Open in app ↗

◐❘   ◯ Search   🔔   👤✨

How to create a data science portfolio in 10 simple steps
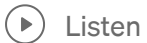
Everything that's helped me create mine

👤 Ali Uzman · Follow

Published in Bootcamp

3 min read · Mar 3, 2024

▶ Listen   ⬆ Share   ••• More

> **Dear reader, stay at least 30 seconds. Don't destroy writers' read ratio :)**

Photo by Samuel Raita on Unsplash

Creating a great portfolio is an art in itself. There's no **doubt** about that.

An updated professional portfolio is a strong way to showcase the consistency of your *style* and the *message* you represent.

It's an opportunity to tell a potential client how what you did worked, as well as how you'd approach the problem now, knowing what you know.

And even if you're not good at remembering every detail when sharing those accomplishments with others during an interview.

The portfolio provides an easy way to not only jog your memory but supplement what you're talking about.

Mostly everyone talks about the advantages of having portfolio projects. But nobody is telling you how to start.

Here are 10 steps you can follow to get started.

1. Find a business problem to solve

It's the start of any big or small project. You need to have a business problem that you are going to solve.

It can be a real-world problem people face in front of you or you can solve the problem in a business model that you think can be automated or solved using data science and AI.

In our case let's go with *(it's just an example to relate all steps)*.

Problem: Recommending soda drinks to users interested in buying different soda drinks based on their preferences.

2. Decide the ML area you want to work on

Right now AI is still a mystery box, in which there are many areas to explore. As a beginner, you need to start from somewhere.

By exploring ways you can finalize the path, which suits you better.

If I move with the above problem I would say, I will choose **'Recommendation Systems'** an area in ML.

I believe it help me and you solve the above business problem effectively.

3. Finding a dataset for your project

After the initial steps, we will go to our dataset.

You can do this **step in two** ways. The first one is easy, you can go to Kaggle and get and use many free datasets available for your projects *(i.e. soda drinks datasets).*

The second is, that instead of using Kaggle datasets, you should go to your local business or people who you know are facing problems.

This will give you experience in talking to end users, building and tweaking datasets, and much more.

Either way, you need to collect your data for your project.

4. Create an educational Jupiter notebook

Then you must set up your project properly. To show your project to clients or hiring persons.

For example: Create the main notebook and a Python file called **utils.py** that will act as the libraries that you'll call in the notebook.

5. Create the structure of the notebook

You should use 6 or more sections to divide your notebook professionally.

6 sections: introduction, data preparation, modeling, evaluation, deployment, and conclusion.

6. Explore & perform the necessary data transformations

Explore the data properly and perform all the necessary ML techniques and data transformations.

Explain each step with the help of short comments.

For example: You are going to use pandas for data preparation and bokeh for visualizations.

7. Select an algorithm

Select and highlight the algorithm you are going to use for a particular problem.

Also, explain briefly the math and its implementation.

For example: Use LightGCN from the Recommenders repo.

8. Evaluate the model and explain the results

For example: In our case use precision at k and recall at k from the Recommenders evaluation module.

9. Deploy the model

Deploy the model on your local server and create a simple API for your project.

For example: You can use cherrypy to deploy the model and show the response of a query.

10. Give the final touch

Write a short concise summary of your project.

Write a discussion of future work or the scope of the project.

And also provide references for a particular project or research if you can.

This is the end. Some note down here =).

> **P.P.S.S. Put a 🤍 in the comments if this content resonates with you.**
>
> **Ali.**

You can **Subscribe** to me here for more.

Cheers!

Data Science   Machine Learning   Programming   Technology

Self Improvement

Written by Ali Uzman

3.2K Followers · Writer for Bootcamp

Founder → Codx Labs & Startertracks.com / Editor → Write The 1% / I share stuff about startups, education, and my solopreneurial journey.

More from Ali Uzman and Bootcamp

👤 Ali Uzman

5 Free Websites to Earn $30K/Month as an Expert

Become an Online Expert to Monetize Your Expertise

✦ · 5 min read · Oct 26, 2023

👏 1K   💬 12

Ameer Omidvar in Bootcamp

Apple's all new design language

My name is Ameer, currently the designer of Sigma. I've been in love with design since i was a kid. It was just my thing. To make things...

5 min read · Feb 14, 2024

4.7K 48

Yutong Xue in Bootcamp

FAANG UX case study: Our story and 5 tips for you

Designers, case studies don't have to be boring. Turn them into stories of you going on a journey to solve a problem. Use them to showcase...

