NEW TERMINOLOGY

DVCS => distributed version control system

CVCS => centralized version control system

SHA => secure hash algorithm

GIT => an example of DVCS (Version control is a software that helps a team manage changes to its source)

GITHUB => is an online space to work collaboratively on projects and keep code safe.

VBA => stands for Visual Basics for Applications. It is a tool for programming, editing and running application code by application (i.e. Excel)

GUI => graphical user interface

Modeling=> data modeling is the process of producing a descriptive diagram of relationships between various types of information that are to be stored in a database. The ability to think clearly and systematically about the key data points to be stored and retrieved, and how they should be grouped and related, is what the data modeling component of data science is all about.

Forecasting=> forecasting is to predict or estimate (a future event or trend). For businesses and analysts forecasting is determining what is going to happen in the future by analyzing what happen in the past and what is going on now. The predictability of an event or a quantity depends on several factors including: - how well we understand the factors that contribute to it; - how much data are available; - whether the forecast can affect the thing we are trying to forecast.

VS Code => stands for Visual Studio Code. It is a source-code editor developed by Microsoft for Windows, Linux and macOS.

Python => a programing language

APIs => stands for Application Programming Interface. It is a set of routines, protocols, and tools for building software applications. Basically, an API specifies how software components should interact. Additionally, APIs are used when programming GUI components. A good APIs makes it easier to develop a program by providing all the building blocks. A programmer then puts blocks together.

RESTful APIs => A RESTful API breaks down a transaction to create a series of small modules. Each module addresses a particular underlying part of the transaction.

Web scraping => Web scraping refers to the extraction of data from any website.

Data mining => Data mining refers to the process of advance analysis of extensive data sets.

JSON => stands for JavaScript Object Notation. It is a lightweight format for storing and transporting data. It is often used when data is sent from a server to a web page (self-describing and easy to understand)

NumPy => stands for Numerical Python. It is a Python's library adding support for large, multi-dimensional arrays and matrices, along with a large collection of high-level mathematical functions to operate on these arrays.

Pandas => stands for Python Data Analysis Library. It is a Python's library for data manipulation and analysis.

CSV file => Comma Separated Values

TSV file=> Tabulator Separated Values

Matplotlib => is a plotting library for the Python programming language and its numerical mathematics extension NumPy. It provides an object-oriented API for embedding plots into applications using general-purpose GUI toolkits.

Beautiful Soup => Beautiful Soup is a Python package for parsing HTML and XML documents (including having malformed markup, i.e. non-closed tags, so named after tag soup). It creates a parse tree for parsed pages that can be used to extract data from HTML, which is useful for web scraping.

SQL => stands for Structures Query Language and it is used to communicate with a database.

PostgreSQL => is Relational Database Management System that uses and extends the SQL language combined with many features that safely store and scale the most complicated data workloads.

MongoDB => is a document-oriented NoSQL database used for high volume data storage. Instead of using tables and rows as in the traditional relational databases, MongoDB makes use of collections and documents.

ETL => is an IT process and it stands for Extract, Transform, Load, three database functions that are combined into one tool to pull data out of one database and place into another.

HTML => stands for Hypertext Markup Language and it is a standard markup language for documents designed to be displayed in a web browser. It allows the user to create and structure sections, paragraphs, headings, links and blockquotes for web pages and applications. It is not a programing language (meaning it doesn't have the ability to create dynamic functionality).

CSS => stands for Cascading Style Sheet (same as HTML it is technically not a programing language - they are page structure and style information). CSS are used to format the layout of web pages.

Bootstrap => is an open source JavaScript framework developed by the team at Twiter, it is a combination of HTML, CSS and JavaScript code designed to help build user interface components. It is also called Front-End framework. It contains HTML and CSS based design templates for typography, forms, buttons navigation, and other interface components, as well as optional JavaScript extensions.

Dashboarding=> A data dashboard is an information management tool that visually tracks, analyzes and displays key performance indicators (KPI), metrics and key data points to monitor the health of a business, department or specific process. They are customizable to meet the specific needs of a department and company.

JavaScript => is a programing language commonly used in web development, often abbreviate as .js. It is a scripting language, used to enhance HTML pages and is commonly found embedded in HTML code.

JavaScript is an interpreted language. It renders web pages in an interactive and dynamic fashion. This allows the pages to react to events, exhibit special effects, accept variable text, validate data, create cookies, detect a user's browser, etc.

AJAX => stands for Asynchronous JavaScript and XML. It is not a programing language or a tool, but a concept. It is a client-side scripts that communicates to and from a server/database without the need for a postback or a complete page refresh. It is a technique for creating fast and dynamic web pages. AJAX allows web pages to be updated asynchronously by exchanging small amounts of data with the server behind the scene (this means that it is possible to update parts of a web page, without reloading the whole page).

XML => stands for Extensible Markup Language and it is a markup language that defines a set of rules for encoding documents in a form that is both human-readable and machine-readable.

D3 => stands for Data Driven Documents. It is a JavaScript library for producing dynamic, interactive data visualizations in web browsers.

Leaflet => is an open-source JavaScript library for mobile-friendly interactive maps.

Tableau => it is a powerful and fastest growing data visualization tool used in Business Intelligence Industry. It helps in simplifying raw data into the very easily understandable format. Data analysis is very fast with Tableau and visualizations crated are in the form of dashboards and worksheets.

Hadoop => it is a java based open-source framework. Hadoop is used for storing and processing big data. In Hadoop data is stored on inexpensive commodity serves that run as clusters. It is a distributed file system allows concurrent processing and fault tolerance. Hadoop MapReduce programming model is used for faster storage and retrieval of data from its nodes.

Big Data Analysis => Big data is a term that describes the large volume of data – both structured and unstructured – that inundates a business on a day-to-day basis. But it's not the amount the data that's important. It's what organization go with the data that matters. Big data can be analyzed for insights that lead to better decisions and strategic business moves. Big Data as 3-Vs: Volume, Velocity and Variety. https://www.sas.com/en_us/insights/big-data/what-is-big-data.html

Machine Learning => is an application of artificial intelligence (AI) that provides systems the ability to automatically to automatically learn and improve from experience without being explicitly programmed, Machine Learning focuses on the development of computer programs that can access data and use it learn for themselves.