

# Week 2: Project Forming

## DATA 510: Data Science Capstone

Lucas P. Cordova, Ph.D.

2026-05-18

Second block, week 2: capstone requirements and milestone placement, a second pass on project brainstorming and PRIDE-sharpened research questions with peer review, introduction to the Data-Driven Scrum operating model for solo and team work, and a Canvas snapshot of project domain, stakeholders, questions, and teaming intent.

### Table of contents

<b>1 Learning Objectives</b>	<b>1</b>
<b>2 Part 1: Capstone Requirements (Context)</b>	<b>2</b>
<b>3 Part 2: Where We Are on the Calendar</b>	<b>3</b>
<b>4 Part 3: Activity 1 — Brainstorming Iteration and PRIDE</b>	<b>4</b>
<b>5 Part 4: Data-Driven Scrum Operating Model</b>	<b>6</b>
<b>6 Part 5: Activity 2 — Project Direction Snapshot</b>	<b>8</b>
<b>7 Wrap-Up</b>	<b>9</b>

### 1 Learning Objectives

#### 1.1 Today's Objectives

#### 1.2 What You Will Leave With

By the end of this session, you will be able to:

1. **Place** tonight’s work on the semester milestone line.
2. **Revise** last week’s project thinking: keep, sharpen, or drop ideas based on new evidence and conversation.
3. **Apply PRIDE** to draft or stress-test at least one capstone-scale research question.
4. **Explain** how **Data-Driven Scrum** artifacts and meta-project peers fit solo work and self-selected teams.
5. **Submit** a short **project direction snapshot** on Canvas (domain, stakeholders, questions, solo or proposed teammates).

## 2 Part 1: Capstone Requirements (Context)

### 2.1 What You Are Building

The DATA 510 capstone is a **semester-long culminating project**: you propose, build, evaluate, and communicate a consequential data science outcome that draws together your MS coursework as **one coherent story**, not a pile of disconnected homework.

The project is **flexible by design**. You should focus effort on a problem you (or a small team you choose) care about, in a domain where you can sustain curiosity through August. That passion is an asset, not a loophole.

...

You still owe a **defensible, integrated** capstone: engineering, analytics or ML, visualization and communication, ethics, and research design should all show up somewhere in the plan. Silo pitches (“only a dashboard,” “only a notebook”) get pushed to deepen before approval.

### 2.2 Team Size and Meta-Project Peers

Dimension	Rule
Execution	<b>Solo</b> , or a <b>self-selected team of 2 to 3</b>
Scope bar	Multi-person teams need <b>noticeably higher</b> scope than a comparable solo project; <b>my approval</b> at the proposal
Meta-project cluster	<b>Assigned</b> by me: your project sits beside <b>two peer capstones</b> for the whole term

