IST722: Class Exercise 1

**This is an individual assignment.**

**Before you begin, please make sure you’ve read and understand 1) our class honor code, 2) course policies on late work and 3) participation policies as posted on the syllabus. “I didn’t know” is not an excuse.**

**You should cite your sources in a standard format like MPA or APA and include a list of works cited.**

|  |  |
| --- | --- |
| Your Name: | Abhijith Anil Vamadev |
| Your Email: | avamadev@syr.edu |

# Instructions (Refer Unit 1)

Answer each of the following questions as concisely as possible. More is not necessarily better. Please justify your answer by citing your sources from the assigned readings from our textbooks, our class lectures, or online if directed to do so. Be sure to cite in text and include a list of works cited. Place your answer below each question. When you’re finished, print out this document and bring it to class as part of your participation grade.

# Questions

[1] Why is data in organizations the more important asset?

- Data is considered an asset because it brings value and money for the organization. Without data you can’t identify customer, their needs. Data helps to identify what products you should and should not put for sale. Data helps to identify expenses and profitability. Each organization level has different requirements for data.

[2] What are the 4 characteristics of a data warehouse? Provide your own novel examples of each.

* + 1. Subject Orientation - Data is domain specific. Built around business entities/processes. Used by all functional areas.
    2. Non-volatility – Data are never removed or changed, data always grows with time.
    3. Integrated – Centralized in one place, holds the entire organization data. Single version of the truth.
    4. Time-variant – Flow of data through time. Data reflected as a point in time.

[3] You probably noticed we made a "copy" of the data from the source system to the data warehouse. Can you think of three reasons why the data must be a copy and you cannot just use the original data?

1. Answer: In Case errors occurs, keeping a backup enables us to go back to our original data. Provides security and saves up cost and time, helps to maintain consistent copies of data after major updates to the data. Different versions of the data can be used by multiple users at any point in time. Also, it helps to avoid slowing down the system.

[4] What is the difference between business intelligence and data warehouse?

1. Business intelligence – Analytical and decision-support capabilities of the data warehouse. Informed decision-making, the visual or front-end that the users see.
2. Data warehousing – as a foundation, by which we base the business intelligence.

[5] How do Inmon and Kimball approaches to Data Warehousing differ? Be brief.

|  |  |  |
| --- | --- | --- |
|  | Inmon | Kimball |
| Father of | Data warehouse | Business intelligence |
| Invented | Data warehouse | Dimensional models |
| Data warehouse is | Normalized tables | Dimensional models |
| Purpose of data warehouse | Data integration | Query |
| Approach: | Data-first (iterative, bottom-up). Relational modelling. Entity-relational model. Data marts – Not believed to be part of the data warehouse.  Inmon advocates the hub-and-spoke architecture (e.g., the Corporate Information Factory), while | Process-first (waterfall, top-down). Dimensional modelling. Fact tables and dimensions. Data marts – believes they are part of the data warehouse).  Kimball promotes the data mart bus architecture with conformed dimensions. |

WORKS CITED: From video lectures.