Archive of Cider Pomology Archive No 20/2/1 21/8/2012

Stephen Parrett for the Bulmer Foundation 2010

Apple Density Trial Data Analysis

Introduction

The following report is the result of an initial analysis of the data obtained from an apple density trial in order to determine varietal variations in apple quality. In particular it will look at the difference and relationships between mass, diameter, volume, sugar content and density in apple varieties. Two main hypotheses will be considered:

a) Smaller apples are denser than larger apples.
b) Denser apples have higher sugar content as a percentage of mass than less dense apples.

It will also look at the statistical differences in density and sugar content (BRIX) between varieties.

Mean Comparison

The results contain data from 6 varieties of apple; Browns, Dabinet, Ellis Bitter, Harry Masters Jersey, Michelin and Yarlington Mill. Two separate data sets were collected comprising of 3 lots of 5 apples per variety for each data set. The following bar graphs plot the mean results for each variety from both sets of data for mass (g), diameter (mm), volume (ml), BRIX (°Bx), and density (g/cm^3).

Figures 1.1 to 1.3 compare the difference between varieties in their mean diameter, mass and volume. As evident by the graphs, Ellis Bitter is consistently the largest variety with Michelin smallest. As we would expect, those apples with the largest diameter then go on to have the largest mass and volume. It is also notable that there is very little variance between the varieties in diameter apart from Ellis Bitter. This is also true to a lesser extent for mass and volume. If the hypothesis is accurate then we would expect the largest apples such as Ellis Bitter to be the least dense and have a low BRIX, whereas the smaller apples such as Michelin should be the most dense and have a higher BRIX.

