Principal Component Analysis & It's Application using Python.

Presentation Outline

- Background
- PCA Overview
- Interest Rate Overview
- Application of PCA in Python
- Analysis

Background

- The objective of this presentation is to use mathematical tools to simplify the complex structure to extract meaningful information.

- The complexity sometimes leads to mismanagement of portfolio and in order to avoid such instance, we need to arrive at a solution which provides efficient ways to manage them.

- PCA and Interest rate term structure combination will help us to understand this in a much better way.

PCA Overview

- PCA concept that has been borrowed from linear algebra which is a branch of mathematics is the concept of Principal Component Analysis (hereinafter also referred to as ‘PCA’).

- PCA as a concept is useful for measuring risk arising from a set of correlated market variables.

- PCA is a model which involves transforming a set of observations into a set of uncorrelated variables called as the PCs.

PCA Overview

- Input Required for Calculation

- **Standard Deviation** : The standard deviation of data set is a measure of how spread out the data is. In mathematical terms, it is the square distance of each point from its mean.

- **Variance:** Variance is another method of spread dataset and it's almost identical to standard deviation. The only thing is that we do not apply a square root to above observation.

- **Covariance:** Standard Deviation and Variance only operates on one dimension, let's say we have dataset X{1,2,3,.....,6} & another dataset as Y{2,3,4,5....44} how X & Y vary with respect to each other, their co-movement is Covariance.

- **EigenVectors**: We can multiply two matrices together provided they are of compatible sizes. Eigen Vector is special of this cases let’s look at them with few example and try to understand them further. Let's look at below example.

$$\begin{pmatrix} 2 & 3 \\ 2 & 1 \end{pmatrix} \times \begin{matrix} 3 \\ 2 \end{matrix} = \begin{matrix} 12 \\ 8 \end{matrix} = 4 \times \begin{matrix} 3 \\ 2 \end{matrix}$$

- **Eigen Values:** This is nothing but acting as scalar multiple associated with Eigen Vector, as in above example resultant multiple got scaled up by **four time** of original vector.

PCA Overview

- Properties of PCA & Logic

- Orthogonality is a property in linear algebra which is used to test if two vectors are perpendicular to one other i.e. the two vectors are uncorrelated with one another.
- This transformation behaves in a way such that the first PC explains the largest possible variance, and this accounts for the majority of the variability in the data.
- PCs represent the directions of the data that explain the maximum amount of variance, i.e. the vectors that capture most of the information that is embedded in the data.

- Dimensional reduction implies, we attempt to capture the essence of a multivariate dataset into a fewer number of variables that would explain the required result.
- The eigenvectors of the covariance matrix are the direction of the axes where there is more variance i.e. most information. These are the PCs.
- Eigenvalues are the coefficients attached to the eigenvectors; they explain the amount of variance carried by each of the PCs.

Interest Rate Overview

- Basic Understanding

- **Interest Rate Yield:** Yield is basically an annual return on investment, it depends upon the purchase price of bond & interest / coupon / payment received.

