

Building a Career in Data Science

David Furrer

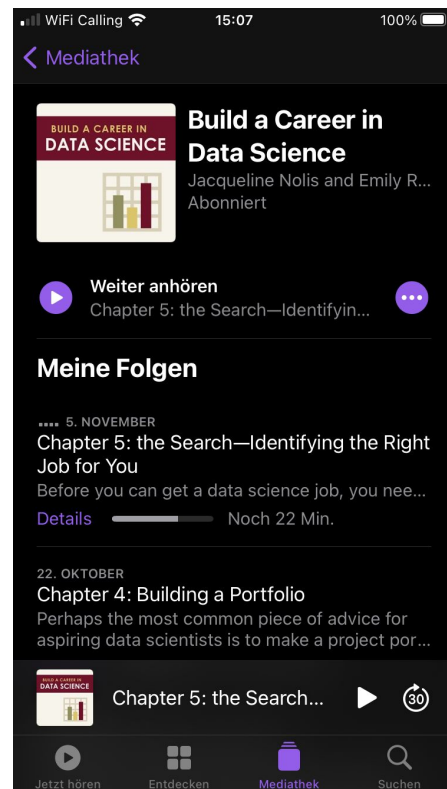
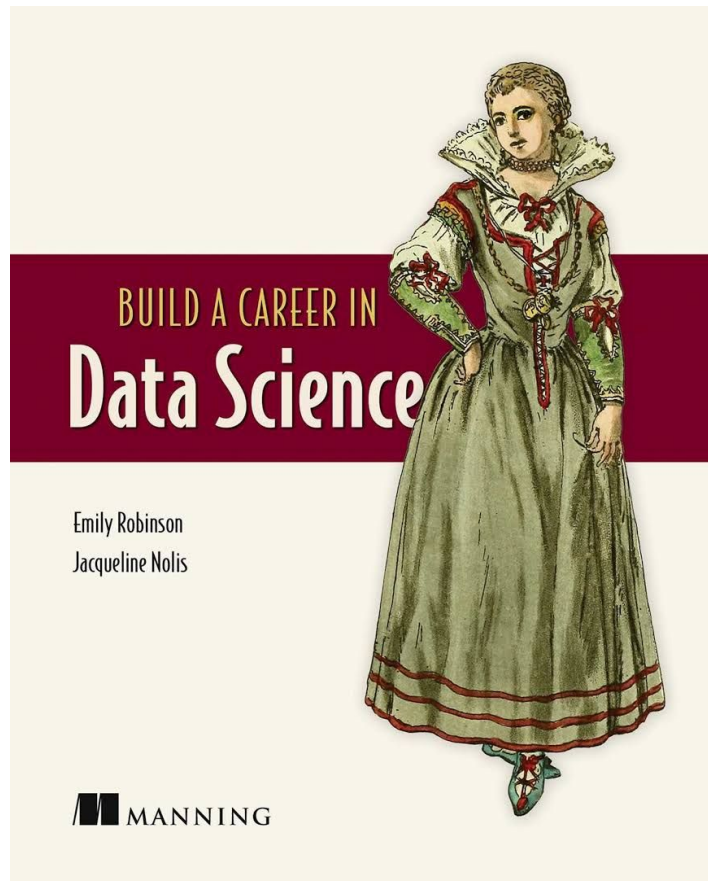
slides:
github.com/davidfurrer/talks

13.11.2020

About me

- 1) Studied Chemistry and Economics
- 2) Finished propulsion (DS track) in June 2019
- 3) Worked as Data Scientist at OTO July 2019-August 2020
- 4) Now: Marketing Data Analyst at Grand Casino Baden

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) What is it like to work as a DS?
- 9) Joining the data science communitycc

1.1) What is Data Science?

