Data Science

Introduction
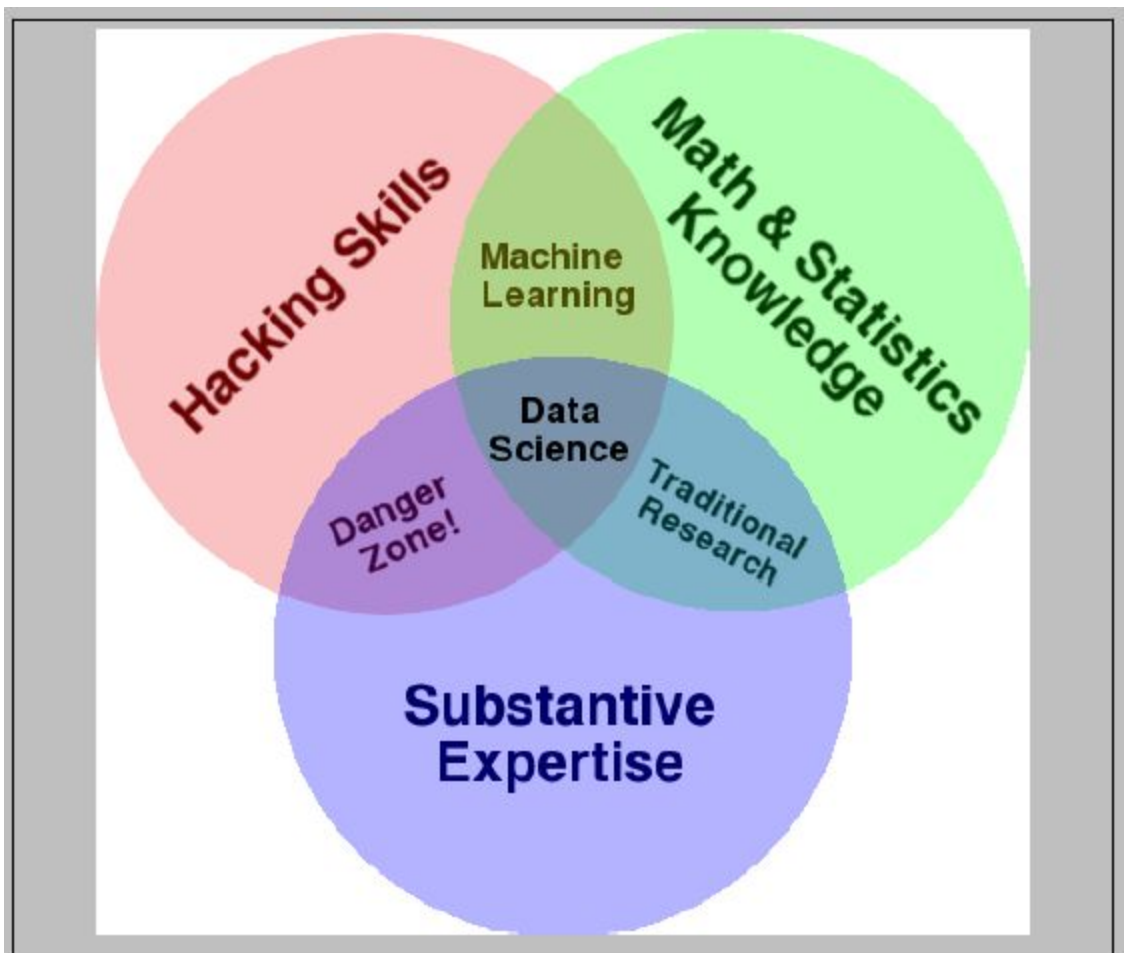