**Databases for Data Science**

**Homework 1**

**Provide answers to the following questions as a PDF or Word file:**

**Q1: Enumerate the four types of data, and give one example of each.**

1. Categorical Data- Plant Species
2. Ordinal Data- Leaf Size
3. Interval Data- Plant Discovery Year
4. Ratio Data- Leaf Growth Rate

**Q2: Apart from the examples given in the slides, give another example of a hierarchical data model.**

An example of a hierarchical model is a football coaching staff. The head coach is the root with three children, the offensive, defensive and special teams coordinators. Then each coordinator has children for each position group such as Quarterbacks, Linebackers and Punters.

**Q3. Differentiate between entities, attributes, and records. Which one(s) belong to a schema (model) and which ones(s) belong to an instance?**

An entity is an object in a schema that has attributes that describe it such as person with attributes of eye color, hair color, etc. A record belongs to an instance and is a collection of attributes for a specific entity.

**Q4. Look at the website at** [**https://www.sports-reference.com/olympics/**](https://www.sports-reference.com/olympics/)**. It provides a lot of information on many entities involved in the Olympic Games. Can you list six entities that you can identify from the information provided on that website.**

(The website was closed but I redirected to <https://www.olympedia.org/>)

Six of the entities on the website are Athletes, Olympic Games, Countries, Sports, Event Year, Event Location, and Events.

**For the next question, you need to install the SQLite database management system on your laptop or PC and then download the chinook database and run it in SQLite DBMS. The following two links will help:**

[How to Install SQLite on Windows, Mac, or Linux (servermania.com)](https://www.servermania.com/kb/articles/install-sqlite/)

[SQLite Sample Database and Its Diagram (in PDF Format) (sqlitetutorial.net)](https://www.sqlitetutorial.net/sqlite-sample-database/)

**Q5. In the SQLite installation and the chinook database that you have downloaded, do the following and answer the questions:**

1. **Perform the command: .tables** 
   1. How many tables are there? 11
2. **Perform the query: select \* from genres;** 
   1. What is the genre ID for jazz? 2
3. **Perform the query: select Name from tracks where GenreId = 2;** 
   1. How many rows are there? 130
4. **Write a note on what you think you did in the above three tasks.**

I determined what data or tables existed in the database. Then queried the genres table to determine what genres where in the database. I then found what tracks existed in the database for my genre of choice, Jazz.

I then used PRAGMA table\_info(genre) to confirm that GenreId is the Primary Key and used PRAGMA table\_info(tracks) to confirm that GenreID existed

I used .tables to figure out what and how the data was stored. I determined it was Music and wanted to see the selection of Jazz tracks that existed but first I needed to know what GenreId was for Jazz. I then was able to use that key to query how many jazz tracks there were.

I determined the schema of the database.