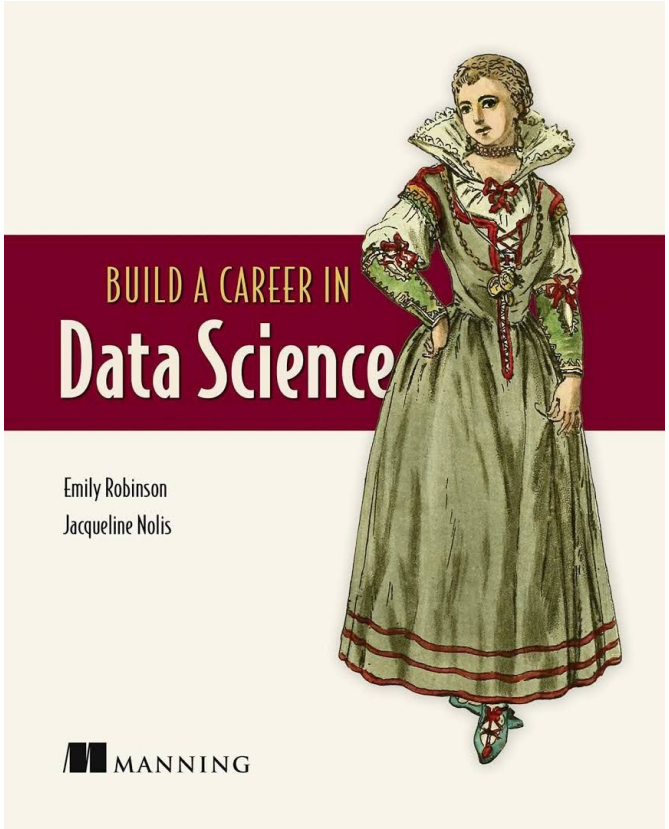
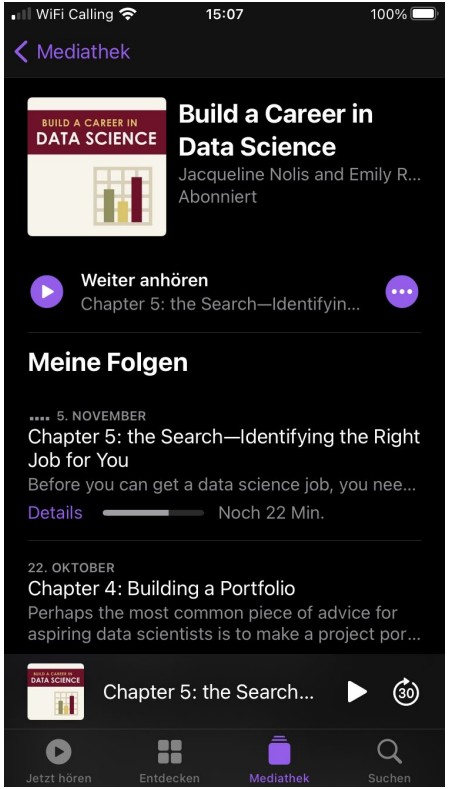
Building a Career in Data Science

David Furrer

13.11.2020

Book and podcast





Overview

1) What is data science
2) Data science companies
3) Getting the skills
4) Building a portfolio
5) Finding the right job
6) CV and cover letter
7) The interview
8) First month on the job
9) Joining the data science community

1.1) What is Data Science?

