context of forecasting.

3. Given two series (red and blue shown below), the adjacent scatter plot shows strong positive correlation.

TSAF2024





TSAF2024

You get the following coefficients:

	Estimate:	Std. Err.	t-value	P-value
Intercept	-3.43193	0.31671	< 2e-16	***
beta	0.31607	0.08013	0.000111	***

Write out the regression in full. What does this regression tell you? What do you conclude?

(b) You decide to check the residuals for the above regression and find that the ADF statistic is -10.4398. What is the relationship between BLUE and RED series?

(c) Suppose you decide that a first order vector error correction model (VECM). Write down the full set of equations of interest. and explain why the VECM may be an appropriate model to better understand the relationship between BLUE and RED.

4. Let's consider a VAR with three endogenous variables, the unemployment rate, the inflation rate, and interest rates.

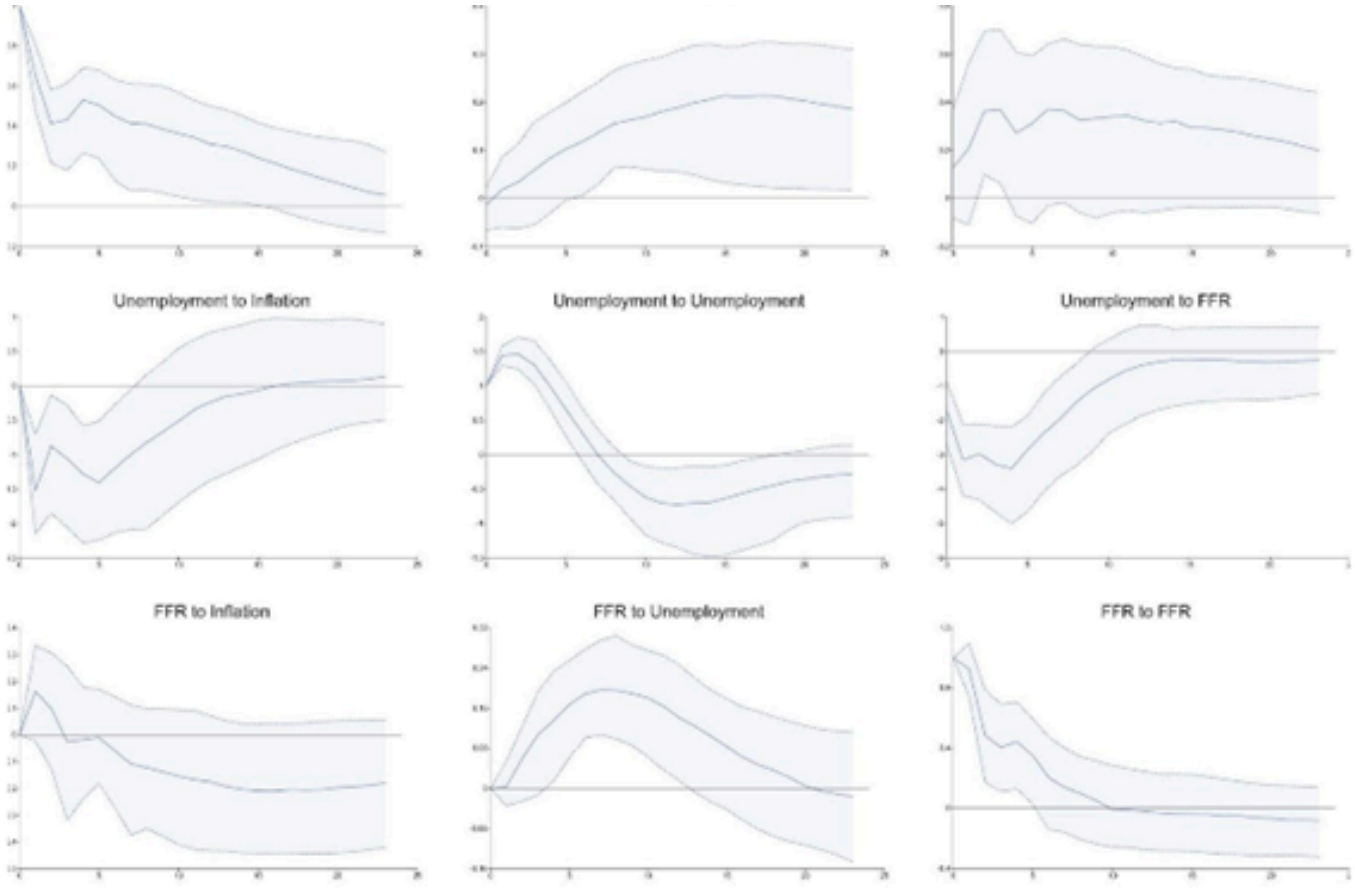
(a) What is the use of a VAR model? How does it differ from the Granger causality model?

