Take Home Project

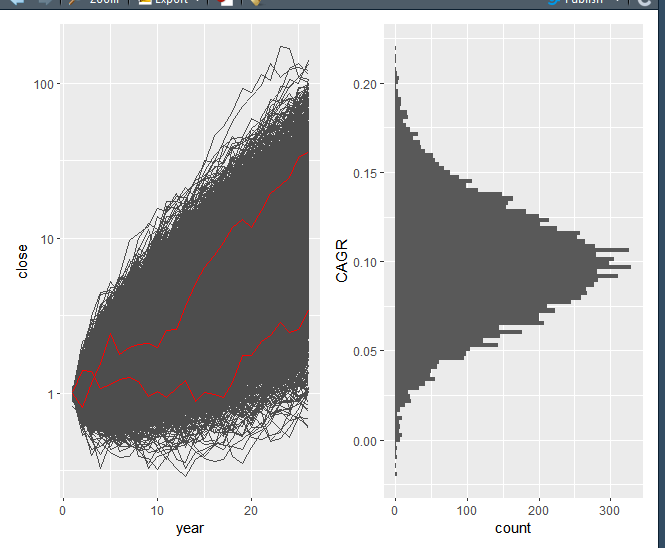
The purpose of this project is to see if you have the ability to learn, understand, and analyze financial data quickly. It is also to see your programming and problem solving skills. The data set is from Kenneth French’s website (<http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/ftp/F-F_Research_Data_Factors_daily_CSV.zip>).   
  
The relevant columns for market returns are circled in the image below:

Table

Description automatically generated

Using the above data-set, please provide the following using either R or Python

1. descriptive statistics and analysis on the distribution of stock market returns ( things like histograms, mean, median, standard deviation)
2. visualize how $1 invested at the beginning would grow over-time
3. bootstrap simulation of 25 years with 10,000 iterations ( or however much your computer can handle in RAM )
   1. sample returns of 25 years from the main dataset with replacement ( feel free to use the daily equivalent of years if you do not group into years. Or better yet, show both )
   2. show a visualization of the bootstrap results by drawing #2 for every possible path of this bootstrap



* 1. calculate the CAGR (<https://www.investopedia.com/terms/c/cagr.asp>) and show a histogram distribution from this bootstrap
  2. the red lines are the 5th and 95th percentiles ( feel free to add anything that you may find meaningful)
     1. perhaps “realized” #2 observations in a different color ( various 25 year returns the market actually experienced )

1. Assuming that stock-market returns are independent and random, how much confidence does one have that investing 100% in the market will result in a CAGR > 0%?
   1. Please use your own data and not the attached image above
2. Using the results of the bootstrap, please explain what might one expect from investing 100% in the stock-market?
   1. If one were to make adjustments to the investment, how might we measure the impact of the adjustment?
      1. Try adjusting the strategy to be 50% Market, 50% Risk-free asset ( use the RF column and Mkt-Rf column to back-calculate the necessary information )
      2. Which strategy is better? Why?
   2. please attach commented code to generate the bootstrap and statistical analysis

Thai Binh Tran

1.

Basic statistics

A screenshot of a computer

Description automatically generated with medium confidence

Free Risk return histogram

Chart

Description automatically generated

Market return histogram.

The graph seems a bit too concentrated in the interver (-0.5;0.5) with nothing outside. However, the minimum and maximum suggest that there are values in that region, but the frequency must be low.

Chart

Description automatically generated

2.

Free Risk return from July 1926 to July 2021

Chart, line chart

Description automatically generated

Market return from July 1926 to July 2021

Chart, line chart

Description automatically generated

The plots suggest that $1 (1 dollar) investing given the free risk rate will be above $20 after roughly 96 years while $1 (1 dollar) investing in the market will return an absurd amount of over $10,000 (detail calculation is in the file). I have some trouble wrapping my head around this because it seems unreal to me (I think my calculation is wrong at some place). I went back to check but I cannot find out why (Maybe it is still somewhere). However, if I understand it correctly, using the yearly compound interest, given the yearly return is 10%. We have:

(1)

In which,

A: the final amount

P: the principal

r: interest rate

n: number of times compounding each period

t: time

A 10% return a year is a bit common to hear, so I guess it might be correct that $1 might return $10,000 given continuously compounding over nearly 100 years.