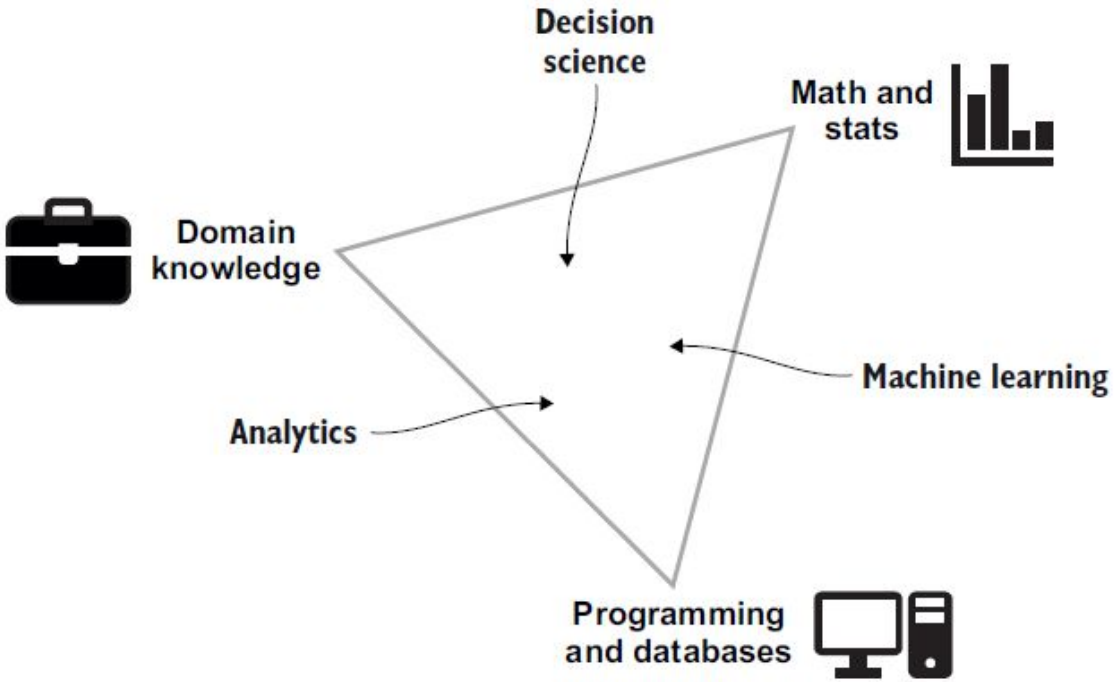


Figure 1.1 The skills that combine to make data science and how they combine to make different roles

1.2) Three levels of knowledge in DS

1) That the technique exists

2) How to apply the technique

3) How to choose which technique to try

1.3) Core skills of Data Science

1) SQL

2) Git

3) Translating business question into data question

4) pandas

1.4) Different types of DS jobs

1) Data Analyst (dashboards and reports)

2) Machine Learning Engineer (creates models)

3) Decision Scientist (analyses that produce recommendations)

2.1) Data Science Companies

Table 2.1 A summary of companies that hire data scientists

Criteria	MTC	HandbagLOVE	Seg-Metra	Videory	GAD
	Massive tech	Retailer	Startup	Mid-tech	Defense
Bureaucracy	Lots	Little	None	Some	Lots
Tech stack	Complex	Old	Fragile	Mixed	Ancient
Freedom	Little	Lots	TONS	Lots	None
Salary	Amazing	Decent	Poor	Great	Decent
Job security	Great	Decent	Poor	Decent	Great
Chances to learn	Lots	Some	Lots	Lots	Few