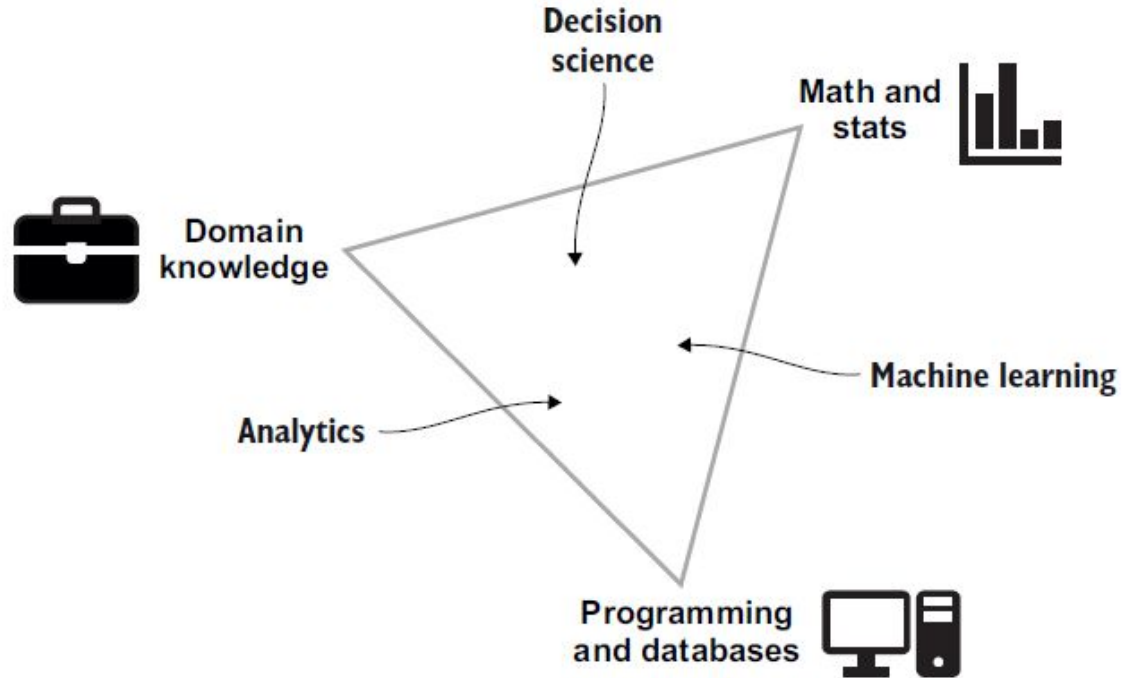


Figure 1.1 The skills that combine to make data science and how they combine to make different roles

