PROJECT REPORT

(PS04CAST53)

ON

MUTUAL FUND PORTFOLIOS AND MAKING INVESTMENT DECISIONS

SUBMITTED BY

Kunal Salunkhe [M.Sc. Applied Statistics | 10]

Nisha Lad [M.Sc. Applied Statistics | 12]

Under the Guidance of

Prof. (Dr.) Jyoti M. Divecha

Department of Statistics
Sardar Patel University
Vallabh Vidyanagar

ACADAMIC YEAR 2023-24

CERTIFICATE

This is to certify that Mr. **Kunal Chatur Salunkhe** Student ID Number: 2022014279 of Semester IV M.Sc. Applied Statistics Department of Statistics, Sardar Patel University, Vallabh Vidyanagar, Anand has successfully completed his group project work on topic **"Mutual Fund Portfolios and Making Investment Decisions"** for the paper code PS04CAST53 and term ending in April 2024.

Project Guide	External Co-Guide	Head of the Department
Prof. (Dr.) Jyoti M. Divecha	Disha Thakkar (ICICI Mutual Fund)	Prof. (Dr.) Jyoti M. Divecha

DATE: 24/04/2024

CERTIFICATE

This is to certify that Miss. **Nisha Rajendra Lad** Student ID Number: 2022014265 of Semester IV M.Sc. Applied Statistics Department of Statistics, Sardar Patel University, Vallabh Vidyanagar, Anand has successfully completed her group project work on topic **"Mutual Fund Portfolios and Making Investment Decisions"** for the paper code PS04CAST53 and term ending in April 2024.

Project Guide	External Co-Guide	Head of the Department
Prof. (Dr.) Jyoti M. Divecha	Disha Thakkar (ICICI Mutual Fund)	Prof. (Dr.) Jyoti M. Divecha

DATE: 24/04/2024

ACKNOWLEDGEMENT

Before delving into the intricacies of our project, we must express our heartfelt appreciation for the invaluable contributions of those who have been instrumental since its inception. Crafting this project has been a formidable academic endeavor, and it is with deep gratitude that we acknowledge the unwavering support, patience, and guidance extended by all involved.

It brings us immense joy to present this comprehensive project report on "**Mutual Funds Portfolios and Making Investment Decisions.**" We are privileged to have had a dedicated team of project guides who have accompanied us since the project's outset. The fruition of this project stands as a testament to the relentless diligence and determination invested by us, alongside the invaluable assistance of our project guide.

In addition, we extend a special note of thanks to our alumni, whose invaluable contributions in providing data and resource materials significantly enriched our project. This project is a tribute to all those individuals whose unwavering support and guidance were pivotal in its realization.

As we embark on this journey of exploration and analysis, we recognize that this project would not have been possible without the collaborative efforts and support of each individual involved. We express our sincere gratitude to all who have played a part, no matter how big or small, in shaping this endeavor.

Signature

Kunal C. Salunkhe

M.Sc. Applied Statistics

Signature

Nisha R. Lad

M.Sc. Applied Statistics

INDEX

1. INTRODUCTION

1.1 Mutual Fund

A mutual fund is a type of investment vehicle consisting of a portfolio of stocks, bonds, or other securities. A mutual fund is a financial vehicle that pools assets from shareholders to invest in securities like stocks, bonds, money market instruments, and other assets.

Mutual funds are operated by professional money managers, who allocate the fund's assets and attempt to produce capital gains or income for the fund's investors. Mutual funds give small or individual investors access to professionally managed portfolios of equities, bonds, and other securities. Each shareholder, therefore, participates proportionally in the gains or losses of the fund.

Mutual funds are divided into several kinds of categories, representing the kinds of securities they invest in, their investment objectives, and the type of returns they seek.

Why do people buy mutual funds?

Mutual funds are a popular choice among investors because they generally offer the following features:

- Professional Management. The fund managers do the research for you. They select the securities and monitor the performance.
- Diversification or "Don't put all your eggs in one basket." Mutual funds typically invest in a range of companies and industries. This helps to lower your risk if one company fails.
- Affordability. Most mutual funds set a relatively low dollar amount for initial investment and subsequent purchases.
- Liquidity. Mutual fund investors can easily redeem their shares at any time, for the current net asset value (NAV) plus any redemption fees.

Avoiding fraud

By law, each mutual fund is required to file a prospectus and regular shareholder reports with the SEC. Before you invest, be sure to read the prospectus and the required shareholder reports. Additionally, the investment portfolios of mutual funds are managed by separate entities known as "investment advisers" that are registered with the SEC. Always check that the investment adviser is registered before investing.

Advantages Of Mutual Fund

1. **Professional Management:** Investors may not have the time or the required knowledge and resources to conduct their research and purchase individual stocks or bonds. A mutual fund is managed by full-time, professional money managers who have the expertise, experience and resources to actively buy, sell, and monitor investments.
2. **Low Cost:** An important advantage of mutual funds is their low cost. Due to huge economies of scale, mutual funds schemes have a low expense ratio. Expense ratio represents the annual fund operating expenses of a scheme, expressed as a percentage of the fund's daily net assets.

Disadvantages of Mutual Fund

1. **Entry or Exit Load:** Some mutual funds may charge either entry or exit load or both. They levy this charge primarily to maintain their operations and pay staff salaries. Sometimes, the charge may go up to a high 3% of the net investment amount. However, it mostly remains around 1%. While the load might seem one of the significant disadvantages of mutual funds, funds charging a high load usually offer much higher returns than the average mutual funds. Hence, while the load certainly eats into your profit, you must analyze the funds past performance before deciding.

Portfolio

In simple term Portfolio can be defined as combination of securities that have Return and Risk characteristic of their own. Portfolio may or may not be aggregate characteristics of their individual parts. Portfolio is the collection of financial or real assets such as equity shares, debentures, bonds etc. Portfolio is a combination of assets or it consists of collection of securities. These holdings are the result of individual preferences, decisions of the holders regarding risk, return and the host of other considerations.

1.2 Types of Mutual Funds

1) **Equity Fund**

   Equity funds are those mutual funds that primarily invest in stocks. You invest your money in the fund via SIP or LUMPSUM which then invests it in various equity stocks on your behalf.

   The consequent gains or losses accrued in the portfolio affect your funds.

What Is Equity?

Equity, typically referred to as shareholders' equity (or owners' equity for privately held companies), represents the amount of money that would be returned to a company's shareholders if all of the assets were liquidated and all of the company's debt was paid off in the case of liquidation. In the case of acquisition, it is the value of company sales minus any liabilities owed by the company not transferred with the sale.

In addition, shareholder equity can represent the book value of a company. Equity can sometimes be offered as payment-in-kind. It also represents the pro-rata ownership of a company's shares.

Equity can be found on a company's balance sheet and is one of the most common pieces of data employed by analysts to assess a company's financial health.

2) **Debt Fund**

   A debt fund is a mutual fund scheme that invests in fixed instruments, such as Corporate and Government Bonds, corporate debt securities, and money market instruments etc. A debt fund is a mutual fund scheme that invests in fixed income instruments, such as Corporate and Government Bonds, corporate debt securities, and money market instruments etc. that offer capital appreciation. Debt funds are also referred to as Income Funds or Bond Funds.

Who should invest in a debt fund?

Debt funds are ideal for investors who want regular income, but are risk-averse. Debt funds are less volatile and, hence, are less risky than equity funds. If you have been saving in traditional fixed income products like Term Deposits, and looking for steady returns with low volatility, debt mutual funds could be a better option, as they help you achieve your financial goals in a more tax efficient manner and therefore earn better returns.

How debt funds are different from other mutual fund schemes?

In terms of operation, debt funds are not entirely different from other mutual fund schemes. However, in terms of safety, they score higher than equity mutual funds. For instance, when the market falls, the NAVs of your equity funds fall sharply, whereas in case of debt funds, the fall is not as sharp. Having said that, debt funds can offer only moderate returns, while equity funds, which are highly risky, offer high returns over longer time horizon.

3) **Hybrid Fund**

To simplify hybrid funds meaning it can be said that hybrid funds are a combination of equity and debt investments that are designed to meet the investment objective of the scheme.

What Is a Hybrid Fund?

A hybrid fund is an investment fund that is characterized by diversification among two or more asset classes. These funds typically invest in a mix of stocks and bonds. They may also be known as asset allocation funds.

Understanding Hybrid Funds

Hybrid funds offer investors a diversified portfolio. The term hybrid indicates that the fund strategy includes investment in multiple asset classes. In general, it can also mean that the fund uses an alternative mixed management approach.

Hybrid funds are commonly known as asset allocation funds. In the investment market, asset allocation funds can be used for many purposes. These funds offer investors an option for investing in multiple asset classes through a single fund.

Hybrid funds evolved from the implementation of modern portfolio theory in fund management. These funds can offer varying levels of risk tolerance ranging from conservative to moderate and aggressive.

- **Balanced funds** are also a type of hybrid fund. Balanced funds often follow a standard asset allocation proportion, such as 60/40.
- **Target date funds** or lifecycle funds also fit into the hybrid category. These funds invest in multiple asset classes for diversification. Target date funds vary from standard hybrid funds in that their portfolio portions begin with a more aggressive allocation and progressively rebalance to a more conservative allocation for use by a specified utilization date.

- A **blend fund** (or blended fund) is a type of equity mutual fund that includes a mix of both value and growth stocks. These funds offer investors diversification among these popular investment styles in a single portfolio.

In all cases, hybrid funds will include some mix of two or more asset classes. In risk-targeted and balanced funds, allocations will typically remain at a fixed proportion. In funds targeting a specified utilization date, the proportion of asset mix will vary over time. In all of the funds, the investment manager may actively manage the individual holdings within each asset category to respond to changing market conditions and potential capital appreciation opportunities

4) Solution Oriented Fund

Solution Oriented mutual fund is designed to help an investor achieve composite goals like marriage, child education, retirement planning are achieved.

SEBI has recently rolled out a new type of mutual fund called Solution Oriented Mutual Funds. This innovative category allows you the flexibility to tailor your mutual fund portfolio to meet specific future needs, such as funding your child's education or marriage or preparing for your retirement.

Even though this concept may seem novel, options to invest in what could now be classified under this category were previously available through equity or balanced advantage fund schemes. This distinct category, however, enables fund managers to employ specialised strategies to achieve higher returns.

With Solution-Oriented Mutual Funds, you can switch between equity mutual funds and debt funds depending on your needs. Fund managers can adjust their strategies based on your age, offering a more personalized investment experience. Some of these funds also offer the perk of tax savings. It's important to note that these funds are designed for the long haul and come with a mandatory lock-in period of 5 years.

This new approach provides a structured way to invest, ensuring that your investments align closely with your life's significant milestones. Whether you're planning for your child's future or your own retirement, Solution Oriented Mutual Funds offer a strategic path to meet those financial goals with greater precision.

Objective of the Study

1. Explore factors influencing the performance of mutual funds, such as economic indicators, interest rates, and market conditions.

2. Analyze the risk-return trade-off for different mutual fund portfolios.

3. To identify the best portfolio of securities.

4. To help the investors to decide the effective portfolio of securities.

1.4 Data Description

Source of Data

https://www.kaggle.com/datasets/ravibarnawal/mutual-funds-india-detailed

DATA: 20 Variables, 814 Observations (study of 39 banks)

Click Here For Data File in CSV Format

Sr.no	Variable Name	Type	Nature
1	Scheme_name	Object	Independent
2	Min_sip	Int64	Independent
3	Min_lumpsum	Float64	Independent
4	Expense_ratio	Float64	Independent
5	Fund_size_cr	Float64	Independent
6	Fund_age_yr	Int64	Independent
7	Fund_manager	Object	Independent
8	Sortino	Float64	Independent
9	Alpha	Float64	Independent
10	SD	Float64	Independent
11	Beta	Float64	Independent
12	Sharpe	Float64	Independent
13	Risk_level	Int64	Dependent
14	Amc_name	Object	Independent
15	Rating	Int64	Dependent
16	Category	Object	Independent
17	Sub_category	Object	Independent
18	Returns_1yr	Float64	Dependent
19	Returns_3yr	Float64	Dependent
20	Returns_5yr	Float64	Dependent

Variable Description:

Scheme_name: Name of the mutual fund scheme.

Min_sip (Systematic Investment Plan): When you set aside a small amount of money to invest in mutual fund.

Min_lumpsum: Min amount 5000 required to start. Lumpsum means putting large amount of money into the fund all at onces.

Expense_ratio: Think of expense ratio as fee you pay each year for investing in mutual fund. And it is calculated as a percentage.

$$\textbf{Expense\_ratio: } \frac{\boldsymbol{Management\ Fees}}{\boldsymbol{Total\ Investment\ in\ a\ fees}}$$

Fund_size_cr: the total amount of money that a mutual fund manager must oversee and invest.

Fund_age_yr: years since inception of scheme.

Fund_manager: A fund manager is responsible for implementing a fund's investment strategy and managing its trading activities.

Sortino: Sortino ratio measures the risk-adjusted performance as it means the investment or portfolio is generating more return for downside risk taken. Investor generally prefer higher sortino ratio.

$$\textbf{Sortino: } \frac{R_p - R_f}{\sigma_d}$$

Where, R_p= portfolio return, R_f= risk free rate, σ_d = downside standard deviation

Alpha: Alpha is the risk adjustive measure of how a security in comparison to the overall market average return. High alpha indicate better performance for fund.

Alpha = Expected return of the market - Actual return fund manager

SD: A standard deviation is a number that can be used to show how much the returns of a mutual fund scheme are likely to deviate from its average annual returns. Low SD indicate lower volatility for the fund.

$$\textbf{Sd} = \frac{\sqrt{\Sigma(X-\overline{x})^2}}{n-1}$$

Where, **X** = The Yearly mutual fund returns (Investment returns),
$\overline{x} = Average\ Return$, **n** = Total number of observations

Beta: Beta in a mutual fund is often used beta less than 1 indicate lower volatility for the fund. Beta ratio more than 1 indicate more volatility for the fund.

Beta: $\frac{\text{Fund return} - \text{Risk free rate}}{Benchmark\ return - Risk\ free\ rate}$

Sharpe: Sharpe Ratio of a mutual fund reveals its potential risk-adjusted returns

$$\textbf{Sharpe: } \frac{R_p - R_f}{\sigma_d}$$

Where, R_p= average rate of return, R_f= the best available rate of return of risk free security, σ_d = standard deviation

Risk_level: 1- Low risk, 2 - Low to moderate, 3 - Moderate, 4 - Moderately High, 5 - High, 6 - Very High

Amc_name: Mutual fund house managing the assets.

Rating: 0-5 rating assigned to scheme.

Category: The category to which the mutual fund belongs (e.g. equity, debt, hybrid)

Sub_category: It includes category like Childrens Funds, FoFs Domestic, etc.

Returns_1yr: The return percentage of the mutual fund scheme over 1 year.

Returns_3yr: The return percentage of the mutual fund scheme over 3 year.

Returns_5yr: The return percentage of the mutual fund scheme over 5year.

1.5 Computing Aids

Computing Tools for Data Analysis in Our Project

Python Code: Click Here For Python Code

✓ Python (Version 3.10.12)
✓ R- Programming (Version 4.3.2)
✓ MS-Excel (Version 2019)
✓ SPSS (Version 29)

In our project, we used a bunch of smart computer tools to look at and understand our data better. We used Python to do a lot of different things with our data because it can do so many things and make tasks easier. R software was another tool we used, and it helped us with math stuff and making charts. Excel was useful too, as it's something most people are familiar with, and it helped us organize and show our findings. And then there's SPSS, which helped us dig even deeper into our data to find important stuff. All of these tools worked together to help us figure out what our data meant and make smart decisions based on it.

1.6 Exploratory Data Analysis

Exploratory Data Analysis (EDA) is an approach for data analysis that employs a variety of techniques (mostly graphical) to

- maximize insight into a data set;
- uncover underlying structure;
- extract important variables;
- detect outliers and anomalies;
- test underlying assumptions;
- develop parsimonious models; and
- Determine optimal factor settings.