3.1) Getting the skills

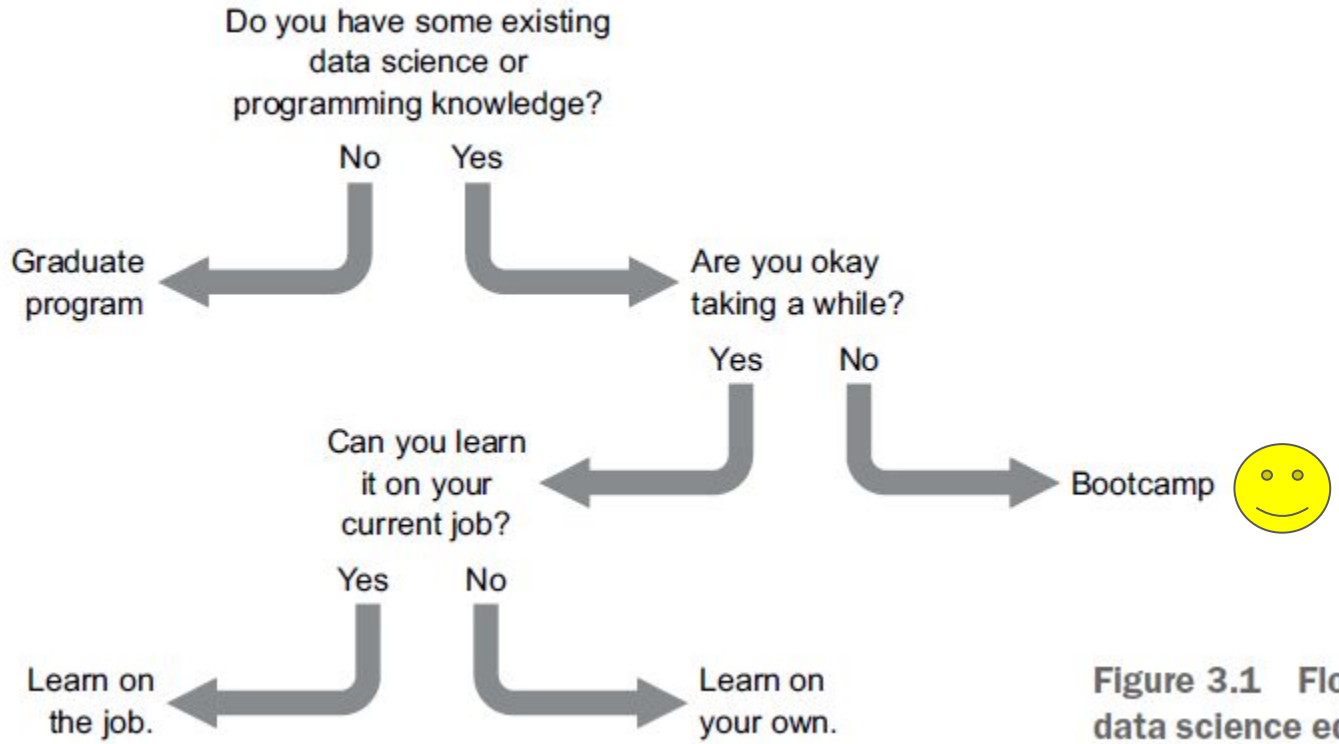


Figure 3.1 Flow of deciding which data science education route to take

4.1) Building a portfolio

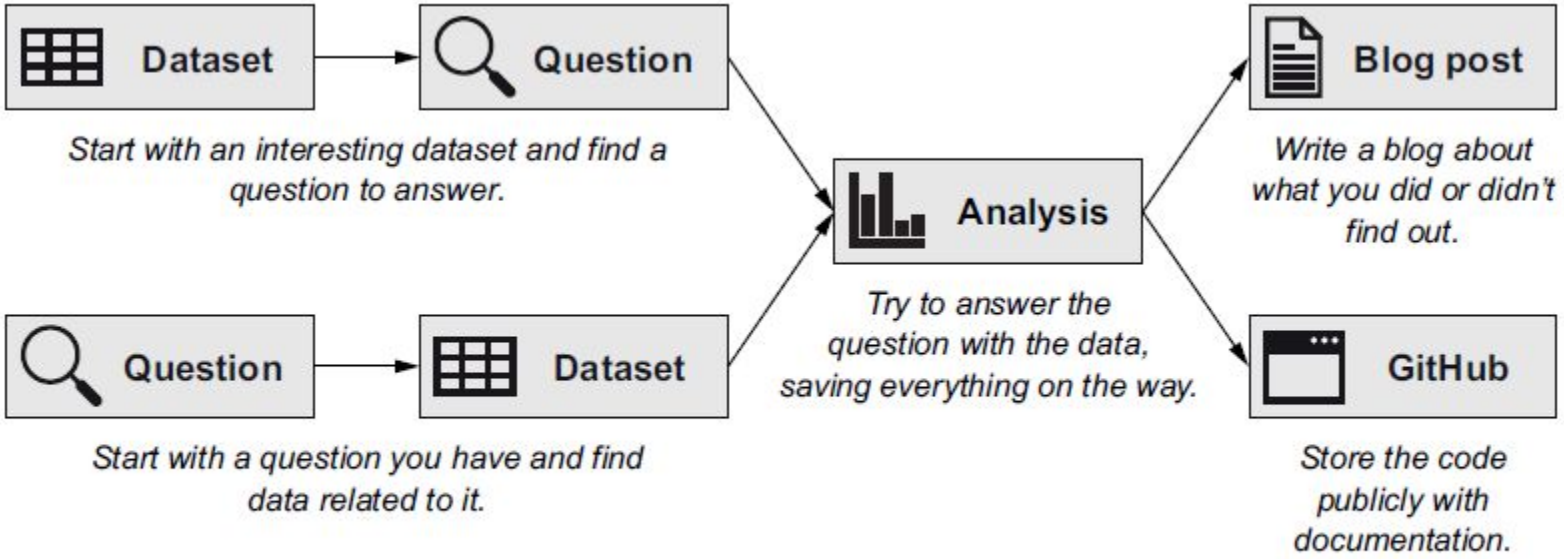


Figure 4.1 The flow of creating a data science project

4.2) Project ideas

1) Tidy up final project and put it on your github (e.g.: https://github.com/davidfurrer/bokeh-recommendation-app)
2) Tidy tuesday (https://github.com/rfordatascience/tidytuesday)
3) Kaggle
4) Go to stats website of you local government and make a blog post about some data
5) Scrape some data
6) Twitter (#RStats, #datascience)
7) Medium newsletter
8) Python podcasts (Talk Python To Me, Podcast.__init__)

5.1) Finding the right job

1) Don’t just look for “data scientist”
2) Also look for “data analyst”, “Machine Learning Engineer”, “data”, “research”
3) Don’t just look in Zurich. E.g.: Zug and Baden or other nearby town also have some interesting jobs

6.1) CV

SARA JONES

San Francisco, CA · 534-241-6264
sarajones@gmail.com · linkedin.com/in/sarajones · sarajones.github.io · github.com/sarajones

EXPERIENCE

JUNE 2019 – PRESENT, SAN FRANCISCO, CA
DATA SCIENCE FELLOW, AWESOME BOOTCAMP

- Built a web application in Python that recommends the best New York City neighborhood to live in based on someone's budget, lifestyle preferences, and work
- Analyzed 2,200 New York Times business articles (obtained via API) using natural language processing (TFIDF and NMF), visualizing how topics changed over time

