

## Exam Revision Questions (B)

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:

*Ravi is the operator of a business named WestProCorp which provides database development consulting services to other businesses throughout Western Australia and the world. He wishes to record information regarding the services he provides such that he can analyse past work and predict what may be expected in the future. Each client he serves has a name, number of employees, latitude and longitude point location as well as contact phone number (e.g. 0444 444 444). To each client he provides zero or more services, which are undertaken on a particular date and take a fraction of (or whole amount of) hours, have a particular cost associated with the service and also a short description of the service provided.*

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 **BusinessStuff**. 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 BusinessStuff.Service (serviceId INT NOT NULL
    AUTO_INCREMENT, numHours FLOAT DEFAULT 0, serviceCost
    FLOAT DEFAULT 0, description VARCHAR(100) DEFAULT
    "Service", serviceDate DATE DEFAULT "2024-01-01",
    clientName VARCHAR(50), PRIMARY KEY(serviceId), FOREIGN
    KEY (clientName) REFERENCES
    BusinessStuff.Client(clientName) ON DELETE CASCADE);

CREATE TABLE BusinessStuff.Client (clientName VARCHAR(50)
    DEFAULT "Client", numEmployees INT DEFAULT 0, location
    POINT, contactPhone VARCHAR(12) DEFAULT "0400 000 000",
    PRIMARY KEY(clientName));
```

Perfectly OK to use the ID field rather than a clientName – 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 the tables have been created, write out the relevant commands to create the following record data within each of the tables, a client named “University of WA” with 1000 employees located at 115.82 (E) and -31.98 (S) with phone number 0412 345 678 which has received a “ER Diagram” service for 3.5 hours on the 24th of April 2024 for a cost of \$500 as well as a “Server Setup” service for 6 hours on the 25th of April 2024 for a cost of \$450.99.

## Answer Two

```
INSERT INTO BusinessStuff.Client VALUES ("University of WA",
    1000, ST_GeomFromText('POINT(115.82 -31.98)'), "0412 345
    678");
```

```
INSERT INTO BusinessStuff.Service VALUES (DEFAULT, 3.5, 500,
    "ER Diagram", "2024-04-24", "University of WA");
```

```
INSERT INTO BusinessStuff.Service VALUES (DEFAULT, 6, 450.99,
    "Server Setup", "2024-04-25", "University of WA");
```

## Question Three

Assuming the tables have been created and populated with the relevant data as per the questions above, create a view named **AllBusinessData** which includes the command to join together the data, such that a single view of all the data is created that consists of not only every service provided to any client but the associated client information, only if the client has been provided a service. Ensure that the client's phone number attribute is given an alias of **examClientPhone**. Ensure that all data is presented in a human-readable format, that is using numbers and letters in English. Finally, write a command to show the command used to create the view.

## Answer Three

```
CREATE VIEW BusinessStuff.AllBusinessData AS (SELECT
    BusinessStuff.Client.numEmployees,
    BusinessStuff.Client.contactPhone AS examClientPhone,
    ST_AsText(BusinessStuff.Client.location),
    BusinessStuff.Client.location, BusinessStuff.Service.*
    FROM BusinessStuff.Client INNER JOIN BusinessStuff.Service
    ON BusinessStuff.Client.clientName =
    BusinessStuff.Service.clientName);
```

```
SHOW CREATE VIEW BusinessStuff.AllBusinessData;
```

## Question Four

Assuming the view was created successfully in Question Three, utilise it to retrieve any services that have been rendered within a 10km radius of the Perth CBD located at -31.95 (S) and 115.87 (E) in the month of February, 2024. Sort the results by the

shortest to longest service provided and then by the largest to smallest number of employees that the business has.

### **Answer Four**

```
SELECT * FROM BusinessStuff.AllBusinessData WHERE
    (ST_Distance_Sphere(location,
    ST_GeomFromText('POINT(115.87 -31.95)')) <= 10000) AND
    (MONTH(serviceDate) = '02') AND (YEAR(serviceDate) =
    '2024') ORDER BY numHours ASC, numEmployees DESC;
```

### **Question Five**

Briefly describe the difference between a **Map** and a **Filled Map** visual in Power BI and how different (numeric) data values can be represented in each of the two types of visuals. Explain a consideration you should have (that is, what to be careful of) when using either of these types of maps (that is, one consideration for each type of map). Which type do you think better represents (numeric) data and why?

### **Answer Five**

Both Map and Filled Map visuals in Power BI display a map of the world and allow data to be displayed corresponding to different regions throughout the world, assuming this information is within the data set. Map visuals display data as different sized 'bubbles' where the bubble is displayed in the middle of the location and the size of the bubble is larger for larger values of data. Filled Map visuals colour in or shade the area and display different colours depending on the data values. Maps need to ensure that the different sizes of the bubble are different enough to make it clear which locations have larger and smaller values, whereas filled maps need to ensure that the colour gradient used is distinct enough to ascertain the differences between values. I think that filled maps make it easier to see differences between (summarised) values for numeric data, hence it is preferred as it makes it easier to compare and contrast data.

### **Question Six**

Briefly describe what an ethical issue is that you could encounter in the field of Business Intelligence, that is, through dealing with business data. Provide an example (real or hypothesised) as to how the issue could occur.

### **Answer Six**

Ethical issues are those where user data is used in a manner which is not what the user would expect or has agreed to, yet may still be used in a way that is still acceptable legally – that is for 'less than proper' or questionable purposes. An example of this may be when a user purchases a product off an online store and provides an email address for receiving an invoice, however the owner uses this email address to sign up the user for an email marketing newsletter.

### **Question Seven**

In your Power BI project connecting to the Polygon API, two parameters were created. Identify which of these parameters required you to ensure that it was not converted to a different data type, which data type it would be converted to (and why you would

want it to be that data type) and what data type you had to keep it as. Briefly describe why it had to be in that particular data type that you kept it as.

### **Answer Seven**

The DatePart parameter was used to gather information about stock prices in the project for a particular date. Power BI would attempt to convert this to its Date format, which is useful if we wished to examine dates in a more advanced manner (e.g. by quarter). However, to ensure compatibility with the Polygon API which was expecting the date in a particular format, it had to be kept in the Text format with dates represented in YYYY-MM-DD format.

### **Question Eight**

Write the commands used to create two MongoDB collections named **customerCollection** and **supplierCollection**. Each collection should have a single document inserted into it, which details whether or not the customer/supplier **isCurrent**, the **name** of the customer/supplier and a **contactEmail**. Create sample data for each of the documents (any reasonable example will do), ensuring that at least two of the three attribute values differ between the two collections.

### **Answer Eight**

*The use of .pretty() is optional.*

```
db.createCollection("customerCollection")

db.createCollection("supplierCollection")

db.customerCollection.insert({isCurrent: true,
    name: "Jim Jones", contactEmail: "jim@jones.net"})

db.supplierCollection.insert({isCurrent: false,
    name: "Food Corp", contactEmail: "corp@example.com"})
```

*End of Exam Revision Questions (B)*