EDA is not identical to statistical graphics although the two terms are used almost interchangeably. Statistical graphics is a collection of techniques--all graphically based and all focusing on one data characterization aspect. EDA encompasses a larger venue; EDA is an approach to data analysis that postpones the usual assumptions about what kind of model the data follow with the more direct approach of allowing the data itself to reveal its underlying structure and model. EDA is not a mere collection of techniques; EDA is a philosophy as to how we dissect a data set; what we look for; how we look; and how we interpret. It is true that EDA heavily uses the collection of techniques that we call "statistical graphics".

Before you begin exploratory data analysis, it's important to understand a few key terms:

Value: A data value is a piece of information, such as a number or a date.

Variable: A data variable is a characteristic that you can measure, such as weight or income.

Distribution: The distribution of a dataset is how the dataset is spread out. You can visualize a dataset's distribution by observing its shape on a graph.

Outlier: An outlier is a data value that is significantly different, including much higher or lower, from the rest of a dataset.

Data model: A data model is a method of organizing data and relationships between values in a dataset.

1.7 Trade-off

Risk and Return Trade Off

The most fundamental principal of finance literature is that there is a trade-off between risk and return. The risk-return relationship requires that the return on a security should be commensurate with its riskiness. If the capital markets are operationally efficient, then all investment assets should be providing a rate or return that is consistent with the risks associated with them. The risk and return are directly variable, i.e. an investment with higher risk should produce higher return.

In the investing world, the dictionary definition of risk is the possibility that an investment actual return will be different than expected. Technically, this is measured in statistical by standard deviation. Risk means investor have the possibility of losing some or even all of their original investment.

Low levels of uncertainty (low risk) are associated with low potential returns. High levels of uncertainty (High risk) are associated with highest potential returns. The risk/returns are the balance between desire for the lowest possible risk and the highest possible return. This is demonstrated graphically in the chart below.

The common misconception is that higher risk equals greater return. The risk/return trade-off tells us that the highest risk gives us the possibility of higher returns. There are no guarantees just as risk means highest potential returns, it also means highest potential losses.

1.8 Cluster Analysis

Cluster analysis is multivariate methods deal with the analysis of data of more than two variables. It is data reduction technique such as cluster analysis reduces data in row. The cluster method was chosen for several reasons due to the nature and specificity of the data, as well as the advantages of cluster analysis. Cluster analysis belongs to the group of methods of classification analysis. It allows to divide a set of observations into different cluster or groups, so that the one observations in one cluster are sufficiently similar to each other. The main advantages of cluster analysis is that observation can be classified into several variables of any nature, and the required number of clusters can be determined depending on the goals of the study. There are different types of cluster analysis. However, this model is based on the k-means clustering method. The purpose of k-means is to group similar observations into clusters denoted by k. One of the drawbacks of k-means clustering is the selection of the optimal number of k (clusters). Plot dendrogram of choosing k, which guarantees the correctness of the chosen number of clusters.

Steps to perform k-means clustering:

1) Choose a value of k
2) Randomly assign each observation to an initial cluster, from 1 to K.
3) For each of the K clusters, compute the cluster centroid. Assign each observation to the cluster whose centroid is closest. Here, closest is defined using Euclidean distance.

If the value of k is pre-determined, we need a decision to fuse ith objects in cluster. This decision is taken on the basis of the value of $K_{(i)}$ where,

$$K_{(i)} = \frac{K\{Sum(i) - Min[Sum(i)]\}}{Max[Sum(i)] - Min[Sum(i)]]} + 1$$

Let the number of objects in l-th cluster can be n_l $(l = 1,2, \dots k)$.The mean of jth variable of this cluster is $\overline{X_{lj}}(j = 1,2, \dots p)$, where

$$\overline{X_{lJ}} = \frac{1}{n_l}\sum_{j=1}^{n_l} X_{lj}\,, \quad l = 1,2, \dots, k$$

Therefore, distance between i-th object and l-th cluster can be measured by d_{il} , were

$$d_{il} = \left[\sum_{j=1}^{p}(X_{ij} - \overline{X_{lJ}})^2\right]^{1/2}$$

Now, the error component due to partitioned of n objects into $p_{n,k}$ parts is given by,

$$E[p_{n,k}] = \sum_{i=1}^{n}[d_{il(i)}]^2$$

- Where $l(i)$ indicates the l-th cluster in which i-th object is merged, $d_{il(i)}$ is the Euclidean distance between i-th object and the l-th cluster in which i-th object is present and $E[p_{n,k}]$ is the error. The clustering of objects from one cluster to another cluster is done in such a way that $E[p_{n,k}]$ after transfer becomes minimum.
- Consider that the statistic to be found out to observe the change in $E[p_{n,k}]$ after transfer of object is $R_{l(i)m}$, where

$$R_{l(i)m} = \frac{n_m {d_{im}}^2}{n_m + 1} - \frac{n_l {d_{il}}^2}{n_l - 1} \text{ , } l \neq m = 1,2, \dots p$$

Here, n_m is the number of objects in m-th cluster, n_l is the number of objects in l-th cluster. If $R_{l(i)m}$>0, the transfer of i-th object from l-th cluster to m-th cluster is not fruitful and $E[p_{n,k}]$ increase.

But, $R_{l(i)m}$ <0 indicates that the transfer is effective. In that case clustering is to be renewed. Due to this re-clustering, there is a change in the error component $[p_{n,k}]$. The new error component is measured by,

$$E[p'_{n,k}] = E[p_{n,k}] + R_{l(i)m}$$

Here, $p'_{n,k}$ is the number of partitioned of objects into k clusters. The value of $R_{l(i)m}$ is to be calculated after each transfer of object and the process is continued until $R_{l(i)m}$>0.

2. RESEARCH METHODOLOGY

2.1 Dataset

```
## Import all important libraries
import pandas as pd
import numpy as np
import statsmodels.api as sm
```

```
## Import the Mutual fund data
df_new = pd.read_csv("/content/Mutual Fund New.csv")
df_new
```

	scheme_name	min_sip	min_lumpsum	expense_ratio	fund_size_cr	fund_age_yr	fund_manager	sortino	alpha	sd	beta	sharpe	risk_level	amc_name	rating	category	sub_categor
0	Aditya Birla SL Active Debt Multi-Mgr FoF-Dir ...	100	5181.818182	0.27	10.0	10	Kaustubh Gupta	0.32	2.24	9.39	0.01	0.24	3	Aditya Birla Sun Life Mutual Fund	3	Other	Fol Domes
1	Aditya Birla SL Arbitrage Fund	1000	5181.818182	0.36	4288.0	10	Lovelish Solanki	1.33	1.53	0.72	0.56	1.10	1	Aditya Birla Sun Life Mutual Fund	3	Hybrid	Arbitra Mutual Fun
2	Aditya Birla SL Asset Allocator FoF-Dir Growth	1000	5181.818182	0.53	157.0	10	Vinod Narayan Bhat	3.44	2.67	10.58	0.67	1.42	5	Aditya Birla Sun Life Mutual Fund	3	Other	Fol Domes
3	Aditya Birla SL Bal Bhavishya Yojna â€“ Dir Gr...	500	5181.818182	0.76	637.0	4	Atul Penkar	2.18	-6.37	14.99	0.85	0.90	6	Aditya Birla Sun Life Mutual Fund	2	Solution Oriented	Childre Fun

2.2 Data Preprocessing

```
## check the null values
df_new.isnull().sum()
```

Output: In our data not present any null values.

```
## Check the variables name
df_new.columns
```

Output:

Index(['scheme_name', 'min_sip', 'min_lumpsum', 'expense_ratio', 'fund_size_cr', 'fund_age_yr', 'fund_manager', 'sortino', 'alpha', 'sd', 'beta', 'sharpe', 'risk_level', 'amc_name', 'rating', 'category', 'sub_category', 'returns_1yr', 'returns_3yr', 'returns_5yr'], dtype='object')

The formulas provided are commonly used in finance and investment analysis to assess the performance and risk of investment portfolios. Here's a brief explanation of each formula:

Portfolio Risk: The sum of standard deviation (sd) and beta for each investment in the portfolio. Standard deviation measures the volatility or dispersion of returns from the mean, while beta measures the sensitivity of an investment's returns to market movements. Adding these two measures provides an estimate of the overall risk of the portfolio.

Sharpe Ratio: The ratio of excess return (return minus risk-free rate) to standard deviation. It measures the risk-adjusted return of an investment or portfolio. A higher Sharpe ratio indicates better risk-adjusted performance.

Sortino Ratio: Similar to the Sharpe ratio, but it only considers the downside risk, which is defined as the standard deviation of negative returns or returns below a certain threshold (e.g., risk-free rate or target return). It provides a measure of risk-adjusted return that focuses on minimizing downside risk.

These formulas are commonly used to evaluate investment portfolios because they provide insights into both the performance and risk characteristics of the portfolio, helping investors make informed decisions about their investments.

```
## Adding new data variables
## Calculate Portfolio Risk
df_new['portfolio_risk'] = df_new['sd'] + df_new['beta']

### Calculate Sharpe Ratio
df_new['sharpe_ratio'] = df_new['alpha'] / df_new['sd']

### Calculate Sortino Ratio
df_new['sortino_ratio'] = df_new['alpha'] / df_new['sortino']
```

Output

portfolio_risk	sharpe_ratio	sortino_ratio
9.40	0.238552	7.000000
1.28	2.125000	1.150376
11.25	0.252363	0.776163
15.84	-0.424950	-2.922018

```
## Calculate the summary of the data
df_new.describe()
```

Output:

	returns_1yr	returns_3yr	returns_5yr	portfolio_risk	sharpe_ratio	sortino_ratio
count	814.000000	814.000000	814.000000	814.000000	814.000000	814.000000
mean	3.921376	18.524693	9.490577	11.094742	0.837470	1.027999
std	6.675502	11.951072	3.310660	7.427464	1.230465	4.701076
min	-19.700000	3.300000	-4.100000	0.720000	-0.682274	-31.500000
25%	1.500000	6.300000	7.100000	3.350000	0.068552	0.234804
50%	4.400000	18.500000	9.490000	12.610000	0.306439	0.854545
75%	5.600000	27.000000	11.500000	17.257500	1.026287	1.531219
max	130.800000	71.400000	23.200000	45.860000	5.851852	115.000000

- In our data, high 1st year return 130% & low 1st year return -19%
- If you want high return - high risk then sharpe ratio will be low When low risk - low return then then sharpe ratio will be high
- If market down & you want lower volatility & higher returns then we'll choose higher sortino ratio.

(Reference : Disha Thakkar (ICICI Mutual Fund, Anand Gujarat))

2.3 Exploratory Data Analysis

1. Bar Plot for Category Vs Schemes_count
2. Violin Plot for Category Vs Expense Ratio
3. Bar Plot for amc_name Vs Number of schemes
4. Multiple Bar Plot for Category Vs Return
5. Bar Plot for sub-category Vs Mutual funds
6. Pie chart for Risk level
7. Correlation heat map for all variables

1. Bar Plot for Category Vs Schemes_count

From Fig 2.3.1 Category and Schemes_count plot can shows that 80 mutual fund in Other category, 116 mutual fund in. Hybrid category, 28 mutual fund in Solution oriented category, 308 mutual fund in Equity category, 282 mutual fund in Debt category. Equity is the highest scheme count in our mutual fund data

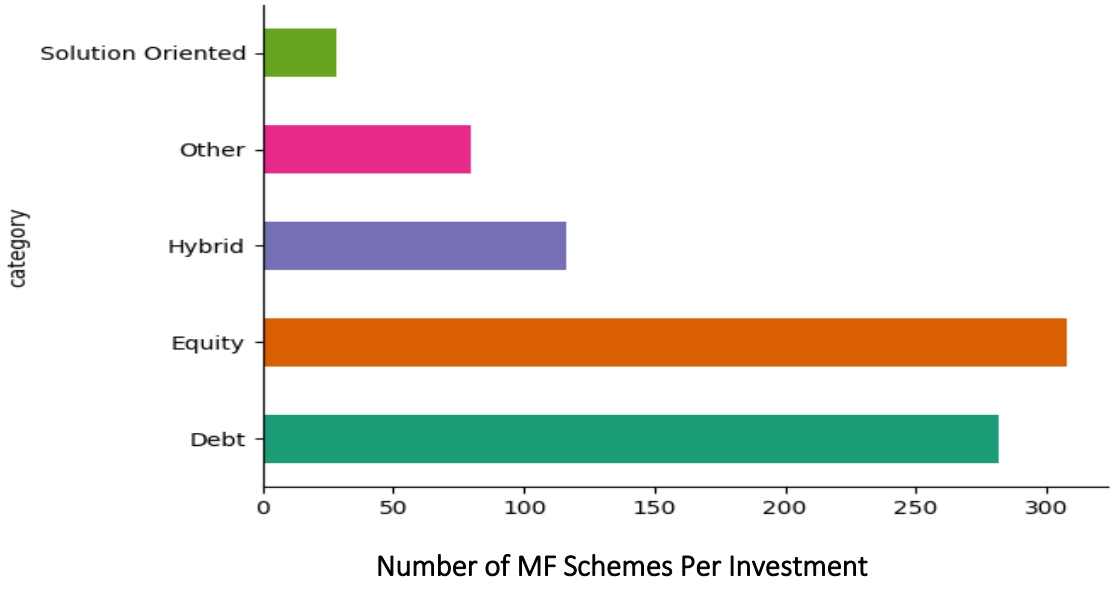

Fig 2.3.1

2. Violin Plot for Category Vs Expense Ratio

We need to check which category show high expense ratio.

Then we have plotted the violin plot, A violin plot is a type of data visualization that combines the features of a box plot and a kernel density plot. It is used to visualize the distribution of numeric data across different categories or groups.

Here are some key features and uses of violin plots:

Shape and Width: The violin plot resembles a violin or a mirrored density plot. The width of the "violin" at any given point represents the density of data at that value. Wider sections indicate a higher density of data points, while narrower sections indicate lower density.

Central Line: The central line within the "violin" represents the median of the data distribution.

Interquartile Range (IQR): The thicker portions of the "violin" represent the interquartile range (IQR), which encompasses the middle 50% of the data. This is similar to the box in a box plot.

Kernel Density Estimation: The smoothed curve inside the violin plot represents the kernel density estimate of the underlying data distribution. It gives a visual representation of the probability density function of the data.

Grouping and Comparison: Violin plots are particularly useful for comparing the distributions of multiple groups or categories within a dataset. They allow you to quickly assess differences in central tendency, spread, and shape across different groups.

Outlier Detection: Like box plots, violin plots can also indicate the presence of outliers in the data, especially in the thin tails of the "violin".

Violin plots are commonly used in exploratory data analysis to visualize the distribution of data and identify patterns or differences between groups. They are especially useful when comparing multiple groups or when dealing with large datasets where individual data points may not be easily discernible.

```
## Plot the violin Plot category Vs expense_ratio'

sns.violinplot(df_new, x='expense_ratio', y='category', inner='box',
palette='Dark2')
```

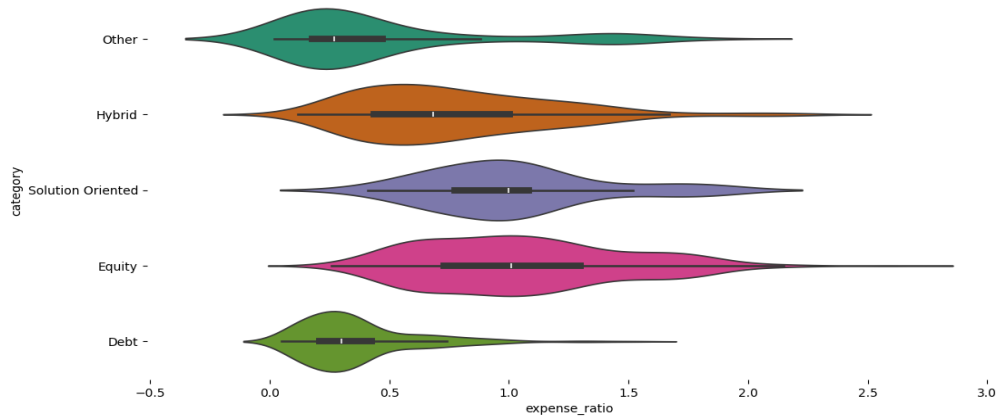

Fig 2.3.2

From Fig 2.3.2 Violin plot we can say that Equity has show high (**2.59%**) Expense Ratio. Expense ratio as fee you pay each year for investing in mutual fund. And it is calculated as a percentage.

