# The final assignment for 2021-2022, for 2023 please see the moodle

In this assignment you will be exploring the factor models in different ways.

## A short introduction to factor models

First we have a very brief review of the problem. In this problem, we consider $n$ major factors that can explain most of the variability of the asset returns. So let us consider factors $F_1$,..., $F_n$. Let us also consider the return of a specific asset that is given by $r$. Then we consider the following regression

$$r = \beta_1 F_1 + \ldots + \beta_n F_n + \epsilon \ .$$

There are several factor models that can be used to explain the return of the assets.

### CAMP

This is a famous model that is based on the Markowitz problem. The model explains that we are only rewarded for the systematic risk and not the risk that can be diversified. In the CAPM model, we only consider one factor which is the market portfolio. For that reason, we can consider the market index. So similar to the first assignment consider 10 assets from different groups, the monthly data. We consider the S&P500 index as the market portfolio.

- You must run 10 regression for the 10 assets and find the so-called $\beta$ from

  $$r - r_f = \beta(r_m - r_f) + \epsilon.$$

  Report the coefficients, performance measures and also compare the 10 assets in terms of their $\beta$.

- In addition, use the same analysis for a portfolio of equally-weighted 10 assets and report coefficients, and the performance measures. Compare them with the previous 10 assets.

## Fama and French 3 factor model

Fama and French are two economists who have developed multifactor models. In particular, they have developed a three-factor model based on three factors that they introduce on the French webpage:

$$r - r_f = \beta_1(r_m - r_f) + \beta_2 SMB + \beta_3 HML + \epsilon.$$

Here

- $r_f$ is the risk free rate
- $r_m$ is the market return
- $SMB$, is a portfolio of "**S**mall [market capitalization] **M**inus **B**ig"
- $HML$, for "**H**igh [book-to-market ratio] **M**inus **L**ow"

The data for all factors are accessible through Keneth French webpage that is provided below.

## Logistic regression

Now we study a classification problem by using logistic regression by considering the following changes. Let us introduce $r_l$ as follow

$$r_l = 1 \; if \; r > r_f \; , and \; r_l = 0, \; otherwise.$$

It means $r_l$ indicates the returns that are larger than risk free $r_f$ by labeling them as 1 and the non-positive returns as 0. Now using the logistic regression, regress $r_l$ by using two different set of explanatory variables.

1. Consistent with the CAMP just use S&P500 index
2. Consistent with Fama and French 3 factor modes, use $(r_m - r_f), \; SMB, \; HML.$

Like the previous part for the two regressions:

- Report the coefficients, performance measures and also compare the 10 assets in terms of their coefficients.

- In addition, use the same analysis for a portfolio of equally-weighted 10 assets and report coefficients, and the performance measures. Compare them with the previous 10 assets.

Summary of the task

| | CAPM | FF3 |
|---|---|---|
| Regression | On $r$ of each single 10 assets & the equally weighted portfolio | On $r$ of each single 10 assets & the equally weighted |
| Logistic regression | On $r_l$ of each single 10 assets & the equally | On $r_l$ of each single 10 assets & the equally |

## Webpages:

French website to download data for $r_f$, $r_m$, $SMB$, $HML$

https://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html

Yahoo finance for choosing the 10 assets
https://uk.finance.yahoo.com/

# The assignment

You have to complete the following tasks in a few steps

1. Find and report the results of the regression for 10 assets as well as the equally weighted portfolio, including the coefficients and the performance measures, and compare them
2. Find and report the results of the logistic regression for indicator of the excess return of 10 assets as well as the equally weighted portfolio, including the coefficients and the performance measures, and compare them

## Details

- First you have to choose 10 assets from yahoo finance with at least 10 years of monthly prices available. Use a fair share of at least 4 sectors, for instance tech, food, manufacturing, banking/finance, etc
- Find the data for the factors that you need including the S&P500 index from yahoo finance and Fama/French 3 factors from French webpage

- By using Pandas create a data frame that includes the data set
- Write a function that finds and returns the coefficients and performance measures

You have to write a report, within the given world counting scope, to explain how this is archived. If necessary you can do it by presenting part of the code in the report. In addition, you need to present your code which will be tested to validate your results.