Digital Fluency

Preamble

After looking at the market needs, the Digital Fluency course is designed to bring you closer to fulfilling the scale gap. The learning from this course will help you gain competitive advantage and to showcase your conceptual understanding of some of the most in-demand technologies like AI, BDA and IoT. At the same time, demonstrate that you are equally focused on building essential soft skills, which are much needed for professional success.

1. Context

"I Voted" – sticker/badge by Facebook

Facebook successfully tied the political activity to user engagement when they came out with a social experiment by creating a sticker allowing its users to declare "I Voted" on their profiles.

This experiment ran during the 2010 midterm elections and seemed useful. Users who noticed the button were likely to vote and be vocal about the behavior of voting once they saw their friends were participating in it. Out of a total of 61 million users, then, 20% of the users who saw their friends voting, also clicked the sticker. The Data science unit at Facebook has claimed that with the combination of their stickers that motivated close to 60,000 voters directly, and the social contagion, which prompted 280,000 connected users to vote for a total of 340,000 additional voters in the midterm elections.

For the 2016 elections, Facebook expanded its involvement into the voting process with reminders and directions to users' polling places.
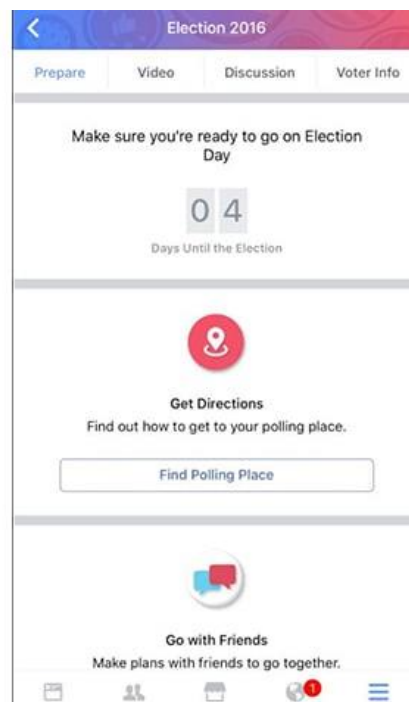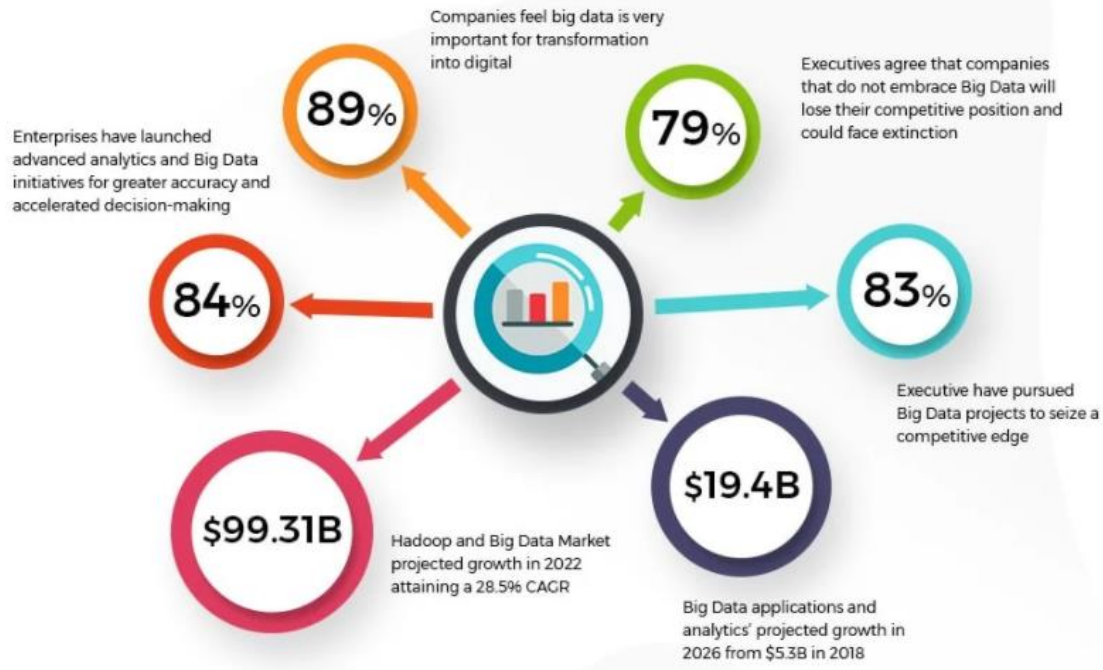
2. Key Concepts

Big Data Analytics

After completing this pathway, you will:

- Understand the importance of Big Data Analytics (BDA) in different fields
- Get an overview of the concepts of BDA
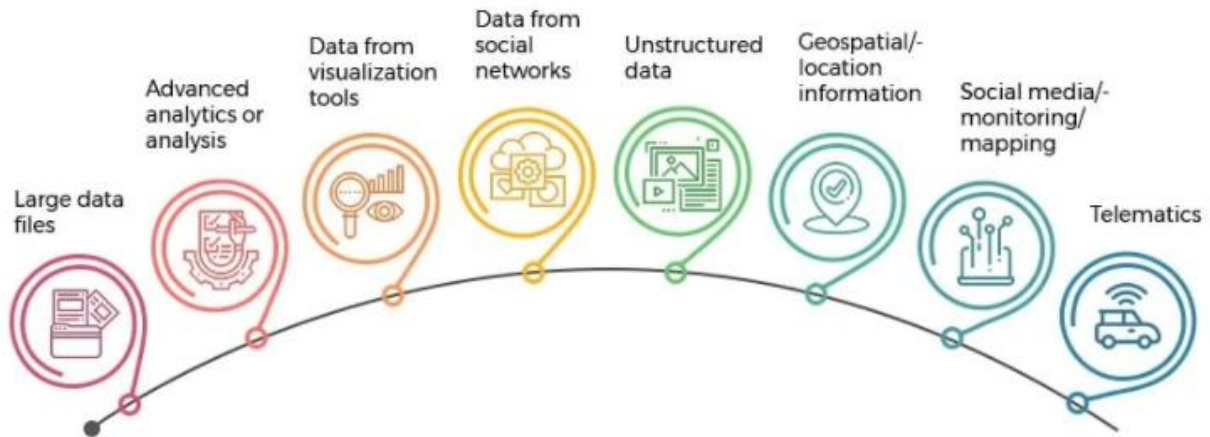- Learn how to perform data analysis in Excel using pivot tables and pivot charts

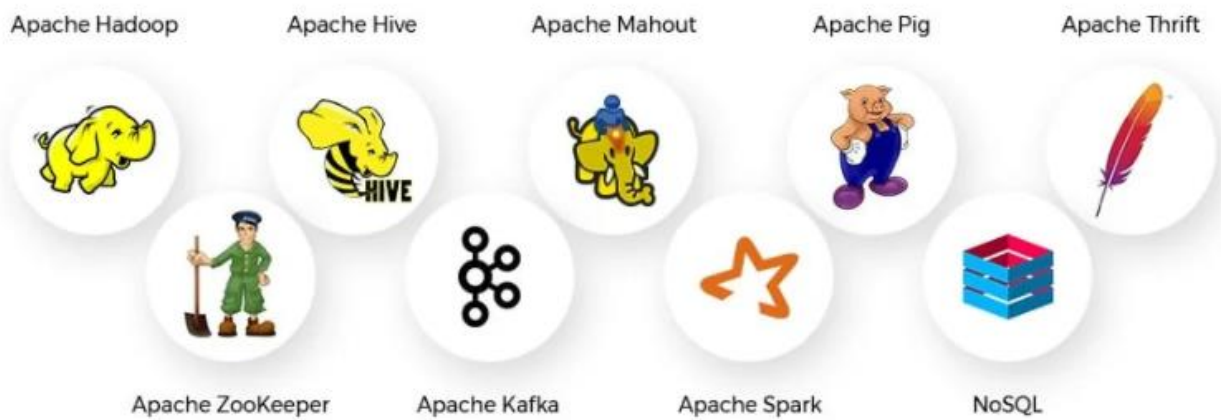### Big Data Analytics – Introduction
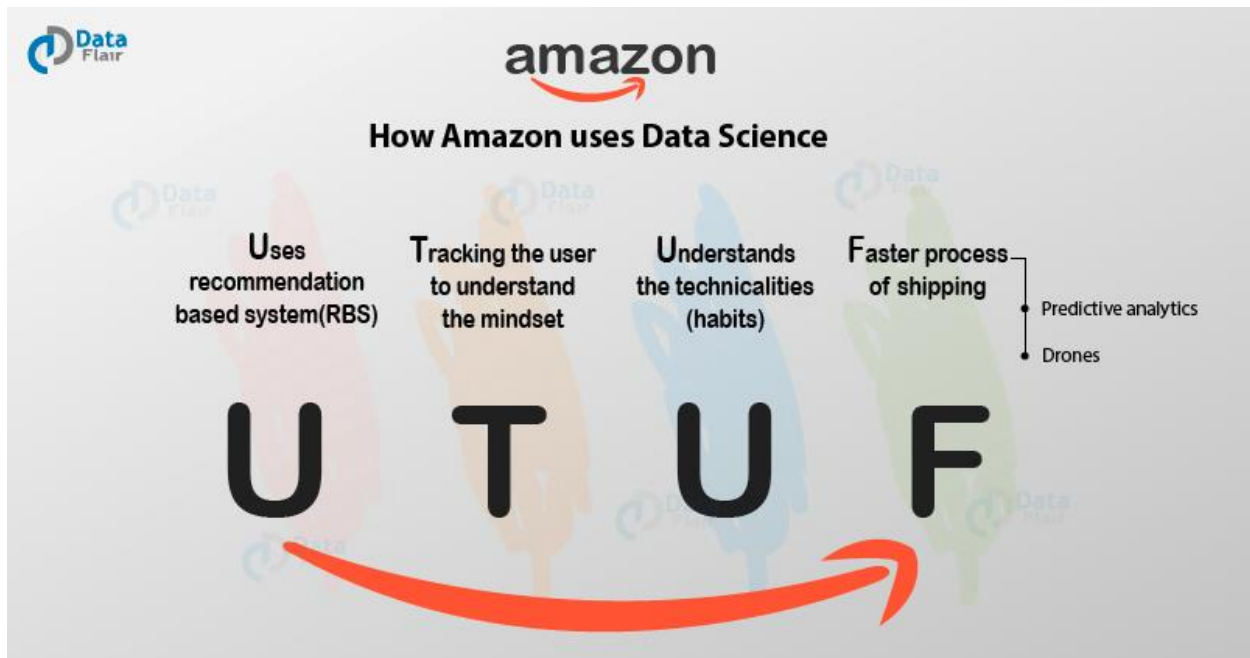

