Landing Your First Awesome Data Science Job

The tricks you need up your sleeve

November 2020

This talk is for…

- Novice data scientists (aiming for a beginner or student data science position)
- You have a BA/BSc/MA/MSc in ……. (exact sciences/statistics/computer science/math/engineering/life sciences/economy)
- Anyone with a strong passion and curiosity for data and knowledge discovery

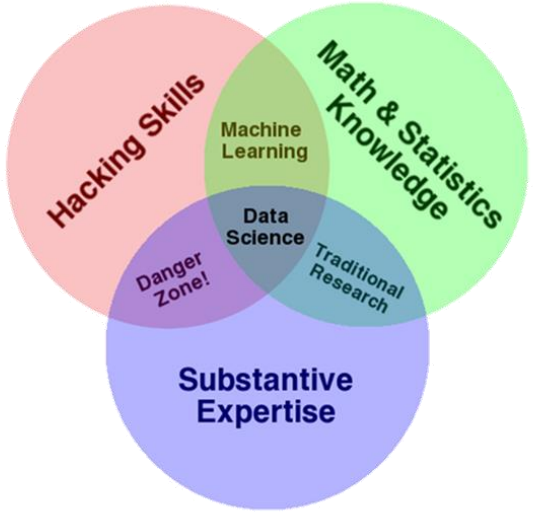
Who am I?

- Adi Sarid
- I studied mathematics, statistics and operations research in the Technion for my BA
- Continued to operations research positions in the IDF (Navy and the Planning Directorate)
- Did my MSc in operations research in the Math department and PhD in the Industrial Engineering department, Tel-Aviv university
- I am a partner in Sarid Research Institute, heading the data science and operations research department
- I mostly work with R
- I'm a certified RStudio instructor, I teach statistics in Tel-Aviv university, and provide R corporate training courses

- But this talk is not about me, it's about how to **get you into the data science Industry**

Some Context: What is Data Science?

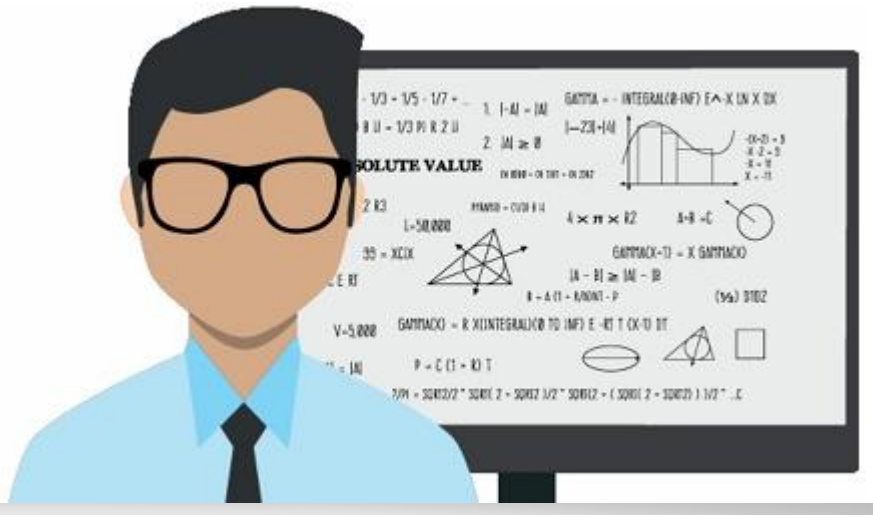
- (This is the Wikipedia definition, 2020-11-01):
  - *Data science is an inter-disciplinary field that uses scientific methods, processes, algorithms and systems to extract knowledge and insights from many structural and unstructured data. Data science is related to data mining, machine learning and big data.*
- Vague, isn't it? here's another, vague, definition (Conway's data scientist Venn diagram):
  - *Where hacking meets math, statistics, and expertise*