(b) How do we choose the number of lags in a VAR model?

(c) Suppose, you decide on a VAR(2) model. Write down the system of equations for a recursive VAR(2) modeling the unemployment rate, the inflation rate, and interest rates.

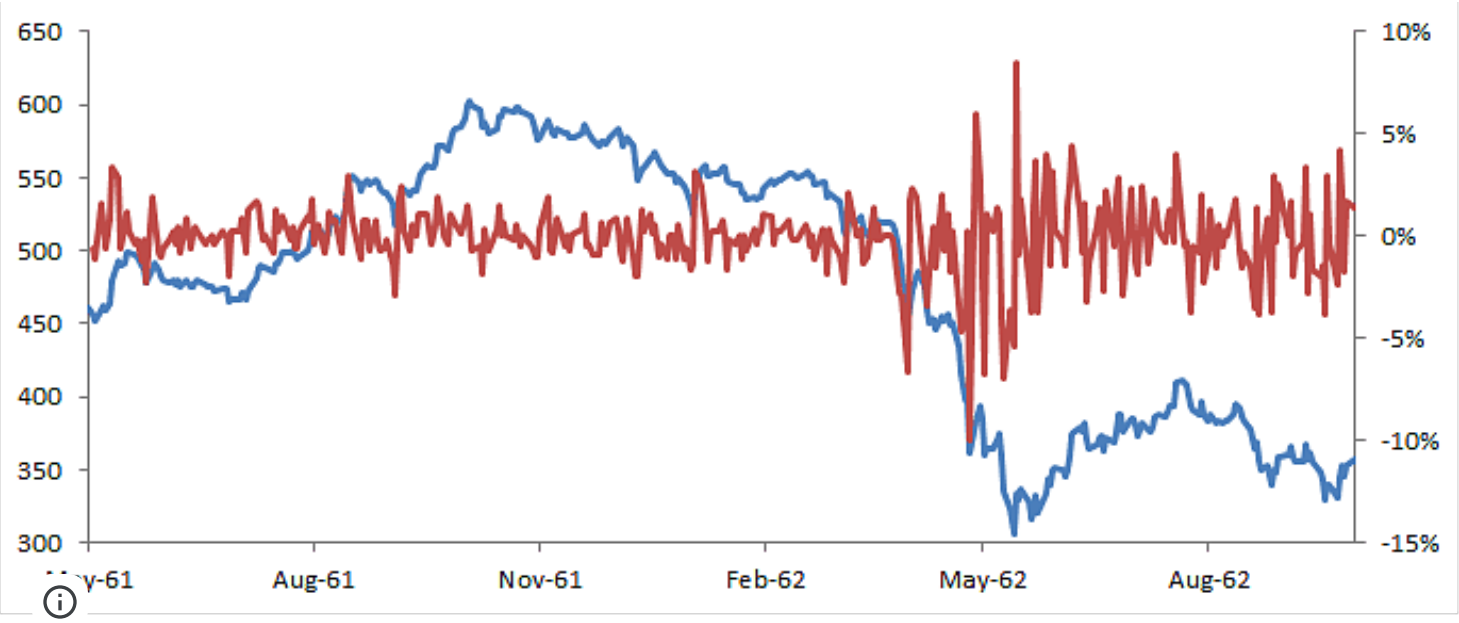
After estimating the model, here are the impulse reponse functions (IRF):

TSAF2024



(d) From the IRF, can you tell whether the stationarity requirement for the VAR(2) model is satisfied? Give an economic interpret of the IRF above.

