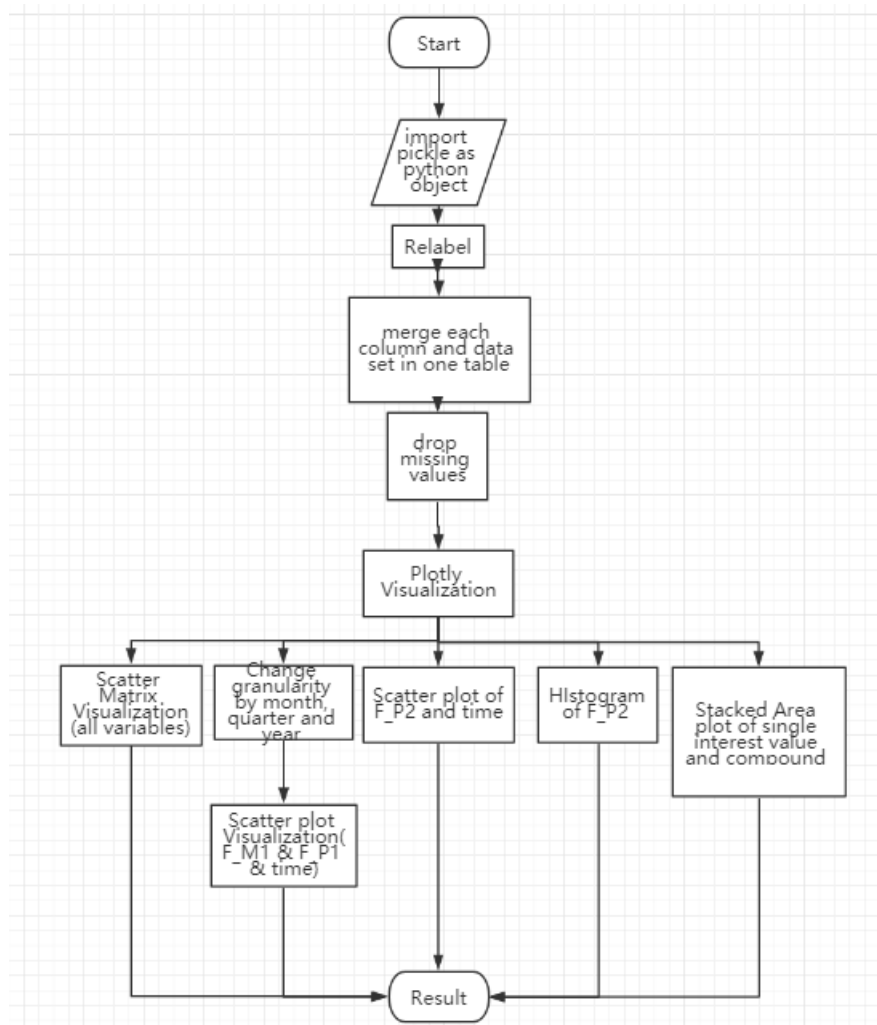


Contents

1. Process	2
2. Columns descriptions (relabel)	2
3. Difficulties and solutions	2
4. Visualizations & explanations	3
a) All variables	3
b) F_M1	4
c) F_P1	6
d) F_M1 & F_P1	8
e) F_P2	9

1. Process



2. Columns descriptions (relabel)

F_M1: 'M1' of 'F_pickle' file

F_M2: 'M2' of 'F_pickle' file

F_P1: 'P1' of 'F_pickle' file

F_P2: 'P2' of 'F_pickle' file

F_M1: integer 'M1' of 'M_pickle' file (I manually repeated it throughout the whole timestamp)