Data Analytics Process

Sources of Big Data


Big Data Tools and Technologies

3. Case Study

How Amazon uses Data Science

Uses recommendation-based system (RBS)

Through this technology, it gathers data from their customers (Can also be called Big Data). The more data they have the better it is for them because once they understand what the user wants, they then streamline the process and try to encourage the customers to purchase the products. RBS seeks and predicts the "rating" or "preference" a user would give to an item.
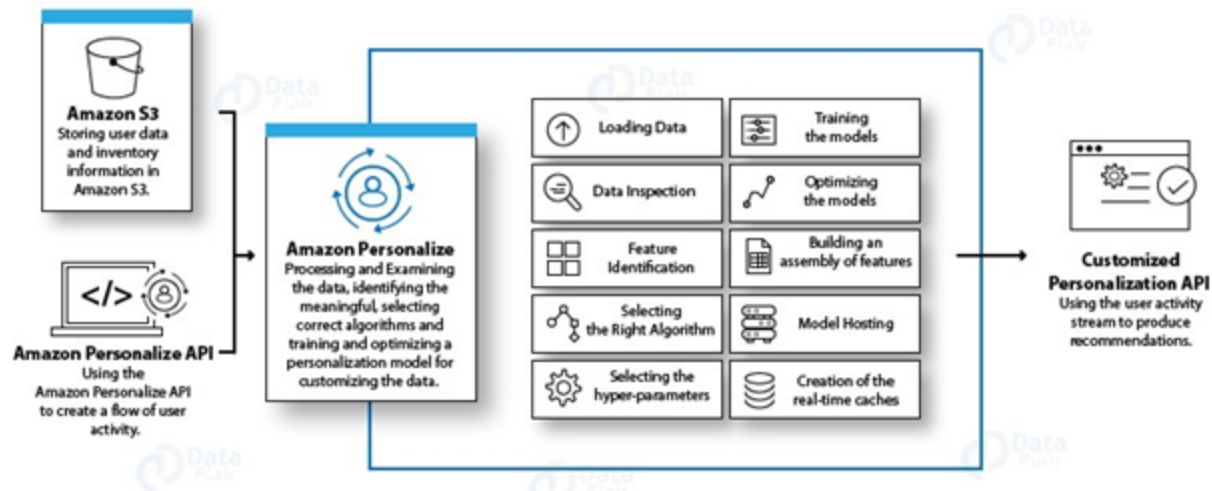
Database Management for Data Science

After completing this Pathway, you will:

1. Explain Database and its different types
2. Advantages of Database

Introduction to Data

Data are individual facts, statistics, or items of information, often numeric, that are collected through observation. In a more technical sense, data are a set of values of qualitative or quantitative variables about one or more persons or objects, while a datum is a single value of a single variable.
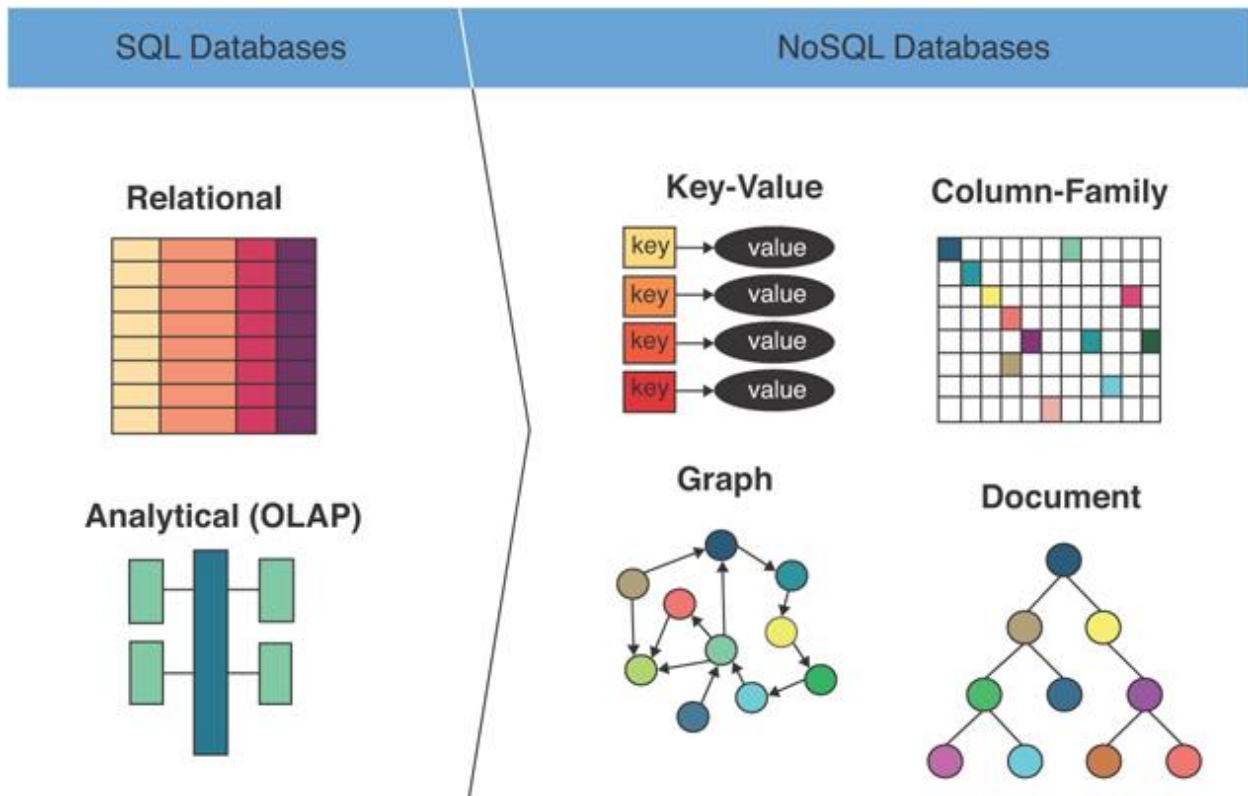
Definition of Database

A database is defined as a structured set of data held in a computer's memory or on the cloud that is accessible in various ways.
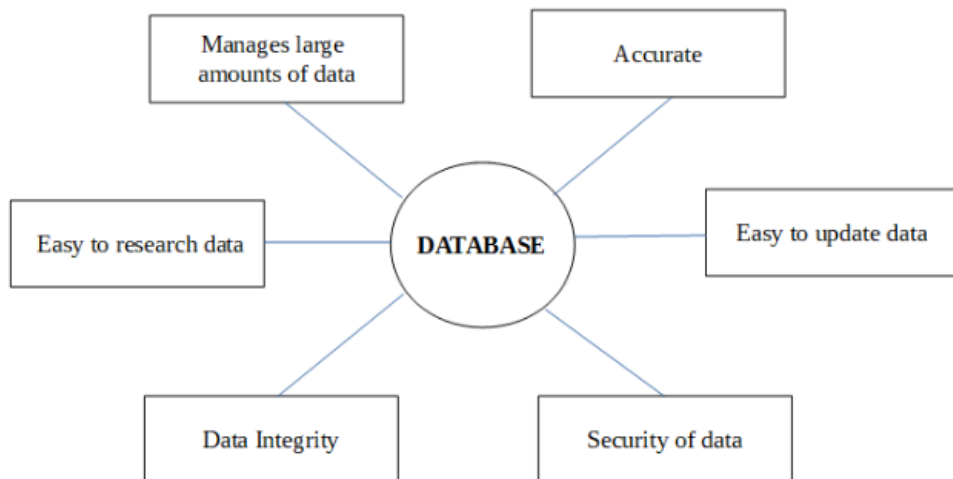
Database Management Systems (DBMS) refer to the technology solution used to optimize and manage the storage and retrieval of data from databases.

