1. Given a matrix $X$ of distinct rows and a vector $w$ of the number of times that each row should occur, reconstruct the original matrix.

2. Run the following code:

```r
gender <- factor(c(rep("female", 91), rep("male", 92)))
table(gender)
gender <- factor(gender, levels=c("male", "female"))
table(gender)
gender <- factor(gender, levels=c("Male", "female"))
  # Note the mistake
  # The level was "male", not "Male"

table(gender)
rm(gender)  # Remove gender
```

Explain the output from the final `table(gender)`.

3. Solve the following system of two linear equations in S-PLUS and report the result:

$$
\begin{align*}
3x + 4y &= 12 \\
x + 2y &= 8
\end{align*}
$$

4. The function `apply` will map a function to each row or column of a matrix or other array, returning the results in an appropriate object, which is a better alternative than loops.

The form is: `apply( array, margin, function, ...)`

Try to look at the help session in S-PLUS or any other source to learn this function as well as some other similar functions `lapply`, `sapply` and `tapply`.

The following questions are based on a built-in dataset in S-PLUS called `auto.stats`.

a. Create a vector containing the maximum value for each numeric variable in the matrix by using the function `apply` (remember to ignore those missing values).

b. Suppose we want to create a vector containing the mean value for each numeric variable:
Instead of using the function `apply`, try to get this result by using the function `sapply`.

**c.** Suppose we create a category object called `cargrp` by dividing the weight of the cars (8th column) into 3 intervals (0-2500, 2500-3500, 3500-5000):

```r
> cargrp <- cut(auto.stats[,8], c(0, 2500, 3500, 5000),
labels = c("Small", "Median", "Large"))
```
Calculate the median gas consumption (2nd column) by using the function `tapply`.

**d.** Calculate the mean gas consumption for different-sized cars, as described by the variable `cargrp`, broken down by the 1978 repair rating (3rd column) (Use the function `tapply`). You should get the following results:

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>NA</th>
</tr>
</thead>
<tbody>
<tr>
<td>Small</td>
<td>NA</td>
<td>NA</td>
<td>25.33333</td>
<td>25.75000</td>
<td>31.12500</td>
<td>26.00</td>
</tr>
<tr>
<td>Median</td>
<td>21</td>
<td>22.00</td>
<td>19.53333</td>
<td>22.66667</td>
<td>17.33333</td>
<td>20.25</td>
</tr>
<tr>
<td>Large</td>
<td>NA</td>
<td>16.25</td>
<td>15.33333</td>
<td>16.57143</td>
<td>NA</td>
<td>NA</td>
</tr>
</tbody>
</table>

**5.** Complete the following table and comment on the results. Please specify which version of Excel you used and which software you used to get the last column.

<p>| Inverse t distribution (Value of x with Pr(|X| &gt; x) = p) |
|-------------------|-------------------|-------------------|-------------------|-------------------|-------------------|</p>
<table>
<thead>
<tr>
<th>p</th>
<th>d.f.</th>
<th>Exact Value</th>
<th>MS Excel (9 d.p.)</th>
<th>S-PLUS (6 d.p.)</th>
<th>Other software (please specify)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.001</td>
<td>2</td>
<td>31.5991</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1E-6</td>
<td>2</td>
<td>999.999</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.001</td>
<td>5</td>
<td>6.86883</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1E-6</td>
<td>5</td>
<td>28.4785</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.001</td>
<td>10</td>
<td>4.58689</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1E-6</td>
<td>10</td>
<td>10.5165</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
6. Produce the following graph in S-PLUS. You are required to produce a graph looks as close as possible as this one.