Welcome back to our workshop activities! Today we will finish the coding part of the workshop. With your partner, claim a project to work on from the list below. Remember that you will be presenting this, so make sure your code is organized. Once you are done with your code, download it and then upload it to Github.

If you are new to coding, I encourage you to try the exercises on your own before turning to your peers for help. When working in groups, talk to each other about how you're solving these problems, but **you should write up your final answers individually**.

List of DataCamp Projects

Complete one of the following guided projects in DataCamp and then discuss with your partner what you learned from them.

1. **Dr. Semmelweis and the Discovery of Handwashing**: In 1847, the Hungarian physician Ignaz Semmelweis made a breakthough discovery: he discovers handwashing. Contaminated hands was a major cause of childbed fever and by enforcing handwashing at his hospital he saved hundreds of lives. In this project, you will reanalyze the medical data Semmelweis collected!

2. **Who Is Drunk and When in Ames, Iowa?**: Using data collected from the State of Iowa, you will group, summarize, and visualize data on breath alcohol tests in Ames, Iowa, (home of Iowa State University) from 2013-2017. Some questions you will answer include, "What is the highest recorded value?" and "When do breath alcohol tests occur most?"

3. **A Visual History of Nobel Prize Winners**: The Nobel Prize is perhaps the world's most well known scientific award. Every year it is given to scientists and scholars in chemistry, literature, physics, medicine, economics, and peace. The first Nobel Prize was handed out in 1901, and at that time the prize was Eurocentric and male-focused, but nowadays it's not biased in any way. Surely, right? Well, let's find out! What characteristics do the prize winners have? Which country gets it most often? And has anybody gotten it twice? It's up to you to figure this out.

4. **Games**: Each partner will explore a different game.

    - R – **Level Difficulty in Candy Crush Saga**: Candy Crush Saga is a hit mobile game developed by King (part of Activision—Blizzard) that is played by millions of people all around the world. In this Project, you will get to work with a real Candy Crush dataset and use this data to estimate level difficulty. This Project assumes you can manipulate data frames using dplyr and make plots using ggplot2.
    - Python – **Exploring the History of Lego**: The Rebrickable database includes data on every LEGO set that has ever been sold; the names of the sets, what bricks they contain, what color the bricks are, etc. It might be small bricks, but

this is big data! In this project, you will get to explore the Rebrickable database and answer a series of questions related to the history of Lego!

Your own EDA

Claim one of the following datasets and do some EDA. Make sure you at least answer 4 interesting questions for your data.

1. Penguins data set
2. Pokemon go data set
3. Books data set
4. Olympics data set