**MICROSOFT EXCEL**

**History**

Microsoft Excel is a spreadsheet software program developed by Microsoft Corporation. It was first released for the Macintosh in 1985 and for the Windows operating system in 1987. The program is a part of the Microsoft Office suite of productivity software and is widely used for data analysis, financial modelling, and business intelligence.

The first version of Excel, called Multiplan, was released in 1982 for the Macintosh, CP/M (Control Program for Microcomputers), and MS-DOS operating systems. It was a basic spreadsheet program that was not very popular. However, Microsoft continued to develop the program and released Excel 2.0 in 1987, which was the first version to run on the Windows operating system. This version of Excel included more advanced features such as graphing, macros, and the ability to handle large amounts of data.

Excel 3.0, released in 1990, was the first version to support the VBA (Visual Basic for Applications) programming language, which allowed users to create custom functions and macros. This version also included support for 3-D charts and a new look and feel. Excel 4.0, released in 1992, added support for the Windows 3.1 operating system and improved the program’s ability to work with large amounts of data.

Excel 5.0, released in 1993, was the first version to support the Windows 95 operating system and included new features such as the ability to create pivot tables and the ability to create and edit charts and graphs directly within the spreadsheet. Excel 95 also introduced the “IntelliSense” feature, which provided context-sensitive help and suggestions as the user typed formulas.

Excel 97, released in 1996, was a major update that included support for the Windows 95 and Windows NT operating systems. It also included new features such as “AutoComplete” which automatically completed formulas as the user typed them, and “AutoCalculate” which automatically recalculated formulas as data was entered. Excel also introduced the “Analysis ToolPak” which provided a set of advanced data analysis tools.

Excel 2000, released in 1999, was the first version to support the Windows 98 and Windows NT operating systems. It included new features such as “Smart Tags,” which provided context-sensitive help and suggestions as the user typed, and “DataPilot,” which allowed users to create pivot tables and charts more easily.

Excel 2002, released in 2001, was the first version to support the Windows XP operating system. It included new features such as “List,” which allowed users to create and manage lists of data, and “AutoFilter,” which allowed users to filter and sort data in a more efficient way.

Excel 2003, released in 2002, was a major update that included new features such as “XML Data Import” which allowed users to import data in XML format, and “Data Validation” which allowed users to set data validation rules to ensure data integrity.

Excel 2007, released in 2006, was the first version to support the Windows Vista operating system and included a new “Ribbon” interface which made it easier to find and use Excel’s features. It also included new charting and data visualization features, improved support for working with large amounts of data, and new collaboration and sharing features.

Excel 2010, released in 2010, introduced new features such as “Sparklines” which allowed users to create small, in-cell charts to show trends in data, and “Slicers” which allowed users to filter data in pivot tables and pivot charts.

Excel 2013, released in 2013, introduced new features such as “Flash Fill” which automatically fills in data based on patterns, and “Quick Analysis” which provided a set of tools for data analysis.

Excel 2016, released in 2015, introduced new features such as “Power Query” which allowed users to easily connect to and import data from external sources, and “Power Pivot” which allowed users to create pivot tables and pivot charts with much larger data sets.

Excel 2019, released in 2018, introduced new features such as “Power BI” which integrated data visualization and business intelligence capabilities into Excel, and “Natural Language Formulas” which allowed users to create formulas using plain language.

Excel for Microsoft 365, which was released in 2020, added new features such as “Ideas” which provided suggestions for data analysis and visualization, “Data Types” which allowed users to define the type of data in a column, and “Dynamic Array Formulas” which allowed users to create formulas that automatically adjust to the size of the data.

**Advantages of Microsoft Excel**

**Easy data input:** MS Excel is frequently used for data entry or analysis since there is no limit to the quantity of information that can be saved in a spreadsheet, making it simple to enter as much data as you wish. In Excel, filtering data is simple and highly useful.