Types of database
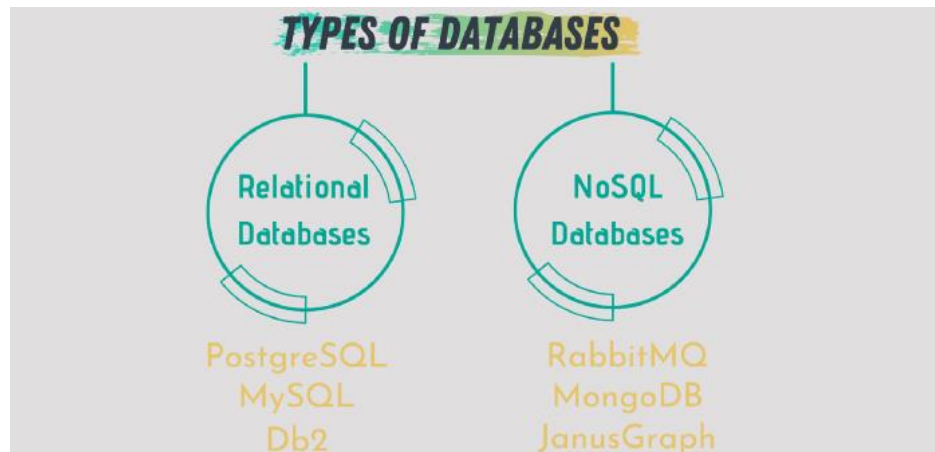
Why use databases (Advantages)?

4. Lab Session

## Visualization of data using Excel/spreadsheet:

A Pivot Table is a powerful tool to calculate, summarize, and analyze data that lets you see comparisons, patterns, and trends in your data. PivotTables work a little bit differently depending on what platform you are using to run Excel.

Sometimes it's hard to see the big picture when your raw data hasn't been summarized. Your first instinct may be to create a PivotTable, but not everyone can look at numbers in a table and quickly see what's going on. Pivot Charts are a great way to add data visualizations to your data.
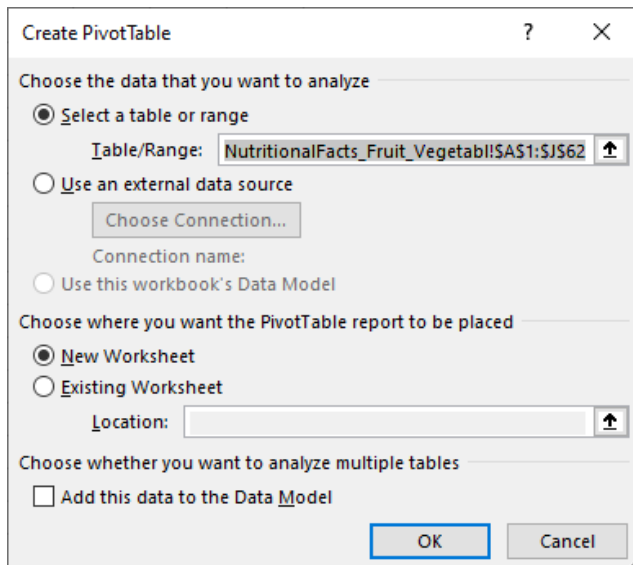
Data set:

Let us take a data set of nutritional facts of fruits, vegetables and seafood as an example and build a Pivot Table and Pivot Chart.

1. Download the data set.
2. Select a cell in the Excel.
3. Click Insert → PivotChart → PivotChart & PivotTable.
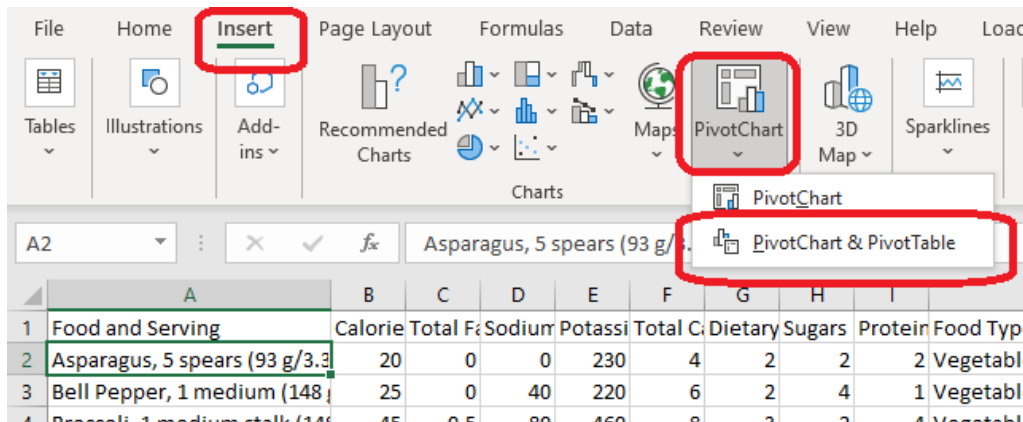
4. Under **Choose the data that you want to analyze**, select **Select a table or range**.

