

FIT3003 – Business Intelligence and Data Warehousing

Week 12 – Revision

Semester 2, 2022



Week 1

The Big Picture

Operational
Database



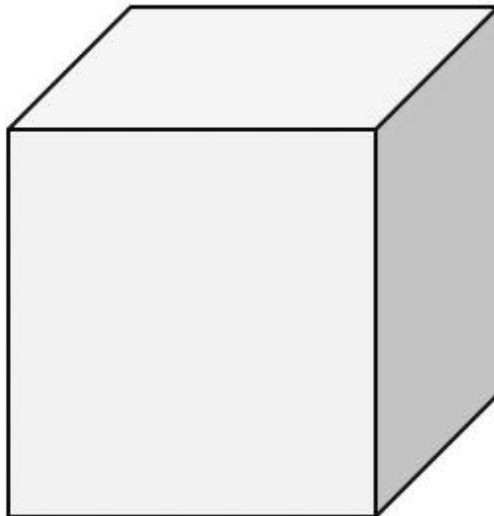
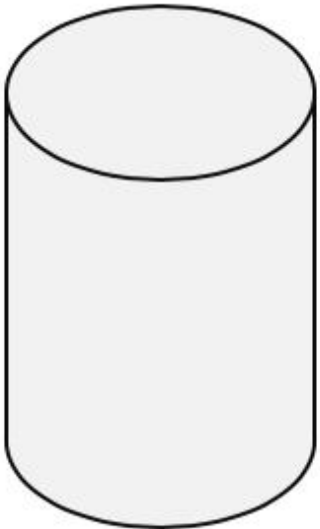
Data
Warehouse



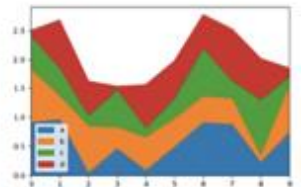
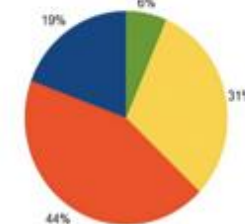
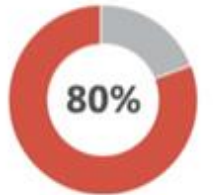
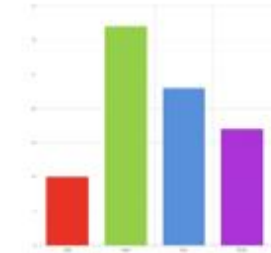
OLAP



Business
Intelligence



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



Basic Data Warehousing and SQL Revision

- Why do we need data warehousing? Why not the operational database?
- What is DW used for? What is the operational DB used for?
- What is ETL? What is transformation?
- Why do we need to explore the data in the operational database?
- Why do we need data cleaning?
- What is transaction?
- What is OLAP?
- What is the relationship between OLAP and BI? What is BI?

Basic Data Warehousing and SQL Revision

- SQL:
 - Create table, PK/FK
 - Insert into
 - Update, Delete
 - Select, Join
 - Group by, Count, Sum
 - Alter

Week 2

Basic Data Warehousing

- What is star schema? What is fact measure?
- How to identify fact measures? How to identify dimensions?
- Why must fact measures be numerical values?
- What is an aggregated value? Why must fact measure be an aggregated value?
- How to create a dimension? What is the purpose of using group by when creating the fact table?
- Where is the PK (FK)?

Simple Star Schema

- What is a two-column table methodology?
- What is category?
- How can we check whether a simple star schema is correct?
- Why do we need a tempfact? Is tempfact compulsory?
- Why do we need to create a dimension manually?
- Why must we update tempfact?

Simple Star Schema

- How to create a dimension? Why do we use Select Distinct in create dim?
- What are attributes in dimensions?
- How many different ways in creating a dimension?
- When creating a fact table, why don't we use the dimension tables?
- Why do we use the tables in the operational database when creating a fact table?

Week 3

Bridge Tables

- What is a bridge table? Why is a bridge table needed?
- What are the three versions of Product-Supplier bridge tables?
- What is weight factor and ListAGG? Where are these two attributes located?
- How to create a dimension with weight factor and ListAGG?
- Is ListAGG needed? When is weight factor needed?
- How to calculate the fact measure after averaging with weight factor?

More Complex Processes

- What is the problem of AVG in the fact?
- How to solve the AVG problem?

Week 4

Multi-Facts

- Can we have multiple fact measures? Do we need to separate into multi facts?
- What is subject-oriented in data warehousing?
- Can dimensions be shared?
- How to determine multi-fact?

Data Cleaning and Data Exploration

- What types of mistakes are there?
- How to explore data? How to find mistakes?
- How to correct mistakes? Should all mistakes be corrected?

Week 5

Data Warehousing Architecture

- What is data warehouse architecture?
- Why do we need multiple level of granularities of star schema?
- What is level of granularity?
- What is drill-down?
- Why is data warehouse architecture needed?
- How can we lower down the granularity of a star schema?
- Can we have multiple star schemas on the same level?
- Can we have more than three levels?
- Can we determine what level a star schema is?
- Why sometimes adding a dimension does not increase the level of granularity?

Week 6

Temporal Data Warehousing

- What is temporal data warehousing?
- Why is a temporal data warehousing needed?
- How to implement a temporal star schema?
- How to calculate the fact measure correctly?

- Temporal attributes vs. temporal dimensions
- What is SCD?
- What are SCD types?

Week 7

Snowflake Schema

- Why use two separate dim? Why not combine into one dim? Is it for drilling down?
- What is hierarchy in dimension? Why use hierarchy?
- What is normalization in dimension?
- Why not have a combined dimension? What is the difference between hierarchy and non-hierarchy in dimension modelling?

Determinant Dimensions

- What is a determinant dimension?
- When to use a determinant dimension?
- How to identify whether a dimension is a determinant dimension?

Week 8

OLAP

- CUBE, ROLL UP
- Grouping, Decode
- Partial Cube, Partial Roll Up
- Rank, Dense Rank, Percent Rank, Row Number
- Cumulative and Moving Aggregate
- Partition

BI Reporting

- What is the purpose of having Business Intelligence (BI) reporting tool?
- What is the connection between OLAP and BI?
- How can BI assist decision-makers in understanding the data from data warehouse?
- How to plot data from an OLAP query to become an informative BI report?

Week 9

Business Intelligence

- What is Business Intelligence (BI) Systems?
- Types of dashboards
- Dashboards versus Reports
- BI Navigation Methods
- Trends in BI Systems
- Self-Service BI vs. Cloud BI

Week 10

Business Intelligence – Power BI

- Connect and shape data in Power BI
- Create a data model in Power BI
- Visualize data and create a dashboard

Week 11 & 12

Active Data Warehousing

- What is active data warehouse?
- Active Data Warehousing vs. Passive Data Warehousing
- Why is Active Data Warehousing important?
- Active DW Architecture
- The challenges faced in implementing Active DW
- Active DW vs. real-time DW

Exam

Exam

- Time: 2 hours 10 minutes
- Closed book, online (through eAssessment platform)
 - For more information on eExam:
<https://www.monash.edu/students/admin/exams/electronic-exams>
- Format:
 - 7 case studies
 - Each question worth 10 marks – total 70 marks

Exam Preparation

- Sample Exams
- Pre-Exam Consultations
- Online Lectures
- Tutorial Practice
- Assessments
- Mock Exam

SETU

- Student Evaluation of Teaching and Unit (SETU)
- Your feedback is very important for us to improve this unit

**THANK YOU
and
GOOD LUCK!**