Fama and French factors

Angelo Calabrese
Jacopo Liberati

22 December 2023

Download the Fama-French portfolios and the Fama-French factors from Kenneth French's home- page. Use the equally weighted portfolios, monthly data for 1960-today. Construct excess returns of the portfolios by subtracting the risk-free rate. Reproduce the figures below (taken from J.H.Cochrane, asset pricing, chapter 20). Run CAPM regressions on the excess returns, where Mkt-RF is taken to be the market excess return. Plot the results (a scatter plot of predicted against actual average excess returns). Run Fama-French regressions (adding SMB and HML to the CAPM regressions, which already are excess returns). Plot the results (a scatter plot of predicted against actual average excess returns. Finally, run a Fama-McBeth regression to obtain the market price of risk for the three factors and error terms. Carefully watch your econometric procedure, especially with regard to the problem of overlapping observations.

INTRODUCTION TO THE ECONOMIC PROBLEM:

The project aims to analyze and compare the returns of 25 portfolios created by Fama and French, based on two key variables: Market Equity, classified from 1 to 5, and Book to Market, also classified from 1 to 5. The methodological approach involves conducting two different regressions: the Capital Asset Pricing Model (CAPM) and the Fama-French Model.

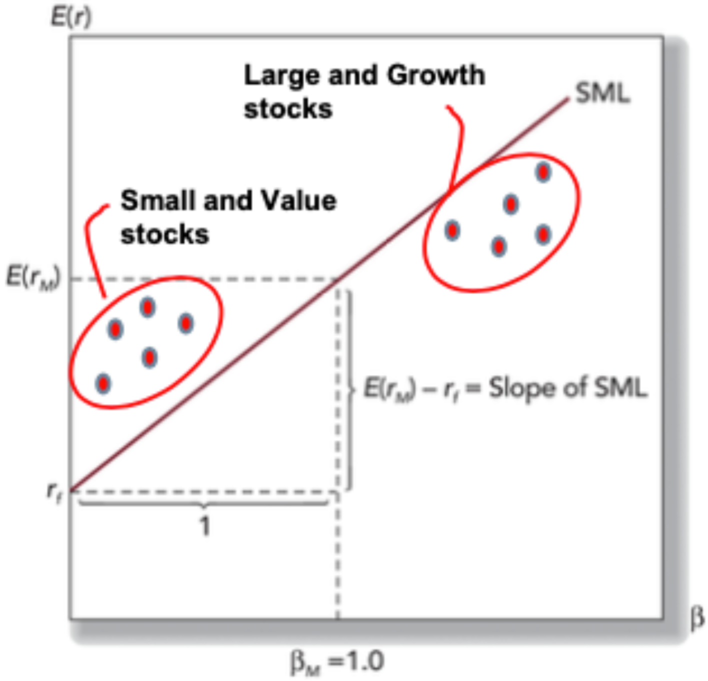


Figure 1

This is the SML of CAPM, where we have:
- CAPM beta on the X Axis
- Expected Return on the Y Axis

The Empirical Observation is that the returns of some securities are not aligned on the SML, but are systematically different.
In particular:

- **Small Companies**(those with a small Market Equity) tend to have returns that are higher (above) than the prediction of the Security Market Line.
Their Expected Return is higher than predicted by the CAPM and it means they are Underpriced.
- **Large Companies**s (those with a large Market Equity) tend to have returns that are smaller (below) than the prediction of the Security Market Line.
Their Expected Return is lower than predicted by the CAPM and it means they are Overpriced.
We get similar results for:
- **Value Companies** (like Small Companies)
Companies for which the Price is LOW relative to fundamental value.
We look at the BOOK / MARKET Ratio and we can see it's HIGH. A company for which this ratio is small (below 1), means that they are trading at a price that is below what the value of the company is in the books.
It means the market is discounting these companies a lot.
- **Growth Companies** (like Large Companies)
Companies for which the Price is HIGH relative to fundamental value. Sometimes, we look at the BOOK / MARKET Ratio and we can see it's LOW. A company for which this ratio is large (above 1), means that they are trading at a price that is above what the value of the company is in the books. It means the market is giving more value to the company than what is written in the books, maybe because the market expects the company to grow in the future.

1 Point 1

PART 1 : Download the Fama-French portfolios and the Fama-French factors from Kenneth French's homepage. Use the equally weighted portfolios, monthly data for 1960-today. Construct excess returns of the portfolios by subtracting the risk-free rate.

We started importing data on the returns of 25 portfolios. Portfolios are divided by size and Book to market ratio.
The CSV file was obtained through Excel, deleting all other rows related to different informations on portfolios or different time horizon. We have isolated the monthly returns from the 60s to the present of Equal Weighted portfolios.

```
df_portfolios = pd.read_csv('data3_from60s.csv', delimiter=";")
```

We imported the factors downloaded from the Kenneth French website. And we always cleaned them via Excel

```
df_factors = pd.read_csv('FACTORS.csv')
```

Subsequently, we proceeded to merge the datasets and calculate the excess returns by subtracting the corresponding risk free rate from each return.

```
df_conc = pd.concat([df_portfolios,df_factors], axis=1)
```

	Average Equal Weighted Returns -- Monthly	Mkt-RF	SMB	HML	RF	SMALL LoBM	ME1 BM2	ME1 BM3	ME1 BM4	SMALL HiBM	...	ME4 BM1	ME4 BM2	ME4 BM3	ME4 BM4	ME4 BM5	BIG LoBM	ME5 BM2	ME5 BM3	ME5 BM4	BIG HiBM
0	196001	-6.98	2.09	2.78	0.33	-5.0724	-4.3209	-0.9161	-1.1473	-0.2120	...	-5.9297	-4.4354	-6.5718	-7.5740	-3.0181	-7.1846	-5.3218	-5.4634	-5.8698	-6.3710
1	196002	1.17	0.51	-1.93	0.29	-2.6624	1.8279	-1.2024	-1.2312	0.1868	...	1.5322	1.2533	1.4103	0.3697	1.1865	2.7073	-0.0177	0.3525	-3.6834	-4.3663
2	196003	-1.63	-0.49	-2.94	0.35	-2.5718	-5.4262	-4.0587	-1.8771	-3.1571	...	-1.0845	-0.5644	-4.2525	-4.9659	-6.3761	-0.4542	-2.0169	-1.2930	-3.5961	-6.4523
3	196004	-1.71	0.32	-2.28	0.19	-4.4383	-1.6109	-2.5659	-3.4976	-3.2945	...	-0.4395	-1.3129	-2.5987	-2.0342	-3.0367	-0.2519	-1.1388	-1.8835	-2.8847	-3.8717
4	196005	3.12	1.21	-3.70	0.27	-5.2209	-1.7687	-3.0192	0.8474	0.2879	...	3.1820	1.5590	0.8145	1.7989	0.0827	5.7467	3.9846	0.3955	-0.9206	-1.5458
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...

