

Файл “**data2.xlsx**” містить щомісячні дані про фінансові показники та ETFs із серпня 2011 року по вересень 2021 року. Використовуючи ці дані, виконати наступні завдання:

1. Для даних з вкладки “*hedge\_fund\_series*” порахувати такі статистики:

- (a) математичне сподівання;
- (b) волатильність;
- (c) Коефіцієнт Шарпа. *Технічно коефіцієнт Шарпа рахується як відношення математичного сподівання до волатильності надлишкового прибутку. Тут ми використовуємо total returns, але умовно будемо називати цей mean-volatility ratio коефіцієнтом Шарпа.*

Порахуйте ці статистики за період 1 рік (annualize).

2. Для даних з вкладки “*hedge\_fund\_series*” порахувати статистики, пов’язані з кінцевими ризиками (tail-risk):

- (a) Асиметрія;
- (b) Надлишковий ексцес (понад 3);
- (c) VaR (0,05) – 5-ий квантиль прибутковості;
- (d) CVaR (0,05) – середнє значення прибутків на рівні або нижче 5-го квантиля;
- (e) Максимальна просадка – знайдіть дати максимального/мінімального/відновлення в межах періоду максимальної просадки. (maximum drawdown)

Ці статистики не потрібно перераховувати за рік.

3. Для даних з вкладки “*hedge\_fund\_series*” реалізувати регресію кожного індекса з SPY (він знаходиться у вкладці “merrill\_factors”). Додайте вільний член (intercept). Порахувати наступні статистики на основі регресії:

- (a) Market Beta
- (b) Treynor Ratio
- (c) Information ratio

Перерахуйте ці статистики за період 1 рік (annualize).

4. Відносна продуктивність. На основі попередніх статистик, дайте відповіді на питання:

- (a) які ви помітили відмінності між SPY та серіями хедж-фондів?
- (b) який фонд на Вашу думку працює краще: HDG чи QAI?

5. Обчислити кореляцію цих активів.

- (a) Відобразіть кореляцію у вигляді heatmap.
- (b) Які ряди мають найвищу та найнижчу кореляції?

6. Використовуючи лінійну регресію, відтворіть індекс HFRI за шістьма факторами, переліченими у вкладці "merrill\_factors". Додайте вільний член:

$$r_t^{\text{hfri}} = \alpha^{\text{merr}} + x_t^{\text{merr}} \beta^{\text{merr}} + \epsilon_t^{\text{merr}} \quad (1)$$

$$\hat{r}_t^{\text{hfri}} = \hat{\alpha}^{\text{merr}} + x_t^{\text{merr}} \hat{\beta}^{\text{merr}}$$

Друге рівняння є лише позначенням для гіпотези лінійної регресії.

- (a) Чому дорівнює вільний член та бета коефіцієнти?
- (b) Чи реалістичні розміри позицій у бета-версіях чи вони вимагають величезних long і short позицій?
- (c) Чому дорівнює R-squared?
- (d) Обчисліть волатильність  $\epsilon_t^{\text{merr}}$  (похибок).

7. Розглянемо авторегресію. Починаючи з  $t = 61$  місяця вибірки, виконайте наступне:

- A. Використайте дані за попередні 60 місяців, щоб натренувати лінійну регресії (1). Це дасть оцінки параметрів регресії, залежних від часу  $\tilde{\alpha}_t^{\text{merr}}$  та  $\tilde{\beta}_t^{\text{merr}}$
- B. Маючи ці параметри, а також залежну змінну  $x_t^{\text{merr}}$ , обчислити значення 
$$\tilde{r}_t^{\text{hfri}} = \tilde{\alpha}_t^{\text{merr}} + (x_t^{\text{merr}})' \tilde{\beta}_t^{\text{merr}}$$
- C. Перейдіть до  $t = 62$ , але тепер використовуйте  $t = 2$  до  $t = 61$  для оцінки. Зробіть те саме, що у пунктах A та B, і продовжуйте цей процес для усього ряду даних. Таким чином, ми запускаємо ковзаючу 60-місячну регресію для кожного моменту часу.

Наскільки добре працює реплікація поза вибіркою по відношенню до цільової змінної?

8. Ми оцінили реплікації з вільним членом. Спробуйте оцінити лінійну регресію на повній вибірці, але цього разу без вільного члену  $\hat{\beta}_t^{\text{merr}}$

$$\bar{r}_t^{\text{hfri}} = \bar{x}_t^{\text{merr}} \beta^{\text{merr}}$$

Порахуйте:

- 1) коефіцієнт бета регресії. Порівняти з  $\hat{\beta}_t^{\text{merr}}$

- 2) середнє значення передбаченого значення  $\overline{r}^{\text{hfri}}_t$ . Порівняйте з середнім значенням HFRI?
- 3) Обчислити кореляції передбаченого значення  $\overline{r}^{\text{hfri}}_t$  з HFRI.

9. Merrill обмежує вагу кожного активу в регресії для реплікації HFRI. Спробуйте обмежити ваші ваги.

(a) Використайте невід'ємні найменші квадрати (NNLS) замість OLS. *Можна використати `LinearRegression` у `scikit-learn` із параметром `positive=True`.*

(b) Використайте Generalized Linear Model для того щоб накласти окремі інтервальні обмеження на кожен бет. *Можна використати GLM з `statsmodels`.*

10. Розкладемо кілька інших рядів, щоб побачити, чи вони поведуться так, як впливає з назви.

(a) Побудуйте лінійну регресію для HEFA використовуючи ті самі фактори, що ми брали для передбачення HFRI.

(b) Розкладіть TRVCI за тими самими факторами, що ми брали для передбачення HFRI. Індекс TRVCI відстежує фонди венчурного капіталу. Який з наших факторів найкраще описує венчурний капітал?

(c) TAIL — це ETF, який відстежує SPY, але також купує пут-опціони для захисту від спадів на ринку. Обчисліть статистики в питаннях 1-3 для TAIL. Чи має він високу кореляцію з SPY одночасно з меншим tail risk?