**CPRE/SE 414: Software Tools for Big Data Analytics**

**Fall 2024**

**Project 1**

**Due: Friday, Oct. 11 (at 11:59PM)**

**Nguyen Do, Andrew McMahon**

* As much as Date is the lowest temporal granularity, one would like to reason about the usual hierarchical structure (month, quarter, year) in both calendar and fiscal terms. Regardless of which aspect is being analyzed, if some naming is needed, both English and Spanish versions should be available.

A screenshot of a computer

Description automatically generatedA screenshot of a computer

Description automatically generatedCreate the dimension Date with hierarchy date -> month -> quarter -> year for both calendar and fiscal year in English and Spanish:

* In addition to the facts about internet sales, analysts may be interested in facts related to sales quota which, in turn, are tied with the employees and respective dates. And both employees and sales will need to be correlated with sales territories.

Dim Fact Sale Quota should tie to employees and date, employee should be correlated with sale territories. Dim Fact Internet Sale should tie to employee and date also:

A computer screen shot of a computer

Description automatically generated

A screenshot of a computer

Description automatically generated

* Products should be available in both subcategories and categories, and among the relevant features is their weight.

Create Dim Product with weight and hierarchy for categories and subcategories:

A screenshot of a computer

Description automatically generated

* When it comes to customers, both spatial and temporal properties will be needed for investigating particular features, at all the available levels/granularities

Dim Customer that has hierarchy at different levels for spatial and temporal properties:

A screenshot of a computer

Description automatically generated

* For some reason, the customer also wants to be able to investigate between participating

organization and the currencies that they are using.

Create the Dim Organization relate to Currency:

A screenshot of a computer

Description automatically generated

* Some geo-analytics may need to be performed for resellers too.

Create Dim Reseller have geo-analytics:

A screenshot of a computer

Description automatically generated

From the Fact of Internet Sales: We care about quantity of orders, standard cost, freight, number of sales. From Fact about Sales quota, we only care about the amount.

A screenshot of a computer

Description automatically generated

A screenshot of a computer

Description automatically generatedSuccessfully deployed Cube:

A screenshot of a computer

Description automatically generated