MATU9D2 Practical Statistics
Solutions to Minitab Questions in Practical 1.

Question 4 : Table is recreated as a check that we have the correct data!!

Tabulated statistics: Social Class, Voting
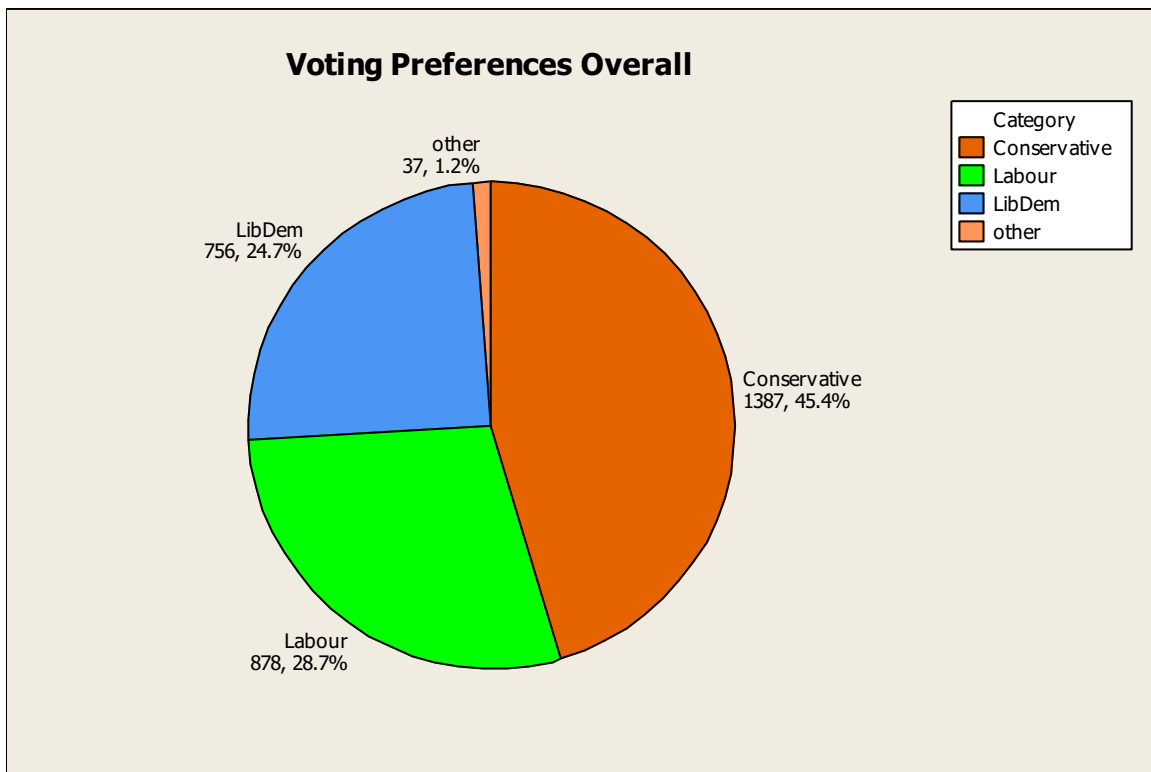
```
Rows: Social Class   Columns: Voting

        Conservative  Labour  LibDem  Other   All

A                110      11      37      2    160
