**Aim:** For Emily’s process, need to implement a function that takes an arbitrary table of data samples (x\_i,y\_i) and returns Mean, Confidence intervals, and a Plot of those data points, for each x\_i.

**Usability Overview:**

* Flask Application : Allows you to upload a CSV file, Processes the data, and generates a plot with mean and confidence intervals, a table of means and confidence intervals, and JSON data with the same information.
* It randomly takes table of sample from the the given data, and produce plots for those data sample.
* Web Interface: Provides a way to upload your CSV file and displays the results.

Fig1: Interface


Fig1:This is the Interface of the Web Application

From Fig1: As we can see, it lets us to choose a file from Local Computer and uploading it produces the results of that data with a plot.

**A table with numbers and symbols

Description automatically generated with medium confidence**

**Fig2: Results of Mean and CI**

From Fig2: As we can see it arbitrarily selected sample from the given data.

**1.) Steps to Use :**

* We can either use Command Prompt or terminal from the Docker Desktop application.
* First set the directory to flask-app.
* **‘docker build -t flask-app’** Run this line to build the Docker, it will download all the essentials to run the code.
* **‘docker run -p 5001:5001 flask-app’** .This line runs the entire code.
* Navigate to ‘http://localhost:5001’ to access the web interface.
* Upload the ‘sample.csv’ using ‘Choose file’ option .
* The application will generate and display the plot, table, and JSON results.

**2.) Reproducing the Demonstration Case:**

* Create a ‘sample.csv’ file with multiple Y\_i columns corresponding to the X\_i column.
* Save ‘sample.csv’ in the ‘static’ directory of the project.
* Run the web application and upload the ‘sample.csv’

**3.) Tools Used for Development:**

* Python: Programming language used for the application.
* Flask: Web framework for building the application
* Pandas: Data manipulation library
* NumPy: Numerical operations library
* SciPy: For Statistical Functions.
* Matplotlib: Plotting library for generating plots.
* Docker: For containerizing the application and ensuring consistent development environments.

**4.) Improvements:**

* I would deal with the missing data in a better way instead of filling it with zero. I would have used KNN to predict the missing data or some other data cleaning process.
* Would develop a more user-friendly front end using HTML, CSS to make the interface more visually appealing.
* Would Integrate interactive plotting libraries like Plotly to allow users to interact with the data plots directly.
* Would Optimize the code for handling larger datasets efficiently.