**Data retrieval is easier:** In an Excel spreadsheet, data can be located and retrieved with ease.

**Mathematical formulae that are Simple to Use:** MS Excel's formula feature has made calculations simpler and faster.

**Password protection:** You can protect your Excel files using passwords to secure your data. Moreover, Excel offers security features, protecting your files from malicious attacks and mistakes. At the file level, you can encrypt files using the password. That way, every time users want to access one such file, they must enter the password. At the workbook level, it protects them from unauthorized access and unwelcome modifications. Whereas, protecting Excel files at the worksheet level locks cells so that users cannot change critical data.

**Provides multiple spreadsheets:** You can save multiple spreadsheets in a single Excel file, making it more practical for use.

**Supported by many third-party applications:** One can use this software on android, iOS, Windows, and Mac. There is also an online version of this software that you can use and also collaborate with other users.

**Availability of templates:** There are many templates available on MS Excel. For example, you can use different templates for tables and graphs in it.

**Demerits of Microsoft Excel**

**Difficult to learn:** One limitation of MS Excel is that it can be difficult to learn. This is because there are many different elements, such as formulas, functions, and keyboard shortcuts.

**Slowdowns:** If the data is too huge, it might slow down the Excel program, especially if all the data is included in one file. Attempting to divide the data down into smaller files may result in some of it being lost or missing.

**Difficult to consolidate:** The usage of an Excel spreadsheet to generate reports makes the consolidation process lengthy. End-users must often collect data from various files, summarize it, then send their report.

**Difficult to Troubleshoot:**It is not rare to have interconnected spreadsheet data scattered across multiple folders, offices, workstations, or even geographic locations. Even though the locations of all connected files can be pinpointed, it can take a long time to track the logic of formulas from one related cell to another.

**Power BI**

**History**

The idea of Power BI started in 2006 within the SQL Server Reporting Services (SSRS) Team at Microsoft under the leadership of Amir Netz. The project code named "Gemini" as it was known then harnessed the power of SQL Server Analysis Services (SSAS) and made it become available as an in-memory engine. Project Gemini was later renamed Power Pivot in 2009.

In order to compete with the then visualization capabilities of Tableau, Microsoft released Power View as part of SharePoint in 2012. To better enhance their market space in visualizations, Microsoft released Data Explorer (later renamed Power Query) in early 2013, but like Power Pivot, this was only as an add-in in Microsoft Excel.

To make it easier to access data and bring it into Power Pivot, Microsoft enhanced and improved the capabilities of Data Explorer, and late in 2013 it was renamed as Power Query. Power Query was only available as an add-in in Excel.

To ensure reports built could be easily distributed and stored in a central repository accessible to anyone with access, Microsoft announced the new Power BI Service that could be accessed via PowerBI.com in January 2015. At this time, the word Power BI was just a phrase that was being used to describe the complexity in the Power Pivot / SharePoint solution used at that time. After the January 2015 announcement, Power BI remained in Preview mode until July 2015 when it became generally available to all.

**Advantages of Power BI**

* It has an assortment of visualizations - Power BI’s visualizations are interactive and intuitive. It features drop-and-drag functionality, letting even relatively new users modify the looks of a report. Administrators can purchase and download custom visuals from Microsoft Marketplace which include maps, charts, and graphs.
* It has powerful collaborations- Power BI lets users collaborate in real-time, and with Power BI Service, multiple people can work on the same file at the same time.
* Connectivity- Power BI connects to multiple data sources, including Excel, SQL Server databases, Google Analytics, social media analytics, Azure sources, and cloud-based sources. Power BI can also access multiple Big Data sources.
* Excel integration- Users can upload and view Power BI data in Excel; this is especially beneficial when using tabular reports.
* Robust data capacity- Power BI can handle approximately 100 million rows of data.
* It is ideal for key performance indicators (KPI) and dashboards.

**Demerits of Power BI**

* Not designed for complex tables- Power BI has trouble handling multiple links between tables. Users need to ensure that additional data models have unique fields, or Power BI could end up confusing the tables.
* Difficult to configure visuals- Power BI does not give numerous alternatives to arrange your representations according to your prerequisites. Consequently, clients have restricted choices for what they can change in visuals.
* DAX formula can be rigid- Power BI is built around DAX as an expression language. DAX can be challenging. Concatenating more than two elements requires nesting statements.
* Swarmed user interface- Power BI’s user interface appears busy with icons and options, sometimes blocking the view of dashboards and reports.

