

## **Project 3: Data Querying and Cleaning**

**Modality: On Demand**

**Duration: 14 Days**

You've covered a lot of ground in the past three weeks, learning about making queries with SQL basics, control flow and data structures in Python. Useful Python libraries: numpy, matplotlib, and pandas. Basic R data structures: vectors, matrices, lists, and data frames. Graphics using R. In your second project you're going to be applying these skills in to explore baseball statistics.

In the process you'll practice a vital data science skill the ability to research new tools and read documentation. You may choose to do this project in either R or Python.

### **Course Outline:**