

DATA 225

Database Systems for Analytics

Fall 2023
Section 21
Instructor: Ron Mak

Assignment #10

Assigned: Monday, November 13
Due: Monday, November 20 at 5:30 pm
Team assignment, 125 points max

OLAP and Rollup operations

The purpose of this assignment is to give your team experience writing SQL queries that perform OLAP and ROLLUP operations on your dimensional model's star schema. (You can modify your star schema from Assignment #9 to accommodate this assignment.)

OLAP

For each of the following operations, use sample data in your dimensional model and write and execute two SQL queries. The two queries should use different dimensions from each other.

- drill up
- drill down
- slice
- dice

For each OLAP operation, display “before” and “after” query output in a dataframe.

Example: For drill up and drill down, choose one level of aggregation, such as quarterly, as your “before” query. Then drill down and show monthly results, and drill up and show yearly results. If the granularity of your fact table is daily, then monthly, quarterly, and yearly results are different levels of aggregation.

You can use artificially generated data in your dimension and fact tables.

GROUP BY . . . WITH ROLLUP

Using at least one of your dimensions, perform a query that uses a **GROUP BY . . . WITH ROLLUP** operation. Display the result in a dataframe. You can simply display the dataframe with the **None** cells.

What to submit

Your Python notebook containing the queries. Use your project team's databases on the class's shared remote database server.

Rubric

Criteria	Max points
Two queries involving different dimensions for each: <ul style="list-style-type: none">• drill up• drill down• slice• dice	100 <ul style="list-style-type: none">• 25• 25• 25• 25
ROLLUP <ul style="list-style-type: none">• One SQL query using <code>GROUP BY ... WITH ROLLUP</code>	25 <ul style="list-style-type: none">• 25