Finance Data Project

For Data Analysis

Created By

Rudraneel Chakraborty

Project Description

This data project is focused on exploratory data analysis of stock prices. Keep in mind, this project is just meant to practice visualization and pandas’ skills, it is not meant to be a robust financial analysis or be taken as financial advice.

It’s focused on bank stocks and see how they progressed throughout the financial crisis all the way to early 2023.

This project used the YFinance library to get stock information for the following banks:

* Bank of America
* Citi Group
* Goldman Sachs
* JPMorgan Chase
* Morgan Stanley
* Wells Fargo

The Financial Stock Market values for a stock depend on public sentiment as well as the Condition of the Financial Market for that year.

From December 2007 to June 2009, The Great Recession happened in the US, which resulted in stock value crashes. Here to analyse the financial stock data we took 6 Bank Stocks and will check how that was react though out the Years.

DATA DESCRIPTION

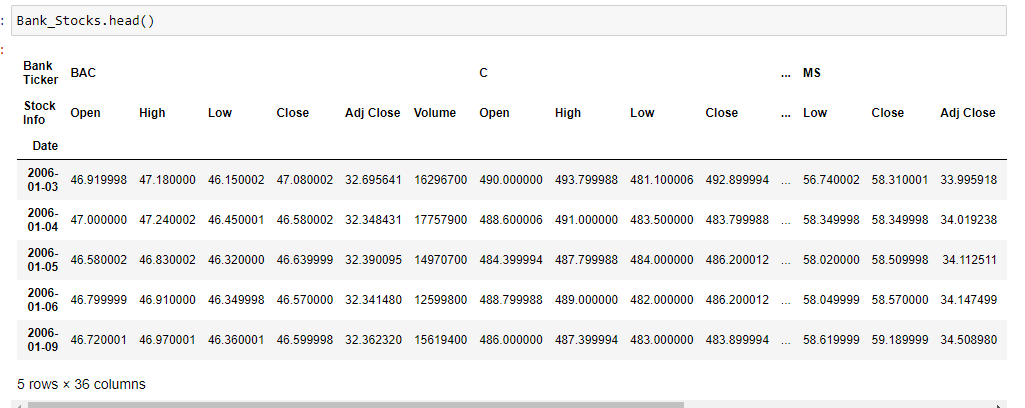
The dataset downloaded through YFinance Library to analyse on Stock Market data from early 2006 to early 2023.

Below following modules are used to do DATA VISULIZATION, ANALYSIS:

* **NUMPY** or numeric python is basically built to work on numerical data
* **PANDAS** is basically built on NUMPY to do certain jobs on the basis of numeric data.
* **MATPLOTLIB.PYPLOT** to implement the graphs such as Bar plot, Histogram, Box plot, Count plot, Bell curve to analyse our data.
* **SEABORN** is used to do critical statistical analysis in very simple &amp; effective way, which gives better results comparing to MATPLOTLIB.
* **DATETIME** module supplies classes for manipulating dates and times.
* **PLOTLY** graphing library makes interactive, publication-quality graphs.
* **Cufflinks** binds the power of plotly with the flexibility of pandas for easy plotting.
* **YFinance** is a Python library that provides convenient access to the Yahoo Finance API.

CONTENT OF DATASET

The dataset contains several parameters which are considered important during analysing. The parameters are shown in the data which is given below:



The Data picked from 2006-01-01 to 2023-01-01, where the ticket\_list contains the ticker for 6 banks:

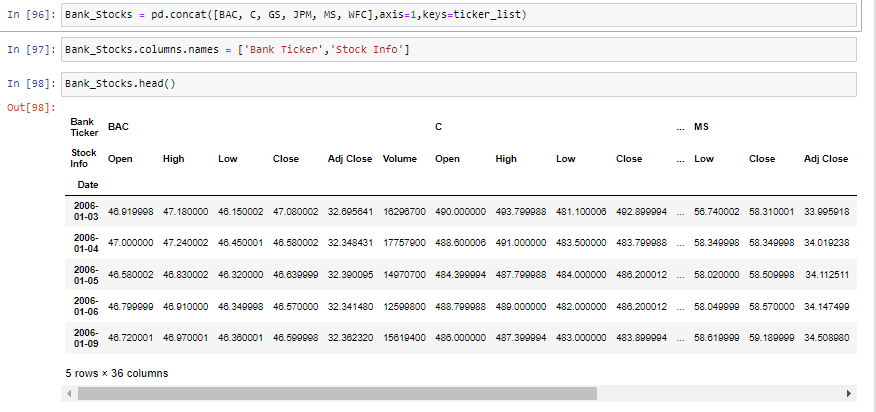
**start = dt.date(2006,1,1)**

**end = dt.date(2023,1,1)**

**ticker\_list = ['BAC','C','GS','JPM','MS','WFC']**

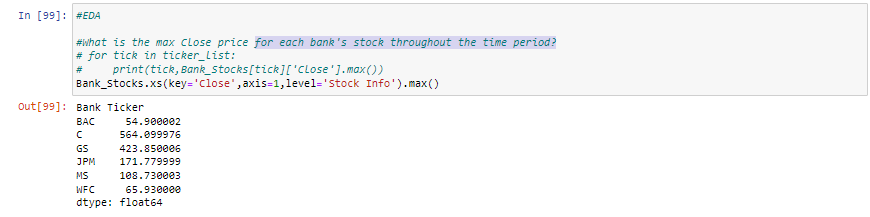
Once the data has been downloaded for those tickers on mentioned start date and end date then Pandas Concat Function called to make those in a single DataFrame by passing the axis value as 1, and the key would be ticker\_list.





Exploratory data analysis (EDA)

In EDA, we will check the values for the stocks on little precisely. So, let’s first have a look the max Close price for each bank's stock throughout the time period.



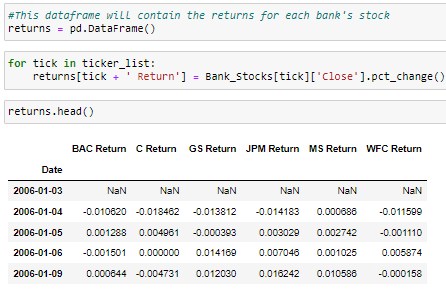
As, we can see from 2006 to 2023 the mentioned Values are the Maximum price of Closer on those Stock.

The Values showing about is in Doller and the whole analysis happen on base of dollar as we are looking for US Stocks Data.

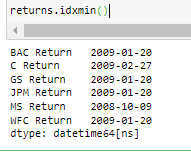
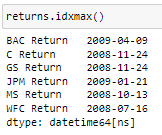
Now, we will check returns for each bank’s stock. The Formula to calculate the return is below:



Below is the same functionality implemented though Pandas.



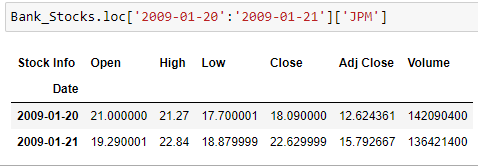
Now Let’s figure out from the PCT\_Change(), In whole timeframe which was the days that gets the highest and lowest single day return.

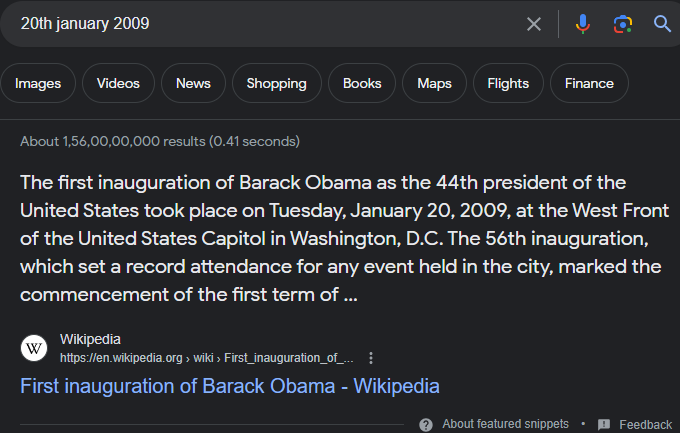
From min and max, we can see it’s targeting the Recession Time period for highest and lowest single day return. And Even id we check that JP Morgan Chase having highest and lowest in just one day.

What happen in one day that drops the price and again rise the price?

So, out of Curiosity if we look into the below data, we would see that the price fluctuates so much in one day that it didn’t fluctuate on the whole-time frame.