3. **Bar Plot for amc_name Vs Number of schemes**

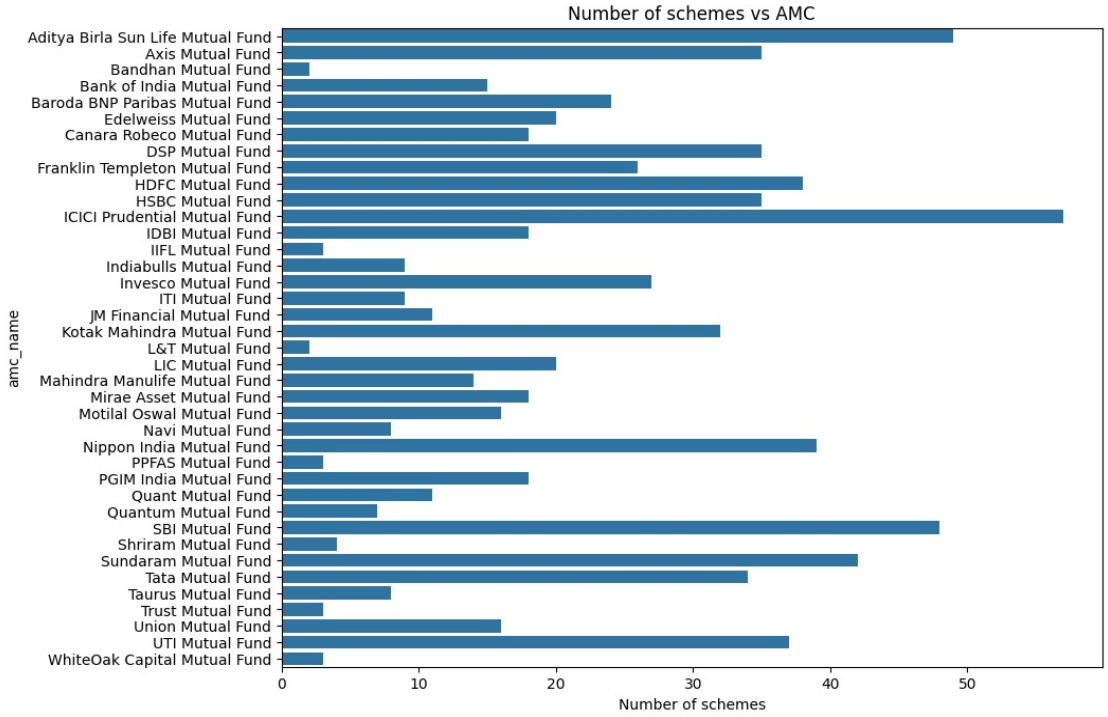

Fig 2.3.3

From Fig 2.3.3, Mutual fund data there are 39 amc (assets managing company) name and 814 schemes. From this plot we can say that ICICI Prudential Mutual Fund has around 57 schemes running.

4. **Multiple Bar Plot for Category Vs Return**

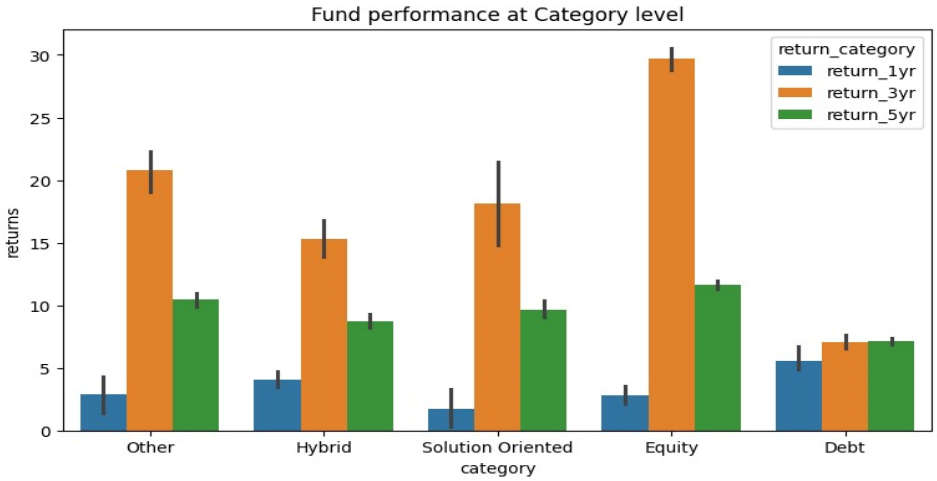

Fig 2.3.4

From Fig 2.3.4, the above chart, Equity category funds gives the highest returns as compared to another category.

5. **Bar Plot for sub-category Vs Mutual funds**

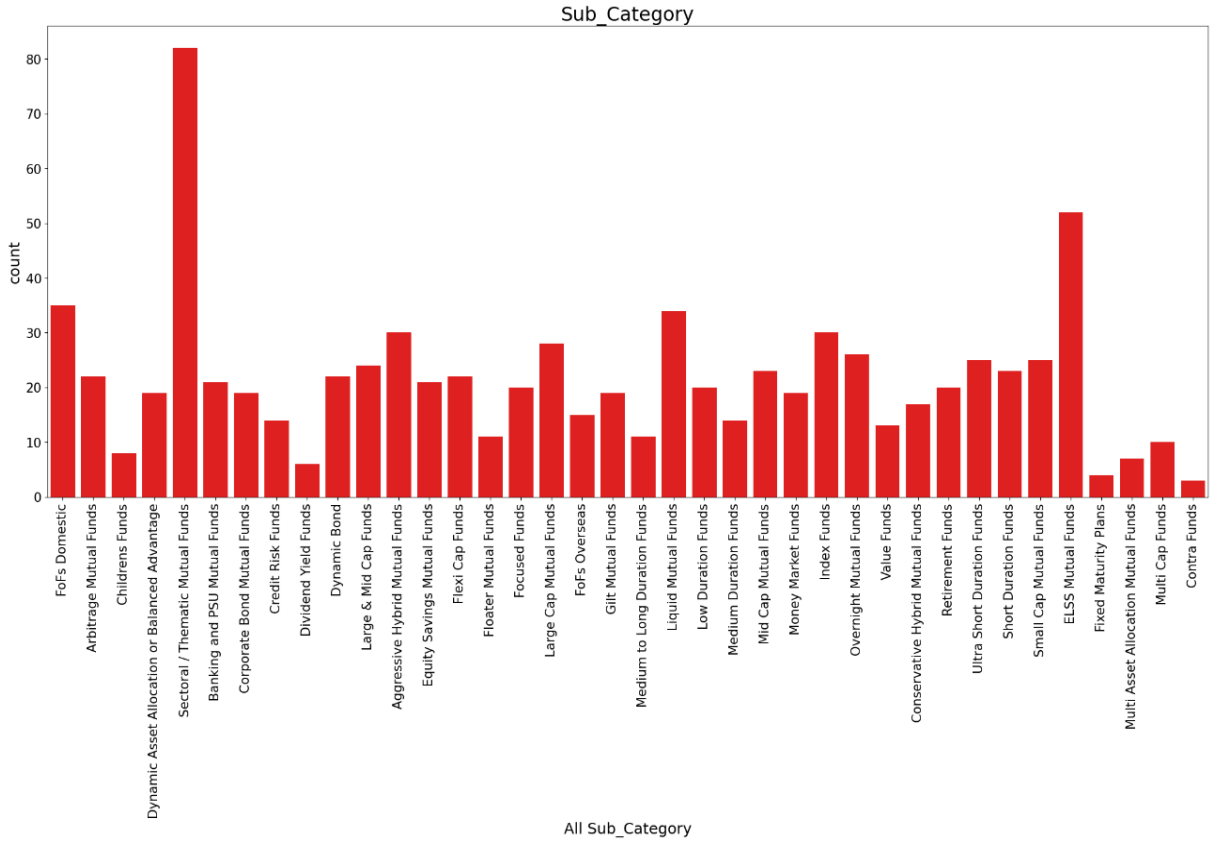

Fig 2.3.5

From Fig 2.3.5, There are 38 sub-category of mutual fund in this plot can show sectoral/Thematic mutual fund has higher count than other mutual fund.

6. **Pie chart for Risk level**

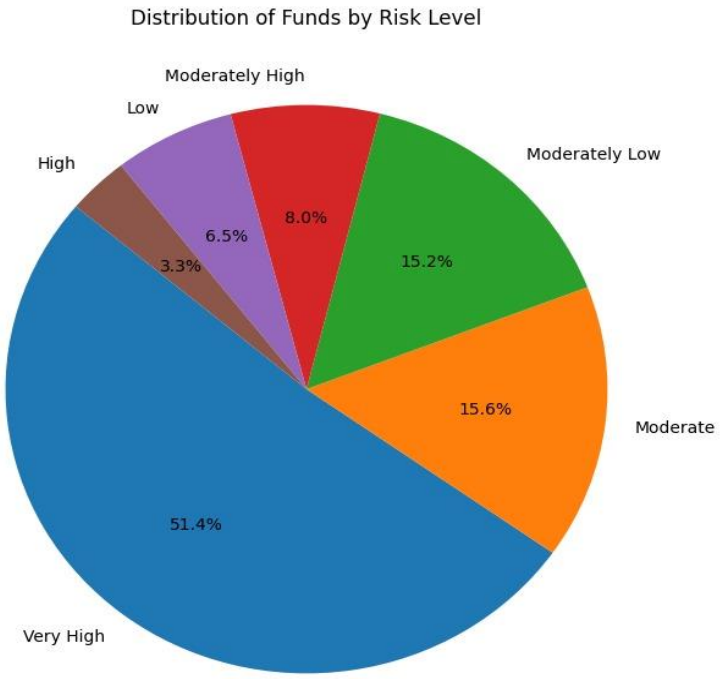

Fig 2.3.6

From Fig 2.3.6, Here most of the schemes have risk level Very High. i.e. 51.4%

7. **Check the correlation between all variables**

```
    ## Check the correlation between all variables
import seaborn as sns
import matplotlib.pyplot as plt

# Drop non-numeric columns
numeric_df = df_new.select_dtypes(include=['float64', 'int64'])

# Compute the correlation matrix
correlation_matrix = numeric_df.corr()
```

```
# Create a heatmap using Seaborn
plt.figure(figsize=(14, 12))
sns.heatmap(correlation_matrix, annot=True, cmap='coolwarm', fmt=".2f",
linewidths=.6)
plt.title('Correlation Heat Map')
plt.show()
```

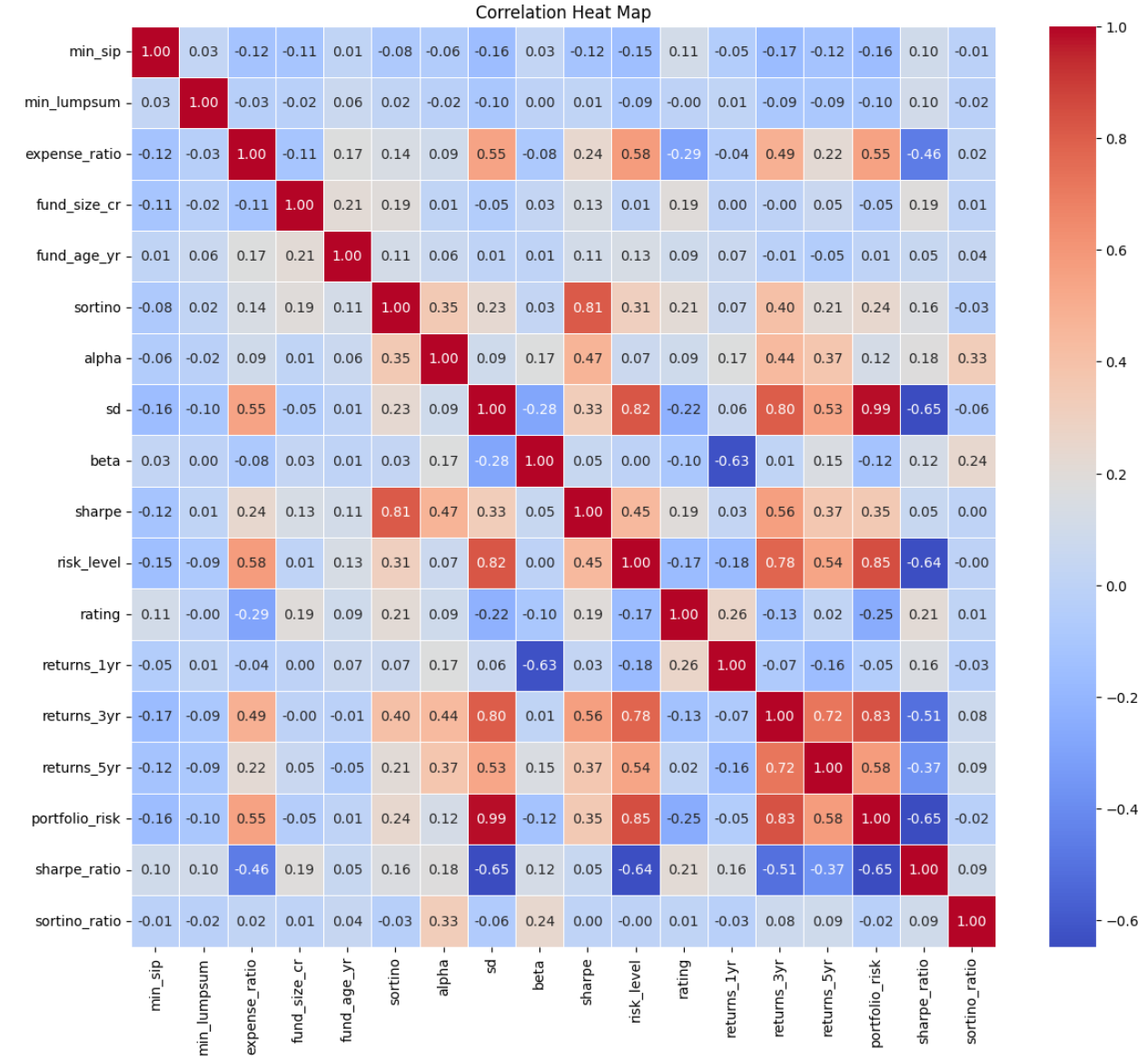

Fig 2.3.7

- Correlation heat map (Fig 2.3.7) show that high correlation among standard deviation & portfolio_risk, returns_3yr

- Moderate correlation among standard deviation & risk_level, sortino & sharpe, returns_3yr & returns_5yr, returns_3yr & portfolio_risk, risk_level & portfolio_risk.

Interpretation: Heatmap show that

Each square shows the correlation between the variables on each axis. Correlation ranges from -1 to +1. Values closer to zero means there is no linear trend between the two variables. The close to 1 the correlation is the more positively correlated they are; that is as one increases so does the other and the closer to 1 the stronger this relationship is. A correlation closer to -1 is similar, but instead of both increasing one variable will decrease as the other increases. The dark red because those squares are correlating each variable to itself (so it's a perfect correlation). For the rest the larger the number and darker the color the higher the correlation between the two variables. The plot is also symmetrical about the diagonal since the same two variables are being paired together in those squares.