Figure 1.1

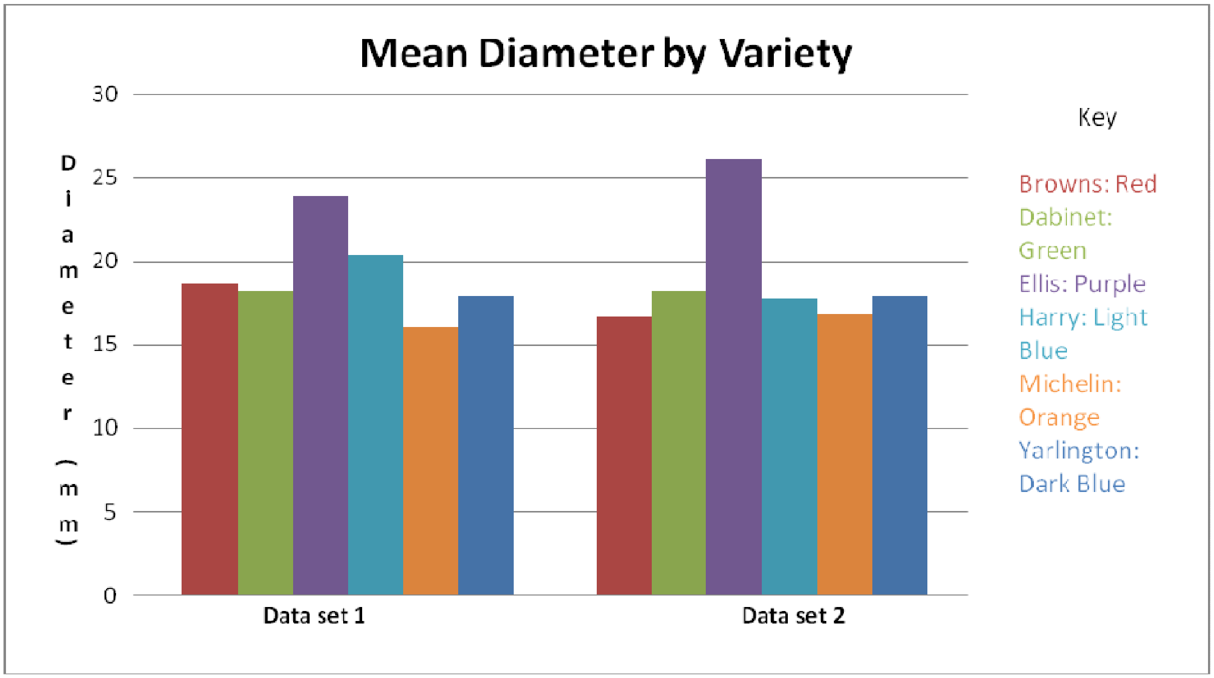


Figure 1.2

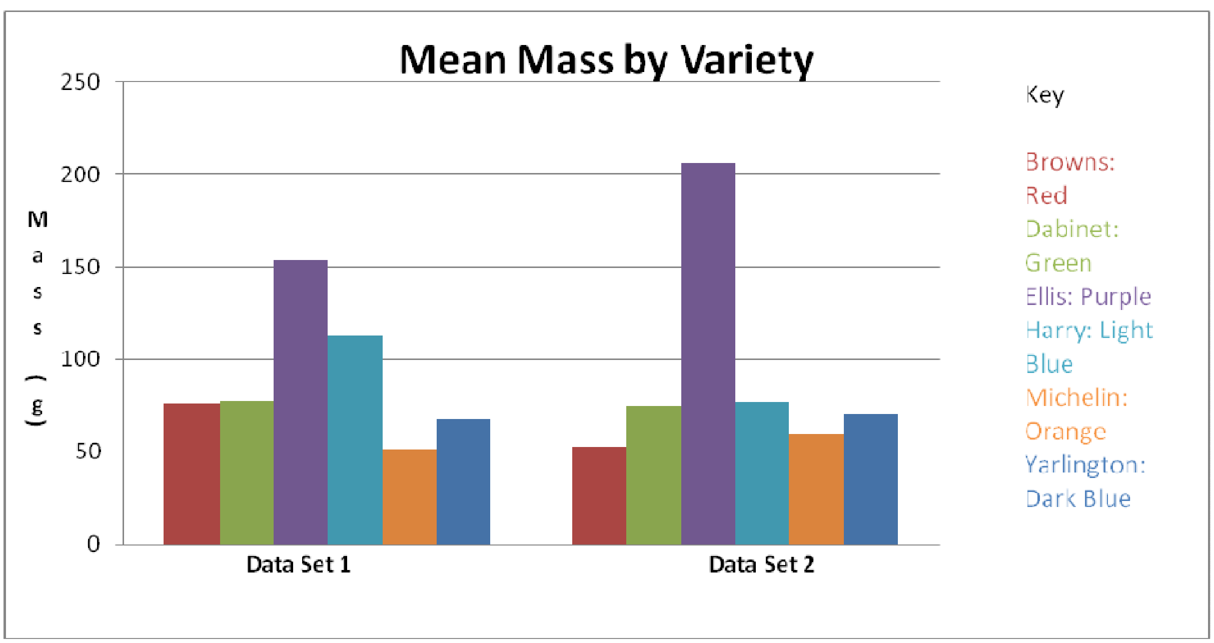


Figure 1.3

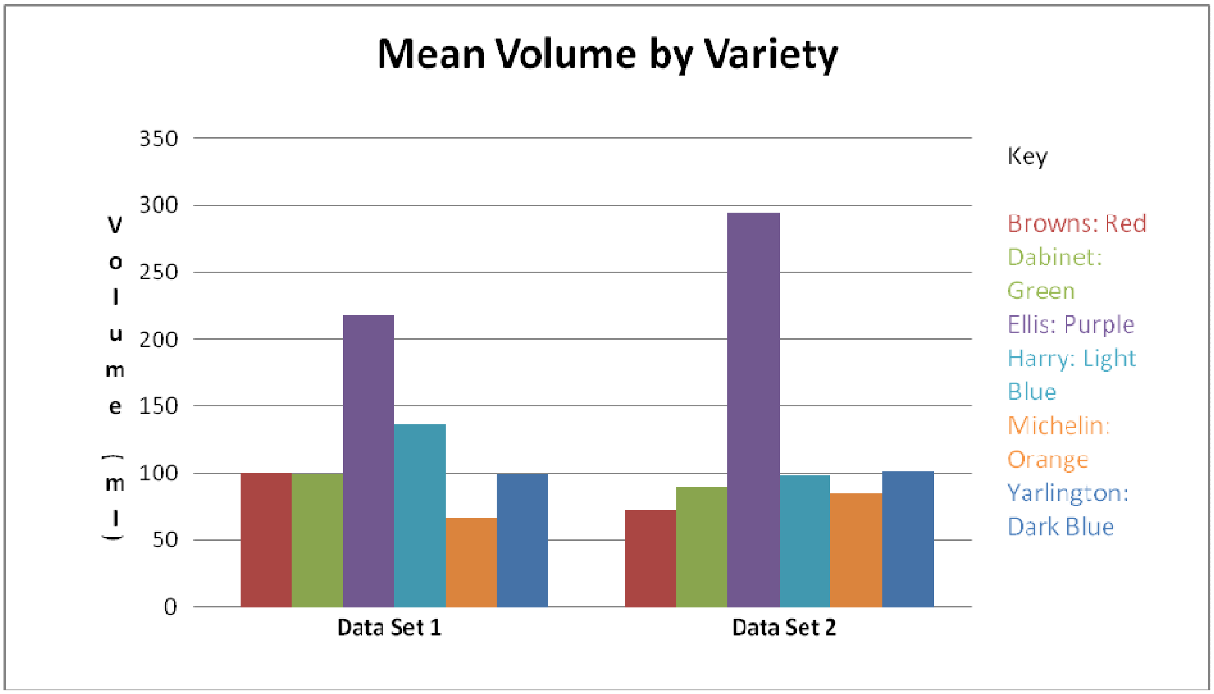


Figure 1.4 shows the mean density for each variety. Here there is a clear difference between the data sets in that Harry Masters comes out as the densest for data set 1, whereas Dabinet is the densest for data set 2. As expected, Ellis Bitter has a low density however; Michelin which has a low mass, also has a low density in data set 2, but the second highest density for data set 1. Yarlington, being generally one of the smaller