Lockdown to R-Full-stack

Richard Nacianceno

2020-07-29

Contents

Overview

This is my journey during Covid-19 lockdown in the US and how I was able to learn R programming through Coursera - JHU Data Science Specialization (Johns Hopkins University, 2013). I will provide the "ah-ha" moment during this journey.

The program consists of 10 courses that covers the following:

1. **The Data Scientist's Toolbox**
2. **R Programming**
3. **Getting and Cleaning Data**
4. **Exploratory Data Analysis**
5. **Reproducible Research**
6. **Statistical Inference**
7. **Regression Models**
8. **Practical Machine Learning**
9. **Developing Data Products**
10. **Data Science Capstone**

On July 29, 2020, I'm currently at #10 - Data Science Capstone - Week 1. The entire data science specialization can be completed in less than 11 months based on your desire to learn. I've put in an average of 5 hours per day in each courses. The reason I wanted to learn these materials so fast are as follows:

- Personal growth
- Love dissecting data for insights
- The course syllabus is well though-out
- Professors bringing their expertise in their respective field
- Professors abilities to communicate and teach the courses is simple and concise.

Professor Roger D. Peng, PhD, Professor Brian Caffo, PhD and Professor Jeff Leek, PhD - from the bottom of my heart - thank you so much.

Expect To Have Lucid Dreams

"I had a dream about you. It's been a whilesince I could remember any of my dreams,and still, this one has left me with such strong impression.(Ninkovic, 2020)"

That quote from Aleksandra is one of the many stories that occurred to me while taking these courses. It is not easy. You need perseverance and continuously seeking knowledge. There were nights that turned to days and nights to day and back to night because I could not figure out how to solve a problem. My brain was constantly looking for a solution. Once you solved it, you then moved on to the next challenge with the same twist of time. It gets better (**?**).

There was no way for me to complete this program without the help of google.com, Yihue Xie, PhD - books, Professor Handley Wickham, PhD - books and to all the people that contributed to stackoverflow - Thank You.

Retrospect

In retrospect, I would add the following on top of the specialization first which are as follows:

1. Get Linux Ubuntu 20.04 LTS as your main environment. I initially started with Microsoft Windows 10. The experience was awful. After the Windows 10 update, it would not let me write on my disk while running RStudio. This added to the sleepless nights. Linux Ubuntu 20.04 LTS is the most stable environment for me.

2. Learn the bash through Cousera course The Unix Workbench by Sean Kross which provides a powerful an insight of the inner working of linux accompanied with Git/Github for version control.
3. Learn blogdown or bookdown packages for proper documentation. This will help in organizing your thought process and create a flow for your internal data. It will give you a stricter rule where to save an assignment rather than having them all over your hard drive. There were many days that I have to find saved assignment in my hard-drive. Now, that I have the power of blogdown or bookdown together with github it provides a central place where I can organize my work.

4. Document and complete your specialization using either bookdown or blogdown packages.

I wish I can go back in-time and start the process with these 3 steps first before jumping into Data Science Specialization.

Chapter 1

Beginning

It was March 2020, I was still traumatize that Andrew Yang drop-out for the presidential race. The other headline news during these time frame was about Syrian Civil War, Coronavirus pandemic, European migrant crisis, and the presidential race in the United States (Wikipedia, 2020).

I was very optimistic that United States will not have a pandemic in any of its states or properties. At the same time, I was amazed at the many dashboards that were created to monitor the Covid-19. Curiosity got the best of me and I decided to sign-up for the Coursera - Johns Hopkins University Data Science Specialization (Johns Hopkins University, 2013).

