

0.1 Plots of single attributes

0.1.1 sulphates

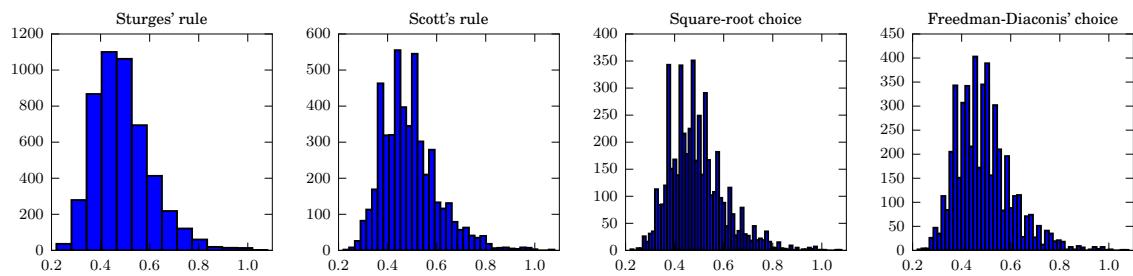


Figure 1: Histograms of attribute *sulphates* using different binning methods

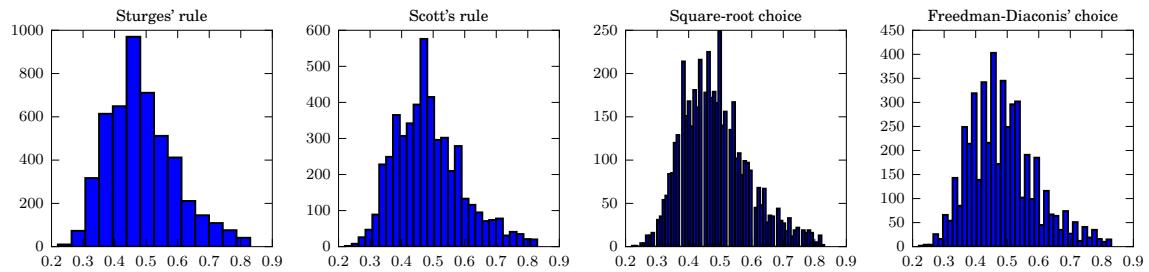


Figure 2: Histograms of attribute *sulphates* with outliers further than 3 standard deviations from the mean filtered

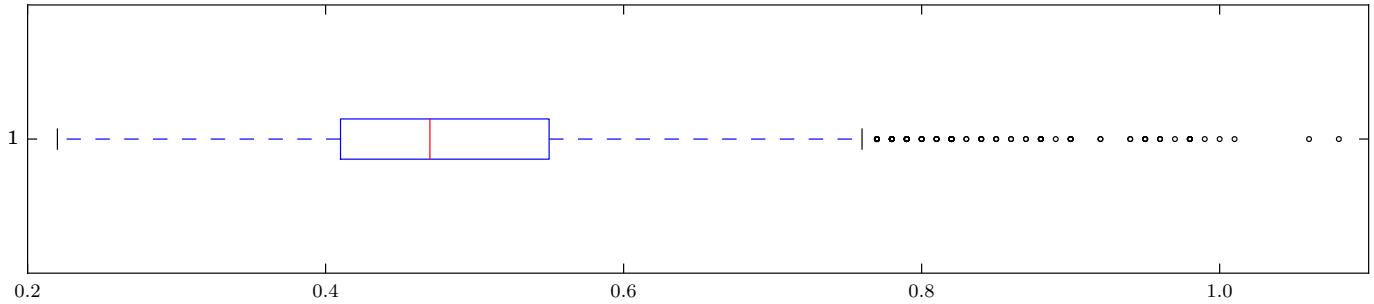


Figure 3: Boxplot of attribute *sulphates*

0.1.2 density

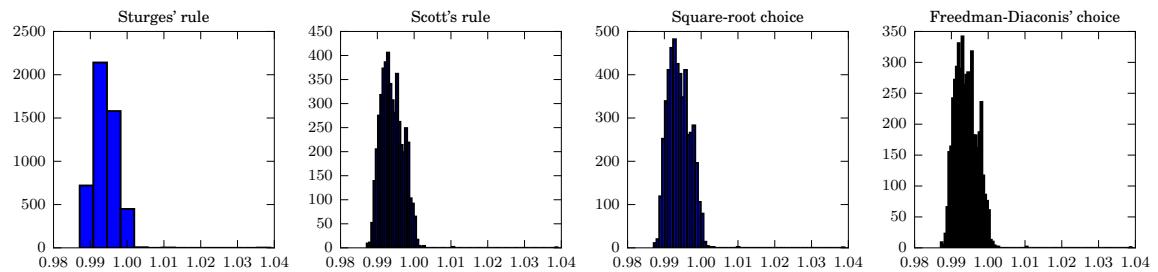


Figure 4: Histograms of attribute *density* using different binning methods

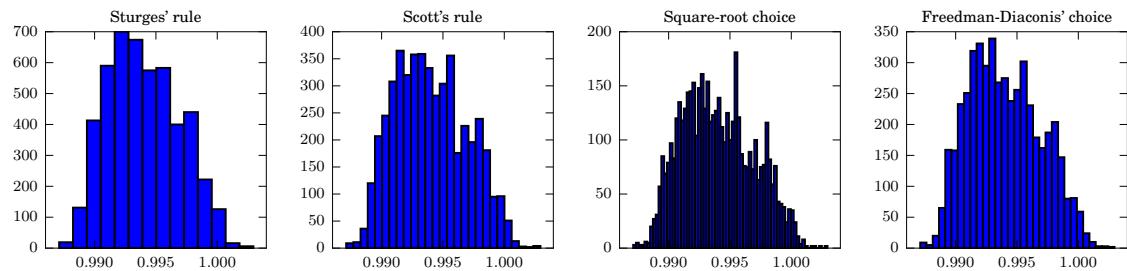


Figure 5: Histograms of attribute *density* with outliers further than 3 standard deviations from the mean filtered

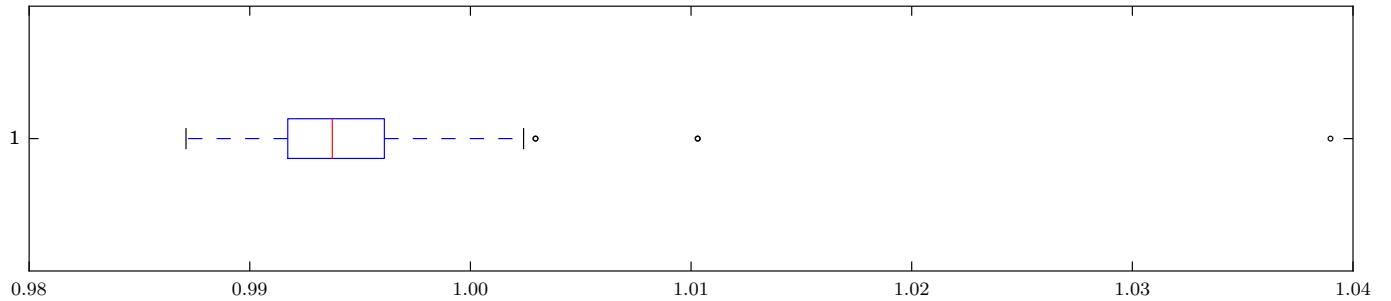


Figure 6: Boxplot of attribute *density*

0.1.3 pH

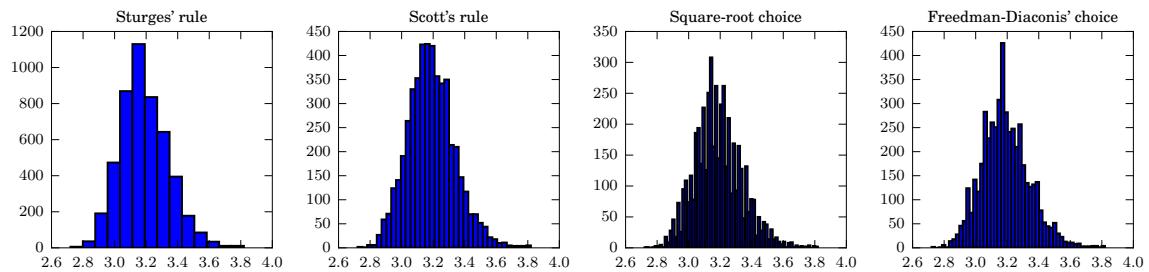


Figure 7: Histograms of attribute pH using different binning methods

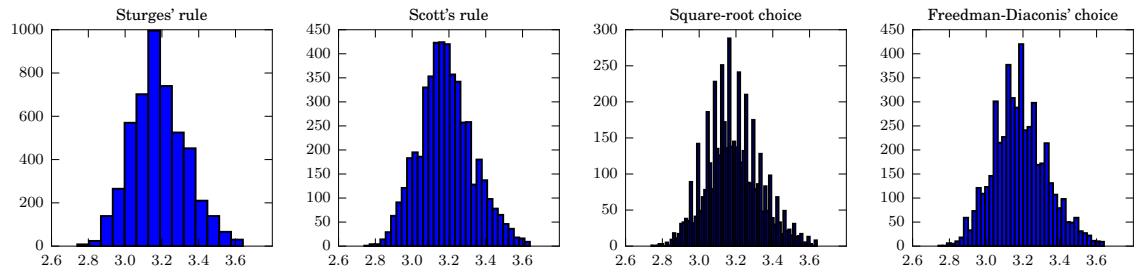


Figure 8: Histograms of attribute pH with outliers further than 3 standard deviations from the mean filtered

