



TesseractAcademy

Data Science Project Assessment Questionnaire

Join **Tesseract Academy's** mission. Are you a CEO, manager, or entrepreneur? We help decision-makers like you understand data science, AI, and blockchain.

You know you want to embark on a data science journey, but you are still not sure about some of the obstacles that you might encounter along the way.

Data science can be described as an exercise in risk management.

Failing to plan is planning to fail, and this questionnaire is built in order to help you understand all the risks associated with your project and how to mitigate them.

The data science project assessment questionnaire

The purpose of this document is to help you clarify your objectives, and factors that can affect the success of a data science project.

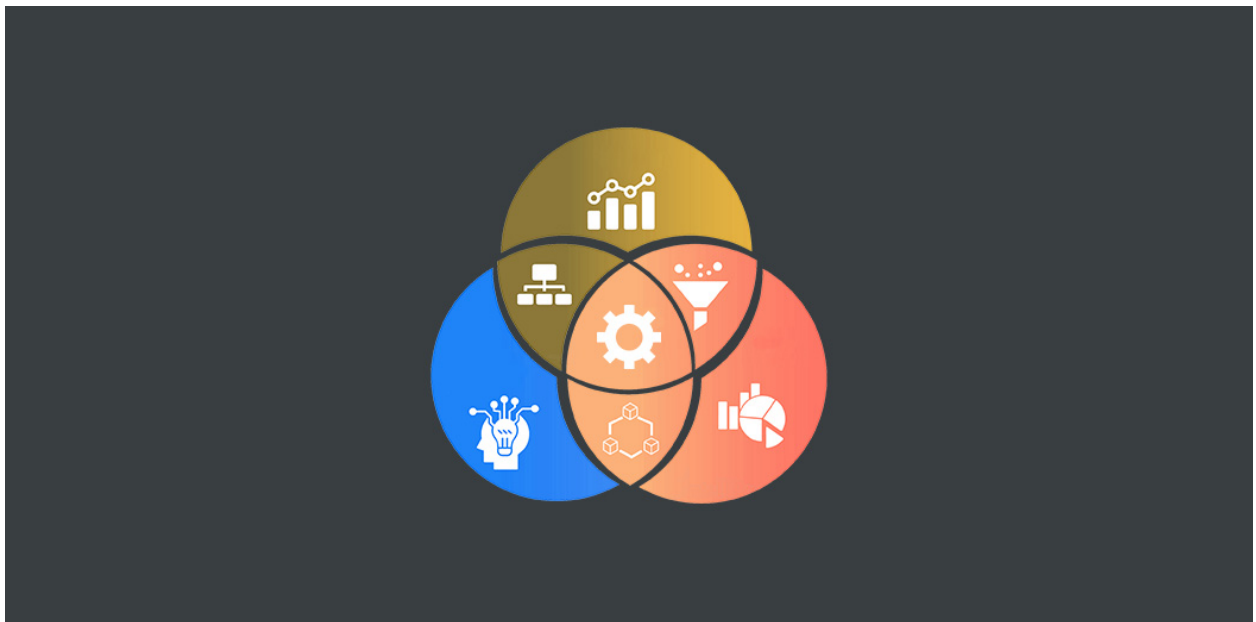
It is important to understand that a data science project is often an exercise in risk management. The problematic part in data science is the “science” bit. Quite often, it is impossible to know how well something will work in advance until you try things out. The only cases where you can be confident of results before you embark on a project are:

1. *You have worked on the same case with very similar data in the past.*
2. *There is an extensive body of literature on this type of problem.*

In any other case, the results cannot be guaranteed. Hence, the right project structure should focus on organizing an “attack plan”, where each step should produce one of either two outcomes:

1. *ful completion of the project.*
2. *Learn something about the problem (e.g., maybe a family of methods is not appropriate, which can lead to improved exploration).*

In order to better manage the project, expectations, and risk, the questionnaire below can help you list all the important requirements. This questionnaire is both for you and the data scientist, and it will help you build a common understanding of the requirements, the challenges, and how to mitigate risks. It is recommended that you do at least one round of back-and-forth between you and the data scientist, in order to make sure that there is appropriate understanding on both sides.



The questionnaire

Success criteria

What would you consider the successful implementation of the project? Share your own thoughts.

Are there any benchmarks in performance? Please try providing a numerical answer. A good answer would be (anything above 60% accuracy is good, based on a benchmark X).



Risk factors

Is the data of appropriate quality? If not, what are the issues?

Is the data of appropriate size? If not, how much do you think this will affect performance? How will you mitigate that risk?

If the data is of high quality and size, then what could possibly prevent a good model from being built (e.g., maybe the domain is particularly difficult).



Timelines

How time-critical is the project?

If the project is difficult, and no approach can reach the desired accuracy, which of the following plans seems the most attractive to you and why?

- ☐ Keep on trying more advanced approaches (e.g., ensemble modeling).
- ☐ Fix data issues and try again (e.g., collect more data)
- ☐ Simply use a model which is “good enough”, even if it doesn’t reach the desired performance.



Best/worst possible case

What would the best possible outcome be for you?

Let's say that the goals cannot be achieved. E.g., the model doesn't work to the desired performance or fails completely. What would you do in this case?



***We are on a mission to educate and help decision-makers understand
and implement data science and AI.....***

Join Tesseract Academy