B                420      99     237      8    764
C1               321     119     143     12    595
C2               353     363     205      9    930
D                183     286     134      6    609
All             1387     878     756     37   3058

Cell Contents:      Count
```
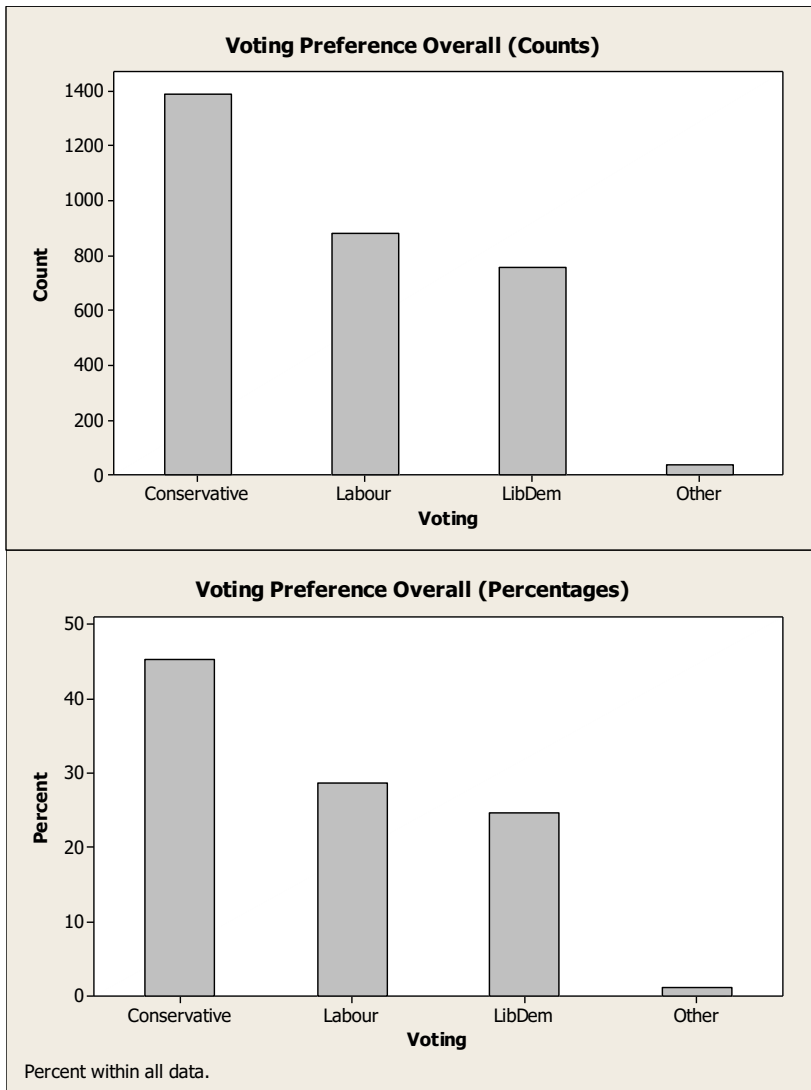
Question 5

