

Initial Data Analysis Report

1. Dataset Info

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

RangeIndex: 196 entries, 0 to 195

Data columns (total 76 columns):

#	Column	Non-Null Count	Dtype
0	id2	196 non-null	int64
1	universi	196 non-null	object
2	region	196 non-null	object
3	cluster	196 non-null	int64
4	birth_ye	196 non-null	int64
5	sample_c	196 non-null	int64
6	h_18	196 non-null	int64
7	h_19	196 non-null	int64
8	h_20	196 non-null	int64
9	h_21	196 non-null	bool
10	h_22	196 non-null	int64
11	h_23	196 non-null	int64
12	h_24	196 non-null	int64
13	h_25	196 non-null	bool
14	h_26	196 non-null	int64
15	h_27	196 non-null	int64
16	h_28	196 non-null	int64
17	h_29	196 non-null	bool
18	h_30	196 non-null	int64
19	h_31	196 non-null	int64
20	h_32	196 non-null	int64
21	h_33	196 non-null	int64
22	h_34	196 non-null	int64
23	h_35_a	196 non-null	int64
24	h_36	196 non-null	int64
25	h_37	196 non-null	int64
26	h_38	196 non-null	int64
27	h_39	196 non-null	int64
28	h_40	196 non-null	int64
29	h_41	196 non-null	int64
30	h_42	196 non-null	int64
31	h_43_1	196 non-null	int64
32	h_43_2	196 non-null	float64
33	h_44_1	196 non-null	int64
34	h_44_2	196 non-null	int64
35	h_44_3	196 non-null	int64
36	h_45	196 non-null	int64
37	z_52	196 non-null	int64
38	z_53	196 non-null	bool
39	z_54	196 non-null	int64
40	z_56_1	196 non-null	int64
41	z_56_2	196 non-null	int64
42	z_56_3	196 non-null	int64

```

43 z_56_4      196 non-null  int64
44 z_56_5      196 non-null  int64
45 z_56_6      196 non-null  int64
46 z_56_7      196 non-null  int64
47 h_57        196 non-null  int64
48 h_58        196 non-null  int64
49 h_59        196 non-null  int64
50 h_60        196 non-null  int64
51 h_61        196 non-null  int64
52 h_62        196 non-null  int64
53 h_63_1      196 non-null  float64
54 h_63_2      196 non-null  int64
55 h_64_1      196 non-null  float64
56 h_64_2      196 non-null  int64
57 sumScore    196 non-null  int64
58 tertiles    196 non-null  int64
59 n_82        196 non-null  int64
60 n_83        196 non-null  int64
61 n_84        196 non-null  int64
62 n_85        196 non-null  int64
63 n_86        196 non-null  int64
64 n_87        196 non-null  int64
65 n_88        196 non-null  int64
66 n_89        196 non-null  int64
67 n_90        196 non-null  int64
68 n_91        196 non-null  int64
69 n_92        196 non-null  int64
70 n_93        196 non-null  int64
71 z_55        196 non-null  object
72 cross_Elemantry 196 non-null  int64
73 cross_intermediate 196 non-null  int64
74 sex_Boy     196 non-null  int64
75 sex_Girl    196 non-null  int64

```

```

dtypes: bool(4), float64(3), int64(66), object(3)
memory usage: 111.1+ KB

```

2. Descriptive Statistics

Feature	count	mean	std	min	25%	50%	75%	max
id2	196.0	1112062643.72	6724732.04	1101118501.0	1106118379.25	1112218006.5	1118217803.25	1123128702.0
cluster	196.0	11.9	6.71	1.0	6.0	12.0	18.0	23.0
birth_ye	196.0	82.13	2.83	76.0	80.0	82.0	84.0	87.0
sample_c	196.0	5.46	2.87	1.0	3.0	5.5	8.0	10.0
h_18	196.0	471.84	64.21	360.0	420.0	465.0	540.0	720.0
h_19	196.0	557.3	55.62	420.0	540.0	557.0	600.0	720.0
h_20	196.0	4.85	1.77	0.0	5.0	6.0	6.0	6.0
h_22	196.0	841.04	52.6	660.0	840.0	840.0	870.0	960.0
h_23	196.0	846.15	41.54	720.0	840.0	840.0	870.0	960.0
h_24	196.0	5.11	1.7	0.0	5.0	6.0	6.0	6.0