**Tableau**

**History**

Tableau Software was founded in 2003 by Chris Stolte, Pat Hanrahan, and Christian Chabot. Chris Stolte was a Ph.D. candidate at Stanford University under Professor Pat Hanrahan, and Christian Chabot was at Stolte's dissertation defense in 2003. Together, the three developed and patented Tableau's foundational technology called VizQL, which visually expresses data by translating drag-and-drop actions into data queries through an intuitive interface.

Tableau’s release v1 (April 2005) connected to structured data in SQL databases (MS Access, MS SQL Server, MySQL) and the two major cube databases (Hyperion Essbase and MS SSAS). Release v1 also connected to Excel and CSV files, which are often used in ad hoc ways. Release v2.0 included quick filter widgets, which allowed consumers to conduct useful analytical operations. Release v3.0 (April 2007) included both Dashboards and the Format Pane.

Release v1.0 (April 2005) could connect to multiple tables, including database views that joined data tables together. However, each sheet could only use a single table. The ability to join tables in Tableau was added in v2.0 (June 2005) with both a dialog and the RAWSQL function in the calculation language. Working with multiple tables got a significant boost with cross data source actions in v5.0 (May 2009). Since users who do not understand joins still need to work with multiple tables, Blending was added in v6.0 (Nov 2010), which allowed users to drag and drop multiple tables on one sheet. The Data Tab was added in v8.2 (June 2014) to give people who understand joins a better experience than a dialog. The Data Tab was integrated with the connection experience and supported the scenario of adding additional data sources during exploration.

The Table & Join Recommendations in v10.3 (May 2017), was Tableau’s first exploration of Machine Learning (ML) technology to provide computer assistance. Tableau’s exploration of ML has accelerated, including with “Einstein Discovery in Tableau” in Tableau 2021.1 (March 2021).

**Advantages of Tableau**

* Data visualization- Tableau supports complex computations, data blending, and dashboarding for the purpose of creating visualizations that deliver insights which cannot easily be derived by a spreadsheet.
* Ability to quickly create interactive visualizations- Using drag and drop functionalities of Tableau, the user can create an interactive visualization within minutes. The interface can handle endless variations while also limiting the user from creating charts that are against data visualization best practices.
* Tableau can handle large amounts of data- Tableau can handle millions of rows of data and different types of visualization can be created with a large amount of data without impacting the performance of the dashboards.
* Mobile Support and Responsive Dashboard- Tableau Dashboard has a great reporting feature that allows the user to customize the dashboard specifically for a certain device such as a mobile phone or laptop. Tableau automatically makes adjustments to ensure that the right report is served to the right device.

**Demerits of Tableau**

* Tableau is expensive, especially for enterprise-level licenses or advanced features. The cost of licensing fees can be an obstacle for individual users or small businesses.
* Tableau provides limited data manipulation capabilities, since it is mainly a visualization tool.
* It lacks advanced statistical analysis capabilities. Users who require advanced analytics need to integrate tableau with other analytical tools.
* Tableau struggles with handling extremely large datasets. Large data sources may require extensive processing and optimization to ensure smooth performance.
* Mastering the advanced features and functionalities of Tableau may require some training. Users with limited technical expertise may find it challenging to fully utilize the tool's capabilities.

**Structured Query Language (SQL)**

**History**

SQL history dates back to the early 1970s where it was developed by IBM researchers Donald D. Chamberlin and Raymond F. Boyce as part of the System R project. They created a language called SEQUEL (Structured English Query Language) to interact with the relational database management system they were working on.

Recognizing the need for a standardized database language, the American National Standards Institute (ANSI) and the International Organization for Standardization (ISO) formed a committee to develop a standardized version of SQL. In 1986, the first SQL standard, known as SQL-86, was published.

SQL quickly gained popularity due to its simplicity, standardization, and compatibility across different database systems. Its ability to manage and manipulate relational databases made it an essential tool for organizations and developers worldwide.

