Amrita Vishwa Vidyapeetham

Department of Artificial Intelliegence

21AIE304 Big Data and DataBase Management Systems
Practice Problems -4

Date: 15-11-2023

> **_Instructions_**
> a. Please compile the practice problems into a Word document for PP1 and save it in your OneDrive folder.
> b. Create a section titled 'Practice Problem 4.' For each question, provide the associated code in text format immediately below the question.
> c. Kindly refrain from including any screenshots.

1. Use sales_data.csv for the following questions

   Get the most popular day/s of the week

   Get the most popular day of the month

   Get the cummulative sales of each customer

   Get the the total quantity sold for each product and rank them

   Get the purchase frequency for each customer and rank them

   Get the rolling average of sales quantity for each product

   Get the running total for each customer's purchase amount

   Estimate the age of each customer based on their first purchase date.

   Identify the gap in days between the availability of a product and its purchase.

   Find customers who made a purchase in the last 3 months.

   Calculate the number of months since each customer's first purchase.

   Detect seasonal patterns in customer purchases, such as spikes during holidays or specific seasons.

   Identify High-Value Customers with Weekly Ranks: Rank customers based on their weekly total sales within each customer partition. This can help you identify high-value customers on a weekly basis.

2. Use weather_data.csv to achieve following

Convert the timezone to american timezone

Get to know the current timezone and convert the timestamp column to standard time (UTC timezone)