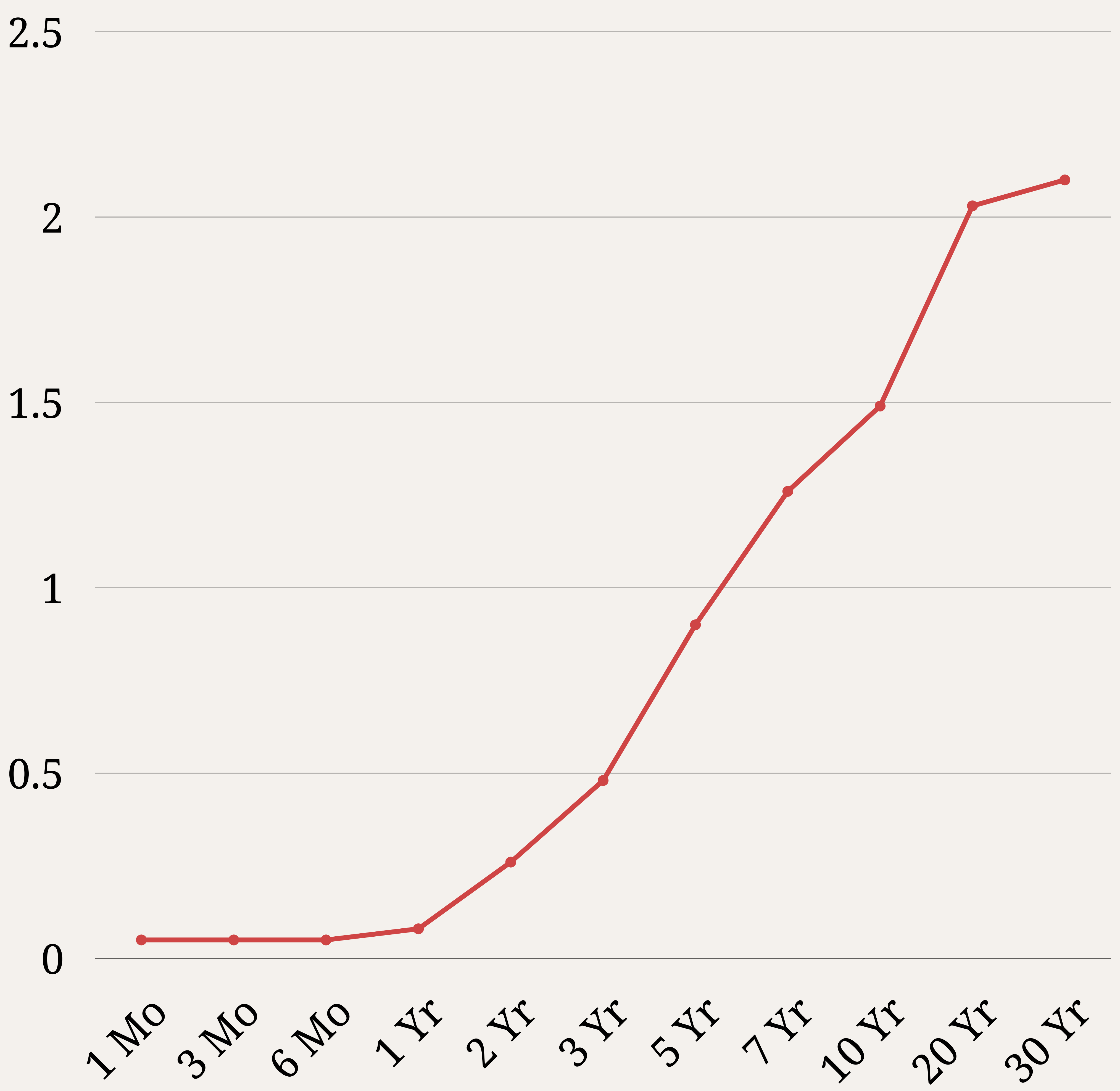
  - **Current Yield**: It's an annual return earned on the price paid for the bond.

    - Face Value of Bond: 1000, Annual Return: 6%
    - Interest payment would be 60 which is 6% current yield.
    - If bond purchase at discount rate of 900.
    - Yield in that case would be 6.6%

- **Yield to Maturity:** Total return an investor receives by holding the bond until it matures.
    - It is generally more meaningful for investors.
    - Investors can compare bonds with varying nature
    - Different maturities / coupon rates / credit quality

What is Yield Curve ?

- The yield curve is a line graph that plots the relationship between yields to maturity and time to maturity for bonds of the same asset class and credit quality.



https://www.treasury.gov/resource-center/

- **Factors, which decides the shape of Yield Curve**: Studies conducted by various people on Interest rate behavior concluded that there are three main factors which decides movement of yields.
  - Level: This is also known as parallel shift across term structure.
  - Slope : This will account for steepness of the curve.
  - Curvature: Attributed to account for bowing of the interest rate curve.

Application of PCA in Python

- https://github.com/abhaydd22/PCA

Data_US_Treasury

- **Range:** Jan'17 to Jun21

- **Terms:** 1 Mo,3 Mo,6 Mo,1 Yr,2 Yr, 3 Yr,5 Yr,7 Yr,10 Yr,20 Yr,30 Yr.