Question 6

Voting Preference Overall (Counts)

Voting Preference Overall (Percentages)

Percent within all data.

Look identical except for vertical scale so if looking at one group – whether you look at percentages or counts we get the same picture

Largest group in this class vote Conservative (although it is not the overall majority!!)

Question 7

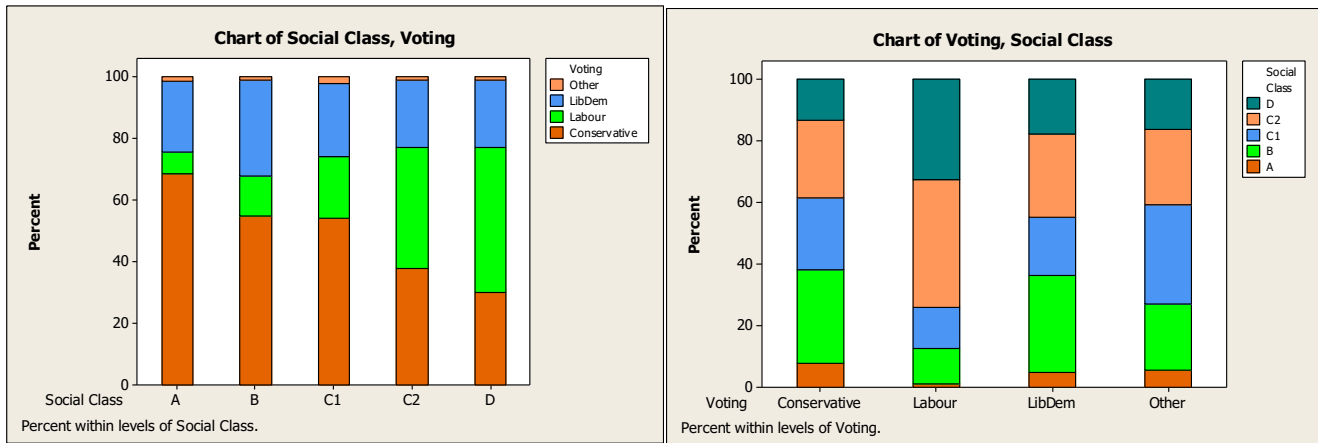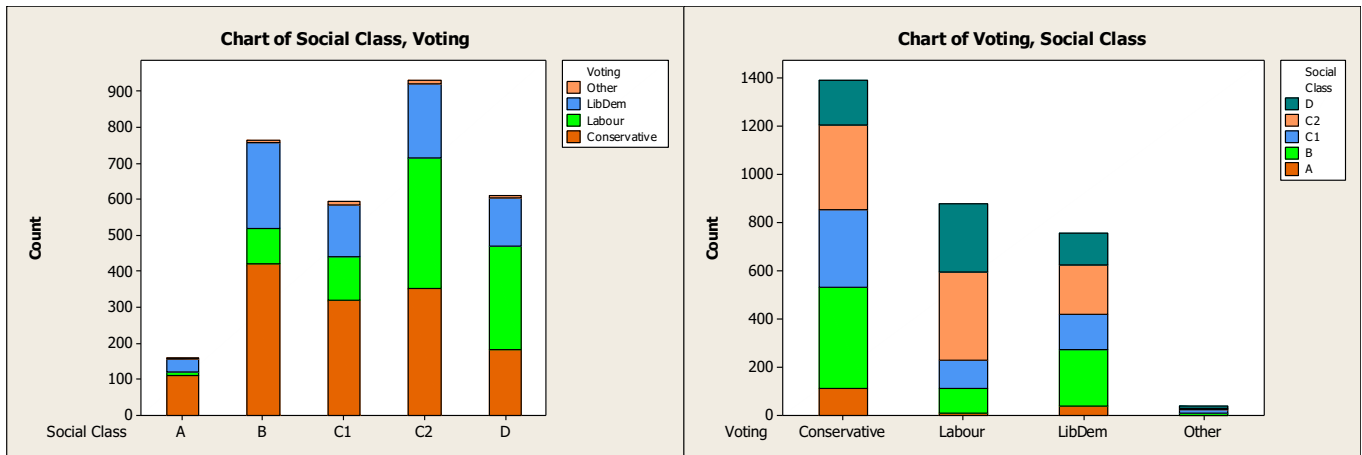## Chart of Voting, Social Class
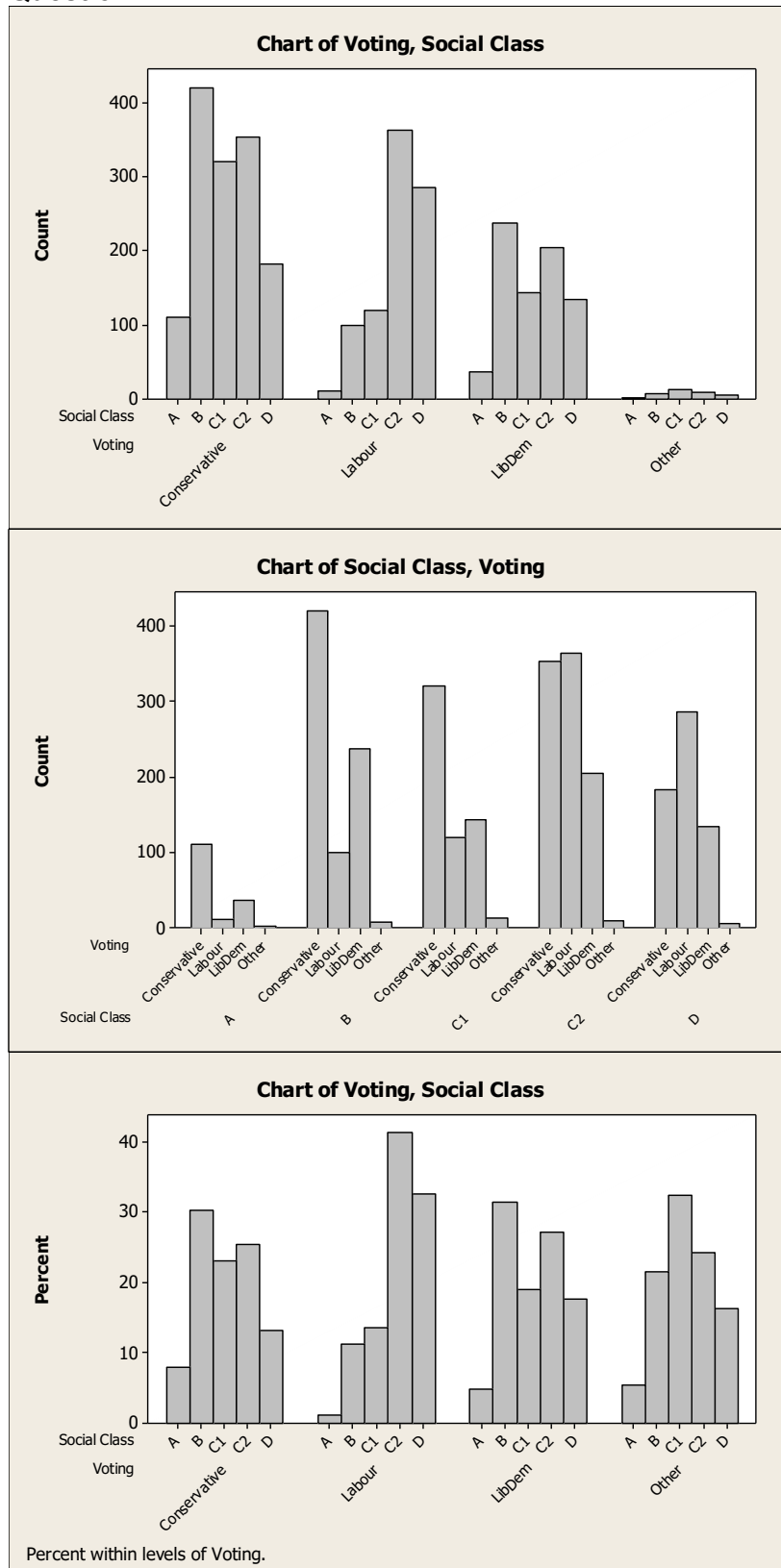
When comparing different groups with different sample sizes, then percentages should be used. Not counts as shown to the left.