5. In **Table/Range**, verify the cell range.
6. Under **Choose where you want the PivotTable report to be placed**, select **New worksheet** to place the PivotTable in a new worksheet or **Existing worksheet** and then select the location you want the PivotTable to appear.
7. Select **OK**.

   New sheet gets created and **PivotChart Fields** pane is displayed.
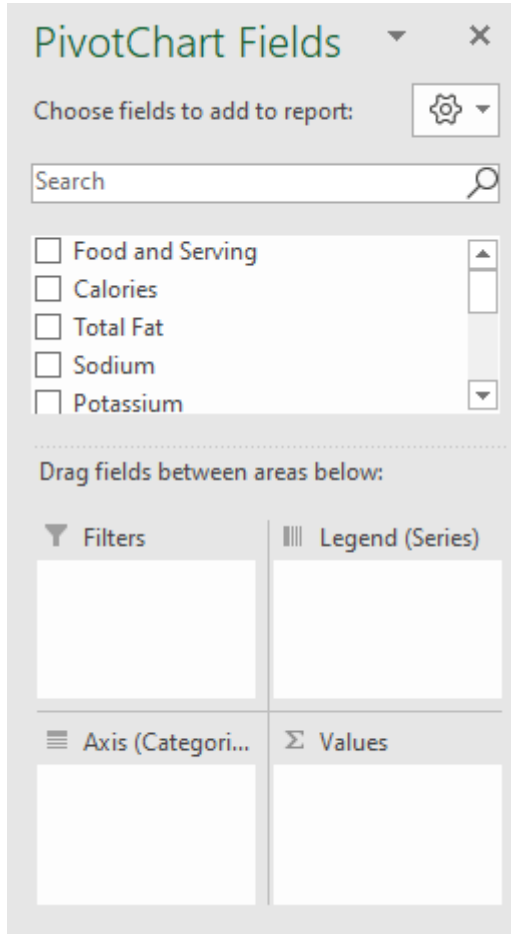
8. Drag and drop the required fields as appropriate. In this example, drag and drop:
   a. **Food Types** to the **Filters**
   b. **Food and Serving** to the **Rows**
   c. **Calories, Total Fat, Sugars,** and **Proteins** to the **Values**.
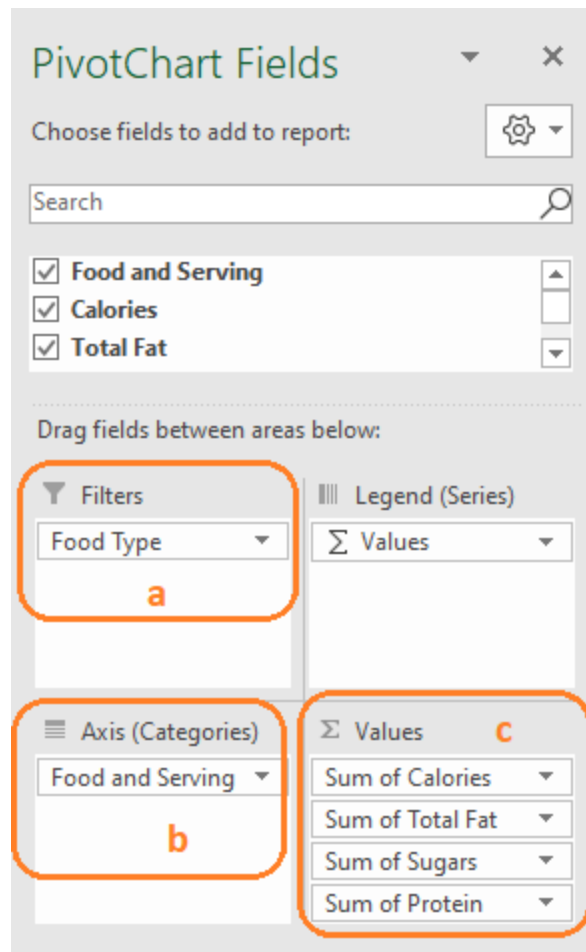
Note: **Values** gets automatically added to **Legend (Series)**.