F_M2: 'M2' of 'M_pickle' file

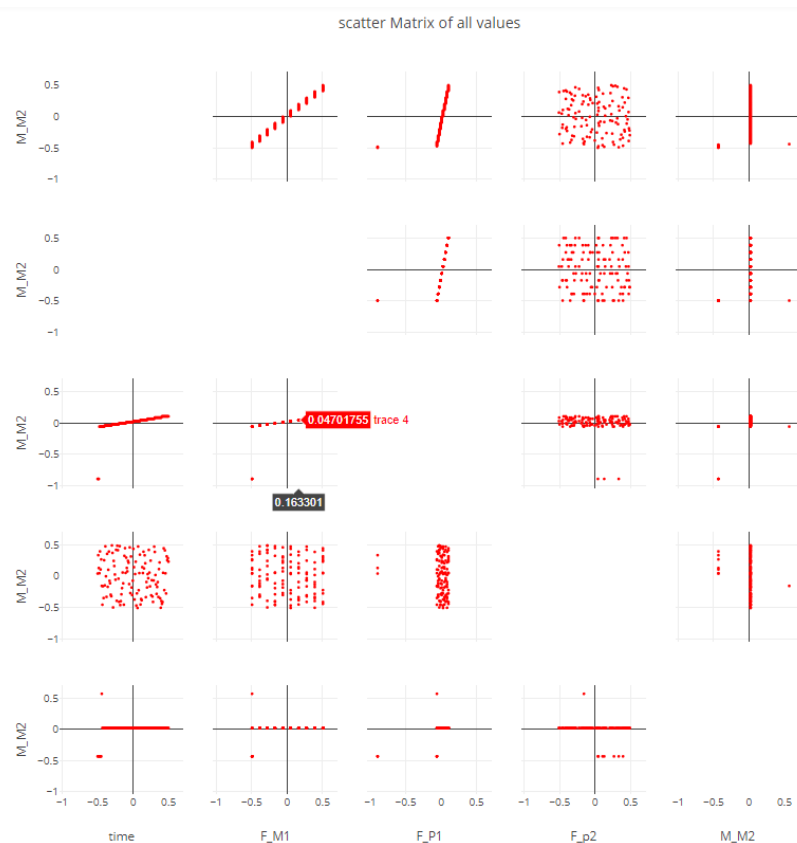
3. Difficulties and solutions

1. Integrate pandas object into Plot.ly

- a) I searched sample codes and learned Plot.ly for almost 10 days to figure out the how to code in Plot.ly.
2. Visualization creativities
 - a) M_M1 and M_P1 goes up as time goes by, so the change rate transformation could be added.
 - b) M_P2 looks like a random variable without obvious relationship with time. It looks like growth rate, so the single interest accumulated value and compounded interest value of it can be added to see any interesting patterns with stacked area plot and percentage stacked area plot.
 - c) Change frequencies with month, quarter and year.

4. Visualizations & explanations

a) All variables

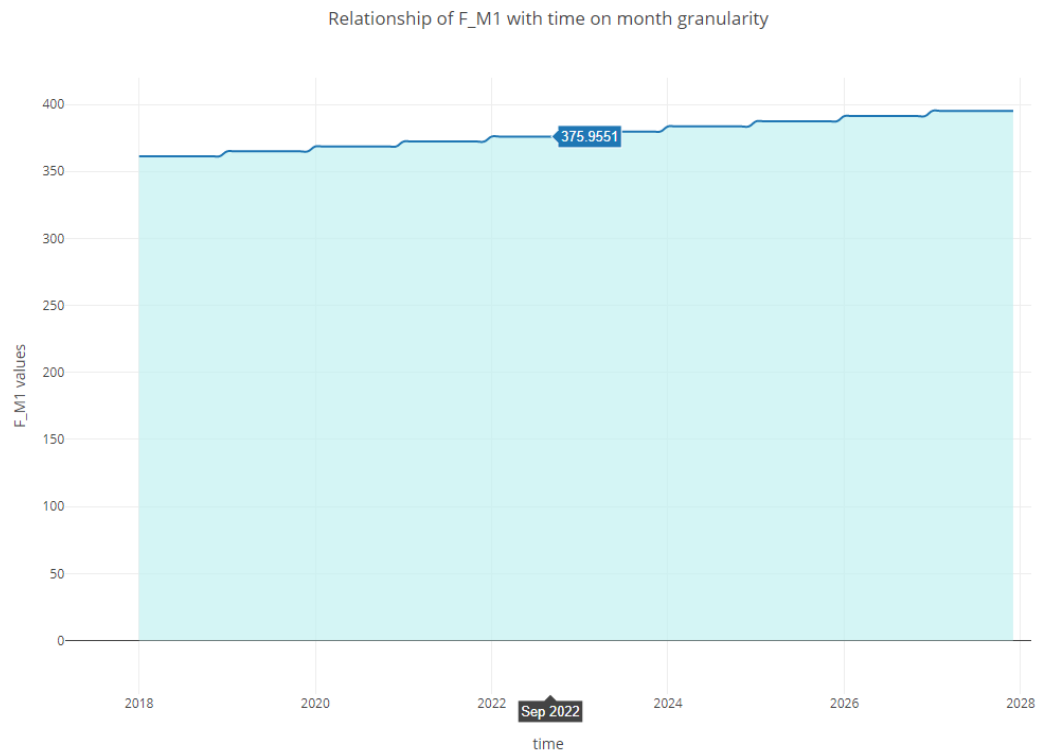


<https://plot.ly/~leiwen/20/scatter-matrix-of-all-values/>

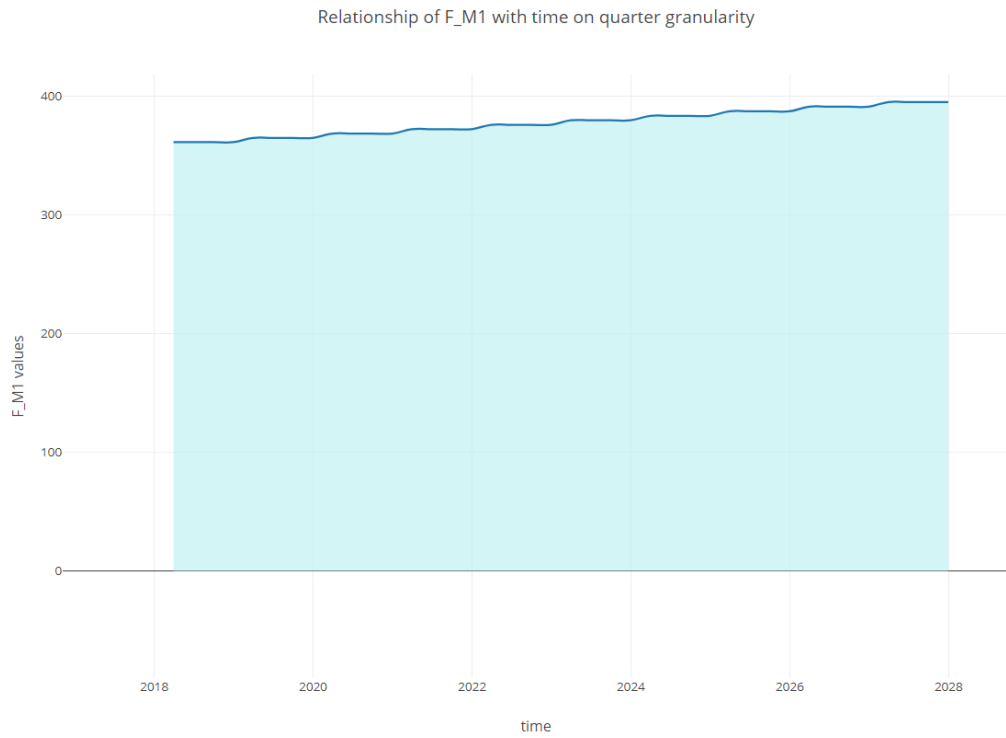
Taking one glimpse at all the variables, they appear with different patterns with time and interaction with each other. In my opinion, F_M1, F_P1 and F_P2 are

most valuable variables to visualize, while others are worthless constants. Let's dig into further the data.