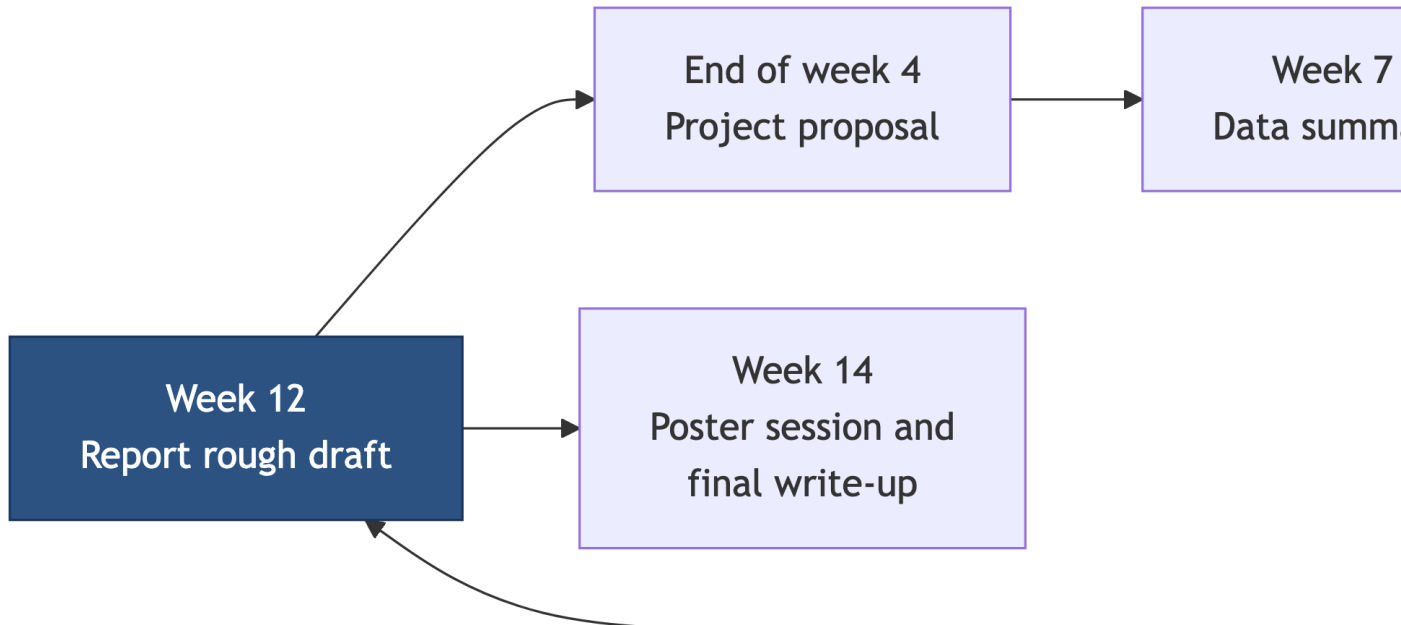
You are **not** merged into one mega-team with those peers. You remain owners of your backlog and milestones. You **do** follow each other's boards and weekly summaries and give structured feedback every week, the way parallel teams stay visible in real organizations.

### 2.3 Data and Approval

Data sources must be **selected and approved by me** early. If access, license, or ethics look shaky, surface that in tonight's snapshot and in next week's charter work.

## 3 Part 2: Where We Are on the Calendar

### 3.1 Milestone Stream



**You are here:** project forming tonight. The **project proposal** locks direction, data plan, ethics, and methods and is due **by the end of week 4**. The **data summary** shows ingestion is stable enough to stop firefighting data; it is due **week 7**.

...

**Data-Driven Scrum** is how you steer week to week toward those graded milestones with visible process.

## 4 Part 3: Activity 1 — Brainstorming Iteration and PRIDE

### 4.1 Activity 1 Overview

**Goal:** Move from last week’s wide exploration to a **defensible direction** you can charter and propose. You may **keep and sharpen** an idea, **merge** threads you heard from classmates, or **scrap** a direction that no longer survives feasibility or ethics checks.

**Bring:** Your submission from last week’s brainstorming session.

### 4.2 Step 1 — Solo Revisit

On your own, answer in writing (notebook or device; not submitted separately):

1. **What changed since last week?** New data leads, dead ends, conversations, or readings from research methods block.
2. **One sentence domain:** What world or organization does this capstone live in?
3. **Keep, pivot, or drop:** Which north-star idea still deserves semester-scale effort? If you drop one, name what replaced it and why.

If last week’s idea still feels vague, that is fine. Name what you need to learn in the next two weeks to choose.

### 4.3 Step 2 — Short Exchanges

Talk with **two or three** different classmates (fewer than week one; deeper this time).

Exchange:

- Domain and who might care (stakeholders).
- One draft **research question**, not a tool list.
- Biggest **risk** (data access, ethics, scope, evaluation).

Capture one line per conversation for yourself. You will reuse this in the Canvas snapshot.

## 4.4 Conversation Norms (Same as Week One)

### Do

- Offer a concrete angle when you can (source, metric, stakeholder, method).

### Avoid

- Shutting down an idea in the first thirty seconds.
- Promising a permanent team tonight unless you and the other person are ready to name your team tonight.