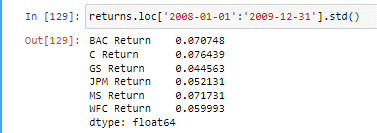
So, there definitely a reason behind this, And if we look on the day 20-January-2009, we would see below result:



**Which shows us that the Market Value is totally dependent on Public’s Sentiment.**

Now, Let’s move and check which one from this was having the riskiest stock to purchase on the timeframe from 2008 to 2009, As the data up above is targeting The Great Recession Time, we will check on that time itself which was riskier among those 6 banks.

To, know about risk among those, we will check the Standard Deviation on that timeframe for 6 Bank Stocks. If a Stock having very high standard deviation on returns that means that stock is riskier.

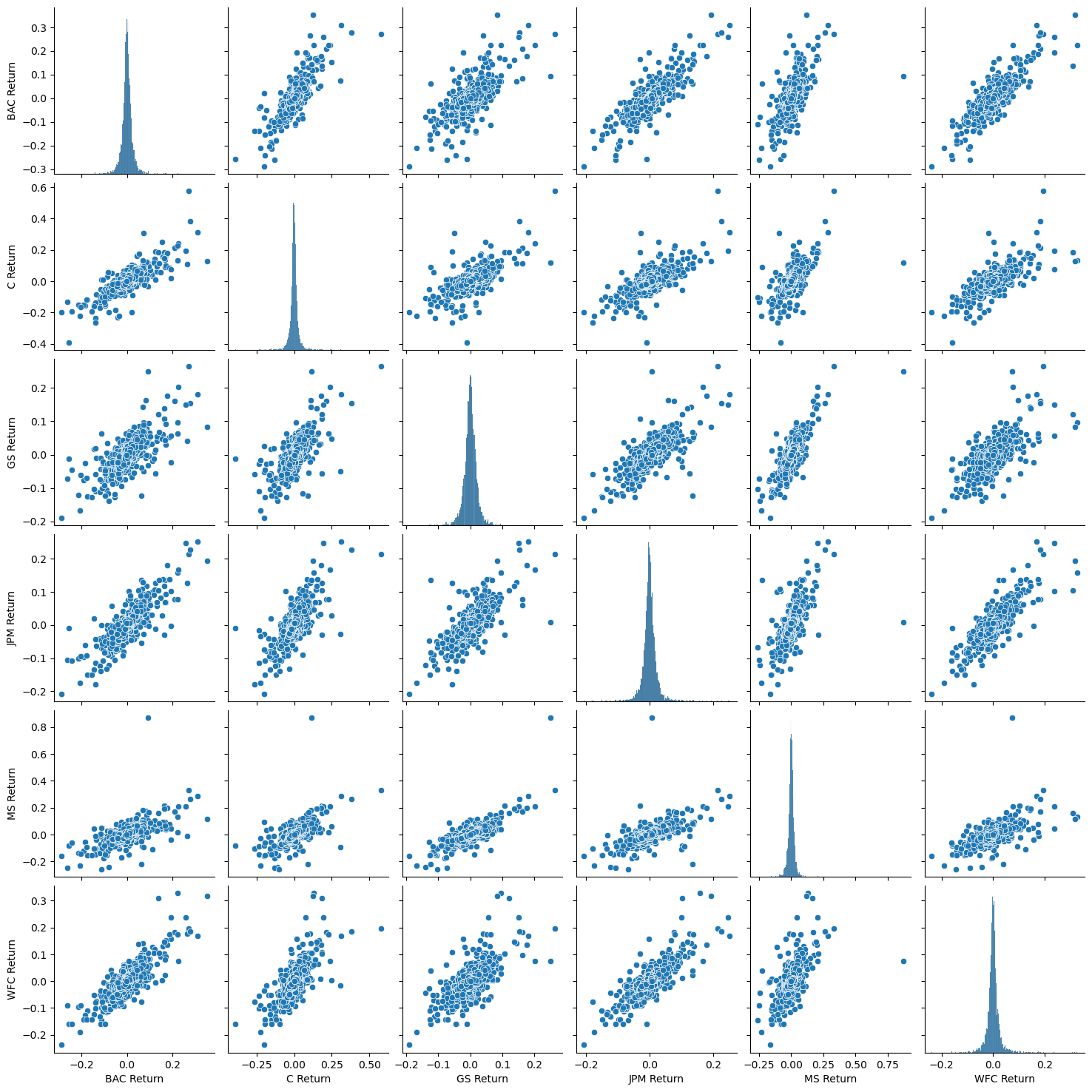


As we can see among those stocks BAC, C & MS having higher STD, which means those was the risker on buy that time because the price was fluctuating from high to low.