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

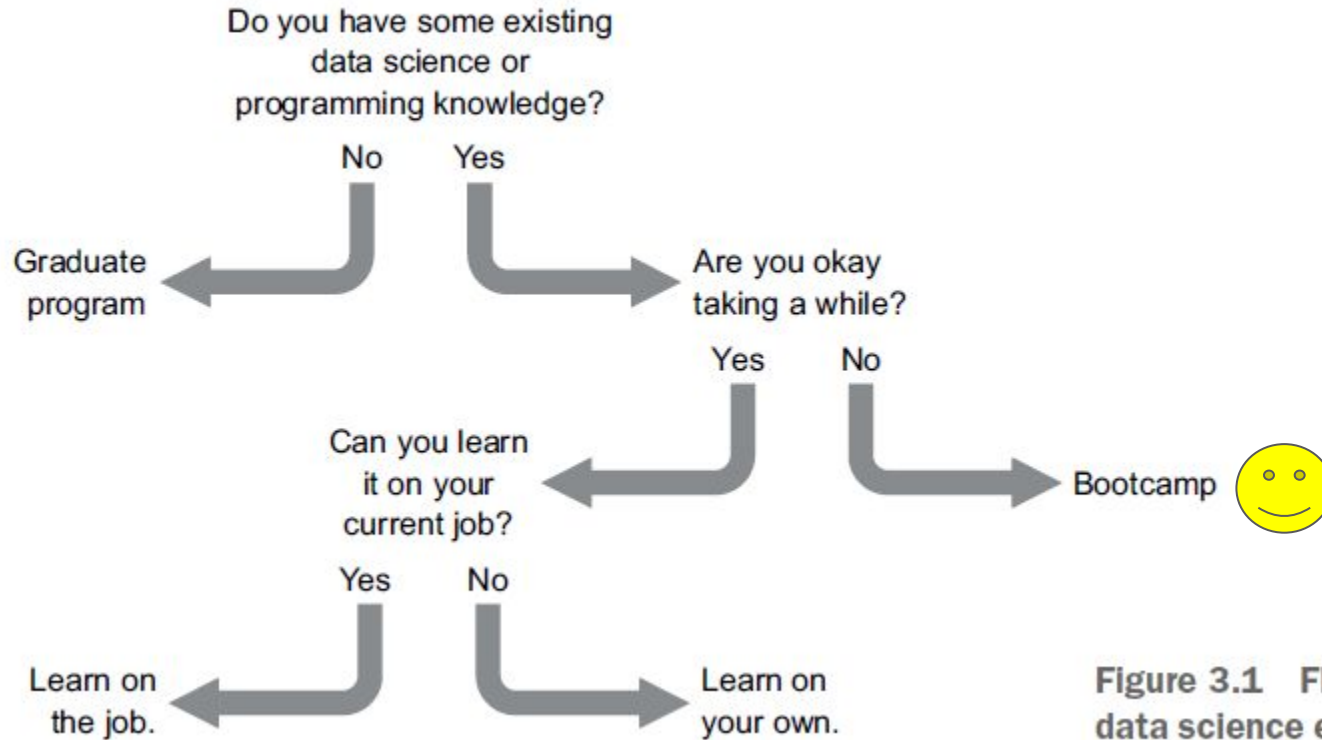


Figure 3.1 Flow of deciding which data science education route to take

4.1) Building a portfolio

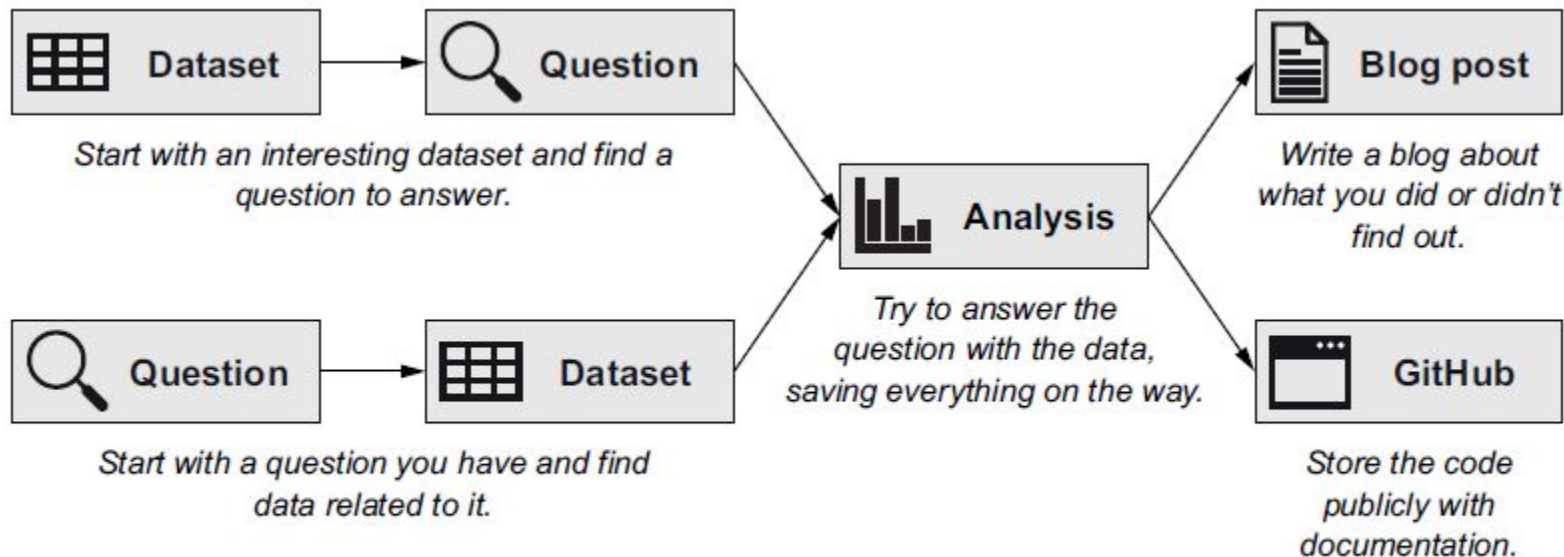


Figure 4.1 The flow of creating a data science project

4.2) Project ideas

- 1) Tidy up final project and put it on your github (e.g.: github.com/davidfurrer/bokeh-recommendation-app)
- 2) Tidy tuesday (github.com/rfordatascience/tidytuesday)
- 3) Kaggle
- 4) Go to stats website of your 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 towns 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