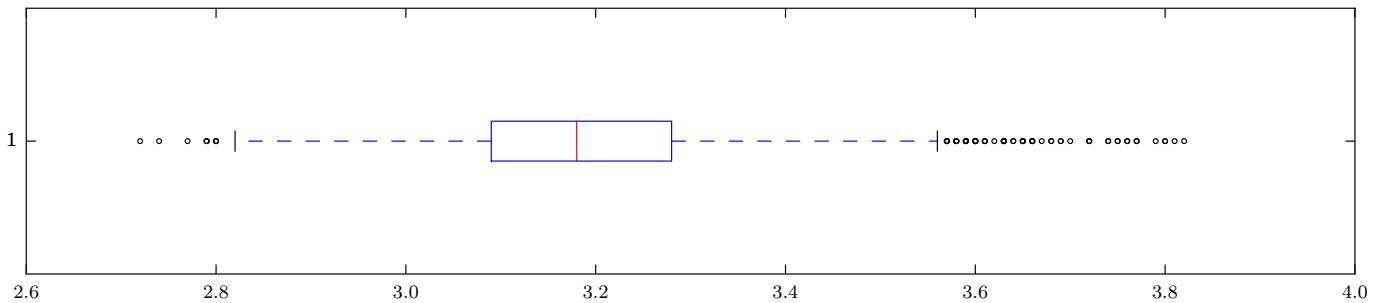


Figure 9: Boxplot of attribute pH

0.1.4 residual sugar

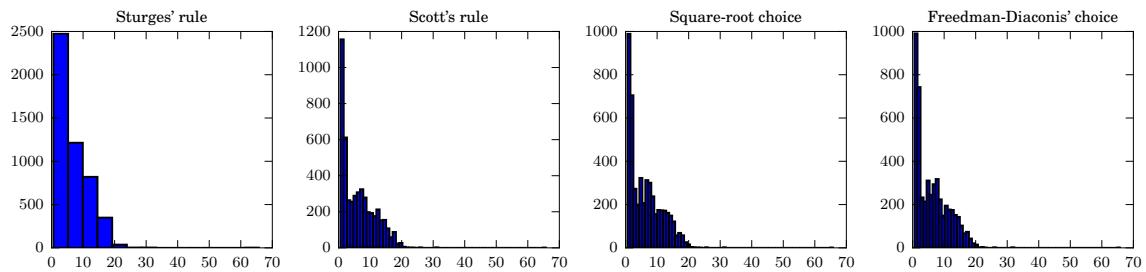


Figure 10: Histograms of attribute *residual sugar* using different binning methods

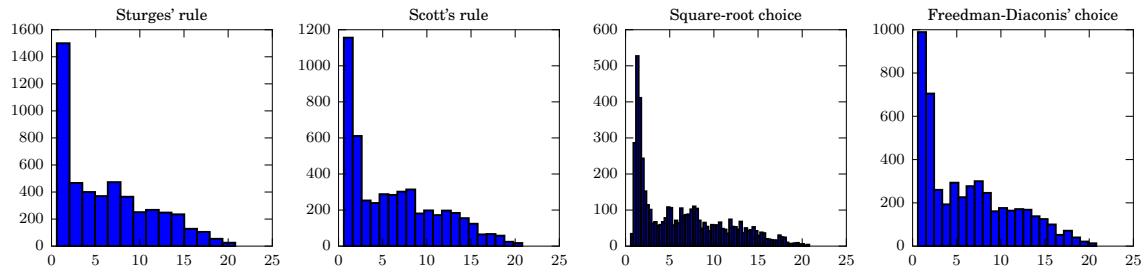


Figure 11: Histograms of attribute *residual sugar* with outliers further than 3 standard deviations from the mean filtered

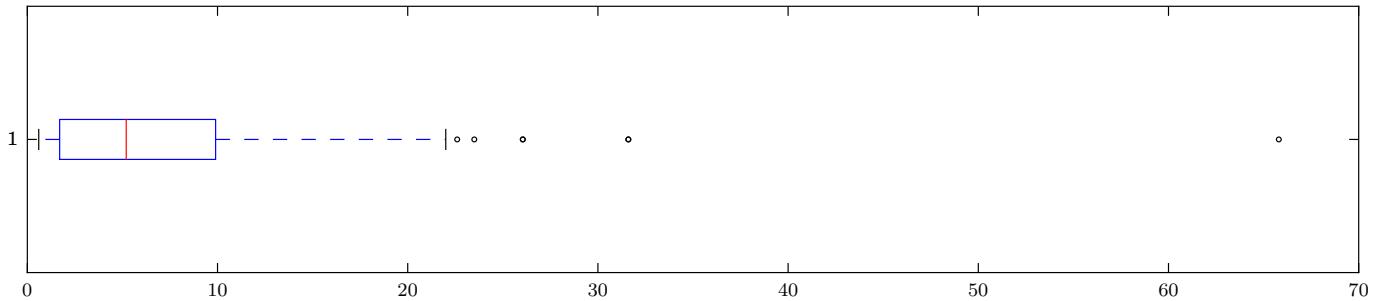


Figure 12: Boxplot of attribute *residual sugar*

0.1.5 citric acid

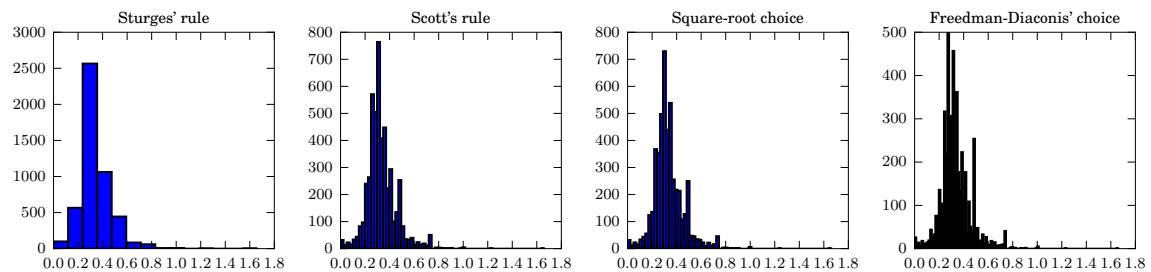


Figure 13: Histograms of attribute *citric acid* using different binning methods

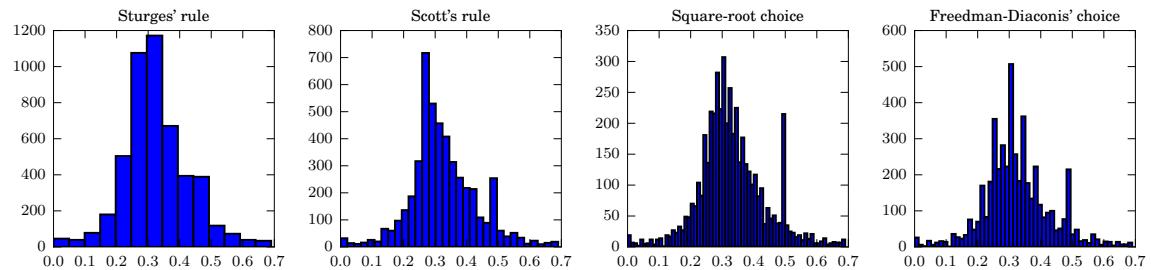


Figure 14: Histograms of attribute *citric acid* with outliers further than 3 standard deviations from the mean filtered

