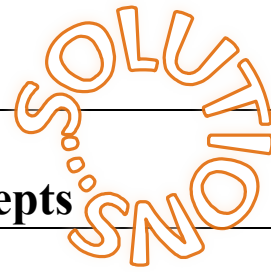


## Tutorial 2

### Star Schema Concepts



Discuss the following questions:

1. In the Student Enrolment case study, why is **Total Score** a fact measure? Is it to answer the question such as “what is the total score of the Java unit last semester?”

Total Score will be used to calculate the Average Score. Because if Average Score is used as the fact measure, there will be problems known as “Average of Average” problem. In order to avoid this problem, Total Score and Number of Students are kept as the fact measures.

2. Every crime (burglary, assault, theft, etc) is recorded in the **policy database**, including where and when it happened, as well as the victims. What are the possibly fact measures and dimensions?

Possible fact measure: Number of crimes

Possible dimensions: Suburb, Month, and Types of crimes (e.g. burglary, theft, assault, etc)

Victims will not be used as a dimension because victim is the highest granularity, whereas data warehousing focuses on lower granularity; Global view, rather than individual view.

3. In a typical **café** that sells coffee, and other kinds of hot drinks, as well as some cakes/muffins, what are the possibly fact measures and dimensions for a café data warehouse?

Possible fact measures: Total amount, Number of Items

Possible dimensions: Item Name (e.g. latte, cappuccino, long black, hot chocolate, muffin, etc), Day (e.g. Mon, Tue, Wed, etc)