**Getting to the Core of the business problem**

* This section will help us become a great data scientist instead of just a good one. Thinking from the business point of view and being able to add value is more valuable than just being able to code and memorize machine learning models.

**1) What is the business goal for this project?**

- What it means: To get a high level understanding of what the project is. How the outputs should look and the metrics for measuring success.

- Why we ask: This helps frame and direct the entire project. Redirects our thinking toward the best solutions. To get the main goals and if possible push for specific goals.

- Who might have the answer: Key stakeholders. Project managers guide you to the source.

**2) Can this project be solved or enhanced through the use of data, and Data Science?**

- What it means: To understand if the data science team is needed.

- Why we ask: For efficiency because data scientists are usually busy.

- Who might have the answer: Talk with key stakeholders and our guidance is key to answer this question, since we are the data experts and know what we can do with the data we have access to.

**3) What is the Minimum Viable Product (MVP)?**

- What it means: MVP is a version of the product or solution with the minimal amount of features to satisfy the goal of the project. Test the waters and prove out the idea to the business.

- Why we ask: So we don't go too far on a project that is not viable or useful to the business. Higher likelihood of solution going into production if all key stakeholders agree on the MVP. If the project is a failure, it will fail fast and not waste resources.

- Who might have the answer: All teams involved because it is essentially a mini solution for our project. Data science team will play a big part because we are the data experts.

**4) When does the project need to be implemented or completed?**

- What it means:

- Why we ask: This information allows us to think about what is and is not possible, and for allocation of resources to the project. Also have to know the timeline for MVP to be completed.

- Who might have the answer: Data science manager.

**5) Are there any considerations around ethics or compliance?**

- What it means: We have to be aware around the permissions we have for certain data sources.

- Why we ask: We have to be aware of potential bias in historical data, because our models will make predictions that will propagate the bias.

- Who might have the answer:

**Deciding on the right Data Science solution**

**1) What work has been done before in this area?**

- What it means: Has the company already done work like this and who has done it.

- Why we ask: Efficiency. We have to know if we have to start from scratch.

- Who might have the answer: Key stakeholder and managers of any data teams. Take any leads that you can.

**2) Who are we relying on, and who is relying on us?**

- What it means: Who do we rely on before we can start our work, and who is relying on us before they can start their work.

- Why we ask: Figure out the workable solution for a dependent team so they we are not delivering something that is not useful. Figure out timeline and how frequently you should update dependent teams on progress.

- Who might have the answer: Dependent teams and teams that we may need data from.

**3) What data is available, where is it, and what state is it in?**

- What it means: Determining what data we can use and how much effort required to make it useable. Also can figure out if there is data from external sources that we can tap into.

- Why we ask: Because this affects timelines and we need to report back to the data science manager. We will also be able to figure out what is required of us to make the data usable for our project.

- Who might have the answer:

**4) What does the data tell us?**

- What it means: We can explore the data and feed back insights to key stakeholders.

- Why we ask: We can give insights to key stakeholders to see if the data is answering the business goals of the project.

- Who might have the answer:

**5) What data is valuable, and what is not - and can we enhance the data further?**

- What it means: Narrow down data only to that which has value to the overall business goal.

- Why we ask: We’ll want to see if we can enhance the data in any way through our human intuition and from what we have learned from speaking to key figures in the business about this project. Try to put yourself in the shoes of the end user and try to empathize with their needs.

- Who might have the answer:

**6) What tool or approach should we use, and why?**

- What it means: Determine what tool to use when there can be many options. Sometimes the approach to use is obvious or there may only be limited options.

- Why we ask: This allows us to justify certain approaches we take with our projects. Documenting our approaches and insights along the way allows us to give others insights on their projects if we’re ever asked.

- Who might have the answer:

**7) How do we sync Data Science success metrics with the overall business success metrics?**

- What it means: Our success metrics within the models won’t mean a lot to key business decision makers, so we have to frame the results in a way they’ll understand. Use examples and use the metrics they care about the most.

- Why we ask: This will increase our likelihood of having our solutions implemented and can add value to the business.

- Who might have the answer:

**8) How will this project or solution change over time?**

- What it means: Better to think about how a project can change when you are building it and not down the line. Also can get a feel for how well our solution will perform as time goes by.

- Why we ask: We can put in performance tracking on models to verify if they are accurate over time, or may need to be updated at some point. This can be looked over if we are laser focused on getting our solutions into production, but this can be really important.

- Who might have the answer: