Mutual Fund Analysis

Data Loading

```
import pandas as pd
df=pd.read_excel("mutual_funds_dataset.xlsx")
df.head()
```

	scheme_name	min_sip	min_lumpsum	expense_ratio	fund_size_cr	fund_age_yr	fund_manager	sortino	alpha	sd	...	risk_level	amc_n
0	Quant Small Cap Fund	1000	5000	0.64	3301	10	Sanjeev Sharma	3.71	19.16	24.75	...	6	Qu Mu F
1	Quant Infrastructure Fund	1000	5000	0.64	822	10	Vasav Sahgal	3.44	27.24	19.24	...	6	Qu Mu F
2	Quant Tax Plan- Direct Growth	500	500	0.57	2779	10	Vasav Sahgal	3.50	17.63	19.74	...	6	Qu Mu F
3	Quant Multi Asset Fund	1000	5000	0.56	634	10	Vasav Sahgal	3.23	19.52	18.63	...	6	Qu Mu F
4	Quant Flexi Cap Fund	1000	5000	0.58	1044	10	Vasav Sahgal	3.65	15.93	19.16	...	6	Qu Mu F

5 rows × 22 columns

```
print(df.describe())
```

```
           min_sip  min_lumpsum  expense_ratio  fund_size_cr  fund_age_yr  \
count    30.000000    30.000000      30.000000     30.000000    30.000000
mean    626.666667  3923.333333       0.551333   6264.966667     8.733333
std     407.416870  1991.305815       0.246475   7518.172272     2.132399
min       0.000000     0.000000       0.080000     89.000000     4.000000
25%     150.000000  5000.000000       0.370000    971.750000     9.000000
50%     750.000000  5000.000000       0.580000   3300.500000    10.000000
75%    1000.000000  5000.000000       0.737500   8547.750000    10.000000
max    1000.000000  5000.000000       1.000000  29953.000000    10.000000

          sortino      alpha       beta     sharpe  risk_level     rating  \
count  30.000000  30.000000  30.000000  30.000000   30.000000  30.000000
mean    4.068667  11.373333   0.869333   1.945333    5.766667   3.433333
std     0.925965   5.019094   0.309013   0.199909    0.817200   2.028815
min     2.290000   5.290000   0.530000   1.370000    2.000000   0.000000
25%     3.597500   7.762500   0.750000   1.880000    6.000000   3.000000
50%     4.020000   9.940000   0.840000   1.945000    6.000000   4.000000
75%     4.405000  14.067500   0.917500   2.075000    6.000000   5.000000
max     7.270000  27.240000   2.360000   2.300000    6.000000   5.000000

       returns_1yr  returns_3yr  returns_5yr  composite_score       rank
count    30.000000    30.000000    30.000000        30.000000  30.000000
mean      7.263333    41.146667    16.315739         0.551269  15.500000
std       7.340369    10.866708     3.822083         0.039859   8.803408
min     -16.000000    14.500000     9.490726         0.510861   1.000000
25%       3.750000    35.425000    13.925000         0.524507   8.250000
50%       7.500000    42.200000    15.600000         0.541482  15.500000
75%      11.050000    45.000000    19.325000         0.569098  22.750000
max      23.200000    71.400000    23.200000         0.667338  30.000000
```

```
df['category'].unique()
```

```
array(['Equity', 'Hybrid', 'Other', 'Debt', 'Solution Oriented'],
      dtype=object)
```

```
df.isnull().sum()
```