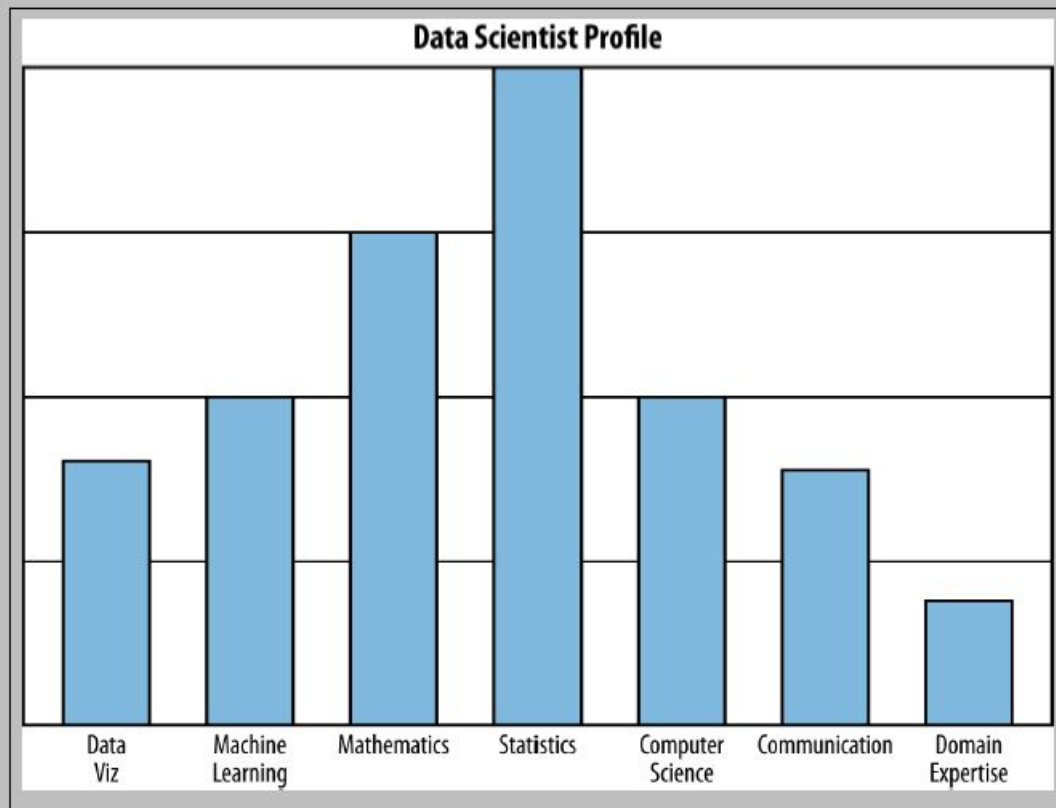
Figure 1-1. Drew Conway's Venn diagram of data science