AUGUST 2017 – JUNE 2019, SAN FRANCISCO, CA
INVESTMENT CONSULTANT, BIGCO

- Created a forecasting model in Python that boosted quarterly revenue by 10%
- Automated generating weekly market and industry trend reports

SEPTEMBER 2016 – JUNE 2017, NEW ORLEANS, LA
INTRODUCTION TO STATISTICS TEACHING ASSISTANT, COOL UNIVERSITY

- Led weekly review sessions of sixty students, earning a 4.86/5 rating in evaluations
- Created and open-sourced study guides that have been downloaded over 1,500 times

JUNE 2016 – AUGUST 2016, NEW ORLEANS, LA
ECONOMICS RESEARCH ASSISTANT, COOL UNIVERSITY

- Conducted an in-person experiment on decision-making with 200 participants, using cluster analysis to analyze the results in Python
- Published the resulting paper in the Journal of Awesome Economics

EDUCATION

JUNE 2017, NEW ORLEANS, LA
BA ECONOMICS, STATISTICS MINOR COOL UNIVERSITY
GPA 3.65/4.0
Relevant Coursework: Linear Algebra, Introduction to Regression and Statistical Computing, Experimental Design, Econometrics, Elements of Algorithms and Computation

SKILLS

- Python
- SQL
- Machine learning
- Git
- Pandas
- Seaborn
- Scikit-learn
- NumPy

Figure 6.1 Example résumé for an aspiring data scientist

David Furrer

Nordstrasse 58B, 8200 Schaffhausen · +41774101343
furrer.david1@gmail.com · linkedin.com/in/david-furrer/ · github.com/davidfurrer · medium.com/@fulowa

