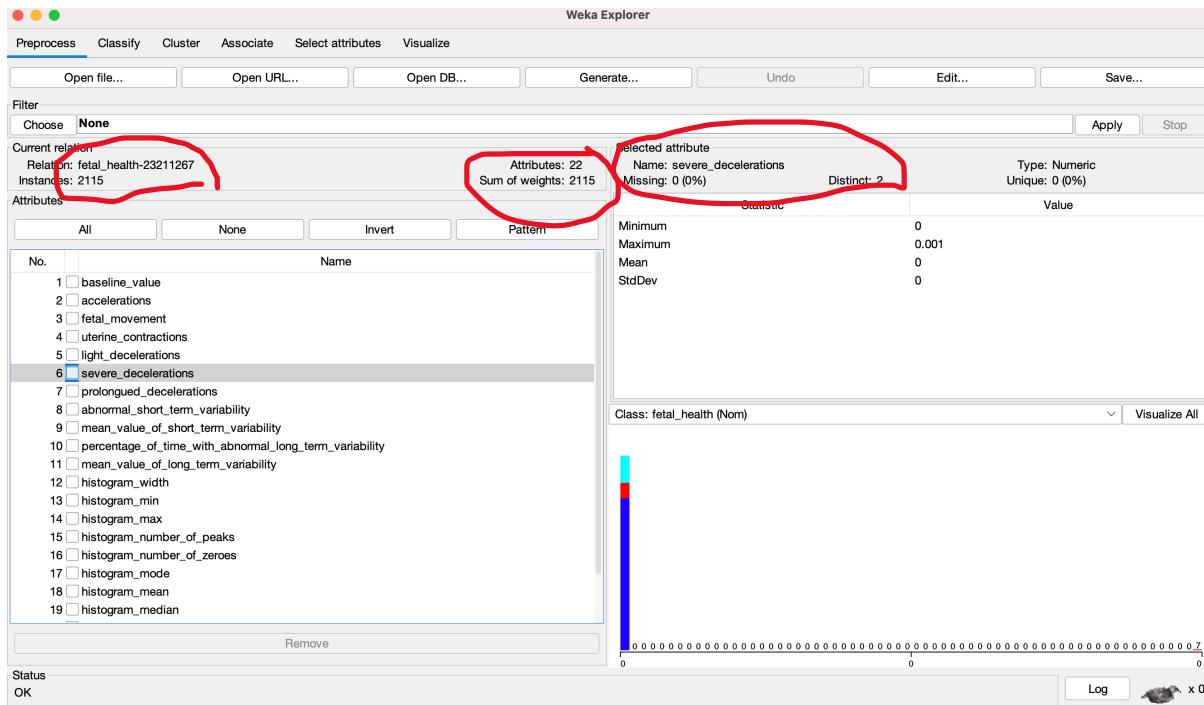


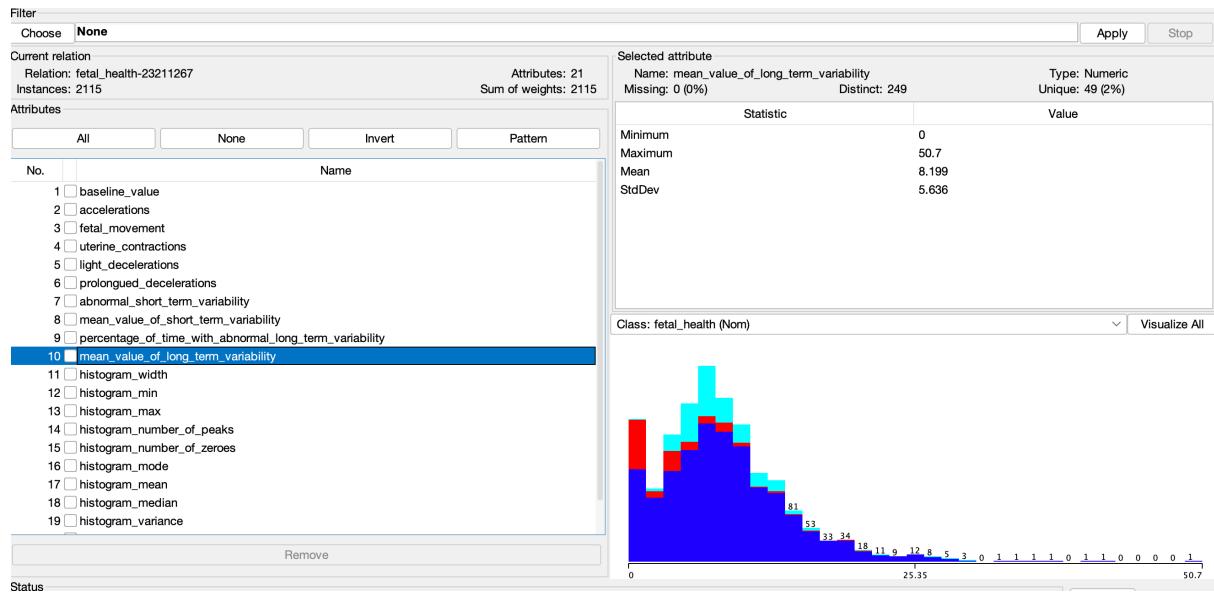
Fetal Health Dataset Analysis

Section 1.1

- Note: Please see uploaded python scripts for incremental output csv files and console output outlined below. In addition, I have taken one of the earlier output files and updated with excel functions to ensure code matches manually calculated results.
- I loaded the dataset into Weka and exported to csv format.
- There were 2115 records or instances loaded from the *.arff file with 22 attributes or features.
- Straight away I could see 'severe_decelerations' was a very limited feature with only 2 distinct values where a non-zero value only occurs 7 times as noted by the python scripts output below (z sum = 7.0 after normalising to range [0, 1]) so I removed this feature from further analysis in Weka and python script runs.
- I think the histogram features would normally be dropped but some of them appear to be quite varied so this makes for a good data sample.



'mean_value_of_long_term_variability' has the most distinct values at 249 which was something to note.



Then I wrote Python code to clean and normalise the data especially the numerical features. I then calculated the z value for each and ordered by z sum as the final program output to see the key features from this process.

Project

- Assignment-1 (~/Library/CloudStorage/GoogleDrive-conor.heffron)
 - plots
 - venv
 - evaluation.py
 - main.py
 - normalise.py
 - out.csv
 - z_sums.csv
 - z_values.csv
 - z_values_analysis.csv
 - z_values_analysis.xlsx
- External Libraries
- Scratches and Consoles

Run main

z_i_light_decelerations	268.2
z_i_percentage_of_time_with_abnormal_long_term_variability	226.1
z_i_histogram_variance	135.7
z_i_histogram_number_of_zeroes	68.2
z_i_prolongued_decelerations	67.4
z_i_fetal_movement	34.6
	dtype: float64

```

30 # Normalise all other numeric features
31 normalise.numerical_features(c, df)
32
33
34 if __name__ == '__main__':
35     # Main run configurations
36     decimals = 3
37     include_histogram_features = True
38     display_plots = False
39     save_plots = True
40     img_dir = 'plots/'
41
42     # Run assignment 1 code
43     assignment_1(include_histogram_features, df)

```

Note: From top left above marked with red:

- plots output directory

- python scripts,
- incremental output files
- manual excel analysis for comparison
- Main configurations for python run script