9. You can see the Pivot Table and Pivot Chart as shown below.
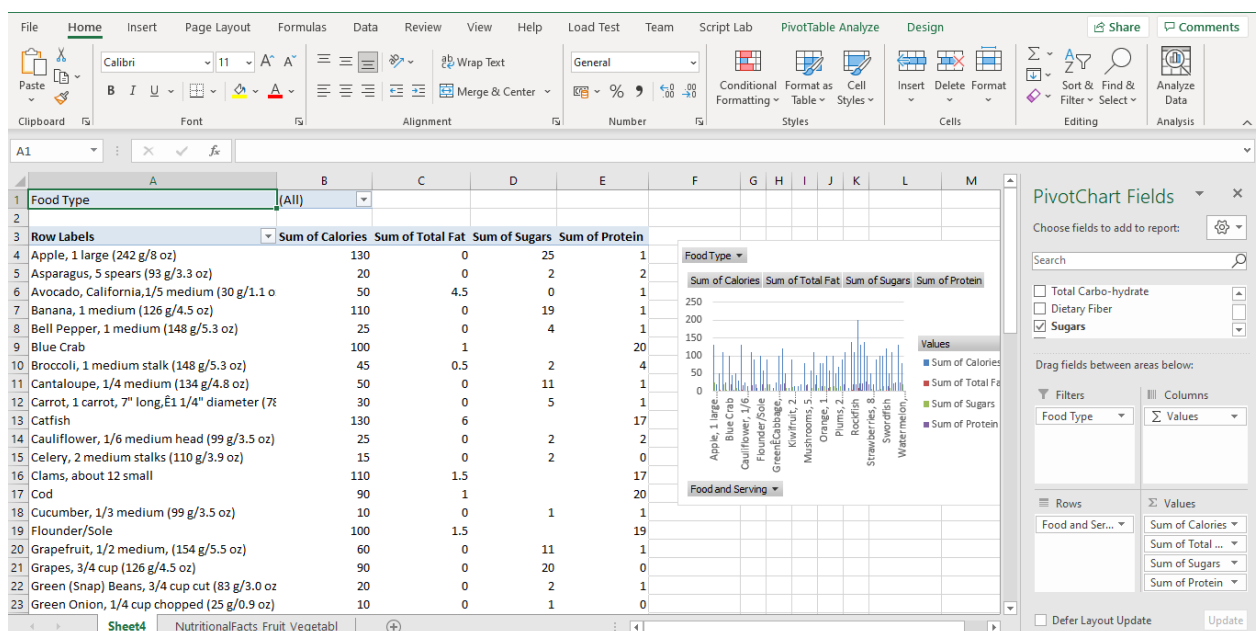
Note: Appropriate filter can be selected under **Food Types** to visualize the different types of food.

5. Frequently Asked Questions

1. What is data science?

Data science is an interdisciplinary field that uses scientific methods, processes, algorithms and systems to extract knowledge and insights from noisy, structured and unstructured data, and apply knowledge and actionable insights from data across a broad range of application domains.

2. What is the need for Data Science?

The reason why we need data science is the ability to process and interpret data. This enables companies to make informed decisions around growth, optimization, and performance. Demand for skilled data scientists is on the rise now and in the next decade.

3. What is Data Science useful for?

Data science is a process that empowers better business decision-making through interpreting, modeling, and deployment. This helps in visualizing data that is understandable for business stakeholders to build future roadmaps and trajectories. Implementing Data Science for businesses is now a mandate for any business looking to grow.

4. How Facebook Uses Data Analytics To Understand Your Posts?

With 1.2 billion people uploading 136,000 photos and updating their status 293,000 times per minute on Facebook, it contributes to unstructured data (information which isn't easily quantified and put into rows and tables for computer analysis).

Textual analysis - A large proportion of the data shared on Facebook is still text. Facebook uses a tool it developed itself called DeepText to extract meaning from words we post by learning to analyze them contextually. Neural networks analyze the relationship between words to understand how their meaning changes depending on other words around them. It learns for itself based on how words are used. It can easily switch between working across different human languages and apply what it has learned from one to another.It can easily switch between working across different human languages and apply what it has learned from one to another.

