EXCEL Homework 2

With the decreasing cost of providing technology and the increasing human activity on the internet comes an enormous increase in the amount of data collected. Researching a company you may like to interview with can often provide insight into the kinds of questions they might ask you in an interview. How does their business use data for its primary value? In other words why did they build their software?

# To empower you on your next interview here is a list of currently popular terms that describe how some of databases are being used today to fulfill their primary goal. Google them and collect a 1-3 sentence definition or description of each and include at least one use case for each. A use case is an example of how somebody will use it. A transactional database use case is: A person drives up to an ATM to withdraw cash from their bank account.

1. Data warehouse

An integrated repository of data from multiple sources used for reporting and data analysis. They usually contain both historical and up-to-date data.

Use case example:

1. Data mining, finding patterns in large data sets for analysis.
2. Predictive analysis, finding patterns in data which allow predictions of future outcomes.
3. Transactional , or Real Time database

A transactional database is one whose data actions (transactions) are divided into discrete operations with a distinct beginning and end. This allows for transactions to be reversed or rolled back if needed.

A real time database is a transactional database whose speed of transaction processing is sufficient to allow immediate interaction with the database while maintaining the integrity of the data in the database.

Use case example:

1. An airline reservation system where a customer buys tickets on a flight.
2. Analytical database

A database, typically read-only, which contains historical data useful for some particular analysis. The database is updated periodically as needed rather than immediately as in a real time database.

Use case example:

1. A company may use this database to analyze sales performance over time.
2. Time series database

This is a database optimized to process time series data, data points which are ordered or indexed by time.

Use case example:

1. Collect real time data from sensors and store for later analysis.
2. Flat file database

Data stored in a single file, typically stored in fixed length records or delimited data, either binary or plain text.

Use case example:

1. Store application configuration data locally.
2. Spatial database

A database optimized to store and retrieve data which represents objects in a given space. The store objects are typically geometric or geographic in nature.

Use case example:

1. A geographic information system which stores and displays geographic data.
2. Cloud database

This is a database which is stored on a cloud computing platform. Database operations are performed by means of services provided by the cloud platform. Database security and maintenance are handled by the cloud platform and not by the users who access the database.

Use case example:

1. A commerce web site displays items for sale and handles purchases of those items.

# Here is a list of some popular database products being used today. Google them to discover how they distinguish themselves from the others. List at least one example of who is using it for what purpose, and one example of why it was chosen over the others.

1. [Oracle](https://en.wikipedia.org/wiki/Oracle_Corporation)

Object-relational database optimized for large databases. Runs on multiple platforms and cloud.

User: PayPal uses it to store payment transaction data such as customer and payment information.

1. [MySQL](https://en.wikipedia.org/wiki/MySQL)

Open-source relational database, provides cheap and easy entry into database management across multiple platforms. Recently acquired by Oracle.

User: Uber uses MySQL to store web-based transactions including trip history, client billing, etc.

1. [Microsoft SQL Server](https://en.wikipedia.org/wiki/Microsoft_SQL_Server)

Microsoft relational database server, integrated with other Microsoft tools and platforms.

User: National Oilwell Varco uses this to store all data gathered at oil rigs including downhole sensor data, daily drilling reports, and worker payroll information.

1. [PostgreSQL](https://en.wikipedia.org/wiki/PostgreSQL)

Free open-source object-relational database, runs on multiple platforms.

User: IMDB.com uses this to store all their movie information.

1. DB2

IBM’s relational database, runs on multiple platforms, support provided by IBM. Provides interfaces for older programming languages such as COBOL and FORTRAN in addition to more modern languages.

User: China Minsheng Banking Corp

Uses DB2 to provide 24/7 support for client banking transactions throughout China.

1. Vertica

Database specialized for data warehouse querying by means of a column-oriented storage structure.

User: Cerner Corporation

Chose Vertica to store their electronic health records to support real-time data analysis of those records. Key factors were performance of the database for large analysis projects as well as concurrency allowing multiple users to access the database.

1. SAP Sybase ASE

Relational database designed for high-performance transaction-based applications.

User: Sony Network Communications

Sony uses this database to manage network subscriber accounts.

1. Teradata

Relational database supporting business analytics. Multiple deployment options (local, cloud-based, hybrid local/cloud). Uses “shared-nothing” architecture where each server node is independent.

