

10 Lecture - CS101

Important Subjective

1. **What is data manipulation, and why is it important?**

Answer: Data manipulation is the process of transforming and preparing data to make it more suitable for analysis or visualization. It involves cleaning, transforming, and aggregating data. It is important because raw data is often messy and inconsistent, making it difficult to analyze. Data manipulation helps to clean and transform data to make it more usable and accurate for analysis.

2. **What are the common tools used for data manipulation?**

Answer: Microsoft Excel, SQL, and Python are some of the common tools used for data manipulation.

3. **What is data cleaning, and what are its objectives?**

Answer: Data cleaning is the process of identifying and correcting errors and inconsistencies in data. The objectives of data cleaning are to improve the quality of the data, reduce errors and inconsistencies, and prepare the data for further analysis.

4. **What are the common techniques used for data transformation?**

Answer: Common techniques for data transformation include merging, filtering, sorting, and aggregating.

5. **What is the difference between data cleaning and data transformation?**

Answer: Data cleaning is the process of identifying and correcting errors and inconsistencies in data, while data transformation involves converting data from one format to another.

6. **What is the purpose of data wrangling in data manipulation?**

Answer: Data wrangling is the process of cleaning and transforming data to make it more suitable for analysis. The purpose of data wrangling is to prepare the data for analysis by cleaning, transforming, and aggregating it.

7. **What is data aggregation, and what are its common techniques?**

Answer: Data aggregation is the process of summarizing data by calculating totals or averages. Common techniques for data aggregation include grouping, sub-setting, and summarizing.

8. **What are the common types of errors in data, and how can they be corrected?**

Answer: Common types of errors in data include missing values, duplicates, and inconsistencies. They can be corrected by identifying the errors, replacing missing values, removing duplicates, and standardizing data.

9. **What is data merging, and how is it useful in data manipulation?**

Answer: Data merging is the process of combining data from multiple sources based on a common variable. It is useful in data manipulation because it allows us to combine data from different sources to create a more complete dataset.

10. **What are the common challenges faced in data manipulation?**

Answer: Common challenges in data manipulation include dealing with missing data, handling errors and inconsistencies, and choosing the appropriate tools and techniques for the data.