# Data Science
# Assignment 1

**Submitted By:**
**Aman Chauhan**
**1805158**

Q1. Differentiate b/w SQL and NoSQL Databases.
Ans.

| | **SQL** | **NoSQL** |
|---|---|---|
| 1) | Databases are categorized as Relational Database Management System (RDBMS). | NoSQL databases are categorized as Non-relational or distributed database systems. |
| 2) | SQL databases have fixed or static or predefined schema. | NoSQL databases have dynamic schemas. |
| 3) | SQL databases display data in the form of tables so it is known as table-based databases. | NoSQL databases display data as collections of key-value pairs, documents, graph databases or wide-column stores. |
| 4) | SQL databases are vertically scalable. | NoSQL databases are horizontally scalable. |
| 5) | SQL databases use a powerful language "Structured Query Language" to define and manipulate the data. | In NoSQL databases, collections of documents are used to query the data. It is also called unstructured query language. It varies from database to database. |
| 6) | SQL databases are best suited for complex queries. | NoSQL databases are not so good for complex queries because these are not as powerful as SQL queries. |

| | | |
|---|---|---|
| 7) | SQL databases are not best suited for hierarchical data storage. | NoSQL databases are best suited for hierarchical data storage. |
| 8) | MySQL, Oracle, Sqlite, PostgreSQL and MS-SQL etc. are examples of SQL databases. | MongoDB, BigTable, Redis, RavenDB, Cassandra, Hbase, Neo4j, CouchDB etc. are examples of nosql databases. |

Q2. Explain the following terms in brief with examples:

1.Discrete and Continuous variables.

**i) Discrete Variables:** It is a count that involves integers. Only a limited number of values is possible. The discrete values cannot be subdivided into parts. So, discrete data can take only certain values. The data variables cannot be divided into smaller parts.

## Characteristics:

- You can count the data. It is usually units counted in whole numbers.
- The values cannot be divided into smaller pieces and add additional meaning.
- It has a limited number of possible values e.g. days of the month.

## Examples:

- The number of students in a class.
- The number of workers in a company.
- Number of languages an individual speaks.

ii) **Continuous Variables:** Continuous data is information that could be meaningfully divided into finer levels. It can be measured on a scale or

continuum and can have almost any numeric value. You can record continuous data at so many different measurements – width, temperature, time, and etc.

## Characteristics:

- In general, continuous variables are not counted.
- The values can be subdivided into smaller and smaller pieces and they have additional meaning.
- It has an infinite number of possible values within an interval.

## Examples:

- The amount of time required to complete a project.
- The height of children.
- The amount of rain, in inches, that falls in a storm.

## 2. Quantitative and Qualitative variables.

i) **Quantitative Variables:** Quantitative Variables can be expressed as a number, so it can be quantified. In simple words, it can be measured by numerical variables.

## Examples:

- Scores of tests and exams e.g. 74, 67, 98, etc.
- The weight of a person.
- The temperature in a room.

ii) **Qualitative Variables:** Qualitative data can't be expressed as a number, so it can't be measured. It mainly consists of words, pictures, and symbols, but not numbers.It is also known as Categorical Data as the information can be sorted by category, not by number.

## Examples:

- Colors e.g. the color of the sea

- Popular holiday destinations such as Switzerland, New Zealand, South Africa, etc.
- Ethnicity such as American Indian, Asian, etc.

3. Dependent and Independent variables

i) **Dependent Variables:** The dependent variable is the condition that you measure in an experiment. You are assessing how it responds to a change in the independent variable, so you can think of it as depending on the independent variable.

## Examples:

- In a study to determine whether how long a student sleeps affects test scores, the dependent variable is the test score.
- In an experiment to determine how far people can see into the infrared part of the spectrum, the light is observed (the response) is the dependent variable.
- You want to compare brands of paper towels, to see which holds the most liquid. The dependent variable would be the amount of liquid absorbed by the paper towel.

ii) **Independent Variables:** The independent variable is the condition that you change in an experiment. It is the variable you control. It is called independent because its value does not depend on and is not affected by the state of any other variable in the experiment.

## Examples:

- In a study to determine whether how long a student sleeps affects test scores, the independent variable is the length of time spent sleeping.

- In an experiment to determine how far people can see into the infrared part of the spectrum, the wavelength of light is the independent variable.
- You want to compare brands of paper towels, to see which holds the most liquid. The independent variable in your experiment would be the brand of paper towel.