Artwork by @Allison_horst

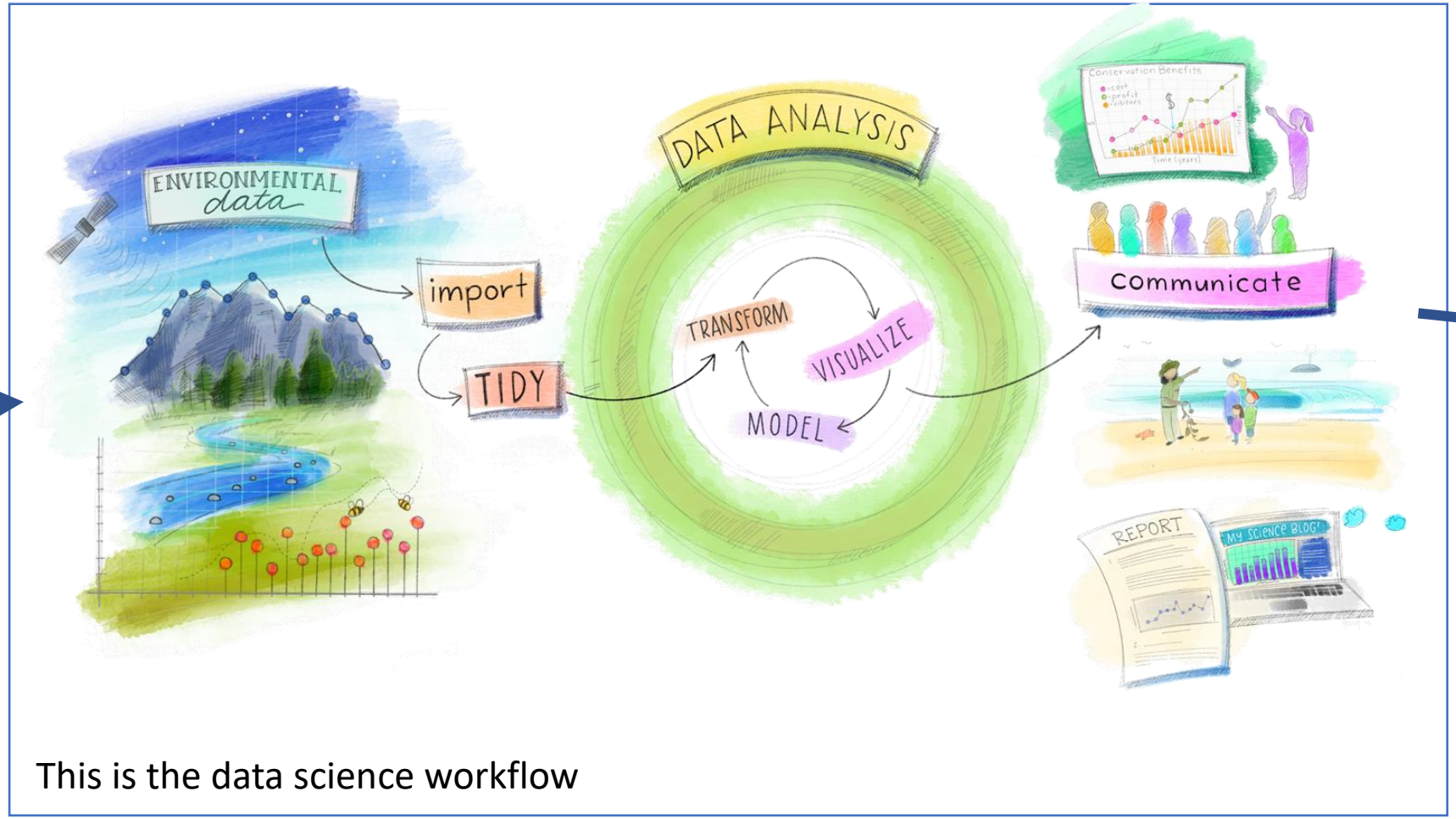
Don’t mix these up

- A data analyst is NOT a data scientist
- A machine learning engineer is also NOT a data scientist



How My Work Looks Like?