Figure 2

Through this data, we calculated the average of all excess returns of each portfolio, with the following results:

Portfolio Type	Average excess returns
SMALL LoBM	0.3558823529411765
ME1 BM2	0.726567320261438
ME1 BM3	0.8672164705882355
ME1 BM4	1.026555816993464
SMALL HiBM	1.2768665359477125
ME2 BM1	0.4324367320261437
ME2 BM2	0.7508473202614379
ME2 BM3	0.8742767320261438
ME2 BM4	0.868157385620915
ME2 BM5	0.9965396078431372
ME3 BM1	0.5003355555555555
ME3 BM2	0.790155816993464
ME3 BM3	0.7194341176470589
ME3 BM4	0.8649896732026144
ME3 BM5	0.9803537254901961
ME4 BM1	0.5859216993464053
ME4 BM2	0.6478716339869282
ME4 BM3	0.7235199999999999
ME4 BM4	0.8360932026143791
ME4 BM5	0.8812754248366013
BIG LoBM	0.5269513725490197
ME5 BM2	0.627783137254902
ME5 BM3	0.6403086274509804
ME5 BM4	0.5851849673202614
BIG HiBM	0.7116938562091503

2 Point 2

PART 2 :Reproduce the figures below (taken from J.H.Cochrane, asset pricing, chapter 20).Run CAPM regressions on the excess returns, where Mkt-RF is taken to be the market excess return. Plot the results (a scatter plot of predicted against actual average excess returns).

2.1 Capital Asset Pricing Model

The CAPM regression is a traditional model that identifies the relationship between the expected return of a security and the expected market return overall. In this case, beta represents the measure of a security's sensitivity to market fluctuations. However, the CAPM model may be limited in fully explaining returns as it considers only market risk. The Capital Asset Pricing Model (CAPM) serves as a fundamental framework in asset pricing, linking expected returns to systematic risk,measured by a Beta value We run the CAPM regression model for the different financial portfolios. The code performs a series of linear regressions to estimate the alpha and beta coefficients for each portfolio in the provided list, using excess market yields as an independent variable and excess portfolio yields as a dependent variable.

```
# List of the portfolios:

portfolios = ['SMALL LoBM', 'ME1 BM2', 'ME1 BM3', 'ME1 BM4', '
    SMALL HiBM', 'ME2 BM1', 'ME2 BM2', 'ME2 BM3','ME2 BM4', 'ME2
    BM5', 'ME3 BM1', 'ME3 BM2', 'ME3 BM3', 'ME3 BM4', 'ME3 BM5', '
    ME4 BM1', 'ME4 BM2', 'ME4 BM3', 'ME4 BM4','ME4 BM5','BIG LoBM'
    ,'ME5 BM2','ME5 BM3','ME5 BM4','BIG HiBM']

# Dictionary in which to save the regressions results:
regression_results = {}

# Iterate over all portfolios
for portfolio_of_interest in portfolios:
# Dependent Variable: portfolio excess return
y_capm = df_excess_returns[portfolio_of_interest]
# Independent Variable: market excess return with first column of
    ones
X_capm = sm.add_constant(df_conc['Mkt-RF'])

# Solve the equation using Matrix Algebra
beta_capm = np.linalg.inv(X_capm.T @ X_capm) @ X_capm.T @ y_capm

# Extract Alfas and Betas
alfa_capm = beta_capm[0]
beta_capm = beta_capm[1]

# Save the results in the Dictionary
regression_results[portfolio_of_interest] = {'Alfa': alfa_capm, '
    Beta': beta_capm}

# Print the results
for portfolio_of_interest, results in regression_results.items():
print(f'Results for the portfolio {portfolio_of_interest}:')
print(f'Alfa (): {results["Alfa"]}')
print(f'Beta (): {results["Beta"]}')
print()
```

The results:

Portfolio Type	Alphas	Betas
SMALL LoBM:	-0.3992237800673871	1.3684319800809024
ME1 BM2:	0.0525291128676915	1.2215176098742468
ME1 BM3:	0.25295971675357615	1.1131793918544401
ME1 BM4:	0.4628634831574324	1.0215446317119468
SMALL HiBM:	0.7136205414122452	1.020735758699055
ME2 BM1:	-0.3561682860681048	1.4291399304529422
ME2 BM2:	0.0726371999139717	1.2290781087954228
ME2 BM3:	0.26047865899689016	1.1123481530982848
ME2 BM4:	0.2897524779960242	1.0482073160709755
ME2 BM5:	0.3332705520838007	1.2020013447418867
ME3 BM1:	-0.2540757249961297	1.3671727420984998
ME3 BM2:	0.1573299519682982	1.1468310395950352
ME3 BM3:	0.13606208597846822	1.057208926696685
ME3 BM4:	0.2951131811331025	1.032751797866006
ME3 BM5:	0.3541415747647969	1.1348454156419365
ME4 BM1:	-0.09985900112540436	1.2427978012956535
ME4 BM2:	0.039752462727626565	1.1020566318749334
ME4 BM3:	0.14996298428036836	1.0394217824497627
ME4 BM4:	0.26638216144058097	1.0324519614762173
ME4 BM5:	0.25527564887454174	1.134460542039124
BIG LoBM:	-0.08037131776912883	1.1006132188979298
ME5 BM2:	0.08350531696592539	0.9863609137494778
ME5 BM3:	0.12317048097306198	0.9371773672935069
ME5 BM4:	0.07807987459783557	0.9189950866620619
BIG HiBM:	0.16673577309166568	0.9875937118538739

A quick glance quickly shows how alphas tended to increase as the B-M ratio increased, regardless of size. Moreover, in all cases of smaller Book-to-Market Ratio (LoBM and BM1) the observed alpha was negative.