## 4.5 Step 3 — PRIDE on Your Question

Use **PRIDE** on the question you are actually considering for the proposal:

Step	Tonight's focus
<b>P</b> Problem and impact	Who is affected? What improves if you succeed?
<b>R</b> Review and gap	What is known? What is unknown and matters for your setting?
<b>I</b> Inquiry	Primary (and optional secondary) <b>research question(s)</b> in plain language
<b>D</b> Data and ethics	What data exist or could exist? Consent, fairness, retention risks?
<b>E</b> Evidence plan	What would count as answering each question (metrics, baselines, design)?

Draft at least **P**, **I**, and one bullet each for **D** and **E**. You do not need a full literature review tonight.

## 4.6 Step 4 — Neighbor Review

Pair with someone **not** in your proposed project team (if you already have one).

Trade drafts. Each reviewer answers:

1. Can you tell **what would count as success**?
2. Is the question **feasible** this term with plausible data?
3. One **strength** and one **revision** (specific, kind, actionable).

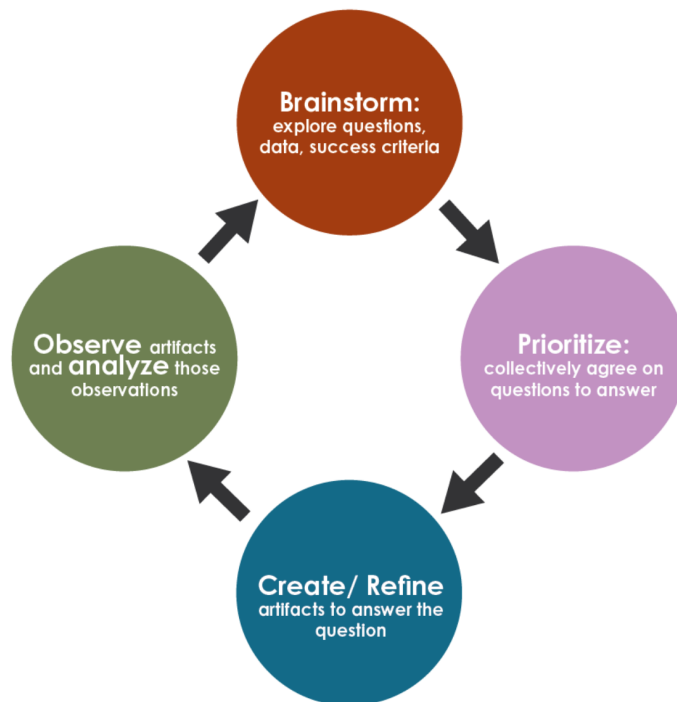
Revise your question once based on the review.

## 5 Part 4: Data-Driven Scrum Operating Model

### 5.1 Why DDS in This Course

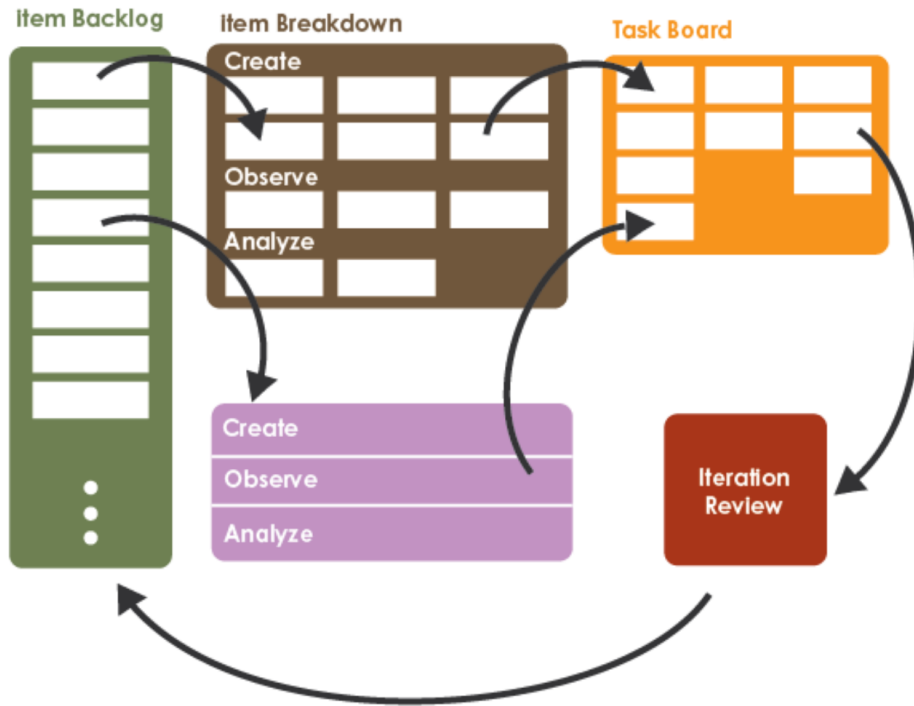
**Data-Driven Scrum** keeps capstone work **transparent**, **prioritized**, and **iterative**: short cycles from questions to experiments, observe results, reprioritize the backlog.

