Developing a fetal classifier.

A data science project by
Kenneth Imade

Photo by Volodymyr Hryshchenko on Unsplash

Objective

- To develop a machine learning model capable of classifying fetus health into three categories; normal, suspect, or pathological
- Evaluate the performance of various machine learning classifiers and choose the best one for classifying fetus health

Overview

Context
Reduction of child mortality is reflected in several of the United Nations' Sustainable Development Goals and is a key indicator of human progress. The UN expects that by 2030, countries end preventable deaths of newborns and children under 5 years of age, with all countries aiming to reduce under-5 mortality to at least as low as 25 per 1,000 live births.

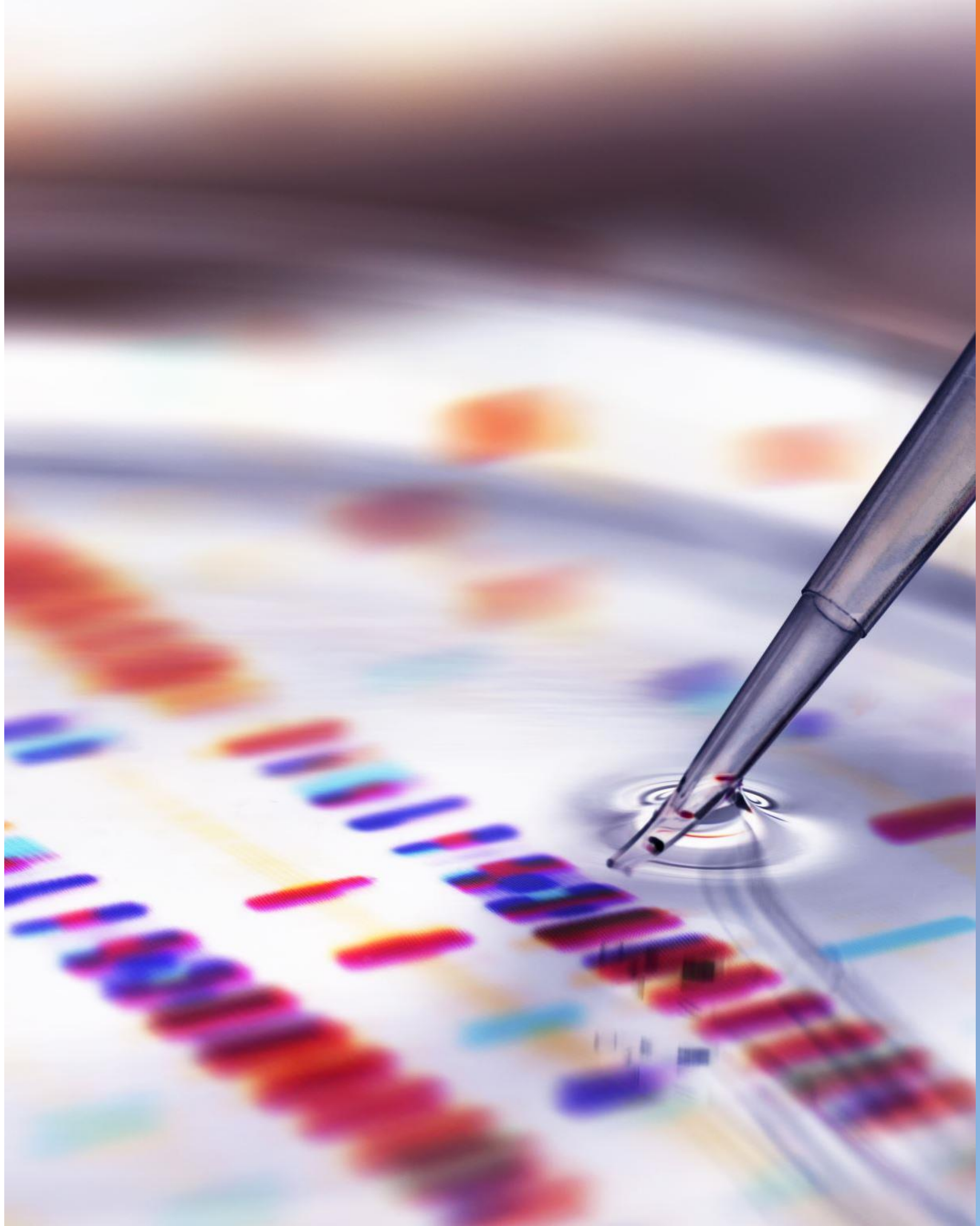
Parallel to notion of child mortality is of course maternal mortality, which accounts for 295 000 deaths during and following pregnancy and childbirth (as of 2017). The vast majority of these deaths (94%) occurred in low-resource settings, and most could have been prevented.

In light of what was mentioned above, Cardiotocograms (CTGs) are a simple and cost accessible option to assess fetal health, allowing healthcare professionals to take action in order to prevent child and maternal mortality. The equipment itself works by sending ultrasound pulses and reading its response, thus shedding light on fetal heart rate (FHR), fetal movements, uterine contractions and more.

Data Source: Fetal Health Classification | Kaggle

Acknowledgement: Ayres de Campos et al. (2000) SisPorto 2.0 A Program for Automated Analysis of Cardiotocograms. J Matern Fetal Med 5:311-318 (link)