varieties, comes out as the least dense. The hypothesis that a small apple is also a dense apple therefore does not always seem to apply.

Figure 1.4

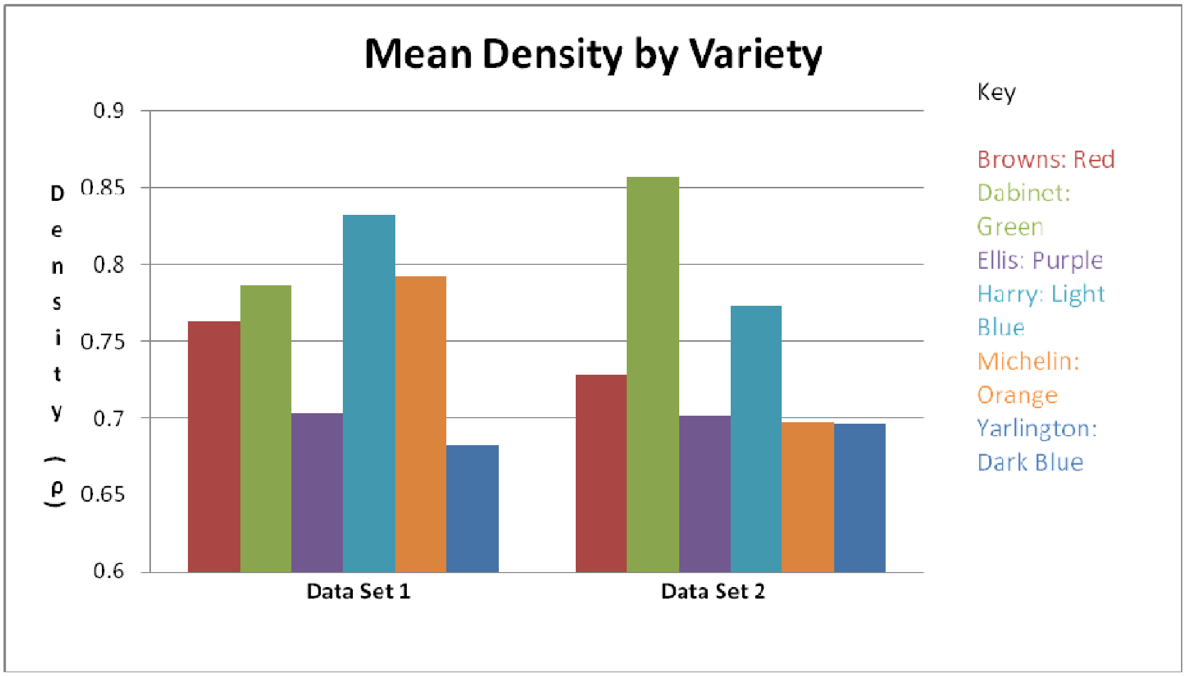
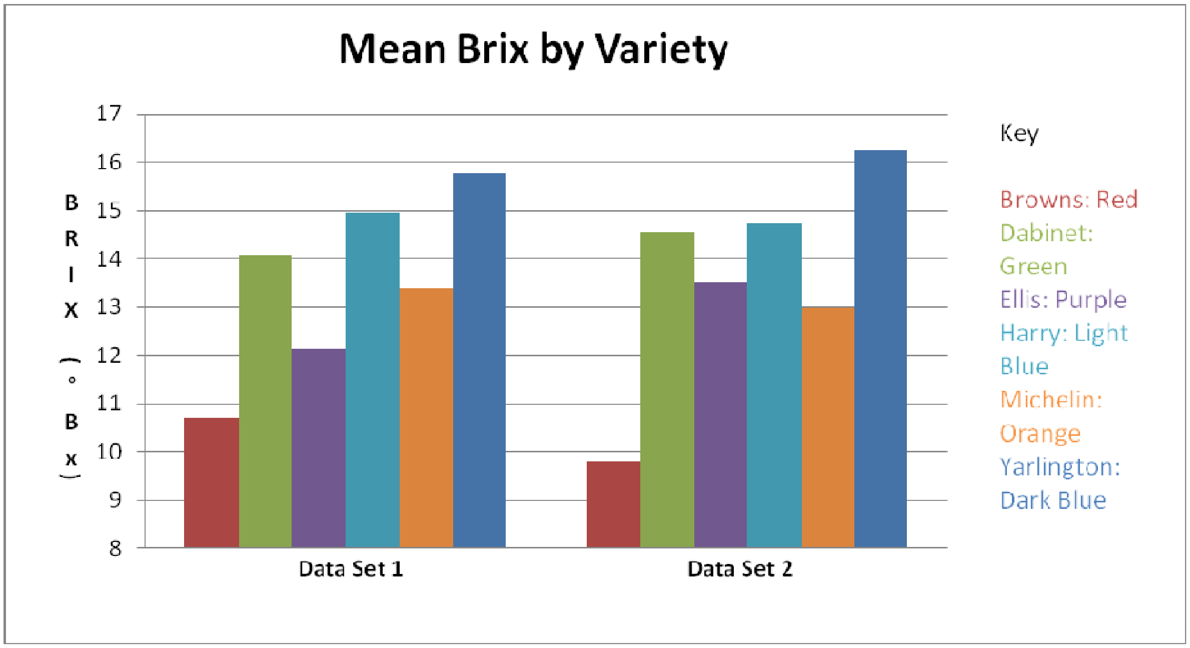


Figure 1.5 compares the mean sugar content in degrees of BRIX between varieties. It shows that despite being the least dense on average, Yarlington apples have the highest levels of sugar. Browns, despite being one of the smaller varieties, contain by far the lowest sugar levels. Again this would go against the hypothesis that a small apple has a higher sugar content.

