1. What are the top three data science programming languages, in order, according to the 2020 Kaggle data science and machine learning survey?

Python, SQL, and R

1. What is the tradeoff between using a GUI vs using a programming language for data science? What were some of the GUIs for data science that we mentioned?

A GUI makes it easier and faster to carry out data science but lacks flexibility. With a general-purpose programming language like Python, we can do almost anything, but it can require much more effort and time.

Some of the GUIs include Excel, Tableau, Alteryx, and Rapidminer.

1. Who are the top three cloud providers for data science and machine learning according to the Kaggle 2020 survey?

Amazon (AWS), Google (GCP), and Microsoft (Azure)

1. What percent of time do data scientists spend cleaning and preparing data?

According to some surveys, it’s between 25% and 75% of their time. Anecdotal evidence suggests a number more like 90% or more of data scientists’ time.

1. What specializations in and around data science did we discuss?

Machine learning engineer

1. What data science project management strategies did we discuss, and which one is the most recent? What are their acronyms and what do the acronyms stand for?

CRISP-DM – Cross-industry standard process for data mining

TDSP – team data science process

1. What are the steps in the two data science project management strategies we discussed?

CRISP-DM: Business understanding, data understanding, data preparation, modeling, evaluation, deployment

TDSP: Business understanding, data acquisition and understanding, modeling, deployment, customer acceptance