Team Members:
Lizbeth, Kaitly, Patricia, Suraya - Cohort 2

## Milestone 2

○ What tools will you use?
- **Frontend Development**: HTML, CSS, Javascript
- **ML**: Jupyter Notebook (Python), Libraries and Frameworks(NumPy, Pandas, Scikit-learn, etc), Github
- **Data visualization**: Plotly, Matplotlib

○ What do you need to learn or get more familiar with in order to execute your project?
- We are going to need to familiarize ourselves with frontend/backend development
- Research more about frameworks such as Scikit-learn and TensorFlow
- Research different machine learning models/algorithms
- Research and identify best evaluation metric (accuracy, precision, recall)
- Research techniques for cleaning and preparing the data
- Learn about handling missing values, outliers, and normalization
- Better understand feature engineering
- Learn about overfitting/underfitting
- Research best data visualization methods for our project

○ How do you plan on learning what's needed in order to execute your project?
- Research needed topics through google, youtube, etc.
- Revisit old lectures/projects to refresh memory
- Get input/feedback from teammates and others
- Look up documentation for libraries and frameworks

## Possible Datasets:
Here are some of the datasets we are considering on using:
1. National Health and Nutrition Examination Survey (NHANES)
   - Source: Centers for Disease Control and Prevention (CDC)
   - Details: NHANES provides extensive data on health and nutritional status, including diabetes prevalence, with breakdowns by ethnicity, including Latino/Hispanic populations.
   - Link: [NHANES](https://www.cdc.gov/nchs/nhanes/index.htm)
2. Hispanic Community Health Study / Study of Latinos (HCHS/SOL)--
   - Source: National Institutes of Health (NIH)
   - Details: A comprehensive study specifically focused on health issues, including Type 2 diabetes, within the Latino/Hispanic population in the U.S.
   - Link: [HCHS/SOL](https://sites.cscc.unc.edu/hchs/)
3. Behavioral Risk Factor Surveillance System (BRFSS)--
   - Source: CDC
   - Details: The BRFSS collects data on health-related risk behaviors, including diabetes rates, and provides breakdowns by ethnicity, including Hispanics/Latinos.
   - Link: [BRFSS](https://www.cdc.gov/brfss/)

4. American Diabetes Association (ADA) - Statistics & Data--
   - Source: ADA
   - Details: ADA compiles research and statistics about diabetes, including ethnic-specific prevalence for Hispanic/Latino populations.
   - Link: [ADA Diabetes Statistics](https://www.diabetes.org/resources/statistics/statistics-about-diabetes)
5. National Institute of Diabetes and Digestive and Kidney Diseases (NIDDK) - Data & Statistics--
   - Source: NIH/NIDDK
   - Details: Provides data related to Type 2 diabetes prevalence, with demographic data including ethnicity (Latino/Hispanic populations).
   - Link: [NIDDK](https://www.niddk.nih.gov/health-information/health-statistics/diabetes-statistics)

Diabetes and Hispanic Americans External link from the HHS Office of Minority Health provides prevalence data on diabetes among Hispanic Americans, including death rates, risk factors, and links to more information.
https://blast.ncbi.nlm.nih.gov/Blast.cgi

**Example of website:**



Website example:

Operation: Diabetes Prevention

Home    Research    Graphs    Ways To Prevent

-Details about graph-

Example graph

-Details about graph-

Example graph

-Details about information above -