**Exercise 3: Fama French Factors**

**Learning Objective**

By the end of this exercise, you should be able to:

* Work with multiple databases through linking table
* Construct SMB/HML style factors

**Part 1. Learn WRDS sample code of Fama-French 3-factor**

In this part, you are going to run the WRDS sample code of constructing two of the Fama-French three factors: SMB, and HML. The sample SAS code is from:

https://wrds-www.wharton.upenn.edu/pages/support/applications/risk-factors-and-industry-benchmarks/fama-french-factors/

For Python users: see ff3\_v2.ipynb

WRDS要点摘要

文本

中度可信度描述已自动生成Futhermore, to be included in the sample for any period July t through June t +1 the firm must have a CRSP stock price for December of year t -1 and June of year t , plus its book value of common equity for year t -1 based on Compustat data. Below we detail the procedure to recreate monthly SMB and HML factors.

**Book Equity** is defined as the Compustat book value of stockholders' equity plus balance sheet deferred taxes and investment tax credit (if available) minus book value of preferred stock.

There were cases when the same firm (permco) had two or more securities (permno) on the same date. For the purpose of ME for the firm, we aggregated all ME for a given permco, date. This aggregated ME was assigned to the CRSP permno that has the largest ME.

SMB is a size factor that is constructed solely based on firms’ market capitalization. HML is a value factor constructed using firms’ book-to-market ratio where the “book” refers to book value of equity and the “market” refers to the market value of equity (market capitalization). We often use the Fama-French three factors as benchmark to compute risk-adjusted returns for assets. At this stage, we only calculate these factors, and we will learn how to use the factor model to conduct risk-adjustment for portfolio-performance in later exercise.

To calculate book-to-market ratio, you need book equity computed based on firms’ accounting data from Compustat database. This is another important database you will often use in the empirical works. An important note is that firm identifiers are different in Compustat and CRSP database. Specifically, firm identifier is PERMNO in CRSP database and it is GVKEY in Compustat database. In order to use CRSP and Compustat data collectively, you need to know the corresponding GVKEY for each PERMNO (in other words, to link PERMNO with GVKEY).

WRDS provide a linking table for PERMNO-GVKEY and you can find it on WRDS website following “Get Data” -> “CRSP” -> “CRSP/Compustat Merged” -> “Linking Table”. You will find the usage of the linking table in the WRDS sample code of Fama-French 3-factor. Please note that, in CRSP/Compustat Merged database, PERMNO is already linked with GVKEY. Using this database, you can download firm accounting data with PERMNO as firm identifier. However, I still suggest downloading accounting data from “Compustat-North America-Annual Updates” and merge with CRSP MSF through linking table. The benefit is that once you are familiar with linking firms in CRSP and Compustat, you will be able to link other databases as well since the linking procedure is similar. For example, if you do some research with analyst data from I/B/E/S database, you will encounter another firm identifier named IBES-TICKER, and you will link IBES-TICKER with CRSP PERMNO in a similar manner.

An additional note in the calculation of HML factor is that, Fama and French impose **a six-month gap** between the accounting data (book equity) and the HML portfolio formation. In particular, **形成期**Fama and French form portfolios at the end of June in year *t*, with the accounting data from year t-1 end (more precisely, fiscal year ending in the calendar year t-1). 在每年六月底利用去年年底的财会数据**持有期**The portfolios are held from July year t to June year t+1, and are **rebalanced annually at the end of each June**. Why not form portfolios at the beginning of each January? This is because firms’ accounting information is not readily available in January. It is important to note that our portfolios (especially for trading strategies) should be formed with public information. By imposing a six-month gap between accounting data from annual report and portfolio formation, we can ensure that the book-to-market ratio is publicly available. Similarly, when we are constructing trading strategies based on firms’ accounting data from **Quarterly Report**, it is standard to impose a one-quarter gap between accounting data and portfolio formation.

阅读文档和基本信息花费20min

正式开始跑代码30min网太烂了跑不了😅

逻辑：

SMB需要 firm’s market value of equity, HML 需要firm’s book value of equity/ firm’s market value of equity。BE需要用到年度财务报表数据，年度财报数据涉及t-1的问题，今年内都用去年年末的数据。

**Part 2. Replication**

First, you need to **replicate monthly returns of SMB and HML factors during January 1970-December 2023.** To assess your replication work, download the Fama French three factors from Professor French’s website and **check the correlation** between your SMB/HML with French’s SMB/HML during January 1970-December 2023. Finally, you should put your monthly SMB/HML together with SMB/HML from French’s data library in a excel file and show me the correlation coefficient. Make sure the correlation is **at least 0.95/0.9** for SMB/HML factors before submitting the results (Ideally you could replicate both factors with correlation above 0.99).

**Part 3. Extend to Fama-French (2015) five factors**

As an additional exercise, please try to replicate Fama-French (2015) five factors (**monthly** time-series). You can find download the original five factor data from WRDS or Prof. French’s website and check the correlation with your replication.