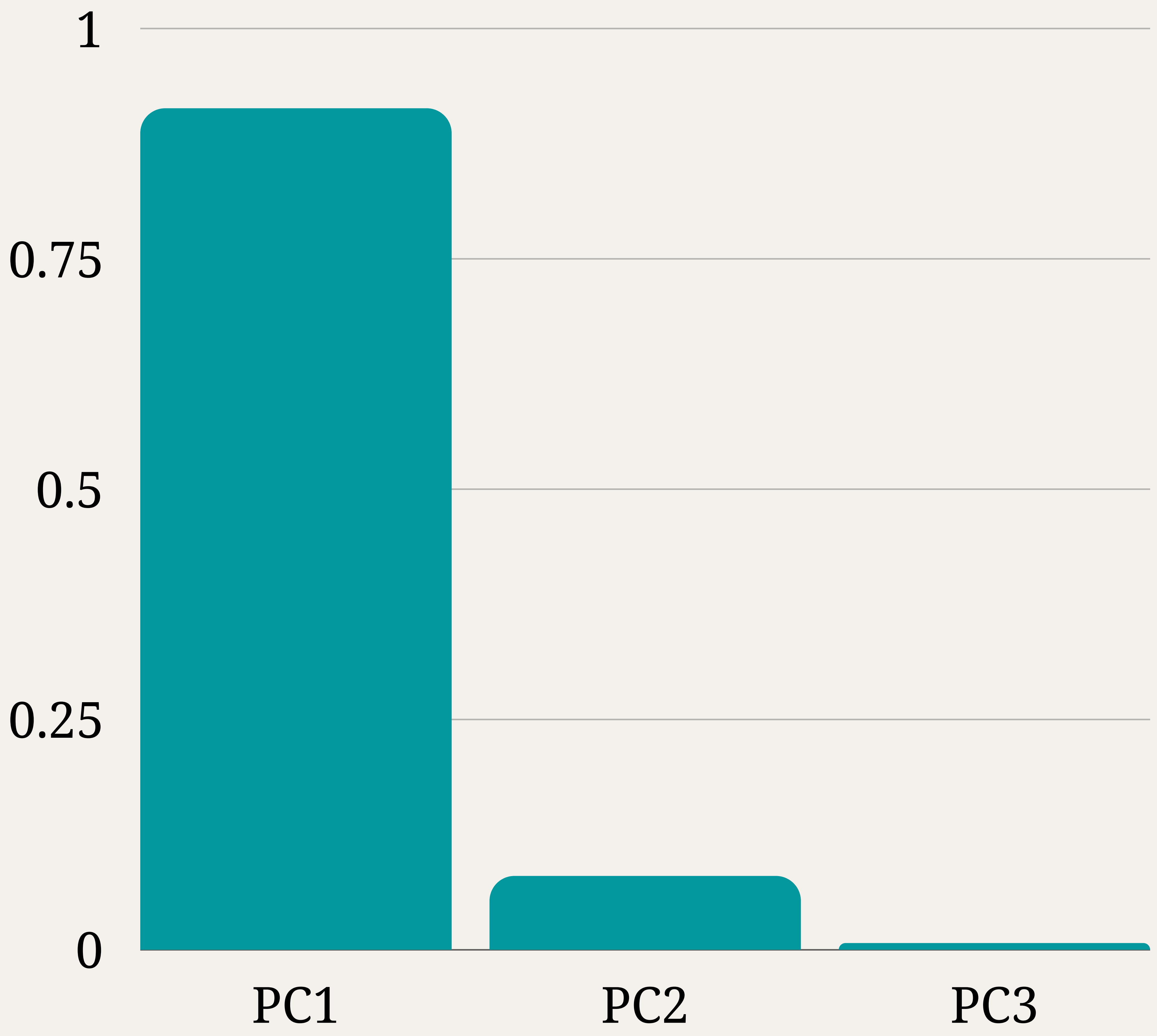
	Date	1 Mo	3 Mo	6 Mo	1 Yr	2 Yr	3 Yr	5 Yr	7 Yr	10 Yr	20 Yr	30 Yr
0	03-01-2017	0.52	0.53	0.65	0.89	1.22	1.50	1.94	2.26	2.45	2.78	3.04
1	04-01-2017	0.49	0.53	0.63	0.87	1.24	1.50	1.94	2.26	2.46	2.78	3.05
2	05-01-2017	0.51	0.52	0.62	0.83	1.17	1.43	1.86	2.18	2.37	2.69	2.96
3	06-01-2017	0.50	0.53	0.61	0.85	1.22	1.50	1.92	2.23	2.42	2.73	3.00
4	09-01-2017	0.50	0.50	0.60	0.82	1.21	1.47	1.89	2.18	2.38	2.69	2.97
...	...	...	...	...	...	...	...	...	...	...	...	...
1116	18-06-2021	0.05	0.05	0.06	0.09	0.26	0.47	0.89	1.22	1.45	1.97	2.01
1117	21-06-2021	0.04	0.05	0.06	0.09	0.27	0.48	0.90	1.25	1.50	2.05	2.11
1118	22-06-2021	0.04	0.04	0.06	0.09	0.25	0.44	0.87	1.23	1.48	2.03	2.10
1119	23-06-2021	0.04	0.05	0.05	0.08	0.26	0.47	0.90	1.25	1.50	2.04	2.11
1120	24-06-2021	0.05	0.05	0.05	0.08	0.26	0.48	0.90	1.26	1.49	2.03	2.10

