**Introduction to R Overview**

R has been selected by the NWU Predictive Analytics program, along with Python and SAS, as a programming language to be used in various courses.  PREDICT 401-DL uses R as a programming language in support of statistical analysis.  At least eight subsequent courses in the Predictive Analytics program will depend upon R. R provides a flexible capability for statistical analysis and graphical display of data.  It is ideal for data visualization and exploratory data analysis while at the same time it supports analyses ranging from elementary to complex.

The syllabus mentions three texts that serve as references for R applications to statistics.  The first by Verzani, *Using R for Introductory Statistics,* is a reference that uses base R and gives a variety of examples.  The book by Stowell, *Using R for statistics*, provides many illustrations of data analysis with R.  Also mentioned is the text by Lander, *R for everyone: Advanced analytics and graphics.*  This latter text illustrates a range of more advanced applications.  All three serve as reference aids for the course.

In addition to the texts, there are a number of tutorials and educational materials listed.  The *Quick Start Guide to R* is a console-based tutorial for those unfamiliar with the language.  SWIRL is similar.  There is also a manual providing optional lessons in R.  Each session optional lessons from this manual are mentioned.  Data sets accompany the lessons.  The lessons are numbered corresponding to chapters in Black, *Business Statistics.*  The lessons draw on course concepts and provide an opportunity to perform data analysis using R.  Suggestions are given with many of the problems.  The texts by Verzani, Stowell, and Lander serve as reference aids.  Answers and code solutions are provided.

The first step in getting started is to install R.  R can be obtained at [*http://cran.r-project.org/* (Links to an external site.)](http://cran.r-project.org/)*.* Instructions for doing so are given in the texts by Verzani and Lander.  RStudio is recommended for new users of R.  It is not required.  Everything students need to do in this course can be accomplished using the standard R console with a plain text editor.  RStudio is an integrated development environment for R.  It is described in the texts by Verzani and Lander along with R packages that should be installed.  Installer packages for RStudio are located at [*www.rstudio.com/* (Links to an external site.)](http://www.rstudio.com/)*.*  The installation process is simple.  RStudio is convenient and easy to use.  There are several other packages that should be installed as well:  ggplot2, coefplot, moments, and asbio.  How to do this is described by Lander.