### 5.2 High-Level Flow of Work



1. **Brainstorm** questions or experiments as backlog items (stories, spikes, hypotheses).
2. **Prioritize** given current data and modeling needs; pull the top item into focus.
3. **Create and refine** pipelines, models, visuals, documentation tied to that item.
4. **Observe** results together and **reprioritize** based on findings and risks.

### 5.3 Core Artifacts You Maintain



- **Backlog**: prioritized items.
- **Item breakdown board (IBB)**: decomposition during refinement.
- **Task board**: visible workflow (to do, in progress, done).
- **Weekly progress report**: iteration review, retrospective, and next week's planned items (often in the repo README per Canvas prompts).

### 5.4 Solo Project: What DDS Looks Like

**Your repo and boards**: You own the backlog and task flow. Weekly README summaries make progress legible to me and to your **meta-project peers**.

**Meta-project cluster (assigned)**: Two other capstone projects read your board and summary each week. You do the same for them. They are **not** co-authors on your code; they are a standing review panel.

**In class**: Standups and backlog refinement finish **during the scheduled meeting** when we run them, including cross-team touchpoints in your cluster.

## 5.5 Team of 2 or 3: What DDS Looks Like

**Shared execution:** One backlog and task board the team maintains together. Division of labor should be visible on the board, not only in private chats.

**Higher scope expectation:** More integration, evaluation, or surface area than a solo project at the same quality bar. Proposal week is where scope gets approved.

**Meta-project cluster (still assigned):** Same as solo: two **outside** projects follow your work weekly. Your teammates are **not** a substitute for meta-project feedback.

## 5.6 Meta-Project Feedback (Reminder)

Good peer comments tie to **visible evidence** (board column, README section, plot, schema sketch), separate curiosity from **blocking** concerns, and surface ethics and engineering risks early.

# 6 Part 5: Activity 2 — Project Direction Snapshot

## 6.1 Activity 2 Overview

**Goal:** One concise submission so I can form **Canvas project groups** and meta-project clusters before chartering and proposal work next week.

**Due:** End of tonight's class on **Canvas** (see assignment page for the upload link).

**Who submits:** If you are proposing a team, **one person** submits for the group; list all members in the form. If you are working **solo**, you submit individually.

## 6.2 What to Submit

Use the **Canvas assignment** instructions. Include:

1. **Project domain** (one short paragraph or tight bullets).
2. **Stakeholders** (who benefits, who might use or care about results).
3. **Research question(s)** (numbered; sharpened after PRIDE and neighbor review).
4. **Teaming: Solo**, or **proposed teammates** (full names as they appear in Canvas).

Optional but useful: one sentence on the biggest **open risk** (data, ethics, scope).

## 6.3 After Tonight and Next Week

**Tonight:** Project direction snapshot on Canvas.

**Next week: Project chartering** exercise (part of the proposal package) and **proposal** activities in class. Come with the same domain and questions unless you document a pivot.

**Week 4: Project proposal** due by the end of class.

Keep last week's brainstorming notes and tonight's PRIDE draft. You will reuse the language in charter sections and backlog items.

## 7 Wrap-Up

### 7.1 Before You Leave

- Canvas snapshot submitted (solo or one per proposed team).
- You can state your **primary research question** aloud in one breath.
- You know the difference between **your team** (optional, 1 to 3) and **meta-project peers** (assigned, 2 other projects).

**Questions on data approval, teaming, or scope:** note them in the Canvas form or email after class.