homework

You are advised, but not obliged, to work on this problem set in groups of up to three people. Groups can change for different problem sets. You are free to use any software you are familiar with. Each group has to hand in one solution sheet (available for each problem set on Canvas) together with the unique calculation file (R/Python/Julia or whatever software you used). Answers in the solution sheet should be typed (or written legibly). The deadline is March 10 midnight. Late submissions will not be accepted by the system. Please round your estimated values when presenting them in the solution sheet to enhance readability.

1. Import `s1_data_hw.txt` into RStudio. This dataset contains the following variables:

name	description	source
SPY	S&P 500 index, in USD	Bloomberg
SXXP	STOXX Europe 600 index, in USD	
NKY	Nikkei 225 index, in USD	
XAU	Gold spot price, USD per troy ounce	
CL1	WTI crude oil, front-month futures contract, in USD	
DFEDTARU	Federal Funds Target Rate - Upper Limit, annualized	
SPDIV	Dividend per "share" of the S&P 500 index, annualized	R. Shiller's website
RF	10-yr US government bond yields, annualized	St. Louis Fed

2. Calculate the monthly logarithmic returns of *SP500, STOXX600, NIKKEI, GOLD* and *WTI*. Convert the risk-free rate to monthly values. Subsample your data to start in January 2011. Report the mean logarithmic return for *STOXX600* and *NIKKEI*. (**2 points**)

3. In separate graphs, plot the time series of the *WTI* price index and of *WTI* returns. Comment briefly on any striking feature of the return series. (**1 point**)

4. Test the null hypothesis that the average monthly return of *STOXX600* is significantly different from zero at the 5% level. Report the t-statistic and whether you reject or cannot reject the null. (**2 points**)

5. Calculate the 95% confidence interval around the mean return of *GOLD*. Can you be 95% sure that *GOLD* has been a better investment (in terms of average return) than the risk-free rate over this period? (**2 points**)

6. Calculate the correlation matrix of the five risky assets' returns. With which asset are *WTI* and *SP500* most correlated, respectively? Does any correlation coefficient surprise you? Hint: The 5 risky assets are SPY, SXXP, NKY, XAU, CL1. (**2 point**)

7. Among the five risky assets, report the asset with the highest excess kurtosis and the corresponding coefficient. Plot a histogram for this asset's returns, overlaying a fitted normal distribution. Is the return distribution approximately normal? (**2 points**)

8. Report which of the five risky assets is farthest from a normal distribution according to the Jarque-Bera test. Report the test statistic and p-value. (**1 point**)

9. In this question, we reconsider the full sample and ask whether equity returns differ systematically across Federal Reserve monetary policy regimes. **(8 points)**

   (a) Using the *DFEDTARU* series, construct a binary dummy variable D_t that identifies Federal Reserve hiking cycles. A hiking cycle **starts** in month t if $DFEDTARU_t > DFEDTARU_{t-1}$. A cycle **ends** in the month of the final rate hike immediately preceding either a rate cut or a pause (no change) lasting 6 months or longer.

   Importantly, let $D_t = 1$ for all months within the window [Start, End], including months where the rate was held constant between two hikes. Report the total number of hiking-cycle months ($\sum D_t$) in your sample. **(2 points)**

   Hint: First, identify all months where $\Delta DFEDTARU > 0$. Then, define a new cycle whenever the gap between two consecutive hikes exceeds 5 months. The cycle ends at the date of the last hike before such a gap or before a rate decrease.

   (b) We should account for dividends, as they constitute a meaningful component of total equity return. At time t you buy the index at the price listed in *SP500*; your actual return at $t+1$ is the new index value plus any dividends collected between the two dates. Construct the total return series by adding $SPDIV/12$ to the index level and dividing by the lagged index level. Report the standard deviation of this total return. **(1 point)**

   Hints: (1) Use `lag()` to lag a time-series object. (2) Use `na.rm=TRUE` when computing statistics in the presence of missing values.

   (c) Select only the months classified as hiking-cycle months. Calculate the average total return of *SP500* and its interquartile range during these periods. **(2 points)**

   Hint: To subset by a dummy variable, use an expression such as `mydata[mydummy==1]`.

   (d) Repeat the same exercise for months outside a hiking cycle (i.e., easing/pause months). **(1 point)**

   (e) Assume the two underlying populations have equal variance. Use the `t.test` function in R to perform a two-sample t-test (not a paired one), testing:

$$H_0 : \mu_{\text{hike}} = \mu_{\text{ease}} \tag{1}$$

   Report the t-statistic, the p-value, and interpret the result. Are average returns during hiking and easing cycles significantly different at the 5% level? Does the sign of the difference align with your prior expectation? **(2 points)**