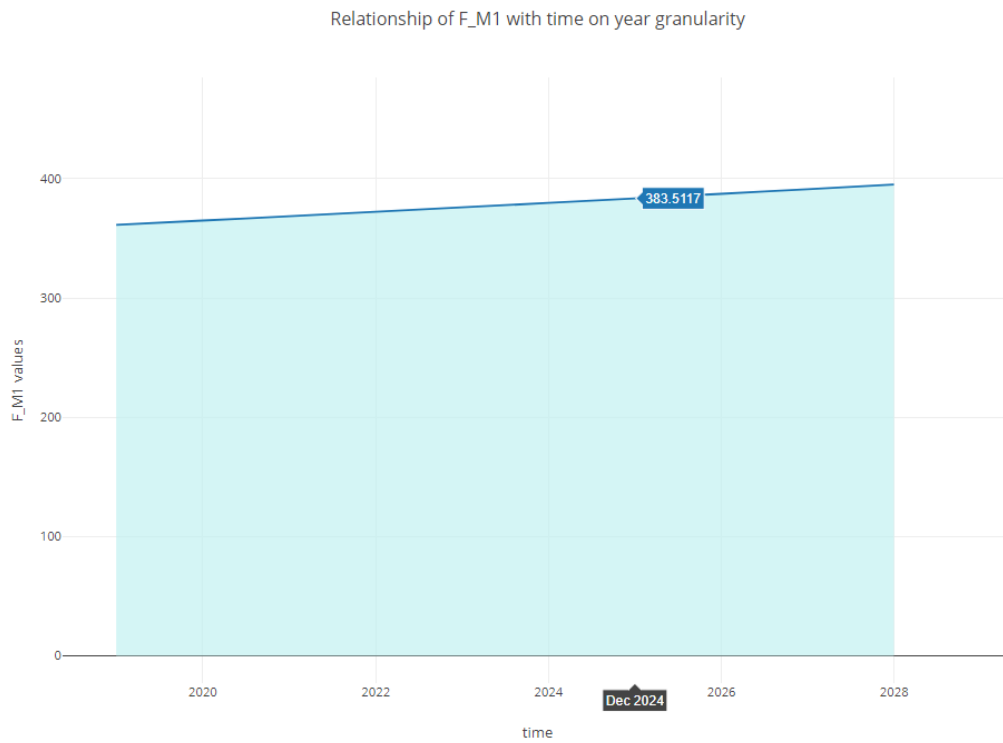
b) F_M1



Relationship of F_M1 with time on month granularity
<https://plot.ly/~leiwen/22/relationship-of-f-m1-with-time-on-month-granularity/>



Relationship of F_M1 with time on quarter granularity
<https://plot.ly/~leiwen/24/relationship-of-f-m1-with-time-on-quarter-granularity/>

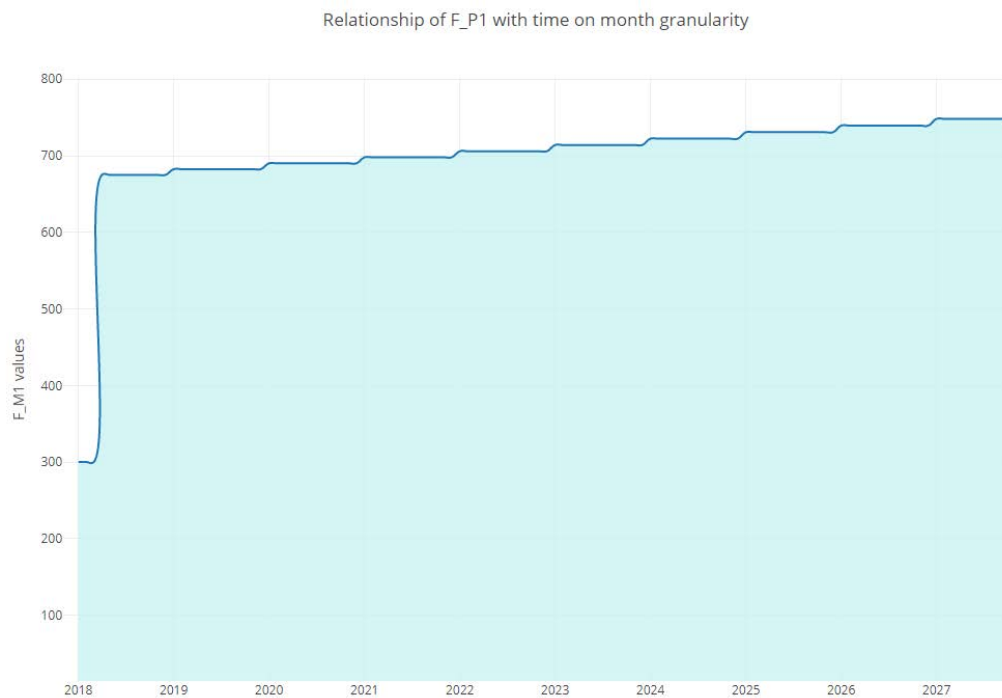


Relationship of F_M1 with time on year granularity
<https://plot.ly/~leiwen/26/relationship-of-f-m1-with-time-on-year-granularity/>

It is obvious that in terms of month or quarter or year granularity, the F_M1 increases at a very smooth and steady speed with time.

