

Exam Revision Questions (A)

The following questions are designed to be similar to the likely content covered and format of the Exam questions with correct answers specified as well. You may wish to use this for helping with your preparation for the Exam.

As the Lecturer is not allowed to attend the Exam, remember in the exam to write down any assumptions that you make for each question!

Please read the blurb below (all names are fictitious and are no relation to any real person or company) to provide context as to answer Questions 1 to 4 of the Exam:

Kim runs the Accounts Payable division at ExampleCorp, a business that produces widgets from multiple factories throughout Perth. As part of this work, Kim wishes to track remittances (notices of payment) from different suppliers, such that she can track when each remittance was made, its total value and the supplier that the remittance belongs to. Kim wishes to be able to analyse information regarding groups of suppliers, not just each supplier, hence she must also be able to track not only who each supplier is, but what industry they are in (e.g. "Construction") and the number of employees each supplier has, such that she is able to create the relevant grouped and aggregated analyses she needs to make business decisions.

Please answer each question, utilising the methods and techniques taught within the learning activities of this unit to answer the question. If something is not stated in the blurb above, it does not apply unless you wish to note so. Explicitly state anything that you do in the question, such as relying on default behaviours. Each question should be considered on its own individually, unless otherwise stated. It is assumed in all questions that you have successfully connected to a database server and are at a stage you can enter a query.

Question One

Write out the relevant commands to create the tables within a relational MySQL database named **SomeDatabase**. Ensure you use the most appropriate data types, specify reasonable default values for all attributes (no defaults are needed for geometric fields and keys) and that appropriate keys have been defined in a best-practice manner.

Answer One

```
CREATE TABLE SomeDatabase.Remittance (remittanceId INT NOT
NULL AUTO_INCREMENT, totalValue FLOAT DEFAULT 0,
remittanceDate DATE DEFAULT "2024-01-01", supplierName
VARCHAR(50), PRIMARY KEY(remittanceId), FOREIGN KEY
(supplierName) REFERENCES
SomeDatabase.Supplier(supplierName) ON DELETE CASCADE);

CREATE TABLE SomeDatabase.Supplier (supplierName VARCHAR(50)
DEFAULT "Supplier", supplierIndustry VARCHAR(50) DEFAULT
"Industry", numEmployees INT DEFAULT 1, PRIMARY
KEY(supplierName));
```

Perfectly OK to use the ID field rather than a supplierName – justify why you think one approach or another is appropriate in terms of unique values for the field/attribute.

Any approach to ON DELETE is also appropriate for the FOREIGN KEY.

Question Two

Assuming that the tables have been created, write out the relevant commands to create the following record data within each of the tables, a supplier named “Genius Corp” in the “Database” industry with 4567 employees along with two remittances that this company has, one for a total value of \$100.50 on 5 September 2023 and another on 10 October 2024 with a total value of \$99.

Answer Two

```
INSERT INTO SomeDatabase.Supplier VALUES ("Genius Corp",
    "Database", 4567);

INSERT INTO SomeDatabase.Remittance VALUES (DEFAULT, 100.50,
    "2023-09-05", "Genius Corp");

INSERT INTO SomeDatabase.Remittance VALUES (DEFAULT, 99.0,
    "2024-10-10", "Genius Corp");
```

Question Three

Assuming the tables have been created and populated with the relevant data as per the questions above, retrieve the number of services that have been rendered by Kim for each year and then sort the results such that 2023 would appear before 2024. The aggregated column should be aliased to **theCount**.

Answer Three

```
SELECT YEAR(remittanceDate), COUNT(totalValue) AS theCount
    FROM SomeDatabase.Remittance GROUP BY YEAR(remittanceDate)
    ORDER BY YEAR(remittanceDate) ASC;
```

Any field that can be assumed to be non-null can be used to obtain the count.

Question Four

Write the command to adjust the data model such that a new column named **gstValue** is added to each remittance, where the value of the GST is calculated each time the data is accessed and is equal to one fifth of the total remittance value.

Answer Four

```
ALTER TABLE SomeDatabase.Remittance ADD COLUMN gstValue FLOAT
    GENERATED ALWAYS AS (totalValue / 5) VIRTUAL;
```

Question Five

GENERATED ALWAYS: 表示这个列是一个计算列，系统会始终根据其他列的值计算这个列的值。
AS (totalValue / 5): 计算表达式

Briefly describe what an R Script Visual is in Power BI and why it would be used compared to a built-in Power BI Visual, through providing two reasons. Briefly describe one downside to using a R Script Visual.

Answer Five

An R Script Visual allows limited analysis to be undertaken and new types of visualisations to be generated in Power BI through writing R code. It uses the same data that is provided to regular Power BI visuals. This type of visual could be used compared to a built-in visual in case a particular type of visual that is desired is not available within the built-in visual library, for additional customisation that is not available within the built-in visuals or to undertake and visualise further analysis beyond what is possible with Power BI. **One downsides may be the the requirement to know how to write R language code, although there are also a few others.**

Question Six

One concept of privacy issues with personal data surrounds what is termed immutable data. Describe what immutable data is and provide an example of a piece of personal information that is immutable and a piece of information that is not. Explain briefly why immutable data is such a big issue with privacy breaches.

Answer Six

Immutable data is information that cannot be reasonably changed (e.g. date of birth), hence data that is not immutable can be reasonably changed (e.g. email address). Immutable data is such a big issue as since it cannot be changed, it makes it more difficult to remediate privacy breaches since the information cannot be reissued, limiting the different amount of solutions to make things right.

Question Seven

Compare and contrast the differences between the **Whole Number**, **Decimal Number** and **Text** data types in Microsoft Power BI's Power Query Editor. Do this by explaining, with examples, what the unique elements are between each type and how these unique factors could aid in the development of a Power BI Dashboard.

Answer Seven

Whole numbers store numeric data that do not contain a value after the decimal point (e.g. 42), whereas decimal numbers store numeric data that may contain a value after the decimal point (e.g. 42.42) and Text data stores strings of text which could also include words and symbols (e.g. "Forty Two"). Numeric types – both whole and decimal - can be aggregated in many ways on a Power BI Dashboard whereas text types can only be used in a categorical manner and aggregated through counting. Different methods exist within Power BI to transform data in any of these three formats.

Question Eight

Write the commands used to create a MongoDB collection named **examCollection**. Into this collection, create two documents: one with a **name** of **Business Intelligence** and **unitCode** of **INMT5526** and another with the **course** of **Master of Analytics**, a **isCurrent** value that is **True** and a **count** of **300**. Ensure you demonstrate appropriate data types. Then, write a command to retrieve all documents from the **examCollection** with a **count** greater than **42**.

Answer Eight

The use of .pretty() is optional.

```
db.createCollection("examCollection")
db.examCollection.insert({name: "Business Intelligence",
    unitCode: "INMT5526"})
db.examCollection.insert({course: "Master of Analytics",
    isCurrent: true, count: 300})
db.examCollection.find({count: {$gt: 42}}).pretty()
```

End of Exam Revision Questions (A)