

## Project Milestone #3

Name: Kanishk Ghai

Student Id: 100815261

Github: <https://github.com/Kan-G1/SOFE4630U-MS3>

### Smart Meter Dataflow Pipeline Design

The Smart Meter pipeline is designed to process real-time sensor data using Google Cloud Dataflow and Pub/Sub. The pipeline performs the following steps:

#### 1. Ingestion:

- Reads streaming data from the input Pub/Sub topic `projects/<PROJECT_ID>/topics/smartmeter_input`.
- The data consists of JSON messages containing sensor readings (ID, pressure in kPa, temperature in Celsius).

#### 2. Transformation:

- **Filtering:** A filter step removes any records that contain missing or None values to ensure data quality.
- **Conversion:** A processing step converts the measurement units:
  - Pressure: Converted from Kilopascals (kPa) to Pounds per Square Inch (psi) using the formula  $P(\text{psi}) = P(\text{kPa}) / 6.895$ .
  - Temperature: Converted from Celsius (°C) to Fahrenheit (°F) using the formula  $T(\text{F}) = T(\text{C}) * 1.8 + 32$ .

#### 3. Output:

- The transformed data is serialized back into bytes.
- It is published to the output Pub/Sub topic `projects/<PROJECT_ID>/topics/smartmeter_output`.