FinalReport

by.Team 404 NOT FOUND

曲星凝 516120910134

陈沛莹 516120910127

孔婷钰 516120910130

李钰　516120910132

黄雅丽 516120990071

I. BRIEF INTRODUCTION

Our main purpose is to predict the expected rate of return for the stock of Alibaba using CAPM, which we have learnt in Financial Management class.

We choose Alibaba as the study object as it is a big company and have a large market share.

II. CAPM

In finance, the **capital asset pricing model** (**CAPM**) is a model used to determine a theoretically appropriate required rate of return of an asset, to make decisions about adding assets to a well-diversified portfolio.

As we assume the financial markets are effective and the investors are effectively dispersed as a whole, the non systemic risk is not a concern. The main risk of individual stocks is systemic risk. If further allowance is assumed to eliminate the systematic risk by dispersion, the expected rate of return for the stock a is

$$E(R_i) = R_f + \beta_i(E(R_m) - R_f)$$

Where:

- $E(R_i)$ is the expected return on the capital asset
- R_f is the risk-free rate of interest such as interest arising from government bonds
- β_i (the beta) is the sensitivity of the expected excess asset returns to the expected excess market returns
- $E(R_m)$ is the expected return of the market

Since the risk-free yield is only referred to the government bond current rate of return, it can be considered as a constant. So the model can be simplified as a linear model of $E(R_i)$ with respect to $E(R_m)$. The point is using the historical data of market returns and stock returns to predict β_i .

III.DATA COLLECTION

We collected the daily closing price of the Alibaba from May 1st 2017 to May 26th 2017, and used the logarithmic rate of return formula Rt=ln (Pt/Pt-1) as the formula for calculating the daily stock return. At the same time, we have collected the three major indexes of the United States: Standard Poole 500 stock index, Dow Jones industrial average, and the Nasdaq composite index during that period of specific values. All these datas are downloaded from financial.Yahoo.com. In the following process, we will use Ridge, Regression and Lasso two linear models to predict the expected return on shares.

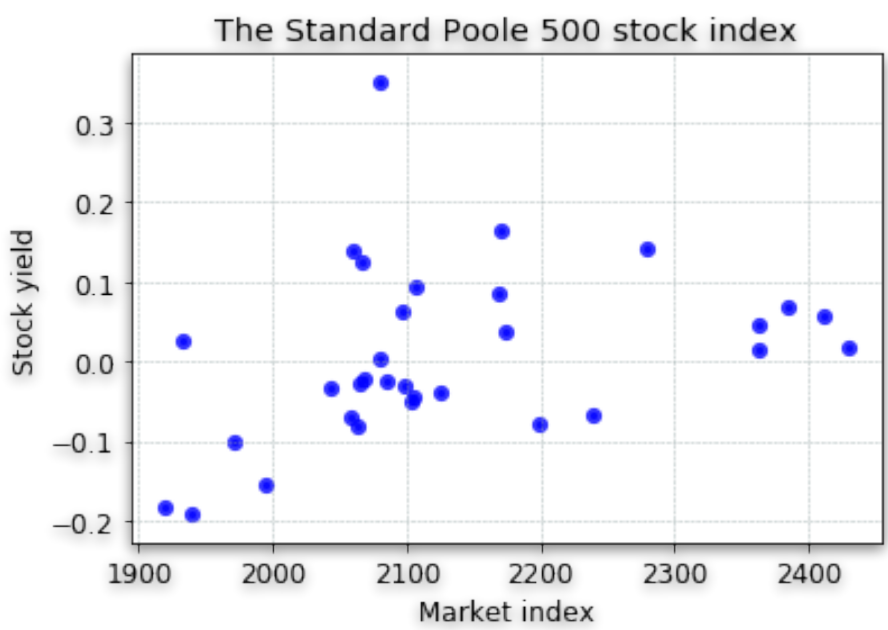
IV.PROGRAMMING

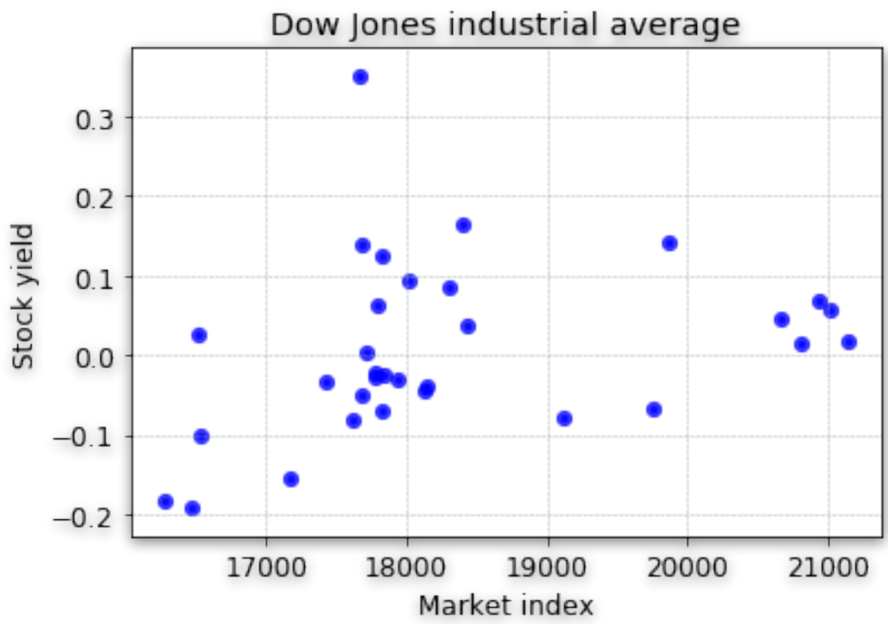
>>> Python library we use:

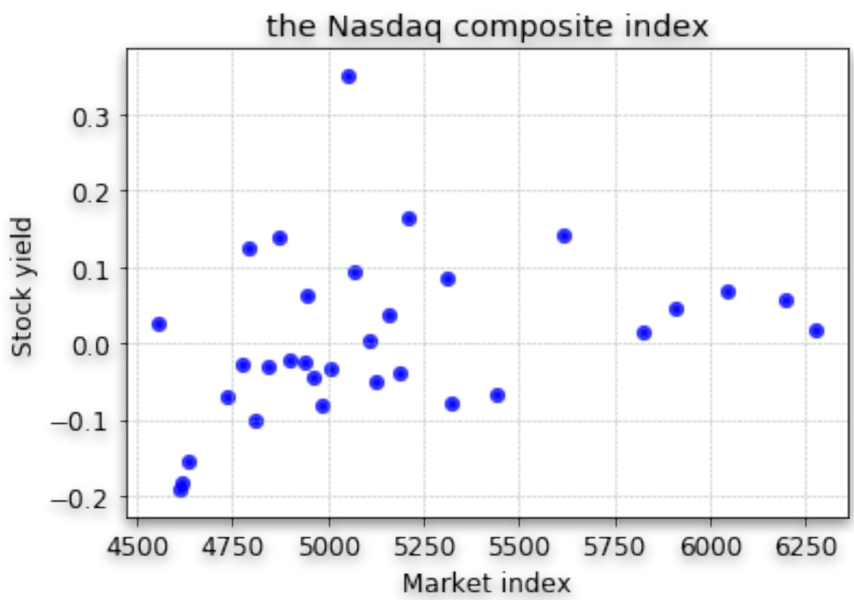
we use linear_model, random, train_test_split, numpy, pandas and matplotlib.pyplot.

>>> Preparation

First, we use the matplotlib, a Python 2D plotting library to draw a scatter diagram showing the relationship between Market Index and Stock Yield, and the results are as follows:







So we can approximately think there is a liner relationship between Market index and Stock yield.

>>> Train Model

We use Ridge Rigression and Lasoo to train the model as these two are liner model.

>>> Test

We can find out that the error variance of Lasso is smaller than Ridge regression and it seems the certainty of forecast is much higher after the contrast. (See more information from our jupyter notebook)

Lasso	The s&p 500 index				Dow-Jones Average				Nasdaq Composite			
Predicting value	0.00309	0.00309	0.00309	0.00309	0.00870	0.00442	0.00860	0.00343	0.01302	0.00186	0.01350	0.00395
Actual value	0.00670	0.01210	0.00290	0.00150	0.00670	0.01210	0.00290	0.00150	0.00670	0.01210	0.0029	0.00150
Error variance	0.0000968				0.0000995				0.00026			

Ridge Regression	The s&p 500 index				Dow-Jones Average				Nasdaq Composite			
Predicting value	0.01426	0.00207	0.01471	0.00460	0.01000	0.00472	0.00990	0.00351	0.01890	0.00113	0.01975	0.00447
Actual value	0.00670	0.01210	0.00290	0.00150	0.00670	0.01210	0.00290	0.00150	0.00670	0.01210	0.0029	0.00150
Error variance	0.00031				0.00012				0.00056			

And the combination of The s&p 500 index and Lasso performs the best in the evaluating.

>>> Output

We can get the beta coefficient from the program, which can reflects the sensitivity of the return on assets to the relative market, which illustratse the extent of the particular asset risk. So it is of great value in the field of economic analysis.

V. Problems in our program

1.We assume that the Rf is constant and can be ignored, which is only true in the short run.But in the short run the data is a bit small to get a perfect result. Our program just need a short time running which is a bit frustrating.

2.The future market index is hard to predict.

VI.Summary

Although the result is far from satisfying and we went through a hard time figuring a proper topic. We enjoy the time brain storming and arguing about the problems we found, from which we have actually learnt a lot about Python and also Financial Management. Specially thanks for our teacher Baoyang, who is so kind and has provided helpful guide and advice during our programming.