5. How Facebook Uses Data Analytics To Understand Your Posts And Recognize Your Face?

Facial recognition - Facebook uses a DL application called DeepFace to teach it to recognize people in photos. It says that its most advanced image recognition tool is more successful than humans in recognizing whether two different images are of the same person or not – with DeepFace scoring a 97% success rate compared to humans with 96%.

6. What is public health analytics?

Public health analytics is the process of obtaining, reviewing, and managing health-related data of entire populations, typically carried out by public health organizations in an effort to maintain population health. Public health data may include mortality reports, demographic data, socioeconomic data, procedural and diagnostic data, and medical claims data, among others. Public health organizations may use analytics to monitor disease trends and

determine patterns in certain populations, guide the implementation of disease control programs, and set priorities for allocating health resources to populations in need.

7. **Which are the 10 companies that are using big data?**
   a. Amazon
   b. Apple
   c. Spotify
   d. Google
   e. Facebook
   f. Instagram
   g. Netflix
   h. Starbucks
   i. American Express
   j. McDonald's

8. **How You Can Leverage the Power of Data to Grow Sales?**

   There is only one place to begin, and this is with data analytics that pertains to your audience. This will give you in-depth insights regarding the demographics of your audience, i.e. their gender, age, income, occupation, where they are based and what language they speak. In addition to this, audience data should inform you of the different devices that your audience is using. Do they mainly access your store from their mobile phone or desktop? If it is the former, do they tend to use Android or Apple devices? This can give you great insights into how your online store is being accessed so that you can target your efforts accordingly.

9. **Where does big data come from?**

   Big data is often boiled down to a few varieties including social data, machine data, and transactional data. Social media data is providing remarkable insights to companies on consumer behavior and sentiment that can be integrated with CRM data for analysis, with 230 million tweets posted on Twitter per day, 2.7 billion Likes and comments added to Facebook every day, and 60 hours of video uploaded to YouTube every minute. Machine data consists of information generated from industrial equipment, real-time data from sensors that track parts and monitor machinery (often also called the Internet of Things).

10. **Who are some of the BIG DATA users?**

    From cloud companies like Amazon to healthcare companies to financial firms, it seems as if everyone is developing a strategy to use big data. For example, every mobile phone user has a monthly bill which catalogs every call and every text; processing the sheer volume of that data can be challenging. The size of Big Data can be relative to the size of the enterprise. For some, it may be hundreds of gigabytes, for others, tens or hundreds of terabytes to cause consideration.

11. **How Data Science differs from Big Data and Data Analytics?**

    Data Science is a field which contains various tools and algorithms for gaining useful insights from raw data. It involves various methods for data modelling and other data related tasks such as data cleansing, preprocessing, analysis, etc. Big Data implies the enormous amount of data which can be structured, unstructured and semi-structured generated through various channels and organisations. The tasks of Data Analytics involve providing

operational insights into complex business situations. This also predicts the upcoming opportunities which the organisation can exploit.

12. How does big data help in decision making for the organization?

Research backs these claims, with studies showing businesses that use big data for making decisions realize up to a 10% increase in profits and a 10% decrease in overall costs. Data can be used to make financial, growth-related, marketing and sales, and customer service decisions that drive your business forward.

13. What is "big data for small business"?

Many small businesses believe they are too small for big data. This is far from the truth as small businesses need big data to succeed, just as much as larger corporations. Data provides businesses with actionable insights needed to become more efficient and profitable.

14. What are Big Data Tools and Software?

    a. Hadoop.
    b. Quoble.
    c. Cassandra. ...
    d. MongoDB. ...
    e. Apache Storm. ...
    f. CouchDB. ...
    g. Statwing

15. How Big Data is changing the Way People Live Their Lives?

The changes in how big data is collected have occurred so rapidly that big data is more prevalent in daily life than you might think. Companies and organizations are collecting information about their targeted audiences. They know what you're watching, what you're reading, and what you're buying.This access to key, personalized data then affects your daily experience in some of the most important and common areas of life. Consider these ways big data is used in your everyday life:

    a. Music, Shows, and Movies
    b. Healthcare and medical services
    c. Shopping and Marketing
    d. Travel and Transportation
    e. News and Information
    f. Education and Employment

Reference:

1. Digital 101 Course offered by Future Skill Prime Platform
2. Why do we need a Database (tutorialspoint.com)
3. Relational Versus Nonrelational Databases | Storing Data in AWS | Pearson IT Certification
4. Overview of PivotTables and PivotCharts (microsoft.com)
5. Introduction to Pivot Tables, Charts, and Dashboards in Excel (Part 1) - YouTube

6. MS Excel - Pivot Table Example 1 Video Tutorials - YouTube
7. MS Excel - Pivot Chart - YouTube