Figure 1.5



Variable Dependencies

In order to test the hypotheses that smaller apples are denser than larger apples and that denser apples have a higher sugar content as a percentage of weight, then it is necessary to ascertain whether there is a statistical correlation between variables. In essence, do apples become less dense as they increase in size and do they have a higher BRIX level the denser they get? Table 1 takes the means of every lot of 5 apples for data sets 1 and 2 and calculates the Spearmans Rank Correlation Coefficient to determine whether any variable is dependent on another. The data highlighted green shows those variables which correlate with each other. As expected, the variables dependent on each other are related to mass, diameter and volume, however there is no correlation between any of these 3 variables with density or BRIX or between density and BRIX. Therefore the hypotheses can be rejected.

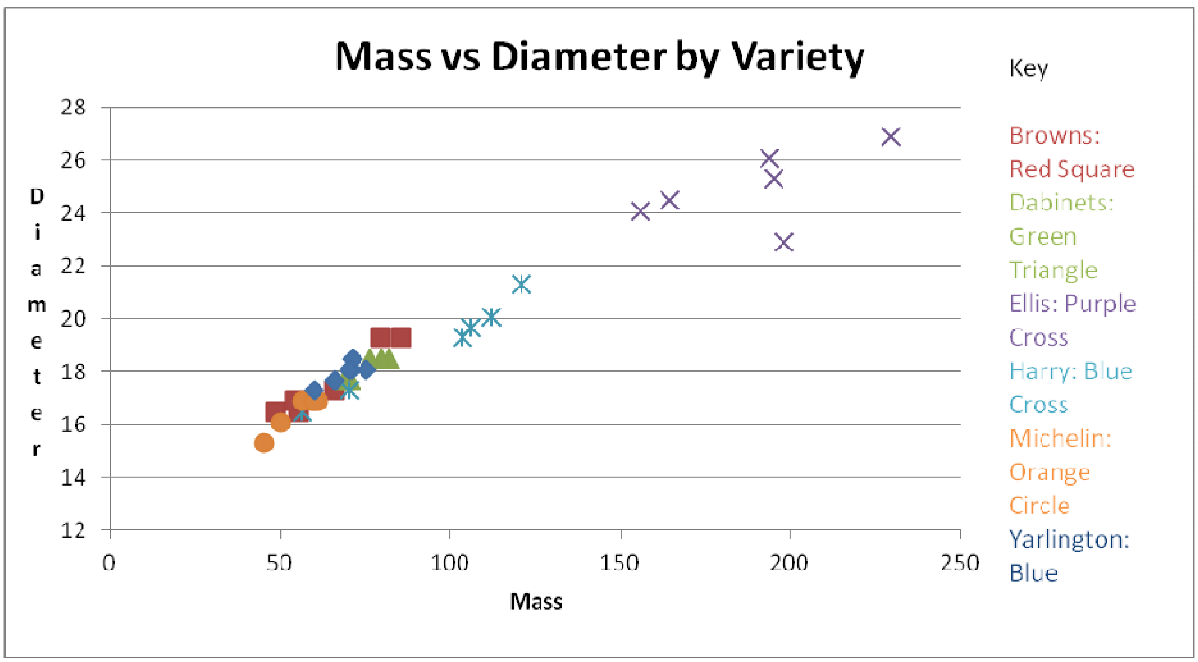
Table 1. Spearmans Rank Correlation Coefficient calculated from the Mean of each lot of 5 apples.

	Density	Volume	BRIX	Mass	Diameter
Density		-1.808	0.023	0.023	-0.036
Volume	-1.808		0.144	0.954	0.95
BRIX	0.023	0.144		0.116	0.115
Mass	0.023	0.954	0.116		0.981
Diameter	-0.036	0.95	0.115	0.981	

Table 1 proves there is no correlation between the size of an apple and its density or sugar content. The following graphs were drawn in order to show the differences in variables between the varieties. The graphs compare the means of each variable for the individual lots of 5 apples. There are 3 lots per data set; therefore there are 6 means in total for each variety plotted on the graphs. These data were chosen so that the graphs would show up any correlations and any interesting clusters of data for each variety.

Figure 2.1 shows the correlation between mass and diameter. The Spearmans Rank correlation coefficient measures this as 0.981 (see table 1), which is a strong positive correlation. This shows that as mass increases so does diameter. This is as expected and similar graphs could be drawn to show correlations between mass and volume and diameter and volume. The graph also shows that Ellis Bitter are much larger than the other varieties, with Harry Masters also standing apart from the others.

Figure 2.1



Figures 2.2 to 2.4 show volume, mass and diameter plotted against density. The similar patterns of the 3 graphs are due to the close correlation between the 3 variables. They show some interesting differences between the varieties. Despite little variation in sizes, Michelin and Dabinets show large variations in density, Ellis Bitters however have similar densities despite variations in their sizes. This shows that apple size affects density differently between varieties. The hypothesis would expect there to be a negative correlation between

size and density in that as mass, volume, and diameter go up, density would decrease. This however is not the case for any of the varieties and spearmans rank (see table 1) shows no correlation between density and any of the other 3 variables in figures 2.2 to 2.4.