Exploratory Data Analysis



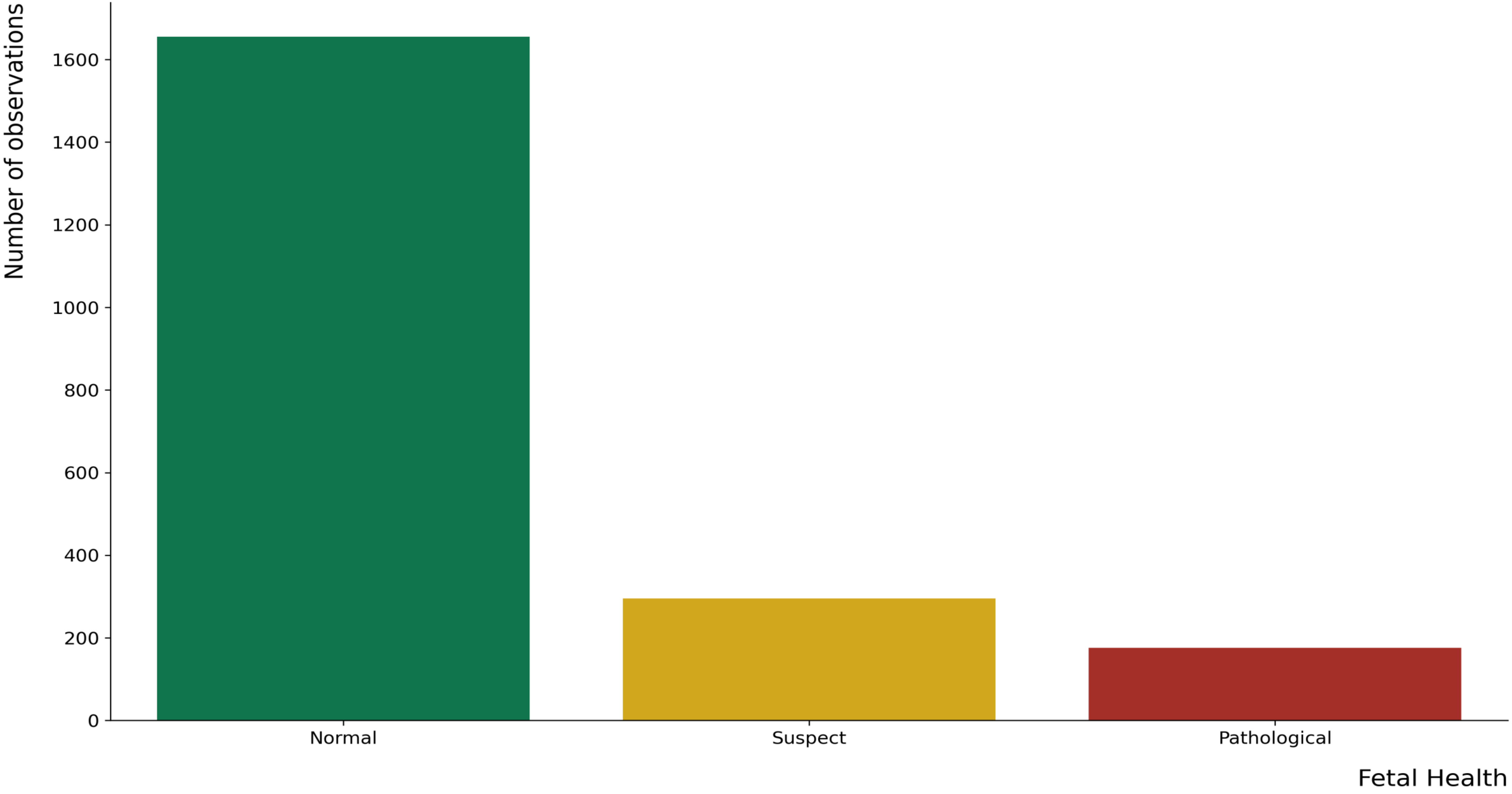
Dataset Statistics

N = 2,126	Mean	Std	Min	25%	50%	75%	Max
Baseline value	133.303857	9.840844	106.0	126.000	133.000	140.000	160.000
Accelerations	0.003178	0.003866	0.0	0.000	0.002	0.006	0.019
Fetal movement	0.009481	0.046666	0.0	0.000	0.000	0.003	0.481
Uterine contractions	0.004366	0.002946	0.0	0.002	0.004	0.007	0.015
Light decelerations	0.001889	0.002960	0.0	0.000	0.000	0.003	0.015
Severe decelerations	0.000003	0.000057	0.0	0.000	0.000	0.000	0.001
Prolongued decelerations	0.000159	0.000590	0.0	0.000	0.000	0.000	0.005
Abnormal short term variability	46.990122	17.192814	12.0	32.000	49.000	61.000	87.000
Mean value of short term variability	1.332785	0.883241	0.2	0.700	1.200	1.700	7.000
Percentage of time with abnormal long term variability	9.846660	18.396880	0.0	0.000	0.000	11.000	91.000
Mean value of long term variability	8.187629	5.628247	0.0	4.600	7.400	10.800	50.700

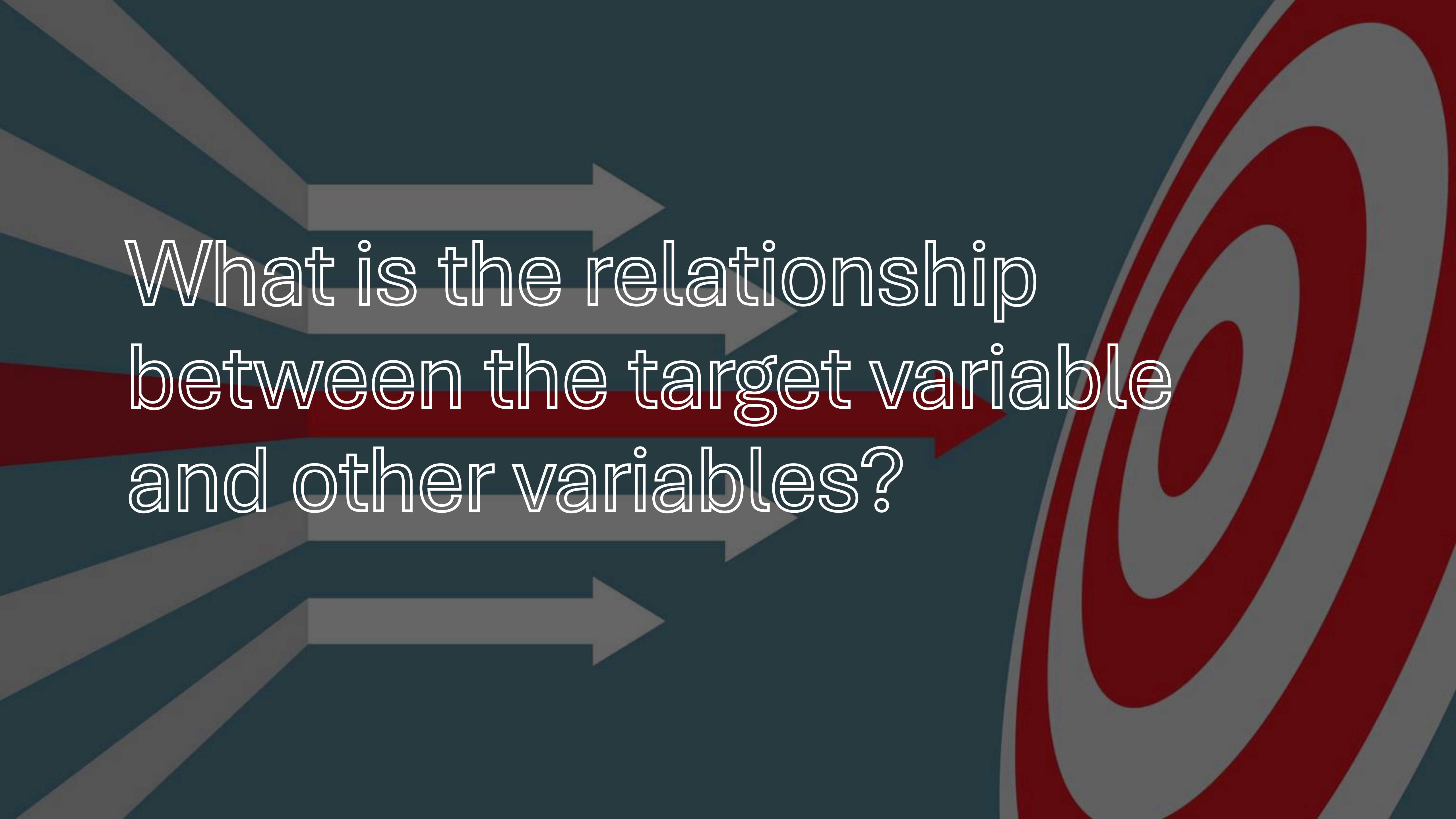
Dataset Statistics Continued

N = 2,126	**Mean**	**Std**	**Min**	**25%**	**50%**	**75%**	**Max**
Histogram width	70.445908	38.955693	3.0	37.000	67.500	100.000	180.000
Histogram min	93.579492	29.560212	50.0	67.000	93.000	120.000	159.000
Histogram max	164.025400	17.944183	122.0	152.000	162.000	174.000	238.000
Histogram number of peaks	4.068203	2.949386	0.0	2.000	3.000	6.000	18.000
Histogram number of zeroes	0.323612	0.706059	0.0	0.000	0.000	0.000	10.000
Histogram mode	137.452023	16.381289	60.0	129.000	139.000	148.000	187.000
Histogram mean	134.610536	15.593596	73.0	125.000	136.000	145.000	182.000
Histogram median	138.090310	14.466589	77.0	129.000	139.000	148.000	186.000
Histogram variance	18.808090	28.977636	0.0	2.000	7.000	24.000	269.000
Histogram tendency	0.320320	0.610829	-1.0	0.000	0.000	1.000	1.000
Fetal health (Target Variable)	**1.304327**	**0.614377**	**1.0**	**1.000**	**1.000**	**1.000**	**3.000**

What is the distribution of the target variable?