2.4 Fundamental Principal - Trade-off

Risk-return trade-off for returns_1yr

```
plt.scatter(df_new['portfolio_risk'], df_new['returns_1yr'],
c=df_new['sharpe_ratio'], cmap='viridis')
plt.colorbar(label='sharpe_ratio')
plt.xlabel('portfolio risk')
plt.ylabel('returns_1yr')
plt.title('risk-return trade-off : returns_1yr')
plt.show()
```

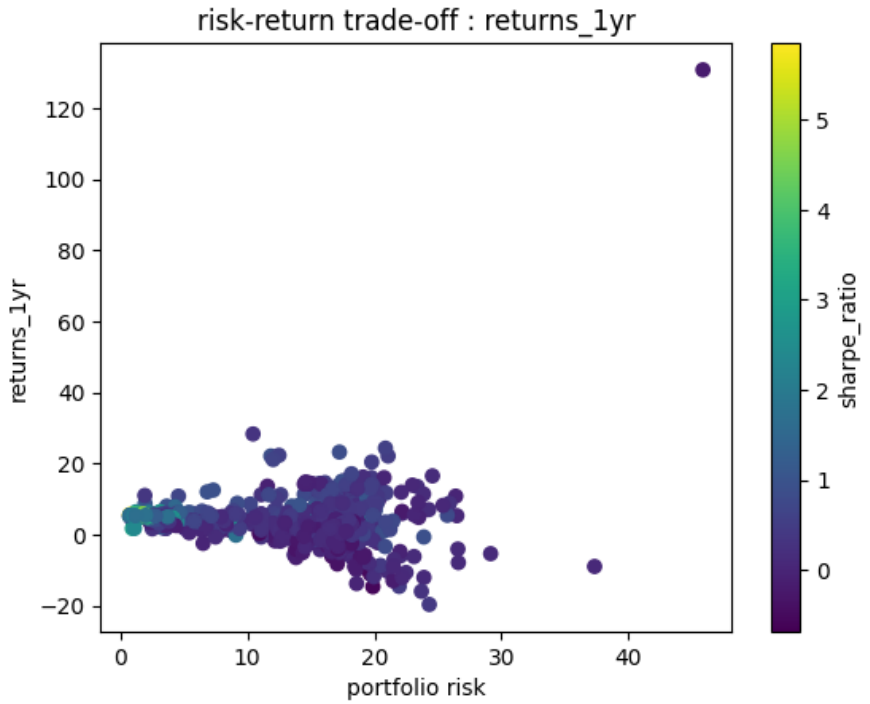

Fig 2.4.1

From Fig 2.4.1, risk-return trade-off: 1st year return.

risk-return trade-off for returns_3yr

```
plt.scatter(df_new['portfolio_risk'], df_new['returns_3yr'],
c=df_new['sharpe_ratio'], cmap='viridis')

plt.colorbar(label='sharpe_ratio')

plt.xlabel('portfolio risk')

plt.ylabel('returns_3yr')

plt.title('risk-return trade-off : returns_3yr')


plt.show()
```

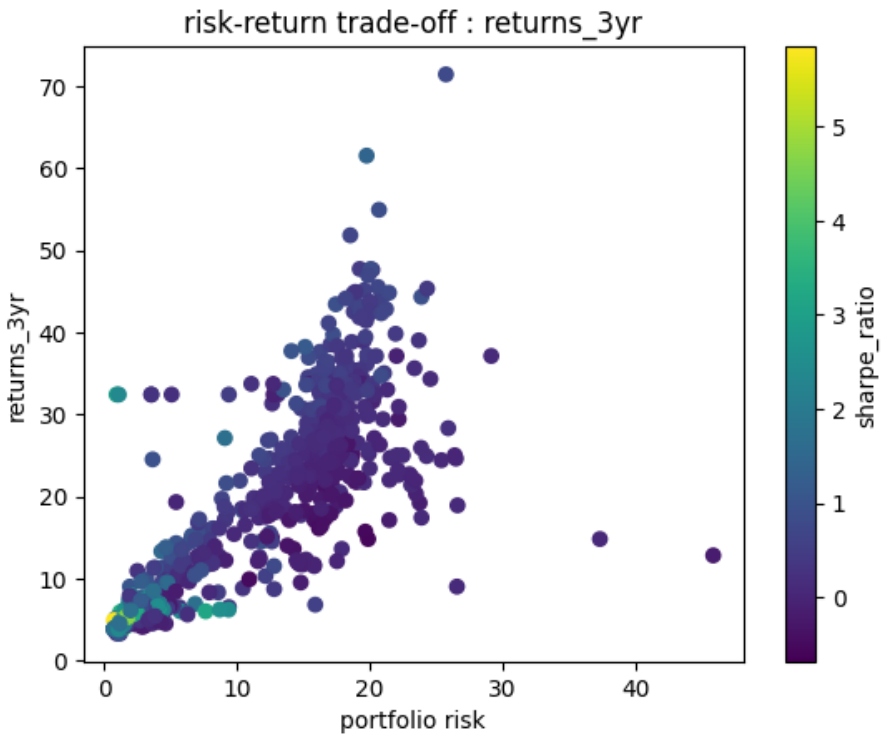

Fig 2.4.2

From Fig 2.4.2, risk-return trade-off: 3rd year return.

risk-return trade-off for returns_5yr

```
plt.scatter(df_new['portfolio_risk'], df_new['returns_5yr'],
c=df_new['sharpe_ratio'], cmap='viridis')

plt.colorbar(label='sharpe_ratio')

plt.xlabel('portfolio risk')

plt.ylabel('returns_5yr')

plt.title('risk-return trade-off : returns_5yr')


plt.show()
```

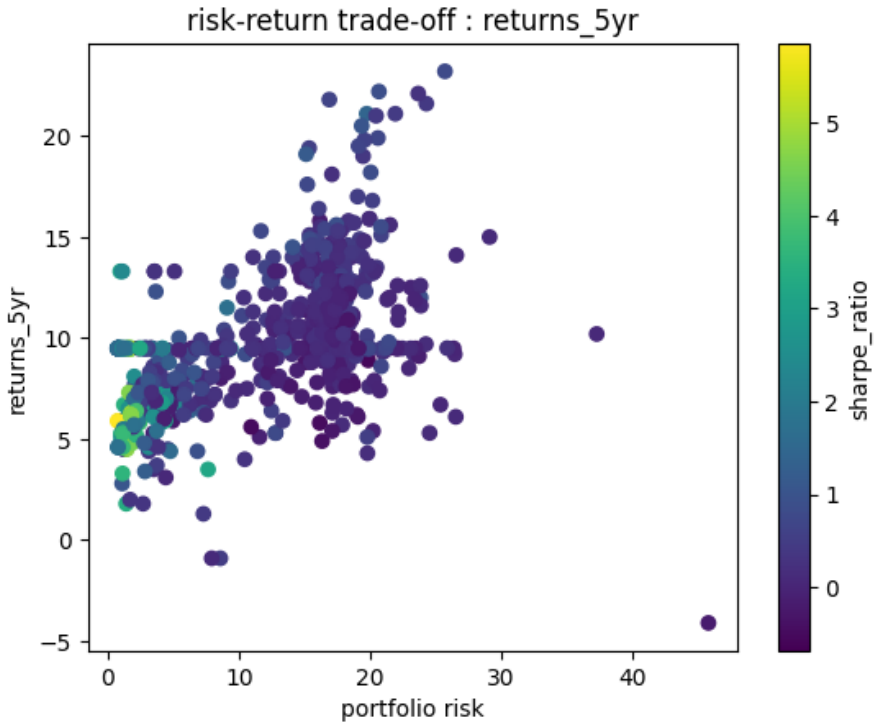

Fig 2.4.3

From Fig 2.4.3, risk-return trade-off: 5th year return.

Calculate mean of 1st year return for each category & check the which category gives the best return.

```
#### Calculate mean portfolio returns for each category

mean_returns_by_category =
df_new.groupby('category')['returns_1yr'].mean()


## mean returns for each category

print(mean_returns_by_category)

plt.plot(mean_returns_by_category)

plt.show()
```

Output

```
category
Debt                5.547872
Equity              2.842532
Hybrid              4.075000
Other               2.891250
Solution Oriented   1.714286

Name: returns_1yr, dtype: float64
```

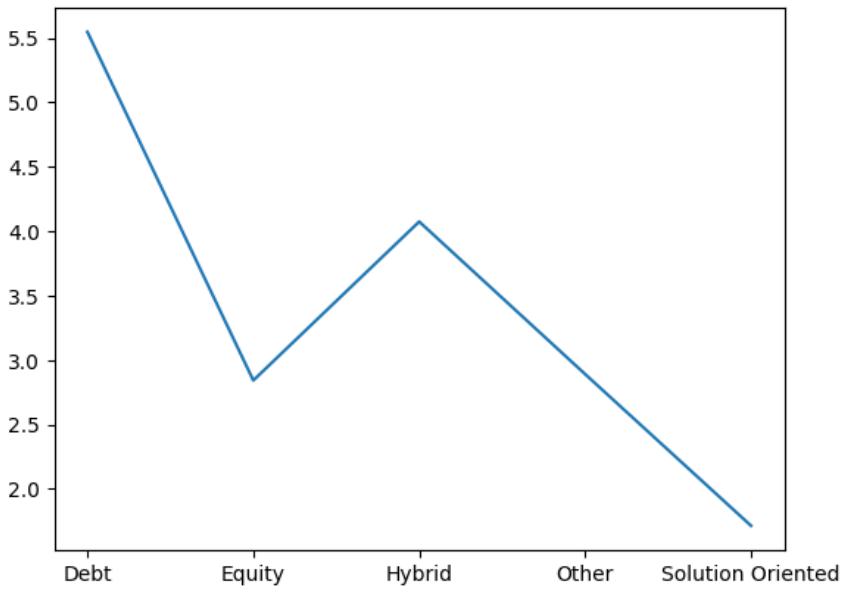

Fig 2.4.4(Mean Profile Plot of 1st Year Returns Category)

Interpretation

The highest return in the "returns_1yr" category is observed in the "Debt" category, with a return value of 5.547872. This indicates that investments classified under the "Debt" category have provided the highest returns over the past year compared to other categories such as Equity, Hybrid, Other, and Solution Oriented.

Calculate mean of 3rd year return for each category & check the which category gives the best return

```
#### Calculate mean portfolio returns for each category

mean_returns_by_category =
df_new.groupby('category')['returns_3yr'].mean()

## mean returns for each category

print(mean_returns_by_category)

plt.plot(mean_returns_by_category)

plt.show()
```

Output

```
category
Debt                   7.055319
Equity                29.701623
Hybrid                15.272414
Other                 20.763750
Solution Oriented     18.167857
Name: returns_3yr, dtype: float64
```

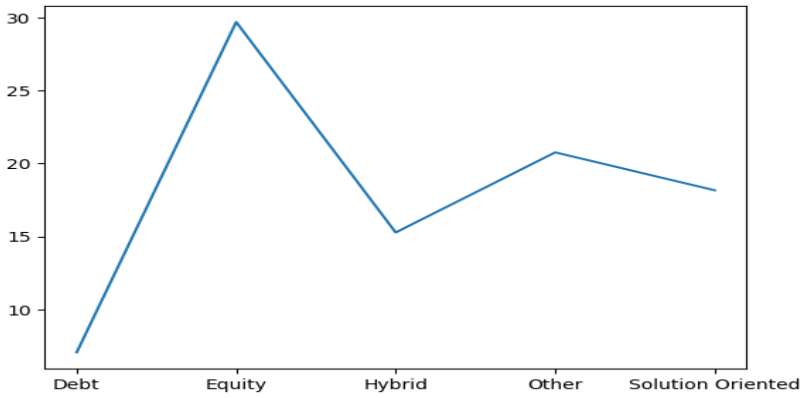

Fig 2.4.5 (Mean Profile Plot of 3rd Year Returns Category)

Interpretation

"Equity" has the highest return among the listed categories, with a return of approximately 29.70%. It suggests that investing in equity securities, such as stocks, has yielded the highest return over the specified five-year period compared to other categories like debt, hybrid, other, and solution-oriented investments.

Calculate mean of 5th year return for each category & check the which category gives the best return.

```
#### Calculate mean portfolio returns for each category
mean_returns_by_category =
df_new.groupby('category')['returns_5yr'].mean()

## mean returns for each category
print(mean_returns_by_category)
plt.plot(mean_returns_by_category)
plt.show()
```

Output

```
category

Debt                    7.145071
Equity                 11.650974
Hybrid                  8.745603
Other                  10.466750
Solution Oriented       9.646071
Name: returns_5yr, dtype: float64
```

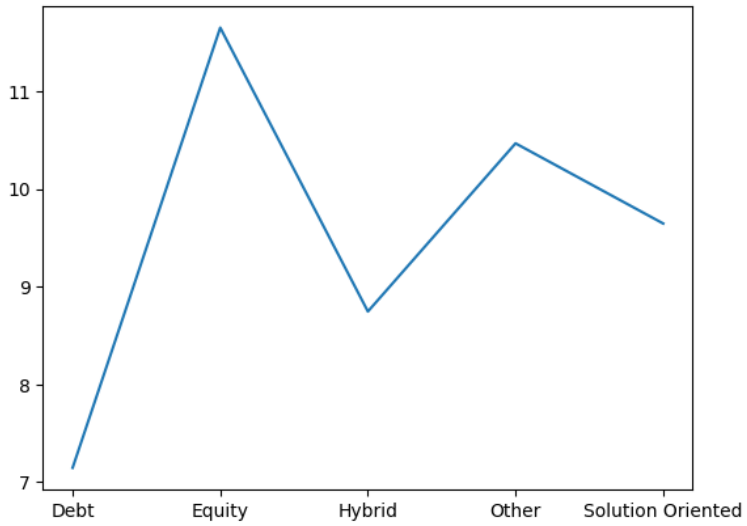

Fig 2.4.6 (Mean Profile Plot of 5th Year Returns Category)

Interpretation

"Equity" has the highest return among the listed categories, with a return of approximately 11.65%. It suggests that investing in equity securities, such as stocks, has yielded the highest return over the specified five-year period compared to other categories like debt, hybrid, other, and solution-oriented investments.

When Sharpe ratio is low which is less than 0.5 then risk return tread off show high risk and moderate-high return.

```
# Filter the DataFrame for schemes with less than 0.5 Sharpe ratio
low_sharpe_df = df_new[df_new['sharpe_ratio'] < 0.5]

# Sort the filtered DataFrame based on 'portfolio_risk' and
'portfolio_returns'
sorted_df = low_sharpe_df.sort_values(by=['portfolio_risk',
'returns_3yr'], ascending=[False, False])

# Extract the top 10 schemes
top_10_schemes = sorted_df.head(10)

# Extract scheme names of the top 10 schemes
top_10_scheme_names = top_10_schemes['scheme_name'].tolist()

# Print the top 10 scheme names
print("Top 10 schemes with high 'portfolio_risk' and 'returns_3yr' and
less than 0.5 Sharpe ratio:")
for scheme_name in top_10_scheme_names:
    print(scheme_name)
```

Output

```
Top 10 schemes with high 'portfolio_risk' and 'returns_3yr' and less
than 0.5 Sharpe ratio:

Bank of India Credit Risk Fund
DSP World Gold Fund
DSP World Mining Fund
PGIM India GEO Fund
Edelweiss Greater China Equity Off-Shore Fund
Aditya Birla SL Banking&Financial Services-DirGrowth
Motilal Oswal Nifty Bank Index Fund
Nippon India Banking&Financial Services-DirGrowth
UTI-Banking and Financial Services Fund
HDFC Infrastructure Fund
```

Interpretation

Here we can say that from the Trade-off with less Sharpe_ratio. i.e. High risk & low return. Which can be dangerous for investors.

