**RESEARCH**

MSSQL:

* MS SQL provides various backup and recovery options, including full, differential, and incremental backups, to ensure data availability in case of system failures or disasters.
* MS SQL Server includes some practical built-in capabilities for machine learning with Python and R, automated reports, and analysis.
* MS SQL can be easily integrated with other Microsoft products, such as Excel, SharePoint, and Power BI, to provide seamless data management and analysis capabilities.
* SQL Server provides high data security and protection.
* Since the snowflake schema is the most generally used and straightforward, it will be employed as a dimensional model.

REFEENCE:

* <https://learn.microsoft.com/en-us/sql/relational-databases/data-collection/management-data-warehouse?view=sql-server-ver16>
* <https://learn.microsoft.com/en-us/sql/sql-server/editions-and-components-of-sql-server-2022?view=sql-server-ver16>

ALTERYX:

* Alteryx is made to handle data from a variety of sources, including databases, cloud-based platforms, spreadsheets, and more.
* Handle large files proficiently with faster runtime.
* It has strong community support, which will help with getting knowledgeable answers.
* With Alteryx's visual workflow design, users can create complex data pipelines with drag and drop feature and don’t having to write any code.
* Can be connect to various databases either cloud or on-premises.
* With the help of the user-friendly reporting tools in Alteryx, users can produce high-quality data-driven reports

REFEENCE:

1. <https://insightoriel.com/what-is-alteryx-advantages-of-alteryx/>
2. <https://www.techrepublic.com/article/alteryx-review/>
3. <https://beinex.com/topics/intelligence-does-not-always-come-in-suit-tie-11-key-alteryx-designer-features-you-should-know/>

GITHUB:

* GitHub's pull requests, which support a systematic code review process, let developers analyse and comment on each other's work.
* GitHub makes it simple for developers to share their work with others and collaborate on projects by allowing them to store their code repositories on its platform.
* GitHub is home to a large and active developer community that offers access to a wide range of open-source software, libraries, and tools.
* GitHub offers dependable hosting for code repositories, so businesses don't have to managing their own server infrastructure.

REFEENCE:

* <https://github.com/features>
* <https://blog.hubspot.com/website/what-is-github-used-for>

ETL:

* ETL (Extract, Transform, Load) procedures can help organisations analyse their data more thoroughly and draw conclusions from it, which can improve decision-making.
* By automating data processing and reducing the need for manual intervention, ETL technologies can help organisations process data more quickly and more affordably.
* ETL gives organisations an easy way to gather, transform, and store data from diverse sources by offering a centralised platform for data management.
* ETL methods make it possible to find and eliminate errors, duplicate data, and inconsistent data, which results in better data quality.
* **ETL provide transparency of your data lineage means** the complete life cycle of your data as it moves from sources to consumption by end-users is clear and easy to map.

REFERENCE**:**

* <https://www.airops.com/blog/etl-vs-elt>
* <https://hevodata.com/learn/importance-of-etl/>

**WHY ARE WE USING ALTERYX AS OUR ETL:**

ETL means Extract, Transform, Load process of data from a source or different sources into a single useable data source. According to my understanding, Alteryx have various tools and connectors to extract data from various sources (cloud, webservices and databases), and to transform that data like; cleaning, formatting, filtering, and aggregation data. These can validate and verify the quality of data during ETL process. Alteryx have GUI (Graphical user interface) that can be used to build workflows that can automate complex data transformations easily as it reduces manual coding. And, using alteryx we can load data to various systems like, databases, cloud and data warehouses.