	ETL ASSIGNMENT
What is a ETL ?	ETL stands for Extract, Transform and Load, which is a process used to collect data from various sources, transform the data depending on business rules/needs and load the data into a destination database.Handling all this business information efficiently is a great challenge and ETL plays an important role in solving this problem.
What is a ELT?	ELT stands for Extract, Load, and Transform.In an ELT process, data is extracted from one or multiple remote sources, loaded into the target data warehouse, and then transformed within the target database
List out the various tools used in ETL	1.Azure Data Factory 2.Google Cloud Dataflow 3.Oracle Data Integrator 4.Hevo Data 5.Microsoft SQL Server SSIS
Draw the mechanism of ETL	image added in the next sheet.pls refer
What is a datalake	A data lake is a central location that holds a large amount of data in its native, raw format. Compared to a hierarchical data warehouse, which stores data in files or folders, a data lake uses a flat architecture and object storage to store the data.
What is a data warehouse	A centralized repository and information system that is used to develop insights and guide decision-making through business intelligence. A data warehouse stores summarized data from multiple sources, such as databases, and employs online analytical processing (OLAP) to analyze data.
OLTP VS OLAP	table added in the next sheet.pls refer
What are all the various analytical tools that can be connected with Datawarehouse?	1.Tableau 2.BigQuery 3.PostgreSQL 4.Amazon dynamoDB 5. IBM Db2 Warehouse
Explain the different stages used in ETL	**1.Extract**: The first stage in the ETL process is to extract data from various sources such as transactional systems, spreadsheets, and flat files.This step involves reading data from the source systems and storing it in a staging area. **2.Transform**: In this stage, the extracted data is transformed into a format that is suitable for loading into the data warehouse. This may involve cleaning and validating the data, converting data types, combining data from multiple sources, and creating new data fields. **3.Load**: After the data is transformed, it is loaded into the data warehouse. This step involves creating the physical data structures and loading the data into the warehouse.

Explain 3 use cases for NOSQL	**1.Netflix:** Netflix uses NoSQL databases to store and manage massive amounts of data, including customer profiles, viewing histories, and content recommendations. **2.Uber:** Uber uses NoSQL databases to handle the massive amounts of data generated by its ride-sharing platform, including driver and rider profiles, trip histories, and real-time location data. **3.Big data analytics:** NoSQL databases are designed to handle large volumes of unstructured data, making them an excellent choice for big data analytics.
What are all the characteristics of a nosql	1.It's more than rows in tables—NoSQL systems store and retrieve data from many formats: key-value stores, graph databases, column-family (Bigtable) stores, document stores, and even rows in tables. 2.It's free of joins—NoSQL systems allow you to extract your data using simple interfaces without joins. 3.It's schema-free—NoSQL systems allow you to drag-and-drop your data into a folder and then query it without creating an entity-relational model. 4.It works on many processors—NoSQL systems allow you to store your database on multiple processors and maintain high-speed performance. 5.uses shared-nothing commodity computers—Most (but not all) NoSQL systems leverage low-cost commodity processors that have separate RAM and disk.
What are all the tools used to analyse nosql	1.Tableau 2.mongoDB compass 3.cluster control 4.Izenda