2.1.1 Alternative way to compute Betas

We also used an alternative way to calculate Alpha and Beta, through the Numpy np.polyfit function. Which confirms the previous results.

```
# Dictionary in which to save values of CAPM Alphas and Betas:
alpha_beta_dict = {}

portfolios = ['SMALL_LoBM', 'ME1_BM2', 'ME1_BM3', 'ME1_BM4', '
    SMALL_HiBM', 'ME2_BM1', 'ME2_BM2', 'ME2_BM3', 'ME2_BM4', 'ME2_
    BM5', 'ME3_BM1', 'ME3_BM2', 'ME3_BM3', 'ME3_BM4', 'ME3_BM5', '
    ME4_BM1', 'ME4_BM2', 'ME4_BM3', 'ME4_BM4', 'ME4_BM5', 'BIG_LoBM'
    , 'ME5_BM2', 'ME5_BM3', 'ME5_BM4', 'BIG_HiBM']

# Iterate over all portfolios
for portfolio in portfolios:
beta, alpha = np.polyfit(df_excess_returns['Mkt-RF'],
    df_excess_returns[portfolio], 1)
```

```
# Save results in the DIctionary
alpha_beta_dict[portfolio] = {'Alpha': alpha, 'Beta': beta}

#Print results
for portfolio, values in alpha_beta_dict.items():
print(f'Portfolio: {portfolio}, Alpha: {values["Alpha"]}, Beta: {
    values["Beta"]}')
```

2.2 Recreate the images

To recreate the images in question, taken from the book J.H.Cochrane, asset pricing, chapter 20, we must plot the excess returns observed versus the beta provided by the CAPM model. Our result is in the figure 2. is very similar to the image to be recreated , and we believe that the differences depend on the time horizon used to calculate it.

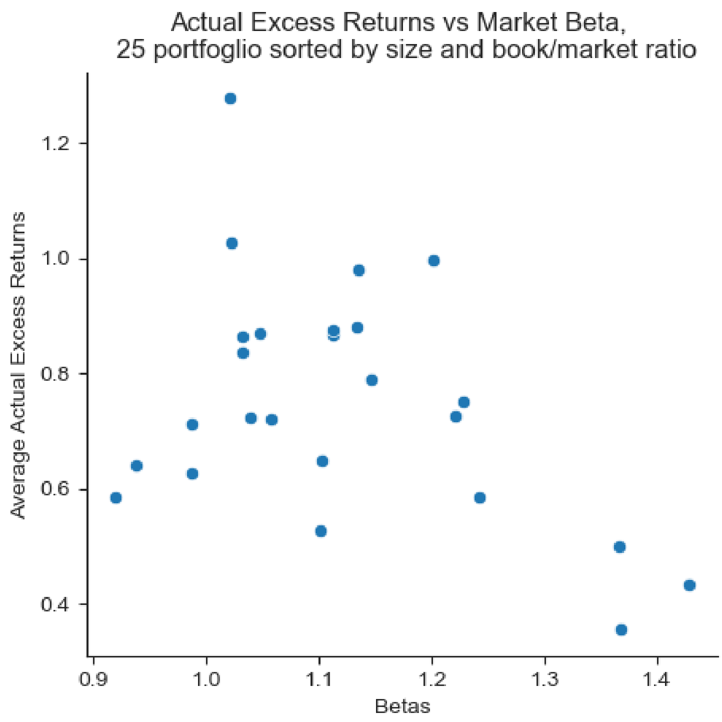
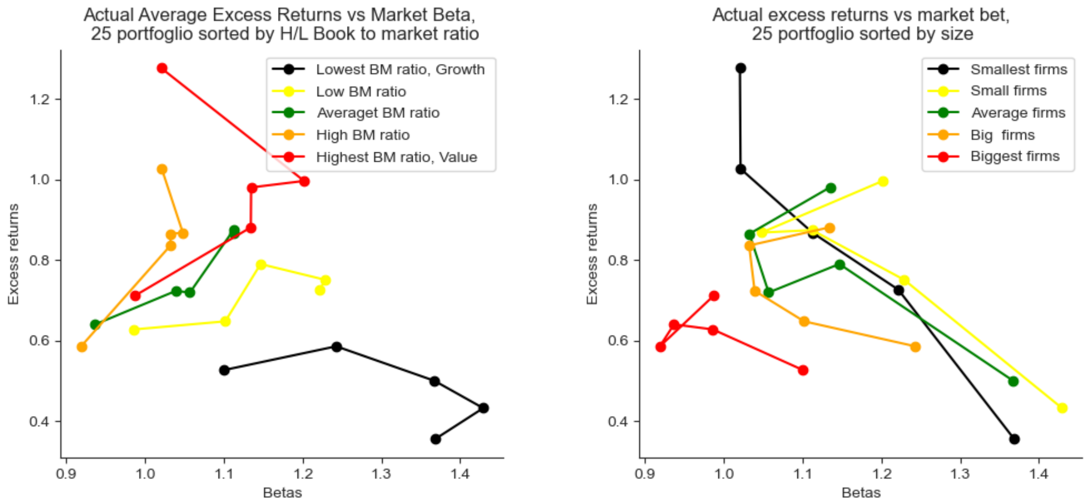


Figure 3: Actual excess returns vs Market Beta

Then we connected portfolios by size and by market to book ratio.



(a) (b)

Figure 4: Graphs recreated

To color and connect the portfolios in question, we created additional lists by subdividing our Betas and Returns by size and Book-to-market ratio. And subsequently plotted again in the plot the points with a different color and a line style, with this type of code :

```
plt.plot(Low_book_market_beta1, LowBMstock1, color = 'black',
    linestyle = '-', marker = 'o', label = 'Lowest BM ratio, 
    Growth ')
```

Where "Low_book_market_beta1" are the betas of all LowBM portfolios, instead LowBM-stock1 are the returns of all LowBM portfolios.

2.3 Actual vs Predicted

The actual excess returns are simply the average of the excess returns observed from the 60s to the present. Instead for the predicted excess return from the model, we used the formula:

$$R_i - r_f = \beta * (R_m - r_f)$$

The axis of the y remains unchanged from the previous graph while the axis of the x must now indicate the excess returns provided by the model.

