This data represents an important asset which the store can use to understand and

improve their business activities and performance.

To take advantage of this data, they need to perform different types of

data analysis and measure their data appropriately.

Let's find out how global superstore can make the most of their data.

And start with a recap of data analytics and

the types of data analytics they can use.

As you should know from previous courses,

data analytics involves analyzing data to derive useful information in valuable insights.

.

There are several key types of data analysis that you've encountered so

far and made use of in other courses.

Let's briefly recap these,

1. descriptive data analysis presents data in a descriptive format.
2. Exploratory data analysis is used to establish a relationship between different variables, and
3. inferential data analysis focuses on a small sample of data to make inferences.
4. Predictive data analysis identifies patterns and data to make predictions about future performance. And
5. causal data analysis explores cause and effect between variables.

Another key question to ask of your data is if it's quantitative or

qualitative,

1. quantitative data refers to numerical data. This is data that can be counted or quantified.

In the case of Global Superstore, this includes the average number of customers who make purchases each day, or the average cost of each purchase made.

1. Qualitative day refers to non-numerical data. This is textual and descriptive data like information about the quality attributes of a product. For example, Global superstores qualitative data includes category names or descriptions of products like furniture or office supplies.

Once you've determined what kind of data you're dealing with,

you then need to organize, identify, and analyze your data.

You can perform these actions using four different measurement scales.

The first of these measurement scales is the

1. nominal scale. This scale describes the property of non-numerical data. It's purely descriptive, which means it just identifies the data. In the case of Global Superstore, they can use this scale to identify products in their stock like a chair or a desk. Each product is one nominal unit of data.
2. ordinal scale. This is a qualitative data type scale which places data in a specific ranked order. However, it doesn't include decisive criteria to determine the difference between the data elements. For example, Global Superstore can rank chairs using ratings values, so they can use a value of one for top quality products, two for very good products, and three for good products and so on. However, there's no precise criteria that determines the measurement between each value.
3. interval scale. This scale includes properties of the nominal and order data scales. Its key feature is that the difference between data points can be clearly identified using specific criteria. The scale can also contain both positive and negative numbers, and zero does not represent an absolute true value. Global superstore can use the interval scale to provide feedback on products from 10 to -10.
4. ratio scale. This scale is a quantitative data type that includes properties from nominal,

ordinal, and interval scales of measurement. It defines the identity of the data classifies the data in order and marks clear intervals. However, it holds an absolute value of zero. Over at global superstore, they can use the ratio scale to mark the weight of products. For example, a small table is 20 kg, a medium sized table is 40 kg while a larger table weighs a total of 60 kg.

In this instance, there's a clear order between variables and an equal distance of 20 kg between each measurement. So, all data points can be measured accurately. You should now be familiar with the basics of data analytics.

 two types of data

you will deal with: Quantitative data,

which refers to numerical data and qualitative data,

which refers to non-numerical data.

**Benefits of data analytics in business**

Over the past decade, data analytics has played a major role in determining how organizations assess risks and seek opportunities. Large corporations and organizations such as Amazon, Microsoft and Apple have invested billions of dollars in advanced data analytics.

It is important to realize that the ways in which data analytics can be utilized relies on many factors such as the requirements of the business, market conditions, competitors and the domain. Nevertheless, there are generic benefits that can be gained from data analytics regardless of the business size and type.

For example, data analytics can be used to:

* Provide refined, detailed and reliable insights that enhance day-to-day business decision making.
* Identify and exploit potential opportunities.
* Discover ways to monetize data.
* Identify threats such as security gaps and breaches.
* Predict and proactively respond to changes and challenges in the market.
* Obtain a panoramic view of the business activities and operations.
* Prevent fraud.
* And reveal new areas in which you can reduce business expenses or cost and increase profits.

It’s also important to remember that nowadays your business is part of the Big Data environment.  This is an environment which includes massive amounts of data, and many kinds of complex data, including structured, semi structured and unstructured data sets. This data comes from a wide variety of data sources such as OLTP databases, commercial data, browsing histories, search engines and social media sites.

For example, there are more than 500 terabytes of data generated every day by Facebook posts including images, comments and videos. Similarly, there are over 20 petabytes of data generated by Google processes every day across its massive computing clusters. These statistics highlight the importance of using artificial intelligence, machine learning and data mining to help businesses make data-informed decisions.

**Data analytics techniques**

There are many techniques that can be used in data analytics to maximize the value of data insights. Some commonly used techniques with relevant examples are listed below.

**Classification technique**

The classification technique assigns data items into categories or classes, which allows you to predict the target class of your data. For example, banks use this technique to identify loan applicants with low, medium, or high credit risks based on their historical credit ratings, employment history, property ownership and so on.

**Association rule**

The association rule identifies the relationships between data elements to determine whether there is a correlation between data. A classic example of the association rule is identifying the relationship between two products in a store. For example, many customers who go to the store to buy a specific product might buy another related product. So, the data shows that there’s a relationship between these two products. This association might be based on data that shows 70% of the total number of the store’s transactions include both products. This data indicates that there is an association link between the two products.

**Outlier detection**

Outlier detection is used to uncover unusual data within a data set. A common example in this analytics model is credit card fraud detection. This technique can be used to check whether the incoming transaction fits with the customer’s profile and behavior. If a request doesn’t fit a customer’s legacy profile and behavior, then it is flagged as an outlier or anomaly detection.

**Clustering analysis**

Clustering analysis looks for similarities within a set of data. If it finds similarities, then it separates related data into clusters of subsets based on their common characteristics.

A good example of clustering analysis is online streaming services. These services use clustering analysis to identify and classify viewers engaged in similar types of streaming activity. The streaming services look at the amount of time a user spent watching videos, and the number of sessions they engaged in per day, week and month.

These metrics can be used within a cluster analysis to identify client habits. The services then know which demographics they should target with their advertisements and the best ways in which to reach them.