Over the years, SQL has undergone various updates and enhancements. New features were introduced in SQL-89, SQL-92, SQL:1999, SQL:2003, SQL:2008, SQL:2011, SQL:2016, and SQL:2019, expanding its capabilities and keeping up with evolving technologies.

**Advantages of SQL**

* SQL is designed to handle large amounts of data efficiently, making it a suitable choice for applications that require scalability and performance.
* SQL works with an efficient speed. This high speed can boost the amount of data retrieval. It can give users a rapid and effective means to obtain, manipulate, or store data. Operations like insertion, deletion, querying, manipulation, and calculations on data via analytical queries in a relational database can be accomplished in a matter of seconds.
* SQL is a standardized language used by all major relational database management systems. It provides a uniform platform worldwide to all its users.
* SQL provides powerful tools for data analysis and reporting. It supports various aggregate functions, grouping, sorting, and filtering capabilities, enabling you to perform complex calculations and generate meaningful reports.
* SQL is easy to learn and use as it does not require extensive knowledge about coding and writing programs. The syntactical rules are simple and easy to comprehend, making it a user-friendly language.

**Demerits of SQL**

* SQL databases typically scale up vertically by increasing hardware investment. This is both costly and time-consuming. They can face challenges when it comes to horizontal scalability. Scaling an SQL database across multiple servers can be complex and may require additional tools or techniques.
* Some SQL database systems, particularly commercial offerings, can be expensive, especially for large-scale deployments. Licensing fees, support contracts, and hardware requirements can contribute to the overall cost of using SQL databases.
* Modifying the structure or schema of a SQL database can be challenging, especially in production environments. Making changes to tables, relationships, or constraints often requires careful planning and potentially downtime for the database.
* SQL queries can sometimes be slow, especially when dealing with large datasets or complex queries. Poorly optimized queries, lack of proper indexing, or inefficient database design can lead to performance bottlenecks and slower response times.
* SQL is primarily designed for relational databases, which may not be suitable for all types of data. If your data does not fit well into a tabular structure with predefined relationships, SQL might not be the most efficient or flexible solution.

**Python**

**History**

Python is an open source, general purpose programming language. Guido van Rossum developed Python based on the defunct ABC programming language and named it after the Monty Python comedy troupe. Python was designed to be simple, readable, and highly extensible through the use of modules. The first version of Python was released in 1991, and the more fully featured Python 2.0 followed in 2000. Both releases have since been discontinued.

Python 3.0 was introduced in 2008, but is not compatible with earlier releases. Even though Python included the 2 to 3 upgrade utility, this decision was highly controversial and created significant issues for the user base. Because the transition was so difficult, there is unlikely to be a release 4 of Python. Instead, new minor releases are planned for each year. The Python Software Foundation currently administers Python, and they continue to work on new features and ongoing performance improvements.

**Advantages of Python**

* Python emphasizes code readability and simplicity, making it easy to learn and understand. Its clean and concise syntax allows developers to express concepts in fewer lines of code compared to other programming languages.
* Python has a large and active community of developers worldwide. This community contributes to the development of libraries, frameworks, and tools, making it easier to find support, share knowledge, and access a vast ecosystem of resources. Additionally, the community-driven nature of Python ensures continuous improvement and updates.
* Python is a programming language which is free of charge and open. The OSI-approved open source license under which Python is developed makes it a language free to use and distribute, including for commercial purposes. This reduces the maintenance costs.
* Python is a cross-platform language, meaning that Python code can run on various operating systems, including Windows, macOS, Linux, and more. This compatibility allows developers to write code once and deploy it across different platforms without extensive modifications.
* Python boasts a vast ecosystem of third-party libraries and frameworks. These libraries, such as NumPy, Panda, and TensorFlow, provide additional functionalities and tools for specific domains, such as data analysis, machine learning, web development, and scientific computing. The extensive collection of libraries enhances productivity and accelerates development.
* Python is a general-purpose language that can be used for a wide range of applications. It is suitable for web development, scientific computing, data analysis, machine learning, automation, and more. Its versatility and flexibility make it a valuable tool across different fields.