Portfolio Type	Actual	Predicted
SMALL LoBM'	0.3558823529411770	0.7551061330085650
ME1 BM2'	0.726567320261438	0.6740382073937480
ME1 BM3'	0.8672164705882360	0.6142567538346600
ME1 BM4'	1.026555816993460	0.563692333836033
SMALL HiBM'	1.2768665359477100	0.5632459945354690
ME2 BM1'	0.4324367320261440	0.7886050180942500
ME2 BM2'	0.7508473202614380	0.6782101203474680
ME2 BM3'	0.8742767320261440	0.613798073029255
ME2 BM4'	0.868157385620915	0.578404907624892
ME2 BM5'	0.9965396078431370	0.663269055759338
ME3 BM1'	0.5003355555555560	0.754411280551687
ME3 BM2'	0.790155816993464	0.6328258650251670
ME3 BM3'	0.7194341176470590	0.5833720316685920
ME3 BM4'	0.8649896732026140	0.5698764920695130
ME3 BM5'	0.9803537254901960	0.6262121507254000
ME4 BM1'	0.5859216993464050	0.6857807004718110
ME4 BM2'	0.6478716339869280	0.6081191712593030
ME4 BM3'	0.7235200000000000	0.5735570157196330
ME4 BM4'	0.8360932026143790	0.5697110411737990
ME4 BM5'	0.8812754248366010	0.6259997759620600
BIG LoBM'	0.5269513725490200	0.6073226903181490
ME5 BM2'	0.627783137254902	0.5442778202889780
ME5 BM3'	0.6403086274509800	0.5171381464779190
ME5 BM4'	0.5851849673202610	0.5071050927224270
BIG HiBM'	0.7116938562091500	0.544958083117486

And plotting this results , we get :

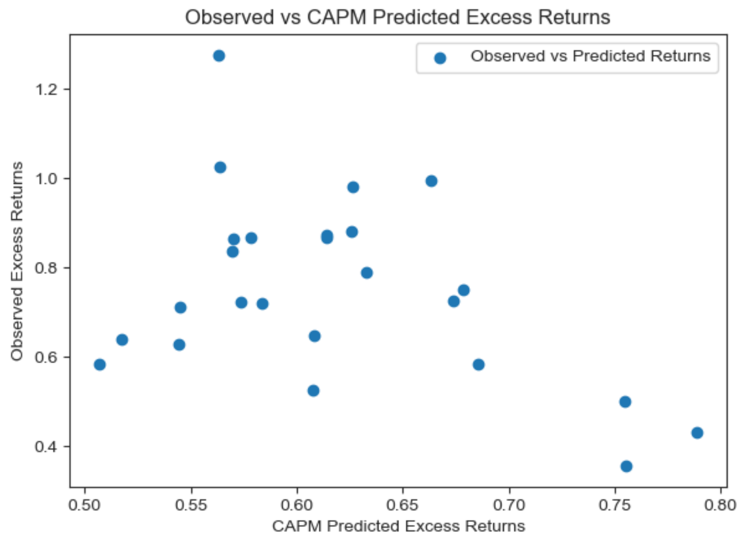
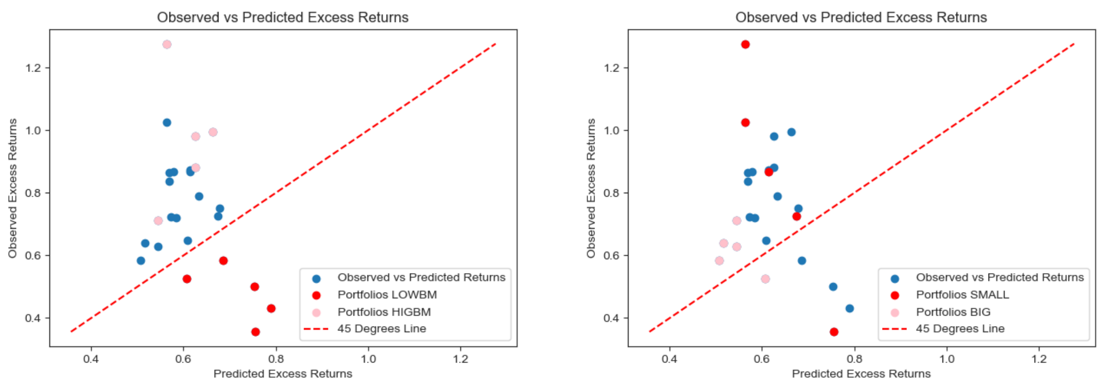


Figure 5

Subsequently we have drawn a 45 degree line, and if the values approach this line it means that the model can correctly predict the returns. In the next 2 figures we highlighted portfolios divided by size and MB ratio.



(a) (b)

Figure 6: Actual vs Predicted for CAPM

While the CAPM boasts simplicity, our findings underscore its lack of accuracy. Notably, LoBM portfolios, expected to yield the highest returns according to the CAPM due to their elevated betas, contradict this projection in our observed data. Surprisingly, they even rank among the poorest performers. Instead, in our right-hand chart we see how Small portfolios record very high alphas, a symptom that there are other risks that the single factor "Excess market return" of the CAPM cannot explain. We think it is a very powerful model, but you also need to know its empirical weaknesses before using it. New models have been developed that use multiple factors, such as the Fama French model, to explain these overperformances and underperformances.

3 Point 3

PART 3 :Run Fama-French regressions (adding SMB and HML to the CAPM regressions, which already are excess returns). Plot the results (a scatter plot of predicted against actual average excess returns.

FAMA – FRENCH:
On the other hand, the Fama-French regression is an extension of the CAPM model that incorporates two additional risk factors: in addition to the Market Risk Premium, the SMB (Small Minus Big) and HML (High Minus Low) factors are introduced.
Fama and French say these different expected returns relative to the CAPM prediction (seen above) are due to different sources of risk that the CAPM is not taking into account.
- Value stocks have higher expected returns because they are more risky than what the CAPM predicts.
- Growth stocks have lower expected returns because they are less risky than what the CAPM predicts.
To do this, we introduce 2 new factors, creating a 3 FACTOR MODEL:
1) The FIRST FACTOR is still the MARKET (Rm - Rf)
2) The SECOND FACTOR is the HML (High minus Low), which is a portfolio which is Long in High Book to Market Stocks (Value) and which is Short in Low Book to Market Stocks (Growth).
3) The THIRD FACTOR is the SMB (Small minus Big), which captures again the Expected Return of Small Stocks (Value) minus the Expected Return of Big Stocks (Growth).

