CSE 817: Data Engineering

Name: _____

CT 2 Date: 06/03/2023

ID: _____

This exam contains 2 pages and 4 questions. Total of points is 20.

1. (5 points) Define **crosstab** and **roll-up** operatiosn in OLAP and how can they be used to analyze data across different dimensions?

....
....
....
....
....
....
....
....
....
....
....
....
....
....
....
....
....
....
....

2. (5 points) Using the given Sales schema, cross tabulate the total sales amount by Product Category and Region. Use standard SQL for cross tabulation. *Sales(SaleID, SaleDate, ProductID, ProductCategory, RegionID, SalesAmount)*

....
....
....
....
....
....
....

3. (5 points) What are the key considerations for designing a scalable and efficient data lake architecture to support big data processing and analytics? Introduce in one or two lines the different components of a typical data lake architecture.

....
....
....
....
....
....
....
....
....
....
....
....
....
....
....
....
....
....
....
....

4. (5 points) Explain an use case of date lakes.

....
....
....
....
....
....
....
....