Tutorial Questions

Attempt both questions either in a notebook or by writing a script, but I would STRONGLY recommend that you practice your workflow for the final exam by writing a script for each answer, and assembling a final pdf document as requested.

Question 1

The following data is sourced from the World Bank. https://data.worldbank.org. We download data on GDP per Capita, Total Population and Intentional Homicides per 100,000 people. The csv files contain the data and metadata on the indices used.

The task is to explore and explain the relationship between homicides and wealth (as measured by GDP per Capita).

1. Load the data for Homicides ("API_VC.IHR.PSRC.P5_DS2_en_csv_v2_4696515.csv"), GDP per Capita ("API_NY.GDP.PCAP.CD_DS2_en_csv_v2_4701206.csv") and Total Population ("API_SP.POP.TOTL_DS2_en_csv_v2_4701113.csv"). You may need to use the 'skiprows' argument to read_csv to correctly format the data.
2. Load the country level meta data from 'Metadata_Country_API_VC.IHR.PSRC.P5_DS2_en_csv_v2_4696515.csv'.
3. Index all the tables by country code.
4. Calculate means for the period 2010 to 2021 for the three variables (homicides, wealth, and population)
5. For the purposes of this exercise, we focus on 'higher income' countries, filter your data to extract data for these countries.
6. Using Seaborn relplot, plot the relationship between Homicides and Wealth. Differentiate between regions (from the metadata table) (hint use the 'hue' argument). Save your chart to a pdf file.
7. Import numpy and repeat the above plot using the log values of the data. (You will have to calculate np.log(X+.01) to calculate logs if some fields have zero values). What do you notice? Save your chart to a pdf file.
8. Explore the relationship between the variables using sns.lmplot on the logged values. What do you conclude? Save your chart to a pdf file.
9. Rank all remaining countries by Population and Wealth. Filter your data to focus on the countries in the top half of the sample on both metrics.
10. Plot Homicides v. Wealth for these large, wealthy countries. What do you conclude?
11. Assemble your charts into a single pdf file, briefly describe your findings and save the document.
12. You may be asked to show your document in class☺

Question 2

The following data is again from the Fama and French data library.
http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html#Research

In this exercise we use the daily data. The purpose of the exercise is to explore the relationship / trade off between market risk and return in different market 'states'.

1. Load the data (F-F_Research_Data_5_Factors_2x3_daily.csv), convert the dates to a pandas date index
2. Calculate the Market return by adding the RF column (risk free rate) back to the 'Mkt-RF' column.
3. Use the pandas 'rolling' command to calculate a rolling mean return for the market over a short (20 day) and long (250 day) window. Create two new binary variables based on the whether each indicator is positive or negative.
4. Resample the daily data to monthly frequency, calculating the mean and standard deviation of the daily return for each month in the sample. Produce a scatter plot of risk v. return (using Seaborn lmplot) and save to a pdf file.
5. Resample your indicator variables to monthly frequency, using resample('M').last() to select the value corresponding to the final day of the month as the resampled quantity.
6. Use the pandas .shift(1) operator to shift your resampled indicator variables forward in time by one period. This means that we can compare market returns during a month to the value of the indicator as at the previous month end (i.e. our shifted indicators are now valid FORECASTS). Combine the resampled and shifted indicators dataframe with the resampled returns dataframe.
7. Produce (in one figure) a seaborn plot (use the col and row keyword arguments) showing the risk/return relationship for 4 different combinations of long and short indicators. Label the four charts as:
    a. Long +, Short + 'Bull Market'
    b. Long +, Short - 'Correction'
    c. Long -, Short + 'Recovery'
    d. Long -, Short – 'Bear Market'
8. Save your chart to pdf.
9. Assemble a document with your plots, briefly describe your conclusions.
10. You may be asked to show your document in class☺