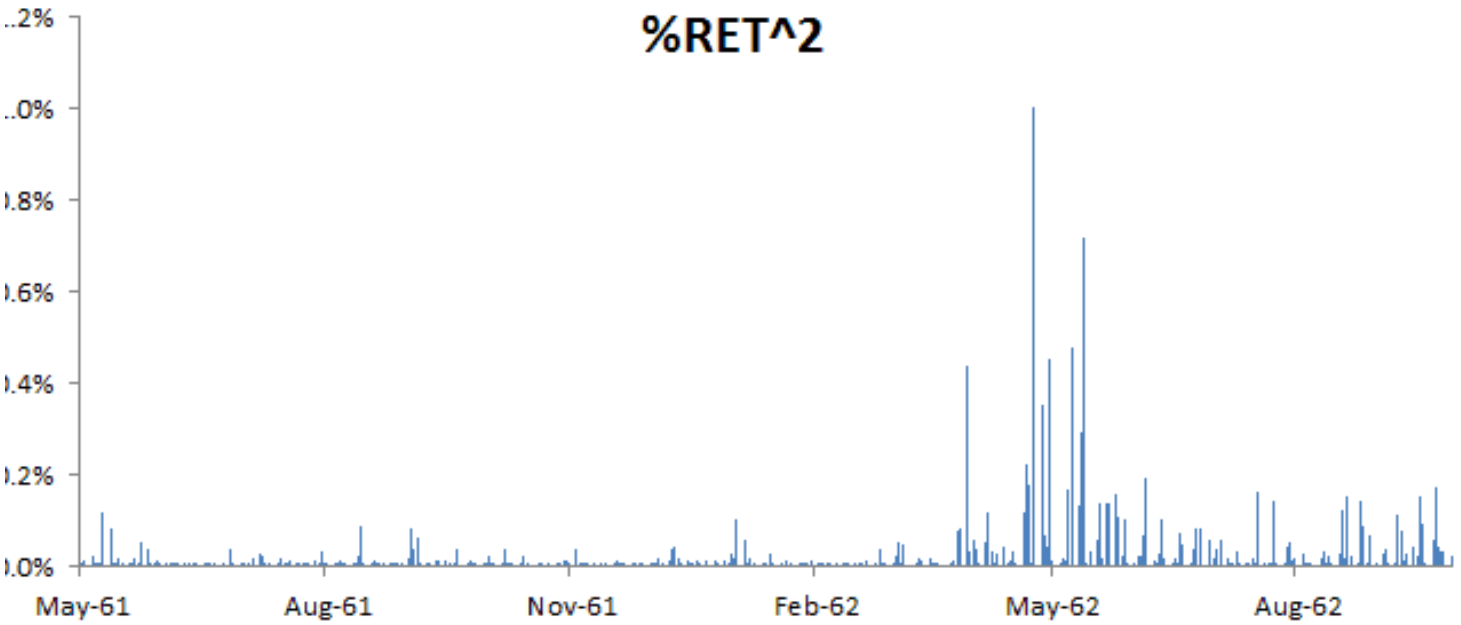
5. Below is a daily financial time series (in Blue) and its logarithmic daily difference (in Red):



TSAF2024

(b) Do you think there are some signs of an ARCH effect? How would you check for the ARCH effect?

You decide the square the returns and plot this:



Below is the ACF and PACF for the squared returns.

(c) Do you think these are useful in deciding which model to use? Explain briefly.

TSAF2024

1	30.27%	10.22%	-10.22%	30.28%	10.22%	-10.22%
2	18.76%	10.22%	-10.22%	10.57%	10.22%	-10.22%
3	32.12%	11.11%	-11.11%	26.55%	10.22%	-10.22%
4	30.60%	11.44%	-11.44%	16.95%	10.22%	-10.22%
5	3.96%	12.35%	-12.35%	-15.34%	10.22%	-10.22%
6	15.77%	13.11%	-13.11%	8.29%	10.22%	-10.22%
7	11.10%	13.13%	-13.13%	-6.12%	10.22%	-10.22%
8	12.14%	13.32%	-13.32%	9.05%	10.22%	-10.22%
9	29.78%	13.42%	-13.42%	30.70%	10.22%	-10.22%
10	26.51%	13.53%	-13.53%	9.99%	10.22%	-10.22%
11	14.58%	14.20%	-14.20%	1.29%	10.22%	-10.22%
12	24.20%	14.71%	-14.71%	3.16%	10.22%	-10.22%

6. (a) When dealing with time series data, one often encounters a problem of autocorrelation. What are the consequences of autocorrelation? How can we test for autocorrelation?

(b) Can you sketch (1) a deterministic trend nonstationary series (2) a stochastic trend nonstationary series (3) a series with a structural break, and (4) a series exhibiting volatility clustering.

(c) How would you model the nonstationary series (1)-(4) above? Explain briefly your answer.