16 min read · Apr 4, 2024

965 17

Ali Uzman

I Found a Passive Income Idea Anyone Can Start

Passive income can be a great way to help you generate extra cash flow, whether you're running a side hustle or just trying to get a little...
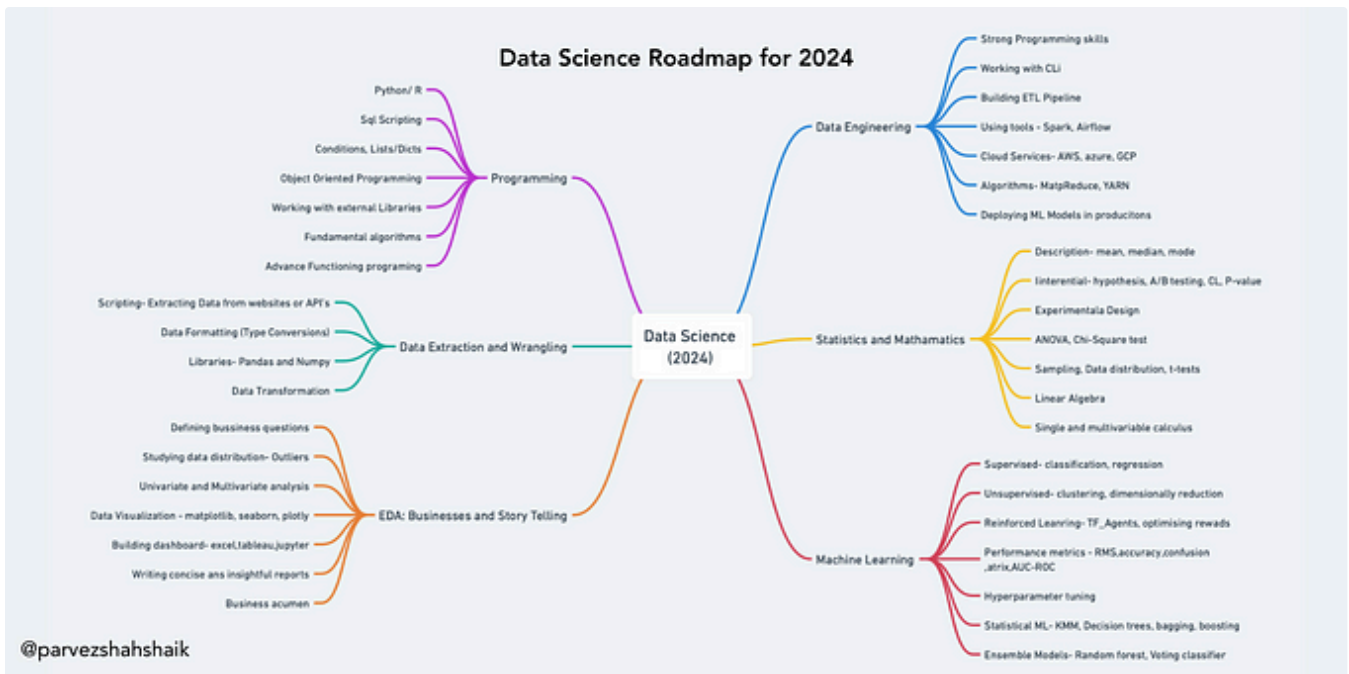
✦ · 4 min read · Dec 27, 2023

👏 1.1K   💬 21   🔖⁺   •••

See all from Ali Uzman

See all from Bootcamp

Recommended from Medium

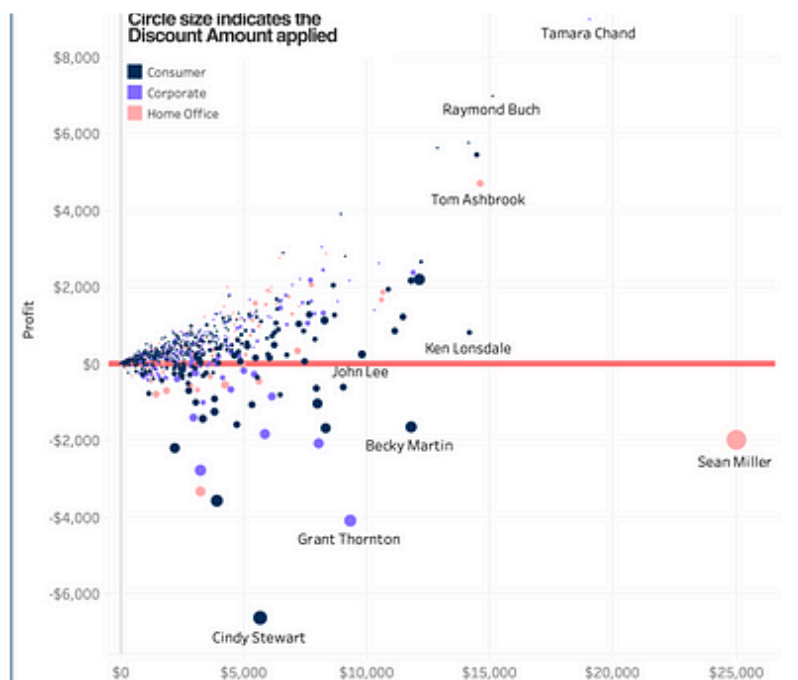🔵 Parvez Shah Shaik in Bootcamp

DataScience RoadMap for 2024

This article aims to guide you in building a strong portfolio showcasing your data science skills. It provides a framework, resources, and...

