**Lesson 8: Foundations of business intelligence**

**Introduction to files**

* **Bit –** the smallest unit of data a computer is able to process.
* **Byte – *8 bits*** is equivalent to a byte. A byte is equivalent to ***one character*** i.e. *letter, number, underscore,* etc.
* **Word –** represent a group of characters i.e. hacker
* **Field –** a group of words e.g. *ominous SHELL.*
* **Record –** a group of ***related*** fields e.g. hackerName, hackerSkill

A ***file*** is a collection of records of the ***same type.***

**Challenges of traditional file environment**

* **Data redundancy –** if the ***same data*** is stored in ***multiple other files***, there is wastage of storage resources.
* **Data inconsistency –** if constraints are not put in place, the ***same field*** may have ***different*** data types or stored information.
* **Poor security –** if no constraints are put in place, anyone who has access to the system also has access to the files. Thus there is greater risk of losing the integrity of this data.
* **Limited data sharing and availability –** transferring files over a network or to a storage device especially hundreds of files, consumes a lot of computer or network resources.

**Relation databases**

Relation database represent data in ***2-dimensional*** tables. These tables will have ***rows*** and ***columns.***

**Capabilities of Database Management Systems (DBMS)**

A database is a central storing system, it’s able to control redundancy of data and provide applications or users with access to this data efficiently.

* System admins are able to ***manage data*** efficiently.
* Provides a ***logical view***of data which often shows the ***relationship*** between different entities.
* The ***physical view*** shows the ***actual data*** and how it is stored, organized and structured on a physical storage device.

**Business intelligence**

Business intelligence refers to the process of ***collecting, examining and interpreting*** large organization ***data,*** often using ***strategy, software tools and techniques.*** It aims to discover ***findings, patterns or valuable information*** that can be used to make more ***informed business decisions***.

**Areas where business intelligence can be applied**

* Data mining
* Budget planning
* Web and ecommerce analysis
* Tracking performance and marketing campaigns.

**Business intelligence infrastructure**

* ***Data warehouse –*** a database that stores all ***interesting data to decision-makers***. This data can be ***current*** and ***historical***, containing information from ***different segments*** in the company.
* ***Data mart –*** a ***group*** of the data ***contained*** in the data warehouse, which is ***summarized*** and is ***focused*** to a ***specific topic, segment or function*** in the organization. E.g. ecommerce’s data mart – will ***compose*** of information ***relating*** to the business ***ecommerce***, i.e. ***web traffic, advertising, online sales, end-users, etc.***
* ***Data mining –*** the process of ***discovering obscure patterns or relationships*** inside ***large*** volumes of ***corporate data***, as well as ***deriving*** ***rules*** in order to ***anticipate*** or ***predict*** these patterns in the ***future***.
* ***Text mining –*** involves the ***extraction*** of ***key*** ***elements*** out of ***unorganized*** data sets, e.g. phone number, emails, call center transcripts, memos, etc.

**Critical success factors to managing an organization’s data**