Figure 2.2

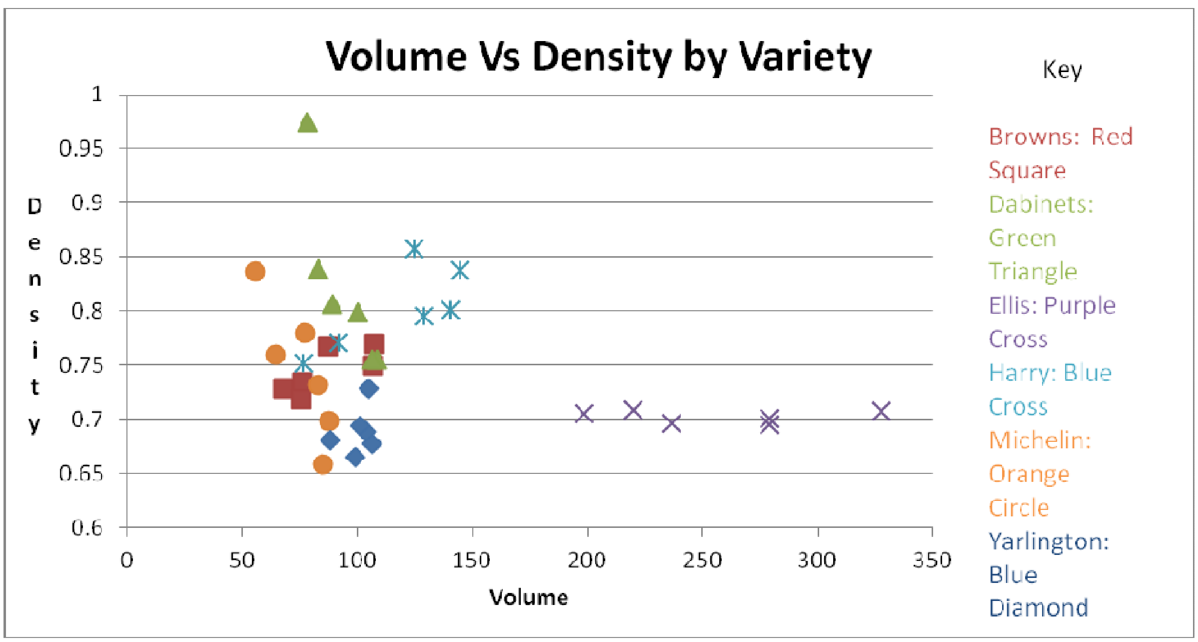


Figure 2.3

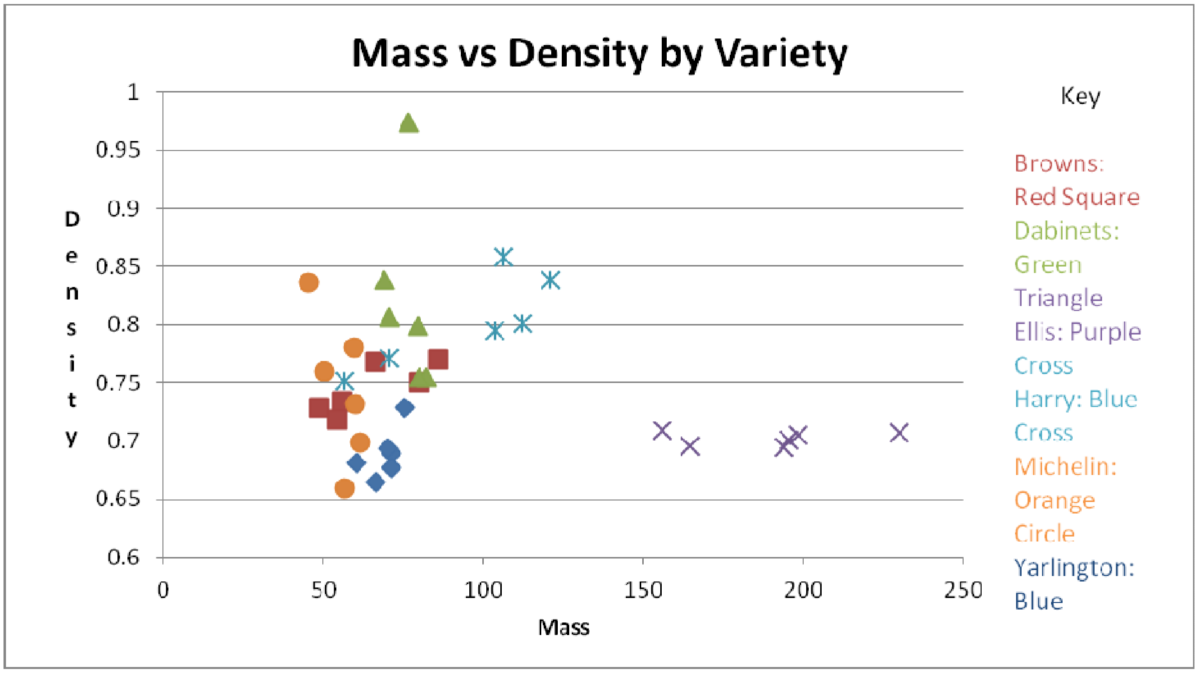
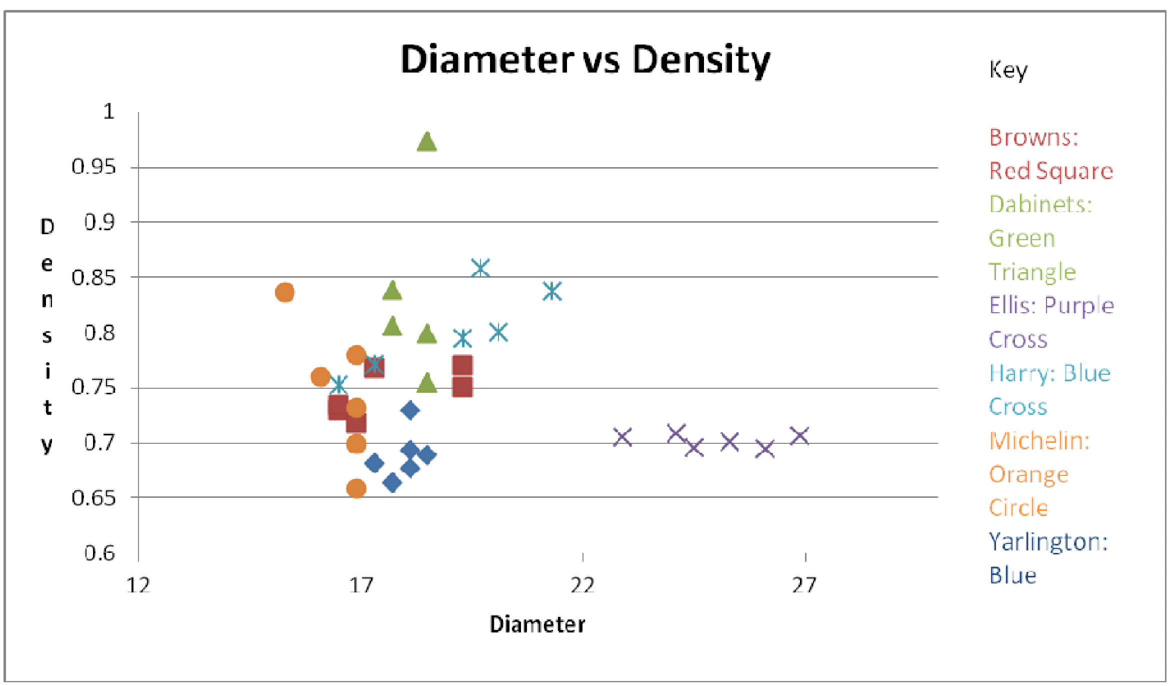

Mass vs Density by Variety
1
0.95
0.9
0.85
0.8
0.75
0.7
0.65
0.6
Density
0
50
100
150
200
250
Mass
Key
Browns: Red Square
Dabinets: Green Triangle
Ellis: Purple Cross
Harry: Blue Cross
Michelin: Orange Circle
Yarlington: Blue

Figure 2.4


Diameter vs Density
1
0.95
0.9
0.85
0.8
0.75
0.7
0.65
0.6
Density
12
17
22
27
Diameter
Key
Browns: Red Square
Dabinets: Green Triangle
Ellis: Purple Cross
Harry: Blue Cross
Michelin: Orange Circle
Yarlington: Blue

Figures 2.5 to 2.7 show volume, mass and diameter plotted against BRIX. Again the similar patterns are a result of the close correlation between the 3 variables. As already established in table 1, there is no correlation between the size (volume, mass or diameter) of an apple and its BRIX level. These graphs generally show that the BRIX levels do not vary significantly within varieties, regardless of size. Here we see that Yarlingtons have a low mass, volume and diameter and a high sugar level and that Browns also have a low mass, volume and diameter but a low sugar level. Ellis Bitters and Michelins have similar BRIX levels but are very different in size.