Main Program Console Out:

```
-----  
Min of baseline_value is: 106  
Max of baseline_value is: 160  
Mean of baseline_value is: 133.30070921985816  
Median of baseline_value is: 133.0  
Standard Deviation of baseline_value is: 9.837149600812499  
Zi sum value baseline_value is: 1074.6  
-----
```

```
Min of accelerations is: 0.0  
Max of accelerations is: 0.019  
Mean of accelerations is: 0.0031787234042553194  
Median of accelerations is: 0.002  
Standard Deviation of accelerations is: 0.003863961682425237  
Zi sum value accelerations is: 367.79999999999995  
-----
```

```
Min of fetal_movement is: 0.0  
Max of fetal_movement is: 0.481  
Mean of fetal_movement is: 0.009526713947990545  
Median of fetal_movement is: 0.0  
Standard Deviation of fetal_movement is: 0.046782681829283676  
Zi sum value fetal_movement is: 34.59999999999994  
-----
```

```
Min of uterine_contractions is: 0.0  
Max of uterine_contractions is: 0.015  
Mean of uterine_contractions is: 0.004373049645390071  
Median of uterine_contractions is: 0.004  
Standard Deviation of uterine_contractions is: 0.002947399462280164  
Zi sum value uterine_contractions is: 616.8  
-----
```

```
Min of light_decelerations is: 0.0  
Max of light_decelerations is: 0.015  
Mean of light_decelerations is: 0.001888888888888889  
Median of light_decelerations is: 0.0  
Standard Deviation of light_decelerations is: 0.0029635828830450834  
Zi sum value light_decelerations is: 268.20000000000005  
-----
```

```
Min of severe_decelerations is: 0.0  
Max of severe_decelerations is: 0.001  
Mean of severe_decelerations is: 3.309692671394799e-06  
Median of severe_decelerations is: 0.0  
Standard Deviation of severe_decelerations is: 5.744822913623554e-05  
Zi sum value severe_decelerations is: 7.0  
-----
```

Min of prolonged_decelerations is: 0.0
Max of prolonged_decelerations is: 0.005
Mean of prolonged_decelerations is: 0.00015933806146572102
Median of prolonged_decelerations is: 0.0
Standard Deviation of prolonged_decelerations is: 0.0005913692853366418
Zi sum value prolonged_decelerations is: 67.4

Min of abnormal_short_term_variability is: 12
Max of abnormal_short_term_variability is: 87
Mean of abnormal_short_term_variability is: 46.95035460992908
Median of abnormal_short_term_variability is: 49.0
Standard Deviation of abnormal_short_term_variability is: 17.162584263361758
Zi sum value abnormal_short_term_variability is: 983.9

Min of mean_value_of_short_term_variability is: 0.2
Max of mean_value_of_short_term_variability is: 7.0
Mean of mean_value_of_short_term_variability is: 1.3339952718676122
Median of mean_value_of_short_term_variability is: 1.2
Standard Deviation of mean_value_of_short_term_variability is: 0.8837077165359826
Zi sum value mean_value_of_short_term_variability is: 345.20000000000005

Min of percentage_of_time_with_abnormal_long_term_variability is: 0
Max of percentage_of_time_with_abnormal_long_term_variability is: 91
Mean of percentage_of_time_with_abnormal_long_term_variability is: 9.823167848699764
Median of percentage_of_time_with_abnormal_long_term_variability is: 0.0
Standard Deviation of percentage_of_time_with_abnormal_long_term_variability is:
18.391154458995572
Zi sum value percentage_of_time_with_abnormal_long_term_variability is: 226.1000000000002

Min of mean_value_of_long_term_variability is: 0.0
Max of mean_value_of_long_term_variability is: 50.7
Mean of mean_value_of_long_term_variability is: 8.19886524822695
Median of mean_value_of_long_term_variability is: 7.4
Standard Deviation of mean_value_of_long_term_variability is: 5.636014802636036
Zi sum value mean_value_of_long_term_variability is: 339.7000000000005

Min of histogram_width is: 3
Max of histogram_width is: 180
Mean of histogram_width is: 70.49267139479906
Median of histogram_width is: 68.0
Standard Deviation of histogram_width is: 38.95078910381832
Zi sum value histogram_width is: 807.9000000000001

Min of histogram_min is: 50
Max of histogram_min is: 159
Mean of histogram_min is: 93.54704491725768
Median of histogram_min is: 93.0
Standard Deviation of histogram_min is: 29.580143601297536
Zi sum value histogram_min is: 844.7

Min of histogram_max is: 122

```
Max of histogram_max is: 238
Mean of histogram_max is: 164.03971631205673
Median of histogram_max is: 162.0
Standard Deviation of histogram_max is: 17.944182664569215
Zi sum value histogram_max is: 768.3
-----
Min of histogram_number_of_peaks is: 0
Max of histogram_number_of_peaks is: 18
Mean of histogram_number_of_peaks is: 4.069976359338061
Median of histogram_number_of_peaks is: 3.0
Standard Deviation of histogram_number_of_peaks is: 2.947638560550025
Zi sum value histogram_number_of_peaks is: 492.3
-----
Min of histogram_number_of_zeroes is: 0
Max of histogram_number_of_zeroes is: 10
Mean of histogram_number_of_zeroes is: 0.3224586288416076
Median of histogram_number_of_zeroes is: 0.0
Standard Deviation of histogram_number_of_zeroes is: 0.7041179104063311
Zi sum value histogram_number_of_zeroes is: 68.20000000000002
-----
Min of histogram_mode is: 60
Max of histogram_mode is: 187
Mean of histogram_mode is: 137.44586288416076
Median of histogram_mode is: 139.0
Standard Deviation of histogram_mode is: 16.403044334341104
Zi sum value histogram_mode is: 1289.5
-----
Min of histogram_mean is: 73
Max of histogram_mean is: 182
Mean of histogram_mean is: 134.60614657210402
Median of histogram_mean is: 136.0
Standard Deviation of histogram_mean is: 15.617939011476754
Zi sum value histogram_mean is: 1198.4999999999998
-----
Min of histogram_median is: 77
Max of histogram_median is: 186
Mean of histogram_median is: 138.09078014184396
Median of histogram_median is: 139.0
Standard Deviation of histogram_median is: 14.48426821225838
Zi sum value histogram_median is: 1184.8
-----
Min of histogram_variance is: 0
Max of histogram_variance is: 269
Mean of histogram_variance is: 18.846808510638297
Median of histogram_variance is: 7.0
Standard Deviation of histogram_variance is: 29.03532873969992
Zi sum value histogram_variance is: 135.7
-----
Min of histogram_tendency is: -1
Max of histogram_tendency is: 1
Mean of histogram_tendency is: 0.32056737588652484
```

```

Median of histogram_tendency is: 0.0
Standard Deviation of histogram_tendency is: 0.6107888873223769
Zi sum value histogram_tendency is: 1396.5

```

z_i_fetal_health	1499.7
z_i_histogram_tendency	1396.5
z_i_histogram_mode	1289.5
z_i_histogram_mean	1198.5
z_i_histogram_median	1184.8
z_i_baseline_value	1074.6
z_i_abnormal_short_term_variability	983.9
z_i_histogram_min	844.7
z_i_histogram_width	807.9
z_i_histogram_max	768.3
z_i_uterine_contractions	616.8
z_i_histogram_number_of_peaks	492.3
z_i_accelerations	367.8
z_i_mean_value_of_short_term_variability	345.2
z_i_mean_value_of_long_term_variability	339.7
z_i_light_decelerations	268.2
z_i_percentage_of_time_with_abnormal_long_term_variability	226.1
z_i_histogram_variance	135.7
z_i_histogram_number_of_zeroes	68.2
z_i_prolongued_decelerations	67.4
z_i_fetal_movement	34.6
z_i_severe_decelerations	7.0

dtype: float64

Process finished with exit code 0

-
- ‘fetal_health’ value is the categorical feature that was originally used to classify the data so I have excluded this as a key feature but more like a classifier field.
 - The 7 key features thereafter and info are highlighted in yellow.
 - The next 3 fields highlighted in orange were non histogram related fields.
 - The features with z sum less than 40 are marked in blue. I removed ‘severe_decelerations’ altogether as I don’t think it’s even close to the threshold to be very useful. However out of the seven non zero values there were 6 records with the same fetal health classifier assigned but this is probably coincidence as the data sample is too small.

