Troubleshooting Guide

August 2022

At the time of publication, all code examples used in the book were tested and verified. However, variations in code execution can result for several reasons discussed below.

1. Using a different version of Python

As specified in the introduction (on page 5), we used Python version 3.7 in this book. If you use either an older or newer version of Python, there may be some issues with code execution. For example, in Python 3.7, the code in Figure 8.1 (on page 188) results in an exception since no unique mode is found in the data (there are 5 modes). However, if you were to execute the code in Figure 8.1 in Python 3.8 or later, an exception is not thrown, since a change was made to the `mode` function in the Python statistics module beginning with Version 3.8.

2. If you use a different version of a Python package

On page 97, we specified that we are using Pandas version 0.24.1 and how you can install that specific version of that package (using `pip install pandas==0.24.1`). The Python Insight box on page 99 reiterates this information and the importance of knowing which versions of software are being used. On page 225, we specify that we are using Seaborn version 0.9.0, and we used matplotlib 3.0.3 in the textbook (as indicated in the reference listed on page 244). One issue that was recently discovered was that there was a change to matplotlib with the `.show()` method. In a more recent version of matplotlib, the default for the `.show()` method causes the figure to quickly display and then disappear. To use the version of matplotlib that was used with the Python code in the book, if you have already installed matplotlib you can first uninstall it (using `pip uninstall matplotlib`) and then install version 3.0.3 (using `pip install matplotlib==3.0.3`). Also note that a warning now results when executing our seaborn code that uses the `distplot()` function (such as the code in Figure 9.19 on page 228) stating that that function is deprecated. The message that is shown is:

> FutureWarning: `distplot` is a deprecated function and will be removed in a future version. Please adapt your code to use either `displot` (a figure-level function with similar flexibility) or `histplot` (an axes-level function for histograms).

Changes such as this also emphasize the importance of knowing which versions of packages you are using and perhaps the need to install an earlier version of a package or else modify code to adapt to changes in packages.

3. Changes to external content

Some of the textbook examples use external content (particularly in Chapter 7 Obtaining Data from the Web Using Python). Although we tried to use sources that were not likely to change, the web page that we referenced in the Python code in Figure 7.20 (on page 166) has been updated in a way that causes the code in that figure to no longer work. However, you can use the following web page (from the Internet Archive), which is a stored version of the webpage that

we used in that example. To make this change, you can use the following URL in lines 6-8 of the code in Figure 7.20.

http://web.archive.org/web/20190418012753/https://gssdataexplorer.norc.org/trends/Gender%20&%20Marriage?measure=happy&print=true&response=Very+happy&breakdown=Total&table=true

Another change in the code that is necessary is to change the index to the tables_found list to 2 in lines 21 and 28 in the code. The reason for this change is that the internet archive site adds a table to the web page at the beginning of the page.

In addition, the dataset referenced in SCU 7.5 (https://data.cityofchicago.org/resource/hemf-zffe.json), is no longer available at the Chicago Data Portal. There are many data sets available at the Chicago Data Portal, and you can use any of them for this exercise. For example, you could use the data set found at https://data.cityofchicago.org/resource/4ijn-s7e5.json - in which case you would replace "hemf-zffe" with "4ijn-s7e5" in the provided solution.

4. Platform independence and dependence

Although Python has been developed to be platform independent and can execute on Windows Computers and Macs, not all programs work on every operating system. For example, Microsoft Access is developed to run on the Microsoft Windows operating system.   Some of our code examples in chapters 5 and 11 use Microsoft Access. For those using Mac computers, one option is to install the Windows OS on your Mac in order to use Microsoft Access. We have posted instructions on how to do this on SAGE's website under the resources section at the following link

https://study.sagepub.com/researchmethods/statistics/kaefer-intro-to-python