Figure 2.5

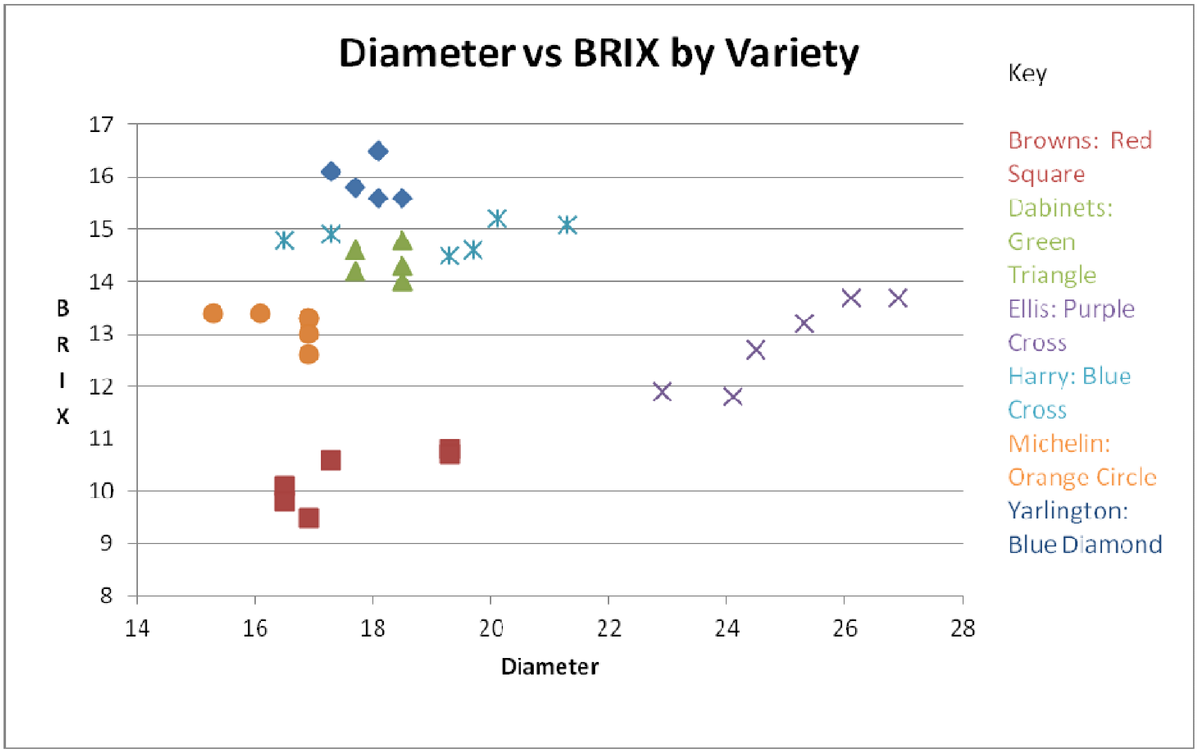


Figure 2.6

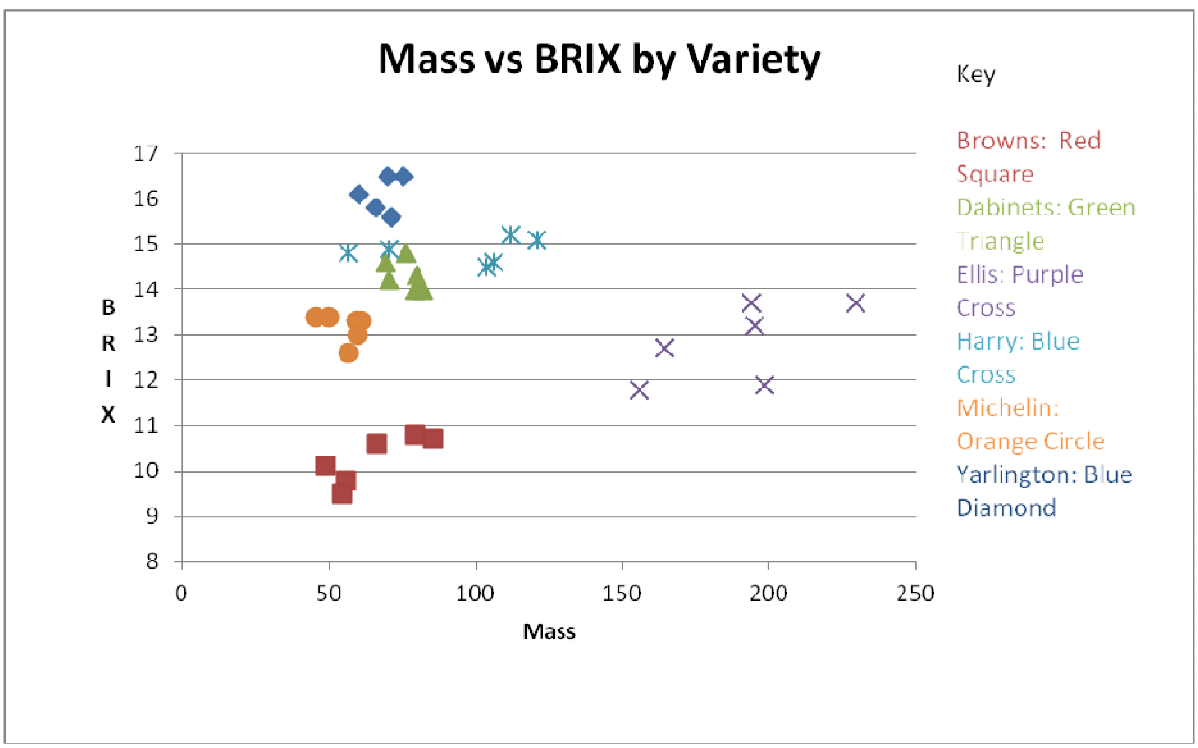


Figure 2.7

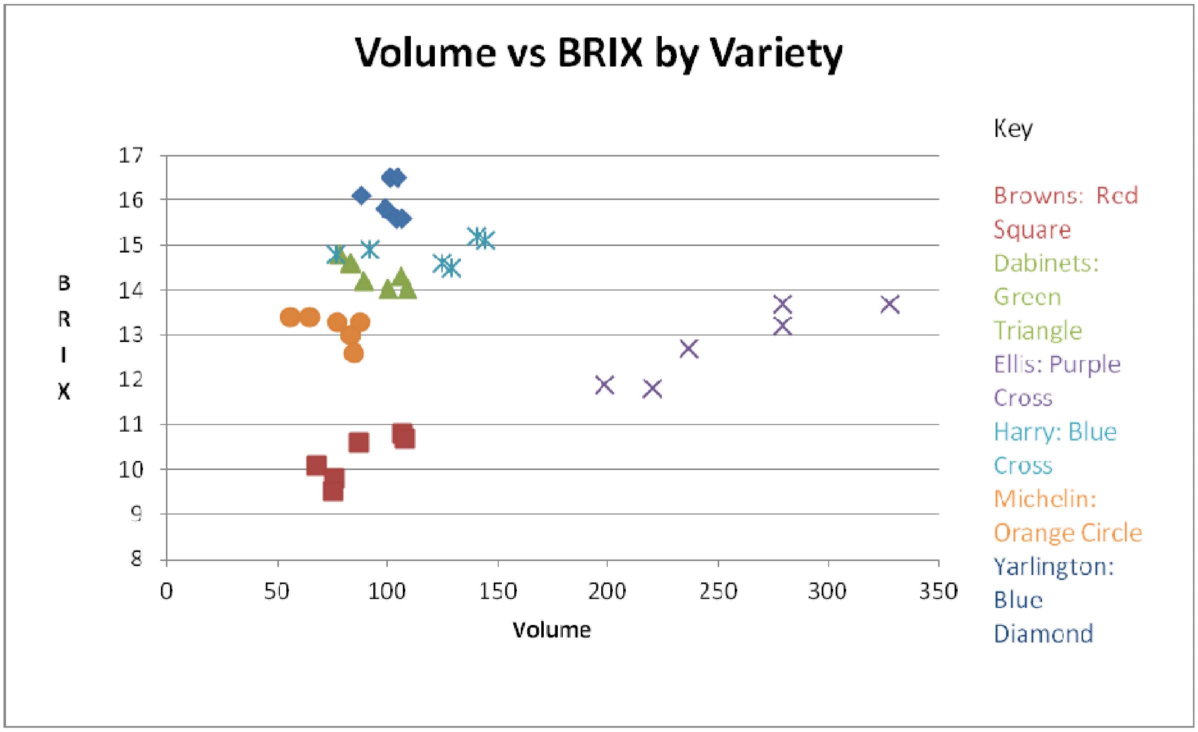
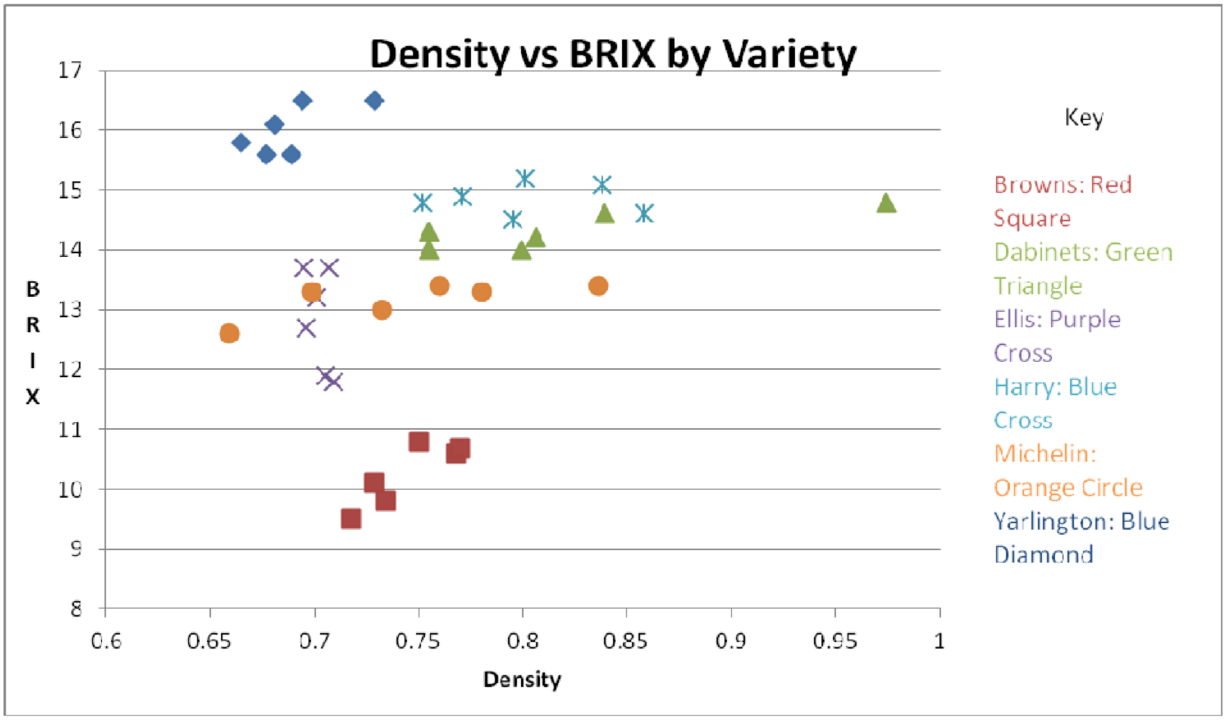


Figure 2.8 show the lack of correlation between density and BRIX. It also shows that certain varieties tend to vary in density more than others. For example, Dabinets and Michelins have a wider range of densities than Browns or Yarlingtons, but all varieties tend to have a similar BRIX to other members of that variety.

Figure 2.8



Student T-Test

We have already established that there are no correlations between apple size and density, or apple density and BRIX level. In order to determine whether there is a statistical difference between the mean densities and BRIX levels of each variety, it is necessary to perform the student t-test. Table 2 shows whether the difference in the mean densities of the varieties in data set 1 are statistically significant. Those numbers coloured green have extremely to very statistically different mean densities. This proves that certain varieties will always on average be more or less dense than others. For example Yarlingtons are significantly less dense than all other varieties apart from Ellis Bitters, although the 2 tailed P value of 0.057 is close to being significantly different.

Table 2. Unpaired T-test results showing 2 Tailed P value comparing each varieties mean density from Data Set 1.

	Browns	Dabinets	Ellis	Harry	Michelin	Yarlington
Browns		0.267	0.0001	0.0005	0.2809	0.0001
Dabinets	0.267		0.0001	0.048	0.8601	0.0001
Ellis	0.0001	0.0001		0.0001	0.001	0.057
Harry	0.0005	0.048	0.0001		0.1497	0.0001
Michelin	0.2809	0.8601	0.001	0.1497		0.0002
Yarlington	0.0001	0.0001	0.057	0.0001	0.0002	

In order to better visualise the difference between varieties densities and BRIX levels, Figure 3.1 plots density against BRIX for individual apple data from data set 1. Here we can see how the densities of certain varieties overlap such as Dabinets and Browns, whereas others such as Yarlingtons and Harry Masters do not, confirming the T-test results from table 2. It also shows that individual apples of a certain variety may have a very different density to another individual of the same variety, but will more than likely have a similar BRIX value. This shows up on the graph as horizontal lines of the same variety. It is very obvious that the BRIX for each variety is significantly different from the other varieties, with no overlapping of BRIX levels at all. This can be confirmed by performing the student t-test on BRIX levels.

