

Project Planning

DS 6011: Capstone Part I
School of Data Science
University of Virginia

Last updated: September 30, 2023



Most people don't plan to fail,
they fail to plan.

~ John J. Beckley

Agenda

Data science projects vs traditional software development

What do do: figuring out stages and sequence

Coordinating tasks

Collaboration

Meetings with Sponsor

Data Science vs Software Projects

Data science projects are different from software development

There is a lot more research and iteration required

Doesn't always fit a tidy schedule (2-week sprints)

DS is based on probability which brings uncertainty

Sprints for Data Science

Even still, activities can be time boxed:

e.g., dedicate 2 weeks for research on certain task

Start by working backward: get clear on what success looks like

Make a list of stages (discovery, lit review, data processing, modeling)

Put the stages on a rough calendar

Timeline

Gantt chart

Task	Dec	Jan	Feb	Mar	Apr	May
Data discovery						
Lit review						
Data processing / mapping						
Model development						
Dashboard & reporting						
Finalize deliverables						

When forming the timeline, take into account:

- capacity of team members
- holidays / time off
- important events (e.g., SIEDS if participating)

Share this timeline with sponsor, faculty mentor to get feedback, buy-in

Timeline management

Gantt chart

Task	Dec	Jan	Feb	Mar	Apr	May
Data discovery						
Lit review						
Data processing / mapping						
Model development						
Dashboard & reporting						
Finalize deliverables						

The chart will change over time and get more detailed

You might add more tasks

Can probably coordinate tasks (e.g., who is owner) without discussing with sponsor

Coordinating Tasks

For managing the project, it will be helpful to have two things in a shared location:

1) Stack of tasks that need to get done:

- name of each task
- task description

2) Tasks in the current sprint

- task name
- task description
- who is working on the task
- owner of task: who is responsible
- status
- any blockers

Collaboration

As the team makes progress / completes sprints, want to share broadly

Team should decide how to do this (and it should go into team contract)

Examples:

- Can update work on current sprint by a certain date and ask members to review
- Can discuss at a meeting
- For blockers, may want to reach out for faster feedback (Teams, email)

Meetings with Sponsor

Prepare a deck for sponsor meeting. Should include:

- Agenda
- Reminder of success criteria (to ensure alignment)
- Timeline (discuss if it needs to change)
- What you've worked on during this sprint:
progress, findings, decisions, blockers
- Plan for next sprint: make sure people have bandwidth and adjust as needed.

Meetings with Sponsor, contd.

Get a sense of how detailed your discussion should be

As you review work, pause, take questions, ask if the details are useful, etc.

It's important to get feedback, keep attention, and have clear next steps