The goal of this model is to enhance the ability to explain returns by considering elements beyond simple market risk. The implementation of the Fama-French regression is justified by its anticipated ability to offer greater precision in analyzing returns compared to the CAPM, thanks to the inclusion of three risk factors instead of just one. This more detailed approach should allow for a deeper understanding of investment dynamics and variables influencing the financial outcomes of the analyzed portfolios.

3.1 Fama-French regressions

The Fama-French model extends the CAPM by incorporating additional factors beyond market risk, namely Size (SMB - Small Minus Big) and Value (HML - High Minus Low), providing a more comprehensive framework for explaining asset returns. Similar to the CAPM we run the fama-french regressions.

```
# List of the portfolios

portfolios = ['SMALL_LoBM', 'ME1_BM2', 'ME1_BM3', 'ME1_BM4', '
    SMALL_HiBM', 'ME2_BM1', 'ME2_BM2', 'ME2_BM3','ME2_BM4', 'ME2_
    BM5', 'ME3_BM1', 'ME3_BM2', 'ME3_BM3', 'ME3_BM4', 'ME3_BM5', '
    ME4_BM1', 'ME4_BM2', 'ME4_BM3', 'ME4_BM4','ME4_BM5','BIG_LoBM'
    ,'ME5_BM2','ME5_BM3','ME5_BM4','BIG_HiBM']

# Dictionary in which to save the regression results:
regression_results_ff = {}
```

```
# Iterate over all portfolios
for portfolio_of_interest in portfolios:
# Dependent Variable: portfolio excess return
y_ff = df_excess_returns[portfolio_of_interest]

# Independent Variable: market excess return, SMB, HML with first
    column of ones
X_ff = sm.add_constant(df_conc[['Mkt-RF', 'SMB', 'HML']])

# Solve the equation using Matrix Algebra
beta_ff = np.linalg.inv(X_ff.T @ X_ff) @ X_ff.T @ y_ff

# Extract Alfas, Betas for Mkt-RF, Betas for SMB, Betas for HML
alfa_ff = beta_ff[0]
beta_mkt_rf_ff = beta_ff[1]
beta_smb_ff = beta_ff[2]
beta_hml_ff = beta_ff[3]

# Save the results in the Dictionary
regression_results_ff[portfolio_of_interest] = {
        'Alfa': alfa_ff,
        'Beta_Mkt-RF': beta_mkt_rf_ff,
        'Beta_SMB': beta_smb_ff,
        'Beta_HML': beta_hml_ff
}

# Print the results
for portfolio_of_interest, results in regression_results_ff.items
    ():
print(f'Results for portfolio {portfolio_of_interest}:')
print(f'Alfa (): {results["Alfa"]}')
print(f'Beta_Mkt-RF (): {results["Beta_Mkt-RF"]}')
print(f'Beta_SMB (): {results["Beta_SMB"]}')
print(f'Beta_HML (): {results["Beta_HML"]}')
print()
```

These are the results:

Portfolio Type	Alpha	Beta Market	Beta SMB	Beta HML
SMALL LoBM:	-0.412568152	1.061386856	1.444106275	-0.147640312
ME1 BM2:	-0.042194738	0.975707542	1.303468644	0.090084758
ME1 BM3:	0.084156379	0.928205601	1.151875654	0.309483351
ME1 BM4:	0.244210184	0.868411319	1.091960652	0.451767691
SMALL HiBM:	0.437052294	0.877235094	1.148943879	0.600969252
ME2 BM1:	-0.289754814	1.177568791	1.047571228	-0.31265709
ME2 BM2:	-0.020974237	1.070276669	0.898926514	0.13853694
ME2 BM3:	0.074894847	1.014742669	0.777044434	0.40249167
ME2 BM4:	0.026824152	0.991730663	0.722363604	0.618383215
ME2 BM5:	-0.036254782	1.154877678	0.866004758	0.888070297
ME3 BM1:	-0.166124256	1.171688638	0.750287997	-0.333025173
ME3 BM2:	0.063513985	1.057842685	0.576257934	0.180132469
ME3 BM3:	-0.04276591	1.026004374	0.457959034	0.424828948
ME3 BM4:	0.034330168	1.030702177	0.466766302	0.645099568
ME3 BM5:	-0.001892867	1.140880825	0.596379871	0.885922324
ME4 BM1:	0.008182928	1.122428365	0.367501598	-0.338607033
ME4 BM2:	-0.056118258	1.092108756	0.21413829	0.23174426
ME4 BM3:	-0.023451738	1.066105134	0.180617808	0.445483126
ME4 BM4:	0.036857812	1.069619818	0.230493839	0.590711869
ME4 BM5:	-0.088207604	1.189278835	0.348651521	0.883527626
BIG LoBM:	0.035659218	1.075883549	-0.089034462	-0.302116438
ME5 BM2:	0.031061926	1.02215334	-0.073653212	0.151037637
ME5 BM3:	-0.004409234	1.010668027	-0.116332625	0.359437476
ME5 BM4:	-0.158766604	1.02694287	-0.084167645	0.650516179
BIG HiBM:	-0.128136813	1.108636372	-0.043009333	0.802030931

To calculate the excess returns of the portfolio we used the formula:

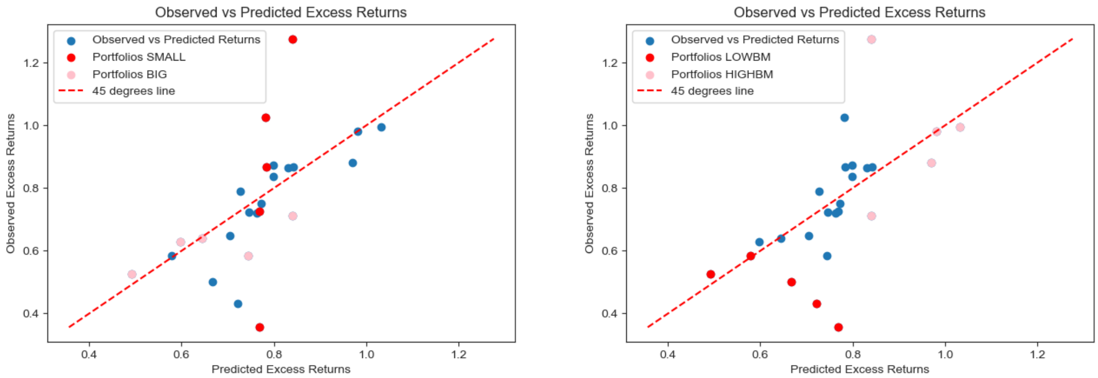
$$R_i - r_f = \beta_M * (R_m - r_f) + \beta_S * (SMB) + \beta_H * (HML)$$