3.

a.

I try to run 10,000 back to 5,000, but my computer can just run mostly efficiently at roundly 1,000 observations, which is pretty low, but more observations would take a lot more time.

Bootstrap return given daily data

Chart, histogram

Description automatically generated

Bootstrap return given annual data

Chart, histogram

Description automatically generated

The two graphs appear quite similar to me with the most concentrating area around 10. There are some small different regarding minimum and maximum return of frequency of each bin. However, there are only 1,000 observations in each case. If the size of the bootstrap goes to 10,000, the two graphs would approach each other.

**Sidenote 1**:

I calculate return given daily data by (1). However, given each day has a different return. I twist the formula to fit the situation.

In which:

A: is the final amount

P: is the principal

r: is the daily return

i: is the subscript represent each day return

n: is the number of days in 25 years.

**Sidenote 2:**

To calculate return by annual data, I need to derive the annual return first. Annual return given daily rate is the same is

However, the daily rate is different and number of day in the market is not 365 day neither. Therefore, I calculate annual return by

In which:

r: is the daily return

i: is the subscript represent each day return

n: is the number of days in a given year.

Table ‘returnByYear’ obtains the annual return but in proportion form. The column name is a bit misleading. For example, Jan 1926. However, it is actually the return at the end of the year 1926. Here is a sample of the table

Annual return calculated by daily return data

A picture containing text, scoreboard, cabinet

Description automatically generated

After calculating annual return, I can calculate return for 25 years by using (2) but with r, i and n representing annual data instead of daily data. The function in the program has ‘typeOfReturn’ with 1 for free risk return and 2 for market return

b.

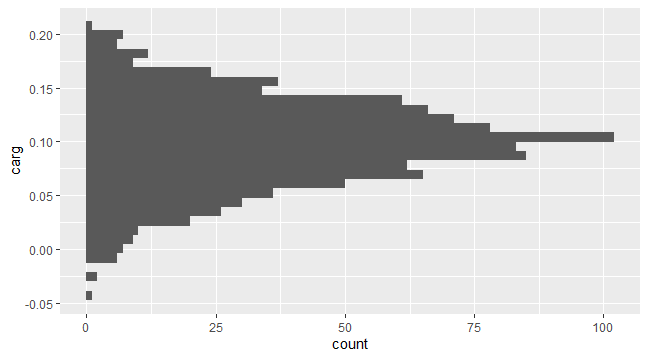
From now, I will use only annual return rate for more efficiency because my laptop is quite old and cannot handle too many calculations.

To continue, I need to create a dataset of $1 growing over any 25 years containing 1000 observations.

Chart, diagram

Description automatically generated

c.



d.

Chart, line chart

Description automatically generated

The 2 red lines are for 5th and 9th percentile

i.

Chart

Description automatically generated

The 3 blue lines are for actual market return given 3 random steaks of 25 years consecutively, which perhaps show that actual return is not that far away from randomly choosing years from bootstrap.

4.

I am not very familiar with calculating hypothesis in bootstrap and with time limitations, I just calculate a student t test given the carg data. The mean and standard deviation of the sample calculated from the bootstrap.

In which:

: is the sample mean

: is the testing hypothesis

s: is the sample standard deviation

n: is the sample size

We can derive a testing hypothesis

With t score is roughly 76 which is above any value in the t table. We can safely say with above 99% confidence that investing in the market in any set of 25 years will result in a CARG > 0%

5.

Return with 50% in market and 50% in risk free

Chart, histogram

Description automatically generated

Return with 100% in the market

Chart, histogram

Description automatically generated

It is a bit difficult to spot the differences between the two graphs. The only different is perhaps the maximum amount. Therefore, I include another table with descriptive statistics.

A screenshot of a computer

Description automatically generated with medium confidence

Similar idea to problem 2, this shows how invest $1 over 25 years period will result differently between 100% market and a mix between risk free and market, using function (2) adjusted for year measurement to achieve the results.

Generally, the results show that investing in the market tends to have better returns overall. On average, returns after 25 years from fully in market is almost double that of a mix between the two. However, in a bad situation, market return might be even lower than a mix bag. Market also has higher volatility too. However, given a long period, investing fully in market tends to be better than that of mix between risk free and market.