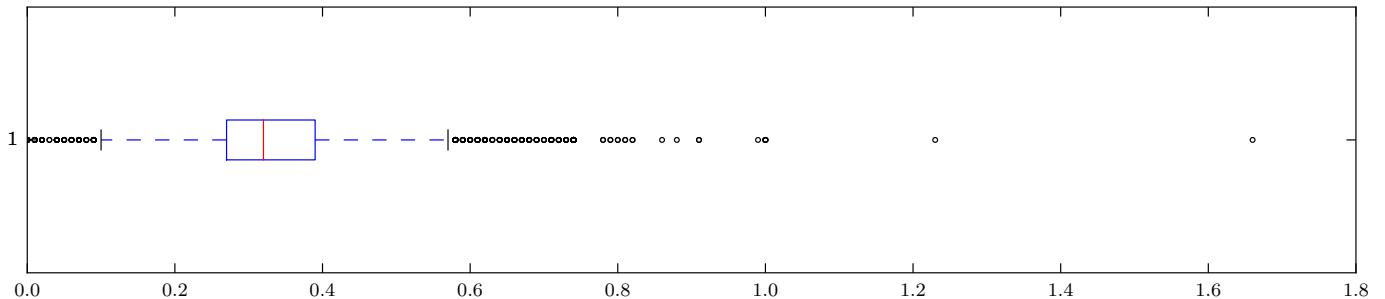


Figure 15: Boxplot of attribute *citric acid*

0.1.6 alcohol

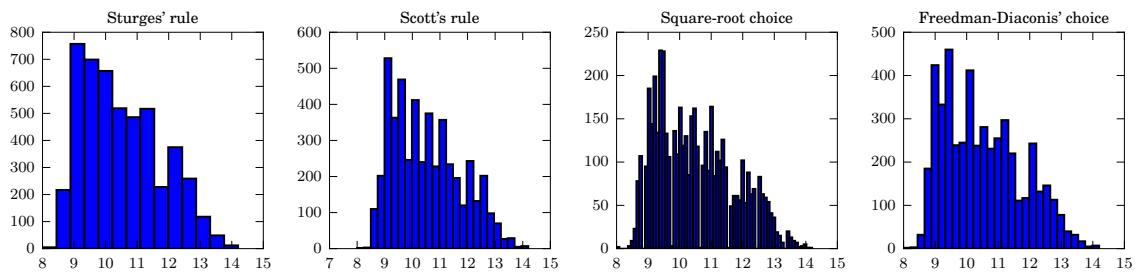


Figure 16: Histograms of attribute *alcohol* using different binning methods

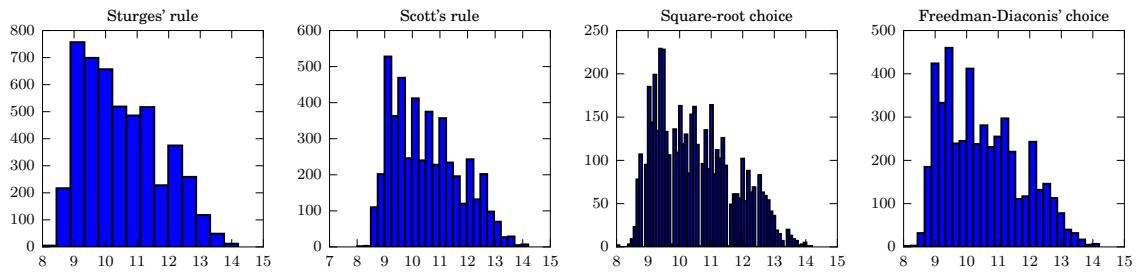


Figure 17: Histograms of attribute *alcohol* with outliers further than 3 standard deviations from the mean filtered

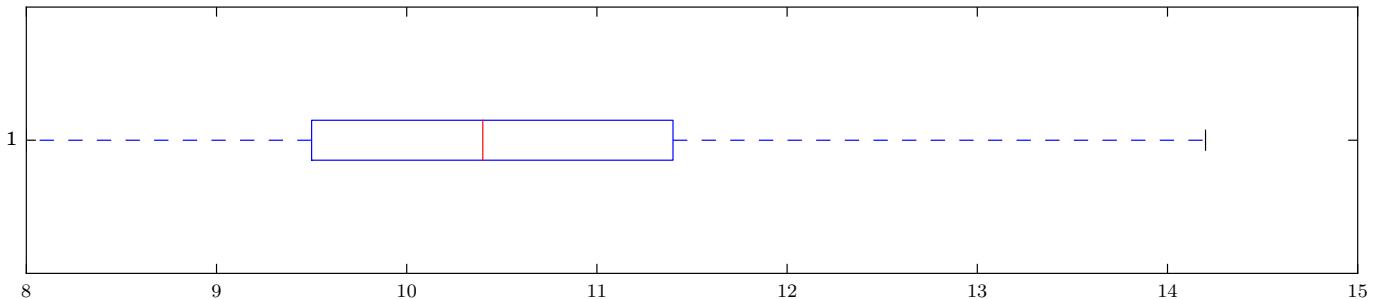


Figure 18: Boxplot of attribute *alcohol*

0.1.7 free sulfur dioxide

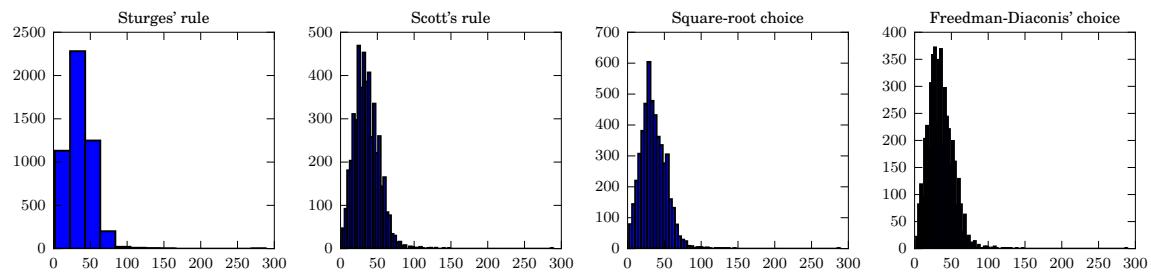


Figure 19: Histograms of attribute *free sulfur dioxide* using different binning methods

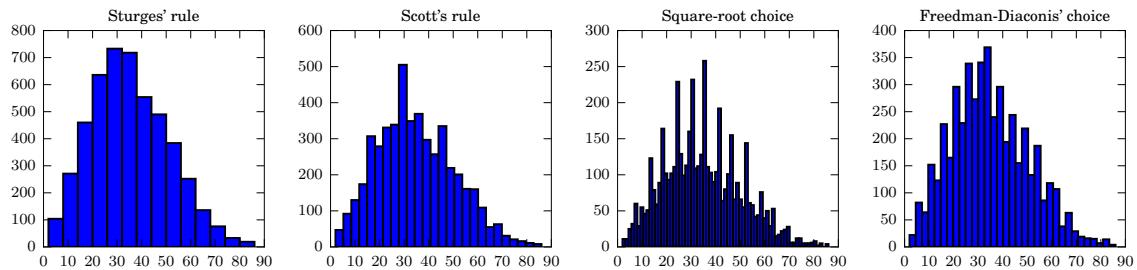


Figure 20: Histograms of attribute *free sulfur dioxide* with outliers further than 3 standard deviations from the mean filtered

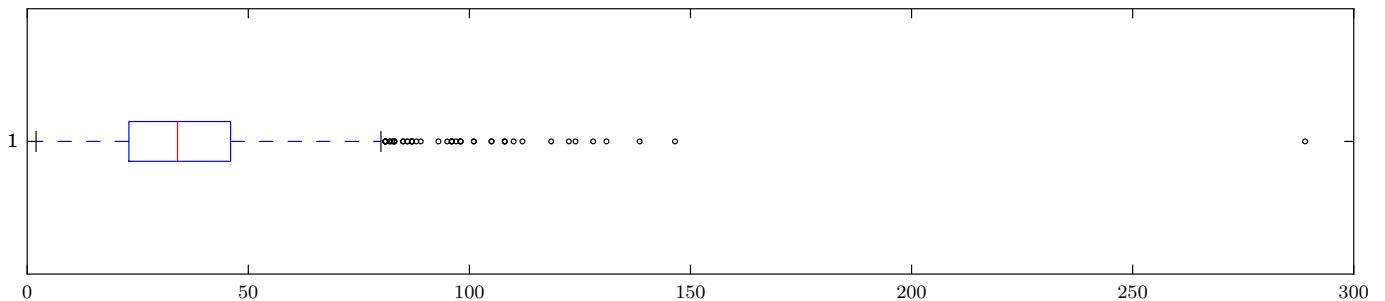


Figure 21: Boxplot of attribute *free sulfur dioxide*

0.1.8 volatile acidity

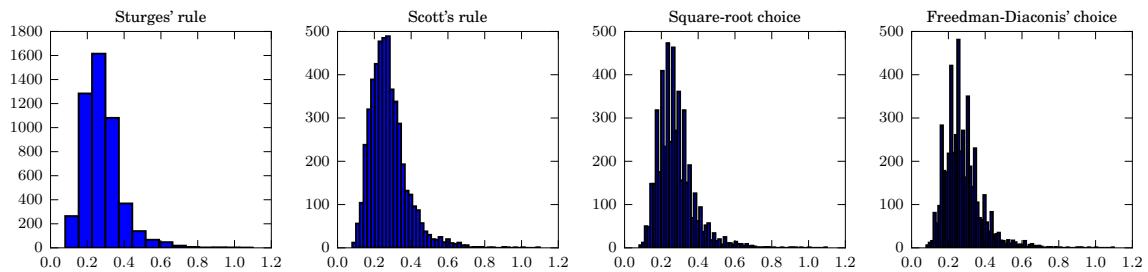


Figure 22: Histograms of attribute *volatile acidity* using different binning methods

