

Question1: Identify 4 different Relational Database Management Systems (RDBMS) and describe some of the difference between them.

Solution: Relational Database Management Systems have been bedrock of the growth of Software Development over last four decades. Without RDBMS, there would have been scaling issues related to Software Products and as we all know, **No-Scale** means **No-Growth**. And, if software industry would not have grown the way it did then today world would have been totally different. We cannot emphasize the importance of RDBMS more and every software student should have clear understanding of what RDBMS are and what types are available/popular in the market. For the purpose of this assignment, we have identified four types of RDBMS and discussed each's features and USP:

Microsoft SQL Server	Oracle DB	SQLite	PostgreSQL
About: Microsoft SQL is a popular RDBMS developed by Microsoft.	About: Oracle Corporation owns Oracle DB and code is not open source. Oracle DB is being the most widely used RDBMS in banking	About: SQLite is popular open source SQL database.	About: PostgreSQL is an open source SQL database and is typically used for web application development.

	industry across the globe.		
USP: a. Customized solutions for Microsoft Infrastructure. b. Excellent Data Recovery Support c. Installation is streamlined.	USP: a. Batch processing of transactions make Oracle DB standout from its competitors. b. Scale is not an issue with Oracle DB. c. Provides functionality in built in DB Management System targetted to specific industry.	USP: a. SQLite can store entire database in a single file. b. Data can be stored locally and therefore no need to connect database to a server. c. No installation needed.	USP: a. Easy to Use. b. Inexpensive c. Reliable d. Large Development Community

	d. Extensive language Support		
Disadvantage: a. Licensing is pretty expensive. For SQL Server 2008, for instance, SQL Server Standard Edition costs \$7,171 per processor. The SQL Server Datacenter edition is \$54,990 per processor, according to the Microsoft website. b. Limited Compatibility as MS SQL-Server is designed to run on Windows-bases servers.	Disadvantage: a. Expensive than Microsoft. b. Steep Learning Curve.	Disadvantage: a. Database size is restricted to 2GB in most cases. b. SQLite is used to handle low to medium HTTP requests.	Disadvantage: a. Slower performance than other open source DBMS such as MySql b. Less popular than MySql

Table1: Comparison of Four RDBMS.

Question 2: Research the popularity of various RDBMS vendors among both small and large businesses and demonstrate how much each is being used. You must note the source of your data.

Solution 2: According to the data published as of September, 2019(Statistics, 2019), the most popular database management systems(DBMS) in the world was **Oracle** – which means it is **the first choice of large organizations across the globe** and it seems obvious too as Oracle DB does not have OS restrictions and it supports extensive programming languages. After Oracle, next most popular RDMS is **MySQL** – **open source DBMS and this trend indicates that small** and medium size enterprises are going with MySQL. Another interesting trend from statistics is that database management system industry is dominated by big tech companies such as Microsoft, Oracle and IBM but statistics is showing that number of free and open sources DBMS such as MySQL, PostgreSQL and Apache Cassandra.