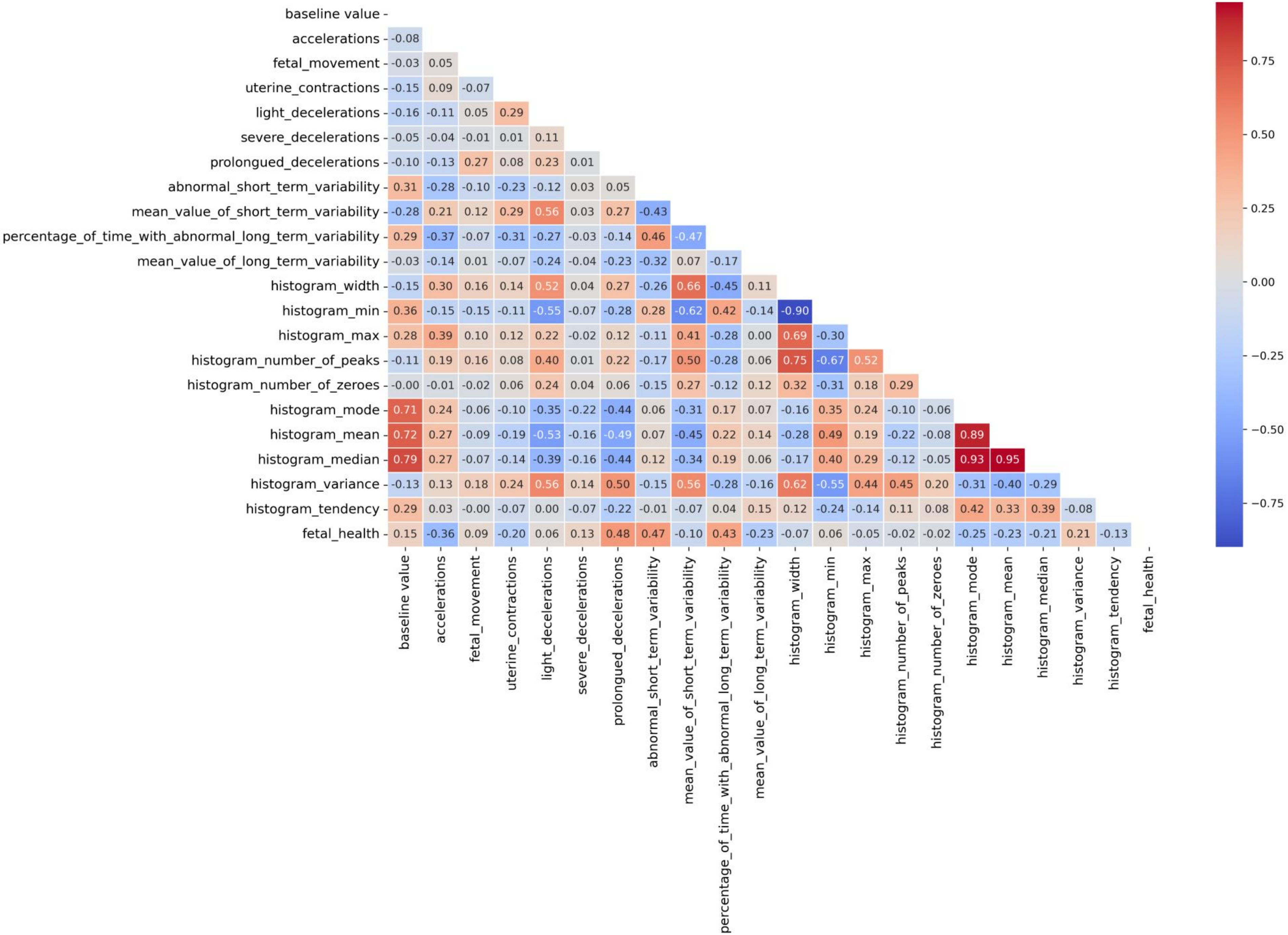


Most Fetuses are classified as normal



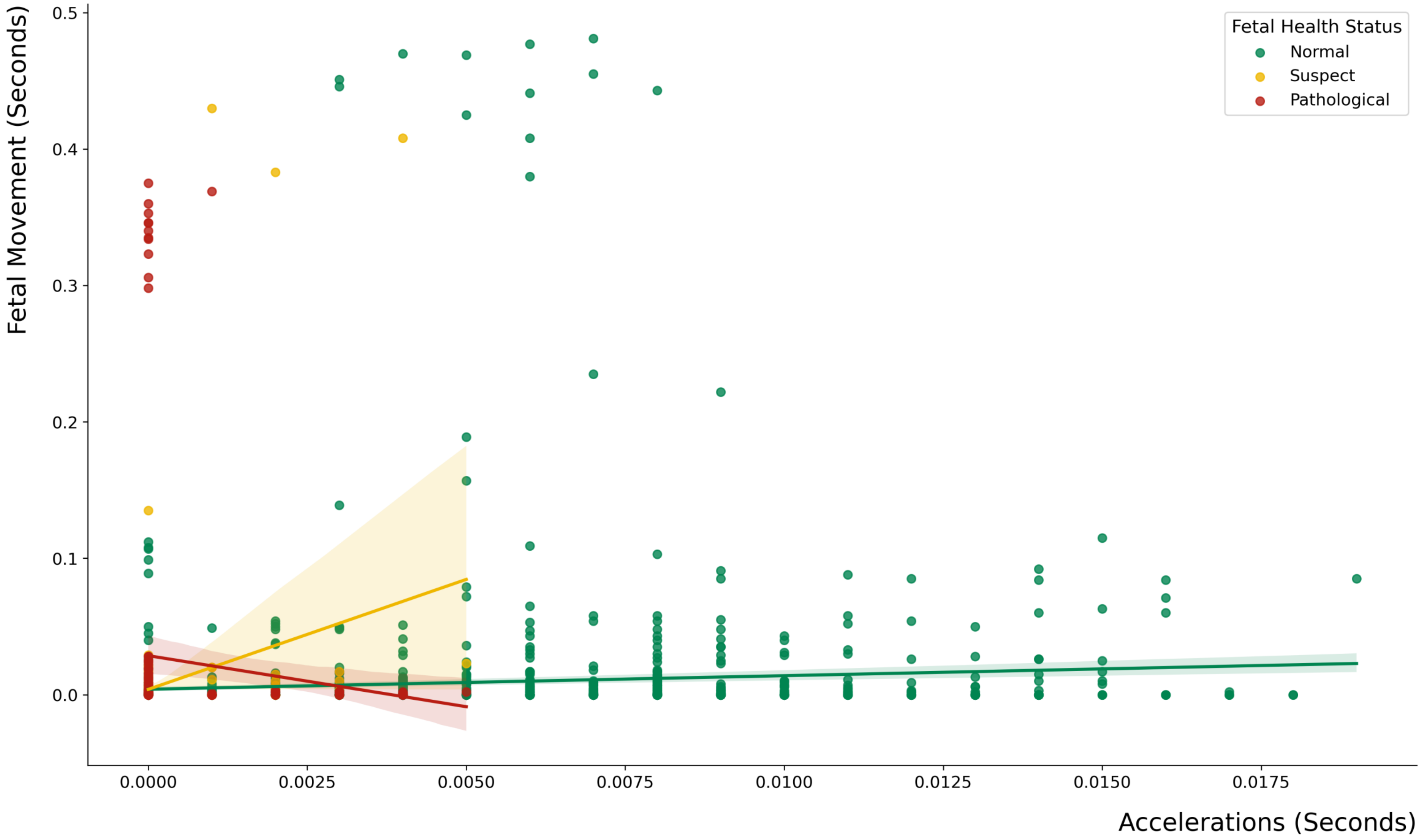
What is the relationship between the target variable and other variables?

