1. Write a simple function to compute the median absolute deviation (used in robust statistics) median $|x - \mu|$ with default $\mu$ the sample median. Compare your answer with the system function mad.

2. The following code uses the for() looping function to plot graphs that compare the relative population growth (here, by the use of a logarithmic scale) for the Australian states and territories.

```r
par(mfrow = c(2,4))
for (i in 2:9) {
  plot(austpop[,1], log(austpop[,i]), xlab="Year", ylab=names(austpop)[i],
  pch = 16, ylim=c(0,10))
}
```

Can this be done without looping? [Hint: Use sapply or lapply]

3. Write a S-PLUS function that accepts two numeric vectors and calculates a two-sample $t$-test to determine if the sample means are equal. You must consider whether the samples are of equal size or not. You may also want to consider the situations for equal variances and unequal variances. The function should return the $P$-value of obtaining the $t$ statistic given the null hypothesis. (Write your own function, don’t use the built-in function in S-PLUS t.test)

4. Here is one way to compute the bivariate standard normal density function over a regular two-dimensional grid defined by x1gr in the x1 direction and x2gr in the x2 direction. The persp() function is then called to draw a perspective plot.

```r
> bivnormplot
function (x1gr, x2gr)
{
  f12 <- matrix(0, nrow = length(x1gr), ncol = length(x2gr))
  i1 <- 0
  for (x1 in x1gr) {
    i1 <- i1 + 1
    i2 <- 0
    for (x2 in x2gr) {
      i2 <- i2 + 1
      f12[i1, i2] <- (1/(2*pi))*exp(-0.5*(x1^2 + x2^2))
    }
  }
  persp(x1gr, x2gr, f12)
  invisible()
}
```
Here is how `bivnormplot()` is used in an example the grid goes from -3 to 3 in 40 steps in each direction:

```r
> bivnormplot(seq(-3,3,len=40), seq(-3,3,len=40))
```

**a.** Write comment statements for each line of the function `bivnormplot()`. Assume somebody who have no experience with S-PLUS is going to read your comment statements and try to understand what is this function doing.

**b.** Instead of using the above function, can you propose another way to compute the bivariate standard normal density function and produce the perspective plot. (Hints: make use of `outer` function in S-PLUS)

5. AAUP Faculty Salary Data

Data description:

Data are from the American Association of University Professors (AAUP) annual faculty survey of American colleges and universities. They include average salary and overall compensations. They include average salary and overall compensation, broken down by full, associate, and assistant professor ranks. The dataset is taken from the March - April 1994 issue of Academe.

Variable description:

<table>
<thead>
<tr>
<th>Name</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>FICE</td>
<td>FICE (Federal ID number)</td>
</tr>
<tr>
<td>College.Name</td>
<td>College name</td>
</tr>
<tr>
<td>State</td>
<td>State (postal code)</td>
</tr>
<tr>
<td>Type</td>
<td>Type* (I, IIA, or IIB)</td>
</tr>
<tr>
<td>AS.Full</td>
<td>Average salary – full professor</td>
</tr>
<tr>
<td>AS.Assoc</td>
<td>Average salary – associate professor</td>
</tr>
<tr>
<td>AS.Assis</td>
<td>Average salary – assistant professor</td>
</tr>
<tr>
<td>AS.All</td>
<td>Average salary – all ranks</td>
</tr>
<tr>
<td>AC.Full</td>
<td>Average compensation – full professor</td>
</tr>
<tr>
<td>AC.Assoc</td>
<td>Average compensation – associate professor</td>
</tr>
<tr>
<td>AC.Assis</td>
<td>Average compensation – assistant professor</td>
</tr>
<tr>
<td>AC.All</td>
<td>Average compensation – all ranks</td>
</tr>
<tr>
<td>N.Full</td>
<td>Number of full professor</td>
</tr>
<tr>
<td>N.Assoc</td>
<td>Number of associate professor</td>
</tr>
<tr>
<td>N.Assis</td>
<td>Number of assistant professor</td>
</tr>
<tr>
<td>N.Inst</td>
<td>Number of instructors</td>
</tr>
<tr>
<td>N.All</td>
<td>Number of faculty – all ranks</td>
</tr>
</tbody>
</table>

* Category I – Doctoral-Level Institutions, Category IIA – Comprehensive Institutions, Category IIB – General Baccalaureate Institutions.

Use S-PLUS to plot the graphs. Using \LaTeX, write up your results and findings in no more than 2 pages (excluding graphs and tables).
a. Calculate the median and median of absolute deviation (MAD) of the average faculty salary (for all ranks) for different types of colleges, broken down by states.

b. Make a scatter plot of the number of faculty (all ranks) against the average salary (all ranks). Comment on the pattern of the plot.

c. Make a scatter plot of the logarithm of the number of faculty (all ranks) against the average salary (all ranks). What is the effect of logarithm transformation?

d. Create a category object called `sizegrp` by dividing the number of faculty (all ranks) (N.All) into 4 intervals (0-60, 60-100, 100-300, 300-2300). Display the average salary (all ranks) on a comparative box plot against `sizegrp`.

e. A person claims that the faculty salary structure in California and Texas are different. Use graphical display and some basic statistics to justify the correctness of this claim. Is there any “outlier” for the faculty salary within these two states, if yes, identify them.