**Assignment: Data Science Scenarios**

In this first assignment, we ask you to think about some real-life process or problem in different problem domains, and how you can improve it using the Data Science process. Think about the following:

1. Which data can you collect?
2. How would you collect it?
3. How would you store the data? How large the data is likely to be?
4. Which insights you might be able to get from this data? Which decisions we would be able to take based on the data?

Try to think about 3 different problems/processes and describe each of the points above for each problem domain.

Here are some of the problem domains and problems that can get you started thinking:

1. How can you use data to improve education process for children in schools?
2. How can you use data to control vaccination during the pandemic?
3. How can you use data to make sure you are being productive at work?

**Instructions**

Fill in the following table (substitute suggested problem domains for your own ones if needed):

| **Problem Domain** | **Problem** | **Which data to collect** | **How to store the data** | **Which insights/decisions we can make** |
| --- | --- | --- | --- | --- |
| Education |  |  |  |  |
| Vaccination |  |  |  |  |
| Productivity |  |  |  |  |

**Rubric**

| **Exemplary** | **Adequate** | **Needs Improvement** |
| --- | --- | --- |
| One was able to identify reasonable data sources, ways of storing data and possible decisions/insights for all problem domains | Some of the aspects of the solution are not detailed, data storage is not discussed, at least 2 problem domains are described | Only parts of the data solution are described, only one problem domain is considered. |