Data Collection and Preprocessing Phase

Date	05-07-2024
Team ID	739733
Project Title	Fetal AI: Using Machine Learning To Predict And Monitor Fetal Health
Maximum Marks	6 Marks

Data Exploration and Preprocessing Report

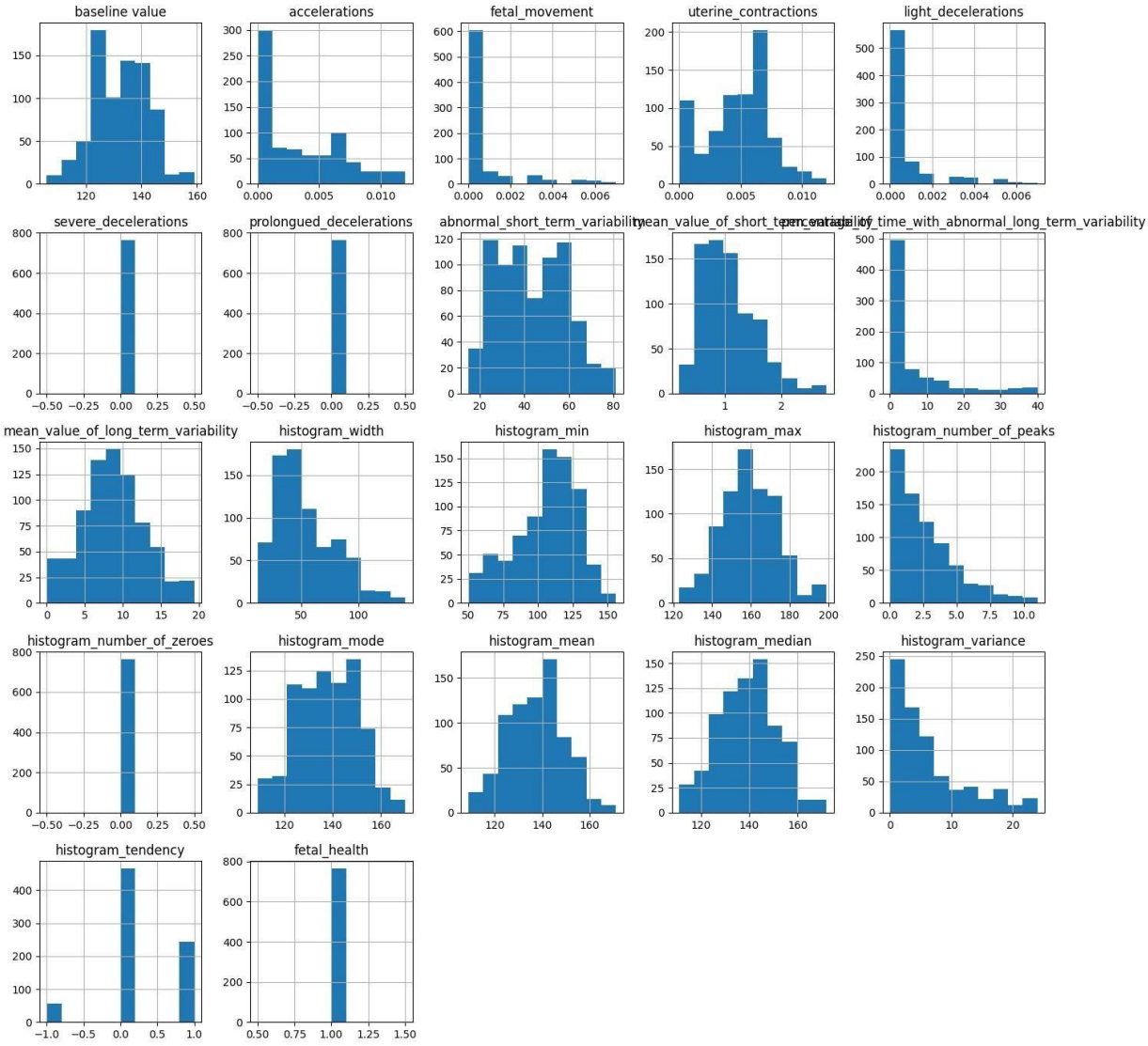
Dataset variables will be statistically analyzed to identify patterns and outliers, with Python employed for preprocessing tasks like normalization and feature engineering. Data cleaning will address missing values and outliers, ensuring quality for subsequent analysis and modeling, and forming a strong foundation for insights and predictions.

