Documentation of Python Code and Query(TASK-2)

1. Introduction

The provided Python program performs data processing and analysis related to voyage events, including calculating durations between events, filtering based on event types, computing distances between geographical coordinates, and visualizing event timelines with durations.

2. Sample Data

The program utilizes sample voyage data containing events (`SOSP` - Start of Sea Passage, `EOSP` - End of Sea Passage), timestamps, geographical coordinates (latitude and longitude), and other relevant identifiers.

3. Conversion to DataFrame

The sample data is converted into a pandas DataFrame (`voyages_df`) to facilitate structured data manipulation and analysis.

4. Conversion of Timestamps to UTC Datetime

A function is implemented to convert the combination of dateStamp and timeStamp columns into UTC datetime objects. This conversion is essential for accurate time-based calculations and plotting.

5. Calculation of Duration Between Events

Another function computes the duration in seconds between successive events based on their UTC datetime values. This calculation uses the `total_seconds()` method to quantify the time difference.

6. Sequencing of Events

To facilitate sequential analysis, the program creates additional columns (`next_event` and `next_utc_datetime`) to identify the subsequent event and its corresponding UTC datetime. This sequencing aids in filtering and calculating event-specific durations.

7. Filtering and Calculation of Sailing and Port Stay Times

Using the sequenced events, the program filters the DataFrame to calculate the total sailing time (`sailing_time_seconds`) and total port stay time (`port_stay_time_seconds`). These calculations are based on specific conditions (`SOSP` followed by `EOSP` and vice versa) to derive meaningful insights into voyage durations.

8. Calculation of Distance Between Geographic Coordinates

A separate function implements the Haversine formula to compute the great-circle distance between pairs of latitude and longitude coordinates. This distance calculation is crucial for understanding geographical movements and distances covered during voyages.

9. Example Usage of Distance Calculation

An illustrative example demonstrates how the distance function is applied to compute the distance between two predefined geographical points (e.g., Port A and Port B). Error handling ensures robustness in case of invalid inputs.

10. Visualization: Plotting Timeline of Events and Durations

The program employs matplotlib to visualize the timeline of voyage events and their corresponding durations. Events are plotted as points along the UTC datetime axis, while durations are represented as horizontal bars. This visualization aids in understanding the temporal distribution and duration of voyage activities.