We get :

Portfolio Type	**Actual**	**Predicted**
SMALL LoBM'	0.3558823529411770	0.76845050478962530
ME1 BM2'	0.726567320261438	0.768762057836721
ME1 BM3'	0.8672164705882360	0.7830600911410031
ME1 BM4'	1.026555816993460	0.7823456325472806
SMALL HiBM'	1.2768665359477100	0.8398142416917854
ME2 BM1'	0.4324367320261440	0.7221915463735855
ME2 BM2'	0.7508473202614380	0.7718215569271105
ME2 BM3'	0.8742767320261440	0.799381884569826
ME2 BM4'	0.868157385620915	0.8413332340652087
ME2 BM5'	0.9965396078431370	1.0327943901028664
ME3 BM1'	0.5003355555555560	0.6664598116503683
ME3 BM2'	0.790155816993464	0.7266418318477389
ME3 BM3'	0.7194341176470590	0.762200027654857
ME3 BM4'	0.8649896732026140	0.8306595056335718
ME3 BM5'	0.9803537254901960	0.9822465928515508
ME4 BM1'	0.5859216993464050	0.5777387711142479
ME4 BM2'	0.6478716339869280	0.7039898918709625
ME4 BM3'	0.7235200000000000	0.7469717383073319
ME4 BM4'	0.8360932026143790	0.7992353901891658
ME4 BM5'	0.8812754248366010	0.9694830290425505
BIG LoBM'	0.5269513725490200	0.4912921544750177
ME5 BM2'	0.627783137254902	0.5967212109588308
ME5 BM3'	0.6403086274509800	0.6447178614576737
ME5 BM4'	0.5851849673202610	0.7439515709330982
BIG HiBM'	0.7116938562091500	0.83983066916439

3.2 Actual vs Predicted

And the following Actual vs Predicted values.



(a) (b)

Figure 7: Actual vs Predicted for the Fama French model

From the observed data, through the addition of SMB (Small Minus Big) and HML (High Minus Low) factors, the Fama-French model offers a better explanation of asset returns, incorporating more relevant risk factors than the CAPM, making it a more comprehensive model and more adherent to the empirical reality of the financial markets. Fama and French's multi-factor model offers an explanation of systemic risk through components other than simple market excess return. These factors allow investors' exposure to different macroeconomic risks to be measured and portfolios to be constructed to protect against these risks, such as exposure to value and growth.

Although SMB and HML are not obvious risk factors per se, it is argued that they may be indicators of more fundamental unknown variables. For example, companies with high market value to book value ratios may be more prone to financial difficulties, while stocks of small companies may be more sensitive to changes in market conditions. Another popular explanation is that large firm markets are more efficient,due to the greater availability of information than small firms, which makes it more difficult for alphas to occur.
We point out that empirical approaches such as Fama and French's model, which use these variables as proxies for sources of extramarket risk, have the problem of not clearly identifying factors that cover a significant source of uncertainty.

According to our work , Fama and French's model proves more robust and better at explaining stock returns than the CAPM, offering a more detailed understanding of the factors that influence risk and returns in financial markets.
However, further examination, such as employing the Fama-McBeth regression, may provide deeper insights into the dynamics of these factors in asset pricing.

4 Point 4

PART 4 : Finally, run a Fama-McBeth regression to obtain the market price of risk for the three factors and error terms. Carefully watch your econometric procedure, especially with regard to the problem of overlapping observations

4.1 Fama-McBeth regression

The Fama-MacBeth regression is a technique used to assess the effect of various factors on asset returns. Unlike the Fama-French regression, which is based on portfolios of stocks, the Fama-MacBeth applies directly to individual asset returns. The procedure involves two main phases:
1. Estimation of Time-Series Betas : Perform a regression of the return of each asset against the risk factors for each available time period. This generates a time series of betas representing the exposure of each asset to the risk factors over time.
2. Estimation of Cross-Sectional Averages: Calculate the averages of the betas for each risk factor across all assets and for each time period. These averages provide an estimate of the average relationship between each risk factor and asset returns.

```
# Time series regressions
X = sm.add_constant(factors)
ts_res = sm.OLS(excessReturns, X).fit()
alpha = ts_res.params[0]
beta = ts_res.params[1:]
avgExcessReturns = mean(excessReturns, 0)
# Cross-section regression
cs_res = sm.OLS(avgExcessReturns.T, beta.T).fit()
riskPremia = cs_res.params
```

The Fama-MacBeth regression allows examining how risk factors influence asset returns over time and assessing the consistency of these relationships across different time windows.

It is a useful methodology for analyzing the stability of risk-return relationships across a broad spectrum of assets. The "annualized risk premia" derived from the Fama-MacBeth regression represent the average annual compensation that investors require for assuming a certain level of risk associated with specific risk factors. Specifically:

	Market	**SMB**	**HML**
Premia	5.5428	2.3834	5.2598
Std. Err.	0.5799	0.4316	0.3915

1. Market Risk Premia (MRP) is 5.5428
This represents the compensation required by investors for bearing market risk. A positive MRP indicates that investors demand an additional return for investing in riskier securities that follow the market behavior.
2. SMB (Small Minus Big) is 2.3834

This represents the premium required for assuming the risk associated with differences in company size (small-cap vs. large-cap). A positive value indicates that investors demand an additional return for investing in small-cap stocks compared to large-cap stocks.
3. HML (High Minus Low) is 5.2598
THIS represents the premium required for assuming the risk associated with differences in the book-to-market value of companies (high book-to-market vs. low book-to-market). A positive value indicates that investors demand an additional return for investing in value stocks (high book-to-market) compared to growth stocks (low book-to-market).
In essence, these premia reflect how investors assess and demand compensation for the risk associated with the market, company size, and value-growth factors. A higher premium implies greater recognition of risk and a demand for higher returns. The standard error associated with the risk premia obtained provides an estimate of the variability or uncertainty of these premia; a low standard error suggests greater precision and confidence in the estimate, while a higher standard error indicates more uncertainty.

