

Initial Data Analysis Report

1. Dataset Info

<class 'pandas.core.frame.DataFrame'>

RangeIndex: 196 entries, 0 to 195

Data columns (total 23 columns):

#	Column	Non-Null Count	Dtype
0	Unnamed: 0	196 non-null	int64
1	h_20	196 non-null	int64
2	h_24	196 non-null	int64
3	h_28	196 non-null	int64
4	Vegetables	196 non-null	int64
5	tea Intake (estekan/day)	196 non-null	float64
6	sumScore	196 non-null	int64
7	tertiles	196 non-null	int64
8	n_90	196 non-null	int64
9	n_91	196 non-null	int64
10	n_92	196 non-null	int64
11	n_93	196 non-null	int64
12	sex_Boy	196 non-null	int64
13	sex_Girl	196 non-null	int64
14	sweets	196 non-null	float64
15	fruits	196 non-null	float64
16	dairy	196 non-null	float64
17	oils	196 non-null	float64
18	active_transport	196 non-null	int64
19	physical_activity	196 non-null	float64
20	leisure_time(hour / week)	196 non-null	float64
21	illnesses	196 non-null	int64
22	n_93_cluster	196 non-null	int64

dtypes: float64(7), int64(16)

memory usage: 35.3 KB

2. Descriptive Statistics

Statistic	Unnamed: 0	h_20	h_24	h_28	Vegetables
count	196.0	196.0	196.0	196.0	196.0
mean	97.5	5.81	6.07	6.32	2.0
std	56.72	1.84	1.72	1.48	0.91
min	0.0	0.0	1.0	0.0	0.0
25%	48.75	6.0	6.0	6.0	1.75
50%	97.5	7.0	7.0	7.0	2.0
75%	146.25	7.0	7.0	7.0	3.0
max	195.0	7.0	7.0	7.0	3.0

3. Missing Values per Column

Unnamed: 0: 0
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