Therefore, F_M1 has a very strong relationship and increases steadily as time goes by.

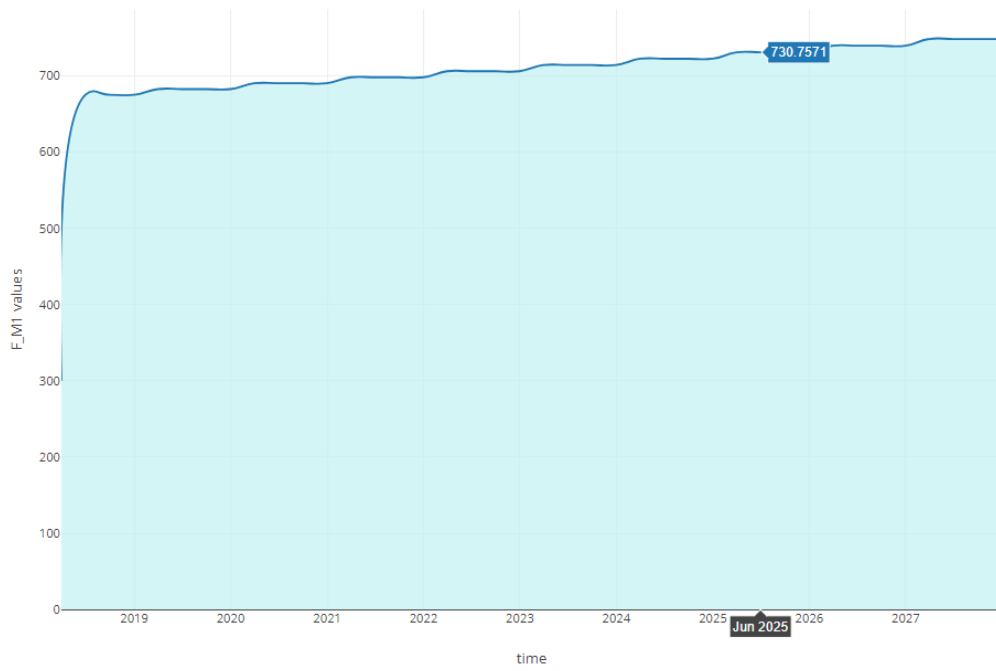
c) F_P1



Relationship of F_P2 with time on month granularity

<https://plot.ly/~leiwen/28/relationship-of-f-p1-with-time-on-month-granularity/>

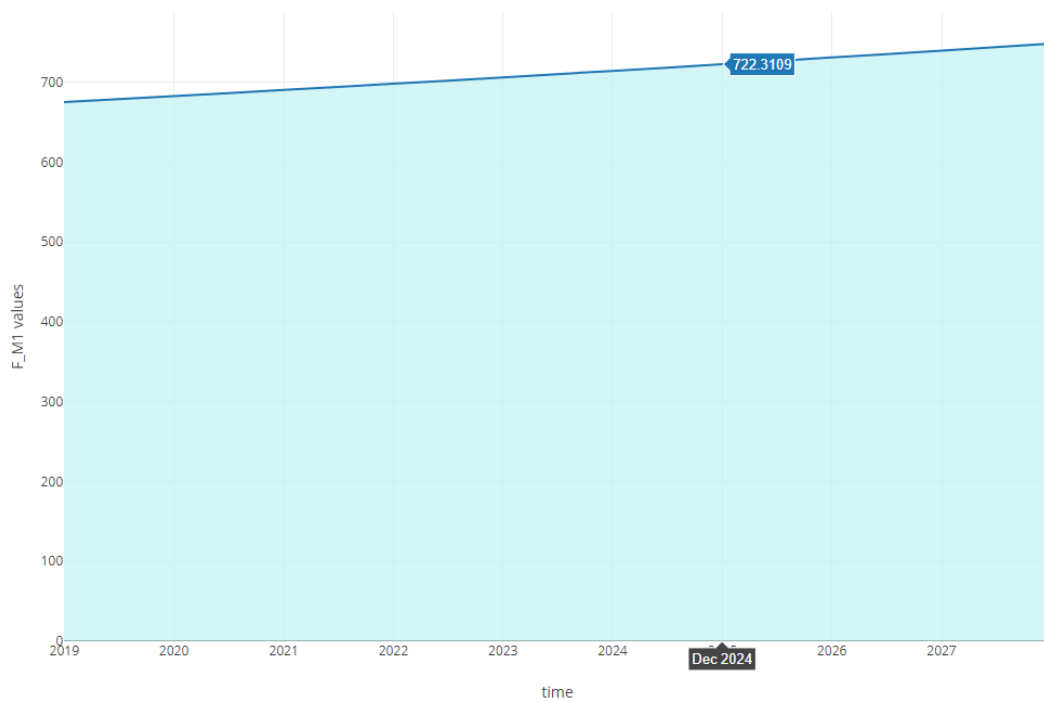
Relationship of F_P1 with time on quarter granularity



Relationship of F_P2 with time on quarter granularity

<https://plot.ly/~leiwen/30/relationship-of-f-p1-with-time-on-quarter-granularity/>