Figures and tables with captions will be placed in `figure` and `table` environments, respectively.

```
par(mar = c(4, 4, .1, .1))
plot(pressure, type = 'b', pch = 19)
```

Reference a figure by its code chunk label with the `fig:` prefix, e.g., see Figure 1.1. Similarly, you can reference tables generated from `knitr::kable()`, e.g., see Table 1.1.

```
knitr::kable(
  head(iris, 20), caption = 'Here is a nice table!',
  booktabs = TRUE
)
```

You can write citations, too. For example, we are using the **bookdown** package (Xie, 2020) in this sample book, which was built on top of R Markdown and **knitr** (Xie, 2015).

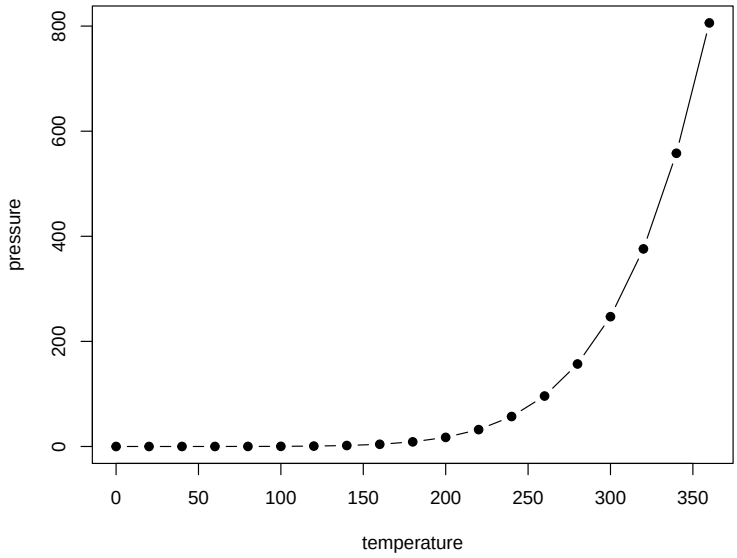

Figure 1.1: Here is a nice figure!

Table 1.1: Here is a nice table!

Sepal.Length	Sepal.Width	Petal.Length	Petal.Width	Species
5.1	3.5	1.4	0.2	setosa
4.9	3.0	1.4	0.2	setosa
4.7	3.2	1.3	0.2	setosa
4.6	3.1	1.5	0.2	setosa
5.0	3.6	1.4	0.2	setosa
5.4	3.9	1.7	0.4	setosa
4.6	3.4	1.4	0.3	setosa
5.0	3.4	1.5	0.2	setosa
4.4	2.9	1.4	0.2	setosa
4.9	3.1	1.5	0.1	setosa
5.4	3.7	1.5	0.2	setosa
4.8	3.4	1.6	0.2	setosa
4.8	3.0	1.4	0.1	setosa
4.3	3.0	1.1	0.1	setosa
5.8	4.0	1.2	0.2	setosa
5.7	4.4	1.5	0.4	setosa
5.4	3.9	1.3	0.4	setosa
5.1	3.5	1.4	0.3	setosa
5.7	3.8	1.7	0.3	setosa
5.1	3.8	1.5	0.3	setosa

Chapter 2

Data Science Toolbox

Here is a review of existing methods.

Chapter 3

Getting and Cleaning Data

Chapter 4

R Programming

We describe our methods in this chapter.

Chapter 5

Applications

Some *significant* applications are demonstrated in this chapter.

5.1 Example one

5.2 Example two

Chapter 6

Final Words

We have finished a nice book.

Chapter 7

Reproducible Research

Chapter 8

Regression Models

Chapter 9

Practical Machine Learning

Chapter 10

Developing Data Products

Bibliography

Johns Hopkins University (2013). Data Science specialization.

Ninkovic, A. (2020). Quotable quote.

Wikipedia (2020). Portal:current events/march 2020.

Xie, Y. (2015). *Dynamic Documents with R and knitr*. Chapman and Hall/CRC, Boca Raton, Florida, 2nd edition. ISBN 978-1498716963.

Xie, Y. (2020). *bookdown: Authoring Books and Technical Documents with R Markdown*. R package version 0.20.