Assignment 3.1

The Center for Research in Security Prices (CRSP) database contains end-of-day and month-end price on primary listings for the NYSE, NYSE MKT, NASDAQ, and Arca exchanges along with basic market indices. The SAS dataset named *DSF* is provided on the server (in CRSP directory on Q drive) and it contains all end-of-day information on all the stocks in the CRSP daily security file (DSF) over the 1926 to 2014 period. Additionally the daily market index file (DSI) has also been provided on the server. The objective of this assignment is to understand and get familiar with the daily CRSP stock file and its variables so that they can be used for further analysis in the future. Please complete the following questions in SAS and submit your SAS code and the output PDF file.

1. Subset the data to 2005–2014 (and required variables) only to minimize processing time and do the following analysis on this dataset.

2. Do a proc contents or right click in library, select view columns and check the label to see the variable descriptions

3. Using the *DSF* SAS data file, please compute the descriptive statistics (N, p25, p50, mean, p75, std. dev) for the return, closing price, bid-ask spread, and trading volume for the full sample period. (Always take the absolute value for prices before computation as some closing prices are marked with a negative sign when the stock doesn't trade over an extended period of time.)

4. To examine the time trends of the variables computed in Question 1, compute and plot the daily averages of the variables over the sample period. Analyze and comment on the time-series patterns of the variables.

5. download daily VIX data from `https://www.cboe.com/publish/scheduledtask/mktdata/datahouse/vixcurrent.csv`. Read into a SAS dataset.

6. Plot and correlate variables computed in Question 3 with VIX and Δ VIX (defined as change in daily VIX)

7. Denote $R_{i,t}$ as the daily return of stock i at time t. From the traditional CAPM (Capital Asset Pricing Model) equation :

$$\mathbb{E}\left[R_{i,t} - r_{f,t}\right] = \beta_i \mathbb{E}\left[R_{m,t} - r_{f,t}\right],$$

where $r_{f,t}$ is a risk-free rate at time t, and $R_{m,t}$ is the market return. To reduce some works in computing the excess returns, instead of the CAPM, we estimate the following market model by linear regression:

$$R_{i,t} = \alpha_{i,t} + \beta_i R_{m,t} + \varepsilon_{i,t}. \tag{1}$$

Given the above market model, we can decompose the total volatility of an individual stock into its systematic volatility and its idiosyncratic volatility as the following:

$$\text{var}\left(R_{i,t}\right) = \beta_i^2 \text{var}(R_{m,t}) + \text{var}(\varepsilon_{i,t}). \tag{2}$$

That is, the total volatility of stock i, $TVOL_{i,t} = \sqrt{\text{var}(R_{i,t})}$, can be decomposed into its systematic volatility, $SVOL_{i,t} = \beta_i \sqrt{\text{var}(R_{m,t})}$, and its idiosyncratic volatility $IVOL_{i,t} = \sqrt{\text{var}(\varepsilon_{i,t})}$.

Based on the information above, provide your answers to following questions.

(a) Using Equation (1) and the ten years of data from 2005-2014 in (*DSF*), compute the betas, β_i, by running a simple linear regression of individual stocks returns on the market returns for each stock over the full sample period. The S&P 500 index return (variable SPRTRN in DSI SAS file) can be used as a proxy for the market return.

(b) For each stock, compute the total volatility, the systematic volatility, and the idiosyncratic volatility using Equation (2) and the betas estimated in part (a).

(c) Using your answers in (a) and (b), sort all stocks based on their *systematic* volatility into quintile portfolios. Then examine and comment on the characteristics (say, mean and median) of the daily returns and the daily volumes of quintile portfolios.

(d) Repeat part (c) by sorting all stocks on their *idiosyncratic* volatility, instead of their systematic volatility. Then examine and comment on the characteristics (say, mean and median) of the daily returns and the daily volumes of quintile portfolios.