Output Verified in Excel Worksheet after manually manipulating incremental csv output files:

A	B	H	M	Q	R	S	U	V	W
fetal_health	z_i_baseline_value	z_i_abnormal_short_term_variability	z_i_histogram_min	z_i_histogram_mode	z_i_histogram_mean	z_i_histogram_median	z_i_histogram_tendency	z_i_fetal_health	
Normal	0.5	0.2	0.1	0.7	0.6	0.6	1	0.8	
Normal	0.4	0.3	0	0.6	0.6	0.5	0.5	0.8	
Suspect	0.6	0.6	0	0.7	0.7	0.7	1	0.5	
Suspect	0.4	0.7	0.7	0.6	0.5	0.5	0.5	0.5	
Suspect	1	0.6	0.9	0.8	0.9	0.9	0.5	0.5	
Normal	0.5	0.6	0.6	0.6	0.7	0.6	0.5	0.8	
Suspect	0.7	0.7	0.7	0.7	0.7	0.7	0.5	0.5	
Normal	0.6	0.6	0.4	0.8	0.6	0.6	1	0.8	
Normal	0.6	0.4	0.5	0.6	0.6	0.6	1	0.8	
Normal	0.3	0.1	0.5	0.5	0.5	0.5	0.5	0.8	
Normal	0.1	0.7	0.4	0.4	0.4	0.3	0	0.8	
Normal	0.6	0.5	0.3	0.7	0.7	0.7	1	0.8	
Pathological	0.4	0.7	0.7	0.6	0.5	0.5	0.5	0.2	
Normal	0.3	0.6	0.2	0.5	0.4	0.5	0.5	0.8	
Normal	0.6	0.2	0.2	0.6	0.6	0.6	1	0.8	
Normal	0.7	0.4	0.8	0.7	0.7	0.6	0.5	0.8	
Suspect	0.8	0.8	0.8	0.7	0.7	0.7	0	0.5	
Suspect	0.6	0.7	0.2	0.7	0.6	0.6	1	0.5	
Normal	0.3	0.3	0.3	0.5	0.4	0.4	1	0.8	
Normal	0.4	0.2	0.1	0.6	0.5	0.5	0.5	0.8	
Normal	0.5	0.3	0.5	0.6	0.6	0.6	0.5	0.8	
Normal	0.7	0.3	0.4	0.7	0.7	0.7	1	0.8	
Normal	0.4	0.5	0.1	0.6	0.5	0.5	1	0.8	
Normal	0.4	0.7	0.1	0.5	0.3	0.5	0.5	0.8	
Normal	0.6	0.5	0.2	0.7	0.7	0.6	1	0.8	
Normal	0.7	0.6	0.1	0.7	0.6	0.6	1	0.8	
Pathological	0.3	0.6	0.2	0.2	0.3	0.3	0.5	0.2	
Suspect	0.4	0.7	0.2	0.7	0.6	0.6	1	0.5	
Normal	0.8	0.5	0.9	0.8	0.8	0.8	1	0.8	
	1074.6	983.9	844.7	1289.5	1198.5	1184.8	1396.5	1499.7	

Section 1.2

- Naïve Bayes is least accurate.
- KNN and decision tree are close in accuracy (over 95% so good options).
- Running with k=1, then k=3 and k=5 improved accuracy each time when updating to use 1/distance weighting.

Naive Bayes

Classifier
Choose J48 -C 0.25 -M 2

Test options
 Use training set
 Supplied test set Set...
 Cross-validation Folds 10
 Percentage split % 66
[More options...](#)

Classifier output
 std. dev. 0.5896
 weight sum 1648
 precision 1 1 1
 0.6866 0.5899

Time taken to build model: 0.01 seconds

==== Evaluation on training set ====
 Time taken to test model on training data: 0.02 seconds

==== Summary ====
 Correctly Classified Instances 1740 82.2695 %
 Incorrectly Classified Instances 375 17.7305 %

==== Detailed Accuracy By Class ====

TP Rate	FP Rate	Precision	Recall	F-Measure	MCC	ROC Area	PRC Area	Class
0.836	0.064	0.979	0.836	0.902	0.679	0.947	0.985	Normal
0.638	0.040	0.587	0.638	0.612	0.576	0.888	0.568	Pathological
0.857	0.147	0.485	0.857	0.619	0.570	0.908	0.601	Suspect
Weighted Avg.	0.823	0.074	0.878	0.823	0.839	0.655	0.937	0.897

==== Confusion Matrix ====
 a b c <-- classified as
 1378 61 209 | a = Normal
 5 111 58 | b = Pathological
 25 17 251 | c = Suspect

Decision Tree

Classifier J48 -C 0.25 -M 2

Test options

- Use training set
- Supplied test set Set...
- Cross-validation Folds 10
- Percentage split % 66

[More options...](#)

(Nom) fetal_health

Start Stop

Result list (right-click for options)

- 23:30:32 - lazy.IBk
- 23:30:46 - lazy.IBk
- 23:31:04 - lazy.IBk
- 18:32:27 - lazy.IBk
- 18:33:01 - lazy.IBk
- 18:33:31 - lazy.IBk
- 18:33:39 - lazy.IBk
- 18:33:54 - lazy.IBk
- 18:34:18 - lazy.IBk
- 18:35:00 - bayes.NaiveBayes
- 18:35:18 - trees.J48**

Classifier output

```

Number of Leaves : 53
Size of the tree : 105
Time taken to build model: 0.05 seconds
==== Evaluation on training set ====
Time taken to test model on training data: 0 seconds
==== Summary ====
Correctly Classified Instances      2066          97.6832 %
Incorrectly Classified Instances   49           2.3168 %
Kappa statistic                   0.936
Mean absolute error               0.0273
Root mean squared error          0.1168
Relative absolute error           11.1509 %
Root relative squared error     33.4123 %
Total Number of Instances        2115

==== Detailed Accuracy By Class ====


|                      | TP           | Rate         | FP           | Rate         | Precision    | Recall       | F-Measure    | MCC          | ROC Area     | PRC Area     | Class        |
|----------------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|
| 0.992                | 0.064        | 0.982        | 0.992        | 0.987        | 0.939        | 0.990        | 0.995        | 0.995        | 0.972        | 0.972        | Normal       |
| 0.937                | 0.001        | 0.988        | 0.937        | 0.962        | 0.959        | 0.995        | 0.995        | 0.995        | 0.972        | 0.972        | Pathological |
| 0.918                | 0.009        | 0.941        | 0.918        | 0.929        | 0.918        | 0.988        | 0.946        | 0.946        | 0.988        | 0.988        | Suspect      |
| <b>Weighted Avg.</b> | <b>0.977</b> | <b>0.051</b> | <b>0.977</b> | <b>0.977</b> | <b>0.977</b> | <b>0.977</b> | <b>0.977</b> | <b>0.938</b> | <b>0.990</b> | <b>0.987</b> |              |


==== Confusion Matrix ====


|      |     |     |                   |
|------|-----|-----|-------------------|
| a    | b   | c   | <-- classified as |
| 1634 | 1   | 13  | a = Normal        |
| 7    | 163 | 4   | b = Pathological  |
| 23   | 1   | 269 | c = Suspect       |


```

k-NN

Classifier J48 -C 0.25 -M 2

Test options

- Use training set
- Supplied test set Set...
- Cross-validation Folds 10
- Percentage split % 66

[More options...](#)

(Nom) fetal_health

Start Stop

Result list (right-click for options)

