

1.0: Essay

Data Design:

Data design is the process of creating a structure for data that is efficient, effective, and easy to use. It involves determining the data requirements, designing the database schema, and executing it. Data design is a crucial element of any software development project because it ensures that the data is managed and available to support the application's functionality. This essay will discuss the basic principles of data design and the best procedures for creating a successful data design.

The first step in data design is to apprehend the data needs. This involves defining the data commodities, attributes, and relationships. Data entities are the objects that are represented in the database, such as customers, orders, or products. Attributes are the characteristics of the entities, such as name, address, or price. Relationships define how the entities are linked to each other, such as a customer maintaining multiple orders.

Once the data necessities have been determined, the next step is to design the database schema. The schema is the blueprint for the database and defines the structure of the tables, columns, and relationships. There are several kinds of database models, including hierarchical, network, relational, and object-oriented. The most commonly used model is the relational model, which uses tables to arrange data and relationships to connect them.

When developing the schema, it is important to consider the performance and scalability of the database. This involves optimizing the table structure, defining indexes, and normalizing the data. Normalization is the process of breaking down large tables into smaller, more manageable tables to decrease monotony and improve data consistency.

Another important part of data design is data integrity. This involves ensuring that the data is correct, uniform, and accurate. There are several ways to ensure data probity, including data validation, referential integrity, and conditions. Data validation checks the data entered into the database to ensure that it meets certain criteria, such as a valid date or a required field. Referential integrity ensures that the relationships between tables are maintained, such as a customer having a valid order. Constraints define rules for the data, such as the maximum length for a field.

In addition to data integrity, security is also a critical aspect of data design. This involves ensuring that the data is protected from unauthorized access and that prudent data is encrypted. Security measures include user authentication, access control, and encryption. User authentication ensures that only authorized users can access the database, while access control defines the level of access each user has. Encryption ensures that sensitive data is protected from hackers and other unauthorized users.

Finally, data design should also consider the usability of the database. This involves designing a user interface that is instinctive and easy to use. Usability features include search functionality, sorting, and filtering. Search functionality allows users to quickly find the data they need while sorting and filtering allow users to organize the data in meaningful ways.

In closing, data design is a critical component of any software development project. It involves defining the data requirements, designing the database schema, and implementing it. Key principles of data design include apprehending the data requirements, designing for performance and scalability, ensuring data integrity and security, and considering usability. By following these best practices, data designers can create a successful data design that supports the application's functionality and is easy to use.

1.1 Brief introduction about chosen case study

Here I have taken as a case study a data set of 5000 companies located in different states of America including their revenue and growth which is publicly available on the Internet. Here is the name of each company and each industry is mentioned in the data, which sector it belongs to. A workers column is also given here From which we can also estimate the size of the company.

As mentioned above, this is the data of 5000 companies and these companies belong to various industries including IT services and marketing health and government services companies. From this, it can be inferred that every industry may face different problems. But all these companies have one thing in common all these companies are offering services in the United States of America. All of the companies are registered in the US states California, Texas, New York, Georgia, Ohio and Virginia etc. The largest number of companies registered in California is about 600. which means that they are provided with a favourable working environment. While the lowest number of companies include New Mexico, South Carolina, Indiana, Oregon, Connecticut, Maine and Alabama In which the number of companies is less than a hundred. There are many reasons of low number of companies in these states of America. Such as law and order problem, high tax issues and most of all the poor condition of infrastructure.

The U.S. government should immediately take these steps in states with less entrepreneurial companies so that business activity can begin to flourish here as well.

- **The government helps promote business in the form of loans, tax credits, and tax deductions**

- The government helps promote labor through protective laws that set, among other things, minimum wage requirements, maximum work hours, and safe working conditions.
- All of these actions increase productivity, which grows the economy. The government should also optimize its infrastructure so that companies do not face any problem in transportation
- The US government should allocate industrial estates exclusively for business in these estates so that companies are attracted in these states

Task 1.2 Analysis of data collection instruments

One example of complementary data collection instruments that can be used together are surveys and focus groups. **Surveys are a quantitative data collection tool that allows for the collection of large amounts of data from a wide range of respondents. They can be administered online or in person and can be used to gather information about attitudes, opinions, and behaviors.**

On the other hand, focus groups are a qualitative data collection tool that allows for in-depth exploration of a particular topic or issue. Focus groups involve bringing together a small group of participants to discuss a particular topic in a structured and moderated setting. The data collected from focus groups can provide insights into the underlying motivations and reasons behind people's attitudes and behaviors.

By combining survey data with focus group data, researchers can gain a more comprehensive understanding of the topic being studied. The survey data can provide quantitative insights into the prevalence of certain attitudes or behaviors, while the focus group data can provide qualitative insights into the underlying reasons for those attitudes or behaviors.

Another example of complementary data collection instruments are interviews and observation. Interviews are a qualitative data collection tool that involves one-on-one conversations with participants to gather information about their

experiences, attitudes, and opinions. Interviews can be structured or unstructured and can be conducted in person or over the phone.

Observation, on the other hand, is a data collection tool that involves the direct observation of people in their natural environment. Observations can be structured or unstructured and can be used to gather information about behaviors, interactions, and other aspects of human behavior.

By combining interview data with observation data, researchers can gain a more comprehensive understanding of the topic being studied. The interview data can provide insights into the experiences, attitudes, and opinions of the participants, while the observation data can provide insights into their actual behaviors and interactions in real-life situations.

Surveys and focus groups are two commonly used data collection methods in business research, and both have their strengths and weaknesses. In order to evaluate their relevance and effectiveness for addressing business questions, it's important to consider the specific research objectives and the nature of the data being collected.

Advantages of Survey Data Collection

- Advantages of surveys is their ability to provide statistically valid data that can be analyzed using statistical software.
- This can be especially useful when the research question requires a large sample size or when the data needs to be analyzed at a high level of detail.

Limitations of Survey Data Collection

Surveys also have some limitations.

- Respondents may not always provide accurate or truthful responses, and there is a risk of response bias, where respondents provide answers that they think are socially desirable or that they think the researcher wants to hear.

- Surveys may not be suitable for exploring complex or nuanced issues, as they tend to rely on predetermined response categories and may not allow for in-depth exploration of individual perspectives or experiences.

Advantages of Focus Data Collection

- Focus groups can provide rich, detailed data that can be used to generate hypotheses or develop a deeper understanding of the research question.
- Clarify and test preconceived notions and findings.

Limitations Of Focus Data Collection

- Focus groups also have some limitations. Because the discussions are guided by a moderator, there is a risk that the data collected may be influenced by the moderator's own biases or perspectives.
- Focus groups may not be suitable for exploring issues that are sensitive or personal, as participants may not feel comfortable sharing their thoughts and experiences in a group setting.

Generally, both surveys and focus groups can be effective tools for addressing business questions, but their appropriateness will depend on the exact research goals and the nature of the data being collected. Surveys may be more appropriate for gathering quantitative data from a large sample size, while focus groups may be more suitable for exploring complex or nuanced issues in greater depth. Ultimately, the choice of data collection method will depend on a range of factors, including the research objectives, the resources available, and the preferences of the research team.