Neethu Raj S

Assignment

1. What is the difference between data analysis and machine learning?

Data Analysis is a process of understanding the data, find patterns and try to obtain inferences due to which the underlying patterns are observed. Machine Learning is when you train a system to learn those patterns and try to predict the upcoming pattern. [Machine learning](https://www.simplilearn.com/tutorials/machine-learning-tutorial/what-is-machine-learning) can be defined as the practice of using algorithms to extract data, learn from it, and then forecast future trends for that topic. Traditional machine learning software is statistical analysis and predictive analysis that is used to spot patterns and catch hidden insights based on perceived data.

A good example of machine learning implementation is Facebook. Facebook’s [machine learning algorithms](https://www.simplilearn.com/10-algorithms-machine-learning-engineers-need-to-know-article) gather behavioral information for every user on the social platform. Based on one’s past behavior, the algorithm predicts interests and recommends articles and notifications on the news feed. Similarly, when Amazon recommends products, or when Netflix recommends movies based on past behaviors, machine learning is at work.

But "data analysis" has different aims and goals in different sectors — and is done by people with very different kinds of formal training — different kinds of "data analysts" will use different software tools and different methods. Deep learning and other predictive models will be appropriate for some kinds of data analysis (web- and app-based companies; certain data-cleaning purposes in scientific research)

2. What is big data?

Big data is data that contains greater variety, arriving in increasing volumes and with more velocity. Simply, big data is larger, more complex data sets, especially from new data sources. These data sets are so voluminous that traditional data processing software just can’t manage them. But these massive volumes of data can be used to address business problems one wouldn’t have been able to tackle before.

Big data analytics helps organizations harness their data and use it to identify new opportunities. That, in turn, leads to smarter business moves, more efficient operations, higher profits and happier customers.

Big data is a collection of data from many different sources and is often describe by five characteristics:

* Volume(Huge amount of data)
* Value(Valuable data)
* Variety(Nature of data & heterogeneous sources)
* velocity(High speed of accumulation of data)
* veracity(Inconsistencies and uncertainty in data)

3. What are the 3 main things we should know before studying data analysis?

* **High school Math is fundamental for Data analysis.**
* **Your critical and analytical skills are very important.**
* **Presentation skills**

4.Most common characteristics used in descriptive statistics?

Descriptive statistics summarizes or describes the characteristics of a data set. Descriptive statistics consists of two basic categories of measures: measures of central tendency and measures of variability (or spread). Measures of central tendency describe the center of a data set. Measures of variability or spread describe the dispersion of data within the set.

5. What is quantitative and qualitative data?

Quantitative data are measures of values or counts and are expressed as numbers. They are data about numeric variables (e.g. how many; how much; or how often).  
Qualitative data are measures of 'types' and may be represented by a name, symbol, or a number code. They are data about categorical variables (e.g. what type).

Generally speaking, quantitative analysis involves looking at the hard data, the actual numbers. Qualitative analysis is less tangible. It concerns subjective characteristics and opinions – things that cannot be expressed as a number. Quantitative Information, involves a measurable quantity—numbers are used. Some examples are length, mass, temperature, and time. Quantitative information is often called data, but can also be things other than numbers. Gender, country name, animal species, and emotional state are examples of qualitative information.

Quantitative and qualitative data provide different outcomes, and are often used together to get a full picture of a population. For example, if data are collected on annual income (quantitative), occupation data (qualitative) could also be gathered to get more detail on the average annual income for each type of occupation.