- 23:30:32 - lazy.IBk
- 23:30:46 - lazy.IBk
- 23:31:04 - lazy.IBk
- 18:32:27 - lazy.IBk
- 18:33:01 - lazy.IBk
- 18:33:31 - lazy.IBk
- 18:33:39 - lazy.IBk
- 18:33:54 - lazy.IBk
- 18:34:18 - lazy.IBk**
- 18:35:00 - bayes.NaiveBayes
- 18:35:18 - trees.J48

Classifier output

```

==== Classifier model (full training set) ====
IB1 instance-based classifier
using 5 inverse-distance-weighted nearest neighbour(s) for classification

Time taken to build model: 0 seconds
==== Evaluation on training set ====
Time taken to test model on training data: 0.38 seconds
==== Summary ====
Correctly Classified Instances      2113          99.9054 %
Incorrectly Classified Instances   2           0.0946 %
Kappa statistic                   0.9974
Mean absolute error               0.0058
Root mean squared error          0.0254
Relative absolute error           2.3695 %
Root relative squared error     7.2677 %
Total Number of Instances        2115

==== Detailed Accuracy By Class ====


|                      | TP           | Rate         | FP           | Rate         | Precision    | Recall       | F-Measure    | MCC          | ROC Area     | PRC Area     | Class        |
|----------------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|
| 0.999                | 0.002        | 0.999        | 0.999        | 0.999        | 0.997        | 1.000        | 1.000        | 1.000        | 1.000        | 1.000        | Normal       |
| 1.000                | 0.000        | 1.000        | 1.000        | 1.000        | 1.000        | 1.000        | 1.000        | 1.000        | 1.000        | 1.000        | Pathological |
| 0.997                | 0.001        | 0.997        | 0.997        | 0.997        | 0.997        | 0.996        | 1.000        | 1.000        | 1.000        | 1.000        | Suspect      |
| <b>Weighted Avg.</b> | <b>0.999</b> | <b>0.002</b> | <b>0.999</b> | <b>0.999</b> | <b>0.999</b> | <b>0.999</b> | <b>0.999</b> | <b>0.997</b> | <b>1.000</b> | <b>1.000</b> |              |

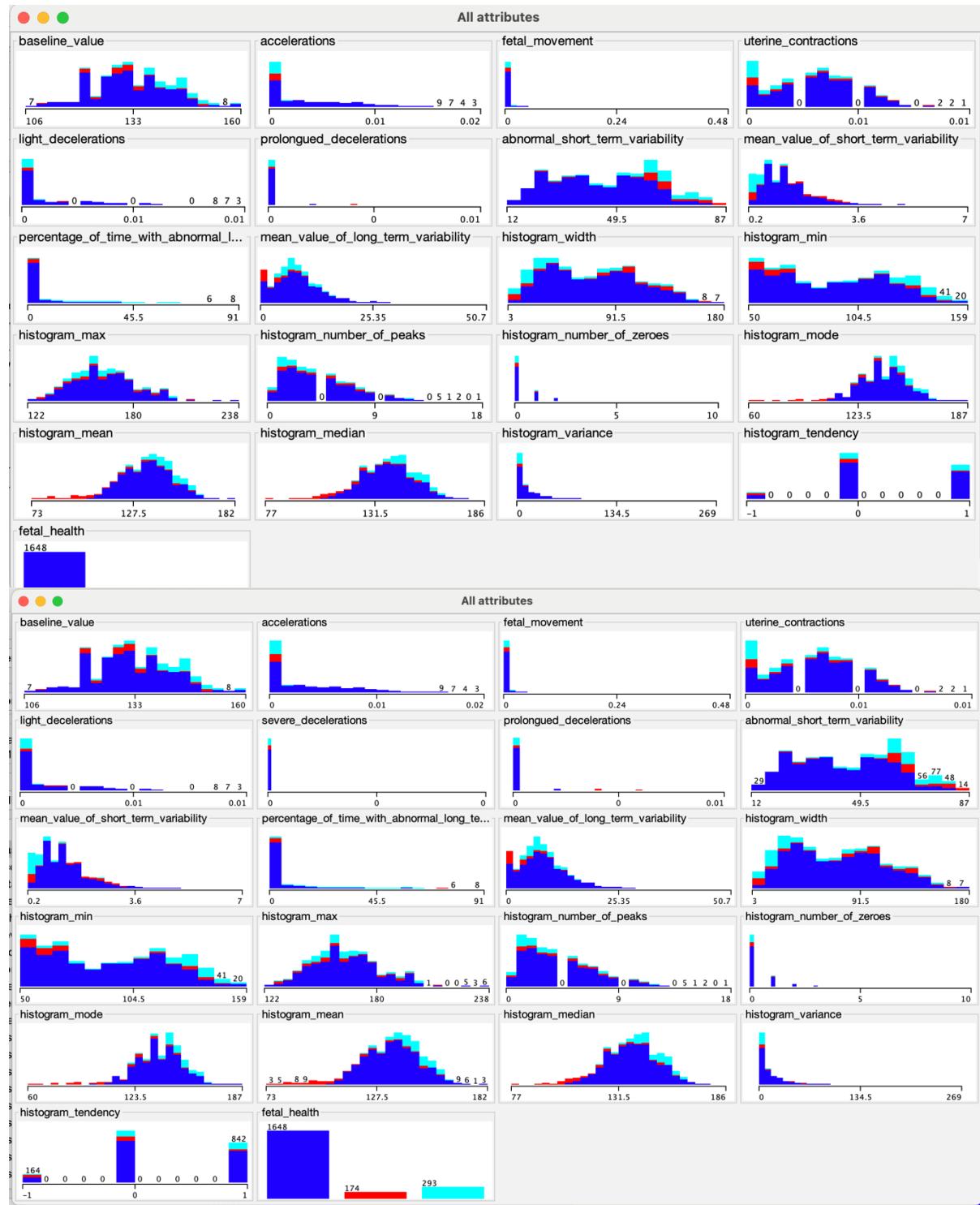

==== Confusion Matrix ====


|      |     |     |                   |
|------|-----|-----|-------------------|
| a    | b   | c   | <-- classified as |
| 1647 | 0   | 1   | a = Normal        |
| 0    | 174 | 0   | b = Pathological  |
| 1    | 0   | 292 | c = Suspect       |

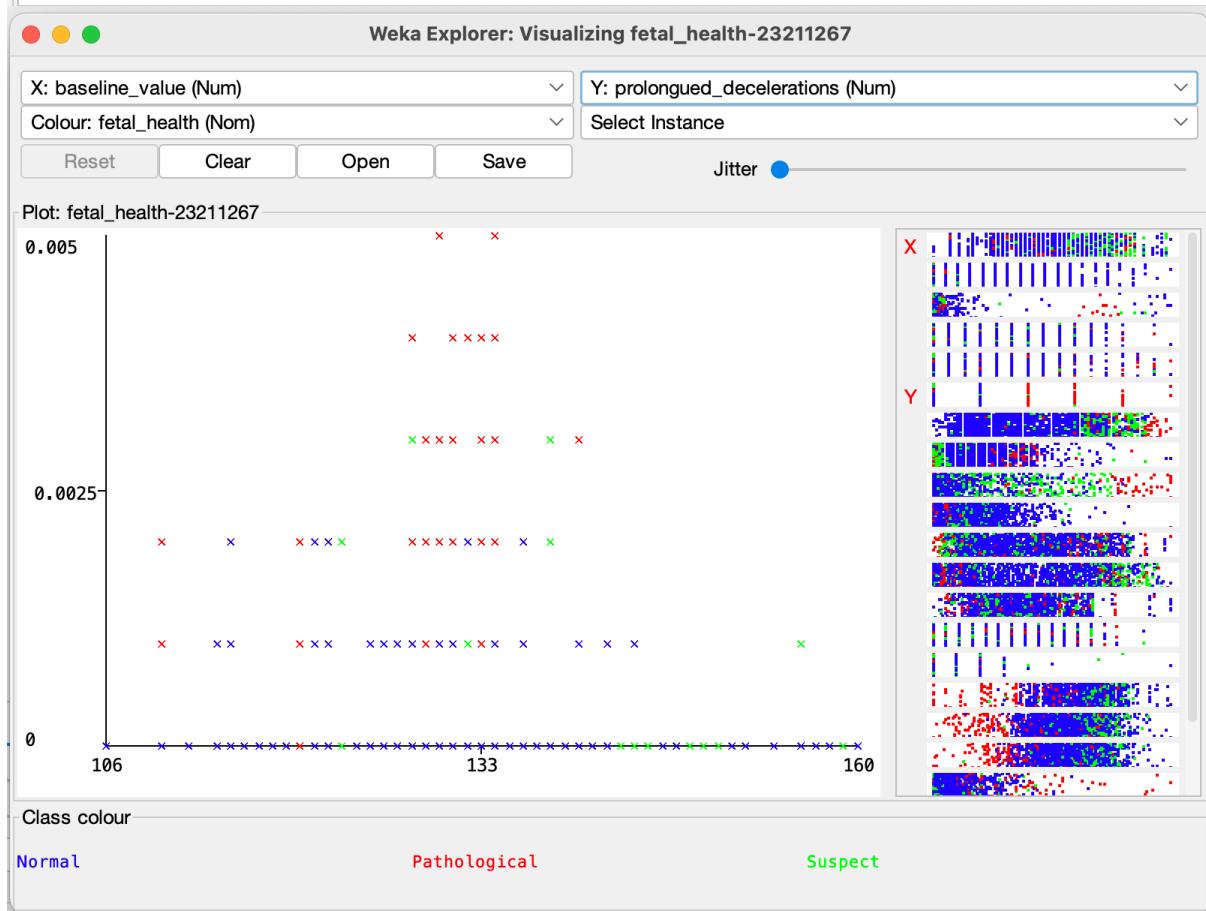

