BIG DATA ANALYSIS USING IBM CLOUD DATABASES

TEAM MEMBER:

THAMIZHMATHY M(822721106045)

CAD_Phase1**: Document Submission**

Problem Title: Big Data Analysis

Problem Statement: Dive into the world of big data analysis with IBM Cloud Databases. Uncover hidden insights from vast datasets, from climate trends to social patterns. Visualize your findings and derive valuable business intelligence. Embark on data-driven adventures, exploring the endless possibilities of big data!

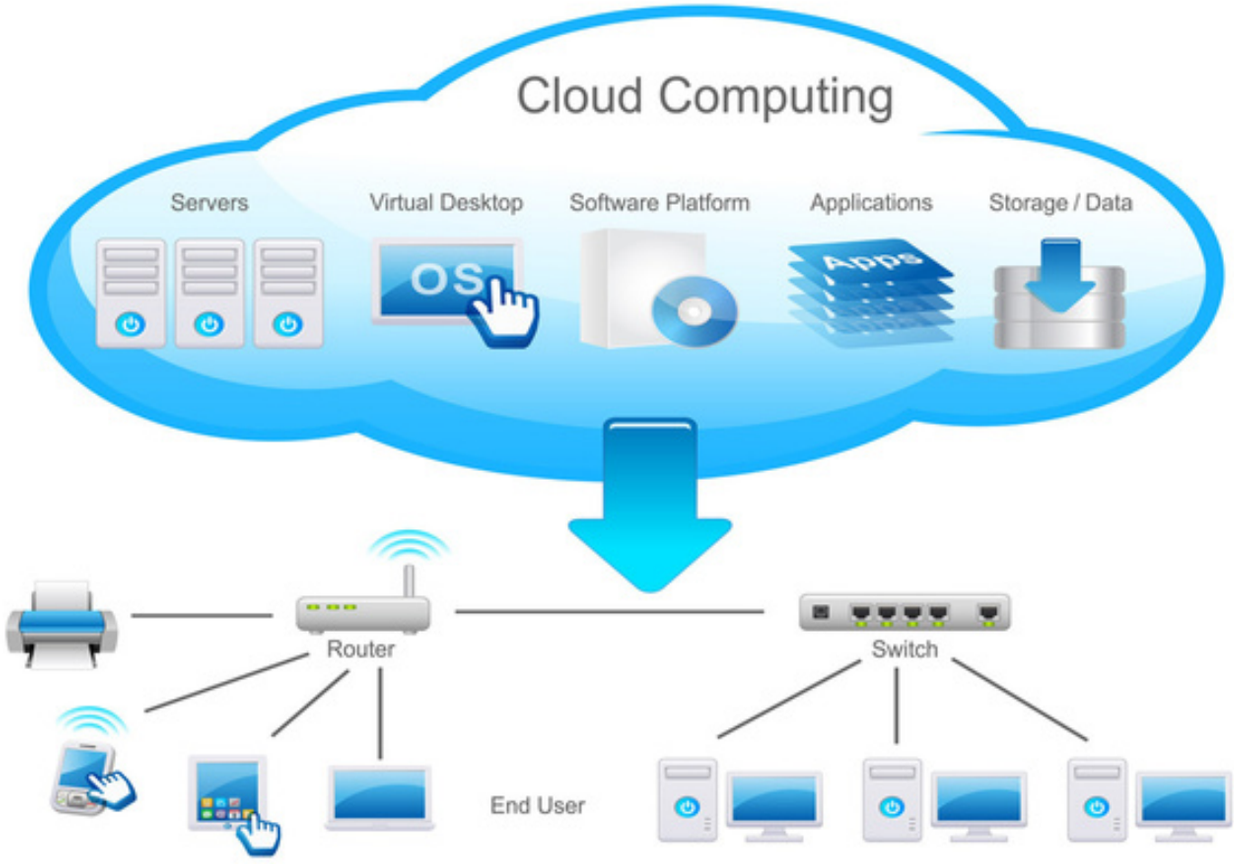
Objective: The project involves delving into big data analysis using IBM Cloud Databases. The objective is to extract valuable insights from extensive datasets, ranging from climate trends to social patterns. The project includes designing the analysis process, setting up IBM Cloud Databases, performing data analysis, and visualizing the results for business intelligence.

Design thinking:

1. **Data Selection:** Data from World Development Indicators and Climate Change Knowledge Portal on climate systems, exposure to climate impacts, resilience, greenhouse gas emissions, and energy use.

