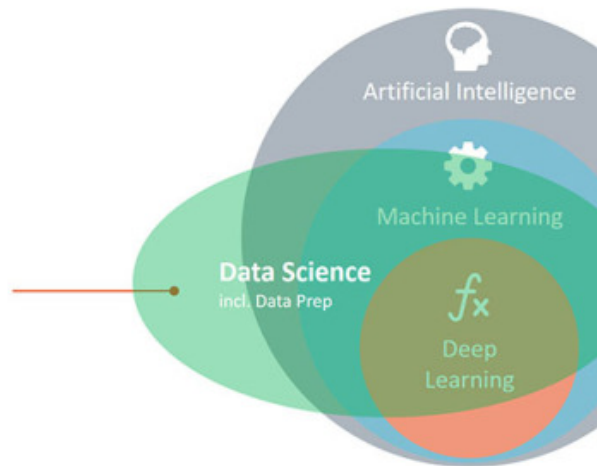**Data Science**

Covers the practical application of advanced analytics, statistics, machine learning, and the necessary data preparation in a business context.

Data Science is the practical application of all those fields (AI, ML, DL) in a business context. "Business" here is a flexible term since it could also cover a case where you work on scientific research. In this case your "business" is science. Which actually is more true than you want to think about.

But whatever the context of your application is, the goals are always the same:

- extracting insights from data,
- predicting developments,
- deriving the best actions for an optimal outcome,
- or sometimes even perform those actions in an automated fashion.

As you can also see in the diagram above, Data Science covers more than the application of only those techniques. It also covers related fields like traditional statistics and the visualization of data or results.   Finally, Data Science also includes the necessary data preparation to get the analysis done. In fact, this is where you will spend most of your time on as a data scientist.

A more traditional definition describes a data scientist as somebody with programming skills, statistical knowledge, and business understanding. And while this indeed is a skill mix which allows you to do the job of a data scientist, this definition falls a bit short. Others realized this as well which led to a battle of Venn diagrams.

The problem is that people can be good data scientists even if they do not write a single line of code. And other data scientists can create great predictive models with the help of the right tools. But without a deeper understanding of statistics. So the "unicorn" data scientist (who can master all the skills at the same time) is not only overpaid and hard to find. It might also be unnecessary.

For this reason, I like the definition above more which focuses on the "what" and less on the "how". **Data scientists are people who apply all those analytical techniques and the necessary data preparation in the context of a business application.** The tools do not matter to me as long as the results are correct and reliable.