```

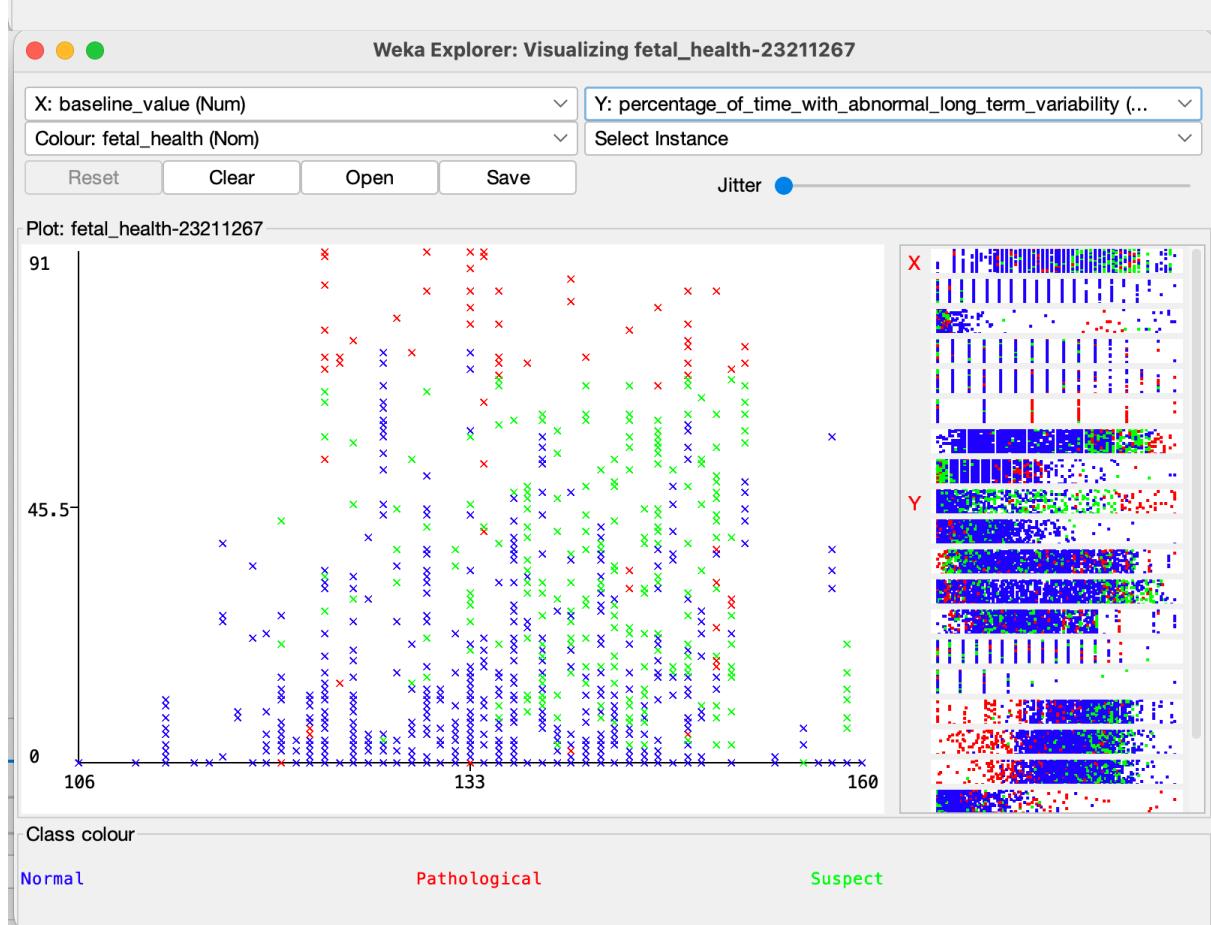
Section 1.3.