EXPERIENCE

Grand Casino Baden, Baden — Marketing Data Analyst
September 2020 - PRESENT

- Custom attribution modeling
- Reporting using Google Analytics, Google Datastudio, BigQuery

OTO, Zurich — Data Scientist
JULY 2019 - PRESENT

- Built a neural network model for NPS prediction using text and audio
- Deployed model to kubeflow pipeline for daily batch processing of calls
- Delivered multiple POCs to customers

Propulsion Academy, Zurich — *Data Science Student*
MARCH 2019 - JUNE 2019

- Data wrangling, experimental design and bootstrapping using R/Python/numpy/pandas
- Machine learning projects using scikit-learn and tensorflow
- Data visualization and graphic design using matplotlib, ggplot2, bokeh and plotly

Novartis, Basel — *Internship*
MARCH 2013 - DECEMBER 2013

- Research chemist in the group of Hans-Jörg Roth, Chemical Libraries (CLI)

EDUCATION

University of St. Gallen, St. Gallen — Economics, *Master of Arts*
SEPTEMBER 2014 - FEBRUARY 2017

- Relevant Coursework: Mathematics, Data Analytics I: Predictive Econometrics, Data Analytics II: Causal Econometrics, Time Series Econometrics

ETH, Zurich — *Chemistry, Bachelor and Master of Science*
SEPTEMBER 2007 - FEBRUARY 2013

- Relevant Coursework: Mathematics I/II (Analysis, Linear Algebra und Statistics), Computer Science I, Mathematics III: Partial Differential Equations

PROJECTS

jobcloud app — *Final Project of Propulsion Academy*

- Using data we received from jobcloud (jobs.ch) we built a bokeh app that recommends similar job ad titles that have a high expected click rate.

SKILLS

- Python
- R
- SQL
- Git
- Pandas
- Seaborn
- Scikit-learn
- BigQuery
- Machine Learning
- Google Analytics
- AWS
- GCP

6.2) Cover Letter

SARA JONES

New York, NY · 534-241-6264

sarajones@gmail.com · linkedin.com/in/sarajones · sarajones.github.io · github.com/sarajones

GREETING

Dear Jared,

INTRODUCTORY PARAGRAPH

I am writing to express my strong interest in applying for the Data Scientist position at Awesome Company. I've enjoyed reading Awesome Company's data science blog since it started 8 months ago. The post on using topic modeling to automatically generate tags for your support articles was immensely helpful in one of my own projects to classify articles in the New York Times business section.

1-2 PARAGRAPHS OF DATA SCIENCE WORK EXAMPLES

I recently graduated from Awesome Bootcamp, a full-time, 3-month Data Science immersive. At Awesome Bootcamp, I designed, implemented, and delivered data science projects in Python involving data acquisition, data wrangling, machine learning, and data visualization. For my final project, I gathered 3,000 neighborhood reviews and ratings from Neighborhood Company. By using natural language processing on the reviews and available listings from Real Estate Company's API, I built a recommendation system that will match you to a neighborhood based on your budget, preferences, and a free-text description of your ideal neighborhood. You can try it out here: myawesomewebapp.com.

Prior to Awesome Bootcamp, I was an Investment Consultant at BigCo. When I joined, my team of six was all using Excel. While exceeding my targets, I began automating common tasks in Python, such as generating a weekly market and industry trends report, saving the team hours each week. I then developed a tailored curriculum to teach them Python. The initiative was so successful the company asked me to develop a full 2-day workshop and flew me out to three other offices to teach it, reaching over 70 consultants.

CLOSING PARAGRAPH

I am confident that my expertise in Python, academic training in Economics and Statistics, and experience delivering business results would make me a great fit for the Data Science team. Thank you for your consideration.