When Sortino ratio is low which is less than 0.5 then risk return tread off show high risk and moderate-high return.

```
# Filter the DataFrame for schemes with less than 0.5 sortino_ratio
low_sortino_df = df_new[df_new['sortino_ratio'] < 0.5]

# Sort the filtered DataFrame based on 'portfolio_risk' and
'portfolio_returns'
sorted_df = low_sortino_df.sort_values(by=['portfolio_risk',
'returns_3yr'], ascending=[False, False])

# Extract the top 5 schemes
top_5_schemes = sorted_df.head(10)

# Plot the scatter plot
plt.scatter(df_new['portfolio_risk'], df_new['returns_3yr'],
c=df_new['sortino_ratio'], cmap='viridis')

# Add color bar
plt.colorbar(label='sortino_ratio')

# Label the points with portfolio_id for top 5 schemes
for i, txt in enumerate(top_5_schemes['scheme_name']):
    plt.text(top_5_schemes['portfolio_risk'].iloc[i],
top_5_schemes['returns_3yr'].iloc[i], txt, fontsize=8)

plt.xlabel('Portfolio Risk')
plt.ylabel('returns_3yr')
plt.title('Risk-Return Trade-off (Top 10 schemes with high portfolio risk
and moderate-high returns and less than 0.5 sortino_ratio)')


plt.show()
```

```
# Extract the top 10 schemes
top_10_schemes = sorted_df.head(10)

# Extract scheme names of the top 10 schemes
top_10_scheme_names = top_10_schemes['scheme_name'].tolist()
```

Output

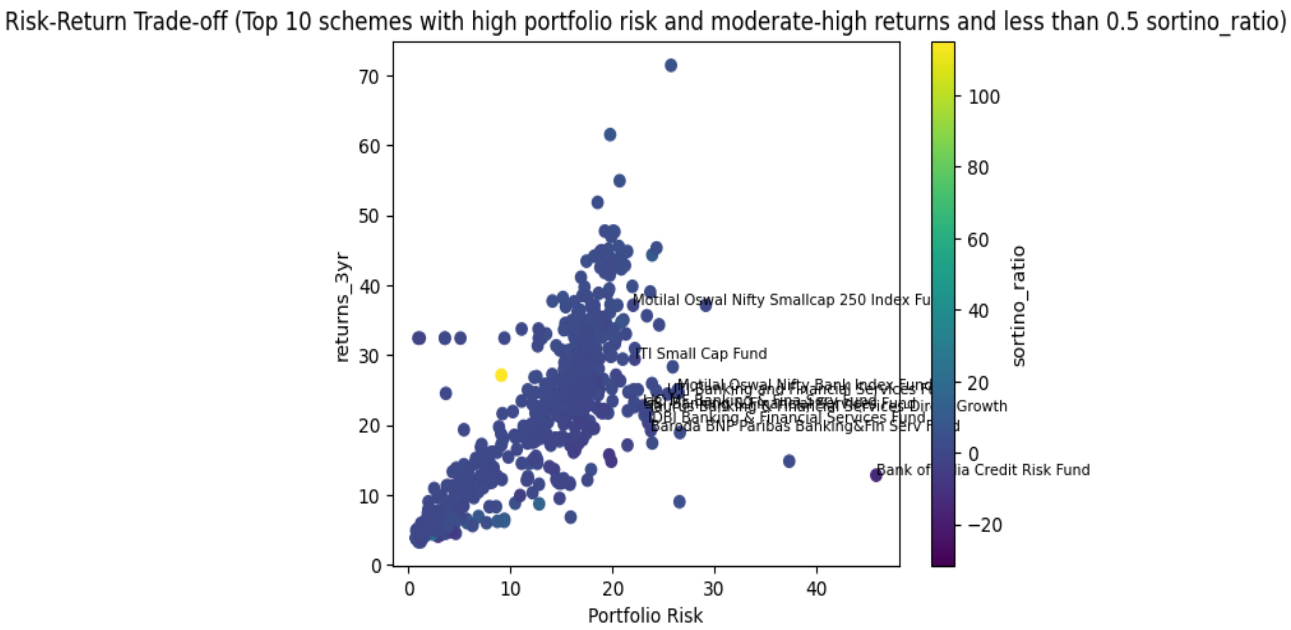

Fig 2.4.7

Top 10 schemes with high 'portfolio_risk' and 'returns_3yr' and less than 0.5 sortino_ratio:

```
Bank of India Credit Risk Fund
Motilal Oswal Nifty Bank Index Fund
UTI-Banking and Financial Services Fund
Baroda BNP Paribas Banking&Fin Serv Fund
IDBI Banking & Financial Services Fund
Taurus Banking & Financial Services-Direct Growth
SBI Banking & Financial Services Fund
LIC MF Banking & Fina Serv Fund
ITI Small Cap Fund
Motilal Oswal Nifty Smallcap 250 Index Fund
```

Interpretation

Here we can say that from the Trade-off with less sortino_ratio. i.e. High risk & low return. Which can be dangerous for investors.

Let's find the schemes of 1st year return low risk with high return as compare to sharpe_ratio.

```
# Filter the DataFrame based on the specified conditions
filtered_df = df_new[(df_new['portfolio_risk'] < 15) &
(df_new['returns_1yr'] > 20) ]

# Sort the filtered DataFrame based on 'portfolio_risk' and
'portfolio_returns'
sorted_df = filtered_df.sort_values(by=['portfolio_risk',
'returns_1yr'], ascending=[True, False])

# Extract the top 5 schemes
top_5_schemes = sorted_df.head(10)

# Extract scheme names of the top 5 schemes
top_5_scheme_names = top_5_schemes['scheme_name'].tolist()

# Plot the scatter plot
plt.scatter(filtered_df['portfolio_risk'],
filtered_df['returns_1yr'], c=filtered_df['sharpe_ratio'],
cmap='viridis')

# Add color bar
plt.colorbar(label='sharpe_ratio')

# Label the points with scheme names for the top 5 schemes
for i, txt in enumerate(filtered_df['scheme_name']):
    if txt in top_5_scheme_names:
        plt.text(filtered_df['portfolio_risk'].iloc[i],
filtered_df['returns_1yr'].iloc[i], txt, fontsize=10)

plt.xlabel('Portfolio Risk')
plt.ylabel('Portfolio Returns')
plt.title('Risk-Return Trade-off (Portfolio Risk < 15 & returns_1yr >
20)')

plt.show()
top_5_scheme_names = top_5_scheme_names[:10]  # Ensure only top 5
scheme names are considered
print(top_5_scheme_names)
```

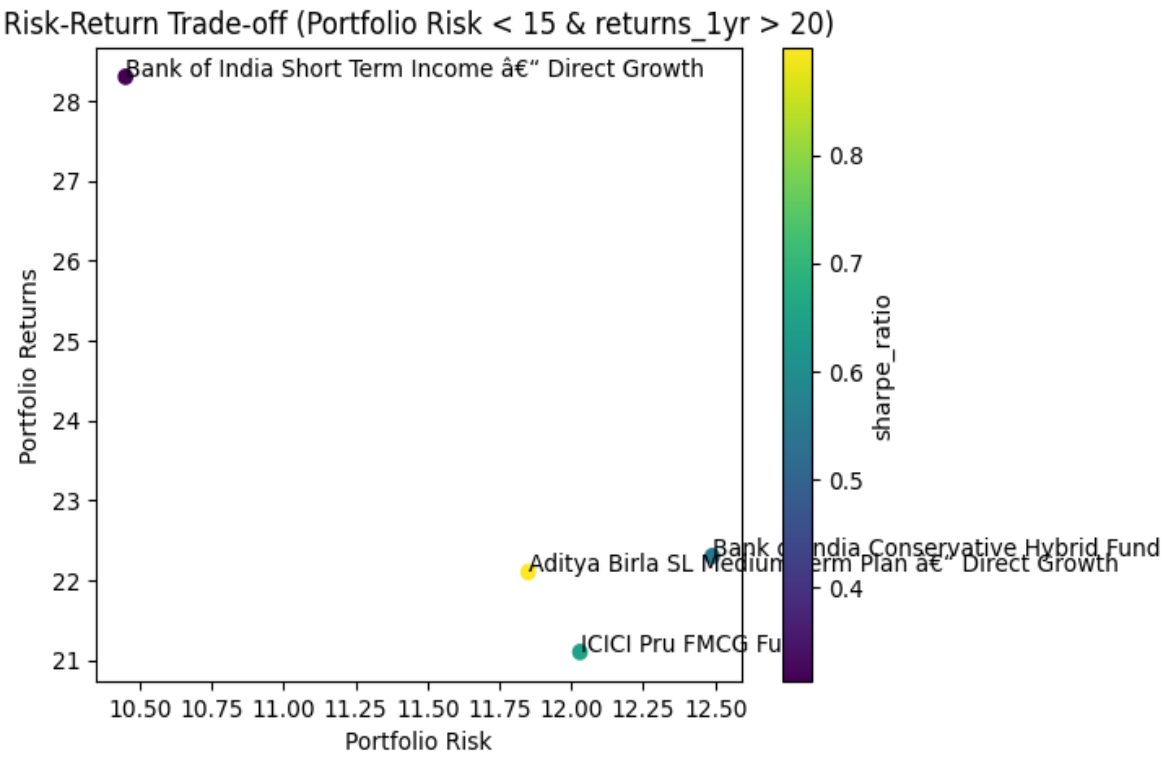

Fig 2.4.8

```
Bank of India Short Term Income â€"Direct Growth

Aditya Birla SL Medium Term Plan â€"Direct Growth

ICICI Pru FMCG Fund

Bank of India Conservative Hybrid Fund
```

Interpretation

- Schemes of 1st year return low risk with high return. So we can invest in those all following fund for 1st year return

  We have extract top schemes:

  - Bank of India Short Term Income â€“ Direct Growth
  - Aditya Birla SL Medium Term Plan â€“ Direct Growth
  - ICICI Pru FMCG Fund
  - Bank of India Conservative Hybrid Fund

Let's find the schemes of 3rd year return low risk with high return as compare to sharpe_ratio.

```
# Filter the DataFrame based on the specified conditions
filtered_df = df_new[(df_new['portfolio_risk'] < 15) &
(df_new['returns_3yr'] > 20) ]

# Sort the filtered DataFrame based on 'portfolio_risk' and
'portfolio_returns'
sorted_df = filtered_df.sort_values(by=['portfolio_risk',
'returns_3yr'], ascending=[True, False])

# Extract the top 5 schemes
top_5_schemes = sorted_df.head(10)

# Extract scheme names of the top 5 schemes
```

```
top_5_scheme_names = top_5_schemes['scheme_name'].tolist()

# Plot the scatter plot
plt.scatter(filtered_df['portfolio_risk'], filtered_df['returns_3yr'],
c=filtered_df['sharpe_ratio'], cmap='viridis')

# Add color bar
plt.colorbar(label='sharpe_ratio')

# Label the points with scheme names for the top 5 schemes
for i, txt in enumerate(filtered_df['scheme_name']):
    if txt in top_5_scheme_names:
        plt.text(filtered_df['portfolio_risk'].iloc[i],
filtered_df['returns_3yr'].iloc[i], txt, fontsize=10)

plt.xlabel('Portfolio Risk')
plt.ylabel('Portfolio Returns')
plt.title('Risk-Return Trade-off (Portfolio Risk < 15 & returns_3yr >
20)')

plt.show()
top_5_scheme_names = top_5_scheme_names[:10]  # Ensure only top 5 scheme
names are considered
print(top_5_scheme_names)
```

Output

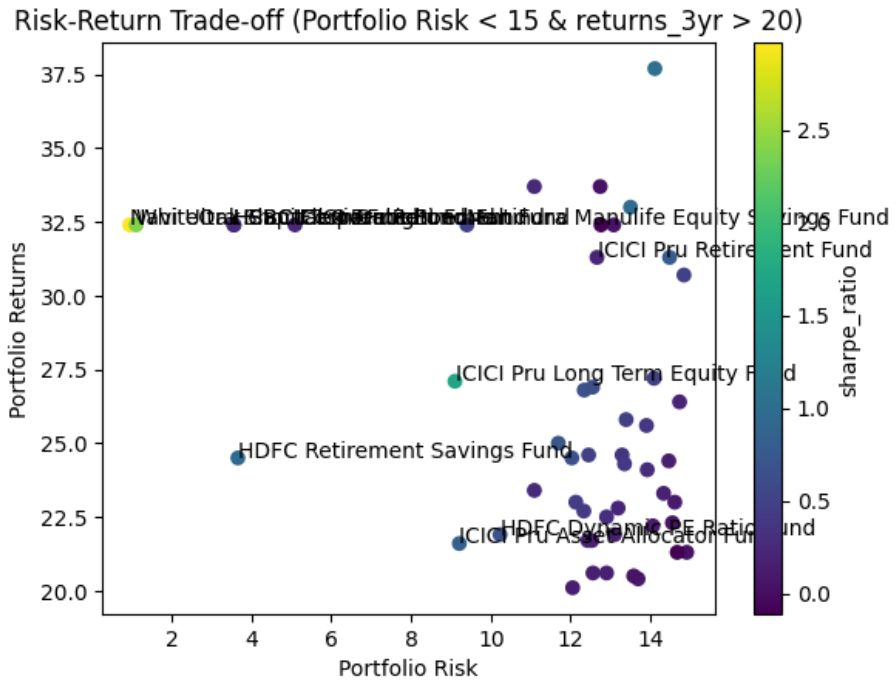

Fig 2.4.9

```
'Navi Ultra Short Term Fund', 'WhiteOak Capital Overnight Fund', 'HSBC
Corporate Bond Fund', 'HSBC Flexi Debt Fund', 'HDFC Retirement Savings
Fund', 'ICICI Pru Retirement Fund', 'ICICI Pru Long Term Equity Fund',
'ICICI Pru Asset Allocator Fund', 'Mahindra Manulife Equity Savings
Fund', 'HDFC Dynamic PE Ratio Fund'
```

Interpretation

- schemes of 3rd year return low risk with high return. So we can invest in those all following fund for 3rd year return

  We have extract top 10 schemes:

  - 1.Navi Ultra Short-Term Fund
  - 2.WhiteOak Capital Overnight Fund
  - 3.HSBC Corporate Bond Fund
  - 4.HSBC Flexi Debt Fund
  - 5.HDFC Retirement Savings Fund
  - 6.ICICI Pru Retirement Fund
  - 7.ICICI Pru Long Term Equity Fund
  - 8.ICICI Pru Asset Allocator Fund
  - 9.Mahindra Manulife Equity Savings Fund
  - 10.HDFC Dynamic PE Ratio Fund

Let's find the schemes of 5th year return low risk with high return as compare to sharpe_ratio.

```
# Filter the DataFrame based on the specified conditions
filtered_df = df_new[(df_new['portfolio_risk'] < 12) &
(df_new['returns_5yr']>12) ]

# Sort the filtered DataFrame based on 'portfolio_risk' and
'portfolio_returns'
sorted_df = filtered_df.sort_values(by=['portfolio_risk', 'returns_5yr'],
ascending=[True, False])

# Extract the top 5 schemes
top_5_schemes = sorted_df.head(10)

# Extract scheme names of the top 5 schemes
top_5_scheme_names = top_5_schemes['scheme_name'].tolist()

# Plot the scatter plot
plt.scatter(filtered_df['portfolio_risk'], filtered_df['returns_5yr'],
c=filtered_df['sharpe_ratio'], cmap='viridis')
```

```
# Add color bar
plt.colorbar(label='sharpe_ratio')

# Label the points with scheme names for the top 5 schemes
for i, txt in enumerate(filtered_df['scheme_name']):
    if txt in top_5_scheme_names:
        plt.text(filtered_df['portfolio_risk'].iloc[i],
filtered_df['returns_5yr'].iloc[i], txt, fontsize=10)

plt.xlabel('Portfolio Risk')
plt.ylabel('Portfolio Returns')
plt.title('Risk-Return Trade-off (Portfolio Risk < 12 & returns_5yr >
12)')

plt.show()
top_5_scheme_names = top_5_scheme_names[:10]  # Ensure only top 5 scheme
names are considered
print(top_5_scheme_names)
```

