Assignment_5_Solution

April 20, 2022

0.1 Problem 1 (100 points)

In this assignment, you will backtest the "momentum" strategy which involves going long on stocks with the highest returns in the past 12 months and shorting the stocks with the lowest returns in the past 12 months, and rebalancing these portfolios every month. Please follow the steps below to implement this backtest:

- **Data** (10 x 2 = 20 points)

    1. Load the "crspm.zip" file into a dataframe called `crsp`. Create a new variable called `mdate` which converts the `date` variable into a monthly period date. Calculate market capitalization (`mktcap`) as number of shares outstanding (`shrout`) times the absolute value of price (`prc`). Create a new variable `mktcap_lag1` that, for each firm, each month, equals the market capitalization (`mktcap`) of that firm in the previous month (assume no duplicates or gaps in the data, i.e. use the `shift()` method for lagging). Keep only `permno`, `mdate`, `ret`, `mktcap_lag1`, and drop all rows that have any missing values in any of those variables. Print the first five and last five rows of the resulting dataset.
    2. Create a new variable called `ret11` which, for each firm, each month, equals the **net** compounded returns of the firm in the past 12 months, excluding the current month (e.g. for December 2010, `ret11` equals the compounded returns from Jan 2010 to Nov 2010). Print a table that gives us the mean and standard deviation of `ret` and `ret11`.

- **Momentum portfolios** (10 x 4 = 40 points)

    3. Drop all rows for which `ret11` is missing. Then, create a variable called `ret11_decile` which, every month, tells us how each firm ranks (which decile) amongst all other firms with respect to their past 11-month returns (`ret`). Firms in decile 1 will form portfolio 1, firms in decile 2 for portfolio 2, etc. Print out a table that tells us how many observations you have for each decile in your dataset.
    4. These portfolios are held for one month, and then they are re-created again (rebalanced) based on the new values of `ret11` for the current month. For this reason, the return of a portfolio (decile) in a given month is based on which firms were in that decile in the *previous* month. To help with this (for steps 5 and 6 below), create a new variable called `portf_nr` which equals the `ret11_decile` that the firm was in the previous month. Print out a table that tells us how many observations you have for each `portf_nr` in your dataset.
    5. Calculate the monthly equal-weighted (EW) returns of each of the 10 portfolios. Calculate the EW returns to the momentum strategy by subtracting returns of portfolio 1 from the returns of portfolio 10. This will be referred to as the "spread portfolio" below.

Store these returns in a new dataframe called `ew_ret` and print the first 5 and last 5 rows of this dataframe.

6. Repeat step 5 using value-weighted (VW) returns. Use `mktcap_lag1` (i.e. the market capitalization at the end of the prior month) as weights. Store these returns in a new dataframe called `vw_ret` and print the first 5 and last 5 rows of this dataframe.

- **Analysis** (10 x 4 = 40 points)

7. Calculate and print the average EW returns for the 10 momentum portfolios as well as the spread portfolio.
   - This should be a 11-by-1 table containing the EW averages
8. Calculate and print the average VW returns for the 10 momentum portfolios as well as the spread portfolio.
   - This should be a 11-by-1 table containing the VW averages
9. Calculate and print risk-adjusted average EW returns (with respect to the Fama-French three-factor model) and their t-statistics for the 10 momentum portfolios as well as the spread portfolio
   - This should be an 11-by-2 table
   - The numbers in this table are "alpha" coefficients from regressing each portfolio's EW returns on the Fama-French three factors (in the first column) and the tstats of these coefficients (in the second column)
10. Calculate and print risk-adjusted average VW returns (with respect to the Fama-French three-factor model) and their t-statistics for the 10 momentum portfolios as well as the spread portfolio
    - This should be an 11-by-2 table
    - The numbers in this table are "alpha" coefficients from regressing each portfolio's VW returns on the Fama-French three factors (in the first column) and the tstats of these coefficients (in the second column)

```
Output for part 1:

           permno    mdate       ret   mktcap_lag1
1         10000.0  1986-02 -0.257143  1.610000e+04
2         10000.0  1986-03  0.365385  1.196000e+04
3         10000.0  1986-04 -0.098592  1.633000e+04
4         10000.0  1986-05 -0.222656  1.517200e+04
5         10000.0  1986-06 -0.005025  1.179386e+04
…           …      …        …            …
2553248   93436.0  2020-08  0.741452  2.666393e+08
2553249   93436.0  2020-09 -0.139087  4.643391e+08
2553250   93436.0  2020-10 -0.095499  4.067015e+08
2553251   93436.0  2020-11  0.462736  3.678235e+08
2553252   93436.0  2020-12  0.243252  5.380286e+08

[2525383 rows x 4 columns]

Output for part 2:
```

```
            ret     ret11
mean   0.011656  0.129892
std    0.191348  0.748724

Output for part 3:

1     229084
2     228849
3     228808
4     228850
5     228905
6     228749
7     228803
8     228855
9     228802
10    229044
Name: ret11_decile, dtype: int64

Output for part 4:

1     223833
2     227185
3     227532
4     227730
5     227686
6     227477
7     227315
8     226930
9     226293
10    225875
Name: portf_nr, dtype: int64

Output for part 5:

portf_nr         1         2         3         4         5         6  \
mdate
1981-01        NaN       NaN       NaN       NaN       NaN       NaN
1981-02  -0.001611  0.018650  0.011000  0.010897  0.012980  0.016318
1981-03   0.073697  0.070473  0.080238  0.082625  0.071313  0.074025
1981-04   0.033108  0.024192  0.020516  0.029522  0.029836  0.027749
1981-05   0.010310  0.011268  0.016465  0.013602  0.026789  0.028721
…            …       …       …         …         …         …
2020-08   0.058369  0.050536  0.066951  0.054317  0.053586  0.043611
2020-09  -0.073705 -0.038189 -0.035999 -0.023008 -0.031162 -0.018343
2020-10  -0.012347  0.030453  0.046157  0.028848  0.014210  0.003614
2020-11   0.404211  0.242672  0.194575  0.185290  0.168104  0.167673
2020-12   0.160956  0.120109  0.099752  0.089930  0.068823  0.058924
```

```
portf_nr         7         8         9        10    Spread
mdate
1981-01        NaN       NaN       NaN       NaN       NaN
1981-02   0.008814  0.007066  0.003615 -0.017493 -0.015882
1981-03   0.068516  0.069465  0.087356  0.079740  0.006043
1981-04   0.028870  0.029622  0.038376  0.031609 -0.001499
1981-05   0.026589  0.022504  0.055683  0.055878  0.045568
…              …         …         …         …         …
2020-08   0.035730  0.041590  0.032555  0.037719 -0.020650
2020-09  -0.023794 -0.023569 -0.012108  0.007355  0.081060
2020-10   0.018308  0.009093  0.010959 -0.025737 -0.013390
2020-11   0.145358  0.165401  0.165987  0.216825 -0.187386
2020-12   0.080126  0.072444  0.099552  0.074695 -0.086260

[480 rows x 11 columns]
```