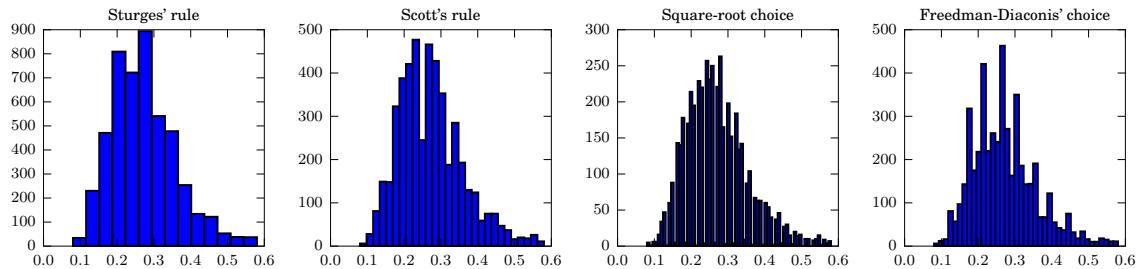


Figure 23: Histograms of attribute *volatile acidity* with outliers further than 3 standard deviations from the mean filtered

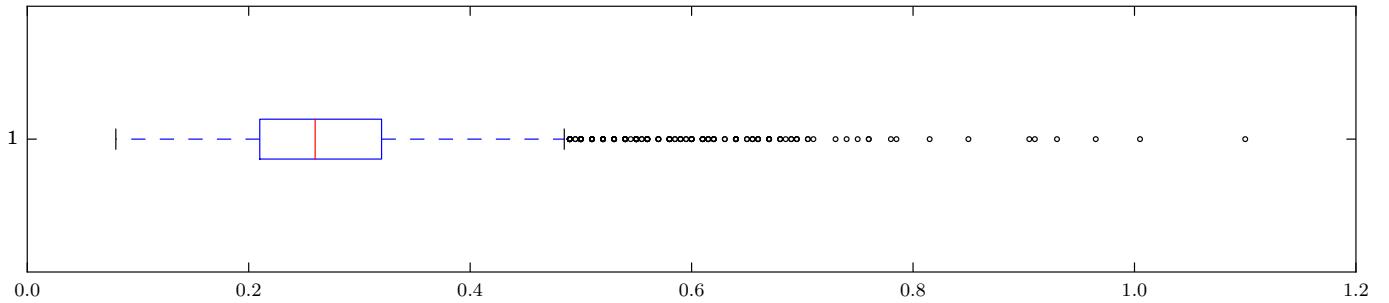


Figure 24: Boxplot of attribute *volatile acidity*

0.1.9 quality

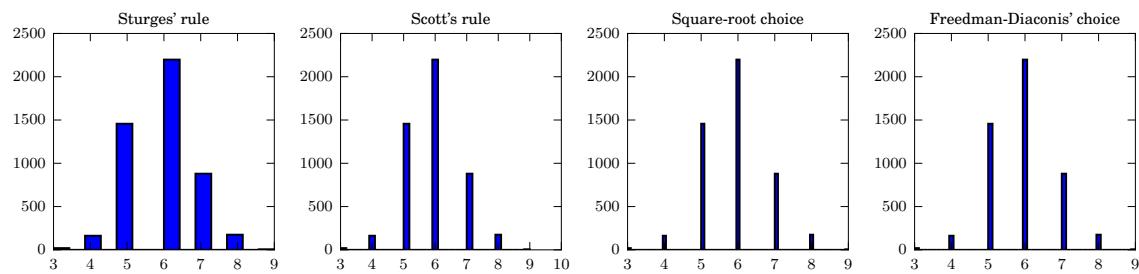


Figure 25: Histograms of attribute *quality* using different binning methods

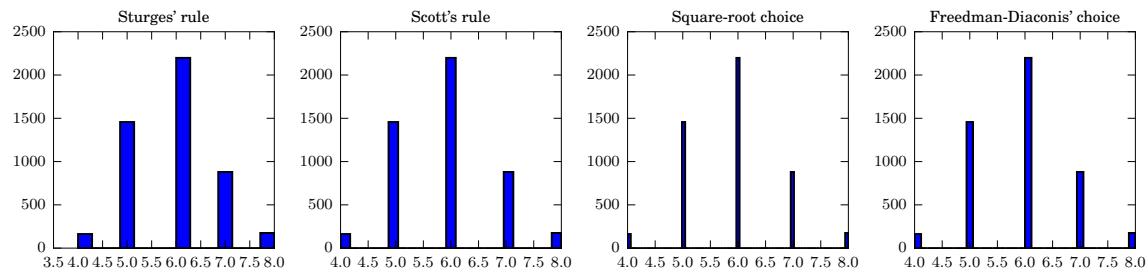


Figure 26: Histograms of attribute *quality* with outliers further than 3 standard deviations from the mean filtered

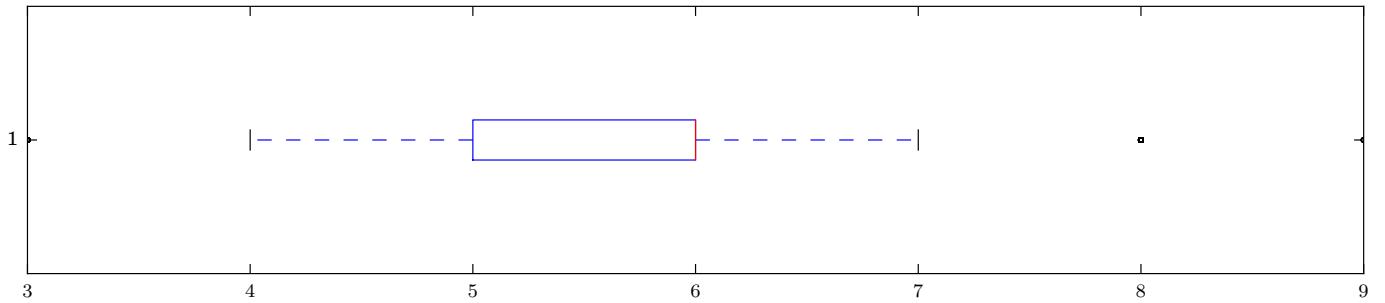


Figure 27: Boxplot of attribute *quality*

0.1.10 total sulfur dioxide

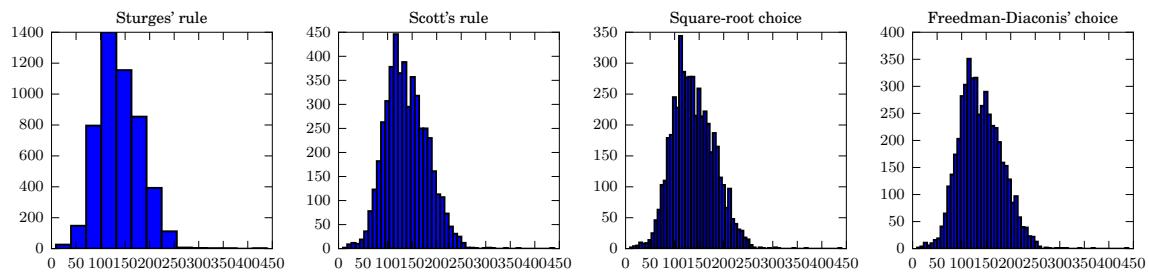


Figure 28: Histograms of attribute *total sulfur dioxide* using different binning methods

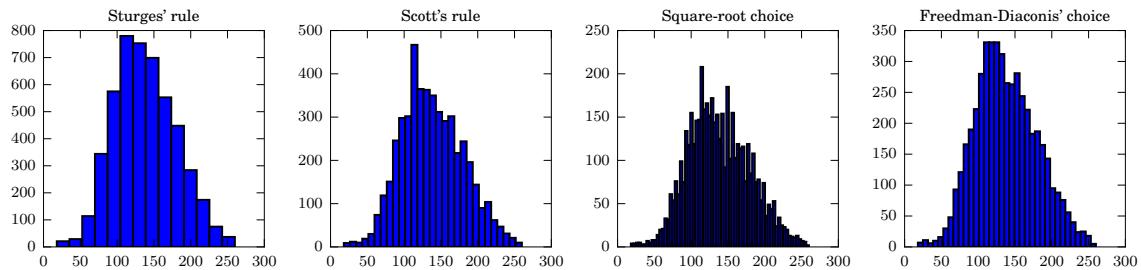


Figure 29: Histograms of attribute *total sulfur dioxide* with outliers further than 3 standard deviations from the mean filtered