Output

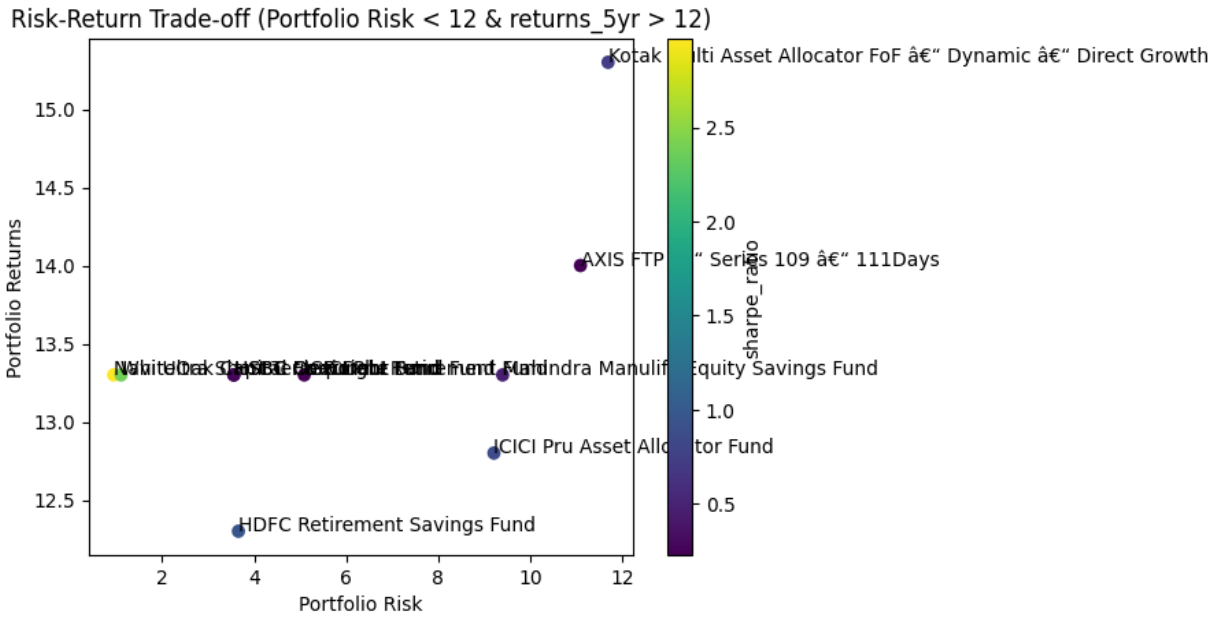

Fig 2.4.10

```
'Navi Ultra Short Term Fund', 'WhiteOak Capital Overnight Fund', 'HSBC
Corporate Bond Fund', 'HSBC Flexi Debt Fund', 'HDFC Retirement Savings
Fund', 'ICICI Pru Retirement Fund', 'ICICI Pru Asset Allocator Fund',
'Mahindra Manulife Equity Savings Fund', 'AXIS FTP â€“Series 109 â€“
111Days', 'Kotak Multi Asset Allocator FoF â€“ Dynamic â€“ Direct
Growth'
```

Interpretation

- schemes of 5th year return low risk with high return. So we can invest in those all following fund for 5th year return

  We have extract top 10 schemes:

  - 1.Navi Ultra Short-Term Fund
  - 2.WhiteOak Capital Overnight Fund
  - 3.HSBC Corporate Bond Fund
  - 4.HSBC Flexi Debt Fund
  - 5.HDFC Retirement Savings Fund
  - 6.ICICI Pru Retirement Fund
  - 7.ICICI Pru Asset Allocator Fund
  - 8.Mahindra Manulife Equity Savings Fund
  - 9.AXIS FTP â€“ Series 109 â€“ 111Days
  - 10.Kotak Multi Asset Allocator FoF â€“ Dynamic â€“ Direct Growth

Let's find highest & lowest return of amc (assets managing company) fund details for

1st, 3rd & 5th year

Let's find highest return of amc fund details for 1st year

```
#Let's Find Highest Return Fund details of 1Year
max_return_1yr = df_new.loc[df_new['returns_1yr'].idxmax()]
max_return_1yr_prcnt =
max_return_1yr[['returns_1yr','amc_name','category']]
max_return_1yr_prcnt
```

Output

```
returns_1yr                 130.8
amc_name      Bank of India Mutual Fund
category                    Debt
Name: 89, dtype: object
```

Let's find lowest return of amc fund details for 1st year

```
#Let's Find Lowest Return Fund Details of 1Year
min_return_1yr = df_new.loc[df_new['returns_1yr'].idxmin()]
min_return_1yr_prcnt =
min_return_1yr[['returns_1yr','amc_name','category']]
min_return_1yr_prcnt
```

Output

```
returns_1yr                -19.7
amc_name      ICICI Prudential Mutual Fund
category                    Equity
Name: 347, dtype: object
```

Interpretation

	amc_name	returns_1yr	category
highest	Bank of India Mutual Fund	130.8	Debt
lowest	ICICI Prudential Mutual Fund	-19.7	Equity

Let's find highest return of amc fund details for 3rd year

```
#Let's Find highest Return Fund Details of 3year
max_return_3yr = df_new.loc[df_new['returns_3yr'].idxmax()]
max_return_3yr_prcnt =
max_return_3yr[['returns_3yr','amc_name','category']]
max_return_3yr_prcnt
```

Output

```
returns_3yr              71.4
amc_name       Quant Mutual Fund
category              Equity
Name: 608, dtype: object
```

Let's find lowest return of amc fund details for 3rd year

```
#Let's Find Lowest Return Fund Details of 3Year
min_return_3yr = df_new.loc[df_new['returns_3yr'].idxmin()]
min_return_3yr_prcnt =
min_return_3yr[['returns_3yr','amc_name','category']]
min_return_3yr_prcnt
```

Output

```
returns_3yr                3.3
amc_name       Indiabulls Mutual Fund
category                 Hybrid
Name: 373, dtype: object
```

Interpretation:

	amc_name	returns_3yr	category
highest	Quant Mutual Fund	71.4	Equity
lowest	Indiabulls Mutual Fund	3.3	Hybrid

Let's find highest return of amc fund details for 5th year

```
#Let's Find highest Return Fund Details of 5year
max_return_5yr = df_new.loc[df_new['returns_5yr'].idxmax()]
max_return_5yr_prcnt =
max_return_5yr[['returns_5yr','amc_name','category']]
max_return_5yr_prcnt
```

Output

```
returns_5yr              23.2
amc_name       Quant Mutual Fund
category              Equity
Name: 608, dtype: object
```

Let's find lowest return of amc fund details for 5th year

```
#Let's Find Lowest Return Fund Details of 5Year
min_return_5yr = df_new.loc[df_new['returns_5yr'].idxmin()]
min_return_5yr_prcnt =
min_return_5yr[['returns_5yr','scheme_name','category']]
min_return_5yr_prcnt
```

Output

returns_5yr -4.1
scheme_name Bank of India Credit Risk Fund
category Debt
Name: 89, dtype: object

Interpretation

	amc_name	returns_5yr	category
highest	Quant Mutual Fund	23.2	Equity
lowest	Bank of India Credit Risk Fund	-4.1	Debt

2.5 Multivariate Technique - Cluster Analysis

We apply the **k – means clustering** for the low risk with high returns grouping. Find the best returns fund. Then we find low risk with best returns fund for 1st, 3rd & 5th year returns.

Returns 1st year

```
## Return 1 Year

data=read.csv(file.choose(),header = TRUE)

View(data)

attach(data)

data1=data[-c(1,2,3,4)]

View(data1)

data=scale(data1)

d=dist(data1)

hc=hclust(d)

plot(hc)
```

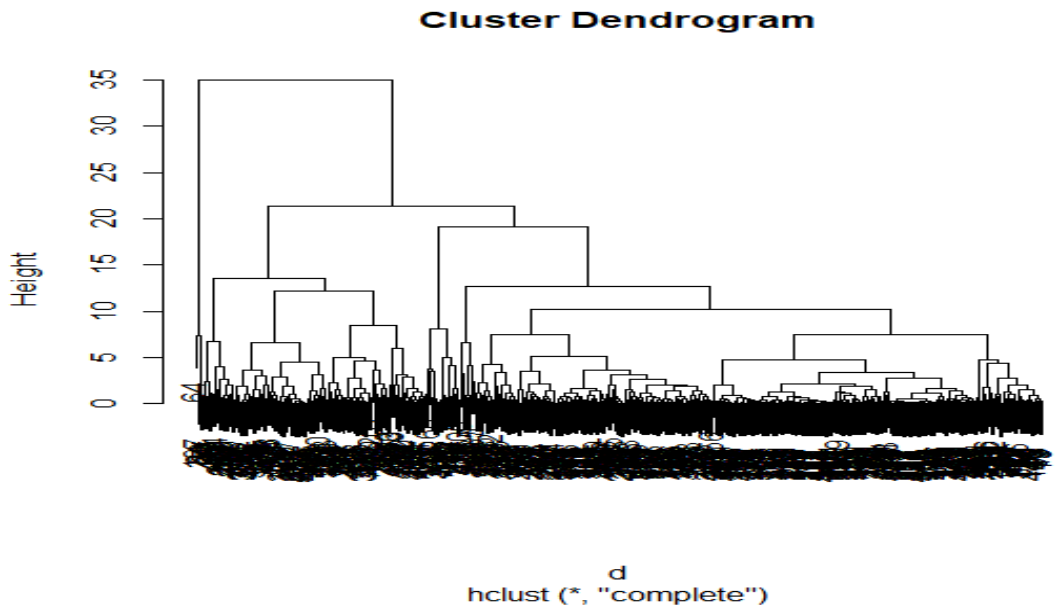

Fig 2.5.1

```
library(NbClust)
nc=NbClust(data1,min.nc=2,max.nc=6,method="Kmeans")
nc
barplot(table(nc$Best.nc[1,]))
```

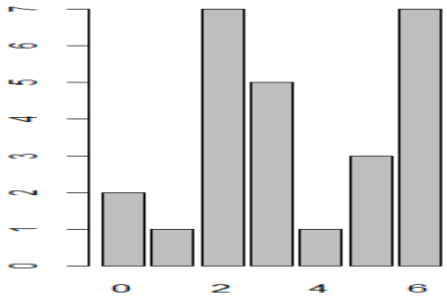

Fig 2.5.2

We have to select 6 cluster in Return 1st year.

Final Cluster Centers

	Cluster					
	1	2	3	4	5	6
returns_1yr	-.85	3.94	5.45	4.11	23.45	11.29
Portfolio_risk	12.68	4.38	1.67	11.14	11.71	12.57
Sharpe_ratio	.03	.79	2.61	.32	.60	.44

Interpretation

We have consider only high return value in the above cluster table, because in our original data we extract only those schemes of portfolio risk less than 15 its shows that low risk in your mutual fund. This cluster table say that,

Cluster 1 having **Low /Negative** Return
Cluster 2 having **Low to Moderate** Return
Cluster 3 having **Low to Moderately High** Return
Cluster 4 having **Low to Moderate** Return
Cluster 5 having **Very High** Return
Cluster 6 having **High** Return

ANOVA

	Cluster		Error			
	Mean Square	df	Mean Square	df	F	Sig.
returns_1yr	727.053	5	1.812	453	401.180	.000
Portfolio_risk	1669.230	5	1.951	453	855.406	.000
Sharpe_ratio	99.138	5	.931	453	106.467	.000

Interpretation

In ANOVA table P_value for returns_1yr, Portfolio_risk and sharpe_ratio is 0.000 which is less than 0.05 so we reject Ho also we can say this variable is important or significant.

Number of Cases in each Cluster

Cluster	1	36.000
	2	139.000
	3	182.000
	4	77.000
	5	4.000
	6	21.000
Valid		459.000
Missing		.000

Retrieved Top Schemes in 5th cluster

Table 1: Cluster 5 (Very High Returns)

Sr.No	scheme_name	fund_size_cr	category	sub_category	returns_1yr	Portfolio_risk
1	Aditya Birla SL Medium Term Plan â€“ Direct Growth	1652	Debt	Medium Duration Funds	22.1	11.85
2	Bank of India Conservative Hybrid Fund	63	Hybrid	Conservative Hybrid Mutual Funds	22.3	12.49
3	Bank of India Short Term Income â€“ Direct Growth	72	Debt	Short Duration Funds	28.3	10.45
4	ICICI Pru FMCG Fund	1156	Equity	Sectoral / Thematic Mutual Funds	21.1	12.03

Returns 3rd year

How many clusters selected for 3rd year return. we can plot dendrogram and select appropriate cluster.

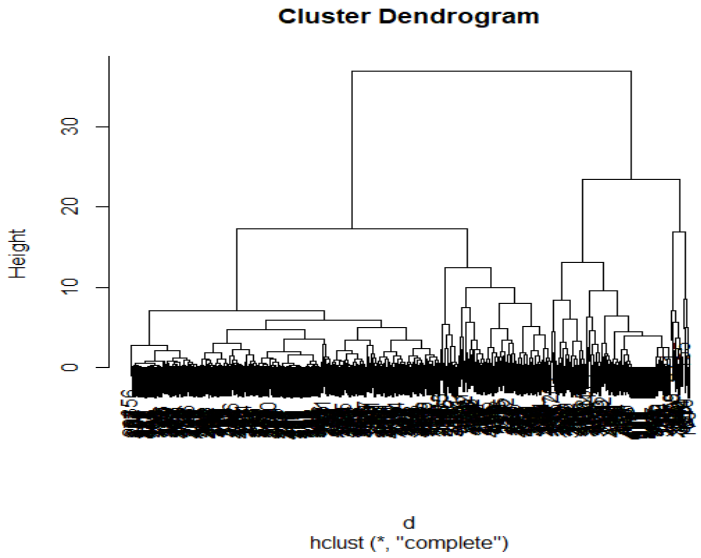

Fig 2.5.3

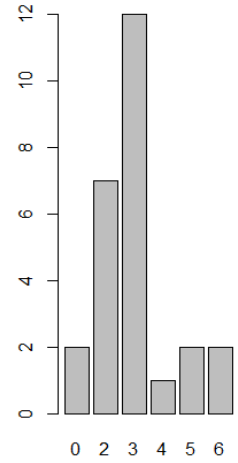

Fig 2.5.4

We have to select 6 cluster in 3rd year returns

Final Cluster Centers

	Cluster		
	1	2	3
returns_3yr	15.70	5.77	26.72
Portfolio_risk	10.36	2.65	11.61
Sharpe_ratio	.27	1.93	.50

Table 2

Interpretation

We have considered only high return value in the above cluster table, because in our original data we extract only those schemes of portfolio risk less than 15 its shows that low risk in your mutual fund. This cluster table say that,

Cluster 1 having **Moderate** Returns
Cluster 2 having **Low** Returns
Cluster 3 having **High** Returns

ANOVA

	Cluster		Error			
	Mean Square	df	Mean Square	df	F	Sig.
returns_3yr	10947.478	2	6.209	456	1763.040	.000
Portfolio_risk	3445.969	2	5.128	456	672.042	.000
Sharpe_ratio	134.920	2	1.420	456	94.992	.000

Interpretation

In ANOVA table P-value for returns_3yr, Portfolio_risk and sharpe_ratio is 0.000 which is less than 0.05 so we reject Ho also we can say that this variable is important or significant.