SIGNOFF

Sincerely,
Sara Jones

Figure 6.2 An example cover letter with highlights showing the different components

DAVID FURRER

Nordstrasse 58B
CH-8200 Schaffhausen
Tel.: 077 4101343
Mail: furrer.david1@gmail.com

Schaffhausen, June 3rd 2020

Economist/Data-Scientist (m/w) (80% - 100%)

Dear Sir or Madam,

I am writing to express my strong interest in applying for the Economist/Data-Scientist (m/w) (80% - 100%) position at Wüest Partner AG.

I recently graduated from Propulsion Academy, a full-time, 3-month Data Science immersive. At Propulsion, I designed, implemented, and delivered data science projects in Python involving data acquisition, data wrangling, machine learning, and data visualization. For my final project, me and my colleague worked together with jobcloud (owner of jobs.ch) to build a recommendation app that can help companies to formulate job ad titles that have a higher expected view count. In the backend of the app the latest NLP techniques (BERT) were used. You can see the project in more detail here: https://github.com/davidfurrer/bokeh-recommendation-app.

After graduation, in July 2019, I started working at OTO, a voice intelligence start-up in Zurich. I am really grateful for the fact that I got to work on a full life-cycle of a model: from training it to bringing it to production. This meant that I was able to see all complications and solutions that arise along this journey. I also realized the importance of having frequent discussions with all stakeholders to keep expectations aligned.

I am confident that my expertise in Python, academic training in Economics and Statistics, and experience in delivering business results would make me a great fit for the Data Science team. Thank you for your consideration.

Sincerely,
David Furrer

7) The interview

1) Startup interview
2) Mid-sized tech interview
3) Not so tech-savvy company interview

7) The interview

1) Startup interview
2) Mid-sized tech interview
3) Not so tech-savvy company interview

Making an effective analysis (2)

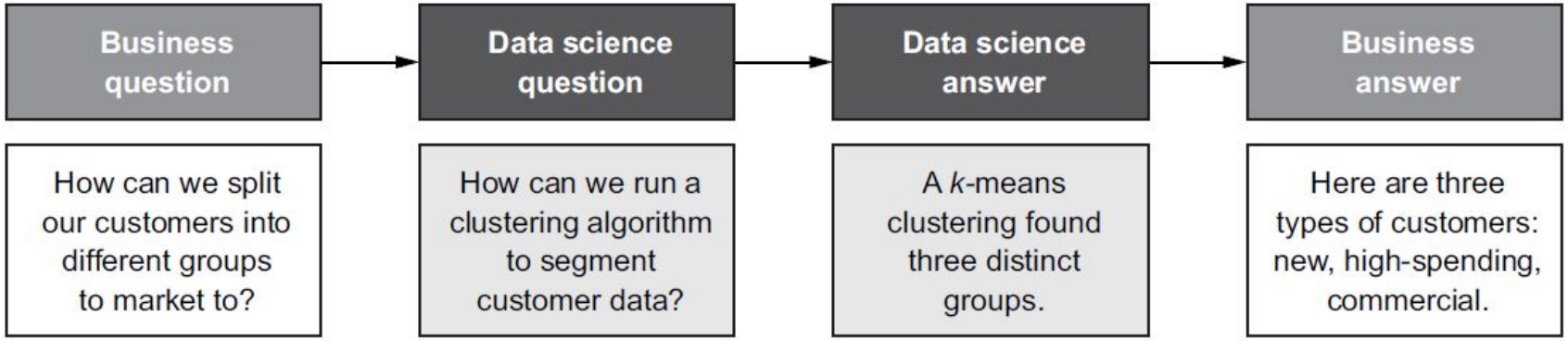


Figure 10.2 The process of answering a business question with data science, devised by Renee Teate

Making an effective analysis (3)

North American Customer Churn Analysis

August McNamara (amcnamara@company.com), May 2020

Objective: to understand why North American customers are joining at a lower rate than in other regions.

