The Scoping Document is your first official milestone in the DS4A Practicum requirement for certification. It completes the **planning phase** of this data scientist project. Please find answers to common questions below and a template here.

Project Funnel

planning phase:
your scoping document will include your ideas, scoping, and business case.

Why write a Scoping Document?
Since you are being trained to be leading data scientists, you are not just a simple analyst or programmer: you also need to communicate your results *and your process*. Any data science project can roughly be split into three parts: (1) a planning/scoping phase, (2) an execution/iteration phase, and (3) a delivery/reporting phase. It is crucial every step of the way you communicate what you are doing to key stakeholders: this can be your boss, your coworkers, your employees, business users waiting for your analysis, designers waiting for your insights, investors waiting for your predictions… etc. This is your chance to practice this. Please refer to this template as an example.

What are the benefits of writing a good Scoping Document?
Doing this well has many benefits:
1. It gives you clarity on what you've done and what you will do;
2. It eases other people's confusion or anxiety about your progress;
3. It gives others the chance to offer advice or ask questions about your work (and therefore improve your work!).

While engineering and data science are fundamentally very different disciplines, data scientists can benefit from learning more about the common engineering principle of the phase-gate process. Top companies like Google and top business programs like Harvard teach this sort of management.

Who will be reviewing them?

Your assigned TAs will be reading them in depth and providing feedback, but Natesh and Haris will also be assessing them.

What is the minimum information to include?

You must include the business problem and why this is relevant. You must include what data you plan on using to solve it, as well as a general description of the methods you will try. You must outline the interface for the application, the timeline for what you plan on delivering for the final report, and any other commentary that is required for the project that may provide the audience with context (e.g., key assumptions, restrictions). You may want to organize your document by sections (the words I have underlined are suggested sections). Please refer to this template as an example; your team can copy and modify it.

What is the ideal information to include?

Ideally, you will provide detail on the business impact on the problem you're solving (e.g., money generated, time saved). You will also provide detail on the datasets available (e.g., size, fields, accuracy, missing data). You will also include more details on the approaches you will try: strengths, weaknesses, and open questions you have. Finally, you will provide detail in your plan on milestones in project delivery:

Version 1: what you know that you can do
Version 2: Version 1 + specific feature
Version 3: add very challenging analysis

For example:

Version 1: Build simple pricing model using historical demand/supply
Version 2: Add seasonality to the pricing model
Version 3: Use competitor's prices in pricing model

How do we submit?

Please have *only one member of your team* submit your final Scoping Document in a *.pdf* format via the submission portal on the DS4A website.