Number of Cases in each Cluster

Cluster	1	118.000
	2	295.000
	3	46.000
Valid		459.000
Missing		.000

Retrieved Top 12 Schemes in 3rd cluster

Table 2: Cluster 3 (High Returns)

Sr.No	scheme_name	fund_size_cr	category	sub_category	returns_3yr	Portfolio_risk
1	Templeton India Equity Income Fund	1350	Equity	Dividend Yield Funds	37.7	14.11
2	L&T Focused Equity Fund	946	Equity	Focused Funds	33.7	11.09
3	Tata Retirement Savings Fund	1259	Solution Oriented	Retirement Funds	33.7	12.74
4	Parag Parikh Tax Saver Fund	1147	Equity	ELSS Mutual Funds	33	13.5
5	HSBC Corporate Bond Fund	157	Debt	Corporate Bond Mutual Funds	32.4	3.56
6	HSBC Equity Hybrid Fund	474	Hybrid	Aggressive Hybrid Mutual Funds	32.4	13.08
7	HSBC Flexi Debt Fund	54	Debt	Dynamic Bond	32.4	3.56
8	ICICI Pru Retirement Fund	152	Solution Oriented	Retirement Funds	32.4	5.09

9	Mahindra Manulife Equity Savings Fund	418	Hybrid	Equity Savings Mutual Funds	32.4	9.4
10	Motilal Oswal Equity Hybrid Fund	385	Hybrid	Aggressive Hybrid Mutual Funds	32.4	12.76
11	Navi Ultra Short Term Fund	11	Debt	Ultra Short Duration Funds	32.4	0.95
12	WhiteOak Capital Overnight Fund	14	Debt	Overnight Mutual Funds	32.4	1.11

Returns 5th year

How many clusters selected for 5th year return. We can plot dendrogram and select appropriate cluster.

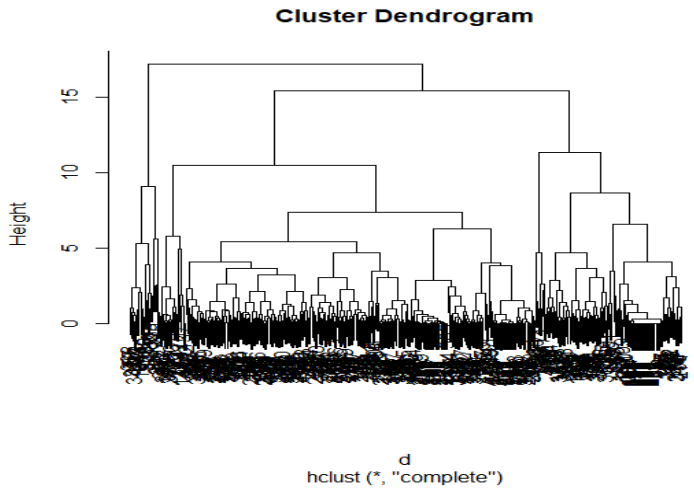

Fig 2.5.5

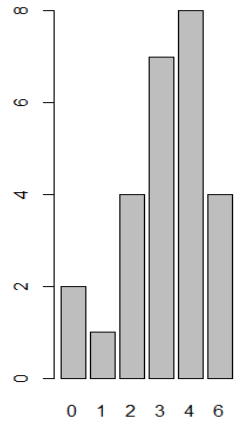

Fig 2.5.6

We have to select 4 cluster in 5th year returns

Table 3

Final Cluster Centers

	Cluster			
	1	2	3	4
returns_5yr	5.86	9.75	7.09	9.41
Portfolio_risk	2.02	1.74	4.82	10.24
Sharpe_ratio	2.61	1.95	.61	.36

Interpretation

We have considered only high return value in the above cluster table, because in our original data we extract only those schemes of portfolio risk less than 12 its shows that low risk in your mutual fund. This cluster table say that,

- Cluster 1 having Low Return
- Cluster 2 having High Return
- Cluster 3 having Moderate Returns
- Cluster 4 having Moderate High Returns

ANOVA

	Cluster		Error			
	Mean Square	df	Mean Square	df	F	Sig.
returns_5yr	330.588	3	2.067	386	159.910	.000
Portfolio_risk	1172.544	3	1.372	386	854.689	.000
Sharpe_ratio	122.181	3	1.118	386	109.295	.000

Interpretation

In ANOVA table P-value for 5th year returns, Portfolio_risk and sharpe_ratio is 0.000 which is less than 0.05 so we reject Ho also we can say that this variable is important or significant.

Number of Cases in each Cluster

Cluster	1	150.000
	2	66.000
	3	109.000
	4	65.000
Valid		390.000
Missing		.000

Retrieved Top Schemes in 2nd cluster

Table 3: Cluster 2 (High Returns)

Sr.No	scheme_name	fund_size_cr	category	sub_category	returns_5yr	Portfolio_risk
1	HSBC Corporate Bond Fund	157	Debt	Corporate Bond Mutual Funds	13.3	3.56
2	HSBC Flexi Debt Fund	54	Debt	Dynamic Bond	13.3	3.56
3	ICICI Pru Retirement Fund	152	Solution Oriented	Retirement Funds	13.3	5.09
4	Navi Ultra Short Term Fund	11	Debt	Ultra Short Duration Funds	13.3	0.95
5	WhiteOak Capital Overnight Fund	14	Debt	Overnight Mutual Funds	13.3	1.11
6	HDFC Retirement Savings Fund	932	Solution Oriented	Retirement Funds	12.3	3.66

3. RESULTS AND DISCUSSION

3.1 Trade-off

From fig 2.4.4, 2.4.5, 2.4.6 calculate mean of each year return for each category & check the category gives the best return for 1st, 3rd & 5th year returns.

1st year returns:

The highest return in the "returns_1yr" category is observed in the "Debt" category, with a return value of 5.547872.

3rd year returns:

"Equity" has the highest return among the listed categories, with a return of approximately 29.70%.

5th year returns:

"Equity" has the highest return among the listed categories, with a return of approximately 11.65%.

When Sharpe ratio is low which is less than 0.5 then risk return tread off show high risk and moderate-high return. ***i.e. High risk & low return*****. Which can be dangerous for investors.**

- Bank of India Credit Risk Fund
- DSP World Gold Fund
- DSP World Mining Fund
- PGIM India GEO Fund
- Edelweiss Greater China Equity Off-Shore Fund
- Aditya Birla SL Banking&Financial Services-DirGrowth
- Motilal Oswal Nifty Bank Index Fund
- Nippon India Banking&Financial Services-DirGrowth
- UTI-Banking and Financial Services Fund
- HDFC Infrastructure Fund

When Sortino ratio is low which is less than 0.5 then risk return tread off show high risk and moderate-high return. *i.e. High risk & low return*. Which can be dangerous for investors.

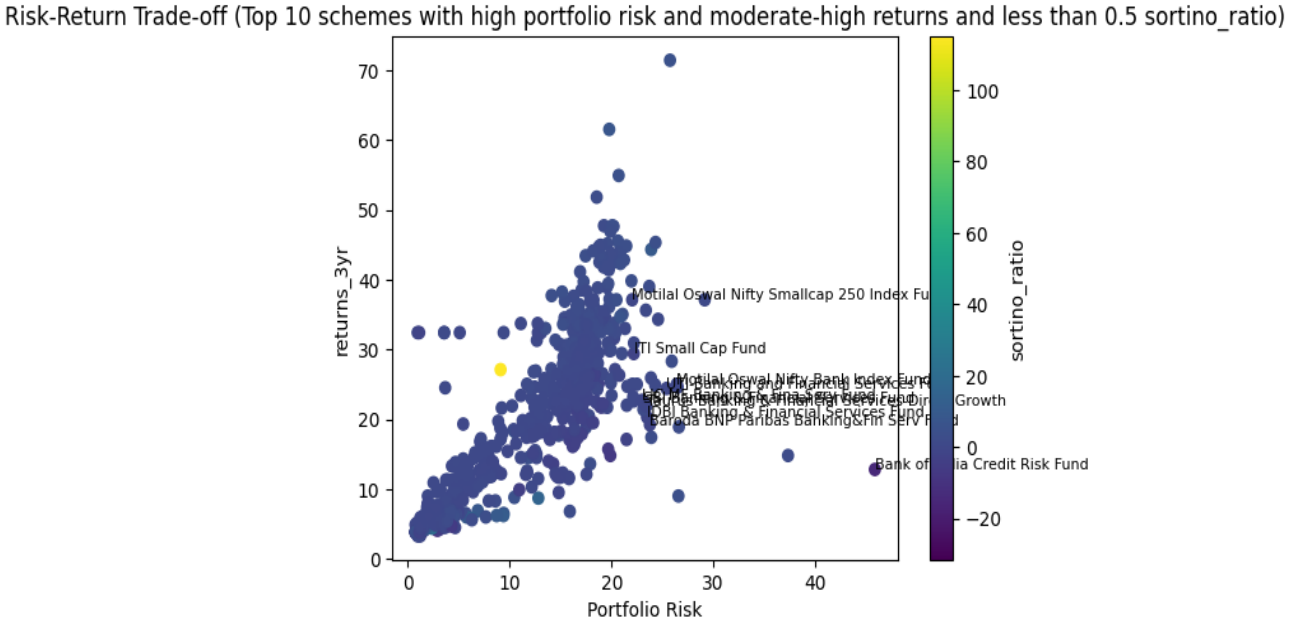

Fig 3.1.1

- Bank of India Credit Risk Fund
- Motilal Oswal Nifty Bank Index Fund
- UTI-Banking and Financial Services Fund
- Baroda BNP Paribas Banking&Fin Serv Fund
- IDBI Banking & Financial Services Fund
- Taurus Banking & Financial Services-Direct Growth
- SBI Banking & Financial Services Fund
- LIC MF Banking & Fina Serv Fund
- ITI Small Cap Fund
- Motilal Oswal Nifty Smallcap 250 Index Fund

From the above mutual funds shows that the high risk. When Sharpe ratio & Sortino ratio is low which is less than 0.5 then risk return tread off show high risk and moderate-high return. ***i.e. High risk & low return*. Which can be dangerous for investors. So we can suggest to mutual fund investors that avoid those all schemes for investment.**

The schemes of 1st year return low risk with high return as compare to sharpe_ratio.

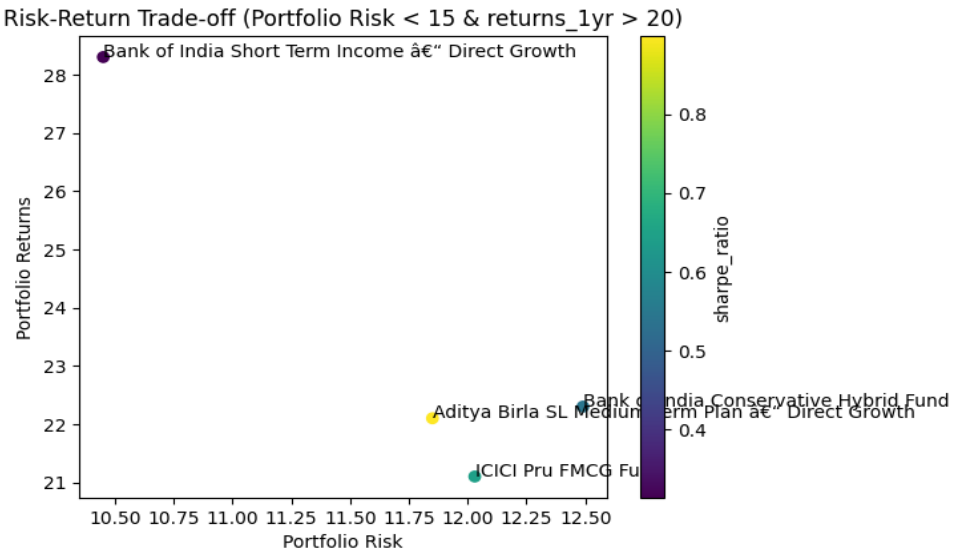

Fig 3.1.2

- Schemes of 1st year return low risk with high return. We have extract top schemes:
    - Bank of India Short Term Income â€“Direct Growth
    - Aditya Birla SL Medium Term Plan â€“Direct Growth
    - ICICI Pru FMCG Fund
    - Bank of India Conservative Hybrid Fund

The schemes of 3rd year return low risk with high return as compare to sharpe_ratio.

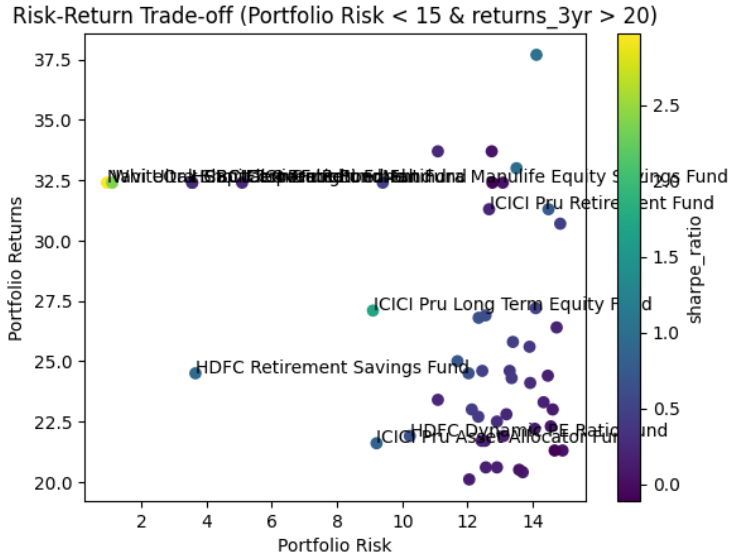

Fig 3.1.3

- Schemes of 3rd year return low risk with high return. We have extract top 10 schemes

    - 1.Navi Ultra Short Term Fund
    - 2.WhiteOak Capital Overnight Fund
    - 3.HSBC Corporate Bond Fund
    - 4.HSBC Flexi Debt Fund
    - 5.HDFC Retirement Savings Fund
    - 6.ICICI Pru Retirement Fund
    - 7.ICICI Pru Long Term Equity Fund
    - 8.ICICI Pru Asset Allocator Fund
    - 9.Mahindra Manulife Equity Savings Fund
    - 10.HDFC Dynamic PE Ratio Fund

From the above mutual funds schemes shows that the low risk. ***i.e. low risk & moderate high returns*****. Which can be beneficial for investors. So, we can suggest to mutual fund investors that invest in those all mutual fund schemes for 3rd year time period.**

The schemes of 5th year return low risk with high return as compare to sharpe_ratio.