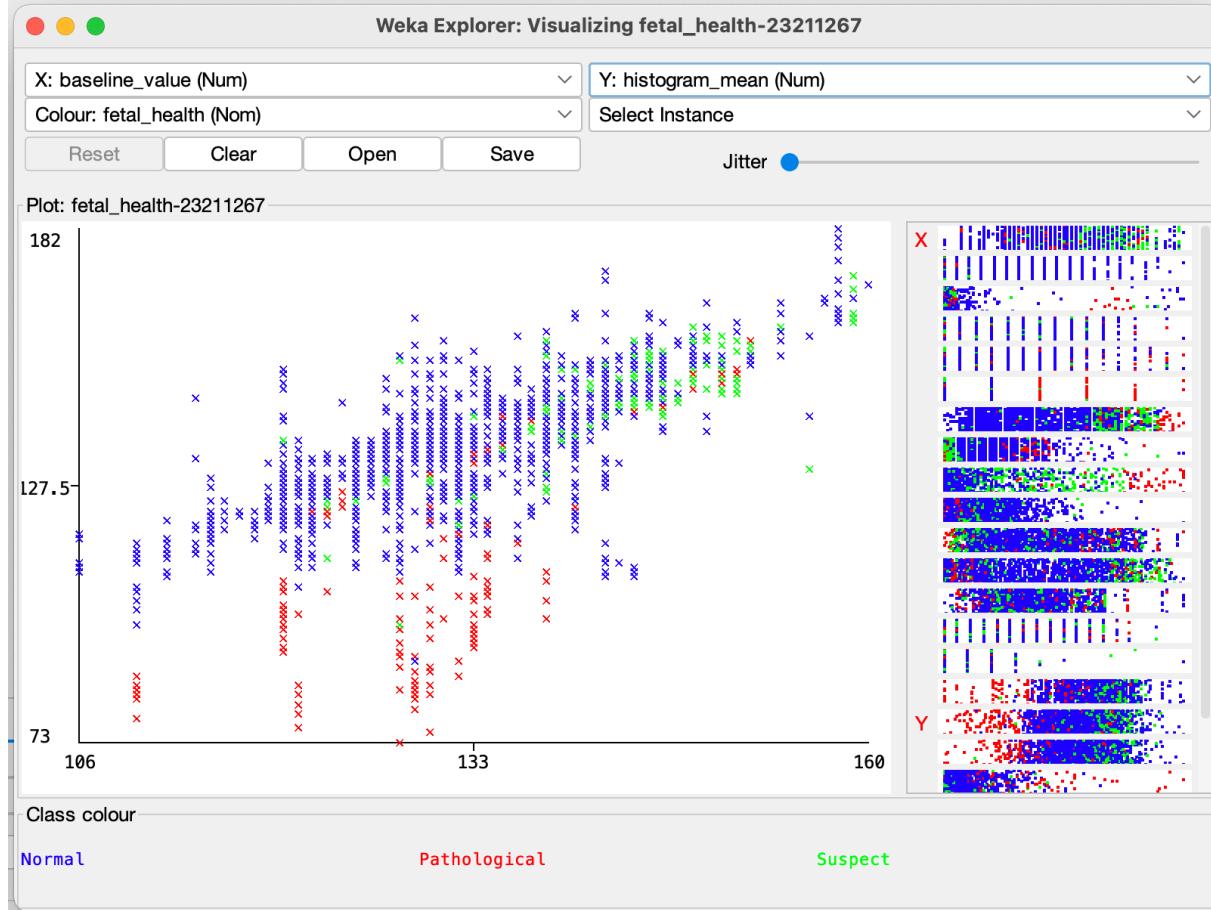
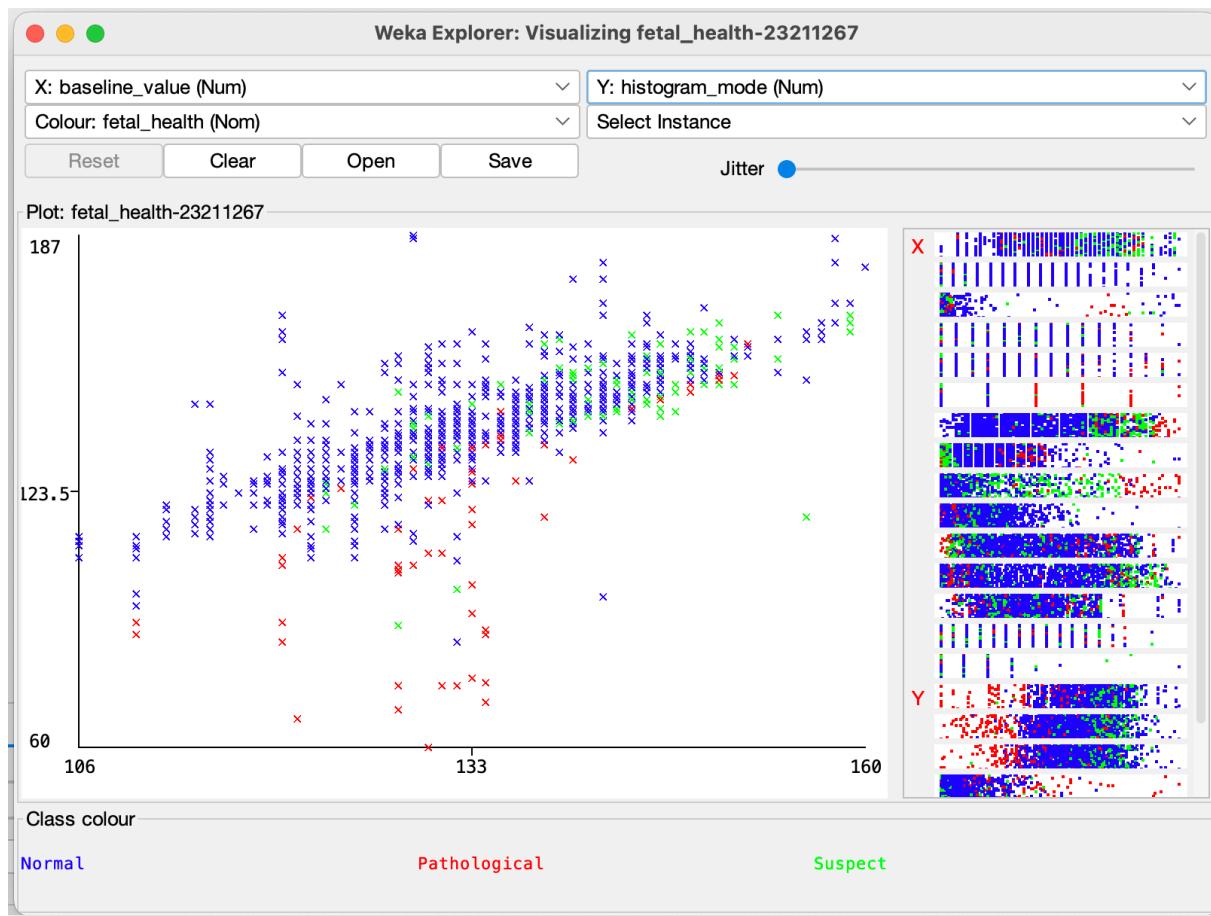
Weka Plots and visualisation notes:

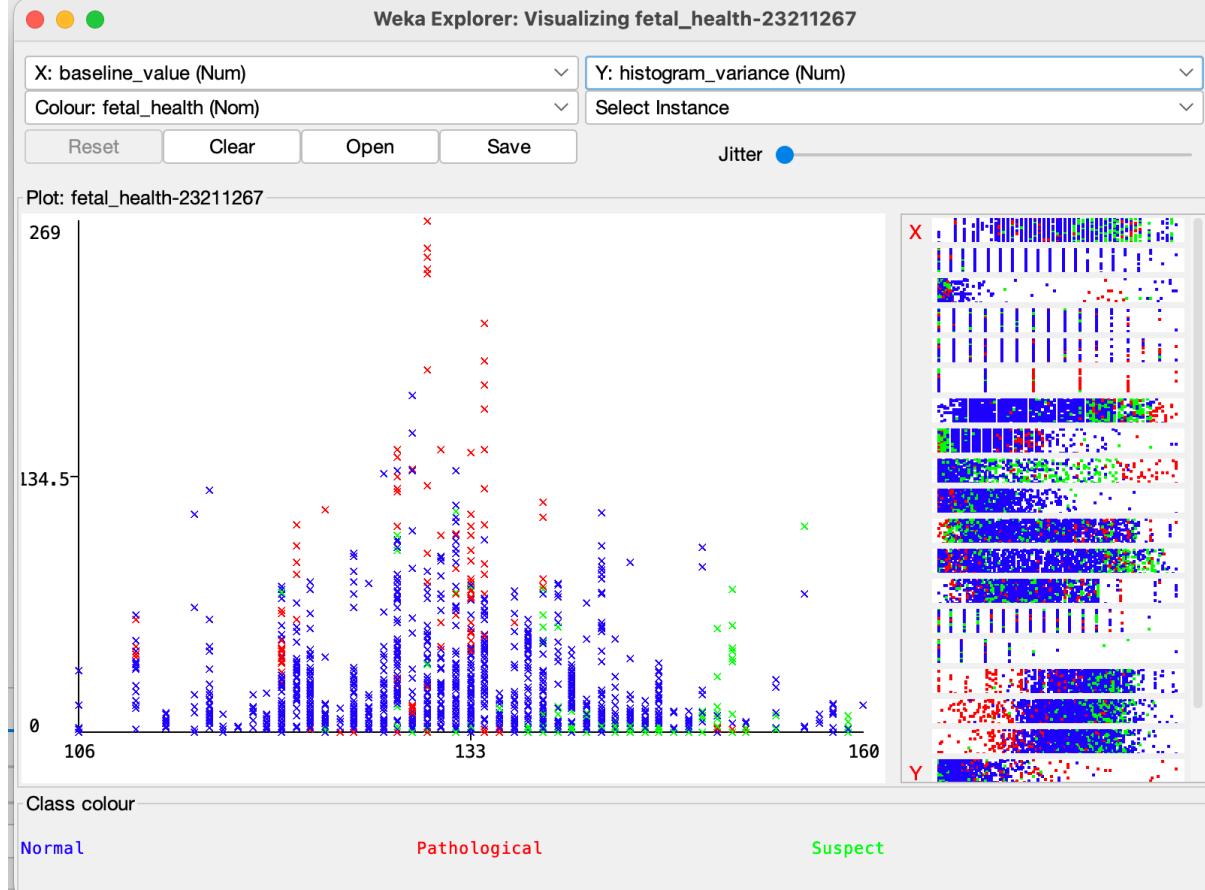
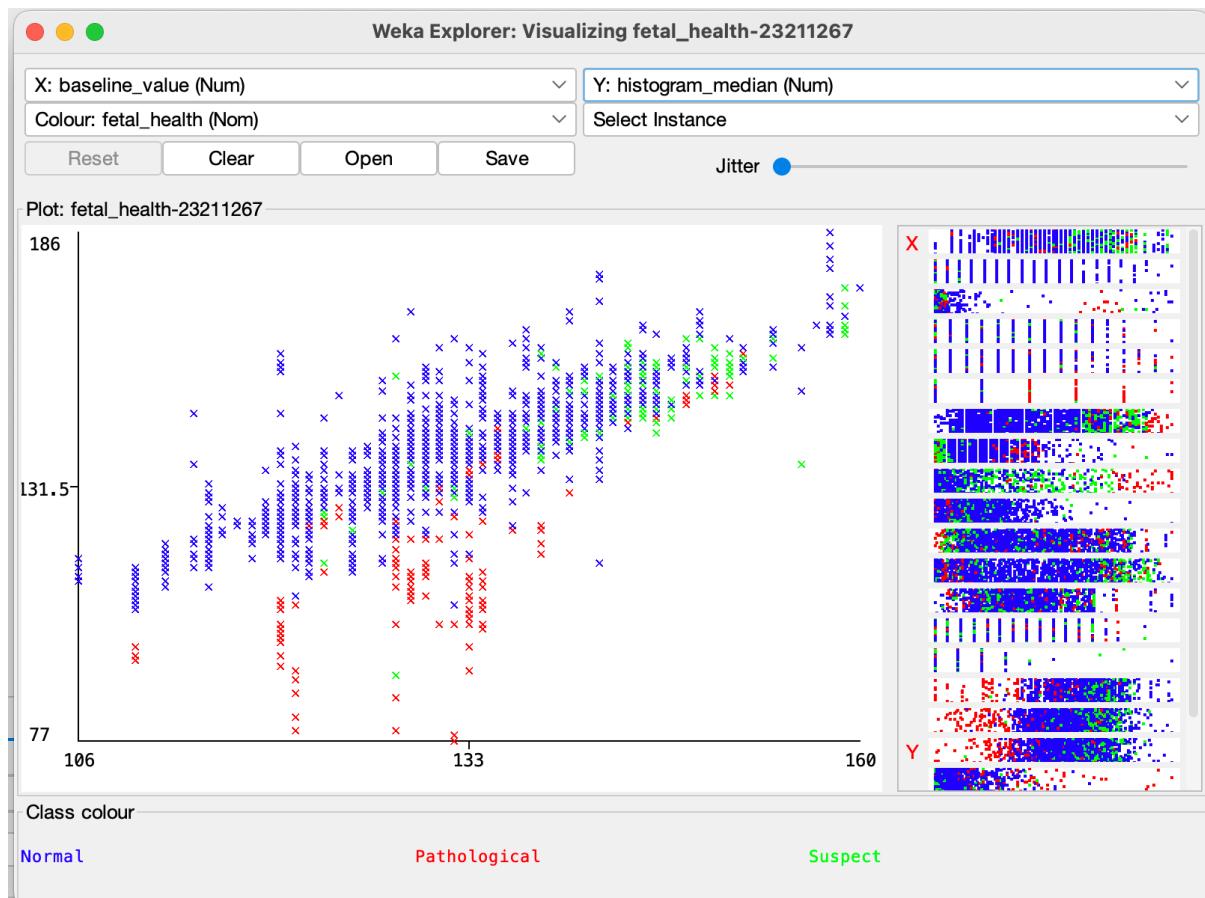


Selected attribute		Type: Numeric Unique: 0 (0%)
Name: severe_decelerations	Distinct: 2	
Missing: 0 (0%)		
	Statistic	Value
Minimum		0
Maximum		0.001
Mean		0
StdDev		0









Section 1.4.

Top 14 features after normalisation of numerical features (top 6 non histogram features in bold / red)

1. histogram_tendency
2. histogram_mode
3. histogram_mean
4. histogram_median
- 5. baseline_value**
- 6. abnormal_short_term_variability**
7. histogram_min
8. histogram_width
9. histogram_max
- 10. uterine_contractions**
11. histogram_number_of_peaks
- 12. accelerations**
- 13. mean_value_of_short_term_variability**
- 14. mean_value_of_long_term_variability**

1.5.

Use the following methods in Weka to find the top 5 features:

- Wrapper + forward search

The screenshot shows the Weka Attribute Evaluator interface. The 'Choose' dropdown under 'Attribute Evaluator' is set to 'CfsSubsetEval -P 1 -E 1'. The 'Search Method' dropdown is set to 'BestFirst -D 1 -N 5'. Under 'Attribute Selection Mode', the 'Use full training set' radio button is selected. The 'Result list' pane shows several evaluation runs, with '20:04:28 - BestFirst + CfsSubsetEval' highlighted. The 'Attribute selection output' pane displays the following text:

```
Attribute selection output
histogram_min
histogram_max
histogram_number_of_peaks
histogram_number_of_zeroes
histogram_mode
histogram_mean
histogram_median
histogram_variance
histogram_tendency
fetal_health

Evaluation mode: evaluate on all training data

== Attribute Selection on all input data ==

Search Method:
  Best first.
  Start set: no attributes
  Search direction: forward
  Stale search after 5 node expansions
  Total number of subsets evaluated: 169
  Merit of best subset found: 0.349

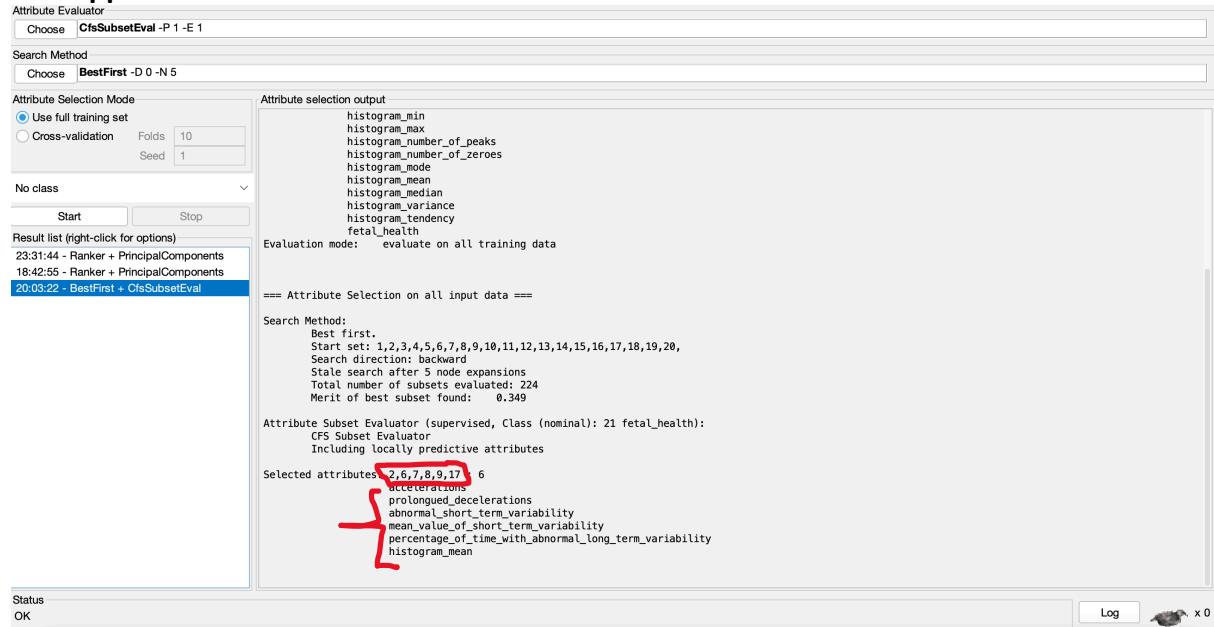
Attribute Subset Evaluator (supervised, Class (nominal): 21 fetal_health):
  CFS Subset Evaluator
  Including locally predictive attributes

Selected attributes: 2,6,7,8,9,17 : 6
  accelerations
  prolonged_decelerations
  abnormal_short_term_variability
  mean_value_of_short_term_variability
  percentage_of_time_with_abnormal_long_term_variability
  histogram_mean
```

