Initial Data Analysis Report

1. Dataset Info

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 196 entries, 0 to 195
Data columns (total 22 columns):
 #   Column                  Non-Null Count  Dtype
---  ------                  --------------  -----
 0   h_20                    196 non-null    int64
 1   h_24                    196 non-null    int64
 2   h_28                    196 non-null    int64
 3   Vegetables              196 non-null    int64
 4   tea Intake (estekan/day)  196 non-null  float64
 5   sumScore                196 non-null    int64
 6   tertiles                196 non-null    int64
 7   n_90                    196 non-null    int64
 8   n_91                    196 non-null    int64
 9   n_92                    196 non-null    int64
 10  n_93                    196 non-null    int64
 11  sex_Boy                 196 non-null    int64
 12  sex_Girl                196 non-null    int64
 13  sweets                  196 non-null    float64
 14  fruits                  196 non-null    float64
 15  dairy                   196 non-null    float64
 16  oils                    196 non-null    float64
 17  active_transport        196 non-null    int64
 18  physical_activity       196 non-null    float64
 19  leisure_time(hour / week)  196 non-null  float64
 20  illnesses               196 non-null    int64
 21  n_93_cluster            196 non-null    int64
dtypes: float64(7), int64(15)
memory usage: 33.8 KB
```

2. Descriptive Statistics

Feature	count	mean	std	min	25%	50%	75%	max
h_20	196.0	5.81	1.84	0.0	6.0	7.0	7.0	7.0
h_24	196.0	6.07	1.72	1.0	6.0	7.0	7.0	7.0
h_28	196.0	6.32	1.48	0.0	6.0	7.0	7.0	7.0
Vegetables	196.0	2.0	0.91	0.0	1.75	2.0	3.0	3.0
tea Intake (estekan/day)	196.0	1.53	1.07	0.0	1.5	1.5	1.5	4.0
sumScore	196.0	20.63	7.99	11.0	14.0	21.0	24.0	49.0
tertiles	196.0	1.22	0.86	0.0	0.0	1.5	2.0	2.0
n_90	196.0	1.54	0.72	0.0	1.0	2.0	2.0	2.0
n_91	196.0	0.65	0.95	0.0	0.0	0.0	1.0	4.0
n_92	196.0	2.35	0.8	0.0	2.0	3.0	3.0	3.0

n_93	196.0	8.44	2.04	0.0	8.0	9.0	10.0	10.0
sex_Boy	196.0	0.74	0.44	0.0	0.0	1.0	1.0	1.0
sex_Girl	196.0	0.26	0.44	0.0	0.0	0.0	1.0	1.0
sweets	196.0	1.38	0.43	0.38	1.12	1.38	1.62	2.62
fruits	196.0	1.97	0.66	0.0	1.5	2.0	2.5	3.0
dairy	196.0	2.42	0.63	0.0	2.0	2.67	3.0	3.0
oils	196.0	1.17	0.61	0.0	1.0	1.0	1.5	3.0
active_transport	196.0	0.46	0.5	0.0	0.0	0.0	1.0	1.0
physical_activity	196.0	1.44	0.6	0.0	1.0	1.25	1.75	2.75
leisure_time(hour / week)	196.0	9.16	2.48	0.0	7.43	9.0	10.43	16.0
illnesses	196.0	4.95	6.09	0.0	0.0	3.0	7.0	32.0
n_93_cluster	196.0	1.73	0.51	0.0	2.0	2.0	2.0	2.0

3. Missing Values per Column

h_20: 0

h_24: 0

h_28: 0

Vegetables: 0

tea Intake (estekan/day): 0

sumScore: 0

tertiles: 0

n_90: 0

n_91: 0

n_92: 0

n_93: 0

sex_Boy: 0

sex_Girl: 0

sweets: 0

fruits: 0

dairy: 0

oils: 0

active_transport: 0

physical_activity: 0

leisure_time(hour / week): 0

illnesses: 0

n_93_cluster: 0