Strong correlation exists between fetal health and other factors



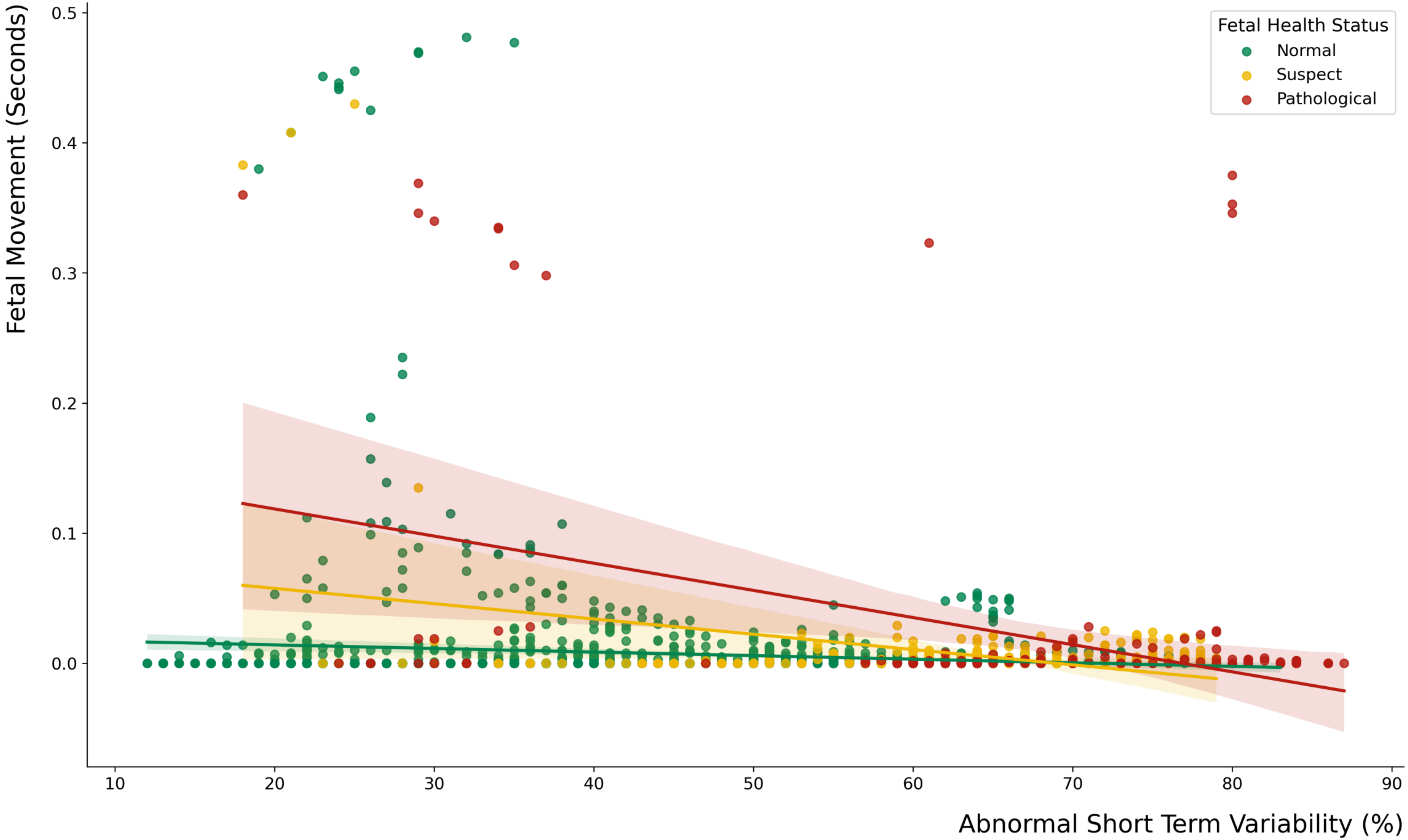
	baseline value	accelerations	fetal_movement	uterine_contractions	light_decelerations	severe_decelerations	prolongued_decelerations	abnormal_short_term_variability	mean_value_of_short_term_variability	percentage_of_time_with_abnormal_long_term_variability	mean_value_of_long_term_variability	histogram_width	histogram_min	histogram_max	histogram_number_of_peaks	histogram_number_of_zeroes	histogram_mode	histogram_mean	histogram_median	histogram_variance	histogram_tendency	fetal_health
baseline value																						
accelerations	-0.08																					
fetal_movement	-0.03	0.05																				
uterine_contractions	-0.15	0.09	-0.07																			
light_decelerations	-0.16	-0.11	0.05	0.29																		
severe_decelerations	-0.05	-0.04	-0.01	0.01	0.11																	
prolongued_decelerations	-0.10	-0.13	0.27	0.08	0.23	0.01																
abnormal_short_term_variability	0.31	-0.28	-0.10	-0.23	-0.12	0.03	0.05															
mean_value_of_short_term_variability	-0.28	0.21	0.12	0.29	0.56	0.03	0.27	-0.43														
percentage_of_time_with_abnormal_long_term_variability	0.29	-0.37	-0.07	-0.31	-0.27	-0.03	-0.14	0.46	-0.47													
mean_value_of_long_term_variability	-0.03	-0.14	0.01	-0.07	-0.24	-0.04	-0.23	-0.32	0.07	-0.17												
histogram_width	-0.15	0.30	0.16	0.14	0.52	0.04	0.27	-0.26	0.66	-0.45	0.11											
histogram_min	0.36	-0.15	-0.15	-0.11	-0.55	-0.07	-0.28	0.28	-0.62	0.42	-0.14	-0.90										
histogram_max	0.28	0.39	0.10	0.12	0.22	-0.02	0.12	-0.11	0.41	-0.28	0.00	0.69	-0.30									
histogram_number_of_peaks	-0.11	0.19	0.16	0.08	0.40	0.01	0.22	-0.17	0.50	-0.28	0.06	0.75	-0.67	0.52								
histogram_number_of_zeroes	-0.00	-0.01	-0.02	0.06	0.24	0.04	0.06	-0.15	0.27	-0.12	0.12	0.32	-0.31	0.18	0.29							
histogram_mode	0.71	0.24	-0.06	-0.10	-0.35	-0.22	-0.44	0.06	-0.31	0.17	0.07	-0.16	0.35	0.24	-0.10	-0.06						
histogram_mean	0.72	0.27	-0.09	-0.19	-0.53	-0.16	-0.49	0.07	-0.45	0.22	0.14	-0.28	0.49	0.19	-0.22	-0.08	0.89					
histogram_median	0.79	0.27	-0.07	-0.14	-0.39	-0.16	-0.44	0.12	-0.34	0.19	0.06	-0.17	0.40	0.29	-0.12	-0.05	0.93	0.95				
histogram_variance	-0.13	0.13	0.18	0.24	0.56	0.14	0.50	-0.15	0.56	-0.28	-0.16	0.62	-0.55	0.44	0.45	0.20	-0.31	-0.40	-0.29			
histogram_tendency	0.29	0.03	-0.00	-0.07	0.00	-0.07	-0.22	-0.01	-0.07	0.04	0.15	0.12	-0.24	-0.14	0.11	0.08	0.42	0.33	0.39	-0.08		
fetal_health	0.15	-0.36	0.09	-0.20	0.06	0.13	0.48	0.47	-0.10	0.43	-0.23	-0.07	0.06	-0.05	-0.02	-0.02	-0.25	-0.23	-0.21	0.21	-0.13	

How does Fetal health change with respect to these variables?

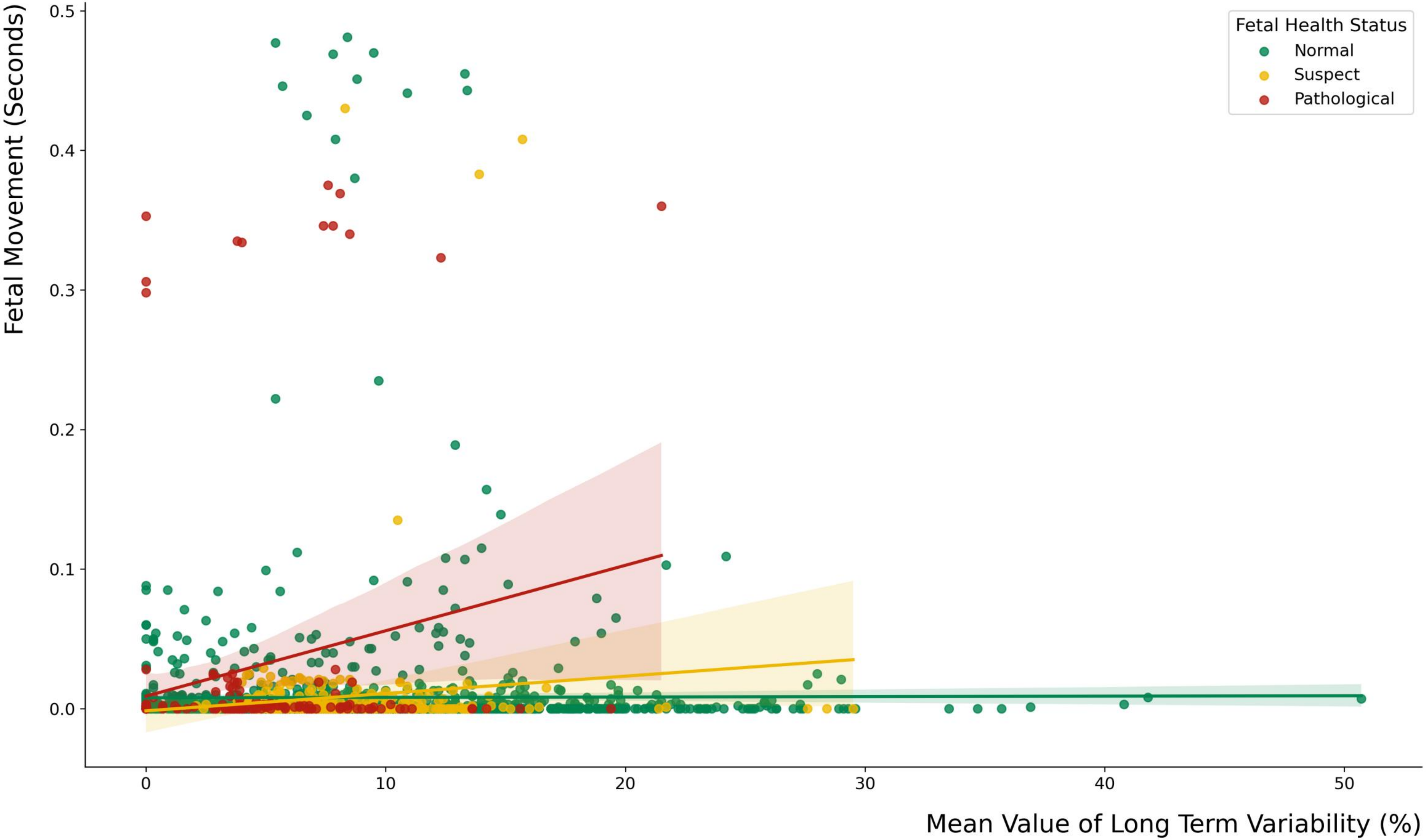


Tendency for both normal and suspect fetuses movement to slightly increase in tandem with their acceleration but relationship flatlines with pathological fetuses

