

Algorithm:

1. Setup Environment:

- System Configuration:

Note down the system's CPU, memory, and other specifications.

- Resource Monitoring:

- Instal python
- Install vs code and required python extension.
- Install required libraries like numpy, pandas, pyarrow using following commands:

pip install pandas

pip install numpy

pip install pyarrow

2. Data Extraction:

- Read Parquet File: Use pyarrow and pandas for efficient Parquet file reading.

Use ***read_parquet()***

- Extract Required Data:

- Total Cost: Sum of all line item costs. (using df[].sum())
- Tax Cost: Sum of all line items where the usage type corresponds to tax.
- Usage Cost: Total cost minus tax cost.

Store all the above calculations in the variables and print as required.

3. Resource Efficiency:

- Memory and CPU Usage: Focus on using efficient data structures and minimizing in-memory data storage.
- This step is completed using the os library which provides the required configuration output.

4. Output:

- Execution Time: Log the time taken for data extraction and compare with previous results.

5. Optimization:

- Review the algorithm to find potential optimizations in reading and processing data.