



Data Collection and Preprocessing Phase

	- representations
Date	05-07-2024
Team ID	740070
Project Title	FetalAI: 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

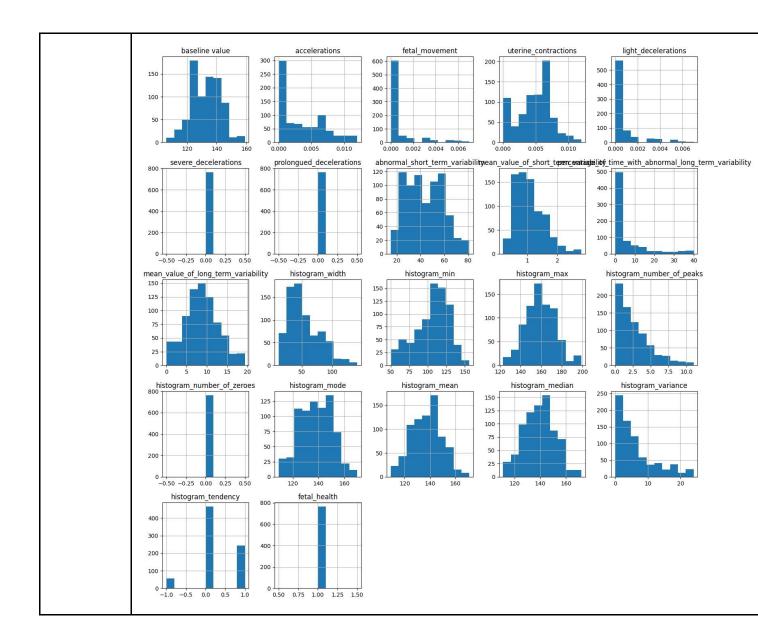






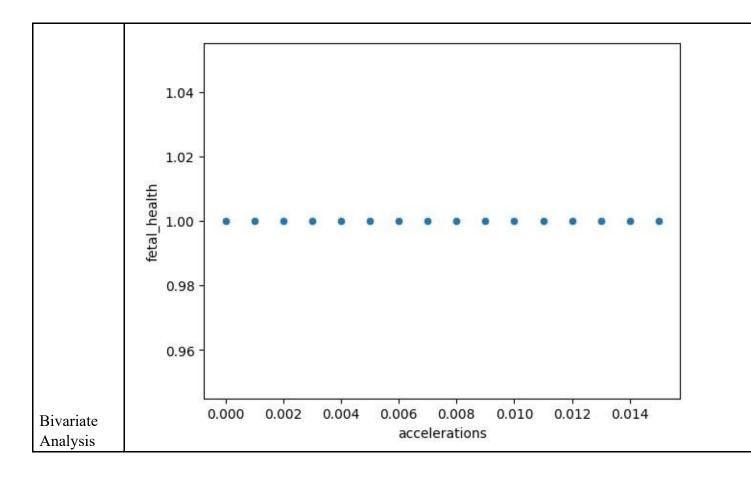
















		21	121022									252								-				- 1.00
	baseline value		0.051										-0.025					0.92		0.93				
	accelerations -	- 0.051	1	0.077	0.079	-0.088			-0.067	0.24	-0.36	-0.4	0.34	-0.084	0.49	0.27		0.25	0.3	0.28	0.38	-0.06		
	fetal_movement	0.02	0.077	1	-0.29	-0.0035			0.35	0.048	-0.047	-0.049	0.17	-0.17	0.03	0.092		-0.036	-0.043	-0.033	0.047	0.08		- 0.75
	uterine_contractions	0.022	0.079	-0.29	1	0.22			-0.14	0.18	-0.11	-0.12	-0.039	0.097	0.082	0.042		0.057	0.026	0.044	0.2	-0.039		
	light_decelerations	0.017	-0.088	-0.0035	0.22	1			-0.069	0.39	-0.15	-0.1	0.35	-0.38	0.036	0.22		-0.063	-0.19	-0.1	0.61	0.24		
	severe_decelerations	-																						- 0.50
	prolongued_decelerations																							
	abnormal_short_term_variability	- 0.27	-0.067	0.35	-0.14	-0.069			1	-0.47	0.27	-0.32	-0.11	0.15	0.038	-0.03		0.24	0.21	0.24	-0.11	0.12		
	mean_value_of_short_term_variability	0.3	0.24	0.048	0.18	0.39			-0.47	1	-0.41	0.056	0.44	-0.44	0.11	0.34		-0.24	-0.32	-0.27	0.55	0.013		
	percentage_of_time_with_abnormal_long_term_variability	- 0.15	-0.36	-0.047	-0.11	-0.15			0.27	-0.41	1	-0.18	-0.3	0.23	-0.18	-0.14		0.07	0.081	0.076	-0.38	-0.036		- 0.25
	mean_value_of_long_term_variability	0.14	-0.4	-0.049	-0.12	-0.1			-0.32	0.056	-0.18	1	0.16	-0.24	-0.091	0.019		-0.21	-0.19	-0.22	-0.068	0.052		
	histogram_width	0.025	0.34	0.17	-0.039	0.35			-0.11	0.44	-0.3	0.16	1	-0.84	0.51	0.71		0.058	0.021	0.051	0.63	0.34		
	histogram_min -	- 0.39	-0.084	-0.17	0.097	-0.38			0.15	-0.44	0.23	-0.24	-0.84	1	0.037	-0.57		0.35	0.41	0.37	-0.49	-0.45		- 0.00
	histogram_max -	- 0.57	0.49	0.03	0.082	0.036			0.038	0.11	-0.18	-0.091	0.51	0.037	1	0.41		0.66	0.69	0.69	0.39	-0.077		
	histogram_number_of_peaks	0.017	0.27	0.092	0.042	0.22			-0.03	0.34	-0.14	0.019	0.71	-0.57	0.41	1		0.047	0.024	0.054	0.46	0.24		
	histogram_number_of_zeroes																							0.25