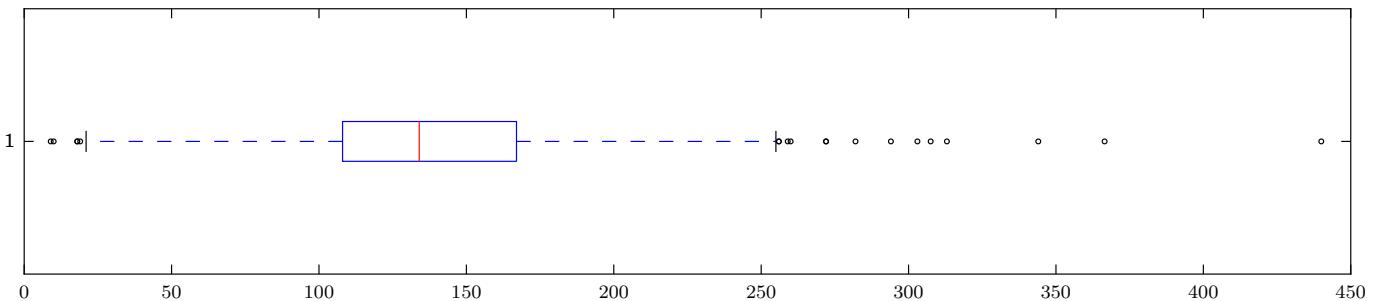


Figure 30: Boxplot of attribute *total sulfur dioxide*

0.1.11 fixed acidity

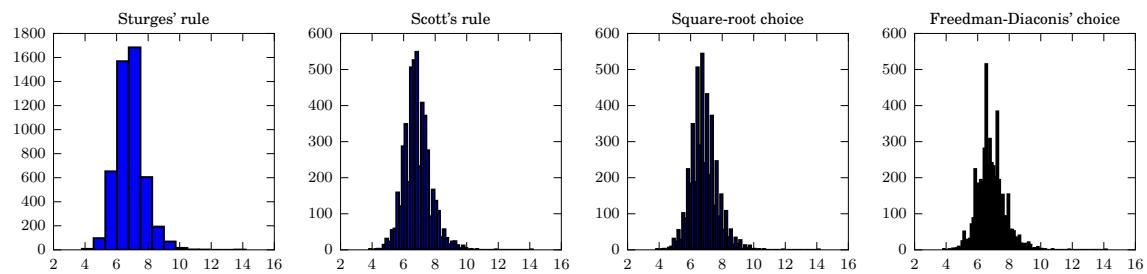


Figure 31: Histograms of attribute *fixed acidity* using different binning methods

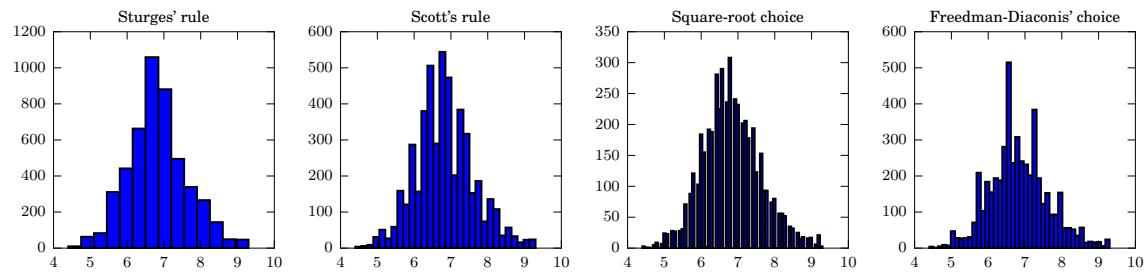


Figure 32: Histograms of attribute *fixed acidity* with outliers further than 3 standard deviations from the mean filtered

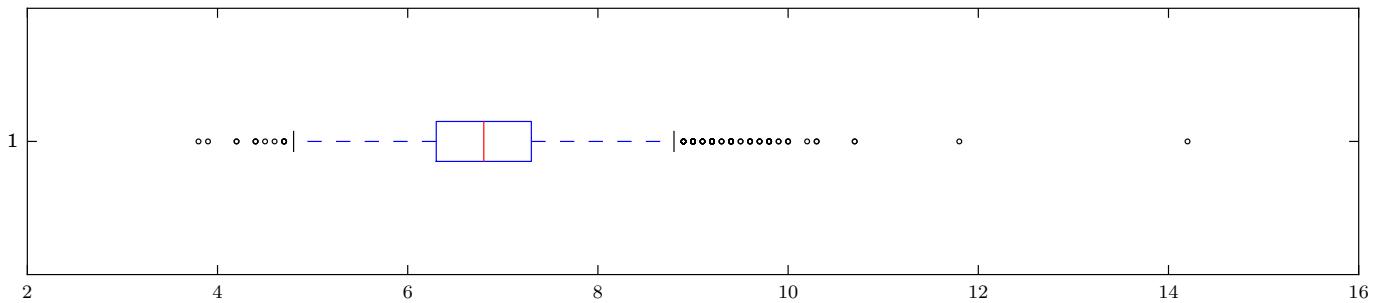


Figure 33: Boxplot of attribute *fixed acidity*

0.1.12 chlorides

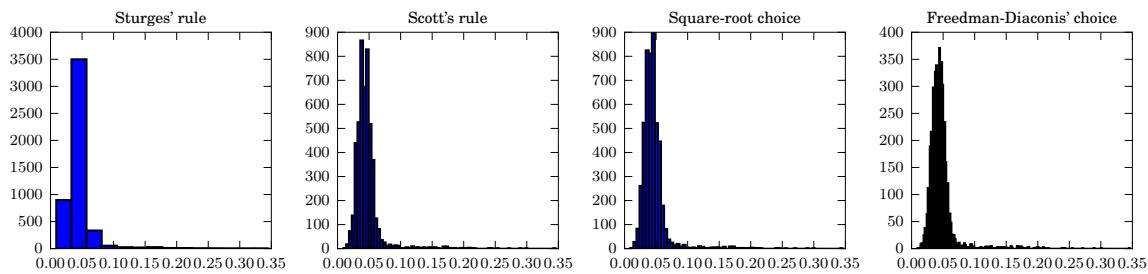


Figure 34: Histograms of attribute *chlorides* using different binning methods

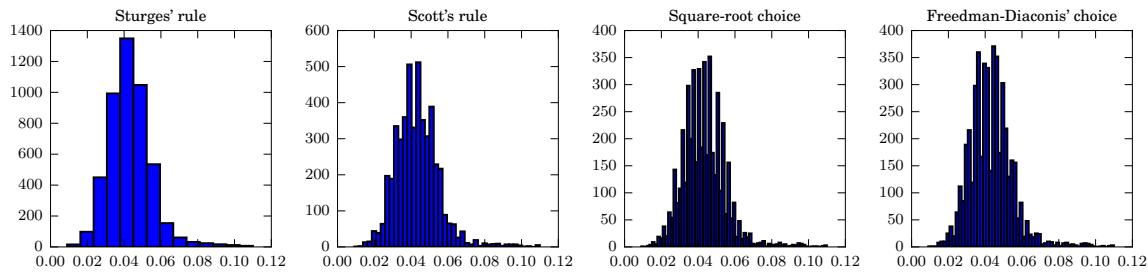


Figure 35: Histograms of attribute *chlorides* with outliers further than 3 standard deviations from the mean filtered

