Genesis Grant

CTEC 128

Big Data and Bias Assignment

**2.2.2. Big Data and Bias**1. Analyze Data

2. How many hours do you typically sleep per night?

3. When you party on the weekends, where do you like to go?

**2.2.6 Cognitive Bias**

Confirmation Bias:

Hailee is doing a project on global warming and its effects on the planet. Instead of looking for sources from both perspectives, she researches only articles and papers that agree with her stance and implement only those sources into her project.

Sampling Bias:

A survey of student food options and preferences is given only to cafetoria employees instead of any other population within the school.

Contextual Bias: Depending on your credit history, residential area, age, etc. banks and credit companies can decide whether they want to lend you money. Certain individuals will have better or worse chances based on the context of themselves.

**2.2.8 Decomposing the Problem**

1. Columns needed to answer statistical questions would include defining facts or characteristics. Maybe the financial earnings of a family, region they live, credit score (if applicable), etc. These data columns can be compared with other varying stats to come up with conclusions. Columns not needed would include specific information about individuals like names, id numbers, address, etc. This information is exclusive to each person, so this data would irrelevant.

2. The calculations that could be used in this project include ranges of data to highlight changes throughout time, we could use averages to compare different numerical data, or ttransfer number data to charts to visualize significant changes.

3. In the data analysis process, the first step is data extraction, where relevant columns are identified and extracted from the dataset. Next, data cleaning is performed to handle missing values and outliers, ensuring that the dataset is accurate and complete. Following this, statistical analysis is conducted to calculate percentages, averages, and correlations, providing insights into the relationships within the data. Finally, visualization techniques such as charts and graphs are used to present the findings in a clear and understandable manner, helping to communicate the results of the analysis effective.