For Tuesday:

Topic:	Introduction to the course and R Studio
Text:	-

Exercises:

1. Set up R and Rstudio.

   Guide to setup R and Rstudio (all following items have a hyperlink)

   - Downloading and setting up R and R studio. Follow the instruction on this webpage.
   - Get to know RStudio.
   - Install and use packages in R.
   - File organizing tips.
   - Create a project and set up working directory (Substitute "earth-analytics" with something more adequate).

2. Make, import and plot a data set.
   The following table displays the prevalence of autism per 10,000 ten-year old children in the United States from 1992 - 2000. (Data from C. J. Newshaffer, M. D. Falb, and J. G. Gurney, "National Autism Prevalence Trends From United States Special Education Data", Pediatrics, 115 (2205): e277-e282.).

Year	Prevalence
1992	3.5
1994	5.3
1996	7.8
1998	11.8
2000	18.3

   (a) Google (or other search engine) your way to find out how to create a data set directly with R.

   (b) Enter the values of the table into an Excel sheet, then import the data set to RStudio. Use google to find out how to import data sets from Excel into RStudio.

   (c) Google your way to find out how to plot the data set. Add labels for each of the axis as well as a title. Again, Google is your friend.

For Thursday:

Topic:	Working in groups
Text:	-

Exercises:

1. Fish Oil and Blood Pressure. The following table displays data from a study with 14 male volunteers with high blood pressure which were randomly assigned to either follow a diet containing fish oil or regular oil. The reduction in diastolic blood pressure displayed in the table.

Fish oil diet:	8	12	10	14	2	0	0
Regular oil diet:	-6	0	1	2	-3	-4	2

   Use google to find a guide to carry out a one-sided/one-sample t-test (you can assume the distribution is normal). And carry out the test. What does the p-value tell you?

2. Solar Radiation and Skin Cancer. In this exercise we will study a data set from a study on the correlation between skin cancer rates and solar radiation (Data from D. F. Andrews and A. M. Herzberg, Data, New York: Springer-Verlag, 1985). The data can be found in this weeks folder named data sets. The data is the yearly skin cancer rates (cases per 100,000 people) in Connecticut, with a code identifying those years that came two years after higher than average sunspot activity and those years that came two years after lower than average sunspot activity.
   First open the data set in Excel. As you can see, the three variables are all displayed in the same column. We must first split up the variables, so they each have a column. Please see this guide, or find another you might like better. Then back in RStuio, import the data set, and make scatterplots of skin cancer rates versus year, for each group separately. Use a search engine to find appropriate guide for doing plots with R. If you get any errors with regards to the margins of the plot, try and search for a solution to the problem online.