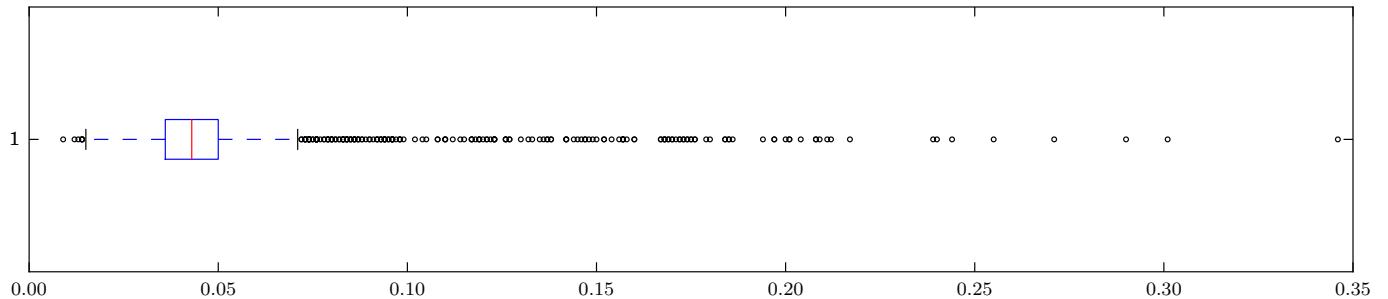


Figure 36: Boxplot of attribute *chlorides*

0.2 Plots for the whole feature set

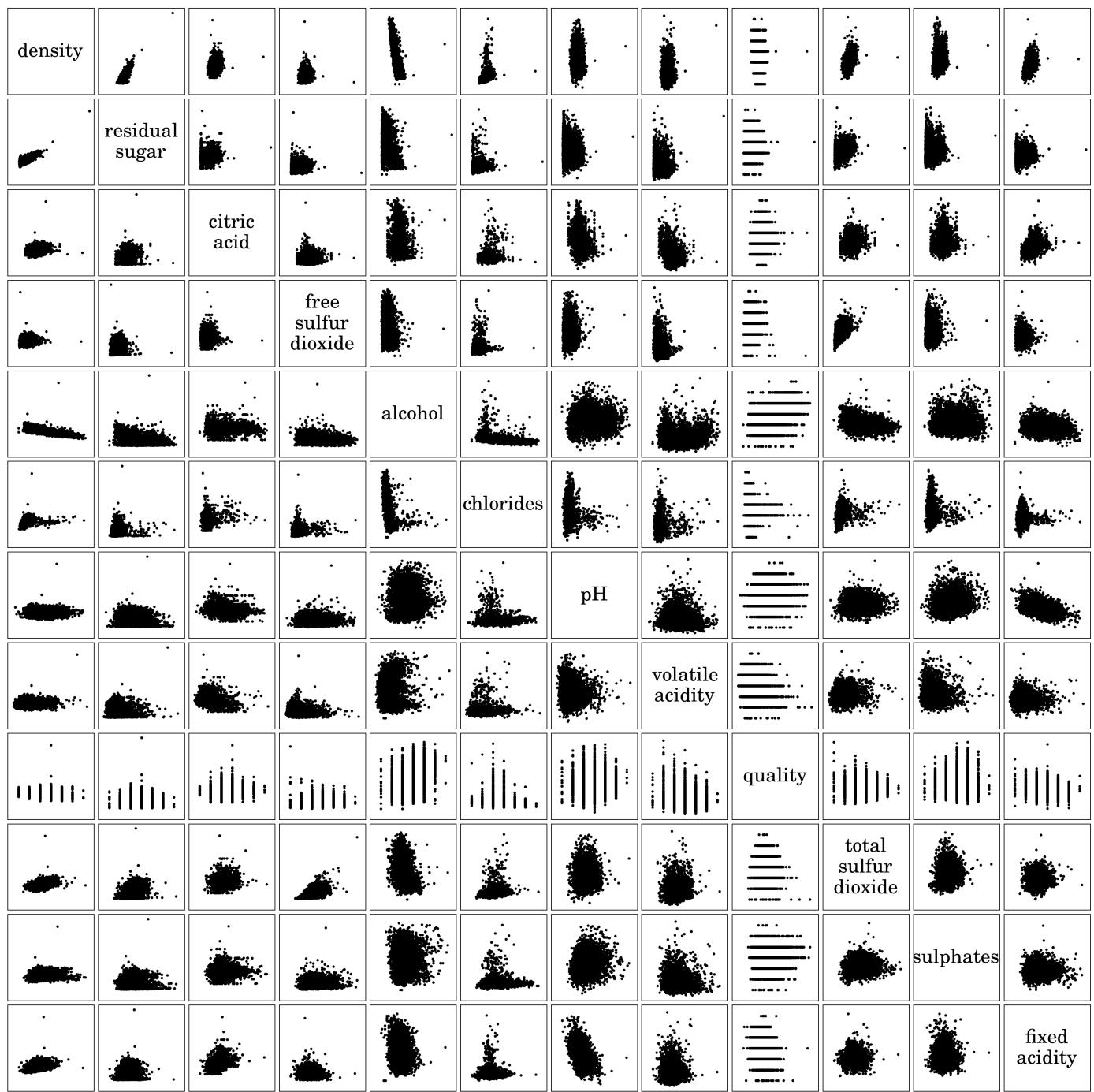


Figure 37: Scatter matrix of the whole feature set

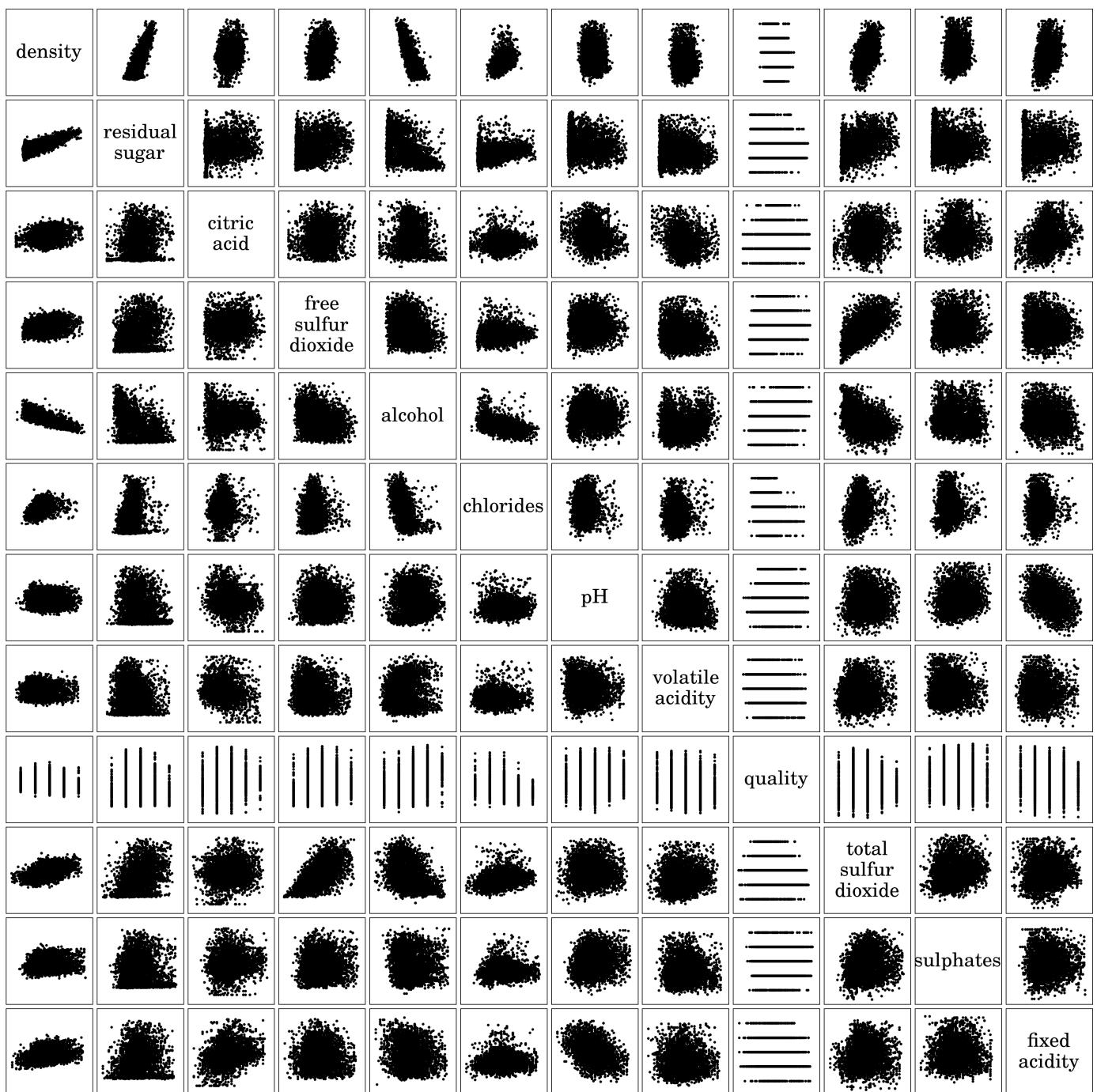


Figure 38: Scatter matrix of the whole feature set with outliers further than 3 standard deviations from the mean filtered