Intra-North America analysis

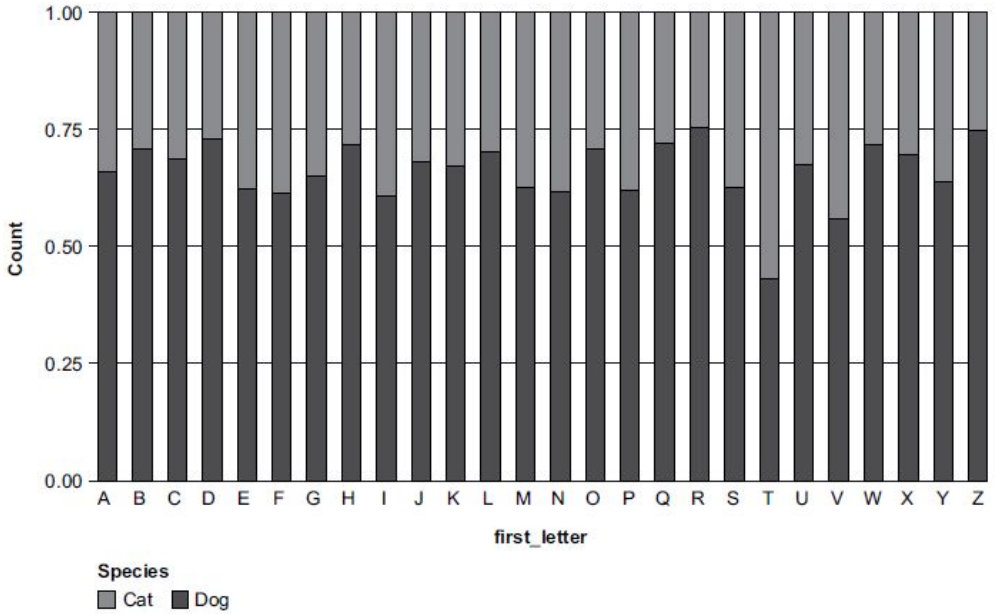
1. Are there attributes within North American customers that relate to new customer acquisition?
    a. Regression model on last month's new North American customers – customer spend and demographic attributes to find importance
    b. Extend part (a) to compare between customers acquired in each month over the past year
2. How much has the acquisition rate been changing over time?
    a. Time series analysis of acquisition rate across the region
    b. Split time series by country / state and look for correlations

Compare North America to other regions

1. How similar are North American customers to other regions?
    a. Generalized linear model with region as an attribute to model acquisition
    b. Create visualization of global map colored by acquisition rate

Figure 10.3 An example analysis plan

Making an effective analysis (4)



