Mini Statistical Analysis Report

Dataset : “Happiness Factors Dataset”

Dataset Overview :

The dataset contains **158 observations** and **5 numerical features**:

- Economy
- Family (Fam)
- Health
- Freedom
- Happiness Score (H_Score)

Description :

	Economy	Fam	Health	Freedom	H_Score
count	158.000000	158.000000	158.000000	158.000000	158.000000
mean	0.846137	0.991046	0.630259	0.428615	5.375734
std	0.403121	0.272369	0.247078	0.150693	1.145010
min	0.000000	0.000000	0.000000	0.000000	2.839000
25%	0.545808	0.856823	0.439185	0.328330	4.526000
50%	0.910245	1.029510	0.696705	0.435515	5.232500
75%	1.158448	1.214405	0.811013	0.549092	6.243750
max	1.690420	1.402230	1.025250	0.669730	7.587000

All columns are numerical and there are **no missing values**, making the dataset clean and suitable for statistical analysis and machine learning preprocessing.

Correlation Analysis of a data

```
Correlation :
           Economy       Fam    Health   Freedom   H_Score
Economy   1.000000  0.645299  0.816478  0.370300  0.780966
Fam       0.645299  1.000000  0.531104  0.441518  0.740605
Health    0.816478  0.531104  1.000000  0.360477  0.724200
Freedom   0.370300  0.441518  0.360477  1.000000  0.568211
H_Score   0.780966  0.740605  0.724200  0.568211  1.000000
```

Strength Meaning

Correlation Value	Strength
0.7 to 1	Strong
0.3 to 0.7	Moderate
0 to 0.3	Weak

High correlation observed between:

Economy ↔ Health ≈ 0.81

This suggests:

- Economically strong regions often have better healthcare systems
- Possible multicollinearity risk in ML models
- May require feature selection or dimensionality reduction later

3