h_26	196.0	1256.86	54.43	1140.0	1230.0	1260.0	1290.0	1440.0
h_27	196.0	1262.72	56.41	1140.0	1230.0	1260.0	1290.0	1440.0
h_28	196.0	5.36	1.39	0.0	5.0	6.0	6.0	6.0
h_30	196.0	2.01	0.84	0.0	1.0	2.0	3.0	3.0
h_31	196.0	1.26	0.77	0.0	1.0	1.0	2.0	3.0
h_32	196.0	1.33	0.81	0.0	1.0	1.0	2.0	3.0
h_33	196.0	0.34	0.66	0.0	0.0	0.0	0.25	3.0
h_34	196.0	0.85	0.87	0.0	0.0	1.0	2.0	3.0
h_35_a	196.0	2.28	0.83	0.0	2.0	2.0	3.0	3.0
h_36	196.0	1.66	0.91	0.0	1.0	2.0	2.0	3.0
h_37	196.0	1.51	0.88	0.0	1.0	1.0	2.0	3.0
h_38	196.0	2.0	0.91	0.0	1.75	2.0	3.0	3.0
h_39	196.0	2.43	0.8	0.0	2.0	3.0	3.0	3.0
h_40	196.0	2.36	0.79	0.0	2.0	3.0	3.0	3.0
h_41	196.0	2.46	0.9	0.0	2.0	3.0	3.0	3.0
h_42	196.0	1.08	0.78	0.0	1.0	1.0	1.0	3.0
h_43_1	196.0	2.39	0.98	0.0	2.0	3.0	3.0	3.0
h_43_2	196.0	1.53	1.07	0.0	1.5	1.5	1.5	4.0
h_44_1	196.0	2.25	1.03	0.0	2.0	3.0	3.0	3.0
h_44_2	196.0	1.64	0.86	0.0	1.0	2.0	2.0	3.0
h_44_3	196.0	1.09	1.06	0.0	0.0	1.0	2.0	3.0
h_45	196.0	0.87	0.79	0.0	0.0	1.0	1.0	3.0
z_52	196.0	2.83	2.15	0.0	1.0	2.0	4.0	7.0
z_54	196.0	1.56	0.63	0.0	1.0	2.0	2.0	3.0
z_56_1	196.0	1.57	1.32	0.0	0.0	2.0	3.0	4.0
z_56_2	196.0	1.66	1.38	0.0	0.0	2.0	3.0	4.0
z_56_3	196.0	1.75	1.42	0.0	0.0	2.0	3.0	4.0
z_56_4	196.0	1.52	1.36	0.0	0.0	1.0	3.0	4.0
z_56_5	196.0	2.12	1.41	0.0	1.0	2.0	3.0	4.0
z_56_6	196.0	1.82	1.45	0.0	0.0	2.0	3.0	4.0
z_56_7	196.0	1.79	1.49	0.0	0.0	2.0	3.0	4.0
h_57	196.0	2.07	1.14	0.0	1.0	2.0	3.0	4.0
h_58	196.0	2.61	1.11	0.0	2.0	3.0	4.0	4.0
h_59	196.0	2.23	1.14	0.0	1.0	2.0	3.0	4.0
h_60	196.0	2.31	1.16	0.0	1.0	2.0	3.0	4.0
h_61	196.0	0.83	1.03	0.0	0.0	1.0	1.0	4.0
h_62	196.0	1.03	1.18	0.0	0.0	1.0	2.0	4.0
h_63_1	196.0	8.31	1.31	4.0	8.0	8.0	9.0	15.0

h_63_2	196.0	1358.07	62.77	1140.0	1320.0	1380.0	1380.0	1530.0
h_64_1	196.0	9.21	1.48	4.0	8.0	9.0	10.0	14.0
h_64_2	196.0	1387.91	65.66	1140.0	1350.0	1380.0	1440.0	1560.0
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_82	196.0	0.75	1.28	0.0	0.0	0.0	1.0	4.0
n_83	196.0	0.59	1.15	0.0	0.0	0.0	1.0	4.0
n_84	196.0	0.42	1.0	0.0	0.0	0.0	0.0	4.0
n_85	196.0	0.3	0.9	0.0	0.0	0.0	0.0	4.0
n_86	196.0	1.24	1.53	0.0	0.0	0.0	3.0	4.0
n_87	196.0	0.84	1.31	0.0	0.0	0.0	1.0	4.0
n_88	196.0	0.51	1.1	0.0	0.0	0.0	0.0	4.0
n_89	196.0	0.31	0.86	0.0	0.0	0.0	0.0	4.0
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
cross_Elemantry	196.0	0.61	0.49	0.0	0.0	1.0	1.0	1.0
cross_intermediate	196.0	0.39	0.49	0.0	0.0	0.0	1.0	1.0
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

3. Missing Values per Column

- id2: 0
- universi: 0
- region: 0
- cluster: 0
- birth_ye: 0
- sample_c: 0
- h_18: 0
- h_19: 0
- h_20: 0
- h_21: 0
- h_22: 0
- h_23: 0
- h_24: 0
- h_25: 0
- h_26: 0
- h_27: 0
- h_28: 0
- h_29: 0

h_30: 0
h_31: 0
h_32: 0
h_33: 0
h_34: 0
h_35_a: 0
h_36: 0
h_37: 0
h_38: 0
h_39: 0
h_40: 0
h_41: 0
h_42: 0
h_43_1: 0
h_43_2: 0
h_44_1: 0
h_44_2: 0
h_44_3: 0
h_45: 0
z_52: 0
z_53: 0
z_54: 0
z_56_1: 0
z_56_2: 0
z_56_3: 0
z_56_4: 0
z_56_5: 0
z_56_6: 0
z_56_7: 0
h_57: 0
h_58: 0
h_59: 0
h_60: 0
h_61: 0
h_62: 0
h_63_1: 0
h_63_2: 0
h_64_1: 0
h_64_2: 0
sumScore: 0
tertiles: 0
n_82: 0
n_83: 0
n_84: 0
n_85: 0

n_86: 0

n_87: 0

n_88: 0

n_89: 0

n_90: 0

n_91: 0

n_92: 0

n_93: 0

z_55: 0

cross_Elemantry: 0

cross_intermidiate: 0

sex_Boy: 0

sex_Girl: 0