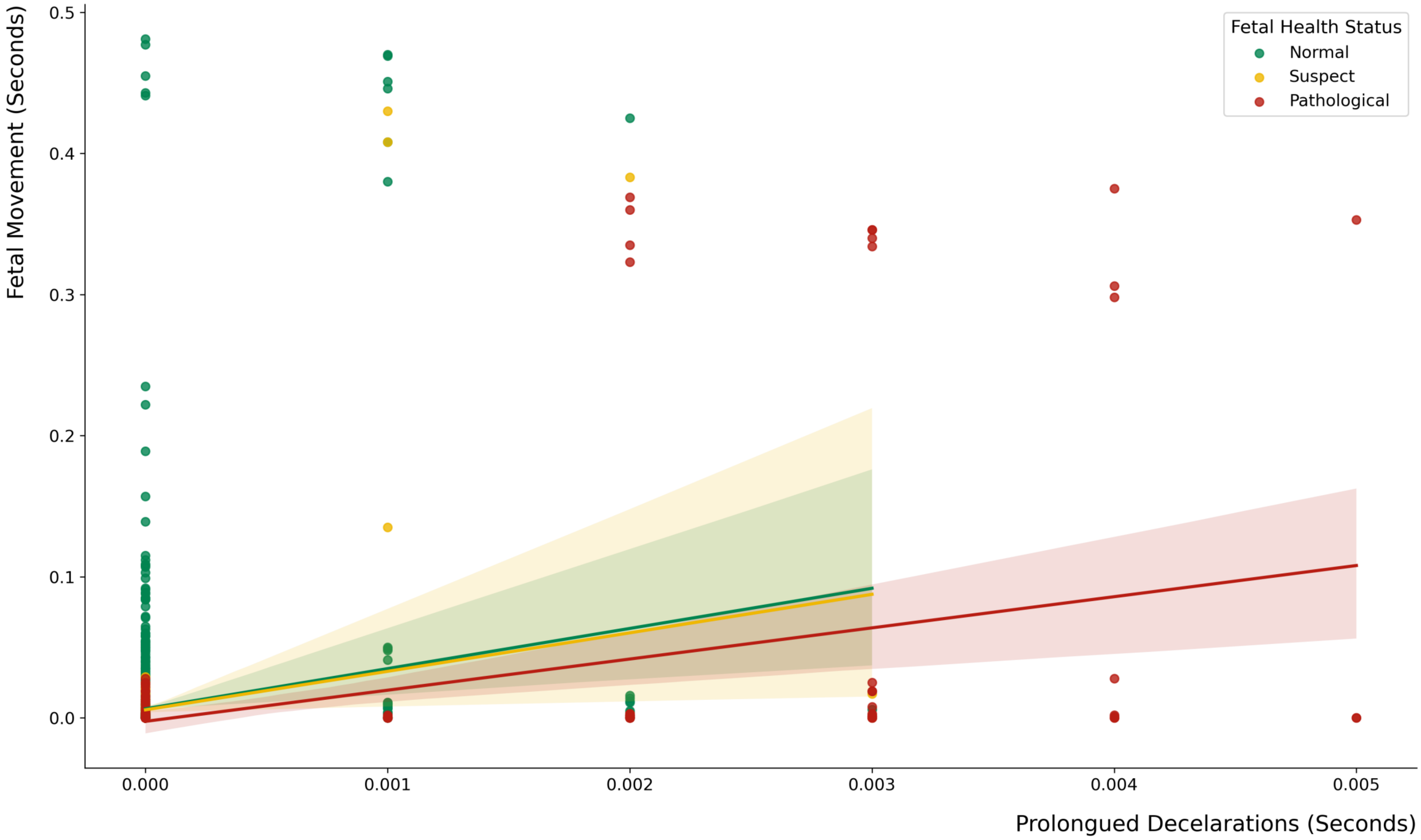
Increased short term fetal heart rate suggests a decrease in fetal movement across all fetal health categories

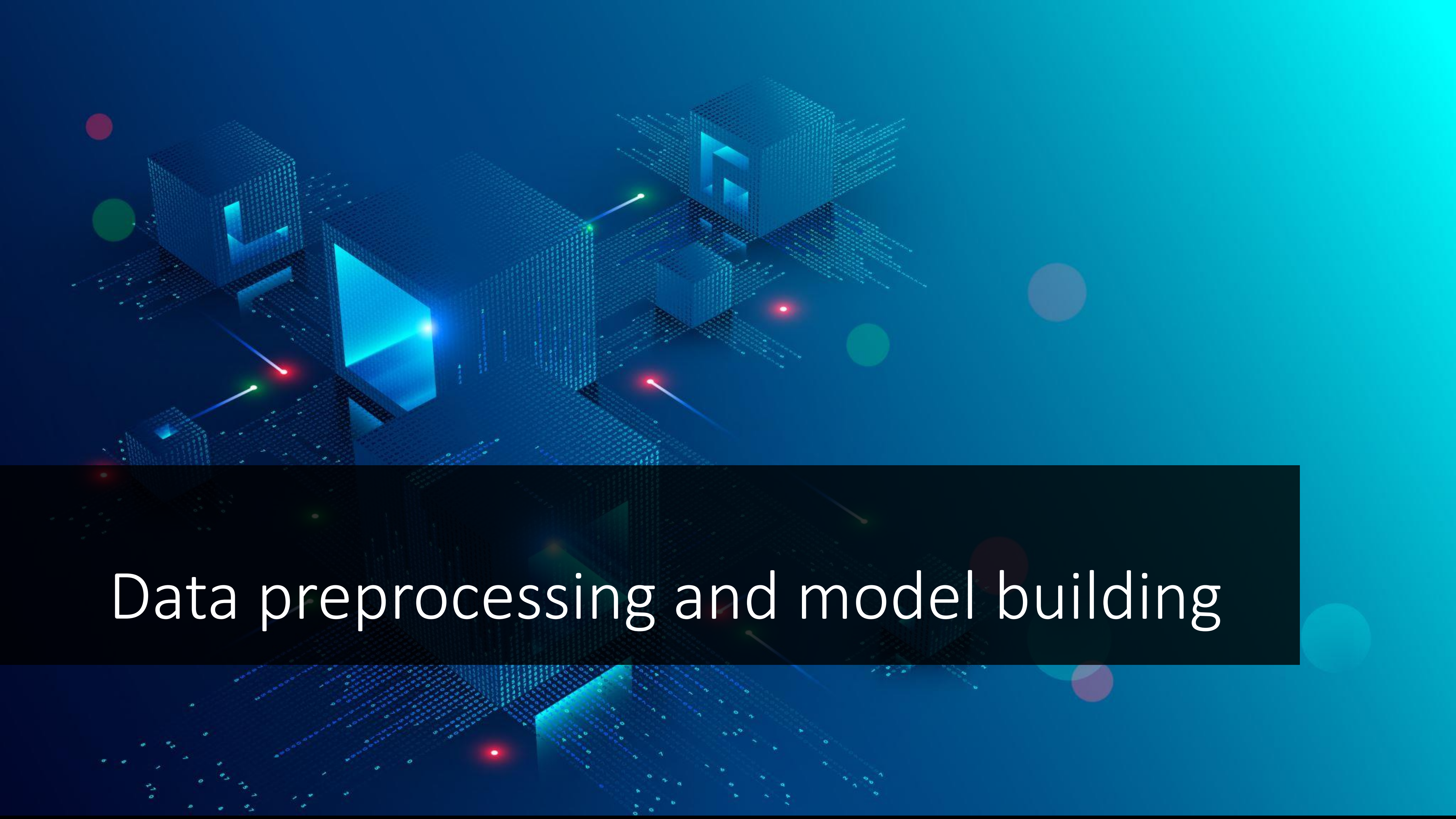


Strong relationship between long term variability and fetal movement in pathological fetuses but relationship is weaker in suspect fetuses and insignificant in normal fetuses