User: 7-Eleven

7-Eleven uses the Teradata database to drive a customer loyalty program through a mobile app.

# The class focuses on learning SQL using Oracle. All of the databases above use SQL. The instructor postulates what you learn in this class is 95-97% transferrable to any of the products listed that in a typical job setting is. Is that true? Or, is the instructor full of beans and gravy.

# Founded in 1918, ANSI (American National Standards Institute) is a private, non-profit organization that administers and coordinates the U.S. voluntary standards and conformity assessment system. Periodically ANSI updates and republishes the “Database Languages – SQL” standards. Each manufacturer claims ANIS compliance with these standards.

# Below is a table showing syntax differences between Oracle SQL and Microsoft SQL Server SQL. Although you may intuit meaning from their names, we have not covered any of these and you are not expected to understand them. Three of the columns are blank. Fill in these three columns.

**SQL Functions: Description and Syntax**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Description** | **Oracle** | **MS SQL Server** | **PostgreSQL** | **MySQL** | **Vertica** |
| Find smallest integer >= n | CEIL | CEILING | CEIL | CEILING | CEILING |
| Modulus | MOD | % | MOD, % | MOD | MOD |
| Truncate number | TRUNC | <none> | TRUNC | TRUNCATE | TRUNC |
| Translate NULL to n | NVL | ISNULL | COALESCE | ISNULL | NVL |
| Return NULL if two values are equal | DECODE | NULLIF | NULLIF | NULLIF | NULLIF |
| String concatenation | CONCAT(str1,str2) | str1 + str2 | CONCAT | CONCAT | CONCAT |
| Capitalize first letters of words | INITCAP | <none> | INITCAP | <none> | INITCAP |
| Find string in string | INSTR | CHARINDEX | POSITION, STRPOS | INSTR | INSTR |
| Find pattern in string | INSTR | PATINDEX | REGEXP\_MATCHES | REGEXP | REGEXP\_INSTR |
| String length | LENGTH | DATALENGTH | LENGTH | LENGTH | LENGTH |
| Pad string with blanks | RPAD, RPAD | <none> | LPAD. RPAD | LPAD, RPAD | LPAD,RPAD |
| Trim leading or trailing chars other than blanks | LTRIM, RTRIM, TRIM | <none> | LTRIM,RTRIM,TRIM | LTRIM, RTRIM,TRIM | LTRIM,RTRIM,TRIM |
| Replace chars in string | REPLACE | STUFF | REPLACE | REPLACE | REPLACE |
| Convert number to string | TO\_CHAR | STR, CAST | ENCODE | CAST | TO\_CHAR |
| Convert string to number | TO\_NUMBER | CAST | DECODE | CAST | TO\_NUMBER |
| Get substring from string | SUBSTR | SUBSTRING | SUBSTR | SUBSTR | SUBSTR |
| Date addition | ADD\_MONTH or + | DATEADD | + | ADDDATE | ADD\_MONTHS, + |
| Date subtraction | MONTHS\_BETWEEN or - | DATEDIFF | - | DATEDIFF | DATEDIFF |
| Last day of month | LAST\_DAY | <none> | <none> | LAST\_DAY | LAST\_DAY |
| Time zone conversion | NEW\_TIME | <none> | <none> | <none> | NEW\_TIME |
| Next specified weekday after date | NEXT\_DAY | <none> | <none> | <none> | NEXT\_DAY |
| Convert date to string | TO\_CHAR | DATENAME, CONVERT | TO\_CHAR | DATE\_FORMAT | CAST |
| Convert string to date | TO\_DATE | CAST | TO\_DATE | STR\_TO\_DATE | CAST |
| Convert date to number | TO\_NUMBER(TO\_CHAR(d)) | DATEPART | DATE\_PART | EXTRACT | DATE\_PART |
| Date round | ROUND | CONVERT | ROUND | <none> | ROUND |
| Date truncate | TRUNC | CONVERT | DATE\_TRUNC | <none> | DATE\_TRUNC |
| Current date | SYSDATE | GETDATE | CURRENT\_DATE | CURDATE | CURRENT\_DATE |
| If statement in an expression | CASE, DECODE, COALESCE | CASE, COALESCE | CASE, COALESCE | CASE | CASE |
| Current user | USER | USER | USER | CURRENT\_USER | CURRENT\_USER |