Team Project

```
$ echo "Data Sciences Institute"
```

Today

1. Team Project guidelines

2. Developing a project plan

3. Git review

Team Project: Overview

Goal

Develop a program that uses data creatively to solve a problem or provide insights that have a positive business impact.

Learning Outcomes

1. Resolve merge conflicts.

2. Describe common problems or challenges a team encounters when working collaboratively using Git and GitHub.

3. Create a program to analyze a dataset with contributions from multiple team members.

Key Questions

- How will you select your dataset?

- How will you make sure all team members can contribute to the project?

- How will you make decisions?

- What is the question you're trying to answer through your data analysis?

- What tasks need to be completed to get to your final output?

Requirements

1. Each team member must **create, review, and merge a pull request**.

2. Each team member must **write one PR description**.

3. Each project must **design, implement, and test regression**.

4. Each team must **create a README** that explains the project, how the team approached their project, and any relevant details to their team project, including the team's Rules of Engagement. Keep in mind that the repo will be on each team member's profile, so be thoughtful about the details you include here.

5. Each team member must **create a video** about their learnings and experience.

Submission

- **One** person from each team must fork the `team_project` Git repository. Other team members should clone the forked repository.

- The folder structure of your forked repository may be modified to suit your needs.

- Follow the Assignment Submission Guide to submit your team project when it is completed. The branch you merge to the **main** branch of your forked repository (**NOT** to the DSI repository) should be called `team-project-1`.
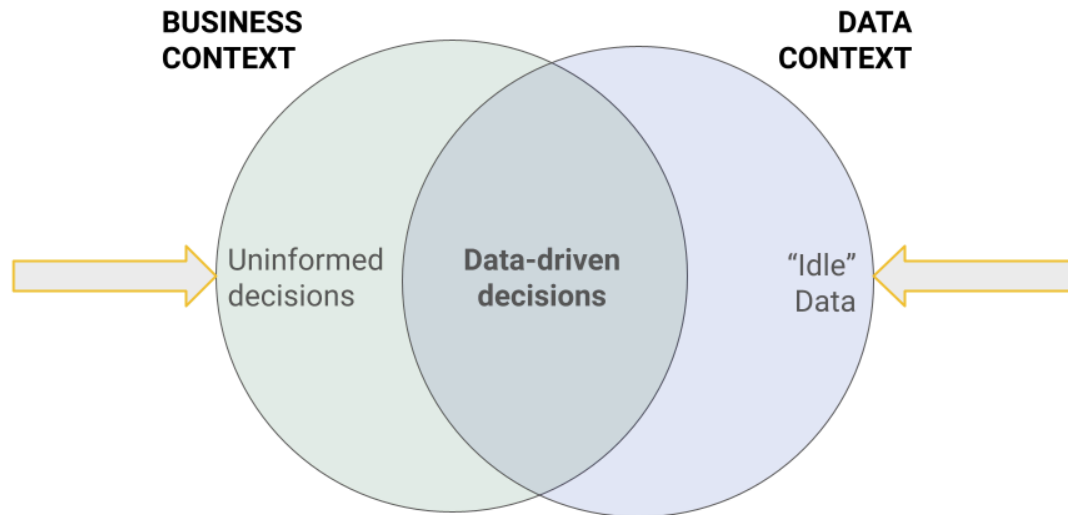
Refer to Git repository for more details.

Developing a project plan

How can we get the most value out of our data, and use it to better inform business decisions?

- We want these contexts to overlap!

The Five Orders of Ignorance

A "meta-model of ignorance" that can help us put the unknowns of projects into context.

0. Zeroth Order Ignorance (0OI): Lack of Ignorance

1. First Order Ignorance (1OI): Lack of Knowledge

2. Second Order Ignorance (2OI): Lack of Awareness

3. Third Order Ignorance (3OI): Lack of Process

4. Fourth Order Ignorance (4OI): Meta Ignorance

The Five Orders of Ignorance

0. Zeroth Order Ignorance (0OI): Lack of Ignorance

- You provably know something

- The ability to prove it is necessary

I know how many departments my company has.

The Five Orders of Ignorance

1. First Order Ignorance (1OI): Lack of Knowledge

- You know that you do not know something

- "Acknowledged ignorance"

I know that I don't know how many people work for my company.

The Five Orders of Ignorance

Second Order Ignorance (2OI): Lack of Awareness

- You do not know that you do not know something

- Metacognition: awareness of what you know and don't know

- This is what we try to eliminate as we work through projects

I don't know that I don't know that my company is about to acquire another company.

The Five Orders of Ignorance

Third Order Ignorance (3OI): Lack of Process

- You do not know of a suitably effective way to find out that you don't know that you don't know something

- This may result in a knowledge trap

- Finding ways to escape from this quickly is critical

- Escaping from this is an important function of software testing: showing areas where you lack knowledge

The Five Orders of Ignorance

Fourth Order Ignorance (4OI): Meta Ignorance

- You do not know about the Five Orders of Ignorance

- Fundamental ignorance of ignorance

The Five Orders of Ignorance

A "meta-model of ignorance" that can help us put the unknowns of projects into context.

- The goal is to move down the ladder to 0OI

- Most production energy should be spent on 2OI

- Most process energy should be spent on 3OI

1. Understand the business context

What are companies or organizations in your industry trying to achieve? What is your organization trying to achieve?

1. Understand the business context

What are companies or organizations in your industry trying to achieve? What is your organization trying to achieve?

Example: We work in finance. Our company is trying to reduce overhead costs.

2. Identify an opportunity

Where can you add business value by enabling better informed decisions?

2. Identify an opportunity

Where can you add business value by enabling better informed decisions?

Example: We notice that office space appears to be under-utilized. We want to explore the idea of reducing office space while maintaining the same level of productivity and comfort for employees.

3. Scope your analysis

What data do you need?

Do you have the resources to acquire additional data?

3. Scope your analysis

What data do you need?

Do you have the resources to acquire additional data?

Example: We want to explore the data on:

1. How often employees work from the office vs. work from home.
2. Whether teams tend to all work in the office on the same days.
3. The cost to acquire new office spaces.
4. The cost to modify existing office spaces.
5. Whether public transit conditions affect employees' choice of work location.
6. How our rate of working remotely compares with other companies and other industries.

4. Develop your solution

1. Integrate your data.

2. What are the key findings?

4. Develop your solution

1. Integrate your data.

2. What are the key findings?

Example:

We can compare the available office space to the space being used under a variety of conditions. This could include breaking the data down by building, floor, team, day of the week, month, and weather. We can investigate the capacitiy of different office layouts, an analyze the tradeoffs involved.

5. Present results and recommendations.

Follow-up studies?

Implementable changes (if applicable)?

5. Present results and recommendations.

Follow-up studies?

Implementable changes (if applicable)?

Example: Perhaps designated work-from-home days per team would help. Perhaps it would make more sense to have less overall office space but more efficient use.

Check: are your insights actionable?

Insights are far more impactful if they are *actionable*.

Check: how robust is your analysis?

How can you quantify the value that you are providing?

Check: what are the caveats or unknowns?

Are you missing data that would provide more insight?

Is there a risk of false correlation?

How do we monitor whether these trends continue?

Brief Git Review