Exercise 4.1

From this table, it is much easier to see that Leeds has the highest sales, and this exceeds the sales of Sheffield which is the city with the largest population. Sales for the rest of the cities seem to decrease with population.

<table>
<thead>
<tr>
<th>£'000</th>
<th>Quarter 1</th>
<th>Quarter 2</th>
<th>Quarter 3</th>
<th>Quarter 4</th>
<th>average</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sheffield</td>
<td>230</td>
<td>220</td>
<td>190</td>
<td>220</td>
<td>220</td>
</tr>
<tr>
<td>Leeds</td>
<td>280</td>
<td>260</td>
<td>220</td>
<td>340</td>
<td>270</td>
</tr>
<tr>
<td>Edinburgh</td>
<td>140</td>
<td>130</td>
<td>130</td>
<td>210</td>
<td>150</td>
</tr>
<tr>
<td>Hull</td>
<td>70</td>
<td>80</td>
<td>70</td>
<td>80</td>
<td>80</td>
</tr>
<tr>
<td>Swansea</td>
<td>60</td>
<td>70</td>
<td>60</td>
<td>80</td>
<td>70</td>
</tr>
<tr>
<td>Plymouth</td>
<td>40</td>
<td>40</td>
<td>30</td>
<td>50</td>
<td>40</td>
</tr>
<tr>
<td>Bolton</td>
<td>30</td>
<td>30</td>
<td>30</td>
<td>30</td>
<td>30</td>
</tr>
<tr>
<td>Luton</td>
<td>20</td>
<td>30</td>
<td>20</td>
<td>30</td>
<td>30</td>
</tr>
<tr>
<td>average</td>
<td>110</td>
<td>100</td>
<td>100</td>
<td>130</td>
<td></td>
</tr>
</tbody>
</table>

Exercise 5.1

1. Copy and paste the code (rubber.r) from the stat100 website into Rcmdr, highlight the code and submit.

2. make rubber the active data set (Data > Active Data set > Select active data set).

3. You can produce the histogram of relative frequencies (i.e., densities) using Graphs > Histogram and select the appropriate options. To fit the density curve you need to type in the following command:

   `lines(density(rubber$loss))`

4. For (c) and (d) of this question you need to calculate the differences between the observations and the mean, and then plot the ecdf of the differences.

   `xbar <- mean(rubber$loss)`
   `plot(ecdf(rubber$loss - xbar))`
```r
#___ rubber.r _________
options(digits=2)
loss <- c(372,206,175,154,136,112,55,45,221,166,164,113,82,32,228,196,128,97,64,
         249,219,186,155,114,341,340,283,267,215,148)
rubber <- data.frame(loss)
summary(rubber$loss)

hist(rubber$loss, prob=T)
lines(density(rubber$loss))

xbar <- mean(rubber$loss)
plot(ecdf(rubber$loss - xbar))

Exercise 5.2
Read in the data file and in Rcmdr select Graphs > Boxplot. Select heights, and Plot by Group: gender.

Exercise 5.3 The median is the midpoint 2.
Exercise 5.4 IQR = 9 − 5 = 4
Exercise 5.5

Exercise 5.6
Simply use Graphs > Histograms and plot by groups.

```r
#__________ wtg.r __________
options(digits=2)
wtg <- read.table("wtg.txt",header=T)
print(summary(wtg))

boxplot(wtgain~group,data=wtg,las=1)
```

Exercise 6.1
The explanatory variable is Flow

Exercise 6.2
The factors are Soil, which has 3 levels: basalt, granite and trap, and Enterprise, which has 3 levels: cattle, sheep and mixed.

Exercise 6.3
- Grade - 4,5,6
- Race - White, Other
- Region - Rural, Suburban, Urban
- Goals - Sports, Popular, Grades
- Grades - 1,2,3,4
Exercise 6.4

1. Don’t forget to declare \texttt{Year} as a factor. In Rcmdr use Data \textgreater Manage variables \textgreater convert numeric to factor.

2. In the first case use \texttt{Graphs} \textgreater \texttt{Scatterplot} and in the second use \texttt{Graphs} \textgreater \texttt{XY conditioning plot}

\begin{verbatim}
#_________ florets.r _______________
florets <- read.table("florets.txt",header=T)
lorets$Year <- factor(florets$Year)

library(lattice)
trellis.device(color=F)
florets.plot1 <- xyplot( Number ~ Temp,data=florets,
groups=Year,auto.key=list(x=0.2,y=0.8) )
print(florets.plot1)

trellis.device(color=F)
florets.plot2 <- xyplot( Number ~ Temp|Year,data=florets)
print(florets.plot2)
\end{verbatim}