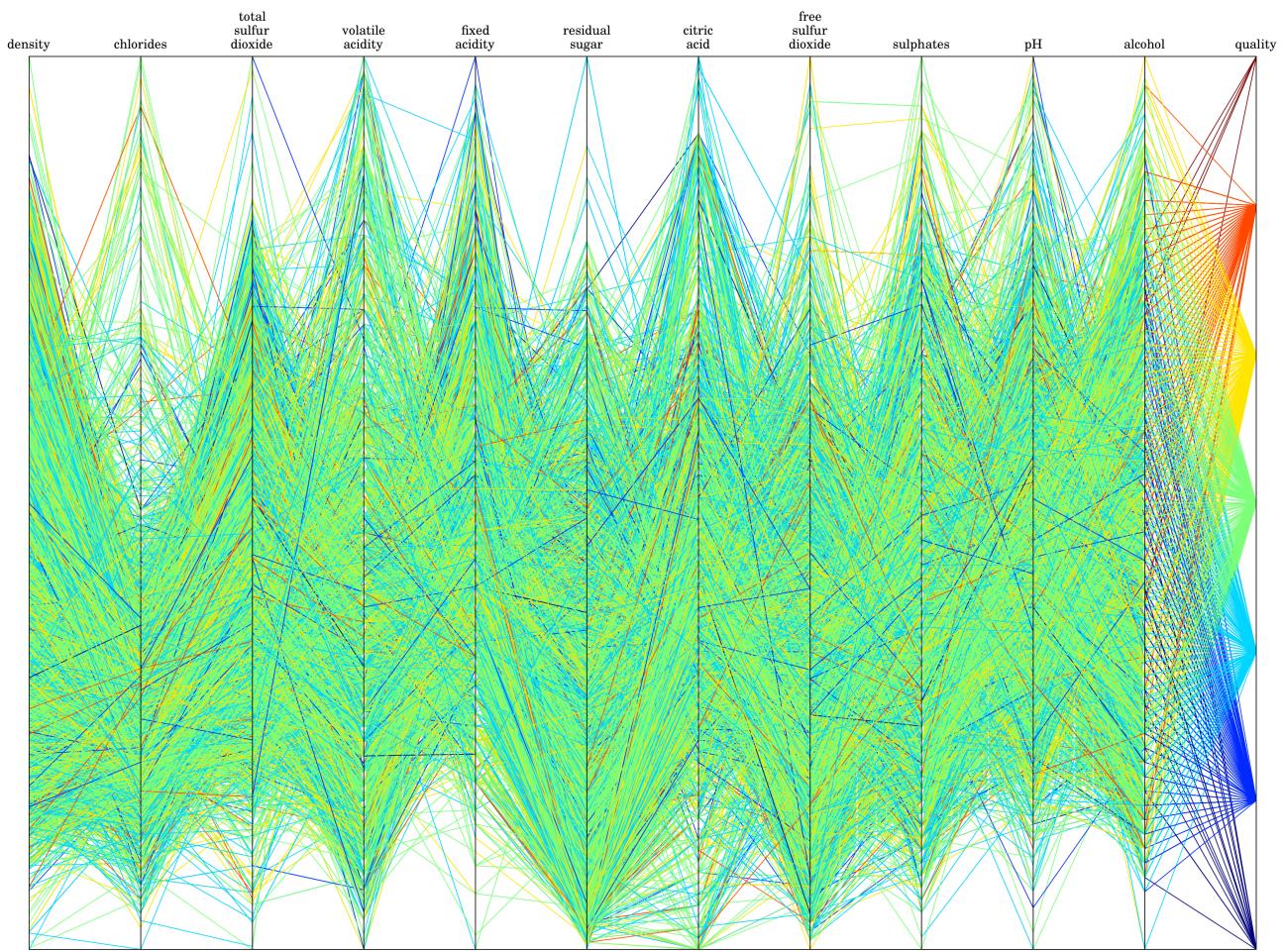


Figure 39: Parallel coordinates representation of the data set

0.2.1 Principal component analysis

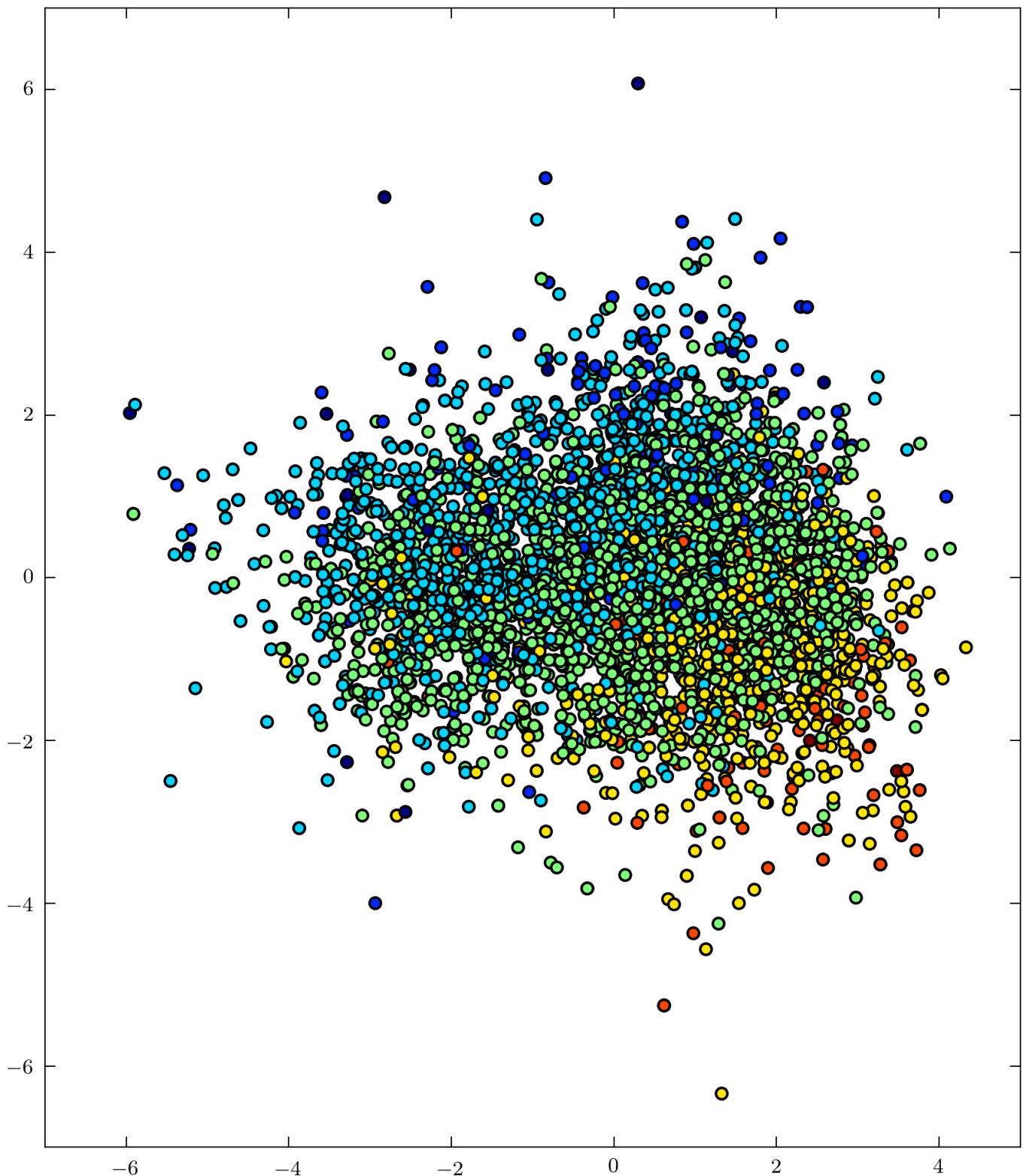


Figure 40: 2D projection of the normalized data set using PCA

| | PC1 | PC2 | PC3 | PC4 | PC5 | PC6 | PC7 | PC8 | PC9 | PC10 | PC11 | PC12 |
|------------------------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|
| Proportion of variance | 0.2746 | 0.1317 | 0.1170 | 0.0923 | 0.0856 | 0.0746 | 0.0654 | 0.0581 | 0.0469 | 0.0293 | 0.0226 | 0.0017 |
| Cumulative variance | 0.2746 | 0.4063 | 0.5234 | 0.6157 | 0.7013 | 0.7760 | 0.8414 | 0.8995 | 0.9464 | 0.9757 | 0.9983 | 1.0000 |