Figure 1-2. Rachel's data science profile, which she created to illustrate trying to visualize oneself as a data scientist; she wanted students and guest lecturers to "riff" on this—to add buckets or remove skills, use a different scale or visualization method, and think about the drawbacks of self-reporting
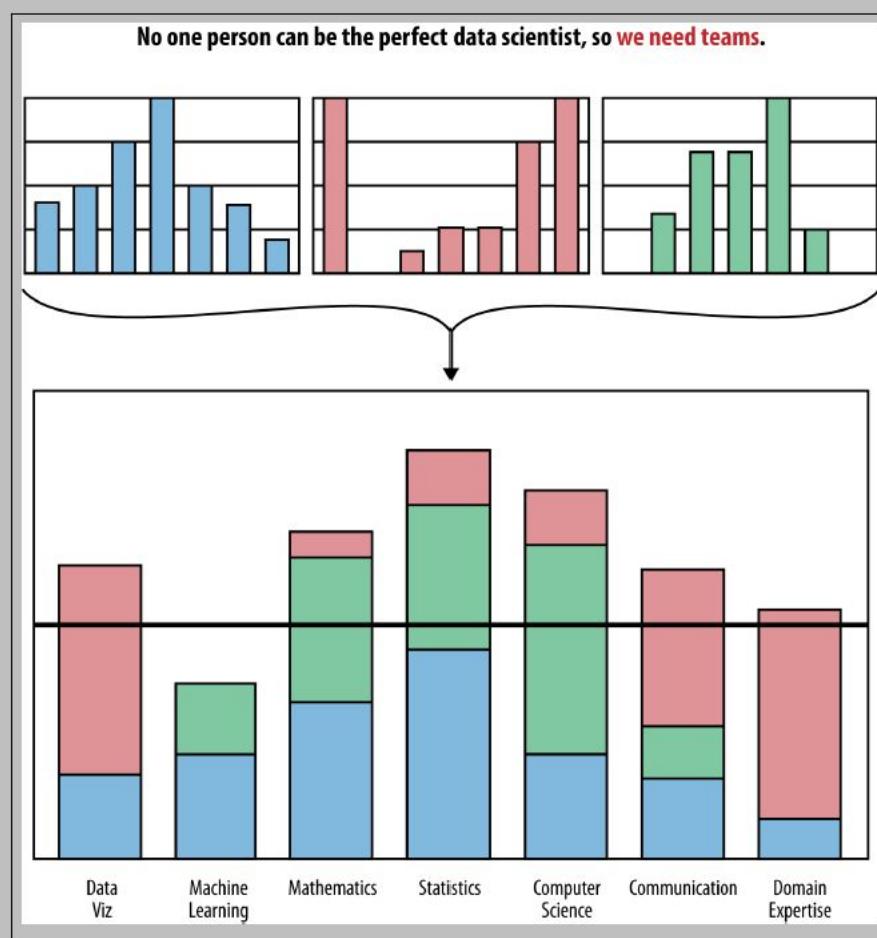
Figure 1-3. Data science team profiles can be constructed from data scientist profiles; there should be alignment between the data science team profile and the profile of the data problems they try to solve
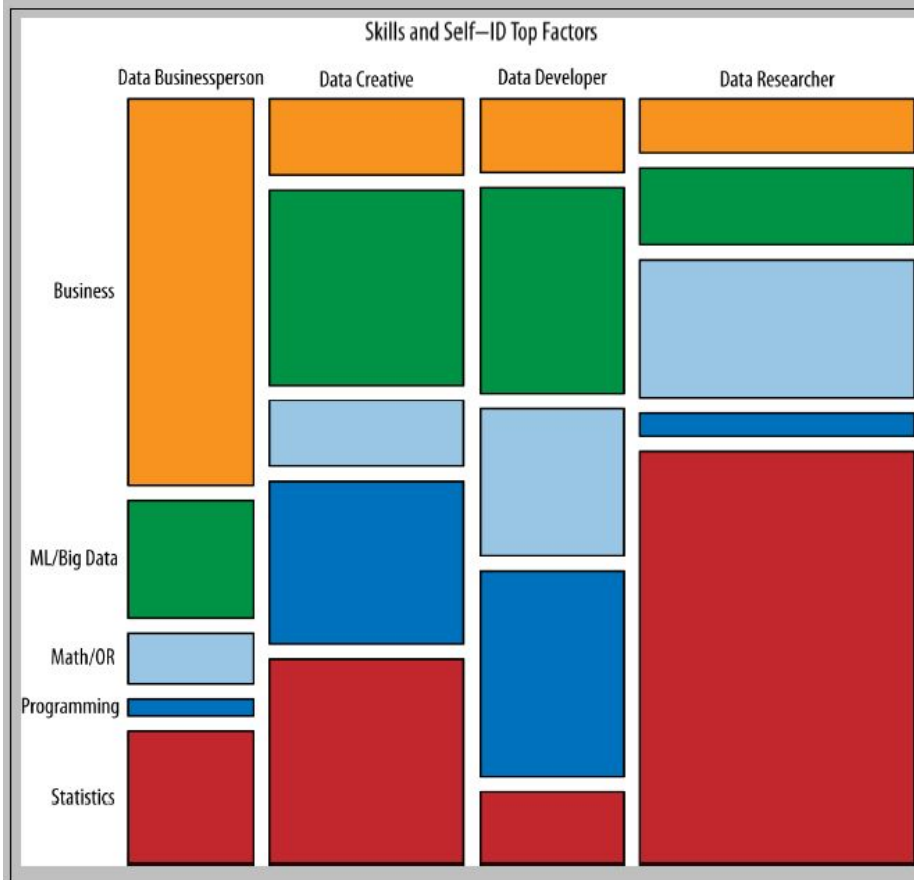
Figure 1-4. Harlan Harris's clustering and visualization of subfields of data science from Analyzing the Analyzers (O'Reilly) by Harlan Harris, Sean Murphy, and Marck Vaisman based on a survey of several hundred data science practitioners in mid-2012