Relationship of F_P1 with time on year granularity



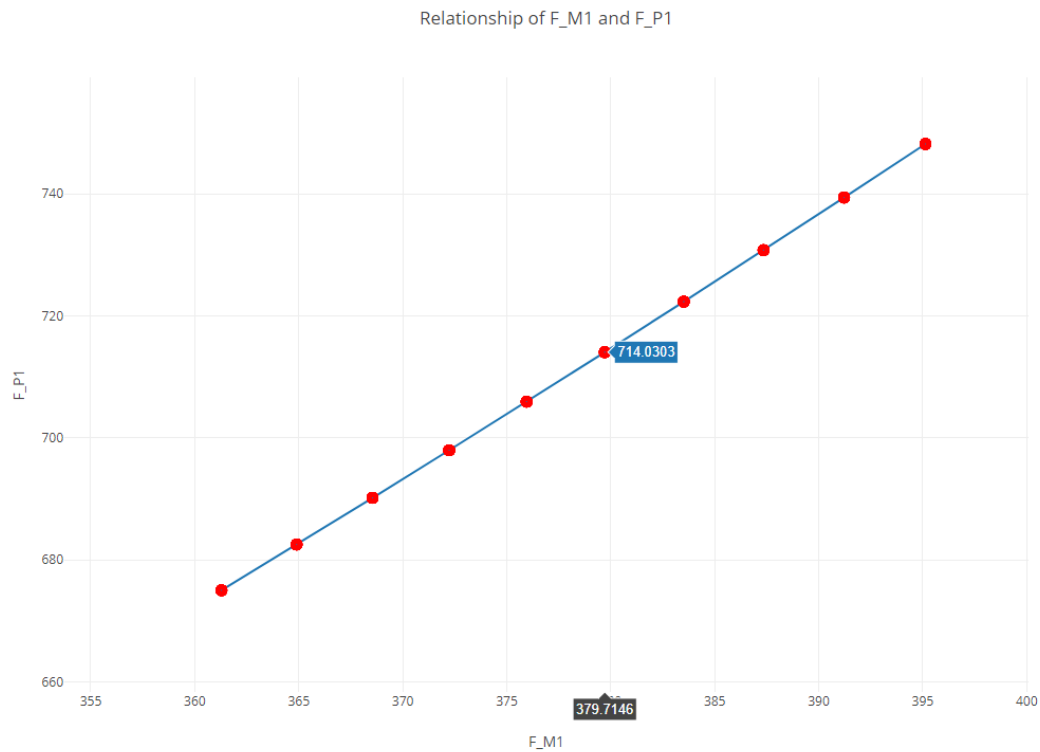
Relationship of F_P2 with time on year granularity

<https://plot.ly/~leiwen/32/relationship-of-f-p1-with-time-on-year-granularity/>

Omitting some special outliers by taking medium, F_P1 has a very strong relationship and increases steadily with time.

d) F_M1 & F_P1

F_M1 & F_P1 Both have a strong linear positive and steady incremental rate with time, so they probably have the same strong linear positive relationship with each other.