Figure 3.1

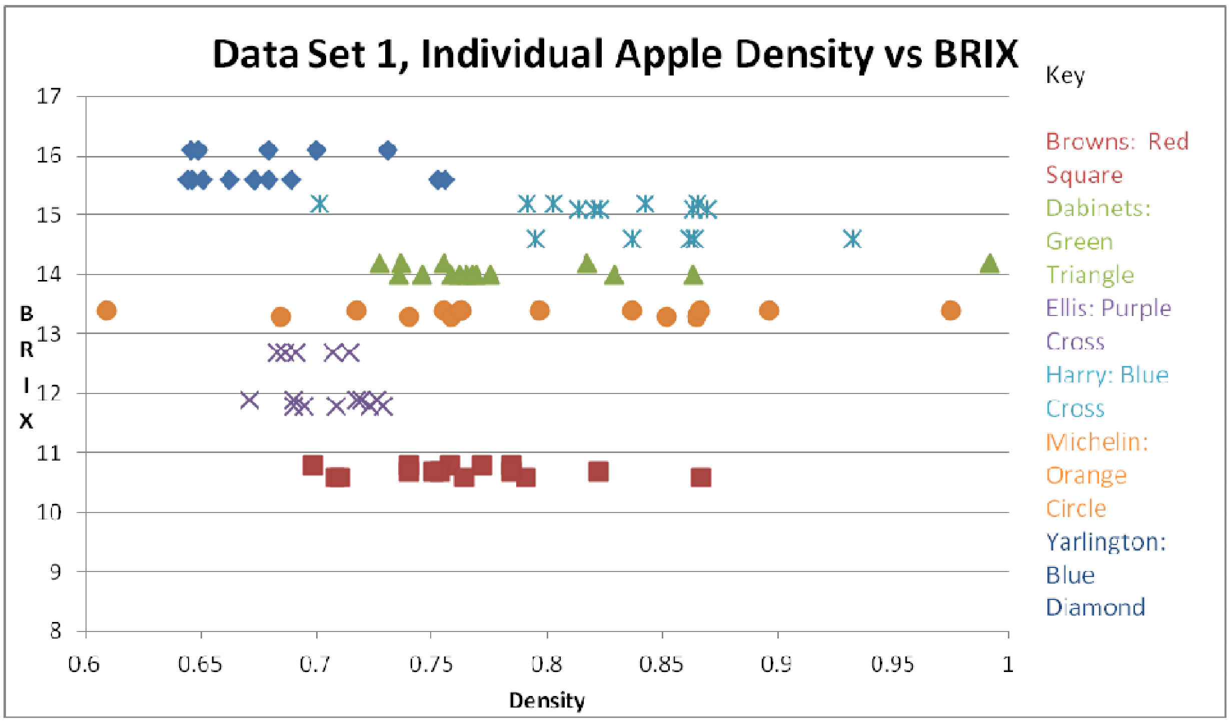
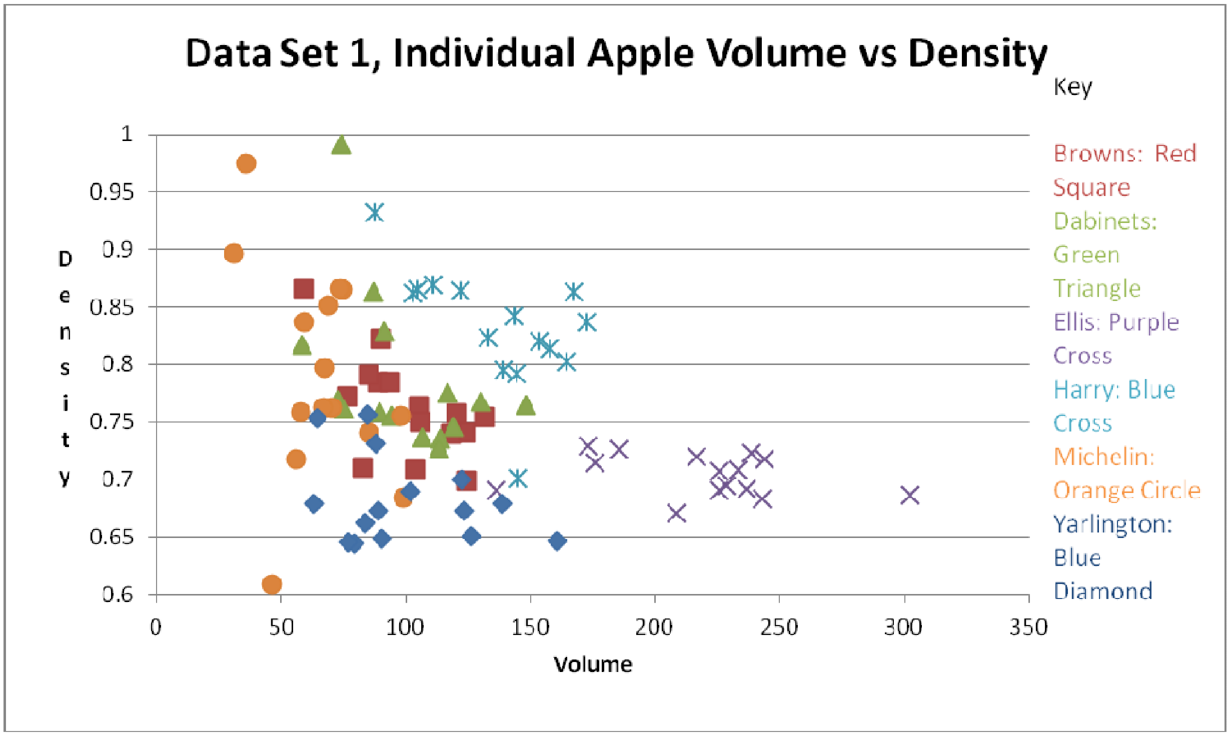


Figure 3.2 shows individual apple volumes plotted against density from data set 1. Here we can see some interesting clusters for each variety, for example Michelins can vary greatly in density, but tend to be of similar size, whereas Ellis Bitters can vary greatly in size, but tend to be of the same density.

Figure 3.2



Conclusions

In conclusion, the two hypotheses that smaller apples are denser apples and that denser apples have a higher sugar content as a percentage of mass have been disproved. There are however some interesting differences between the 6 varieties tested. BRIX levels between apples of the same variety will be similar regardless of size or density (see figure 2.5 and 3.1). Also apples of the same variety may be of similar size but vary in density; this is especially true for Dabinets and Michelins (see figure 3.2). The densest apples on average tend to be Harry Masters which also have the second highest BRIX levels. Yarlington Mill have the highest BRIX levels but are also the least dense. In order to maximise BRIX levels and densities then Harry Masters would be the best apple to use.

Recommendations

This report was designed as an initial look at the data from an apple density trial. The analysis was designed to determine whether there were any correlations specifically between density, BRIX and size in apples. It was also designed to see how apple varieties vary in regards to density, BRIX and size. This was done primarily by looking at the mean data from each lot of apples to plot graphs and also looking at individual apple data from data set 1.

In order to confirm the results obtained, similar graphs could be plotted for individual apple data from data set 2, to look at BRIX levels against density, volume against density and to confirm the t-test results in table 2. Spearmans Rank results could also be obtained for individual apple data in order to confirm the results shown in table 1. Analysis also needs to be done comparing irrigated with non-irrigated apples.

Appendix

Data Used to Create Figures 1.1 to 2.8 and Table 1.

