Take-Home Exercise for Software Engineer

Create an application to process and analyze data generated by laboratory instruments. The application must:

1. Parse output files from various instruments (different parsing logic required).
   - Automatically detect the appropriate parser based on file content.
   - Support both CSV and Excel file formats.
2. Validate data based on provided business logic.
3. Store valid results in an SQLite database.
4. Provide a web interface for user interaction (is OK to skip CSS styling).

Objectives

The solution must:

1. Be implemented in Python.
2. Meet all the outlined requirements.
3. Run without issues (provide clear setup and usage instructions).
4. Follow best engineering practices (e.g., DRY, clear interfaces, modularity, readability, testability, and maintainability).

The application should support the following workflows:

Uploading New Results

1. Upload a results file (use provided sample files in the `samples` folder, which contain valid and invalid files for different experiment types).
2. Validate the data:
   - If errors are found, display them to the user.
   - If no errors are found:
     - Store results in the database.
     - Display the newly uploaded results.

Exploring Stored Results

1. Access an index page showing all stored results, grouped by experiment type.
2. Select an experiment type to view:
   - Overall statistics: median, average, and standard deviation.
   - A table of results with columns: `formulation_id` and `calculated_value`.

Implementation Hints

- Use uv for project setup and dependency management.
- Use a DataFrame library for data manipulation (e.g., Polars).
- Keep the web interface simple: HTML templates, Vanilla JS, and a basic Python web server. Alternatively, use a web framework of your choice.
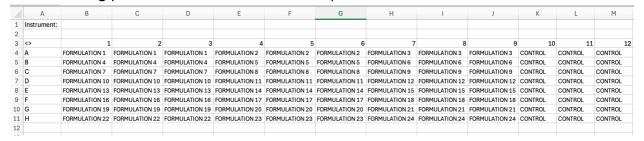
Business Logic for Parsing and Calculating Results

The results are derived from instrument readings measuring the physicochemical properties of formulations. The two experiment types and their calculations are as follows:

Experiment Type: Zeta Potential

- Data: Excel file with readings from a standard 96-well plate.
- Each formulation has 3 replicate readings, and each row includes control readings for normalization.
- Calculate results:
  - Average the triplicate readings for each formulation.
  - Normalize by the average control value of the same row.
- Example:
  - Formulation 1: `Average(A1–A3) / Average(A9–A12)`
  - Formulation 2: `Average(A4–A6) / Average(A9–A12)`
  - Formulation 4: `Average(B1–B3) / Average(B9–B12)`
- A result is valid if it exceeds **10**.

The following picture shows the formulation placement:

	A	B	C	D	E	F	G	H	I	J	K	L	M
1	Instrument:												
2													
3	<>	1	2	3	4	5	6	7	8	9	10	11	12
4	A	FORMULATION 1	FORMULATION 1	FORMULATION 1	FORMULATION 2	FORMULATION 2	FORMULATION 2	FORMULATION 3	FORMULATION 3	FORMULATION 3	CONTROL	CONTROL	CONTROL
5	B	FORMULATION 4	FORMULATION 4	FORMULATION 4	FORMULATION 5	FORMULATION 5	FORMULATION 5	FORMULATION 6	FORMULATION 6	FORMULATION 6	CONTROL	CONTROL	CONTROL
6	C	FORMULATION 7	FORMULATION 7	FORMULATION 7	FORMULATION 8	FORMULATION 8	FORMULATION 8	FORMULATION 9	FORMULATION 9	FORMULATION 9	CONTROL	CONTROL	CONTROL
7	D	FORMULATION 10	FORMULATION 10	FORMULATION 10	FORMULATION 11	FORMULATION 11	FORMULATION 11	FORMULATION 12	FORMULATION 12	FORMULATION 12	CONTROL	CONTROL	CONTROL
8	E	FORMULATION 13	FORMULATION 13	FORMULATION 13	FORMULATION 14	FORMULATION 14	FORMULATION 14	FORMULATION 15	FORMULATION 15	FORMULATION 15	CONTROL	CONTROL	CONTROL
9	F	FORMULATION 16	FORMULATION 16	FORMULATION 16	FORMULATION 17	FORMULATION 17	FORMULATION 17	FORMULATION 18	FORMULATION 18	FORMULATION 18	CONTROL	CONTROL	CONTROL
10	G	FORMULATION 19	FORMULATION 19	FORMULATION 19	FORMULATION 20	FORMULATION 20	FORMULATION 20	FORMULATION 21	FORMULATION 21	FORMULATION 21	CONTROL	CONTROL	CONTROL
11	H	FORMULATION 22	FORMULATION 22	FORMULATION 22	FORMULATION 23	FORMULATION 23	FORMULATION 23	FORMULATION 24	FORMULATION 24	FORMULATION 24	CONTROL	CONTROL	CONTROL
12													

Experiment Type: TNS

- Data: CSV file with triplicate readings for each formulation.

- First 3 rows (labeled "STD") are control readings.

- Calculate results:

  - Compute the average of the control readings.
  - For each formulation, take the average of the triplicate and normalize by the control average.

- A result is valid if it is **positive**.