Introduction to R
Session 8: Introduction to R Packages

The `ggplot2` package is a widely-used data visualization package for enhanced graphics in R. In this session we will use `ggplot2` for plotting data from the salary dataset.

The two main "generic" plotting functions of `ggplot2` are `qplot()` and `ggplot()`. These functions work by trying to 'guess' a useful form of plot, depending on the user-supplied data and desired geometry. In addition to their help pages, detailed information about these (and other) plotting functions in this package can be found at http://ggplot2.org/ .

1. Using either the drop-down menus or the command line (with the `install.packages()` and `library()` commands) install this package and load it into your current R session.

2. Some `qplot()` commands to plot smooth density curves of salary for each gender are given below. Create a similar density plot of salary by rank.

```
qplot(salary, geom="density", fill=gender, data=salary, xlab="Monthly
Salary", ylab="Density")

qplot(salary, geom="density", fill=gender, data=salary, xlab="Monthly
Salary", ylab="Density", alpha=0.5)
```

3. By adjusting the `geom` argument – described in the documentation for `gplot()` – use `qplot()` to plot a histogram of salaries, and a histogram of salaries across rank. Finally, create a histogram of salary by gender within each rank (hint: use the `interaction()` command).

4. The `ggplot()` command extends the basic idea of qplot; the user has to specify 'aesthetic mappings' (via the `aes()` command) that describe how variables in the data are mapped to visual properties, after which other aspects can be added. For example, for a boxplot of salary by rank gender, we can use

```
ggplot(data=salary, aes(x = rank, y = salary, fill = gender) ) +
geom_boxplot()
```

   Another 'add-on' to the plot it to break it break it down by another variable, using `+facet_wrap(~variablename)`. Using the help page for `facet_wrap()` to help you, create boxplots of salary by gender within each rank across fields.

5. [For keen people!] Illustrate the distribution of salary by gender within each rank for the different starting years.