



The University of Fiji

Undergraduate Programme
Department of Computer Science and Mathematics
Semester 1 Year 2019

FINAL EXAMINATION PAPER

ITC251 DATA WAREHOUSING

Important Instructions to Students:

1. Read the questions, instructions carefully and answer the questions.
2. You are allowed an **extra 10 minutes reading time** (during this time, you are not allowed to write with pen) and **180 minutes (3 Hours) to complete this paper.**
3. Silent, non-programmable calculators are allowed, but not supplied.
4. Write all your answers in the answer booklet provided. If you need extra sheets, please ask from the supervisor. Write your name, course code and student identification number in the answer booklet and in the attendance sheet.
5. **This Final Course Examination Paper is worth 100 marks contributes 50% of the total course assessments.**
6. This paper contains one (1) section, A. All questions are compulsory to answer.
7. Write all your answers legibly. The marker has the right to award zero for answers that are not legible.
8. This is a **Closed Book Examination**. The University views plagiarism, cheating and dishonest practice as serious offences and if found, offenders are penalized.

Examination Paper Content (Format):

<u>Section</u>	<u>Question</u>	<u>Content</u>	<u>Marks</u>	<u>Suggested Time</u>
A	1-10	Short Answer Questions	100	180 minutes
		TOTAL	100	180 Minutes

SECTION A: SHORT ANSWER QUESTIONS**[100 Marks]****Question 1: Data Warehouse Concepts and Architectures****[12 Marks]**

- A. List and explain the **three (3)** technology and deployment limitations of traditional Database Management Systems (DBMS).

[3 Marks]

- B. List and explain **four (4)** essential characteristics of a data warehouse with the help of an example.

[4 Marks]

- C. Complete the table below contrasting the characteristics of an operational database and a data warehouse. *(Draw the table in the answer booklet)*

Characteristic	Operational Database	Data Warehouse
Details level		
Orientation		
Records per request		
Normalization level		
Update level		
Data model		

[3 Marks]

- D. List and explain **two (2)** challenges faced in data warehouse projects

[2 Marks]**Question 2: Data Cube Concepts****[11 Marks]**

- A. List and explain the **three (3)** aggregate measure used in a data warehouse with the help of an example.

[3 Marks]

- B. Complete the following table comparing and contrasting the SQL and MDX statements based on statement clauses. *(Draw the table in the answer booklet)*

	Language	
Clause	SQL	MDX
SELECT		
FROM		
WHERE		

[3 Marks]

- C. Complete the following table stating the purpose and description of each data cube operator. *(Draw the table in the answer booklet)*

Operator	Purpose	Description
Slice		
Dice		
Drill-down		
Roll-up		
Pivot		

[5 Marks]

Question 3: Relational Database Concepts for Multidimensional Data

[11 Marks]

- A. What is a grain of a data warehouse? How does the level of granularity affect computation time and the size of a data warehouse?
[2 Marks]
- B. List and explain the **three (3)** types of fact tables in a data warehouse.
[3 Marks]
- C. Compare and contrast between the **Star Schema**, the **Constellation Schema** and the **Snowflake Schema** based on the level of fact and dimension tables.
[3 Marks]
- D. Explain the supply driven design methodology used to develop a data warehouse. Clearly state each required step along with associated tasks.
[3 Marks]

Question 4: Concepts of Data Integration Processes

[14 Marks]

- A. Briefly list and explain **four (4)** constraints on the refresh workflow.
[2 Marks]
- B. What is change data? Explain **two (2)** challenges associated with change data.
[3 Marks]
- C. List and explain the **four (4)** types of change data
[2 Mark]
- D. List and briefly explain **six (6)** data quality problems.
[3 Marks]
- E. What is Parsing in Computer science?
[1 Mark]

- F. Complete the following table stating **two (2)** examples of target strings based on the regular expressions.
(Draw the table in the answer booklet)

Search Expression	Target Strings
"colou?r"	
"tre*"	
"tre+"	
"[abcd]"	
"[0-9]{3}-[0-9]{4}"	
"ba{2,3}b"	

[3 Marks]

Question 5: Architectures and Marketplace

[10 Marks]

- A. Compare and contrast between the **ETL** and the **ELT** architectures.
- B. List at least **four (4)** major features on the Pentaho data integration software.
- C. Explain the advantages of Talend and Pentaho data integration software.
- D. Explain the open core model used by vendors of software's like Talend.

[3 Marks]

[2 Marks]

[3 Marks]

[2 Marks]

Question 6: Relational Database Support Overview

[6 Marks]

- A. List and explain **four (4) reasons** for usage of relational DBMS for data warehouse processing.
- B. What features were missing from DBMS systems to support tasks related to data warehousing.

[4 Marks]

[2 Marks]

Question 7: SQL Subtotal Operators

[6 Marks]

- A. Why do we need subtotal extensions in the Group By clause? List as least **two (2)** reasons.
- B. What is the purpose of the SQL CUBE operator?

[1 Mark]

[1 Mark]

- C. What is the purpose of the SQL ROLLUP operator? [1 Mark]
- D. What is the purpose of the SQL GROUPING SETS operator? [1 Mark]
- E. Briefly list some of the variations of the CUBE, ROLLUP, and GROUPING SETS operators. [2 Marks]

Question 8: Oracle SQL Analytic Functions [7 Marks]

- A. List and explain the **three (3)** drawbacks of using SQL SELECT statement extensions. [3 Marks]
- B. List and explain the **four (4)** different types of Rank functions. [4 Marks]

Question 9: Materialized View Processing and Design [15 Marks]

- A. Why are materialized views important for data warehouses but not important for operational databases? [2 Marks]
- B. Compare and contrast query rewriting for materialized views to query modification for traditional (non materialized) views. [3 Marks]
- C. Briefly explain the matching processes to enable query rewriting. [2 Mark]
- D. List and explain **three (3)** advantages of using traditional views. [3 Mark]
- E. What is the disadvantage of using traditional views? [1 Mark]
- F. Discuss the difference between traditional views and materialized views. [4 Marks]

Question 10: Physical Design and Governance of Data Warehouses [8 Marks]

- A. What are the pros and cons of a MOLAP storage engine? Explain in detail. [2 Marks]
- B. What are the pros and cons of a ROLAP storage engine? Explain in detail. [2 Marks]

C. What are the pros and cons of a HOLAP storage engine? Explain in detail.

[2 Marks]

D. What is Hadoop and what are some of the services it provides?

[1 Mark]

E. Explain the inner working of the MapReduce Framework.

[1 Mark]

THE END