8 min read · Dec 27, 2023

👏 150   💬   🔖⁺   •••

🔵 Kim W.

Sales Data Analysis Project: 10 visualizations crafting the roadmap to business growth

and two things I've learned from this project
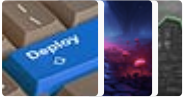
12 min read · Feb 22, 2024

👏 24   💬

🔖⁺   •••

Lists

Predictive Modeling w/ Python
20 stories · 1117 saves

Practical Guides to Machine Learning
10 stories · 1339 saves

Coding & Development
11 stories · 573 saves

ChatGPT prompts
47 stories · 1456 saves

```
*__, a, b, *_ = [1, 2, 3, 4, 5, 6]
print(__, _)
```

What does this print?

· A) Syntax error

· B) `[1] [4, 5, 6]`

· C) `[1, 2] [5, 6]`

· D) `[1, 2, 3] [6]`

· E) `<generator object <genexpr> at 0x1003847c0>`

🔴 Liu Zuo Lin

You're Decent At Python If You Can Answer These 7 Questions Correctly

No cheating pls!!

✨ · 6 min read · Mar 6, 2024

👤 Abbas Ali

Day — 1: 30 Days Machine Learning Project Challenge;
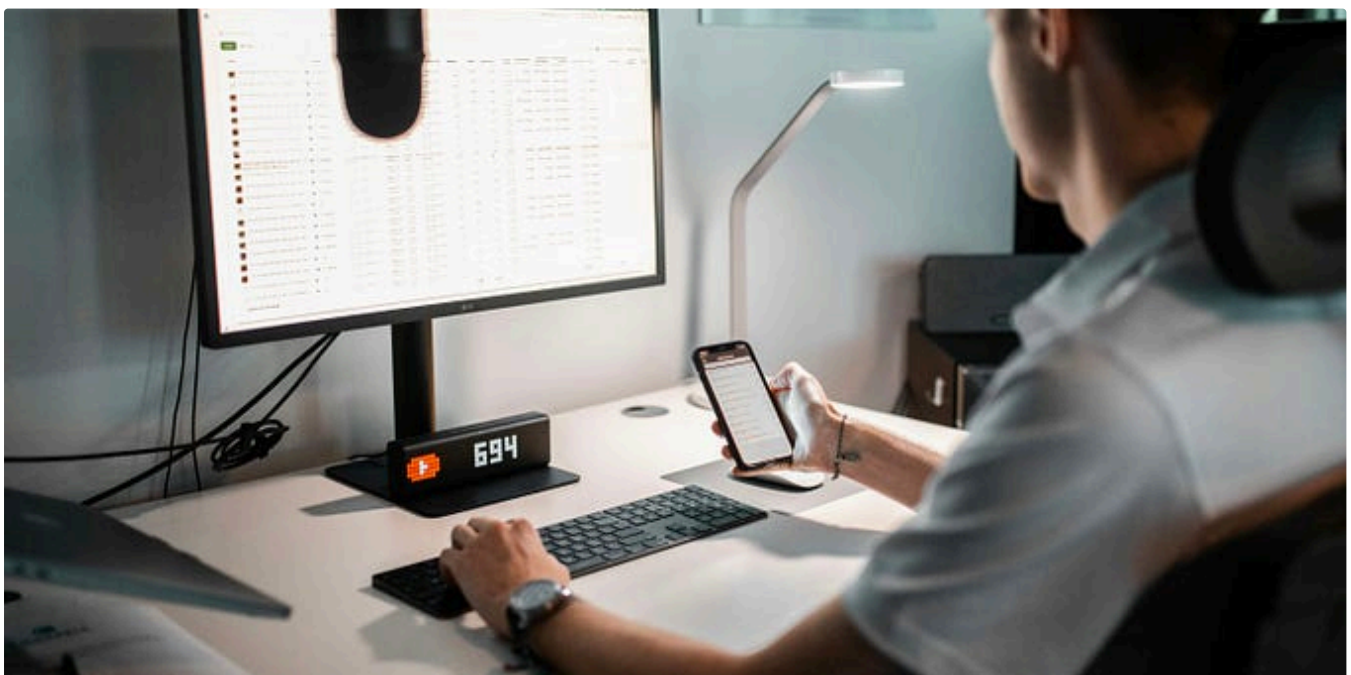
A simple project with a simple Dataset.

3 min read · Mar 3, 2024

Egor Howell in Towards Data Science

7 Subscriptions That Help Me As A Data Scientist

Subscriptions that boost my productivity, knowledge, and focus as a practicing data scientist

✦ · 8 min read · 3 days ago

👏 271   💬 5   🔖⁺   •••

Somnath Singh in Level Up Coding

The Era of High-Paying Tech Jobs is Over

The Death of Tech Jobs.

✦ · 14 min read · Apr 1, 2024

👏 8.5K   💬 247   🔖⁺   •••

[See more recommendations]