



An End-to-End Data Science Project

Deena Gergis
Lead Data Scientist @ Bayer



Sessions' goal

The series of sessions will guide you through the process of completing an **end-to-end Data Science project**.

We will start with a **problem statement** and end with a **deployed product** that our client will be able to use.

We will utilize and connect various technologies, packages and programming paradigms to produce a functional product for our (fictional) client.



Problem statement

Our (*fictional*) client is an IT educational institute. They have reached out to us has reach out with the following:

“IT jobs and technologies keep evolving quickly. This makes our field to be one of the most interesting out there. But on the other hand, such fast development confuses our students. They do not know which skills they need to learn for which job.

“Do I need to learn C++ to be a Data Scientist?” “Do DevOps and System admins use the same technologies?” “I really like JavaScript; can I use it in Data Analytics?” Those are some of the questions that our students ask.

Could you please develop a data-driven solution for our students to answer such questions? They mostly want to understand the relationships between the jobs and the technologies.



And now it's your turn...

Which Data should we use?

Our client doesn't have any internal data sources that could be used for this project.

Find the data source that you will use to build the solution



See you on Sunday!