Analisis Data Obesitas dengan Streamlit

Unggah file CSV dataset obesitas

| ☁ | **Drag and drop file here** Limit 200MB per file • CSV | | Browse files |

| 📄 | ObesityDataSet.csv 254.9KB | ✕ |

Preview Dataset

	Age	Gender	Height	Weight	CALC	FAVC	FCVC	NCP	SCC	SMOKE	CH2O	famil
0	21	Female	1.62	64	no	no	2	3	no	no	2	yes
1	21	Female	1.52	56	Sometimes	no	3	3	yes	yes	3	yes
2	23	Male	1.8	77	Frequently	no	2	3	no	no	2	yes
3	27	Male	1.8	87	Frequently	no	3	3	no	no	2	no
4	22	Male	1.78	89.8	Sometimes	no	2	1	no	no	2	no

Informasi Dataset

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 2111 entries, 0 to 2110
Data columns (total 17 columns):
 #   Column          Non-Null Count  Dtype
---  ------          --------------  -----
 0   Age             2097 non-null   object
 1   Gender          2102 non-null   object
 2   Height          2099 non-null   object
 3   Weight          2100 non-null   object
 4   CALC            2106 non-null   object
 5   FAVC            2100 non-null   object
 6   FCVC            2103 non-null   object
```
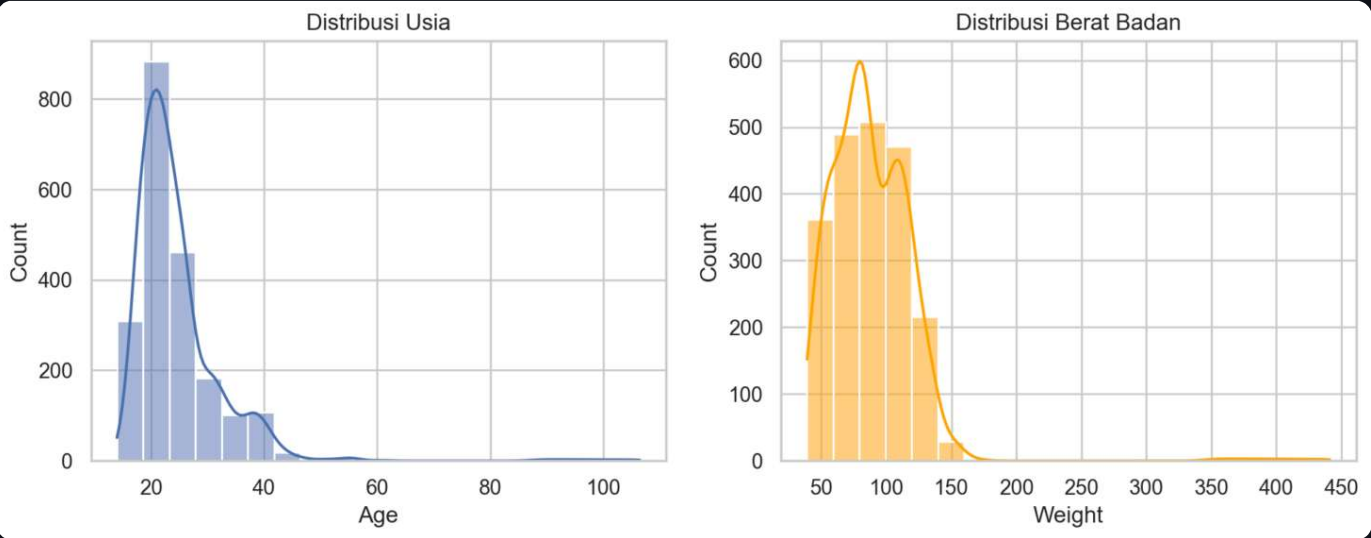
```
 7  NCP                  2099 non-null   object
 8  SCC                  2101 non-null   object
 9  SMOKE                2106 non-null   object
10  CH2O                 2105 non-null   object
11  family_history_with_overweight 2098 non-null   object
12  FAF                  2103 non-null   object
13  TUE                  2102 non-null   object
14  CAEC                 2100 non-null   object
15  MTRANS               2105 non-null   object
16  NObeyesdad           2111 non-null   object
dtypes: object(17)
memory usage: 280.5+ KB
```
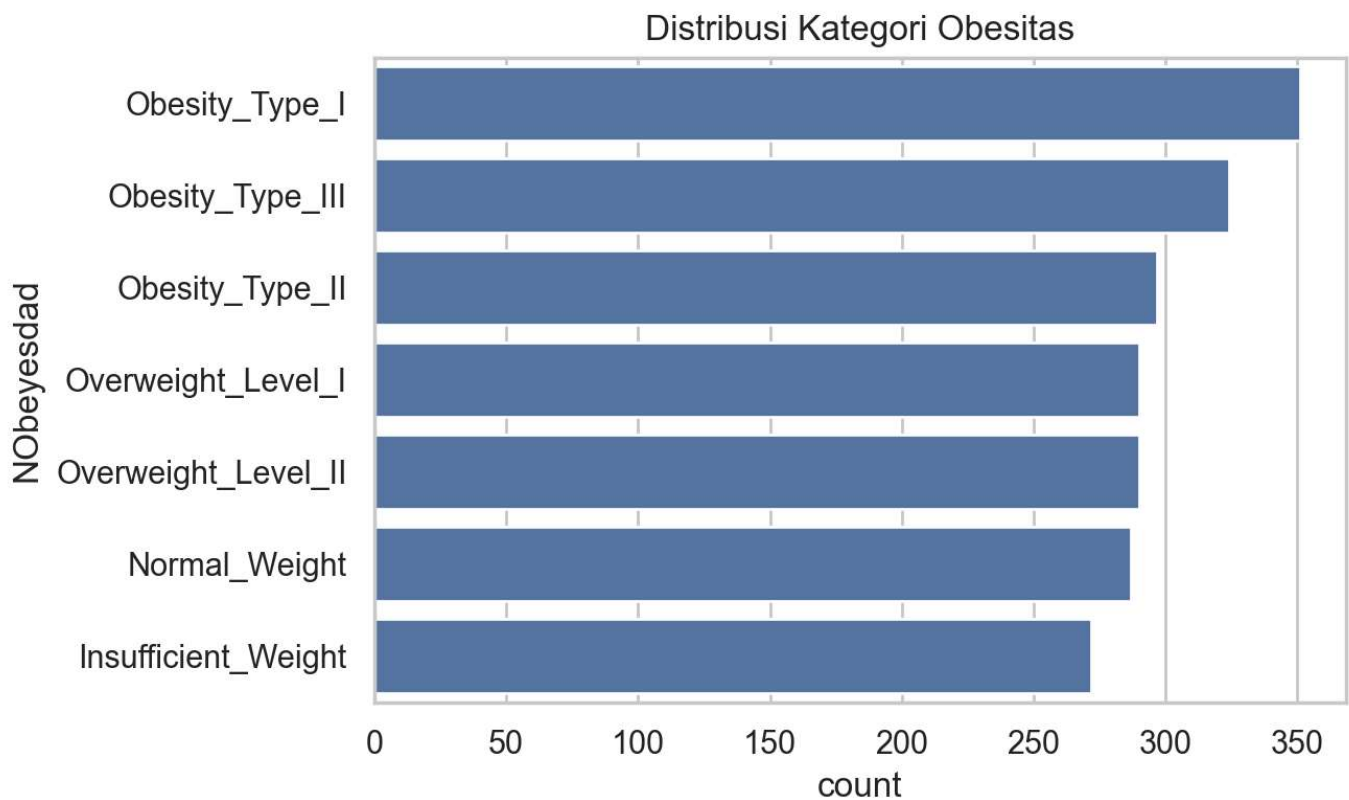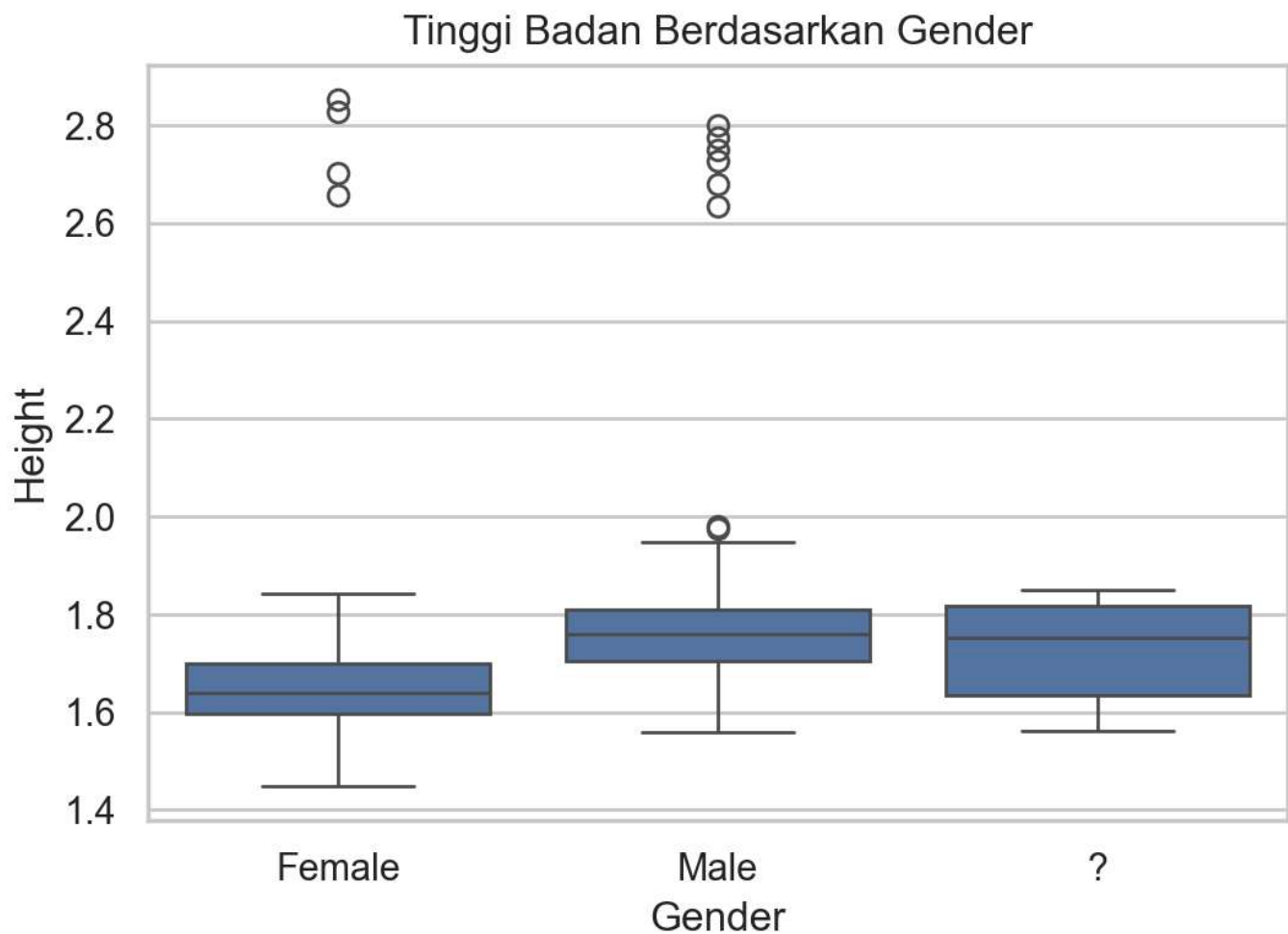
Statistik Deskriptif

	Age	Gender	Height	Weight	CALC	FAVC	FCVC	NCP	SCC	SMOKE	CH2O	fam
count	2097	2102	2099	2100	2106	2100	2103	2099	2101	2106	2105	209
unique	1394	3	1562	1518	5	3	808	637	3	3	1263	3
top	18	Male	1.7	80	Sometimes	yes	3	3	no	no	2	yes
freq	124	1056	58	58	1386	1844	647	1183	1997	2054	441	170

Visualisasi Data

Distribusi Usia — Distribusi Berat Badan

Tinggi Badan Berdasarkan Gender

Distribusi Kategori Obesitas

Preprocessing Data

Data berhasil dibersihkan.

Jumlah data setelah preprocessing: 1182 baris

Modeling dan Evaluasi

Pilih proporsi data uji (%)

20

●

10
50

Hasil Evaluasi Model

	Akurasi	Precision	Recall	F1 Score
Decision Tree	0.9427	0.9437	0.9427	0.9427
Random Forest	0.9740	0.9748	0.9740	0.9740
SVM	0.8672	0.8794	0.8672	0.8686

Perbandingan Performansi Model