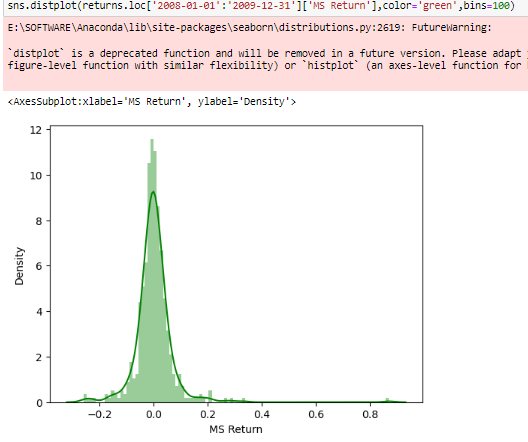
And we will take another step to make more analysis towards the data though Plots on next step.

Data Analysis Through Plots

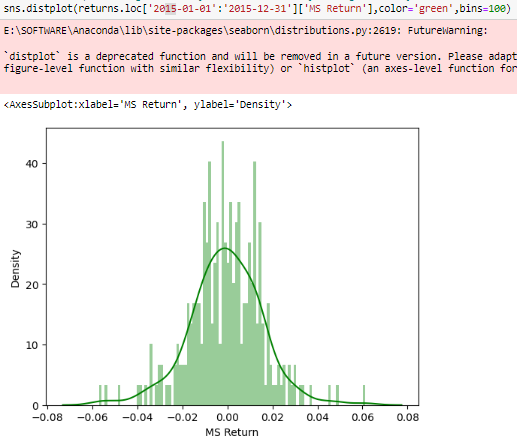
Let’s take a look on PairPlot of those 6 banks from 2006 to 2023 data.



And to make clear picture on 2008-2009 let’s take a look on MS Return and now we will see in distplot.



After seeing the KDE printed a smooth line on Distribution, we can see that the plot Is a right skewed. And if we differentiate the same plot with 2015 data, we would see that the plot will create a Bell Curve.

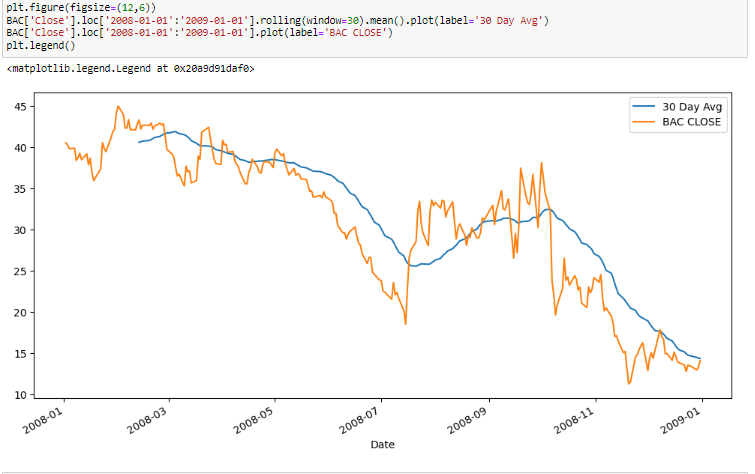


For more visibility and Understanding, Let’s create a Line Plot by using Plotly to clearly understand how 2008’s recession hit the stock market.



So, from this we can see that on 2008-2009 All Bank Stocks ae just crashed, But the GS Bank recovered very fast but others were not able to recover so much faster.

And in below some other plots which provide us some more insights about the Timeperiod, So below is 30 days moving average data with BAC Close on 2008 -2009.



And below is the Bollinger Curve for BAC Bank for 2008-2009 Timeframe.



**So, At the End After Analysis we can tell that, The stock market is totally dependent on Public Sentiment as well as that depend on Market Situation. In the Analysis I targeted on Recession timeframe as that was the time when financial crisis happened and we saw the significant example to see how much that made a difference on Stock Market.**

**I made this analysis on 6 stocks but if we see any other stocks on that time frame, we would see same amount of significant difference though out the Period.**