

Lab Instructions

This page contains links to the individual lab instructions for our Applied Statistics course.

NOTE: All these links take you to HTML versions of the labs. But if you change the link extension to “.pdf”, you will get a more print-friendly version.

With the exception of the first two labs, our labs focus on using RMarkdown as the primary interface and to generate the lab reports. We also promote the practice of creating a new RStudio project for each lab.

Lab 1¹ Introduces you to the RStudio environment and shows you how to carry out basic descriptive statistics.

The lab is intended for complete beginners, and all the needed R syntax is provided. Students work in the RStudio Console and write their answers on the printed pdf version of the lab assignment. This lab uses the `counties` dataset provided with our `hanoverbase` package, originally obtained from OpenIntro², which contains information about the different U.S. counties.

Lab 2³ Shows you how to create graphs and carry out basic one-variable and two-variable analyses.

In this lab, students continue to work in the RStudio Console and we provide them with all the needed R syntax. In this lab students continue to use the `counties` dataset from the first lab. The emphasis is on becoming familiar with the format of the R syntax, the RStudio environment, and using the generated graphs and results to draw conclusions in the context of the data.

Lab 3⁴ Shows you how to use RMarkdown to structure your work and create your report all in one document.

This lab introduces the use of RMarkdown and R code chunks to integrate R code and text into a single document. The emphasis is on getting the students comfortable with the RMarkdown syntax, document structure and R code chunks. The lab project is imported from a prepared GitHub repository. In this and all later labs, student answers are included in the RMarkdown document they produce.

Starting from this lab, students are expected to write some of their own syntax. It is advisable to have the cheatsheet⁵ available from this point on. This is the third and final lab that deals with the `counties` dataset.

Lab 4⁶ Offers you an opportunity to practice and solidify your learning from the first three labs, and introduces more two-variable techniques.

In this lab students learn how to create a new RStudio project, and how to import data from outside sources (CSV etc). It uses a new dataset (`driving`). Students are expected to produce some of the basic R code, while the code for some more advanced features is provided.

Lab 5⁷ Shows you how to customize graphics outputs.

Students learn how to customize the generated graphs by controlling the color, labels and legend. This is the first lab on the `brfss` dataset, generated by collecting together four different years of the annual BRFSS survey from the CDC⁸. This dataset is provided as part of the `hanoverbase` package

²<https://www.openintro.org/stat/data/>

⁵[../cheatSheet.html](#)

⁸<https://www.cdc.gov/brfss/index.html>

in three different sizes. The largest size, with 1.3 million rows, will likely overtax most systems if multiple students are sharing the same server resources.

Lab 6⁹ Adds more graph customizations and shows you how to create new variables from existing ones.

This lab continues the analysis of the `brfss` dataset, including more techniques for working with graphs. Emphasis is placed on the creation of new variables from existing variables. This lab contains more complex code than previous labs.

Lab 7¹⁰ Shows you how to work with linear models in RStudio.

This lab shows students how to import an Excel sheet (a modified version of “gun violence around the world”¹¹) into RStudio.

Lab 8¹² Lets you practice, in a more independent fashion, the data analysis techniques we have learned so far.

This lab is different in that no R syntax is provided and the questions are more open-ended. It is a chance for students to practice what they have learned so far, in preparation for their term projects.

The dataset is the Cars 2015¹³ dataset provided by Lock5Stat.

¹¹https://www.openintro.org/stat/data/?data=gun_violence

¹³<http://www.lock5stat.com/datapage.html>