Rachelle “Chelly” Bell

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
   1. A relational database that is designed for query and analysis rather than for transaction processing. It usually contains historical data derived from transaction data, but it can include data from other sources. It separates analysis workload from transaction workload and enables an organization to consolidate data from several sources. Most companies use data warehousing for validation, tactical reporting or exploration.
      1. For example, Denver consumers buy products differently than New York City consumers. New York folks tend to purchase a candy bar on a whim (city population buying patterns), where Denver folks are less likely to do so (rural population buying patterns). This has been hypothesized for years, but empirical data shows it to be true.
2. Transactional , or Real Time database
   1. A database system providing all features on traditional database system such as data independence and concurrency control, while at the same time enforces real-time constraints that applications may have. Tasks have time constraints, a RTDBS has the added requirement to ensure some degree of confidence in meeting the system’s timing requirements.
      1. The Stock Market Exchange utilizes RTDBS; the fast pace and ever-changing environment making a RTDBS a valued tool. RTDBS are equipped to handle the graphs of the different markets that appear to be very unstable and yet a database has to keep track of current values for all of the markets of the Stock Exchange.
3. Analytical database
   1. A read-only system that stores historical data on business metrics such as sales performance and inventory levels. An analytic database is specifically designed to support business intelligence and analytic applications, typically as part of a data warehouse or data mart
      1. Hospitals and retail business use these types of databases to keep track of patient/customer records and supplies.
4. Time series database
   1. **(TSDB)** is a software system that is optimized for handling time series data, arrays of numbers indexed by time (a datetime or a datetime range). In some fields these *time series* are called profiles, curves, or traces. A time series of stock prices might be called a price curve. A time series of energy consumption might be called a load profile. A log of temperature values over time might be called a temperature trace. Time series data is a sequence of data points collected at regular intervals over a period of time. In short, it is any data that has a timestamp, including Internet of Things, device data, stocks, commodity prices, tide measurements, solar flare tracking, and health information.
      1. Informix TimeSeries; A time series database is optimized to meet the challenges of handling massive amounts of data from thousands or more devices as in a large corporation that employees have company cell phones, laptop and etc.
5. Flat file database
   1. A flat file database is a database that stores data in a plain text file. Each line of the text file holds one record, with fields separated by delimiters, such as commas or tabs.
      1. Universities use to store students information from DOB to ID numbers
6. Spatial database
   1. A database that is enhanced to store and access spatial data or data that defines a geometric space. These data are often associated with geographic locations and features, or constructed features like cities. Data on spatial databases are stored as coordinates, points, lines, polygons and topology. Some spatial databases handle more complex data like three-dimensional objects, topological coverage and linear networks.
      1. A road map utilizes spatial databases.
7. Cloud database
   1. A type of database service that is built, deployed and delivered through a cloud platform. It is primarily a cloud Platform as a Service (PaaS) delivery model that allows organizations, end users and their applications to store, manage and retrieve data from the cloud. A cloud database is a database that has been optimized or built for a virtualized computing environment.
      1. “Amazon Web Services has a variety of cloud-based database services, including both relational and NoSQL databases. Amazon Relational Database (RDS) run either MySQL, Oracle or SQL Server instances.”

# 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)
   1. Pandora; uses with finance department to read real-time data and easy access with all information is in one place for faster more robust reporting.
      1. The chose Oracle because of their engagement through the development process and their ability to grow and adjust to Pandora’s ever-changing needs.
2. [MySQL](https://en.wikipedia.org/wiki/MySQL)
   1. Twitter; open source relational databases to run globally-distributed, highly available and consistent web applications.
      1. Twitter choses MySQL because they are familiar with the functionality as well as it is easily upgraded or downgraded depending on project. “We know the source code”
3. [Microsoft SQL Server](https://en.wikipedia.org/wiki/Microsoft_SQL_Server)
   1. Visus, LLC; uses Ms SQL Server for the following: “Automated aggregation of data from multiple data sources within an enterprise, extension of backend systems (Oracle ERP, AS400, MAS500, and other database systems) to secure B2B portals, Reporting solutions, customer portals for account management, product selection and purchases, business Intelligence Reporting through SSRS and Performance Point and the implementation of SQL Server reduces costs and allows businesses to quickly respond to changing business needs by allowing for the rapid development of solutions. The resulting applications can streamline business operations improving the management and access of information, which fosters better use of business information, improved customer service and ultimately better decisions.”
      1. Microsoft SQL Server is used because there are a wide range of options for users such as small business to large aimed at different audiences and for workloads ranging from small single-machine applications to large Internet-facing applications with many concurrent users.
4. [PostgreSQL](https://en.wikipedia.org/wiki/PostgreSQL)
   1. VMWare; PostgreSQL provides some of the most advanced enterprise capabilities of any open source database, and is backed by a vibrant and innovative community with proven customer success.
      1. PostgreSQL standards support and ACID compliance outweighs MySQL’s speed. PostgreSQL not only got faster, it also added JSON, making it one of the few relational databases to also support NoSQL.
5. DB2
   1. Sicoob; utilizes to reduce reporting time, relatable, reliable and customer service.
      1. Sicoob choose DB2 for their exclusive BLU Acceleration technology as well as its capability to Oracle SQL.
6. Vertica
   1. JP Morgan Chase Bank; “Chase’s IBIE data management platform collects, calculates, aggregates, and distributes client-related reference data on a broad collection of financial metrics for Chase’s wholesale bank and global investment bank generating over 2,000 reports per day.”
      1. I think a link best describes why they chose this operating system.
      2. <https://www.hpe.com/h20195/V2/GetDocument.aspx?docname=4AA5-8726ENW&cc=us&lc=en&es_p=594760>
7. SAP Sybase ASE
   1. National Train Enquiry System (India’s Railway System); utilizes SAP Sybase ASE for its ability to record, track and store millions of concurrent transactions as well as multi-terra bytes of information.
      1. NTES uses SAP Sybase ASE because of its 99.9% accuracy and availability.
8. Teradata
   1. NCR; uses Teradata to improve speed and increase detection of faults.
      1. NCR chose Teradata to increase correlation and predictive algorithms, ultimately reducing downtime for their global customers.