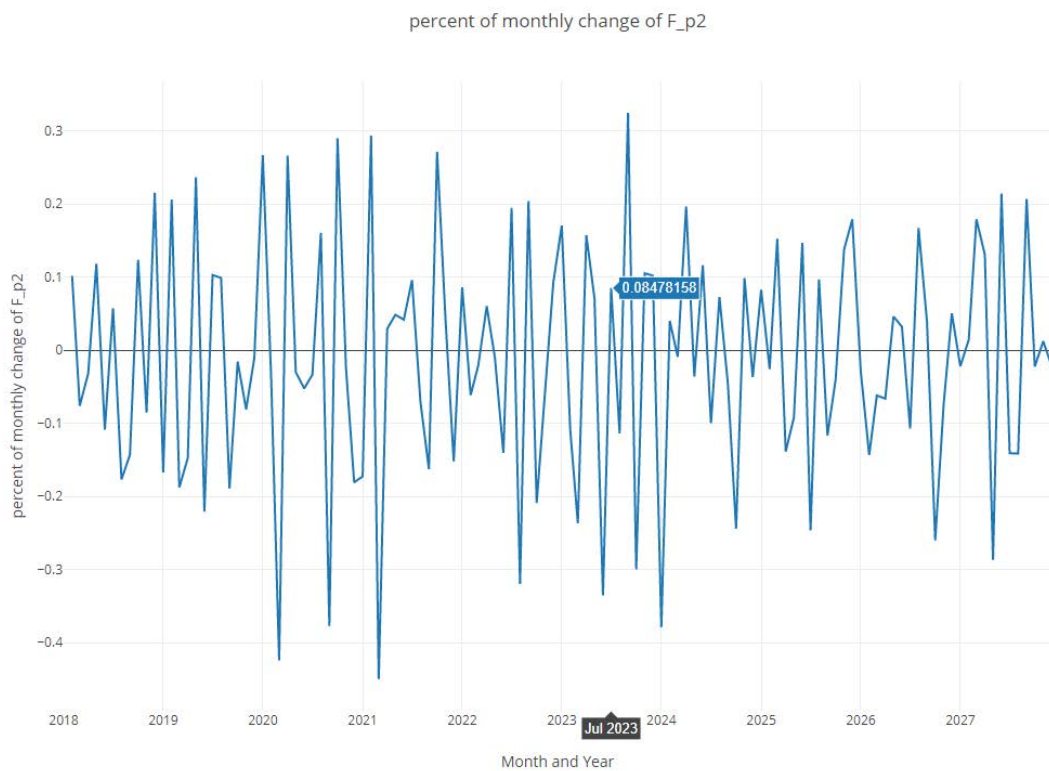


Relationship between F_P1 and F_M1

<https://plot.ly/~leiwen/34/relationship-of-f-m1-and-f-p1/>

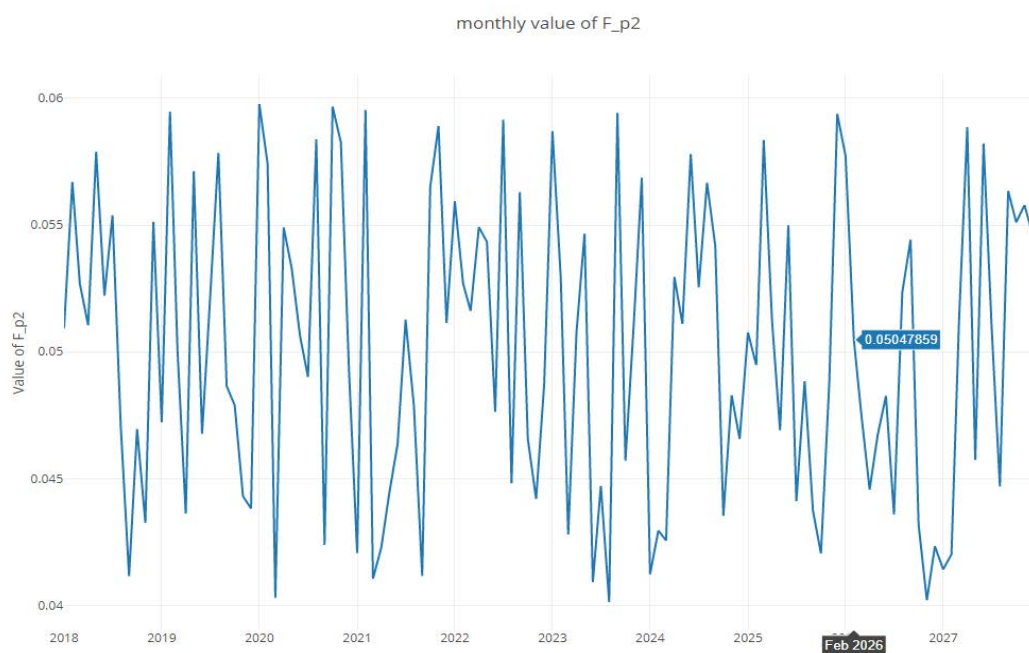
Put F_M1 and P_P1 on x and y axis, it is clear that they are highly linear corelated (omitting the first three observation of F_P1 and plot on year granularity).

e) F_P2



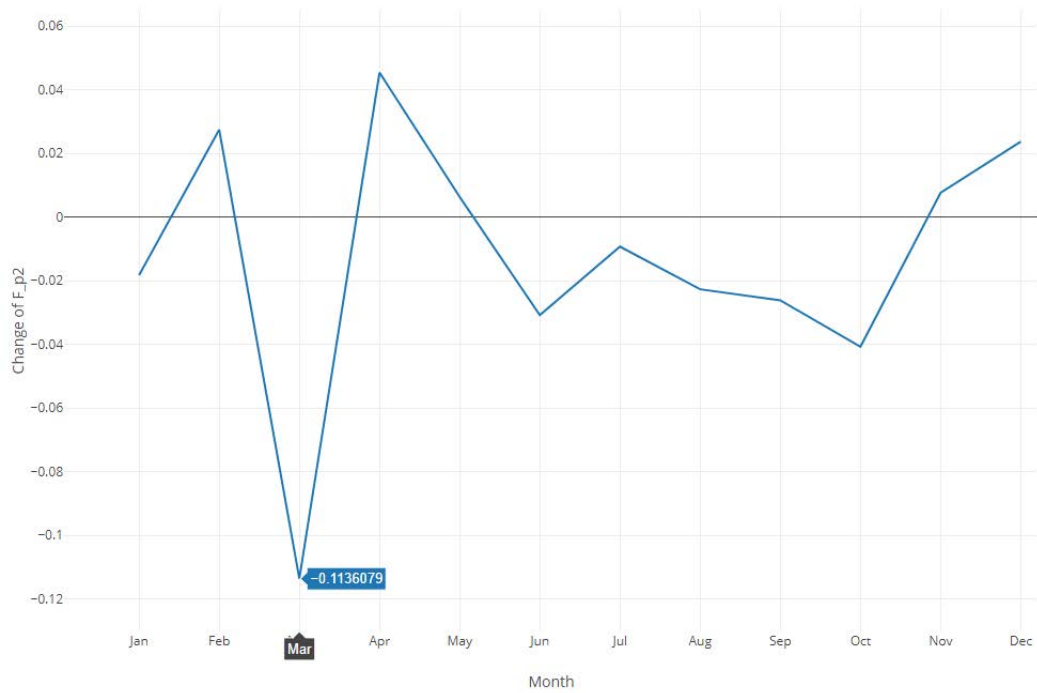
Relationship of change rate of F_p2 with time on month granularity