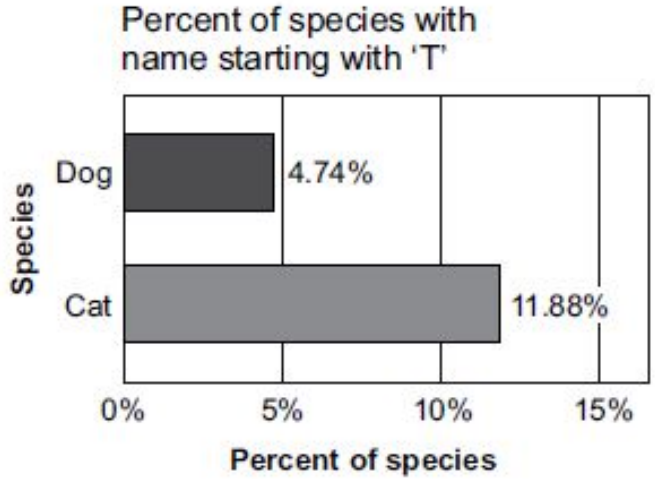


Figure 10.5 Example of a visualization made during an analysis before cleaning

Working with stakeholders

Joining the community

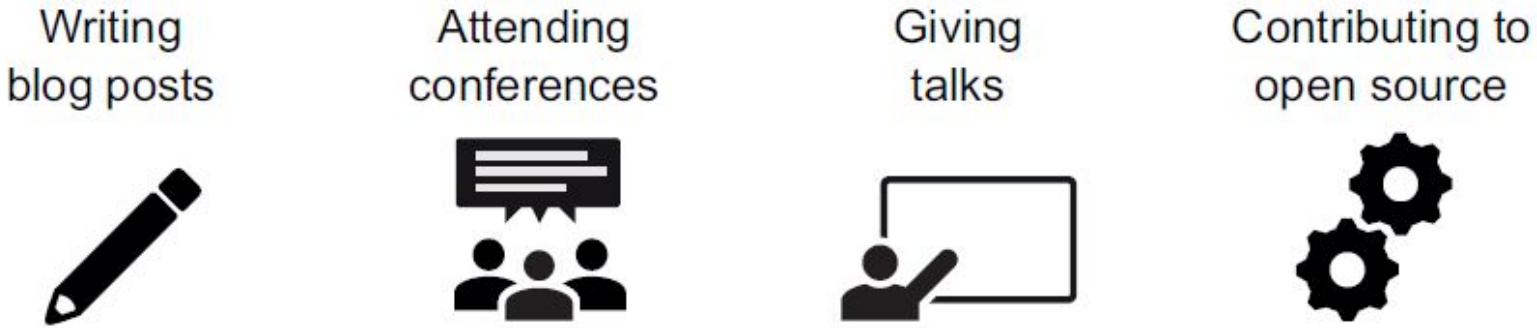


Figure 14.1 Some ways to join the community that are covered in this chapter

Some interview questions

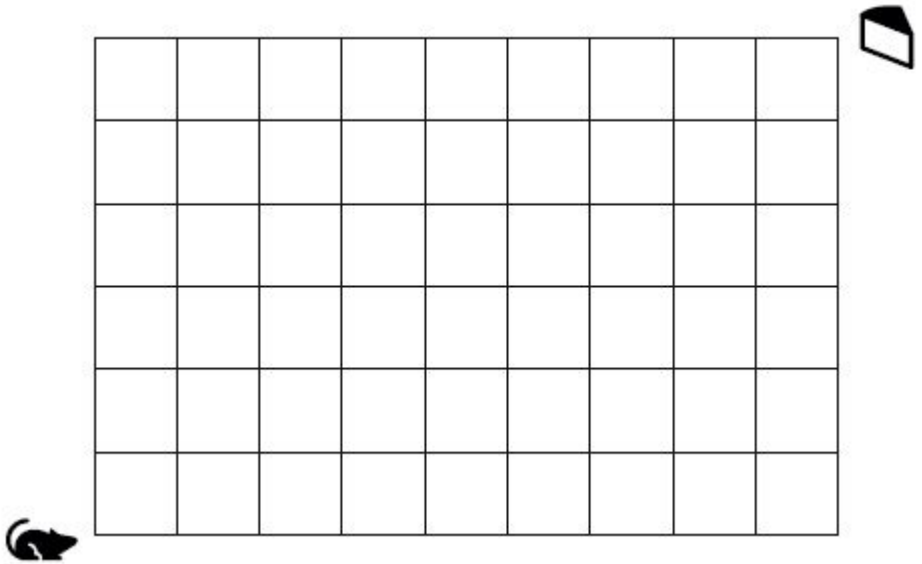
Behavioral questions

- What's the project you worked on that had the biggest impact?
- Can you tell me about a time you found something in the data that surprised you?
- What is the thing that you wanted to change most in your previous job that you couldn't?
- What would you do if you had calculations or results that conflicted with the previous results of a senior person in the company? Would you try to convince them that you were right, and if so, how?
- Tell me about a time you disagreed with a teammate. What was it about, and what did you do?
- What do you do when you don't know how to solve a data science–related problem?

Brain teasers

- What's an estimate for how many mini shampoo bottles are used by all the hotels in the United States in a year?

- Imagine a grid like the one pictured above, with a mouse at the bottom-left corner of the grid. At the top-right corner is a piece of cheese. The mouse can travel only along the lines in the grid and would never move away from it. How many paths are there from the mouse to the cheese?



Slides:

Contact:

furrer.david1@gmail.com

linkedin.com/in/david-furrer/

github.com/davidfurrer

medium.com/@fulowa

twitter.com/fulowa