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

Figure 6.1 Example résumé for an aspiring data scientist

6.2) Cover Letter

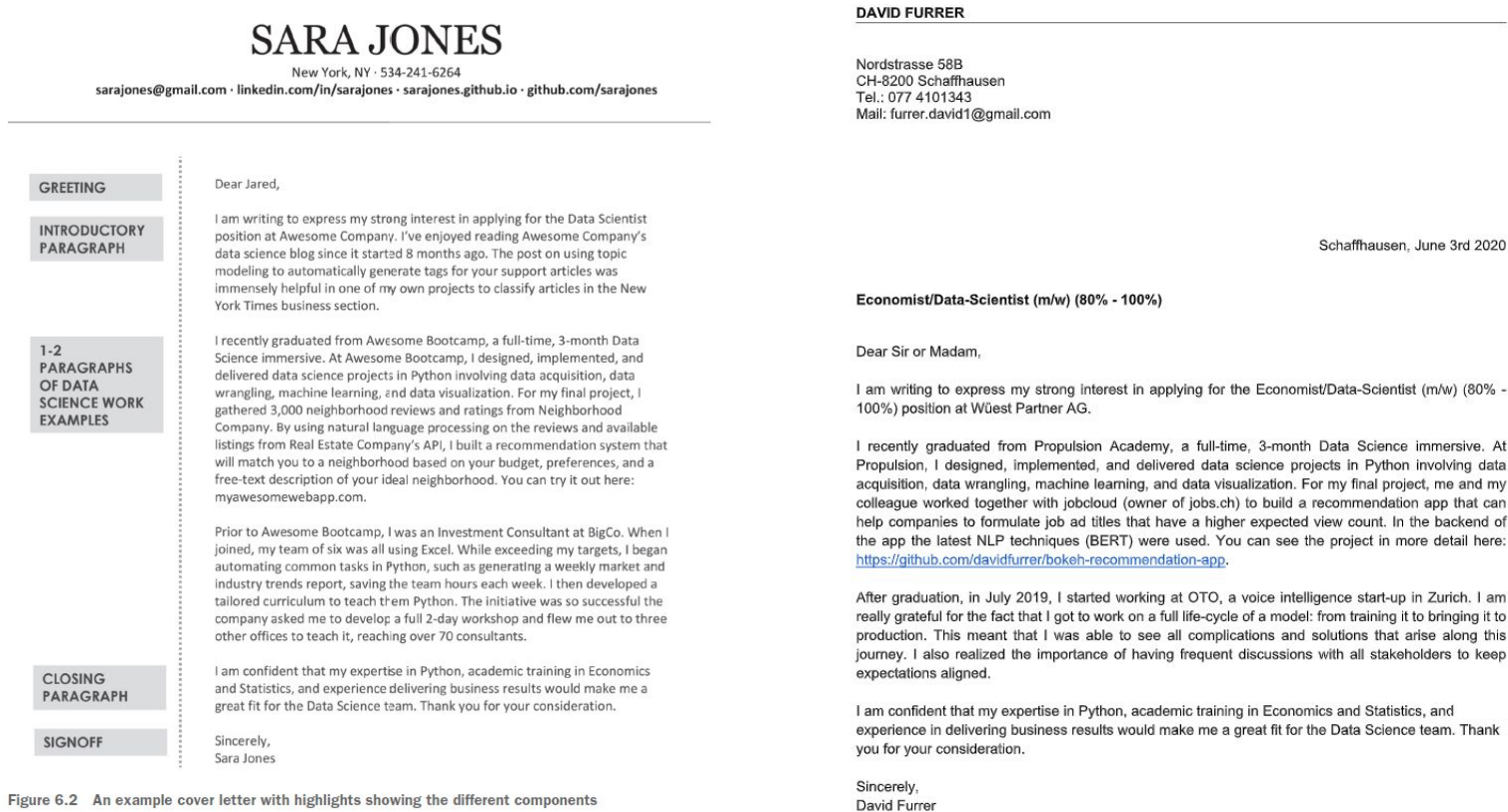


Figure 6.2 An example cover letter with highlights showing the different components

7.1) The interview

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

7.2) 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?

8.1) What is it like to work as a DS?

- You learn A LOT on the job
- Will often use cloud
- Will use git
- It's very interesting
- Can be stressful (in startup especially)

8.2) General Tips

- Don't overpromise
- Managing expectations
- Try to learn from others
- 15 minute rule
- You will get better with time
- Imposter syndrome
- Don't be afraid to learn and switch areas
- Document what you learn (github.com/davidfurrer/cheatsheets)

9.1) Joining the community

Writing
blog posts



Attending
conferences



Giving
talks



Contributing to
open source



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



Slides:

github.com/davidfurrer/talks

Contact:



furrer.david1@gmail.com



linkedin.com/in/david-furrer/



github.com/davidfurrer



medium.com/@fulowa



twitter.com/fulowa