## Sep 5, 2024 | [Milestone 1 project intro](https://www.google.com/calendar/event?eid=NHRldG9uaXA5OWtjOHYzdjdua3JrMmMzYjAga2FzcmFAdW1pY2guZWR1)

Attendees: [Kasra Afzali](mailto:kasra@umich.edu) [Michael Light](mailto:mikalite@umich.edu) [Iris Lin](mailto:irisqlin@umich.edu)

Meeting Minutes: Group Project Kickoff

**Dataset Selection**:

* The group agreed to use **publicly available datasets** from the CDC and government sources.
* The focus will be on **Medicaid/Medicare coverage**, exploring various metrics beyond financials, including coverage, inclusivity, efficiency, risk, etc
* **Avoid Overly Specialized Medical Details:** The group emphasized that the project should aim to uncover **meaningful insights** that do not require deep field expertise, making it accessible for discussion within an **interdisciplinary team**. The focus will be on metrics that provide **high-level understanding**, such as Incidence rates by age group, treatment costs, and coverage percentages, which are relevant to decision-makers in healthcare insurance or government agencies.
* The emphasis should be on demonstrating a **strong approach and technical skills**, rather than diving into overly specialized medical details. The **context** should act as an enabler for showcasing analytical abilities, ensuring that the project remains engaging yet appropriately scoped.

#### **Dataset Feasibility:**

* The group emphasized balancing **complexity and feasibility**:
  + **Linking datasets**: Can the two datasets be converged on common features like time, region, or class?
  + **Level of Detail**: Are the datasets at comparable levels of granularity? If one dataset is detailed (e.g., medical encounters) and another is aggregate (e.g., patient counts per state), the connection must be feasible without losing key insights.
  + **Next Steps**: Identify potential obstacles early, such as missing common fields or mismatched detail levels between datasets.
  + **Long-Term Opportunities**: Is there room to expand the scope of the analysis after the initial insights?

#### **Tools Discussion:**

* The group will use **Python** and common data science libraries, including:
  + **pandas**, **numpy**, **scikit-learn**, **plotly**, **seaborn**, etc., for data manipulation, analysis, and visualization.
  + Use task-specific libraries and tools as needed and/or per discretion of the project mentor

#### **Communication:**

* **Meeting Schedule**: Weekly meetings were scheduled, preferably **early in the week**, with follow-up meetings as needed.
* **Google Calendar**: The group will create a shared calendar for tracking meetings and deadlines.
* **Contact Information**: Group members exchanged emails and phone numbers for efficient communication.

#### **Next Steps:**

* Regroup on **Saturday at 12 PM EST / 9 AM PST**
* **Idea Generation**: Each team member will introduce two project ideas, focusing on how to analyze Medicaid/Medicare data meaningfully.
* **Proposal Preparation**:
  + Fill out the proposal form with bullet points for each idea.
  + Discuss the **challenges and opportunities** of each idea.
  + Commit to one idea and split up the proposal sections among the group.
  + After selecting a final project idea, each team member will be assigned a section to expand upon.
* **Deliverables**:
  + **Two Individual Proposal drafts: 09/07**
  + **Proposal Submission Deadline**: **09/09.**
  + Individual peer feedback x2