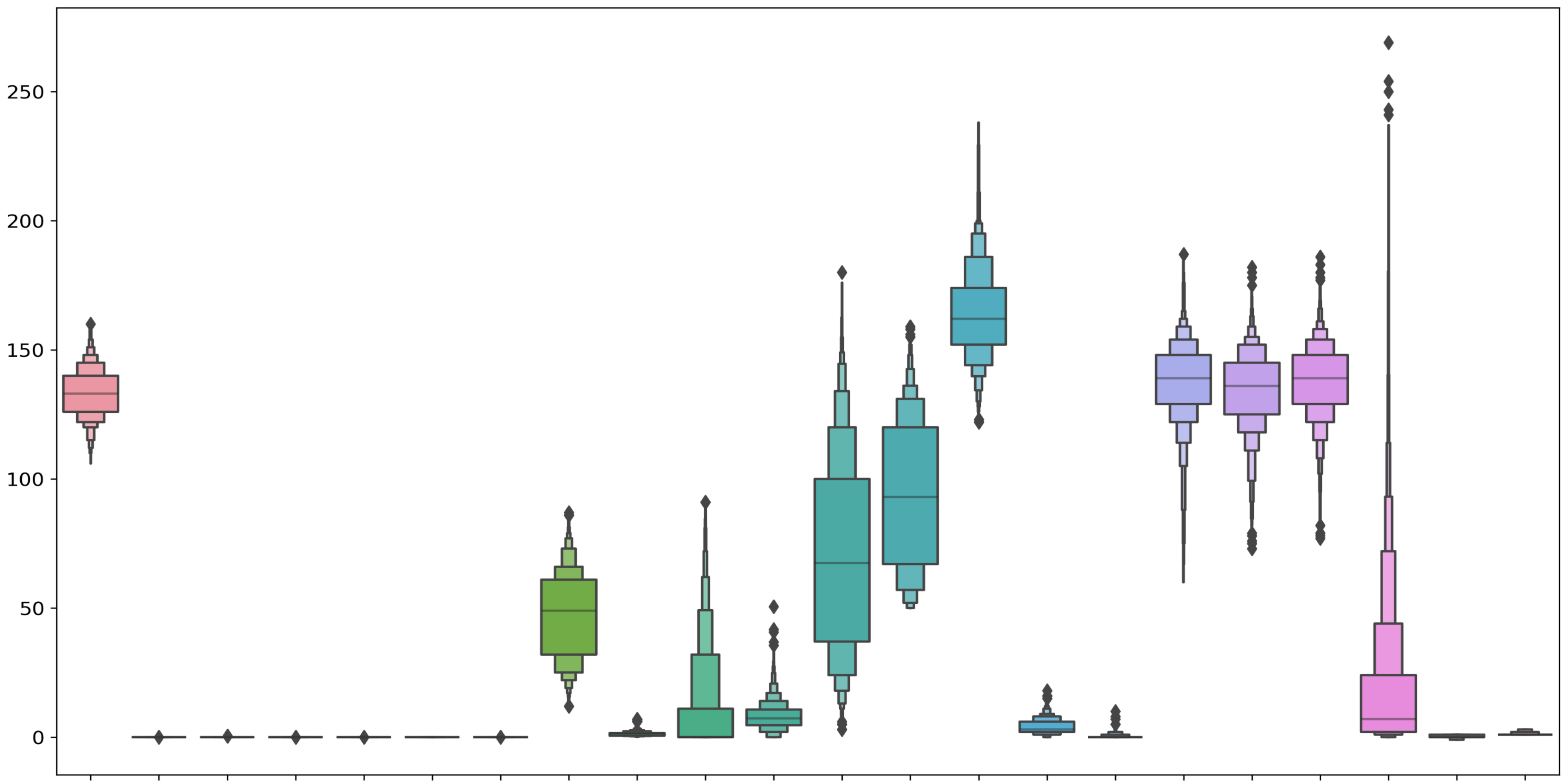
Fetal movement increases alongside prolongued decelerations in both normal and suspect fetuses but does not for pathological fetuses