A red curly brace is drawn around the last five selected attributes: accelerations, prolonged_decelerations, abnormal_short_term_variability, mean_value_of_short_term_variability, and percentage_of_time_with_abnormal_long_term_variability.

1. accelerations
2. prolonged_decelerations
3. abnormal_short_term_variability
4. mean_value_of_short_term_variability
5. percentage_of_time_with_abnormal_long_term_variability
6. histogram_mean

- **Wrapper + backwards search**



1. accelerations
2. prolonged_decelerations
3. abnormal_short_term_variability
4. mean_value_of_short_term_variability
5. percentage_of_time_with_abnormal_long_term_variability
6. histogram_mean

- **Information Gain** Discuss the differences in the selected sets of features.

Attribute Evaluator

Choose **InfoGainAttributeEval**

Search Method

Choose **Ranker -T -1.7976931348623157E308 -N -1**

Attribute Selection Mode

Use full training set

Cross-validation Folds 10
Seed 1

No class

Start Stop

Result list (right-click for options)

```
23:31:44 - Ranker + PrincipalComponents
18:42:55 - Ranker + PrincipalComponents
20:03:22 - BestFirst + CfsSubsetEval
20:04:28 - BestFirst + CfsSubsetEval
20:05:15 - Ranker + InfoGainAttributeEval
```

Attribute selection output

Evaluation mode: evaluate on all training data

==== Attribute Selection on all input data ====

Search Method:
Attribute ranking.

Attribute Evaluator (supervised, Class (nominal): 21 fetal_health):
Information Gain Ranking Filter

Ranked attributes:

```
0.3022    8 mean_value_of_short_term_variability
0.2674    9 percentage_of_time_with_abnormal_long_term_variability
0.2598    7 abnormal_short_term_variability
0.2172    17 histogram_mean
0.2094    19 histogram_variance
0.199     16 histogram_mode
0.1979    2 accelerations
0.1829    18 histogram_median
0.147     1 baseline_value
0.1362    10 mean_value_of_long_term_variability
0.1358    11 histogram_width
0.1287    12 histogram_min
0.1261    6 prolonged_decelerations
0.0844    4 uterine_contractions
0.0697    3 fetal_movement
0.0528    5 light_decelerations
0.0308    20 histogram_tendency
0.0286    13 histogram_max
0.0232    14 histogram_number_of_peaks
0         15 histogram_number_of_zeroes
```

Selected attributes: 8,9,7,17,19 16,2,18,1,10,11,12,6,4,3,5,20,13,14,15 : 20

1. mean_value_of_short_term_variability
2. percentage_of_time_with_abnormal_long_term_variability
3. abnormal_short_term_variability
4. histogram_mean
5. histogram_variance
6. histogram_mode

Section 1.6.

Wrapper search forward and backward produce the same top 5 fields.

IG produces are different and more accurate top 5 feature list for what I have compared to up to this point.

Section 1.7.

Attribute Evaluator

Choose **PrincipalComponents -R 0.95 -A 5**

Search Method

Choose **Ranker -T -1.7976931348623157E308 -N -1**

Attribute Selection Mode

Use full training set

Cross-validation Folds 10 Seed 1

No class

Start Stop

Result list (right-click for options)

23:31:44 - Ranker + PrincipalComponents
18:42:55 - Ranker + PrincipalComponents

Attribute selection output	
0.77263	1 0.321 histogram_min -0.309 mean_value_of_short_term_variability+0.305 histogram_mean -0.301 histogram_width -0.28 histogram.variance...
0.5584	2 -0.321 histogram_mode -0.317 histogram_median -0.303 accelerations+0.303 fetal_health=Pathological -0.294 fetal_health=Normal...
0.4383	3 0.413 fetal_health=Normal -0.348 baseline_value -0.307 fetal_health=Suspect -0.302 histogram_max -0.262 abnormal_short_term_variability...
0.3694	4 0.496 mean_value_of_long_term_variability+0.385 histogram_tendency -0.371 accelerations -0.273 histogram_min+0.271 histogram_number_of_peaks...
0.3515	5 0.456 uterine_contractions+0.361 uterine_tendancy -0.351 uterine_tendancy+0.351 uterine_tendancy+0.341 long_term_variability+0.341 dependency...
0.2701	6 0.586 histogram_tendency+0.319 histogram_max+0.317 fetal_health=Suspect+0.306 mean_value_of_long_term_variability+0.276 fetal_movement...
0.2283	7 -0.468 uterine_contractions -0.435 fetal_movement+0.327 accelerations+0.277 abnormal_short_term_variability -0.276 prolonged_decelerations...
0.1987	8 -0.538 histogram_number_of_zeroes+0.498 fetal_health=Suspect -0.463 fetal_health=Pathological+0.223 fetal_movement+0.221 light_decelerations...
0.1578	9 0.656 histogram_number_of_zeroes+0.428 fetal_movement -0.417 mean_value_of_long_term_variability -0.213 fetal_health=Pathological+0.151 light_deceleration...
0.1293	10 -0.423 accelerations -0.325 fetal_health=Suspect+0.321 light_decelerations -0.305 histogram_tendency -0.29 prolonged_decelerations...
0.1041	11 0.554 uterine_contractions+0.486 abnormal_short_term_variability -0.327 histogram_number_of_peaks -0.319 mean_value_of_short_term_variability -0.284 light...
0.0806	12 -0.698 percentage_of_time_with_abnormal_long_term_variability+0.323 prolonged_decelerations+0.318 abnormal_short_term_variability -0.269 uterine_contractions...
0.0622	13 -0.529 histogram_number_of_peaks -0.453 prolonged_decelerations+0.359 abnormal_short_term_variability+0.273 accelerations+0.243 mean_value_of_long_term_variability...
0.0464	14 0.733 histogram.variance -0.468 mean_value_of_short_term_variability -0.254 fetal_health=Pathological -0.217 histogram_number_of_peaks+0.202 percentage_of...

Selected attributes: 1,2,3,4,5,6,7,8,9,10,11,12,13,14 : 14

Status OK Log  x 0