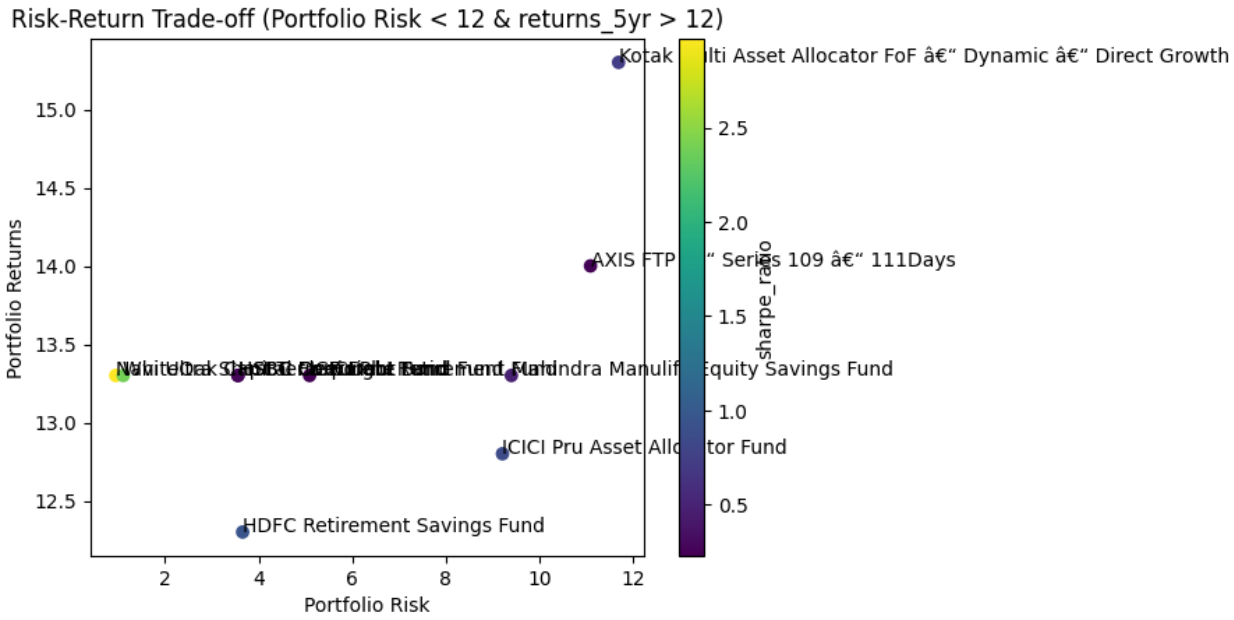

Fig 3.1.4

- Schemes of 5th year return low risk with high return. We have extract top 10 schemes

    - 1.Mahindra Manulife Equity Savings Fund
    - 2.AXIS FTP â€“Series 109 â€“ 111Days
    - 3.Kotak Multi Asset Allocator FoF â€“ Dynamic â€“ Direct Growth
    - 4.HSBC Flexi Debt Fund
    - 5.HDFC Retirement Savings Fund
    - 6.ICICI Pru Retirement Fund
    - 7.ICICI Pru Asset Allocator Fund
    - 8. WhiteOak Capital Overnight Fund
    - 9.Navi Ultra Short-Term Fund
    - 10.HSBC Corporate Bond Fund

From the above mutual funds schemes shows that the low risk. ***i.e. low risk & moderate returns*****. Which can be beneficial for long term investment. So, we suggest to mutual fund investors that invest in those all mutual fund schemes for 5th year long time period.**

3.2 Cluster Analysis

Our aim is that to identify the best portfolio of securities also to help the investors to decide the effective portfolio of securities. This purpose we use the k-means clustering and identify best schemes in mutual fund. Mutual fund dataset there are 39 banks were selected for analysis such as (Aditya Birla Sun Life, Axis, Bandhan, Bank of India, Baroda BNP Paribas, Edelweiss, Canara Robeco, DSP, Franklin Templeton, HDFC, HSBC, ICICI Prudential, IDBI, IIFL, Indiabulls, Invesco, ITI, JM Financial, Kotak Mahindra, L&T, LIC, Mahindra Manulife, Mirae Asset, Motilal Oswal, Navi, Nippon India, PPFAS, PGIM India, Quant, Quantum, SBI, Shriram, Sundaram, Tata, Taurus, Trust, Union, UTI, Whiteoak Capital). Clustering model fit in each particular returns.

1st year returns:

1st year returns cluster model divides schemes into six cluster.

Final Cluster Centers

	Cluster					
	1	2	3	4	5	6
returns_1yr	-.85	3.94	5.45	4.11	23.45	11.29
Portfolio_risk	12.68	4.38	1.67	11.14	11.71	12.57
Sharpe_ratio	.03	.79	2.61	.32	.60	.44

Final cluster centers table shows that,

Cluster 1 having **Low /Negative** Return
Cluster 2 having **Low to Moderate** Return
Cluster 3 having **Low to Moderately High** Return
Cluster 4 having **Low to Moderate** Return
Cluster 5 having **Very High** Return
Cluster 6 having **High** Return

The 1st cluster consists of **36** schemes: **Aditya Birla SL Equity Hybrid â€™95 Fund, Aditya Birla SL Inter Equity â€“Plan A-Direct Growth, AXIS Equity Hybrid Fund and so on**.

The 2nd cluster consists of **139** schemes: **Aditya Birla SL Banking & PSU Debt Fund, AXIS All Seasons Debt Fund, Baroda BNP Paribas Conservative Hybrid Fund and so on.**

The 3rd cluster consists of **182** schemes: **Aditya Birla SL Arbitrage Fund, AXIS Arbitrage Fund, Bank of India Liquid Fund and so on**.

The 4th cluster consists of **77** schemes: **Aditya Birla SL Asset Allocator FoF-Dir Growth, AXIS Equity Saver Fund, Baroda BNP Paribas Aggressive Hybrid Fund and so on.**

The 5th cluster consists of 4 schemes: **Aditya Birla SL Medium Term Plan â€“ Direct Growth, Bank of India Conservative Hybrid Fund, Bank of India Short Term Income â€“ Direct Growth, ICICI Pru FMCG Fund.**

The 6th cluster consists of 21 schemes: **Aditya Birla SL Gold Fund, HDFC Dynamic PE Ratio Fund, ICICI Pru Bharat Consumption Fund** and so on.

So as per all over cluster considering we can say that 1st year returns clustering output that 5th cluster shows important mutual fund schemes because Return is 23.45 which is high compared to other cluster returns, and their corresponding portfolio risk is 11.71, Sharpe ratio is 0.60. Number of cases in 5th cluster is 4 we also retrieved mutual fund schemes names over here

1.	**Aditya Birla SL Medium Term Plan â€“ Direct Growth**
2.	**Bank of India Conservative Hybrid Fund**
3.	**Bank of India Short Term Income â€“ Direct Growth**
4.	**ICICI Pru FMCG Fund**

3rd year returns:

3rd year returns cluster model divides schemes into three cluster.

Final Cluster Centers

	Cluster		
	1	2	3
returns_3yr	15.70	5.77	26.72
Portfolio_risk	10.36	2.65	11.61
Sharpe_ratio	.27	1.93	.50

Final cluster centers table shows that,

Cluster 1 having **Moderate** Return

Cluster 2 having **Low** Return

Cluster 3 having **High** Return

The 1st cluster consists of **118** schemes: **Aditya Birla SL Balanced Advantage Fund, AXIS Balanced Advantage Fund, AXIS FTP â€“ Series 109 â€“ 111Days and so on.**

The 2nd cluster consists of **295** schemes: **Aditya Birla SL Active Debt Multi-Mgr FoF-Dir Growth, AXIS All Seasons Debt Fund, Baroda BNP Paribas Credit Risk Fund and so on.**

The 3rd cluster consists of **46** schemes: **Franklin India Equity Hybrid Fund, HSBC Equity Hybrid Fund, ICICI Pru Exports and Services Fund and so on.**

So as per all over cluster considering we can say that 3rd year returns clustering output that 3rd cluster shows important mutual fund schemes because Return is 26.72 which is high compared to other cluster returns, and their corresponding portfolio risk is 11.61, Sharpe ratio is 0.50. Number of cases in 3rd cluster is 46 we retrieved only top 12 mutual fund schemes names over here

1. Templeton India Equity Income Fund
2. L&T Focused Equity Fund
3. Tata Retirement Savings Fund
4. Parag Parikh Tax Saver Fund
5. HSBC Corporate Bond Fund
6. HSBC Equity Hybrid Fund
7. HSBC Flexi Debt Fund
8. ICICI Pru Retirement Fund
9. Mahindra Manulife Equity Savings Fund
10. Motilal Oswal Equity Hybrid Fund
11. Navi Ultra Short-Term Fund
12. WhiteOak Capital Overnight Fund

5th year returns:

5th year returns cluster model divides schemes into four cluster.

Final Cluster Centers

	Cluster			
	1	2	3	4
returns_5yr	5.86	9.75	7.09	9.41
Portfolio_risk	2.02	1.74	4.82	10.24
Sharpe_ratio	2.61	1.95	.61	.36

Final cluster centers table shows that,

Cluster 1 having **Low** Return
Cluster 2 having **High** Return
Cluster 3 having **Moderate** Return
Cluster 4 having **Moderate High** Return

The 1st cluster consists of **150** schemes: **Aditya Birla SL Arbitrage Fund, AXIS Arbitrage Fund, Bank of India Liquid Fund and so on.**

The 2nd cluster consists of **66** schemes: **HSBC Corporate Bond Fund, Navi Ultra Short-Term Fund, WhiteOak Capital Overnight Fund and so on.**

The 3rd cluster consists of 109 schemes: **Aditya Birla SL Dynamic Bond Fund, AXIS Retirement Savings Fund, Baroda BNP Paribas Credit Risk Fund and so on.**

The 4th cluster consists of **65** schemes: **Aditya Birla SL Active Debt Multi-Mgr FoF-Dir Growth, AXIS FTP â€“Series 109 â€“ 111Days, Baroda BNP Paribas Balanced Advantage Fund and so on.**

So as per all over cluster considering we can say that 5th year returns clustering output that 2nd cluster shows important mutual fund schemes because Return is 9.75 which is high compared to other cluster returns, and their corresponding portfolio risk is 1.74 and Sharpe ratio is 1.95. Number of cases in 2nd cluster is 66 we retrieved only top 6 mutual fund schemes names over here

1. **HSBC Corporate Bond Fund**
2. **HSBC Flexi Debt Fund**
3. **ICICI Pru Retirement Fund**
4. **Navi Ultra Short-Term Fund**
5. **WhiteOak Capital Overnight Fund**
6. **HDFC Retirement Savings Fund**

5. CONCLUSION

In this study, we successfully achieved

Explore factors influencing the performance of mutual funds. Using multiple EDA plot, we check our data in graphical way, we conclude that

Equity category is showing the highest number of schemes (308) of total 814 Schemes.

Asset Management Company (Amc) **ICICI Prudential Mutual Fund** has highest number of Schemes 57 out of 814.

Equity category funds gives the highest returns in all years, 1st year, returns 3rd year, returns 5th year.

In each year returns we have identified schemes which gives low risk with high return. This purpose using Fundamental Principal process the risk-return trade-off for different mutual fund portfolios.

3rd and 5th year high returns mutual fund gives low risk with high return.

scheme_name	min_sip/month	min_lumpsum
Navi Ultra Short-Term Fund	500	5000
HSBC Corporate Bond Fund	500	5000
HSBC Flexi Debt Fund	500	5000
WhiteOak Capital Overnight Fund	500	5000
HDFC Retirement Savings Fund	100	5000
ICICI Pru Retirement Fund	100	5000
ICICI Pru Asset Allocator Fund	1000	5000
Mahindra Manulife Equity Savings Fund	500	5000

To identify the best portfolio of securities. This purpose we use the k-means clustering and identify best schemes in mutual fund.

1st and 3rd year high return mutual fund gives the low risk with high returns**.**

scheme_name	min_sip/month	min_lumpsum
ICICI Pru FMCG Fund	100	5000

3rd and 5th year return gives the low risk with high returns**.**

scheme_name	min_sip/month	min_lumpsum
HDFC Retirement Savings Fund	100	5000
HSBC Corporate Bond Fund	500	5000
HSBC Flexi Debt Fund	500	5000
ICICI Pru Retirement Fund	100	5000
Navi Ultra Short Term Fund	500	5000
WhiteOak Capital Overnight Fund	500	5000

To help the investors to decide the effective portfolio of securities.

If investor wants invest money the period of 1st year & 3rd year then we suggest that **ICICI Pru FMCG Fund** because that it takes low risk with give the best 1st year returns 21.1% and 3rd year returns 24.5%. Also it is Sharpe ratio & Sortino ratio greater than 1. It management fees 1.44%. There fund size 1156 Cr. So here all factor show the satisfaction about best effective portfolio of securities. So, we suggest that **ICICI Pru FMCG Fund** is the best portfolio of securities.

If investor wants invest money the period of 3rd year & 5th year then we suggest that

Scheme Name	3rd year Return	5th year Return	management fees	Fund size in Cr
HDFC Retirement Savings Fund	34.9	14.8	0.78	2696
HSBC Corporate Bond Fund	32.4	13.3	0.35	157
HSBC Flexi Debt Fund	32.4	13.3	0.94	54
ICICI Pru Retirement Fund	32.4	13.3	1.12	152
Navi Ultra Short Term Fund	32.4	13.3	0.73	11
WhiteOak Capital Overnight Fund	32.4	13.3	0.22	14

Because that those takes low risk with give the best 3rd year returns and moderate 5th year returns. Those management fees are also very minimum. So here all factor shows the satisfaction about best effective portfolio of securities. So, we suggest that those all above mutual fund schemes are the best portfolio of securities.

5. REFERENCES

- **Disha Thakkar** (ICICI Mutual Fund, Anand Gujarat)
- **Aman Sir** (HDFC Mutual Fund, Pune Maharashtra)
- **Utkarsh Mishra** (Senior Manager, Bank of Baroda, Mumbai)

Book: Bhuyan, K.C. (2004). Multivariate analysis and its application. New Central Book Agency (P) LTD. (171-193)

Web Link:

- https://youtu.be/X1rbN1VRaxc?si=-arGWyegjWPMMcdl
- https://www.investor.gov/introduction-investing/investing-basics/investment-products/mutual-funds-and-exchange-traded-1#:~:text=What%20are%20mutual%20funds%3F,buy%20shares%20in%20mutual%20funds.
- https://www.investopedia.com/terms/e/equityfund.asp
- https://www.amfiindia.com/investor-corner/knowledge-center/debt-fund.html#accordion4
- https://www.investopedia.com/terms/h/hybridfund.asp
- https://www.angelone.in/knowledge-center/mutual-funds/solution-oriented-#what-are-solution-oriented-mutual-funds-1
- https://www.indeed.com/career-advice/career-development/how-to-conduct-exploratory-data-analysis

6. APPENDIX

Schemes	min_sip	min_lumpsum	expense_ratio	fund_size_cr	fund_age_yr	fund_manager	sortino	alpha	sd	beta	sharpe	risk_level	amc_name	rating	category	sub_category	returns_1yr	returns_3yr	returns_5yr
Aditya Birla SL Active Debt Multi-Mgr FoF-Dir Growth	100	5181.818182	0.27	10	10	Kaustubh Gupta	0.32	2.24	9.39	0.01	0.24	3	Aditya Birla Sun Life Mutual Fund	3	Other	FoFs Domestic	4	6.5	6.9
Aditya Birla SL Arbitrage Fund	1000	5181.818182	0.36	4288	10	Lovelish Solanki	1.33	1.53	0.72	0.56	1.1	1	Aditya Birla Sun Life Mutual Fund	3	Hybrid	Arbitrage Mutual Funds	5.6	4.8	5.5
Aditya Birla SL Asset Allocator FoF-Dir Growth	1000	5181.818182	0.53	157	10	Vinod Narayan Bhat	3.44	2.67	10.58	0.67	1.42	5	Aditya Birla Sun Life Mutual Fund	3	Other	FoFs Domestic	2	18.9	9.7
Aditya Birla SL Bal Bhavishya Yojna â€“ Dir Growth	500	5181.818182	0.76	637	4	Atul Penkar	2.18	-6.37	14.99	0.85	0.9	6	Aditya Birla Sun Life Mutual Fund	2	Solution Oriented	Childrens Funds	-0.7	17.1	9.49
Aditya Birla SL Balanced Advantage Fund	100	5181.818182	0.61	6386	10	Mohit Sharma	3.69	1.99	10.38	0.68	1.39	6	Aditya Birla Sun Life Mutual Fund	4	Hybrid	Dynamic Asset Allocation or Balanced Advantage	4.5	18.6	9.7
Aditya Birla SL Banking&Financial Services-DirGrowth	1000	5181.818182	1.17	2384	9	Dhaval Gala	2.07	1.24	25.53	0.96	0.97	6	Aditya Birla Sun Life Mutual Fund	2	Equity	Sectoral / Thematic Mutual Funds	5.3	24.6	9.2
Aditya Birla SL Banking&PSU Debt Fund	1000	5181.818182	0.37	7994	10	Kaustubh Gupta	1.92	4.46	1.91	1.78	1.09	3	Aditya Birla Sun Life Mutual Fund	4	Debt	Banking and PSU Mutual Funds	4.5	6.8	7.3
Aditya Birla SL CEF â€“ Global Agri Plan-Direct Growth	1000	5181.818182	1.29	15	10	Vinod Narayan Bhat	1.42	2.5	20.18	0.82	1.06	6	Aditya Birla Sun Life Mutual Fund	0	Equity	Sectoral / Thematic Mutual Funds	-10.3	29.4	9.3

Python Code: **Click Here For Python Code**

Click Here For Data File in CSV Format