Further you can plot either way round – see two possibilities to the left.

Chart of Social Class, Voting

Chart of Voting, Social Class

Percent within levels of Voting.

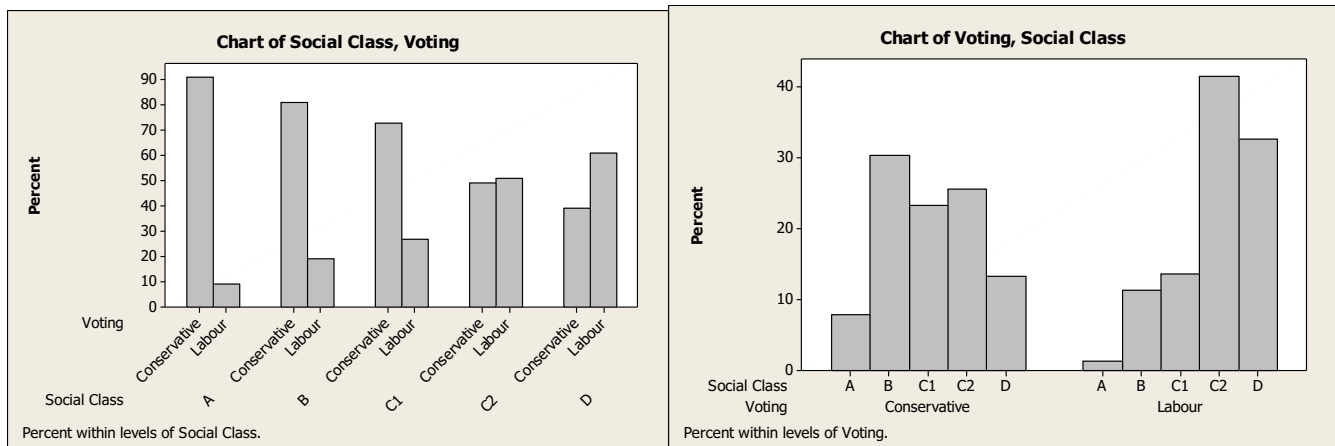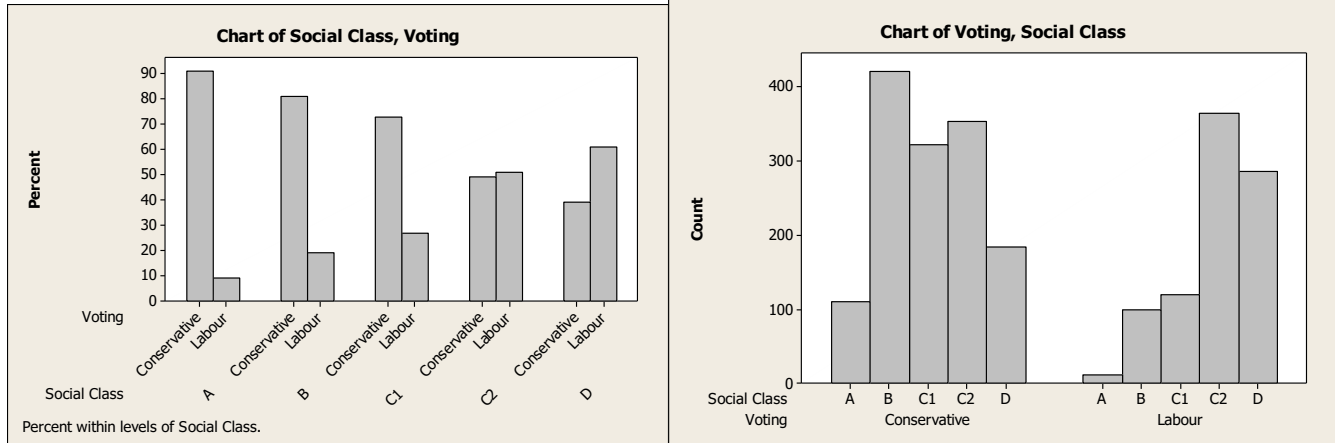Plotting using Percentages - can plot either way round – see one possibilities to the left.

Further do not use the overall percentages because this gives the same results as the counts.

Question 8

Chart of Social Class, Voting

Chart of Voting, Social Class

Chart of Social Class, Voting
Percent within levels of Social Class.

Chart of Voting, Social Class
Percent within levels of Voting.

Question 9

Chart of Social Class, Voting

Percent within levels of Social Class.

Chart of Voting, Social Class

Chart of Social Class, Voting

Percent within levels of Social Class.

Chart of Voting, Social Class

Percent within levels of Voting.

Question 10   (i)
Stem-and-Leaf Display: Yield (kgs)

```
Stem-and-leaf of Yield (kgs)  N  = 30
Leaf Unit = 0.10


   1   14  0
   2   15  0
   3   16  0
   5   17  00
   8   18  000
  11   19  000
  15   20  0000
  15   21  000
  12   22  0000
   8   23  00
   6   24  000
   3   25
   3   26  00
   1   27
   1   28  0
```

Default Scale from Minitab – too many rows

Stem-and-Leaf Display: Yield (kgs)

```
Stem-and-leaf of Yield (kgs)  N  = 30
Leaf Unit = 1.0


   2   1  45
   5   1  677
  11   1  888999
  (7)  2  0000111
  12   2  222233
   6   2  444
   3   2  66
   1   2  8
```