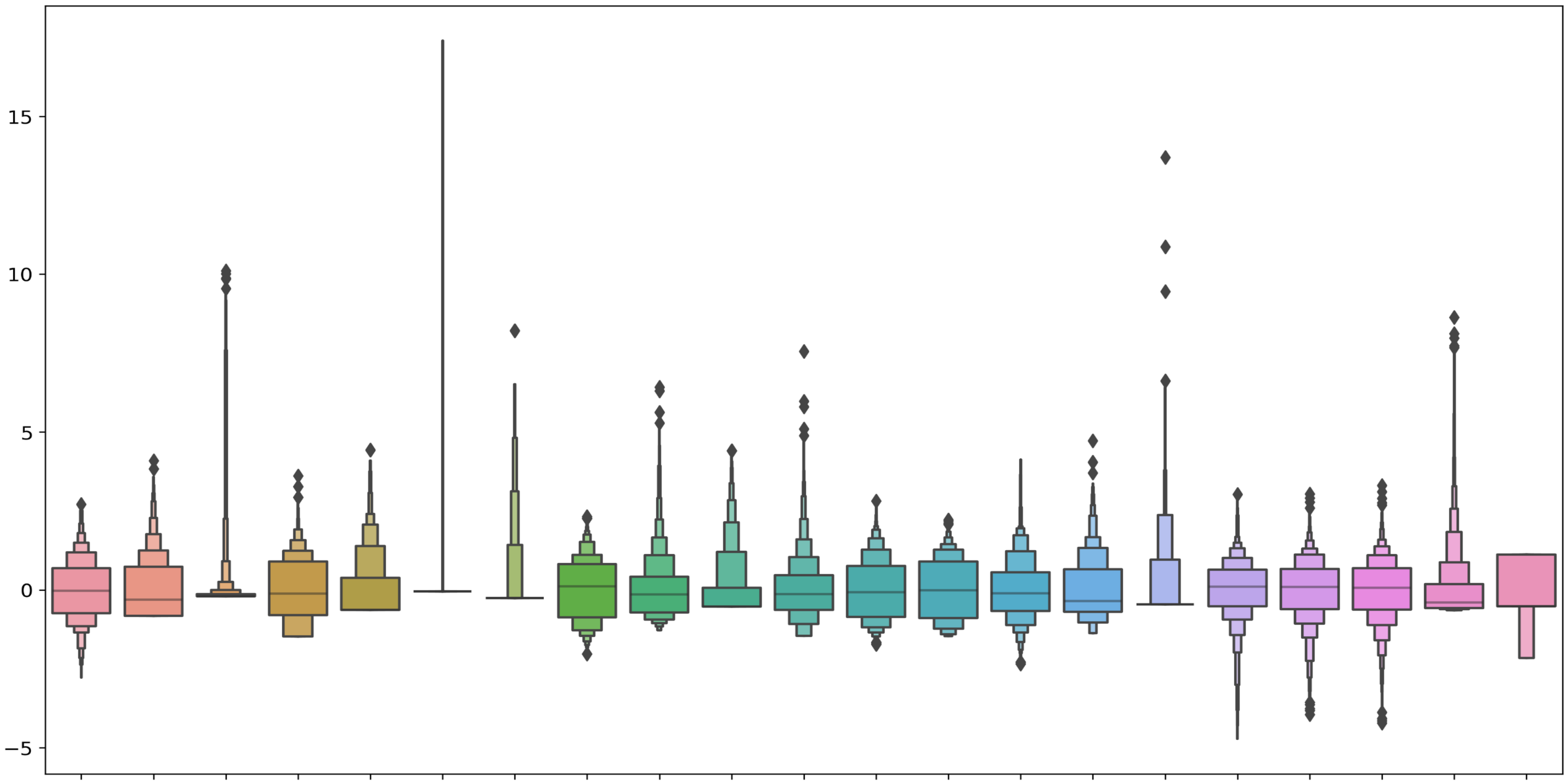


Data preprocessing and model building

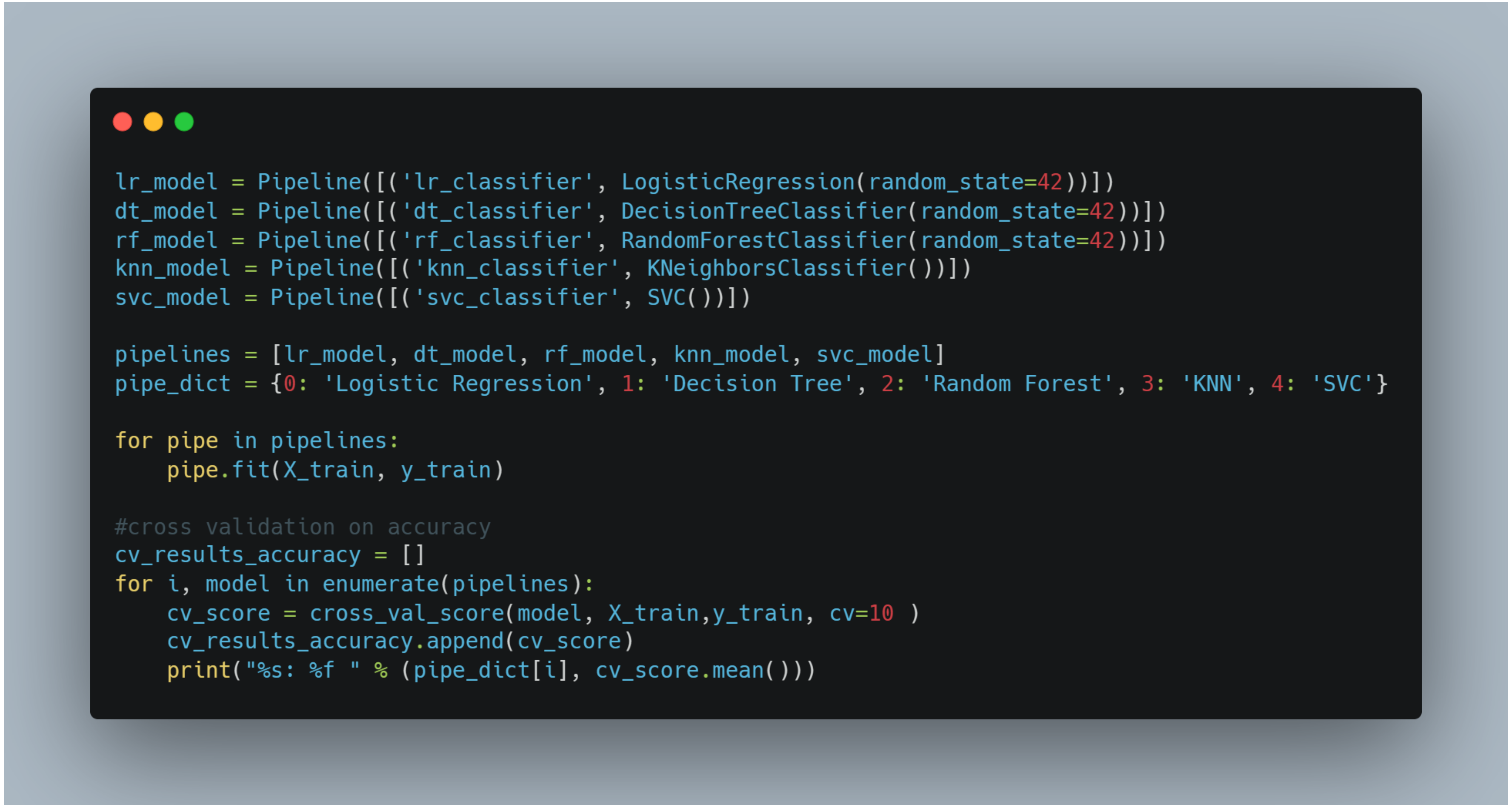
Features in dataset have varied ranges



Feature ranges standardised for model building

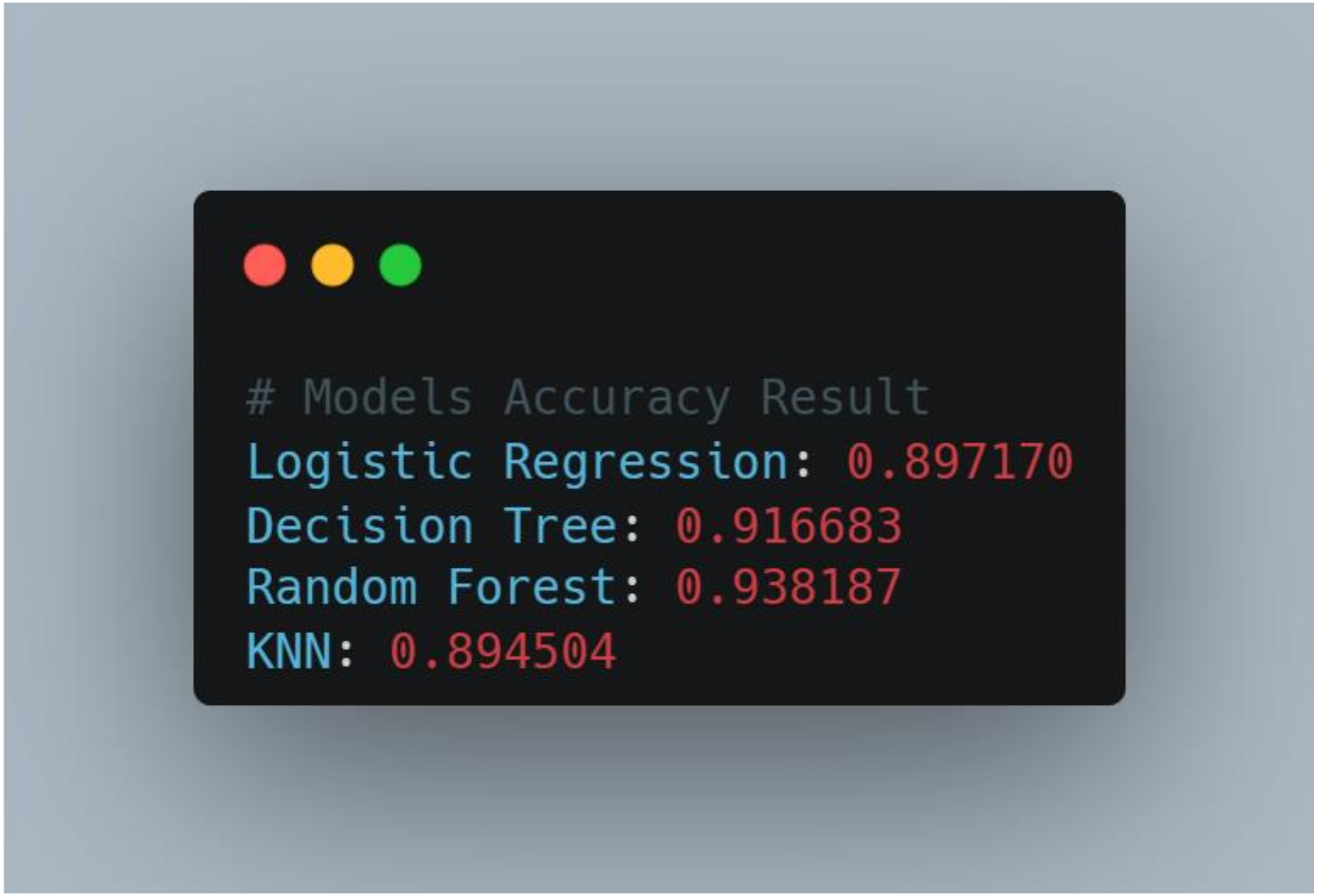


Choosing the best model



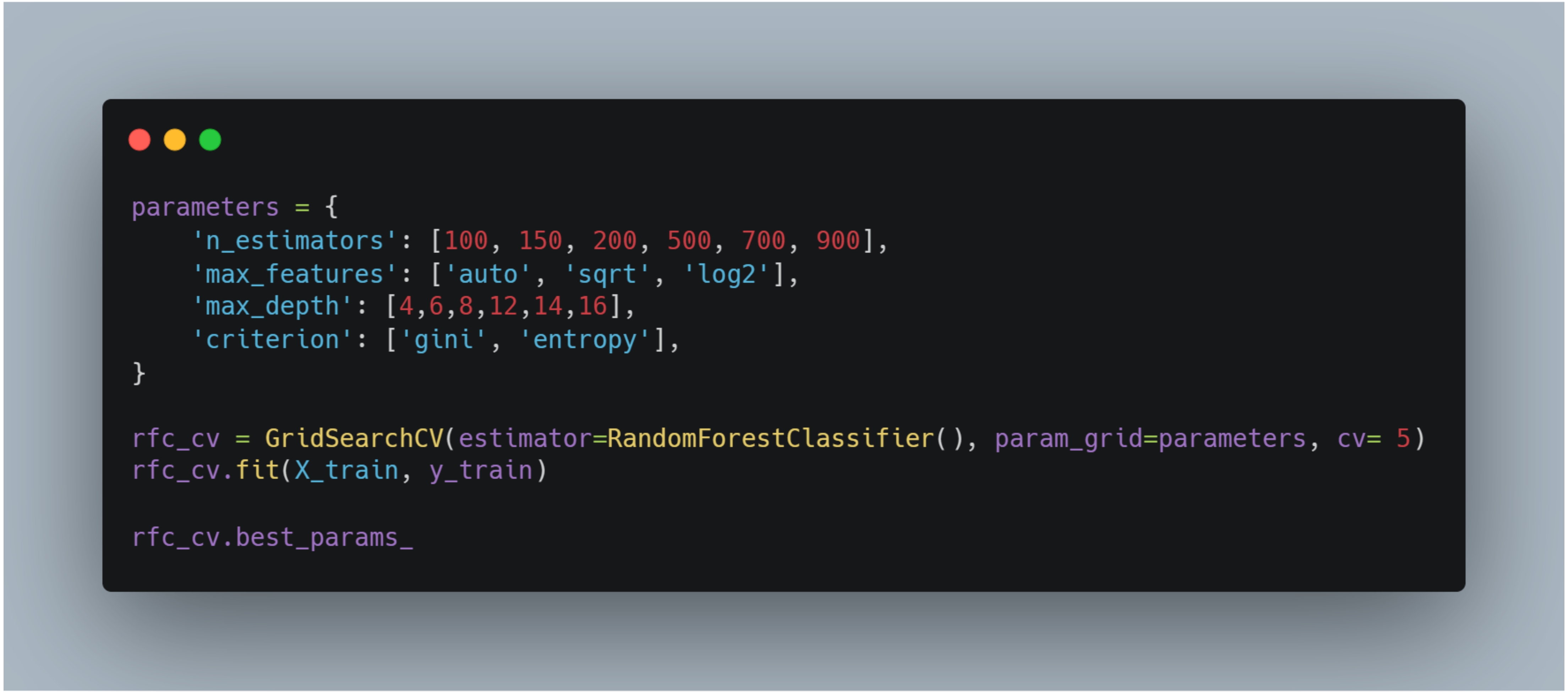
```
lr_model = Pipeline([('lr_classifier', LogisticRegression(random_state=42))])
dt_model = Pipeline([('dt_classifier', DecisionTreeClassifier(random_state=42))])
rf_model = Pipeline([('rf_classifier', RandomForestClassifier(random_state=42))])
knn_model = Pipeline([('knn_classifier', KNeighborsClassifier())])
svc_model = Pipeline([('svc_classifier', SVC())])

pipelines = [lr_model, dt_model, rf_model, knn_model, svc_model]
pipe_dict = {0: 'Logistic Regression', 1: 'Decision Tree', 2: 'Random Forest', 3: 'KNN', 4: 'SVC'}

for pipe in pipelines:
    pipe.fit(X_train, y_train)

#cross validation on accuracy
cv_results_accuracy = []
for i, model in enumerate(pipelines):
    cv_score = cross_val_score(model, X_train,y_train, cv=10 )
    cv_results_accuracy.append(cv_score)
    print("%s: %f " % (pipe_dict[i], cv_score.mean()))
```

Classification models accuracy result



```
# Models Accuracy Result
Logistic Regression: 0.897170