1121 rows × 12 columns

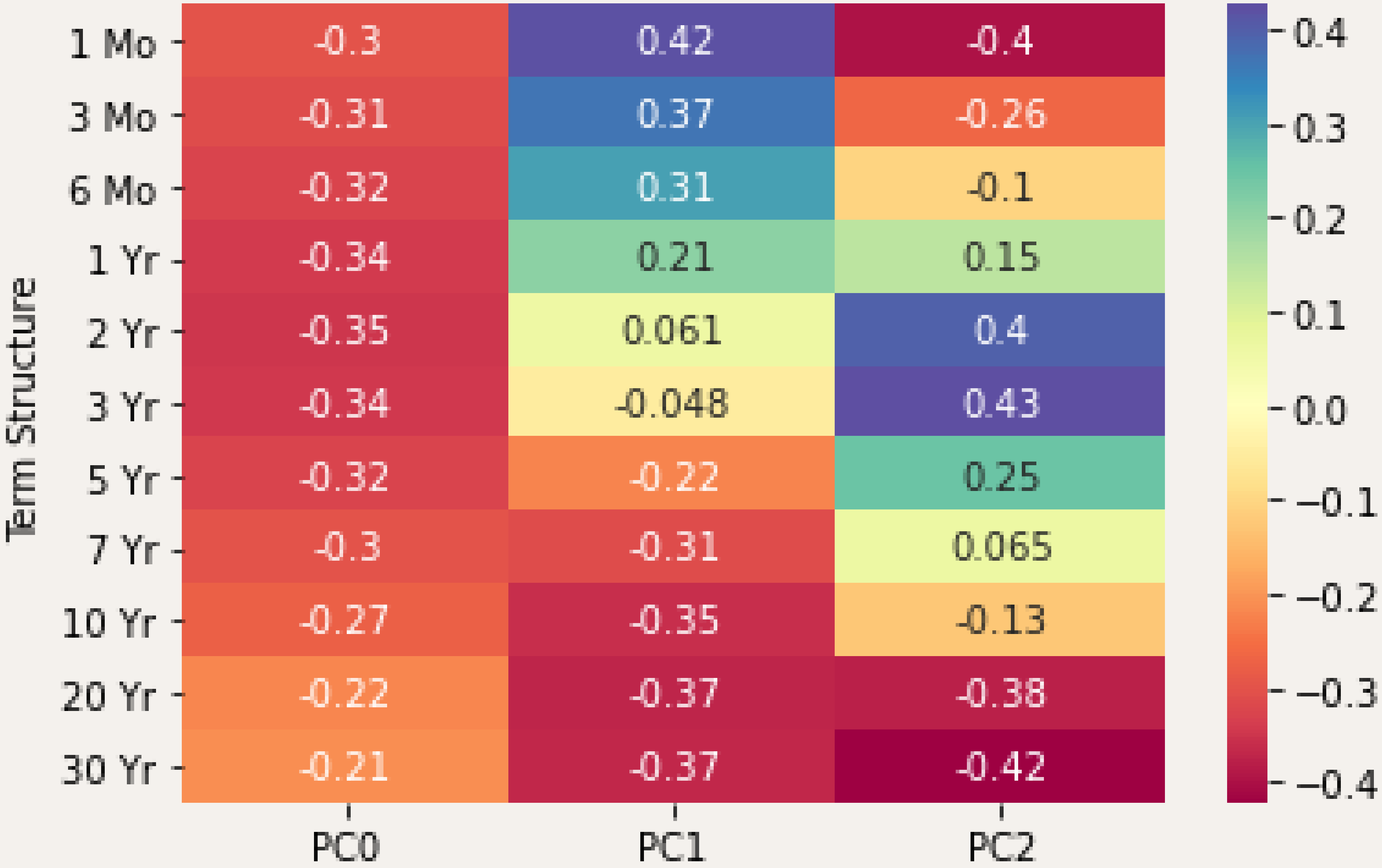
Data_US_Treasury: Basic Statistics

	1 Mo	3 Mo	6 Mo	1 Yr	2 Yr	3 Yr	5 Yr	7 Yr	10 Yr	20 Yr	30 Yr
count	1121.000000	1121.000000	1121.000000	1121.000000	1121.000000	1121.000000	1121.000000	1121.000000	1121.000000	1121.000000	1121.000000
mean	1.153568	1.203916	1.274380	1.334112	1.415656	1.493568	1.669831	1.855540	2.001017	2.321097	2.495406
std	0.857911	0.868350	0.892612	0.915885	0.926115	0.916553	0.872778	0.820953	0.773181	0.646017	0.623108
min	0.000000	0.000000	0.020000	0.040000	0.090000	0.100000	0.190000	0.360000	0.520000	0.870000	0.990000
25%	0.110000	0.140000	0.160000	0.170000	0.200000	0.340000	0.820000	1.250000	1.530000	1.910000	2.080000
50%	1.160000	1.290000	1.450000	1.490000	1.500000	1.560000	1.810000	2.030000	2.210000	2.550000	2.770000
75%	1.930000	1.980000	2.090000	2.160000	2.270000	2.340000	2.430000	2.530000	2.630000	2.840000	3.010000
max	2.510000	2.490000	2.580000	2.740000	2.980000	3.050000	3.090000	3.180000	3.240000	3.370000	3.460000