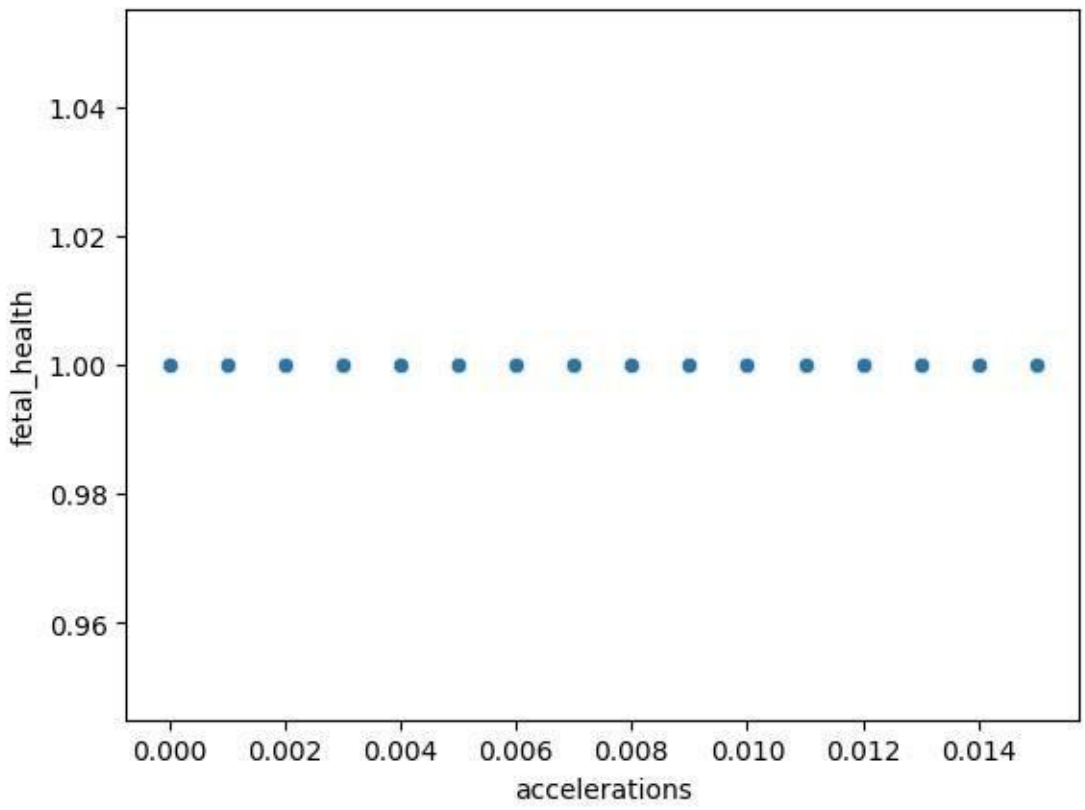
Section	Description
Data Overview	Dimension: 8 rows × 22 columns Descriptive statistics: `df.describe()`

	baseline value	accelerations	fetal_movement	uterine_contractions	light_decelerations	severe_decelerations	prolongued_decelerations	abnormal_short_term_variabil
count	2126.000000	2126.000000	2126.000000	2126.000000	2126.000000	2126.000000	2126.000000	2126.000
mean	133.303857	0.003178	0.009481	0.004366	0.001889	0.000003	0.000159	46.990
std	9.840844	0.003866	0.046666	0.002946	0.002960	0.000057	0.000590	17.192
min	106.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	12.000
25%	126.000000	0.000000	0.000000	0.002000	0.000000	0.000000	0.000000	32.000
50%	133.000000	0.002000	0.000000	0.004000	0.000000	0.000000	0.000000	49.000
75%	140.000000	0.006000	0.003000	0.007000	0.003000	0.000000	0.000000	61.000
max	160.000000	0.019000	0.481000	0.015000	0.015000	0.001000	0.005000	87.000

8 rows × 22 columns

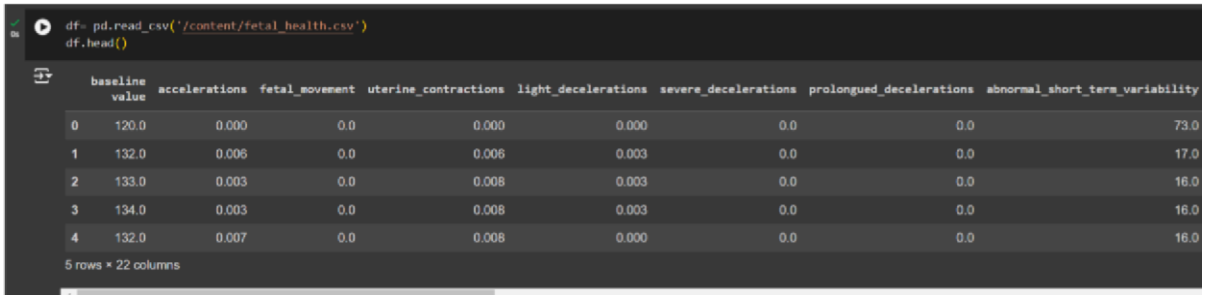
Univariate Analysis



Bivariate Analysis	

Multivariate Analysis	

Data Preprocessing Code Screenshots

Outliers and Anomalies	-
Loading Data	
Handling Missing Data	-
Data Transformation	-
Feature Engineering	Attached the codes in final submission.
Save Processed Data	-

```
df= pd.read_csv('/content/fetal_health.csv')
df.head()
```

	baseline value	accelerations	fetal_movement	uterine_contractions	light_decelerations	severe_decelerations	prolongued_decelerations	abnormal_short_term_variability
0	120.0	0.000	0.0	0.000	0.000	0.0	0.0	73.0
1	132.0	0.006	0.0	0.006	0.003	0.0	0.0	17.0
2	133.0	0.003	0.0	0.008	0.003	0.0	0.0	16.0
3	134.0	0.003	0.0	0.008	0.003	0.0	0.0	16.0
4	132.0	0.007	0.0	0.008	0.000	0.0	0.0	16.0

5 rows × 22 columns