0.3 Correlation coefficients using different functions

0.3.1 Correlation coefficients using Pearson's correlation coefficient

| | sulphates | density | pH | residual sugar | citric acid | alcohol | free sulfur dioxide | volatile acidity | quality | total sulfur dioxide | fixed acidity | chlorides |
|----------------------|---------------|----------------|---------------|----------------|---------------|----------------|---------------------|------------------|---------------|----------------------|---------------|---------------|
| sulphates | 1.0000 | 0.0745 | 0.1560 | -0.0267 | 0.0623 | -0.0174 | 0.0592 | -0.0357 | 0.0537 | 0.1346 | -0.0171 | 0.0168 |
| density | 0.0745 | 1.0000 | -0.0936 | 0.8390 | 0.1495 | -0.7801 | 0.2942 | 0.0271 | -0.3071 | 0.5299 | 0.2653 | 0.2572 |
| pH | 0.1560 | -0.0936 | 1.0000 | -0.1941 | -0.1637 | 0.1214 | -0.0006 | -0.0319 | 0.0994 | 0.0023 | -0.4259 | -0.0904 |
| residual sugar | -0.0267 | 0.8390 | -0.1941 | 1.0000 | 0.0942 | -0.4506 | 0.2991 | 0.0643 | -0.0976 | 0.4014 | 0.0890 | 0.0887 |
| citric acid | 0.0623 | 0.1495 | -0.1637 | 0.0942 | 1.0000 | -0.0757 | 0.0941 | -0.1495 | -0.0092 | 0.1211 | 0.2892 | 0.1144 |
| alcohol | -0.0174 | -0.7801 | 0.1214 | -0.4506 | -0.0757 | 1.0000 | -0.2501 | 0.0677 | 0.4356 | -0.4489 | -0.1209 | -0.3602 |
| free sulfur dioxide | 0.0592 | 0.2942 | -0.0006 | 0.2991 | 0.0941 | -0.2501 | 1.0000 | -0.0970 | 0.0082 | 0.6155 | -0.0494 | 0.1014 |
| volatile acidity | -0.0357 | 0.0271 | -0.0319 | 0.0643 | -0.1495 | 0.0677 | -0.0970 | 1.0000 | -0.1947 | 0.0893 | -0.0227 | 0.0705 |
| quality | 0.0537 | -0.3071 | 0.0994 | -0.0976 | -0.0092 | 0.4356 | 0.0082 | -0.1947 | 1.0000 | -0.1747 | -0.1137 | -0.2099 |
| total sulfur dioxide | 0.1346 | 0.5299 | 0.0023 | 0.4014 | 0.1211 | -0.4489 | 0.6155 | 0.0893 | -0.1747 | 1.0000 | 0.0911 | 0.1989 |
| fixed acidity | -0.0171 | 0.2653 | -0.4259 | 0.0890 | 0.2892 | -0.1209 | -0.0494 | -0.0227 | -0.1137 | 0.0911 | 1.0000 | 0.0231 |
| chlorides | 0.0168 | 0.2572 | -0.0904 | 0.0887 | 0.1144 | -0.3602 | 0.1014 | 0.0705 | -0.2099 | 0.1989 | 0.0231 | 1.0000 |

0.3.2 Correlation coefficients using Spearman's rho

| | sulphates | density | pH | residual sugar | citric acid | alcohol | free sulfur dioxide | volatile acidity | quality | total sulfur dioxide | fixed acidity | chlorides |
|----------------------|---------------|----------------|---------------|----------------|---------------|----------------|---------------------|------------------|---------------|----------------------|---------------|----------------|
| sulphates | 1.0000 | 0.0951 | 0.1402 | -0.0038 | 0.0798 | -0.0449 | 0.0523 | -0.0169 | 0.0333 | 0.1578 | -0.0132 | 0.0939 |
| density | 0.0951 | 1.0000 | -0.1101 | 0.7804 | 0.0914 | -0.8219 | 0.3278 | 0.0101 | -0.3484 | 0.5638 | 0.2700 | 0.5083 |
| pH | 0.1402 | -0.1101 | 1.0000 | -0.1800 | -0.1462 | 0.1489 | -0.0063 | -0.0452 | 0.1094 | -0.0118 | -0.4183 | -0.0540 |
| residual sugar | -0.0038 | 0.7804 | -0.1800 | 1.0000 | 0.0246 | -0.4453 | 0.3461 | 0.1086 | -0.0821 | 0.4313 | 0.1067 | 0.2278 |
| citric acid | 0.0798 | 0.0914 | -0.1462 | 0.0246 | 1.0000 | -0.0292 | 0.0883 | -0.1504 | 0.0183 | 0.0932 | 0.2979 | 0.0327 |
| alcohol | -0.0449 | -0.8219 | 0.1489 | -0.4453 | -0.0292 | 1.0000 | -0.2726 | 0.0340 | 0.4404 | -0.4766 | -0.1068 | -0.5708 |
| free sulfur dioxide | 0.0523 | 0.3278 | -0.0063 | 0.3461 | 0.0883 | -0.2726 | 1.0000 | -0.0812 | 0.0237 | 0.6186 | -0.0245 | 0.1670 |
| volatile acidity | -0.0169 | 0.0101 | -0.0452 | 0.1086 | -0.1504 | 0.0340 | -0.0812 | 1.0000 | -0.1966 | 0.1176 | -0.0429 | -0.0049 |
| quality | 0.0333 | -0.3484 | 0.1094 | -0.0821 | 0.0183 | 0.4404 | 0.0237 | -0.1966 | 1.0000 | -0.1967 | -0.0845 | -0.3145 |
| total sulfur dioxide | 0.1578 | 0.5638 | -0.0118 | 0.4313 | 0.0932 | -0.4766 | 0.6186 | 0.1176 | -0.1967 | 1.0000 | 0.1126 | 0.3752 |
| fixed acidity | -0.0132 | 0.2700 | -0.4183 | 0.1067 | 0.2979 | -0.1068 | -0.0245 | -0.0429 | -0.0845 | 0.1126 | 1.0000 | 0.0947 |
| chlorides | 0.0939 | 0.5083 | -0.0540 | 0.2278 | 0.0327 | -0.5708 | 0.1670 | -0.0049 | -0.3145 | 0.3752 | 0.0947 | 1.0000 |

0.3.3 Correlation coefficients using Kendall's tau

| | sulphates | density | pH | residual sugar | citric acid | alcohol | free sulfur dioxide | volatile acidity | quality | total sulfur dioxide | fixed acidity | chlorides |
|----------------------|---------------|----------------|---------------|----------------|---------------|----------------|---------------------|------------------|---------------|----------------------|---------------|---------------|
| sulphates | 1.0000 | 0.0642 | 0.0958 | -0.0025 | 0.0545 | -0.0264 | 0.0356 | -0.0116 | 0.0264 | 0.1087 | -0.0087 | 0.0626 |
| density | 0.0642 | 1.0000 | -0.0756 | 0.5890 | 0.0615 | -0.6351 | 0.2173 | 0.0066 | -0.2666 | 0.3884 | 0.1855 | 0.3491 |
| pH | 0.0958 | -0.0756 | 1.0000 | -0.1256 | -0.1013 | 0.1026 | -0.0052 | -0.0304 | 0.0844 | -0.0084 | -0.2948 | -0.0379 |
| residual sugar | -0.0025 | 0.5890 | -0.1256 | 1.0000 | 0.0153 | -0.3056 | 0.2367 | 0.0728 | -0.0631 | 0.2933 | 0.0749 | 0.1553 |
| citric acid | 0.0545 | 0.0615 | -0.1013 | 0.0153 | 1.0000 | -0.0200 | 0.0608 | -0.1040 | 0.0146 | 0.0622 | 0.2086 | 0.0223 |
| alcohol | -0.0264 | -0.6351 | 0.1026 | -0.3056 | -0.0200 | 1.0000 | -0.1825 | 0.0235 | 0.3467 | -0.3258 | -0.0732 | -0.4040 |
| free sulfur dioxide | 0.0356 | 0.2173 | -0.0052 | 0.2367 | 0.0608 | -0.1825 | 1.0000 | -0.0548 | 0.0172 | 0.4447 | -0.0169 | 0.1139 |
| volatile acidity | -0.0116 | 0.0066 | -0.0304 | 0.0728 | -0.1040 | 0.0235 | -0.0548 | 1.0000 | -0.1548 | 0.0813 | -0.0296 | -0.0035 |
| quality | 0.0264 | -0.2666 | 0.0844 | -0.0631 | 0.0146 | 0.3467 | 0.0172 | -0.1548 | 1.0000 | -0.1512 | -0.0655 | -0.2449 |
| total sulfur dioxide | 0.1087 | 0.3884 | -0.0084 | 0.2933 | 0.0622 | -0.3258 | 0.4447 | 0.0813 | -0.1512 | 1.0000 | 0.0773 | 0.2571 |
| fixed acidity | -0.0087 | 0.1855 | -0.2948 | 0.0749 | 0.2086 | -0.0732 | -0.0169 | -0.0296 | -0.0655 | 0.0773 | 1.0000 | 0.0654 |
| chlorides | 0.0626 | 0.3491 | -0.0379 | 0.1553 | 0.0223 | -0.4040 | 0.1139 | -0.0035 | -0.2449 | 0.2571 | 0.0654 | 1.0000 |