	0
scheme_name	0
min_sip	0
min_lumpsum	0
expense_ratio	0
fund_size_cr	0
fund_age_yr	0
fund_manager	0
sortino	0
alpha	0
sd	0
beta	0
sharpe	0
risk_level	0
amc_name	0
rating	0
category	0
sub_category	0
returns_1yr	0
returns_3yr	0
returns_5yr	0
composite_score	0
rank	0

dtype: int64

dtype: int64

Data Cleaning

```
mean_returns_3yr=df["returns_3yr"].mean
mean_returns_5yr=df["returns_5yr"].mean
df["returns_3yr"].fillna(mean_returns_3yr,inplace=True)
df["returns_5yr"].fillna(mean_returns_5yr,inplace=True)
df.head()
```

	scheme_name	min_sip	min_lumpsum	expense_ratio	fund_size_cr	fund_age_yr	fund_manager	sortino	alpha	sd	...	risk_level	amc_n
0	Quant Small Cap Fund	1000	5000	0.64	3301	10	Sanjeev Sharma	3.71	19.16	24.75	...	6	Qu Mu F
1	Quant Infrastructure Fund	1000	5000	0.64	822	10	Vasav Sahgal	3.44	27.24	19.24	...	6	Qu Mu F
2	Quant Tax Plan- Direct Growth	500	500	0.57	2779	10	Vasav Sahgal	3.50	17.63	19.74	...	6	Qu Mu F
3	Quant Multi Asset Fund	1000	5000	0.56	634	10	Vasav Sahgal	3.23	19.52	18.63	...	6	Qu Mu F
4	Quant Flexi Cap Fund	1000	5000	0.58	1044	10	Vasav Sahgal	3.65	15.93	19.16	...	6	Qu Mu F

5 rows × 22 columns

Data Normalization

```
from sklearn.preprocessing import MinMaxScaler

columns_to_normalize = ['expense_ratio', 'returns_1yr', 'returns_3yr', 'returns_5yr',
                        'sharpe', 'sortino', 'alpha', 'beta']
df[columns_to_normalize] = df[columns_to_normalize].replace('-', pd.NA).apply(pd.to_numeric)

scaler = MinMaxScaler()
df_normalized = pd.DataFrame(scaler.fit_transform(df[columns_to_normalize]), columns=columns_to_normalize)

# Adjust metrics where lower is better
df_normalized['expense_ratio'] = 1 - df_normalized['expense_ratio']
df_normalized['beta'] = 1 - df_normalized['beta']
```

Composite Scoring

```
weights = {
    'expense_ratio': 0.2,
    'returns_1yr': 0.15,
    'returns_3yr': 0.15,
    'returns_5yr': 0.15,
    'sharpe': 0.1,
    'sortino': 0.1,
    'alpha': 0.1,
    'beta': 0.05
}

df_normalized['composite_score'] = sum(
    df_normalized[col] * weight for col, weight in weights.items()
)
df['composite_score'] = df_normalized['composite_score']
```

Ranking Fund

```
df["rank"]=df["composite_score"].rank(ascending=False)
df_sorted=df.sort_values(by="rank")
df_sorted.head()
```

	scheme_name	min_sip	min_lumpsum	expense_ratio	fund_size_cr	fund_age_yr	fund_manager	sortino	alpha	sd	...	risk_level	amc_n
1	Quant Infrastructure Fund	1000	5000	0.64	822	10	Vasav Sahgal	3.44	27.24	19.24	...	6	Qu Mu F
0	Quant Small Cap Fund	1000	5000	0.64	3301	10	Sanjeev Sharma	3.71	19.16	24.75	...	6	Qu Mu F
2	Quant Tax Plan- Direct Growth	500	500	0.57	2779	10	Vasav Sahgal	3.50	17.63	19.74	...	6	Qu Mu F
3	Quant Multi Asset Fund	1000	5000	0.56	634	10	Vasav Sahgal	3.23	19.52	18.63	...	6	Qu Mu F
5	Quant Absolute Fund	1000	5000	0.56	1013	10	Sanjeev Sharma	3.75	16.98	14.41	...	6	Qu Mu F

5 rows × 22 columns

Exporting Result

```
df_top_30=df_sorted.head(25)
df_top_30.to_excel("Top 25 Mutual Funds.xlsx",index=False)
print("Exported to excel file 'Top 25 Mutual Funds.xlsx'")
```

```
Exported to excel file 'Top 25 Mutual Funds.xlsx'
```