Variety	Mean Density	Mean Volume	Mean Brix	Mean Mass	Mean DM
Browns Data Set 1 Lot B1	0.75	106.7	10.8	79.66	19.3
Browns Data Set 1 Lot B2	0.768	87.16	10.6	66.22	17.3
Browns Data Set 1 Lot B3	0.77	107.7	10.7	85.58	19.3
Browns Data Set 2 Lot B1	0.729	67.8	10.1	48.6	16.5
Browns Data Set 2 Lot B2	0.734	75.82	9.8	55.54	16.5
Browns Data Set 2 Lot B3	0.718	75.6	9.5	54.36	16.9
Dabinets Data Set 1 Lot D1	0.755	108.6	14	81.88	18.5
Dabinets Data Set 1 Lot D2	0.799	99.88	14	79.54	18.5
Dabinets Data Set 1 Lot D3	0.806	89.16	14.2	70.52	17.7
Dabinets Data Set 2 Lot D1	0.75477	106.28	14.3	79.86	18.5
Dabinets Data Set 2 Lot D2	0.839	83.12	14.6	69	17.7
Dabinets Data Set 2 Lot D3	0.974	78.16	14.8	76.34	18.5
Ellis Data Set 1 Lot E1	0.696	236.8	12.7	164.5	24.5
Ellis Data Set 1 Lot E2	0.705	198.2	11.9	198.2	22.9
Ellis Data Set 1 Lot E3	0.709	220	11.8	155.8	24.1
Ellis Data Set 2 Lot E1	0.701	279.06	13.2	195.15	25.3
Ellis Data Set 2 Lot E2	0.695	278.98	13.7	193.8	26.1
Ellis Data Set 2 Lot E3	0.707	327.7	13.7	229.58	26.9
Harry Data Set 1 Lot H1	0.858	124.6	14.6	106	19.7
Harry Data Set 1 Lot H2	0.801	140.4	15.2	111.9	20.1
Harry Data Set 1 Lot H3	0.838	144.4	15.1	120.8	21.3
Harry Data Set 2 Lot H1	0.771	91.7	14.9	70.28	17.3
Harry Data Set 2 Lot H2	0.795	129.02	14.5	103.48	19.3
Harry Data Set 2 Lot H3	0.752	76.54	14.8	56.62	16.5
Mich Data Set 1 Lot M1	0.836	55.92	13.4	45.22	15.3
Mich Data Set 1 Lot M2	0.78	77	13.3	59.5	16.9
Mich Data Set 1 Lot M2	0.76	64.8	13.4	49.98	16.1
Mich Data Set 2 Lot M1	0.659	85.08	12.6	56.5	16.9
Mich Data Set 2 Lot M2	0.699	87.56	13.3	61.2	16.9
Mich Data Set 2 Lot M3	0.732	82.92	13	59.94	16.9
Yarl Data Set 1 Lot Y1	0.689	103.8	15.6	71.26	18.5
Yarl Data Set 1 Lot Y2	0.677	106.5	15.6	71.14	18.1
Yarl Data Set 1 Lot Y3	0.681	88.1	16.1	60.2	17.3
Yarl Data Set 2 Lot Y1	0.66459	98.94	15.8	66.04	17.7
Yarl Data Set 2 Lot Y2	0.694	101.18	16.5	70.16	18.1
Yarl Data Set 2 Lot Y3	0.729	104.74	16.5	75.3	18.1

Individual Apple Data from Data Set 1 Used to Create Figures 3.1, 3.2 and Table 2

Variety	Density	Volume	BRIX
Browns	0.69855	124.4	10.8
Browns	0.77165	76.2	10.8
Browns	0.74032	124	10.8
Browns	0.78442	88.6	10.8
Browns	0.75768	120.5	10.8
Browns	0.86655	59.2	10.6
Browns	0.70837	103.9	10.6
Browns	0.79108	85.2	10.6
Browns	0.7103	82.5	10.6
Browns	0.76381	105	10.6
Browns	0.74005	118.1	10.7
Browns	0.7538	131.6	10.7
Browns	0.82254	89.6	10.7
Browns	0.78442	93.7	10.7
Browns	0.75047	105.4	10.7
Dabinets	0.7355	113.8	14
Dabinets	0.76486	148	14
Dabinets	0.74601	118.9	14
Dabinets	0.75839	89.4	14
Dabinets	0.76923	72.8	14
Dabinets	0.82895	91.2	14
Dabinets	0.76203	74.8	14
Dabinets	0.7753	116.6	14
Dabinets	0.76769	130	14
Dabinets	0.8629	86.8	14
Dabinets	0.99187	73.8	14.2
Dabinets	0.72751	113.4	14.2
Dabinets	0.73664	106.7	14.2
Dabinets	0.81724	58	14.2
Dabinets	0.75506	93.9	14.2
Ellis	0.69104	236.6	12.7
Ellis	0.7144	175.8	12.7
Ellis	0.70695	225.9	12.7
Ellis	0.68651	302.4	12.7
Ellis	0.68285	243.1	12.7
Ellis	0.71956	216.8	11.9
Ellis	0.67082	208.7	11.9
Ellis	0.69016	136.2	11.9
Ellis	0.72629	185.6	11.9
Ellis	0.71751	243.9	11.9

Ellis	0.7232	238.8	11.8
Ellis	0.72906	173.1	11.8
Ellis	0.70848	233.6	11.8
Ellis	0.69013	225.9	11.8
Ellis	0.69453	228.5	11.8
Harry	0.79511	139.1	14.6
Harry	0.83701	172.4	14.6
Harry	0.86371	121.8	14.6
Harry	0.86173	102.7	14.6
Harry	0.93234	87.2	14.6
Harry	0.80255	164.6	15.2
Harry	0.79155	144.4	15.2
Harry	0.70138	145	15.2
Harry	0.8424	143.4	15.2
Harry	0.86507	104.5	15.2
Harry	0.81345	157.6	15.1
Harry	0.86296	167.1	15.1
Harry	0.86925	110.9	15.1
Harry	0.8202	153.5	15.1
Harry	0.82291	132.7	15.1
Michelin	0.97479	35.7	13.4
Michelin	0.75535	98.1	13.4
Michelin	0.8961	30.8	13.4
Michelin	0.8367	59.4	13.4
Michelin	0.71763	55.6	13.4
Michelin	0.74031	85.1	13.3
Michelin	0.68453	98.9	13.3
Michelin	0.86479	74.7	13.3
Michelin	0.85174	68.8	13.3
Michelin	0.75826	57.7	13.3
Michelin	0.8663	73.3	13.4
Michelin	0.76286	70	13.4
Michelin	0.76269	67	13.4
Michelin	0.79674	67.4	13.4
Michelin	0.60907	46.3	13.4
Yarlington	0.75586	84.54	15.6
Yarlington	0.67939	138.8	15.6
Yarlington	0.66228	83.5	15.6
Yarlington	0.67306	88.7	15.6
Yarlington	0.67287	123.5	15.6
Yarlington	0.68898	101.6	15.6
Yarlington	0.75272	64.3	15.6
Yarlington	0.65083	126.3	15.6
Yarlington	0.64614	160.8	15.6
Yarlington	0.64403	79.5	15.6

Yarlington	0.67937	63	16.1
Yarlington	0.69992	122.3	16.1
Yarlington	0.73099	88.1	16.1
Yarlington	0.64583	76.8	16.1
Yarlington	0.64895	90.3	16.1