# 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 | LEAST (CEIL) | CEILING |
| Modulus | MOD | % | Modulo | MOD | MOD |
| Truncate number | TRUNC | <none> | TRUNC | TRUNCATE | TRUNC |
| Translate NULL to n | NVL | ISNULL | NVL | IFNULL | NULL |
| Return NULL if two values are equal | DECODE | NULLIF | DECODE | DECODE | NULLIF |
| String concatenation | CONCAT(str1,str2) | str1 + str2 | (Str1, str2) | CONCAT | (STR1,STR2) |
| Capitalize first letters of words | INITCAP | <none> | INITCAP | <NONE> | INTICAP |
| Find string in string | INSTR | CHARINDEX | <NONE> | CHARINDEX | INSTR |
| Find pattern in string | INSTR | PATINDEX | PATINDEX | PATINDEX | INSTR |
| String length | LENGTH | DATALENGTH | LENGTH | CHAR\_LENGTH | CHARACTERLENGTH |
| Pad string with blanks | RPAD, RPAD | <none> | RPAD | RPAD | RPAD |
| Trim leading or trailing chars other than blanks | LTRIM, RTRIM, TRIM | <none> | LTRIM,TRIM | TRIM,RTRIM | TRIM |
| Replace chars in string | REPLACE | STUFF | REPLACE | REPLACE | REPLACE |
| Convert number to string | TO\_CHAR | STR, CAST | TO\_CHAR | CAST () | TO\_CHAR |
| Convert string to number | TO\_NUMBER | CAST | TO\_NUMBER | CONVERT() | TO\_NUMBER |
| Get substring from string | SUBSTR | SUBSTRING | SUBSTRING | SUBSTRING | SUBSTR |
| Date addition | ADD\_MONTH or + | DATEADD | DATEADD | DATE\_ADD | TIMESTAMPADD |
| Date subtraction | MONTHS\_BETWEEN or - | DATEDIFF | DATESUB | DATE\_SUB | DATEDIFF |
| Last day of month | LAST\_DAY | <none> | LAST\_DAY | LAST\_DAY() | LASTDAY |
| Time zone conversion | NEW\_TIME | <none> | NEW\_TIME | NEW\_TIME | NEWTIME |
| Next specified weekday after date | NEXT\_DAY | <none> | NEXT\_DAY | NEXT\_DAY | NEXTDAY |
| Convert date to string | TO\_CHAR | DATENAME, CONVERT | TO\_DATE | TO\_DATE | TO\_DATE |
| Convert string to date | TO\_DATE | CAST | TO\_DATE | STR\_TO\_DATE | CAST |
| Convert date to number | TO\_NUMBER(TO\_CHAR(d)) | DATEPART | TO\_CHAR | TO\_CHAR | TO\_CHAR |
| Date round | ROUND | CONVERT | ROUND | ROUND() | ROUND |
| Date truncate | TRUNC | CONVERT | DATE\_TRUNC | DATE\_TRUNC | DATETRUNC |
| Current date | SYSDATE | GETDATE | CURRENT\_DATE | CURDATE | SYSDATE |
| If statement in an expression | CASE, DECODE, COALESCE | CASE, COALESCE | CASE, DECODE, COALESCE | CASE, DECODE, COALESCE | CASE,NVL |
| Current user | USER | USER | USER | USER | USER |