Output for part 6:

```
portf_nr         1         2         3         4         5         6  \
mdate
1981-01        NaN       NaN       NaN       NaN       NaN       NaN
1981-02  -0.007998  0.034823  0.032598  0.011224  0.025033  0.030649
1981-03   0.143031  0.085449  0.044863  0.033512  0.048550  0.023256
1981-04  -0.011028 -0.010393  0.023165 -0.015613 -0.008582 -0.021044
1981-05  -0.005517  0.030109 -0.009834  0.004325  0.017044  0.002600
…              …         …         …         …         …         …
2020-08   0.042744  0.060993  0.055256  0.039988  0.063935  0.040496
2020-09  -0.076537 -0.069751 -0.054620 -0.044092 -0.025421 -0.009056
2020-10  -0.045062  0.015458  0.030485  0.013707 -0.003825 -0.031272
2020-11   0.373683  0.275725  0.219271  0.179983  0.143367  0.128766
2020-12   0.088073  0.065939  0.065340  0.058779  0.067205  0.030568

portf_nr         7         8         9        10    Spread
mdate
1981-01        NaN       NaN       NaN       NaN       NaN
1981-02   0.031447  0.000641  0.024081 -0.001883  0.006115
1981-03   0.034338  0.043392  0.060206  0.090606 -0.052425
1981-04  -0.017775 -0.029303 -0.002552 -0.003691  0.007337
1981-05   0.000368  0.020718  0.035657  0.062959  0.068476
…              …         …         …         …         …
2020-08   0.050129  0.071553  0.076713  0.165289  0.122546
2020-09  -0.020338 -0.024782 -0.042851 -0.066764  0.009772
2020-10  -0.022938 -0.018253 -0.008879 -0.048785 -0.003723
2020-11   0.112797  0.104732  0.075896  0.125390 -0.248294
2020-12   0.049820  0.020490  0.049229  0.067837 -0.020236

[480 rows x 11 columns]
```

Output for part 7:

```
portf_nr
1         0.011980
2         0.008312
3         0.009356
4         0.010551
5         0.010858
6         0.012008
7         0.012817
8         0.013812
9         0.014665
10        0.015816
Spread    0.003836
dtype: float64
```

Output for part 8:

```
portf_nr
1         0.000556
2         0.004428
3         0.006921
4         0.009850
5         0.009972
6         0.009068
7         0.010314
8         0.011152
9         0.010892
10        0.014960
Spread    0.014404
dtype: float64
```

Output for part 9:

```
           EW_alphas  EW_alpha_tstats
portf_nr
1         -0.002583        -0.823146
2         -0.004644        -2.644778
3         -0.002771        -2.266313
4         -0.000842        -0.878880
5          0.000096         0.133591
6          0.001456         2.181240
7          0.002512         3.811645
8          0.003563         5.144844
9          0.004273         5.315209
10         0.004444         3.311954
Spread     0.007027         2.030369
```

Output for part 10:

portf_nr	VW_alphas	VW_alpha_tstats
1	-0.015832	-5.462913
2	-0.009966	-4.847471
3	-0.005650	-3.628101
4	-0.001643	-1.408868
5	-0.000825	-1.020090
6	-0.001114	-1.604570
7	0.000491	0.654656
8	0.001369	1.963817
9	0.001285	1.355809
10	0.004613	3.012605
Spread	0.020446	5.380577