<https://plot.ly/~leiwen/36/percent-of-monthly-change-of-f-p2/>



Relationship of F_p2 with time on month granularity

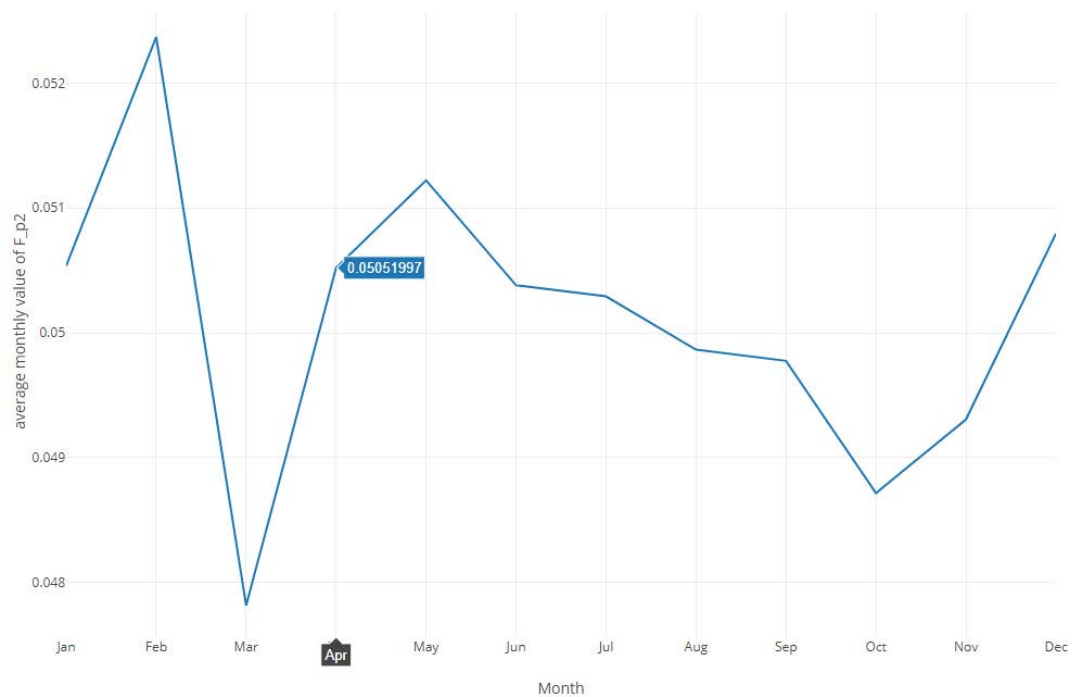
Monthly F_{p2} and its change rate has no apparent relationship with time, and quarterly nor yearly data neither (I do not include the year nor quarter plot in this write-up)



Average percent of monthly change of F_{p2}

<https://plot.ly/~leiwen/40/average-percent-of-monthly-change-of-f-p2/>

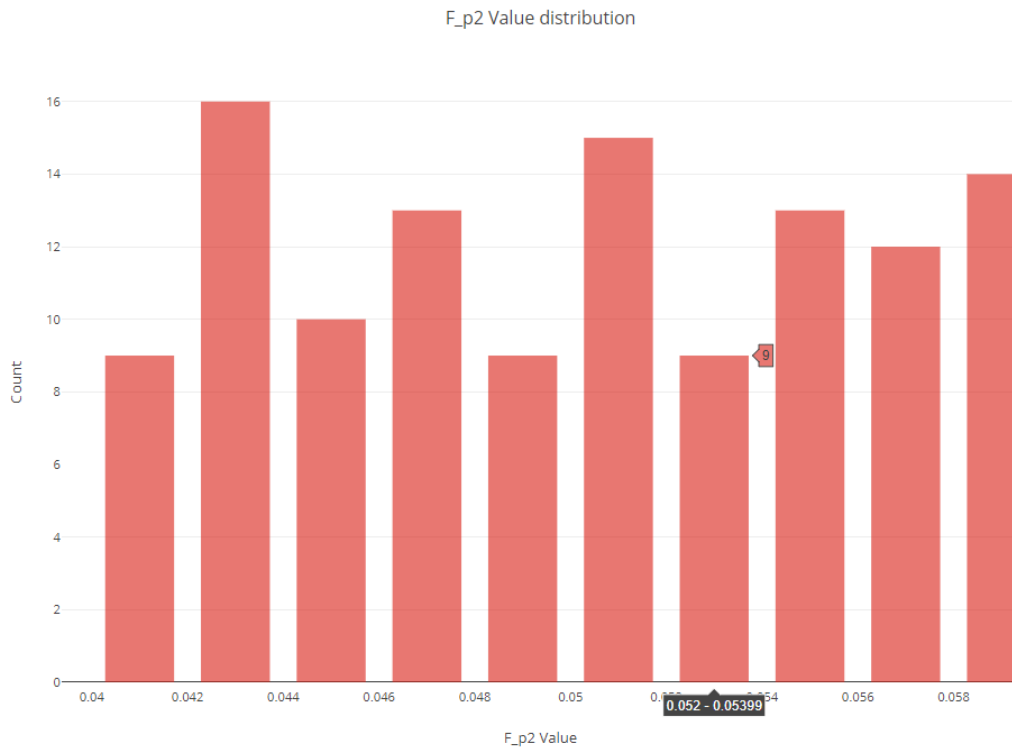
The average change rate of each month has no distinct pattern.



Average monthly value of F_p2

<https://plot.ly/~leiwen/42/average-monthly-value-of-f-p2/>

The average value of each month has no distinct pattern. We can only see average values of F_P2 of first two month (Jan and Feb) among all years are higher than other that of other months and that of March is lowest among them.