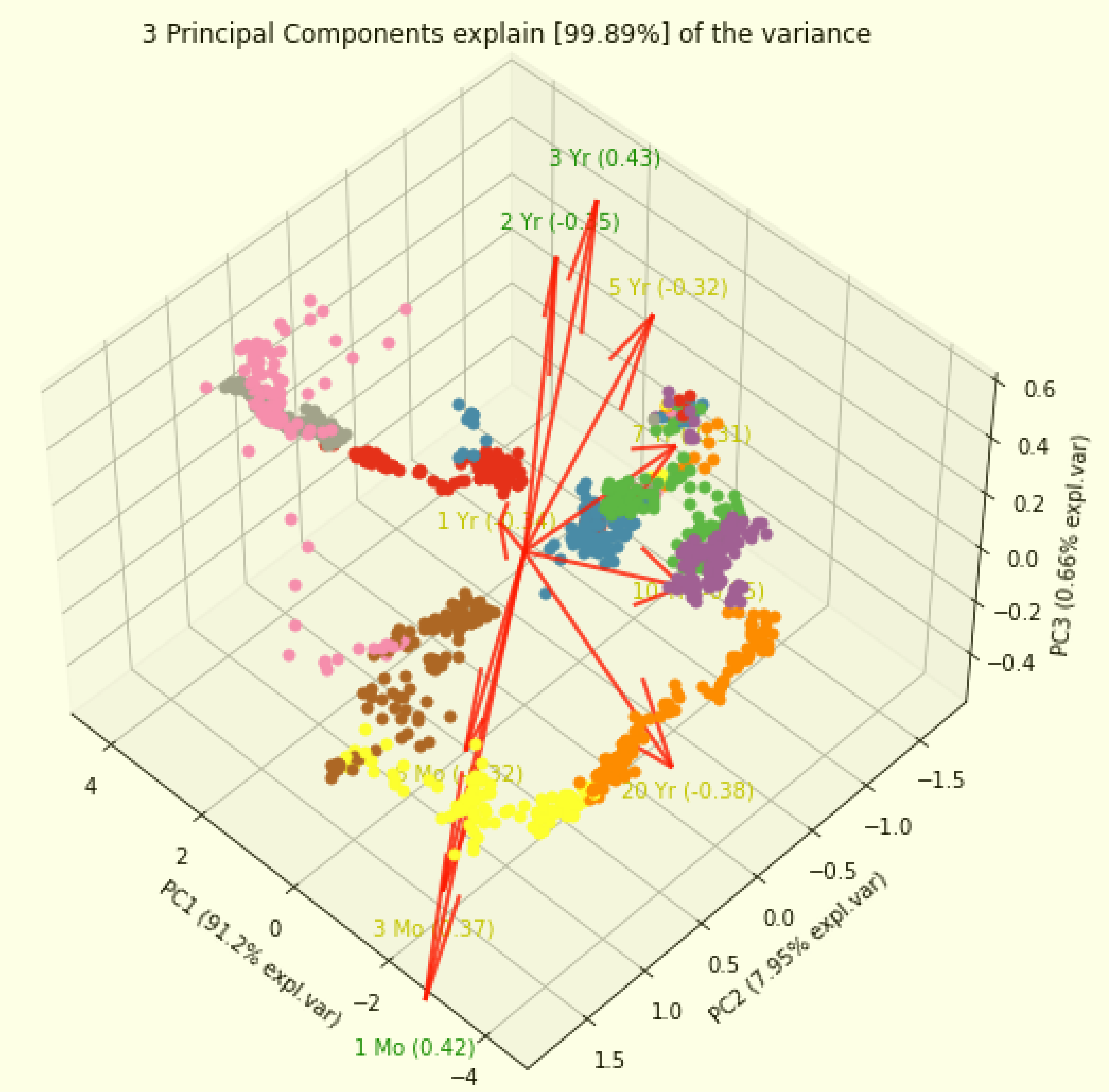
PCA on Data : First three PC's explains 99.87% of variance.



PCA Loadings



PCA Factors:


3 Principal Components explain [99.89%] of the variance
3 Yr (0.43)
5 Yr (-0.32)
20 Yr (-0.38)
1 Mo (0.42)
PC1 (91.2% expl.var)
PC2 (7.95% expl.var)
PC3 (0.66% expl.var)
4
2
0
−2
−4
−1.5
−1.0
−0.5
0.0
0.5
1.0
1.5
0.6
0.4
0.2
0.0
−0.2
−0.4

Analysis of Outcome

- **Effect of PCA on Rates:** Principal component helps in identifying the effect of unit change or shock on the interest rate change will lead to.

	PC0	PC1	PC2
10 Yr	-0.27	-0.35	-0.13

- One unit change in interest rate, for example, let's say for 2 year tenure will lead to following results.

$$\Delta r_t = \sum_{j=1}^{k} a_{jt} \Delta f_j$$

- Change in 10 years rate will have following equation.
  - -0.27 - 0.35 - 0.13 = -0.75%
  - 10 years rate will change from 1.49% to 1.47%

$$\Delta r_{10} = \sum_{j=1}^{3} a_{j10} \Delta f_j$$

- Risk measurement of portfolio:

$$\sigma\,(Portfolio) = \sqrt{w1^2\sigma1^2 + w2^2\sigma2^2 + w3^2\sigma3^2}$$

  - If we multiply PV01 with respective PC (weights) factors against their tenures.
  - Sigma is basically standard deviation of factor loadings.
  - Since, PC's are independent, we don't need to include correlation in the above formula.
  - Finally, we can calculate 1-day VaR of portfolio & it can be calculated as σ (Portfolio) * 2.33.

Acknowledgement

Various papers and webpages helped to gain understanding.

- https://machinelearningmastery.com/calculate-principal-component-analysis-scratch-python/
- https://www.reneshbedre.com/blog/principal-component-analysis.html
- Estimating Term Structure Changes using Principal Component Analysis in Indian Sovereign Bond Market.pdf (Golaka C Nath)
- https://pypi.org/project/pca/