Change Scale – Increment of 2

Question 10   (ii)
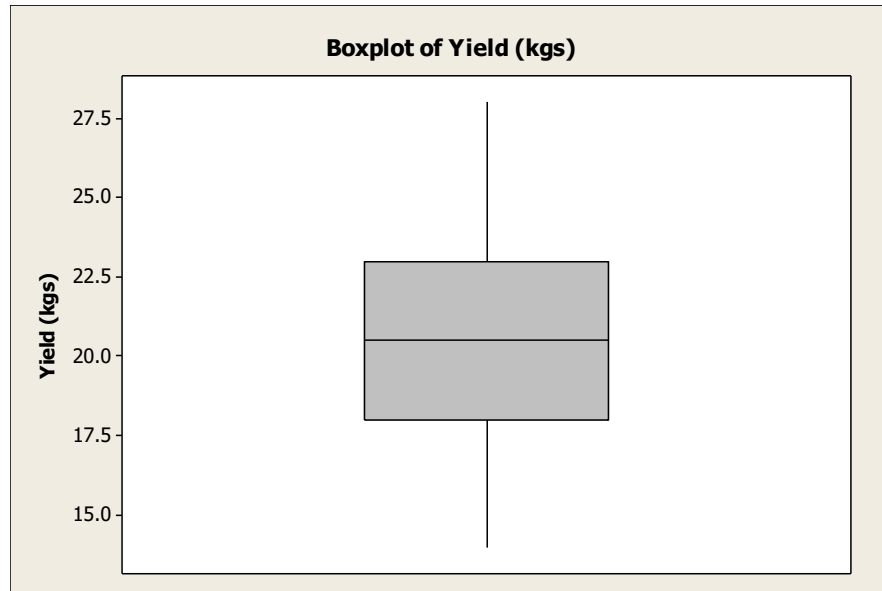**Descriptive Statistics: Yield (kgs)**

```
Variable       N  N*    Mean  SE Mean  StDev  Minimum      Q1  Median       Q3
Yield (kgs)   30   0  20.633    0.605  3.316   14.000  18.000  20.500   23.000

Variable      Maximum
Yield (kgs)    28.000
```

Use Descriptive Statistics and read off Minimum = 14, Q1 = 15, Median = 20.5, Q3 = 23 and Maximum = 28

Question 10   (iii)

Boxplot of Yield (kgs)

Yield (kgs)
- 27.5
- 25.0
- 22.5
- 20.0
- 17.5
- 15.0

Question 11

Use the same exploratory data analysis techniques as Question 10 to provide a Subjective Impression to answer the question

Stem-and-Leaf Display: Rainfall (mm)

```
Stem-and-leaf of Rainfall (mm)  N  = 9
Leaf Unit = 0.10


  1    10  7
  2    10  9
  3    11  1
  4    11  3
 (1)   11  4
  4    11  67
  2    11
  2    12  0
  1    12  2
```

Too many rows

Stem-and-Leaf Display: Rainfall (mm)

```
Stem-and-leaf of Rainfall (mm)   N  = 9
Leaf Unit = 0.10


  2   10   79
 (3)  11   134
  4   11   67
  2   12   02
```
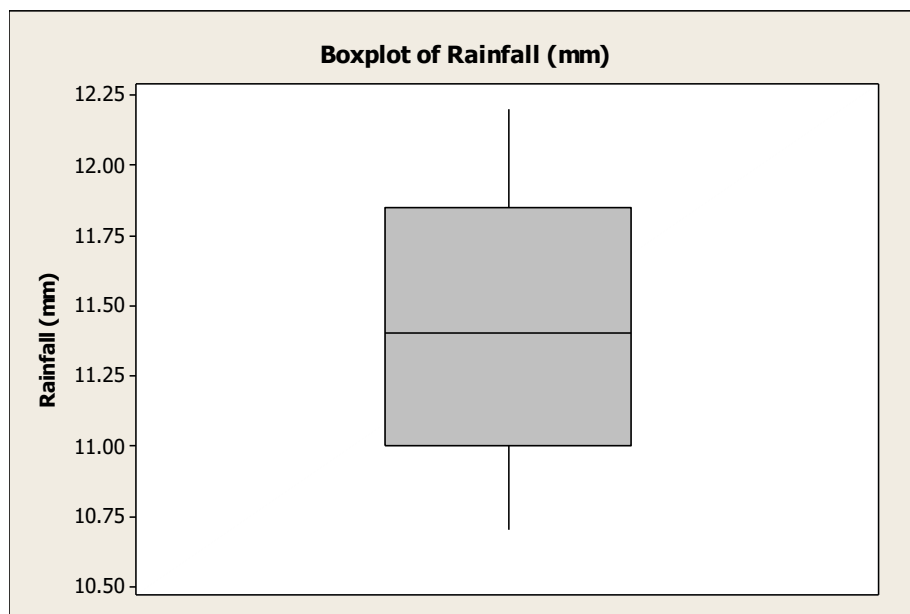
Change Scale – Increment of 0.5

