Team Project

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

Review

- Last week, we focused on several key topics:
    i. How to present this project in your portfolio.
    ii. How to effectively develop a business case for your project.
    iii. How to work together as a team.
    iv. How to hold effective standup meetings.

Project Plan

- You have each received feedback on your Project Plan in the form of a GitHub issue. You should incorporate this feedback into your README this week.
- When submitting your final project, ensure that you have a section in the README for changes you have made to your original project plan.

Final Project Submission

Your final project will be evaluated on the following criteria:

1. Each team member must have created a pull request, and reviewed and merged a different pull request.
2. (a) For **Data Science** teams, your project must include at least one visualization that presents new insights into the chosen dataset.

   (b) For **Machine Learning** teams, your project must include at least one machine learning model that you have developed and implemented to obtain new insights.

Final Project Submission

3. In addition to the project proposal from Week 1, each project's README should describe the final outcome of the project, the key business takeaways, and describe your team's approach to working collaboratively. It should also demonstrate thoughtful consideration of the guiding questions.

Final Project Submission

4. Your project must include ***one completely reproducible feature***. This could be a single visualization, training and storing a model, or using a previously trained model to predict results, and the value of the feature should be described in your README. It could be formatted as a self-contained Jupyter notebook or a script, however anyone new to your project should be able to reproduce it based on the instructions provided, including any software, data, or file dependencies.

Final Project Submission

5. Each team member must record a 3-5 minute video reflecting on your experience. You may each choose where to host your own video, however *it should be public and a link to each team member's video should be included in your project README*. This video is meant to be an asset to your portfolio, and should be available for prospective employers. Your videos should answer the following questions:
    - What did you learn?
    - What challenges did you face?
    - How did you overcome those challenges?
    - If you had more time, what would you add?
    - What strengths do you bring to a team environment?

Keep in Mind

- **Good Code & Structure** → Code should be well-commented, clean, and follow a logical structure. It should be easy to read and maintain.
- **Strong Documentation & Presentation** → The README should be clear, well-written, and explain the dataset, findings, and methodology. The project should be easy to understand for both technical and non-technical reviewers.
- **Application of Module Teachings** → Projects should showcase key technical skills, such as regression modeling, deep learning models, data visualizations, or strong analysis of sampling techniques.
- **Effective Team Collaboration** → Teams should follow best practices for Git (small commits, branches, pull requests) and actively participate in stand-ups and progress updates.

Crafting a Comprehensive Main README File

- **Purpose & Overview:** Introduce the project with essential details, concise description and a project objective.
- **Goals & Objectives:** Articulate what the project aims to achieve. *Include any changes made to the original project plan here*.
- **Techniques & Technologies:** Highlight the tools and methods used.
- **Key Findings & Instructions:** Summarize outcomes and provide setup instructions.
- **Visuals & Credits:** Enhance with visuals; acknowledge contributors.

This Week's Schedule

Day 6 (Tues): Review + Co-work + Standup Meetings.

Day 7 (Wed): Co-work + Standup Meetings.

Day 8 (Thurs): Case Study + Co-work.

Day 9 (Fri): Co-work.

Day 10 (Sat): Project Showcase.

Standups

- Each day, a member of the DSI instructional team will guide your team through a standup.
- Stand-ups are **quick, structured check-ins** that help teams stay on track and remove obstacles.
- The goal is not to compete for who did the most work, it’s to ensure the entire team is working effectively and efficiently.
- This is a great opportunity to help your teammates and **resolve blockers early**.
- Standups should take no more than **10 minutes**.

Communicating & Presenting Your Findings

Business Impact

Examples Across Sectors (Part 1)

Retail Example

"***Optimized inventory management*** with SQL and Tableau. Data was cleaned and aggregated using SQL queries to identify sales patterns. Visualizations in Tableau helped to ***forecast demand, leading to cost savings of 5%.***"

⏭️ *Continued on next slide...*

Examples Across Sectors (Part 2)

Marketing Example

Increased customer engagement by 8% by performing customer segmentation using k-means clustering with Python. Feature engineering was conducted to create meaningful segments based on purchase history and demographics. The optimal number of clusters was determined using the elbow method, which helps identify the point where adding more clusters does not significantly improve the model.

How to Effectively Communicate Business Impact

Showcase your project's business value by applying specialized skills:

- 📊 **Data Science Stream**: Apply the techniques learned in your sampling and visualization modules to deliver impactful data insights.
- ⚙️ **Machine Learning Stream**: Demonstrate model efficiency and algorithm optimization with detailed performance graphs and clear pipeline visualizations.

Ensure your presentations translate complex data into clear, actionable business insights.

Making Complex Concepts Accessible

Communicating complex data science concepts clearly is essential for audiences without technical backgrounds, **including recruiters and HR teams.**

Strategies to Enhance Understandability:

- **Use Analogies and Metaphors**: Bridge understanding gaps with relatable comparisons.
- **Incremental Explanation**: Simplify concepts step-by-step.
- **Visual Aids**: Employ diagrams and infographics for clarity.

Clear communication ensures your projects are comprehensible to all, enhancing your professional appeal and broadening project impact.

Example of Making Complex Concepts Accessible

Remember the "Healthcare Example" above?

Before:

```
"Improved patient outcomes by developing a predictive model using Python and scikit-learn. Data was preprocessed using pandas to handle missing values and standardize features. Logistic regression was selected through cross-validation due to its interpretability and performance."
```

After:

```
“Improved patient outcomes by developing a predictive model using Python and scikit-learn. Think of the predictive model as a way to forecast patient health based on past data, much like predicting the weather. Data was preprocessed using pandas to handle missing values and standardize features. This preprocessing is like organizing and cleaning up messy medical records to ensure all information is complete and consistent. Logistic regression was selected through cross-validation due to its interpretability and performance. Imagine trying  different medical treatments and choosing the one that consistently gives the best results; that’s what cross-validation does for model selection.”
```

This approach effectively communicates complex concepts to both technical and non-technical audiences, ensuring clarity and comprehension across diverse stakeholders.

Final Presentation & Feedback Session

- This day is for showcasing your progress on the project goals so far. It's not about meeting all the requirements yet, but about presenting what your project is about, what you've done, and what you plan to do next. It's also a chance to get feedback and ideas from your fellow participants.
- Each team has a maximum of 5 minutes to present. Put your best foot forward, as you may never get such a diverse audience of this size again. The audience is comprised of your fellow participants, who now have a good background in data science and machine learning, and have extensive expertise in their respective industries. They will serve as judges and provide valuable feedback.

Next Steps Post-Certification

As you transition from the certificate, continue to build and showcase your skills.

Strategies for Continuous Growth:

- **Active Portfolio**: Aim for 5-6 substantial projects to demonstrate a broad range of skills. Focus on quality over quantity.
- **Explore and Learn**: Continue exploring new areas within data science and machine learning to keep your knowledge current and dynamic.
- **Peer Collaboration**: Participate in group projects with fellow participants to enhance your skills and widen your professional network.

Actively applying for jobs should proceed in tandem with these activities.

Questions?