

# Live Coding Assignment - Software Engineer - Data

## Objective

Create a basic REST API endpoint that retrieves the average CPU usage for a specified EC2 instance ARN over a given timeframe. The data is provided in two Parquet files.

## Instructions

### 1. Data Ingestion

- Use the following Parquet files loaded from a third party source (provided in the coding environment):
  - `metrics.parquet` : Contains time-series data of EC2 instance metrics.
  - `resources.parquet` : Contains metadata for EC2 instances.

### 2. API Endpoint

- Implement a single API endpoint with the following requirements:
  - Accepts query parameters:
    - `arn` : EC2 instance identifier (required)
    - `start` : Start timestamp of the timeframe (optional)
    - `end` : End timestamp of the timeframe (optional)
  - Responds with the **average CPU usage** within the specified timeframe for the given ARN.
- If no `start` or `end` is provided, calculate the average over all available records for that ARN.
- If no `arn` is provided, calculate the average over all available ARNs.

### 3. Example Usage

**Request:**

```
GET /api/v1/cpu-usage?arn=<EC2_ARN>&start=<start_timestamp>&end=<end_timestamp>
```

**Response:**

```
{  
  "average_cpu_utilization": <average_cpu_value>  
}
```

## Notes

- Use any programming language, framework and libraries you want.
- Focus on getting the basics to work. Optimize later.

Good luck! 🍀