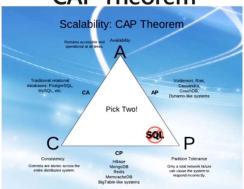


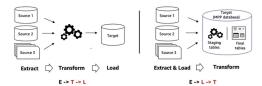
- There are a wider range of consumers using different types of devices to consume or generate data Consumers Data abundance Variety There's a wider variety of data types that need to be processed and stored Responsibilities A data engineers' role is responsible for more data types and technologies

Technologies different in the terms of who utilizes it wide set of tools and technologies



CAP Theorem





Comparison between relational and non relational data



Power BI is a business analytics service by Microsoft.

Components of Power BI ecosystem

Power BI Desktop Power BI Service Power BI Mobile Apps Power BI Gateway Power BI Embedded Power BI Report Server Power BI Visuals Marketplace

A Power BI dashboard is a single page, often called a canvas, that uses visualizations to tell a story.



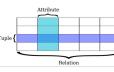


Features

- ACID: atomicity, consistency, isolation, and durability
- Stored Procedures and Relational Databases
- Database Locking and Concurrency
- High-performance, high-reliability
- Scalability

Structure **Relational Model**

RELATIONAL MODEL (RM) represents the database as a collection of relations. A relation is nothing but a table of values. Every row in the table represents a collection of related data values. These rows in the table denote a real-world entity or relationship.



Structure

Non Relational Model

NON RELATIONAL MODEL represents the database as a collection of record. Every row in the table represents a collection of related data values. Each row is called as record. Each cell of data is called as atom. Together they are called a database.

Non-Relational Data Characteristics:

- . Store the data in its original state and format. Doesn't impose a schema on data
- . Store the information for entities in collections or containers
- Other formats you might see include AvroORC

 - Parquet

Types of non-relational and NoSQL data



Image data

Natural language

Data visualization

- Graphic representation of data
- Data visualization refers to the techniques used to communicate data or information by encoding it as visual objects (e.g., points, lines or bars) contained in graphics.



show the data

- - avoid distorting what the data has to say

Characteristics (data visualization should)

Data Analytics

Ownerwise unknown events.
What can Happen?

- present many numbers in a small space
- make large data sets coherent encourage the eye to compare different pieces of data
- · reveal the data at several levels of detail, from a broad overview to the fine structure
- reasonably be serving clear purpose: description, exploration, tabulation or decoration
- have close integration with the statistical and verbal descriptions of a data set

Predictive analytics encompasses a variety of statistical techniques fron data mining, predictive modelling, and machine learning, that analyze current and historical facts to make predictions about future or

4541 **** **



Data Analytics

Descriptive - which talks about what was happened?

Data Analytics

Diagnostic analytics is a form of advanced analytics that examines data or content to answer the question, "Why did it happen?" It is characterized by techniques such as drill-down, data discovery, data mining and correlations.

Prescriptive analytics is a type of data analytics—the use of technology Prescriptive analytics is a type of data analytics—file use of technology to help businesses make better decisions through the analysis of raw data. Specifically, prescriptive analytics factors information about possible situations or scenarios, available resources, past performance, and current performance, and suggests a course of action or strategy. What we can prof. (Suggestion)

Cognitive analytics is a field of analytics that tries to mimic the human brain by draw inferences from existing data and patterns, draws conclusions based on existing knowledge bases and then inserts this back into the knowledge base for future inferences - a self learning feedback loop.

Prediction with action points

Azure SQL Database

Azure SQL Database is a fully managed platform as a service (PasS) database engine that handles most of the database management functions such as upgrading, patching, backups, and monitoring without user involvement.

MBAAS (database as a service)

Deployment options

- Single database PAS Hodel

Data Analytics

SQL Server on Azure Virtual Machines (laaS) enables you to use full versions of SQL Server in the cloud without having to manage any on-premises hardware. SQL Server VMS also simplify licensing costs when you pay as you go.

Azure Synapse Analytics

Azure Synapse is an analytics service that brings together enterprise data warehousing and Big Data

Processing Types

That Based

In a street of the control of the contro

Import big data with simple PolyBase T-SQL queries, 2 and then use the power of MPP to run high-performance analytics.

Ly Synapse SSL To SQL on dema Aparone Sport Ly hybrid data intoregration Stadio (unifica)

Storage Dala Lake

(3) Round Robin



DDL	DML	Tool	Description	Operating system
It stands for Data Definition Language.	It stands for Data Manipulation Language.	Azure Data Studio	A light-weight editor that can run on-demand SQL queries, view and save results as text, JSON, or Excel.	Windows macOS Linux
It is used to create database schema and can be used to define some constraints as well.	It is used to add, retrieve or update the data.			
			Manage a SQL Server instance or database with full GUI support.	Windows
It basically defines the column (Attributes) of the table.	It add or update the row of the table. These rows are called as tuple.			
It doesn't have any further classification.	It is further classified into Procedural and Non- Procedural DML.	SQL Server Data Tools (SSDT)	A modern development tool for building SQL Server relational databases, Azure SQL databases, Analysis Services (AS) data models, Integration Services (IS) packages, and Reporting Services (RS) reports.	Windows
Basic command present in DDL are CREATE, DROP,	BASIC command present in DML are UPDATE, INSERT, MERGE etc.			
RENAME, ALTER etc.			The mssql extension for Visual Studio Code is the official SQL Server extension that supports connections to SQL Server and rich editing experience for T-SQL in Visual Studio Code.	Windows macOS Linux
DDL does not use WHERE clause in its statement.	While DML uses WHERE clause in its statement.			
		Query Tools		

A Power BI dashboard is a single page, often called a canvas, that uses visualizations to tell a story.



Data warehousing workloads

A data warehouse workload refers to all operations that transpire in relation to a data warehouse. The depth and breadth of these components depend on the maturity level of the data warehouse the data warehouse workload encompasses:

The entire process of loading data into the warehouse
Performing data warehouse analysis and reporting

- Managing data in the data warehouse
- Exporting data from the data warehouse

Data warehousing requirements

- · Platform Functions
- Data Visualization
- Analytics
- · Document Management
- Decision Services
- · Integrations

Data Ingestion

Data ingestion is the process used to load data records from one or more sources to import data into azure services

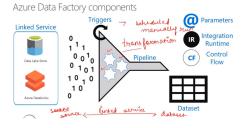


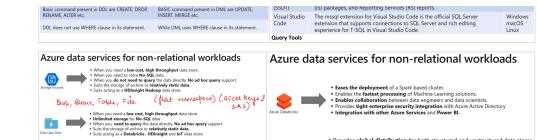
Types of data ingestion



Azure Data Processing Services

- · Azure Databricks
- ADF
- · Synapse Analytics
- · HDInsight





(historchical namespace)