**Demerits of Python**

* Python is much slower than more efficient languages like C and Java. Python is interpreted and dynamically-typed, so the run-time compiler has a lot of work to do. It must constantly validate the type of each variable. This can be a concern for computationally intensive tasks or real-time systems where high performance is critical.
* Python can consume more memory compared to other languages due to its dynamic typing, garbage collection mechanism, and object-oriented nature. This can be a concern for memory-constrained environments or applications that need to handle large datasets.
* Python's dynamic typing allows for flexibility and rapid development but can lead to potential runtime errors that are only discovered during execution.
* Python is used for back-end programming due its high memory usage and slow speed; it is generally not used for front-end programming or mobile application development.
* Python's database access capabilities, while sufficient for many use cases, may not be as extensive as dedicated database languages like SQL. Python relies on external libraries for database connectivity, and these libraries may have limitations or performance trade-offs. It is not the best choice for applications that have complex interactions with a large corporate database.

**R**

**History**

R is a programming language and software environment for statistical computing and graphics. It was created by Ross Ihaka and Robert Gentleman at the University of Auckland, New Zealand, in the early 1990s. The development of R was influenced by the S programming language, which was developed at Bell Laboratories in the 1970s.

Ross Ihaka and Robert Gentleman sought to create an open-source and free alternative to commercial statistical software packages. They aimed to provide a flexible and powerful tool for statistical analysis and data visualization that would be accessible to researchers, statisticians, and data analysts.

The name "R" is derived from the first names of the creators, Ross and Robert. Initially, R was developed as a series of scripts written in the S language. Over time, it evolved into a standalone programming language with its own syntax and functionality.

In 1995, the first public release of R, version 1.0.0, was made available. The early versions of R gained popularity among statisticians and researchers due to its extensive statistical analysis capabilities, graphical capabilities, and its open-source nature.

As R gained traction, a community of developers and statisticians began contributing to its development. Packages were created to extend R's functionality, covering various domains such as machine learning, data visualization, data manipulation, and more. This vibrant community-driven development ecosystem has been a significant factor in R's growth and success.

The R Project for Statistical Computing was officially announced in 1997, providing a centralized platform for the development and distribution of R. The R Foundation for Statistical Computing was established in 2002 as a non-profit organization to support and oversee the development and distribution of R.

**Advantages of R**

* R provides a comprehensive set of built-in statistical functions and packages that enable advanced statistical analysis. It offers a wide range of statistical techniques, including regression analysis, hypothesis testing, time series analysis, clustering, and more.
* R has an extensive ecosystem of packages contributed by the R community. These packages cover various domains such as machine learning, data visualization, data manipulation, geospatial analysis, and bioinformatics. The availability of these packages allows users to leverage existing code and methodologies, saving time and effort. R language has more than 10,000 packages stored in the CRAN repository, and the number is continuously increasing.
* R offers powerful and customizable data visualization capabilities. Its built-in graphics functions, as well as packages like ggplot2 and plotly, allow users to create high-quality plots, charts, and interactive visualizations.
* R is an open-source language, which means it is freely available for anyone to use, modify, and distribute. This accessibility makes R an attractive choice for individuals, researchers, and organizations with budget constraints.
* R is a cross-platform language, meaning that R code can run on various operating systems, including Windows, macOS, and Linux. This compatibility allows users to develop and run R code on their preferred operating system without major modifications.
* R provides extensive tools and packages for data manipulation and cleaning. Packages like dplyr and tidyr allow users to efficiently filter, transform, reshape, and aggregate data. R's syntax and functions make it convenient to handle messy or complex datasets, ensuring data quality and consistency.

**Demerits of R**

* R can have a steep learning curve, especially for beginners with no prior programming experience. The syntax and concepts used in R may be unfamiliar to those new to programming, requiring dedicated time and effort to become proficient.
* R's memory management can be less efficient compared to other programming languages. Large datasets or complex analyses may require careful memory management techniques to avoid performance issues or crashes, as it takes up a lot of memory.
* While R is excellent for statistical analysis and data manipulation, it may not be the best choice for general-purpose programming or tasks outside the realm of data analysis.
* The programs and functions in R are spread across different packages. This makes it slower than alternatives such as MATLAB and Python.