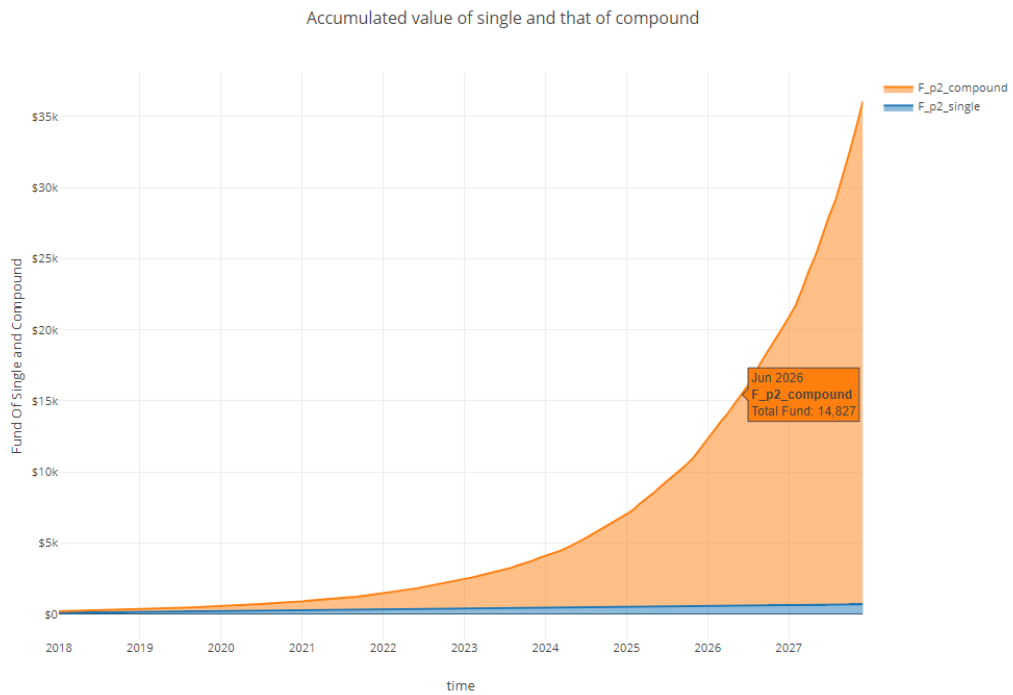


Histogram of F_P2 distribution

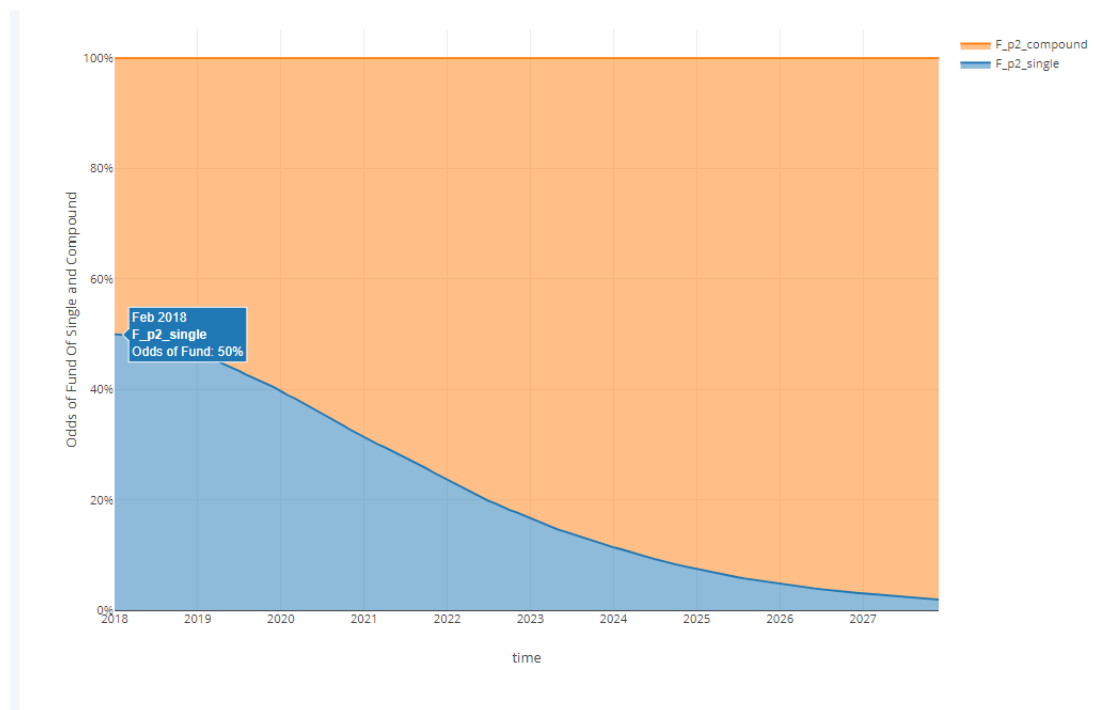
<https://plot.ly/~leiwen/46/f-p2-value-distribution/>

Also, F_P2 distribution has no distinct pattern. Therefore, F_P2 is more like a **random walk** variable, such as return rate (though return rate has negative values), growth rate, unemployment rate or volatility of options (I guess).

Furthermore, if it is a growth rate variable or interest rate variable, we can draw its accumulated value (both single interest and compounded value)



See compounded value of P_F2 grows much faster than single interest value.
<https://plot.ly/~leiwen/48/accumulated-value-of-single-and-compound/>



odds-of-value-of-single-and-value-of-compound
<https://plot.ly/~leiwen/50/odds-of-value-of-single-and-value-of-compound/>

See, the compounded value and the single value are 50:50 at the beginning, but compounded value becomes much more dominant as time goes by.