- Low Standard Error: Indicates that the estimated risk premia are more reliable, and there is greater confidence in their accuracy. Investors can have more confidence that the estimated value accurately reflects the population mean.

- High Standard Error: Suggests that the estimated risk premia are less reliable and may vary more widely from the true population mean. Investors should take this uncertainty into account when interpreting and using these estimates in risk assessment and investment decision-making.

Low standard errors, like those we obtained (0.57, 0.43, and 0.39), indicate that the risk premia estimates obtained from the Fama-MacBeth regression are relatively precise, with less variability around their mean estimates. In other words, there is greater confidence that the estimated values accurately represent the population mean.

We used ChatGPT 3.5 for the following analysis

T-TEST

To go more into the details of the results we obtained for the Alphas and Betas, we have reported also the T-Stat (t-statistic), an indicator of the statistical significance of the coefficients associated with the risk factors (Market, SMB, HML). It assesses whether the estimated coefficient is significantly different from zero, indicating whether the risk factor has a statistically significant impact on asset returns.
- High T-Stat: a T-Stat significantly larger than 2 (or smaller than -2) suggests that the estimated coefficient is statistically significant. In this case, there is more confidence that the risk factor has an impact on asset returns.
- Low T-Stat: if the T-Stat is close to zero, the coefficient may not be statistically significant. This indicates that there is not enough statistical evidence to claim that the risk factor has a significant effect on asset returns.
The presence of high T-statistics for the betas suggests that the effects of the risk factors (Market, SMB, HML) on asset returns are statistically significant. In other words, there is strong statistical evidence that the betas associated with the risk factors differ significantly from zero, indicating that these factors have a significant impact on asset returns. On the other hand, if the T-statistics for the Alphas are close to zero, it may indicate that

the intercept is not statistically significant. This means that, once the risk factors are considered there is no statistical evidence that the intercept is different from zero. In other words, the residual return cannot be considered statistically significant after accounting for the included risk factors in the model.
In practice, this suggests that a significant portion of the variation in asset returns can be explained by the included risk factors, and the residual intercept (alpha) does not contribute statistically significantly to the variability of returns.

J–TEST

Another test we can use to evaluate the validity of the model is the J-test. Specifically, the test checks whether the expected returns predicted by the model match the observed returns. If the value obtained from the J-test is statistically significant, it may suggest that the model fails to adequately explain the asset returns.
In the Fama-MacBeth regression, the J-test is used to assess the model's effectiveness in terms of the predictive ability of covariates with respect to returns. A significant J-test suggests that the explanatory variables in the model have a significant impact on returns. If the J-test value is low and statistically insignificant, it may indicate that the regression model may not be suitable or that the variables used do not significantly explain returns. On the other hand, a high and statistically significant J-test value indicates that the explanatory variables in the model are contributing significantly to the prediction of returns. If the J-test produces a high and significant value like in our case, it suggests that the variables used in the Fama-MacBeth regression model are contributing significantly to the explanation of returns.
The J test in Fama-MacBeth regression is often used to test the Null Hypothesis that pricing errors (alpha) have a mean of zero.
In other words, it tests whether the independent variables included in the model significantly explain expected returns, making pricing errors not significantly different from zero on average. A significant result in the J test may indicate that the risk variables in the model are relevant and contribute to explaining expected returns, while a non-significant result suggests that pricing errors could be considered, on average, as zero.
A J test value of 119.6836 with an asymptotic Chi-Squared distribution of 25 degrees of freedom and a p-value of 0 suggests that we have obtained a statistically significant result. In this context, the J test indicates that the pricing errors (alpha) in our model are not, on average, equal to zero.
The beta coefficients associated with our independent variables are significantly different from zero. This suggests a significant relationship between these variables and expected returns.
Pricing errors, on average, are not equal to zero, as indicated by the relevance of the J test. This could indicate the presence of market inefficiencies or other factors not included in our model that influence expected returns. In summary, a significant result in the J test has implications for the validity of the model and suggests that the risk variables (beta) are contributing significantly to explaining expected returns, and there may be a systematic trend in pricing errors (alpha).

```
J-test:   119.6836
P-value:   0.0000
Size: 1, Value:1   Alpha    Beta(VWM)   Beta(SMB)   Beta(HML)
Coefficients:    -0.4126      1.0614      1.4441     -0.1476
Std Err.          0.1389      0.0380      0.0803      0.0656
T-stat           -2.9693     27.9288     17.9908     -2.2517

Size: 1, Value:2   Alpha    Beta(VWM)   Beta(SMB)   Beta(HML)
Coefficients:    -0.0422      0.9757      1.3035      0.0901
Std Err.          0.1055      0.0319      0.0605      0.0541
T-stat           -0.3999     30.6010     21.5580      1.6664

Size: 1, Value:3   Alpha    Beta(VWM)   Beta(SMB)   Beta(HML)
Coefficients:     0.0842      0.9282      1.1519      0.3095
Std Err.          0.0896      0.0265      0.0602      0.0460
T-stat            0.9397     35.0009     19.1278      6.7228

Size: 1, Value:4   Alpha    Beta(VWM)   Beta(SMB)   Beta(HML)
Coefficients:     0.2442      0.8684      1.0920      0.4518
Std Err.          0.0721      0.0215      0.0515      0.0368
T-stat            3.3865     40.3040     21.2013     12.2666

Size: 1, Value:5   Alpha    Beta(VWM)   Beta(SMB)   Beta(HML)
Coefficients:     0.4371      0.8772      1.1489      0.6010
Std Err.          0.0917      0.0270      0.0592      0.0494
T-stat            4.7641     32.5449     19.3993     12.1563

Size: 2, Value:1   Alpha    Beta(VWM)   Beta(SMB)   Beta(HML)
Coefficients:    -0.2898      1.1776      1.0476     -0.3127
Std Err.          0.0832      0.0222      0.0519      0.0376
T-stat           -3.4839     53.1292     20.1755     -8.3167

Size: 2, Value:2   Alpha    Beta(VWM)   Beta(SMB)   Beta(HML)
Coefficients:    -0.0210      1.0703      0.8989      0.1385
Std Err.          0.0577      0.0200      0.0462      0.0324
T-stat           -0.3635     53.4160     19.4703      4.2820

Size: 2, Value:3   Alpha    Beta(VWM)   Beta(SMB)   Beta(HML)
Coefficients:     0.0749      1.0147      0.7770      0.4025
Std Err.          0.0576      0.0184      0.0562      0.0343
T-stat            1.3014     55.2322     13.8247     11.7219

Size: 2, Value:4   Alpha    Beta(VWM)   Beta(SMB)   Beta(HML)
Coefficients:     0.0268      0.9917      0.7224      0.6184
Std Err.          0.0494      0.0144      0.0386      0.0251
T-stat            0.5425     68.6784     18.7092     24.6206

Size: 2, Value:5   Alpha    Beta(VWM)   Beta(SMB)   Beta(HML)
Coefficients:    -0.0363      1.1549      0.8660      0.8881
Std Err.          0.0605      0.0289      0.0384      0.0291
T-stat           -0.5996     39.8976     22.5254     30.5574

Size: 3, Value:1   Alpha    Beta(VWM)   Beta(SMB)   Beta(HML)
Coefficients:    -0.1661      1.1717      0.7503     -0.3330
Std Err.          0.0767      0.0218      0.0494      0.0392
T-stat           -2.1659     53.8116     15.1882     -8.4896

Size: 3, Value:2   Alpha    Beta(VWM)   Beta(SMB)   Beta(HML)
Coefficients:     0.0635      1.0578      0.5763      0.1801
Std Err.          0.0592      0.0182      0.0435      0.0336
T-stat            1.0722     58.1786     13.2593      5.3689

Size: 3, Value:3   Alpha    Beta(VWM)   Beta(SMB)   Beta(HML)
Coefficients:    -0.0428      1.0260      0.4580      0.4248
Std Err.          0.0583      0.0179      0.0435      0.0317
T-stat           -0.7336     57.2585     10.5256     13.3980
```