Business
Meets
Problem



This is the data science workflow

Business
Solves
Problem
and
Makes
money

Artwork by @Allison_horst

Build knowledge (academic background)

Probability

Statistics

Intro to DS + Advanced stats

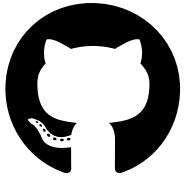
Advanced DS / ML

Is BSc enough? Maybe…, probably not

Build knowledge (tools)

- In terms of tools you should know
  - *Python* (*pandas*) or *R* or both

- You should know *git*
  - If you don't already know, learn it and start using it



- *SQL* might also help

- You should be able to prove all of it (i.e., portfolio, home assignment)

Prepare (Portfolio)

- Building a portfolio is not that complicated as you might think
  - Basic statistics + R + Kaggle/tidytuesday can get you a long way (examples)

kaggle

- If you have a personal blog, it's even better (example)

Prepare (CV)

- Emphasize uniqueness, describe what you bring to the table
- Related previous positions, **business value** you brought
- Think thoroughly on all the tools you know/learned, if you know a certain tool (even at a novice level) don't be afraid to put it (mention your level)
  - R, Python, and respective packages/languages – are tools
  - Cloud services – are tools
  - Linear regression, XGBoost, are not tools (they are algorithms)
- Make it stand out and visually appealing
- Don't lie, make sure you stand behind everything you write
- In this context: a recommended post to read

Prepare (to the specific position)

You must prepare for each interview:

- Read about the company
- Read about relevant research fields (i.e., healthcare, online commerce,…)
- Prepare a list of questions that **you** want to ask about the position, i.e.:
  - What tools are used? how the organization uses data? where the data comes from? what is your boss' background?
  - How can the job contribute to you, is this a good "first job" to start your professional career?
  - What kind of qualifications are you expected to get during your job (on job training? internal courses? external courses?)
  - Project vs. Product
  - Company culture

Prepare (to the specific position)

- Prepare a list of talking points
  - What's important to the other side?
  - What's your elevator pitch? how do you present yourself?
  - What are your added values? i.e., did you learn anything special that might be related to the position (like biology? economy?)
  - Industry or academic projects that are related?
  - How do you interact in a team? are you a "leader"? are you the "go to" person? implementor?
  - What are your expectations?

Prepare (to the specific position)

- Refresh in probability and statistics, e.g.:
  - What is the difference between supervised and unsupervised?
  - The difference between categorical and numerical variables
  - Using linear regression versus logistic regression
  - Data visualizations
  - A lot of online sources to rehearse from
    - google “data science exercises from interviews”

Home assignment

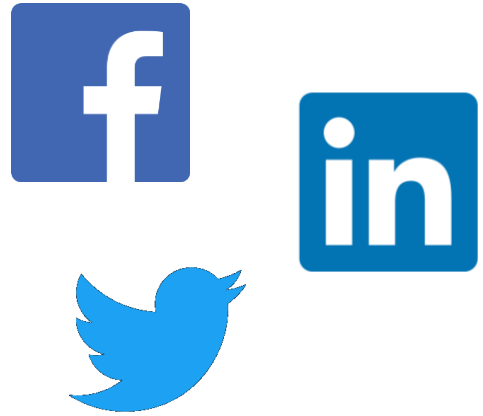
- It goes without saying, but if you want the job, think of it as a graded project in an extremely important course
- Provide very good documentation and working code
- Show what sources you used and how (i.e., stackoverflow, books)
  - It’s a good thing to show that you can learn on your own

Recommendations

- Don’t be shy – ask for recommendations
  - If you don’t have previous work experience, try asking a professor you worked closely with (i.e., in a seminar, some other project)
  - If you did an industry project, ask your industry advisor

Social footprint

- Facebook
  - Connect to relevant groups, and comment/post (e.g. “R להמונים“)
- LinkedIn
- Twitter (+ follow opinion leaders)



Sources

- Books
  - To learn more about R and data science r4ds book
  - Books (programming): https://bookdown.org/home/archive/
  - Books (theory):
    - Beginner: Introduction to Statistical Learning
    - Intermediate-Advanced: The Elements of Statistical Learning
- Blog posts
  - Seven steps to landing your dream job as a data scientist

SARID

TRUSTED INSIGHTS. PROVEN RESULTS.

Thanks!

Adi Sarid

@SaridResearch

https://www.sarid-ins.co.il