Summary for Rainfall (mm)

Descriptive Statistics: Rainfall (mm)

```
Variable          N  N*     Mean  SE Mean  StDev  Minimum      Q1  Median      Q3
Rainfall (mm)  9   0   11.433    0.165  0.495   10.700  11.000  11.400  11.850

Variable        Maximum
Rainfall (mm)    12.200
```


Boxplot of Rainfall (mm)

Subjective Impression : Sample Size small so difficult to tell the 'shape' of the distribution from stem and leaf. However, looks fairly symmetric from Boxplot.

So, if single peaked symmetric with no outliers then appropriate measures of location and spread are 11.43mm and 0.495mm respectively.

Observed mean is 11.433 and sd is 0.495 and maximum is 12.2 so this suggests thet the mean is likely to be different from 12.1mm

Question 12

Stem-and-Leaf Display: Weight (kgs)

```
Stem-and-leaf of Weight (kgs)  N  = 36
Leaf Unit = 0.10


  2   2  23
  3   2  4
  6   2  677
 10   2  8899
 15   3  00111
(4)   3  2233
 17   3  4445555
 10   3  677
  7   3  889
  4   4  01
  2   4  3
  1   4
  1   4  6
```

Too many rows

Stem-and-Leaf Display: Weight (kgs)

```
Stem-and-leaf of Weight (kgs)  N  = 36
Leaf Unit = 0.10


  3   2  234
 10   2  6778899
(12)  3  001112233444
 14   3  5555677889
  4   4  013
  1   4  6
```

Change Scale – Increment of 0.5

Descriptive Statistics: Weight (kgs)

```
Variable       N  N*    Mean  SE Mean   StDev  Minimum      Q1  Median      Q3
Weight (kgs)  36   0  3.2861   0.0919  0.5514   2.2000  2.9000  3.3000  3.6750

Variable      Maximum
Weight (kgs)   4.6000
```
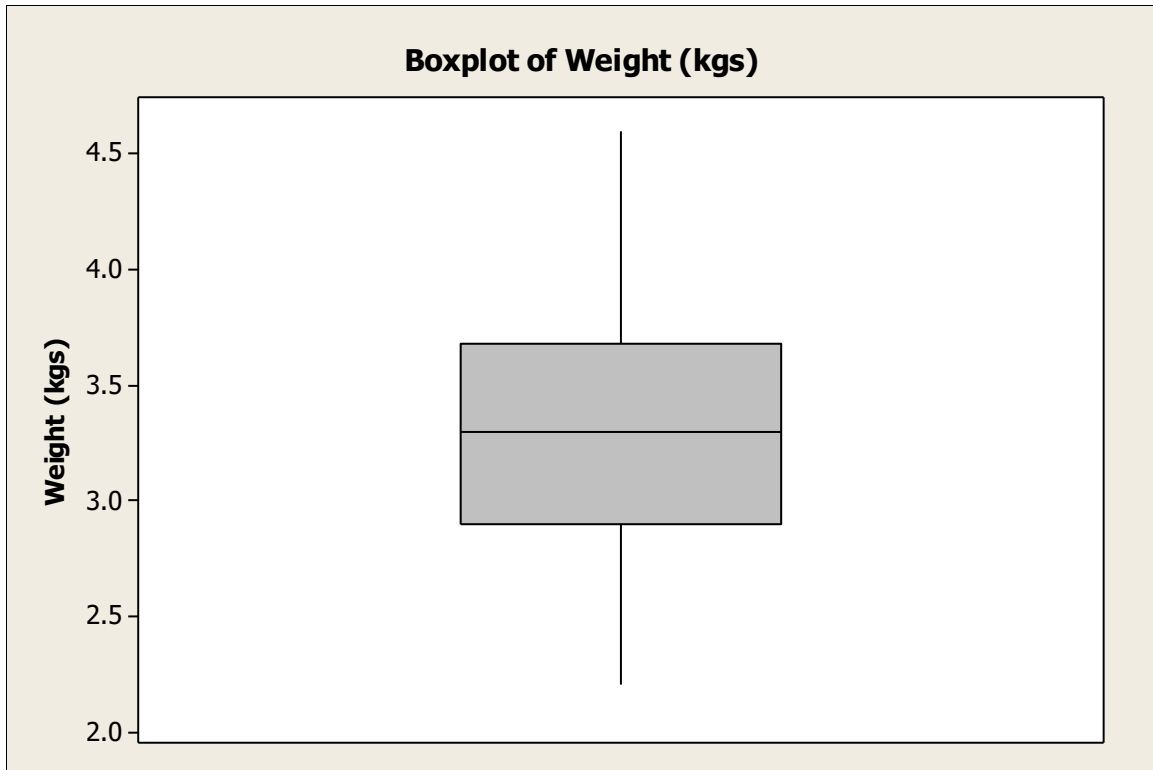
Subjective Impression : Sample Size reasonable this time so we can see that
    (a) Widely spread, (b) single peak, (c) symmetric, (d) no gaps, (e) no outliers