Figure 8

Size: 3, Value:4	Alpha	Beta(VWM)	Beta(SMB)	Beta(HML)
Coefficients:	0.0343	1.0307	0.4668	0.6451
Std Err.	0.0584	0.0175	0.0346	0.0339
T-stat	0.5874	58.9511	13.4741	19.0502

Size: 3, Value:5	Alpha	Beta(VWM)	Beta(SMB)	Beta(HML)
Coefficients:	-0.0019	1.1409	0.5964	0.8859
Std Err.	0.0723	0.0249	0.0566	0.0364
T-stat	-0.0262	45.8502	10.5382	24.3193

Size: 4, Value:1	Alpha	Beta(VWM)	Beta(SMB)	Beta(HML)
Coefficients:	0.0082	1.1224	0.3675	-0.3386
Std Err.	0.0649	0.0204	0.0414	0.0302
T-stat	0.1262	55.1476	8.8828	-11.2272

Size: 4, Value:2	Alpha	Beta(VWM)	Beta(SMB)	Beta(HML)
Coefficients:	-0.0561	1.0921	0.2141	0.2317
Std Err.	0.0635	0.0227	0.0541	0.0353
T-stat	-0.8833	48.0427	3.9597	6.5638

Size: 4, Value:3	Alpha	Beta(VWM)	Beta(SMB)	Beta(HML)
Coefficients:	-0.0235	1.0661	0.1806	0.4455
Std Err.	0.0633	0.0215	0.0455	0.0372
T-stat	-0.3702	49.7002	3.9667	11.9763

Size: 4, Value:4	Alpha	Beta(VWM)	Beta(SMB)	Beta(HML)
Coefficients:	0.0369	1.0696	0.2305	0.5907
Std Err.	0.0714	0.0253	0.0373	0.0438
T-stat	0.5162	42.3146	6.1815	13.4983

Size: 4, Value:5	Alpha	Beta(VWM)	Beta(SMB)	Beta(HML)
Coefficients:	-0.0882	1.1893	0.3487	0.8835
Std Err.	0.0853	0.0275	0.0433	0.0426
T-stat	-1.0343	43.2601	8.0460	20.7410

Size: 5, Value:1	Alpha	Beta(VWM)	Beta(SMB)	Beta(HML)
Coefficients:	0.0357	1.0759	-0.0890	-0.3021
Std Err.	0.0458	0.0150	0.0245	0.0236
T-stat	0.7791	71.8721	-3.6399	-12.8088

Size: 5, Value:2	Alpha	Beta(VWM)	Beta(SMB)	Beta(HML)
Coefficients:	0.0311	1.0222	-0.0737	0.1510
Std Err.	0.0538	0.0153	0.0335	0.0300
T-stat	0.5774	66.6396	-2.1990	5.0330

Size: 5, Value:3	Alpha	Beta(VWM)	Beta(SMB)	Beta(HML)
Coefficients:	-0.0044	1.0107	-0.1163	0.3594
Std Err.	0.0551	0.0168	0.0341	0.0287
T-stat	-0.0800	59.9896	-3.4151	12.5396

Size: 5, Value:4	Alpha	Beta(VWM)	Beta(SMB)	Beta(HML)
Coefficients:	-0.1588	1.0269	-0.0842	0.6505
Std Err.	0.0613	0.0207	0.0289	0.0319
T-stat	-2.5883	49.7087	-2.9142	20.4047

Size: 5, Value:5	Alpha	Beta(VWM)	Beta(SMB)	Beta(HML)
Coefficients:	-0.1281	1.1086	-0.0430	0.8020
Std Err.	0.0865	0.0307	0.0498	0.0427
T-stat	-1.4817	36.0611	-0.8638	18.7714

Figure 9

Contents

1 Point 1 **2**

2 Point 2 **3**

2.1 Capital Asset Pricing Model 4

2.1.1 Alternative way to compute Betas 5

2.2 Recreate the images 6

2.3 Actual vs Predicted 8

3 Point 3 **10**

3.1 Fama-French regressions 10

3.2 Actual vs Predicted 14

4 Point 4 **15**

4.1 Fama-McBeth regression 15

We hereby certify that
– We have written the program ourselves except for clearly marked pieces of code
– We have tested the program and it ran without crashing
Angelo Calabrese and Jacopo Liberati