Example: Hurricane Frances

Hurricane Frances was on its way, barreling across the Caribbean …

A week ahead of the storm's landfall, Linda M. Dillman, Wal-Mart's chief information officer, pressed her staff to come up with forecasts based on what had happened when Hurricane Charley struck several weeks earlier

Example: Predicting Customer Churn

Assume you just landed a great analytical job with MegaTelCo, one of the largest telecommunication firms in the United States. They are having a major problem with customer retention in their wireless business. In the mid-Atlantic region, 20% of cell phone customers leave when their contracts expire, and it is getting increasingly difficult to acquire new customers.
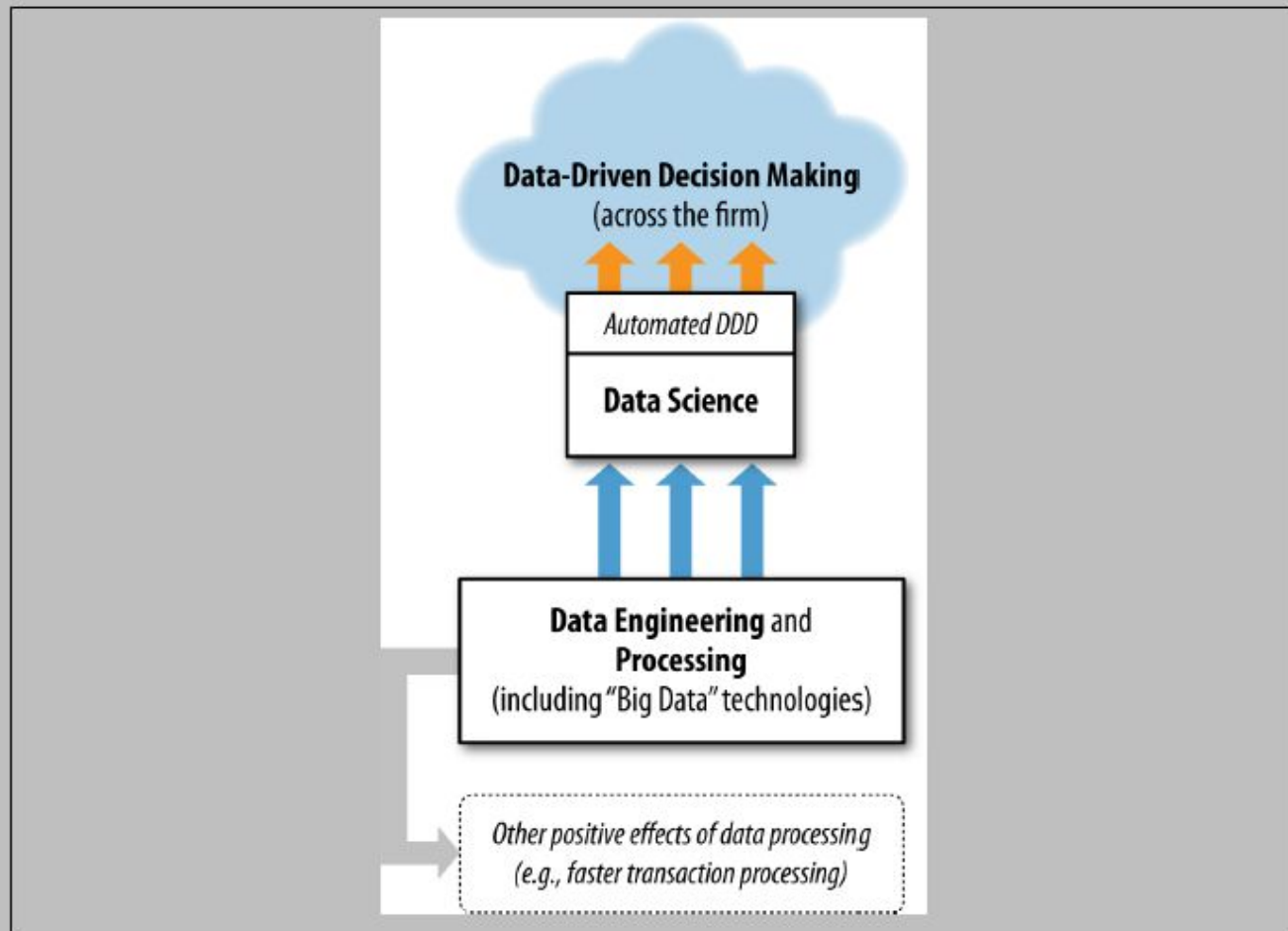
Figure 1-1. Data science in the context of various data-related processes in the organization.

In 2012, Walmart's competitor Target was in the news for a data-driven decision-making case of its own

Like most retailers, Target cares about consumers' shopping habits, what drives them, and what can influence them.

Consumers tend to have inertia in their habits and getting them to change is very difficult.

Decision makers at Target knew, however, that the arrival of a new baby in a family is one point where people do change their shopping habits significantly.

References

1. Doing Data Science, Straight Talk From The Frontline by Cathy O'Neil and Rachel Schutt - Chapter 1
2. Data Science for Business by Foster Provost and Tom Fawcett - Chapter 1