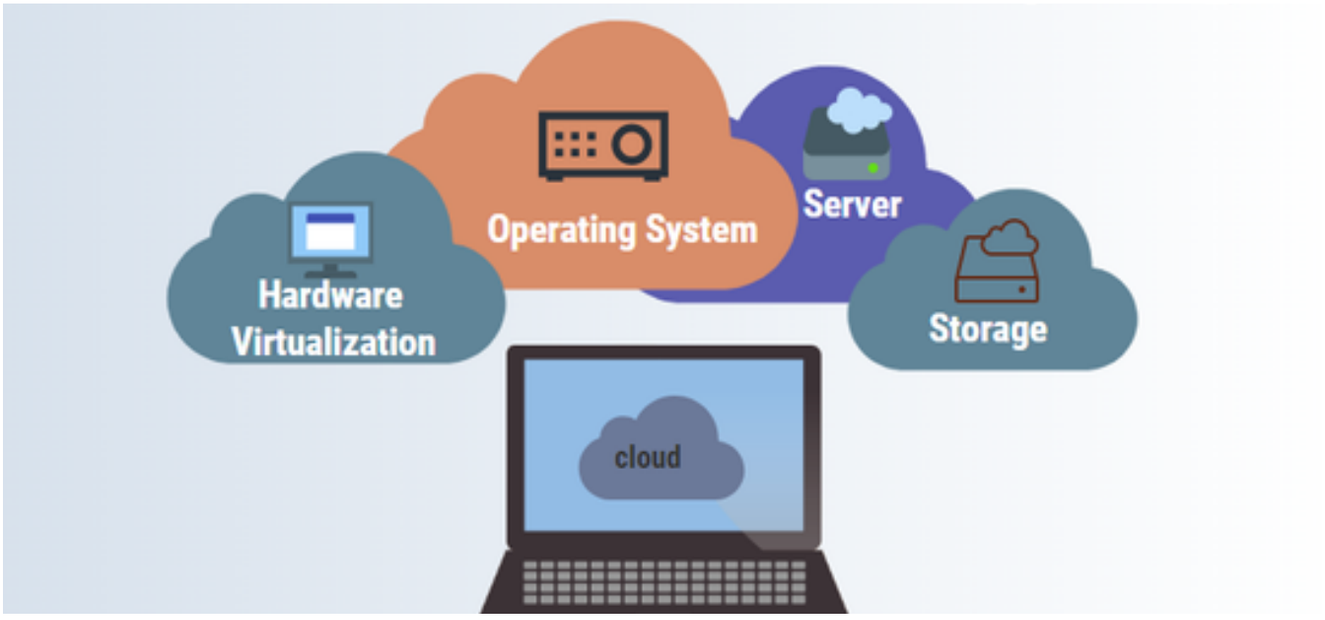
1	Country	Country name	Series code	Series name	SCALE	Decimals	1990	199
2	ABW	Aruba	AG.LND.EL5M.ZS	Land area below 5m (% of land area)	0	1	29.57	
3	ADO	Andorra	AG.LND.EL5M.ZS	Land area below 5m (% of land area)	0	1	0.00	
4	AFG	Afghanistan	AG.LND.EL5M.ZS	Land area below 5m (% of land area)	0	1	0.00	
5	AGO	Angola	AG.LND.EL5M.ZS	Land area below 5m (% of land area)	0	1	0.21	
6	ALB	Albania	AG.LND.EL5M.ZS	Land area below 5m (% of land area)	0	1	4.97	
7	ARE	United Arab Emirates	AG.LND.EL5M.ZS	Land area below 5m (% of land area)	0	1	4.64	
8	ARG	Argentina	AG.LND.EL5M.ZS	Land area below 5m (% of land area)	0	1	1.23	
9	ARM	Armenia	AG.LND.EL5M.ZS	Land area below 5m (% of land area)	0	1	0.00	
10	ASM	American Samoa	AG.LND.EL5M.ZS	Land area below 5m (% of land area)	0	1	17.89	
11	ATG	Antigua and Barbuda	AG.LND.EL5M.ZS	Land area below 5m (% of land area)	0	1	32.42	
12	AUS	Australia	AG.LND.EL5M.ZS	Land area below 5m (% of land area)	0	1	1.14	
13	AUT	Austria	AG.LND.EL5M.ZS	Land area below 5m (% of land area)	0	1	0.00	
14	AZE	Azerbaijan	AG.LND.EL5M.ZS	Land area below 5m (% of land area)	0	1	19.97	
15	BDI	Burundi	AG.LND.EL5M.ZS	Land area below 5m (% of land area)	0	1	0.00	
16	BEL	Belgium	AG.LND.EL5M.ZS	Land area below 5m (% of land area)	0	1	9.24	
17	BEN	Benin	AG.LND.EL5M.ZS	Land area below 5m (% of land area)	0	1	1.16	
18	BFA	Burkina Faso	AG.LND.EL5M.ZS	Land area below 5m (% of land area)	0	1	0.00	
19	BGD	Bangladesh	AG.LND.EL5M.ZS	Land area below 5m (% of land area)	0	1	14.07	
20	BGR	Bulgaria	AG.LND.EL5M.ZS	Land area below 5m (% of land area)	0	1	0.43	
21	BHR	Bahrain	AG.LND.EL5M.ZS	Land area below 5m (% of land area)	0	1	39.03	
22	BHS	Bahamas, The	AG.LND.EL5M.ZS	Land area below 5m (% of land area)	0	1	71.99	
23	BIH	Bosnia and Herzegovina	AG.LND.EL5M.ZS	Land area below 5m (% of land area)	0	1	0.07	

1. **Database Setup:**



3.Data Exploration: Data exploration is a broad process that is performed by business users and an increasing numbers of citizen data scientists with no formal training in data science or analytics, but whose jobs depend on understanding data trends and patterns. Exploration and preparation typically involve a great deal of interactive data analysis and visualization—usually using languages such as Python and R in interactive tools and environments that are specifically designed for this task.

4. Analysis techniques: Big-data analysis is a very active research area with significant impact on industrial and scientific domains where is important to analyze very large and complex data repositories. In particular, in many cases data to be analyzed are stored in cloud platforms and elastic computing clouds facilities are exploited to speed-up the analysis.



5. Visualization: Virtualization is technology that you can use to create virtual representations of servers, storage, networks, and other physical machines. Virtual software mimics the functions of physical hardware to run multiple virtual machines simultaneously on a single physical machine. Businesses use virtualization to use their hardware resources efficiently and get greater returns from their investment. It also powers cloud computing services that help organizations manage infrastructure more efficiently.

6. Business Insights: A business insight combines data and analysis to find meaning in and increase understanding of a situation, resulting in some competitive advantage for your business. This provides more than low-level understanding of an issue, giving you deeper insight into major mechanics related to your particular business.