Learning Activities for a Data Science Course

1. Observation → Data Exploration Walkthrough

- **Activity:** Provide students with a raw, messy dataset (e.g., customer purchases, healthcare visits, or social media posts).
- **Task:** Carefully observe patterns, inconsistencies, and anomalies without jumping to conclusions.
- **Learning Outcome:** Students practice exploratory data analysis (EDA) — descriptive statistics, plots, missing values.
- **Design Link:** You learn to see before solving.

2. Critical Thinking → Hypothesis Testing Challenge

- **Activity:** Present a problem statement such as: *"Do students who study more hours actually score higher in exams?"* with a dataset.
- **Task:** Students must frame null and alternative hypotheses, choose the right statistical test.
- **Learning Outcome:** They evaluate claims logically, rather than assuming correlations.
- **Design Link: Critically think and justify or reject the**

---

3. Collaboration → Group Data Project with GitHub Repository

- **Objective:** To develop students' ability to work collaboratively on data science projects using version control, shared repositories, and structured workflows on Github

- **Task:**

  - Students form small groups and create a shared GitHub repository.
  - Assign roles (e.g., Data Cleaning, EDA, Modeling, Visualization, Documentation).
  - Each member works on their own branch, commits changes with proper messages, and submits pull requests.
  - Teams review and merge contributions into the main branch.
  - Final submission includes a well-documented repository with a clear **README**, organized code, and results.

- **Learning Outcome:**

  - Students gain practical skills in collaborative coding and version control.
  - They learn structured teamwork, code review practices, and project documentation.
  - They understand how diverse contributions can be integrated into one cohesive project, mirroring real-world data science teamwork.

- **Design Link:**

- Reflects **collaboration** in design thinking, where multiple perspectives and roles come together to co-create solutions.
- GitHub acts as the "workspace" for team brainstorming, prototyping, and iteration — just like in design projects.
- Emphasizes that innovation in data science is not individual but *collectively built and refined*.

- **Resources Needed:**
    - GitHub accounts for all students
    - A dataset (e.g., Kaggle dataset on customer churn, COVID-19, or e-commerce transactions)
    - Tools: Python (Jupyter Notebook), Git/GitHub Desktop, Google Colab (optional)
- **Steps:**

1. **Repository Setup:** Instructor creates a central repository and shares it with groups OR each group creates their own repo.
2. **Role Assignment:** Within groups, students take roles such as Data Cleaner, EDA Specialist, Model Builder, Visualizer, Documentation Lead.
3. **Branching:** Each member works on a separate branch (e.g., `eda_branch`, `model_branch`).
4. **Commit & Pull Requests:** Students commit changes regularly with proper commit messages and submit pull requests.
5. **Code Review & Merge:** Team members review code changes before merging into the main branch.
6. **Final Integration:** Groups finalize a project (data cleaning, analysis, visualization, model) and prepare a README file with instructions, results, and insights.

- **Expected Learning Outcomes:**
    - Understand collaborative coding and version control.
    - Learn how to integrate diverse contributions into a single project.
    - Experience real-world workflow of a data science team.
    - Practice clear documentation and communication.
- **Assessment:**
    - Evaluation based on repository activity (commits, branches, pull requests).
    - Quality of collaborative work (clear README, organized folder structure).
    - Peer feedback on collaboration experience.

4. Creativity → Data Visualization Redesign

- **Activity:** Provide a poorly designed graph or dashboard.
- **Task:** Students must redesign it using principles of clarity, aesthetics, and meaningful storytelling.
- **Learning Outcome:** They practice turning numbers into *creative, human-centered insights*.
- **Design Link:** Just like prototyping in design, they learn *creative representation of solutions*.

Divergent Thinking → Multiple Model Approaches

- **Activity:** Pose a predictive problem (e.g., *predict whether a customer will buy a product*).
- **Task:** Ask students to try at least **three different modeling approaches** (logistic regression, decision tree, k-NN).
- **Learning Outcome:** They explore multiple possibilities before narrowing to the best one.
- **Design Link:** Analogous to brainstorming many design ideas before converging on one prototype.
- **Observation:** Students explore churn dataset (EDA).
- **Critical Thinking:** Form hypotheses about factors affecting churn.
- **Collaboration:** Groups share different analyses and insights.
- **Creativity:** Each group builds a dashboard/storyboard to present findings.
- **Divergent Thinking:** Groups test multiple machine learning models before recommending the best solution.

5. Empathy: Present Failed results and reasoning behind it.