	histogram_mode		0.25	-0.036	0.057	-0.063			0.24	-0.24	0.07	.0.21	0.058	0.35	0.66	0.047		1	0.96	0.98	0.024	0.27		
	histogram_mean												0.021					0.96			-0.045			
																								0.50
	histogram_median												0.051					0.98	0.99		0.0087			
	histogram_variance												0.63							0.0087		0.14		
	histogram_tendency ·	- 0.23	-0.06	0.08	-0.039	0.24			0.12	0.013	-0.036	0.052	0.34	-0.45	-0.077	0.24		0.27	0.17	0.23	0.14	1		0.75
	fetal_health		, vo		-	-	- 8		- /	-	- /	- /	÷	-		,		62	-	-	1	- /		
		baseline value	acceleration	fetal_movemen	ne_contraction	light_decelerations	eleration	prolongued_deceleration	nal_short_term_variability	of_short_term_variability	nal_long_term_variability	_of_long_term_variability	histogram_width	histogram_mir	histogram_max	sber_of_peak	oper_of_zeroe	histogram_mode	nstogram_mea	ogram_mediar	gram_variance	tendenc	fetal_health	
		base	ac	fetal	derine_co	light_dec	severe_decelerati	pap panbi	nort_term_	ort_term_	ong_term_	ong_term_	histog	histo	histor	n_number	number	histog	histog	histogra	histogram	histogram	ž	
							ø	prolor		-	E	e e				histogran	histogram							
									abnor	mean_value	e_with_abn	mean_val					-03							
											ge_of_tim													
Multivariate											percenta													
Analysis																								
Outliers and																								
Anomalies	-																							





Data Preprocessing Code Screenshots





	ı								
	€ 0	df= pd.read_d		fetal_health.csv'					
	∓	baseline	accelerations	fetal movement	uterine contractions	light decelerations	severe deceleration	a prolongued deceleration	s abnormal_short_term_variability
	,	0 120.0							
		1 132.0				0.003			
	, 7	2 133.0							
	i 7	3 134.0	0.003	0.0	0.008	0.003	3 0.0	0 0.0	0 16.0
	, 7								
	ı 🗐	5 rows × 22 colu	umns						
Loading Data									
	i					-			
Handling	i		-						
Missing Data	i								
							-		
	ı		-						
Data	i								
Transformation	i								
Transformation	<u> </u>								
	ı								
Feature	i								
Engineering	Attac	ched th	e codes	in final s	submission.				
	1								
Save	i								
Processed Data	l -								