***Framework One - Zumel and Mount, Practical Data Science with R, chapter 1:***

Define the goal The first step in a data science process is to define a measurable and quantifiable goal.

* Why do the stakeholders want to do the project?
* What do they need from it?
* Why is their current solution inadequate?
* What resources do you need?
* How will the result of your project be deployed?

Collect and manage data This step includes identifying the data you need, then exploring and conditioning it. This is often the most time consuming step.

* What data is available?
* Will it help to solve the problem? Is it enough?
* Is the data quality good enough?

Build the model Here is where you try to extract useful insights from the data in order to achieve your goals.

* Which techniques might I apply to build the model?
* How many techniques should I apply?

Evaluate and critique the model Once you have derived a model, you need to determine whether it meets your goals. If not, it’s time to loop back to the modeling step.

* Is the model accurate enough to meet the stakeholders’ needs?
* Does it perform better than “the obvious guess” and any techniques being used currently?
* Do the results of the model make sense in the context of the real-world problem domain?

Present results and document Once you have a model that meets your criteria, you will present your results to your project sponsor and   other stakeholders.

* How should stakeholders interpret the model?
* How confident should they be in its predictions?
* When should they potentially overrule the model’s predictions?

Deploy and maintain the model Finally the model is put into But you still need to ensure that the model will run smoothly. In many cases this requires enhancement of the requirements based on customer feedback or in some cases fixing bugs.

* How is the model to be handed off to “production”?
* How often, and under which circumstances, should the model be revised?

**R**ecommendation

* How can you most effectively present the results of your analysis to your stakeholders (in terms they can understand and in alignment with information they’ll value)?
* Note: A generic template for a recommendation presentation or report might include:
* Objective
  + Background (optional)
  + Scope (optional)
  + Approach (optional)
  + Recommendations
  + Key insights with impact
  + Next steps

Create a report to your client's management in the form of a PowerPoint presentation. Your report should include:

1. A written statement of the goal(s)
2. A well-defined data science process framework and the reasons you are proposing it
3. Descriptions and location of related data sources
4. An explanation of how you will manage the data for the project
5. Any known issues with the data and how you plan to address them
6. A flowchart visualizing the detailed process you will follow, annotated with any potential pitfalls you’ve identified and your proposed solutions to such pitfalls.
7. Any initial insights you can glean from your quick look at the data.

**TIP:**

You can use LucidChart, Vizio, Dia, PowerPoint, et cetera to produce your flowchart.

**TIP:**

Remember that your report is intended for business rather than technical people. Consider the things likely to be important to a business person and the things likely to confuse them. If you believe that a technical detail is likely to be important, think about how best to explain it to a business audience.