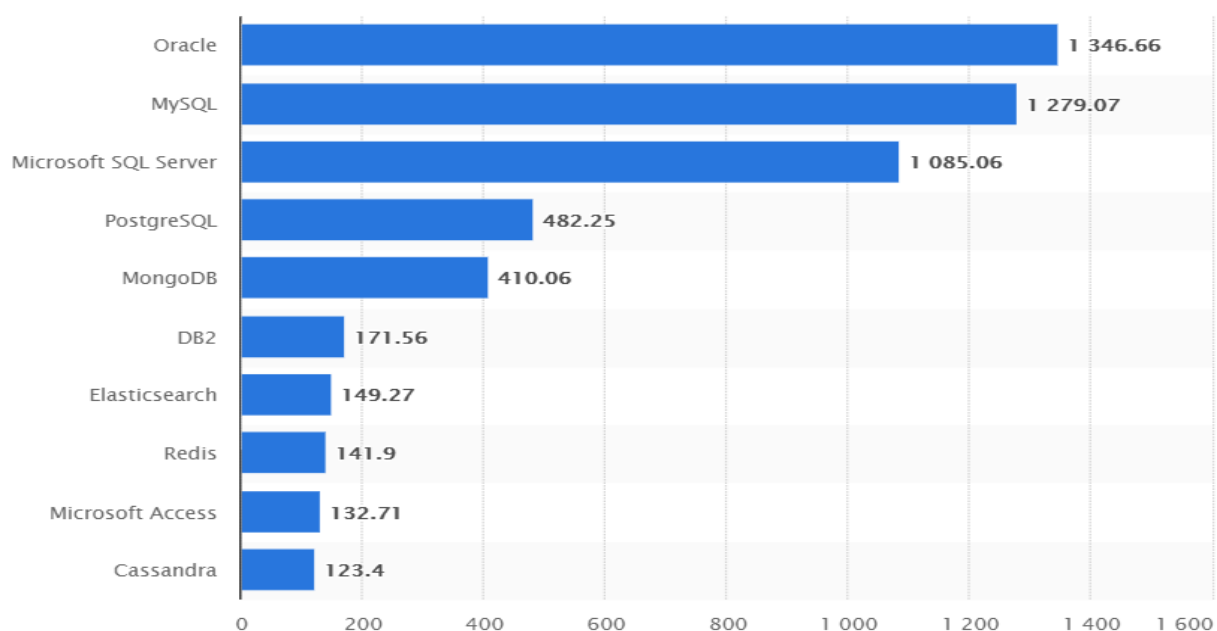


Figure1: Popularity of DBMS worldwide. Adapted from Statistica, 2019.

Question 3: Discuss the differences between relational databases and non-relational databases and highlight what situations are best for each.

Solution:

Relational Database	Non-Relational Database
Data is stored and represented in the form of tables consisting of rows and columns.	No use of Tables. Instead they use the storage mechanism which is suited as per the requirement. For example, as Json Objects etc.
In Relational database, for selecting a row, the entire database is loaded and then the row is selected.	In Non-Relational database, the data to be fetched is referred by an index. The data is stored sequentially adjacent to each other. So, it does not require the entire file to be opened or loaded.
They can be used when dealing with several amount of data since the data is fetched by selecting a row which requires the entire file to be loaded.	Non-relational database should preferably be used in situations where huge amount of data is involved since it fetches data by index.
Example of relational database: MySQL, PostgreSQL	Example of non-relational database: MongoDB

Relational database is best suited for situations where small to average amount of data is involved. Whereas, non-relational database is used for the situations where huge amount of data is involved since the data is stored adjacent to each other. The data can

be picked up by the use of an index. It does not require to open the entire file and hence is more efficient in case of large amount of data.

Question 4: Explain what a “Data Warehouse” is as if you were telling someone who knows nothing about databases. Describe its characteristics as well as why you would use one.

Solution 4: A database is grouping of data in an organized manner. A department or organization has various databases since they are required to store information about everything. For example, an organization needs to store their employee data, customer information, product details etc.

A data warehouse is a database which acts as a central repository or a central storage location for all the data in an organization. The data comes from the different department of an organization like CRM, financial and HR system etc. All the data from these systems are collected and copied into the data warehouse.

Data warehouse does not only contain the internal data of an organization. But it also contains external data that such as weather etc. which can affect the decision making and they can be used for analytical purposes.

They are used for analytical and business purposes.

Reasons for using data warehouse are enumerated below:

- Data warehouse helps to have a single version of data. If the data is stored at different location, the data might differ.

- Data warehouse provides simplicity while accessing the data since the data is normalized and organized in a proper way which helps easy retrieval.
- Data warehouse allows persistence of data in a way that organization requires. Data warehouse stores historical data.

For example, if the organization requires full history of data along with the current version, data warehouse allows to do that and presents the information in desired manner.

REFERENCES

[1] Statistica, 2019. Worldwide Popularity Ranking Database Management Systems. Retrieved from :

www.statista.com/statistics/809750/worldwide-popularity-ranking-database-management-systems/