      So appropriate measures of location and spread are 3.29 kgs and 0.55kgs
      respectively.

This suggests that the mean may be different to 3.6kgs.

(Box and whisker plot – see below)

Boxplot of Weight (kgs)

Question 13

Stem-and-Leaf Display: Group 1 (ng/ml), Group 2 (ng/ml)

```
Stem-and-leaf of Group 1 (ng/ml)  N  = 10
Leaf Unit = 0.010

  1   32   0
  2   33   0
  3   34   0
  5   35   00
  5   36
  5   37   00
  3   38   0
  2   39
  2   40   0
  1   41   0
```

Too many rows

```
Stem-and-leaf of Group 2 (ng/ml)  N  = 10
Leaf Unit = 0.010


  1   28   0
  1   29
  2   30   0
  4   31   00
  5   32   0
  5   33   0
  4   34   00
  2   35
  2   36   00
```

Too many rows

Stem-and-Leaf Display: Group 1 (ng/ml), Group 2 (ng/ml)

```
Stem-and-leaf of Group 1 (ng/ml)  N  = 10
Leaf Unit = 0.10


  3    3   234
 (5)   3   55778
  2    4   01
```

Change increment to 0.5

```
Stem-and-leaf of Group 2 (ng/ml)  N  = 10
Leaf Unit = 0.10


  1    2   8
 (7)   3   0112344
  2    3   66
```

Boxplot of Group 1 (ng/ml), Group 2 (ng/ml)

Descriptive Statistics: Group 1 (ng/ml), Group 2 (ng/ml)

Variable	N	N*	Mean	SE Mean	StDev	Minimum	Q1	Median
Group 1 (ng/ml)	10	0	3.6200	0.0929	0.2936	3.2000	3.3750	3.6000
Group 2 (ng/ml)	10	0	3.2500	0.0820	0.2593	2.8000	3.0750	3.2500

Variable	Q3	Maximum
Group 1 (ng/ml)	3.8500	4.1000
Group 2 (ng/ml)	3.4500	3.6000

Subjective Impression : Sample Sizes small so difficult to tell the 'shape' of the distribution from stem and leaf. However, looks fairly symmetric from Boxplot.

So, if single peaked symmetric with no outliers then appropriate measures of location and spread are means and sd's respectively.

Observed means are 3.62ng/ml and 3.25ng/ml so with fairly small sd's this suggests a difference in the mean drug level in the two groups. This is also reflected in the side-by-side box and whisker plot below.

Boxplot of Group 1 (ng/ml), Group 2 (ng/ml)