Decision Tree: 0.916683
Random Forest: 0.938187
KNN: 0.894504
```

Hyper-parameter tuning

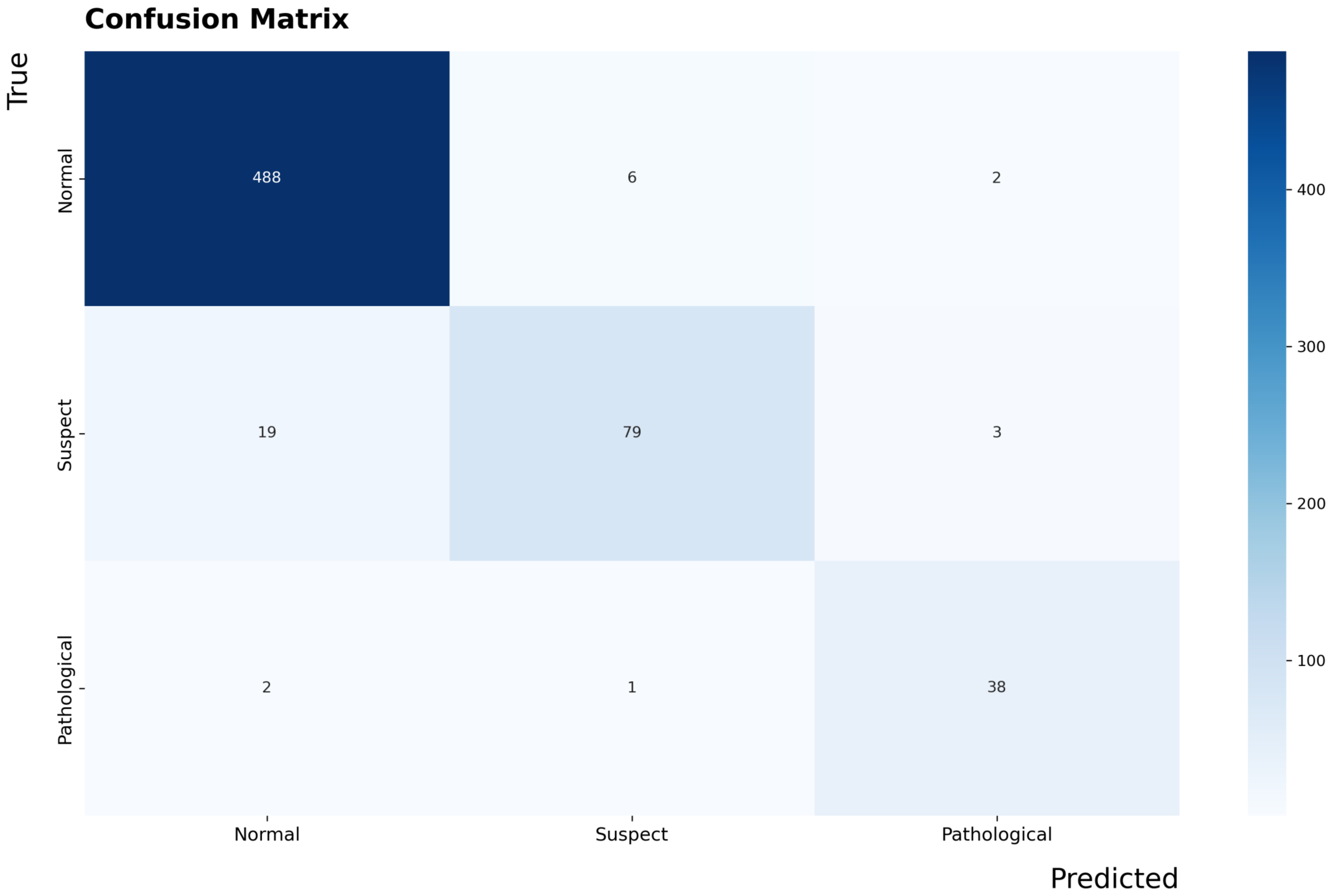


```
parameters = {
    'n_estimators': [100, 150, 200, 500, 700, 900],
    'max_features': ['auto', 'sqrt', 'log2'],
    'max_depth': [4,6,8,12,14,16],
    'criterion': ['gini', 'entropy'],
}

rfc_cv = GridSearchCV(estimator=RandomForestClassifier(), param_grid=parameters, cv= 5)
rfc_cv.fit(X_train, y_train)

rfc_cv.best_params_
```

Results

Classification Metrics

	Precision	Recall	F1-Score	Support
Normal (1.0)	0.96	0.98	0.97	496
Suspect (2.0)	0.92	0.78	0.84	101
Pathological (3.0)	0.88	0.93	0.90	41
Accuracy			0.95	638
Macro Avg	0.92	0.90	0.91	638
Weighted Avg	0.95	0.95	0.95	638


Confusion Matrix
True
Normal
Suspect
Pathological
488
6
2
19
79
3
2
1
38
Normal
Suspect
Pathological
Predicted
400
300
200
100

Conclusion

The best model was found to be the **Random Forest Classifier**. The model is most accurate with the 'Normal' class, with a high number of correct predictions and a relatively small number of false predictions. The model performs reasonably well on the 'Pathological' class. The 'Suspect' class has the most misclassifications relative to its true count, with a notable number of false negatives (predicted as 'Normal').