



# ACHARYA INSTITUTE OF TECHNOLOGY

Affiliated to Visvesvaraya Technological University, Belagavi, Govt. of Karnataka.

Approved by AICTE, New Delhi

## Computer Science & Engineering (Data Science)

---

PROJECT REPORT  
ON  
**From BI to BigData : Explain, Design & Defend**

Subject Name: Big Data Analytics

Subject Code: BAD601

Submitted By : DISHA V SHETTY

1AY23CD018

Submitted To:

Ms. Surbhi



# ACHARYA INSTITUTE OF TECHNOLOGY

Affiliated to Visvesvaraya Technological University, Belagavi, Govt. of Karnataka.

Approved by AICTE, New Delhi

## Computer Science & Engineering (Data Science)

---

### Table of Contents

Task No	Task	Pg No.
1	Big data in Daily Life – Visual Storytelling	3-4
2	BI vs Big Data Play	5-6
3	Architecture Design Challenge	7-8
4	Analytics and Tool Match	9-10
5	Explain Big Data to a 10 year Old	11

### TASK 1 : Big Data in Daily Life – Visual Storytelling

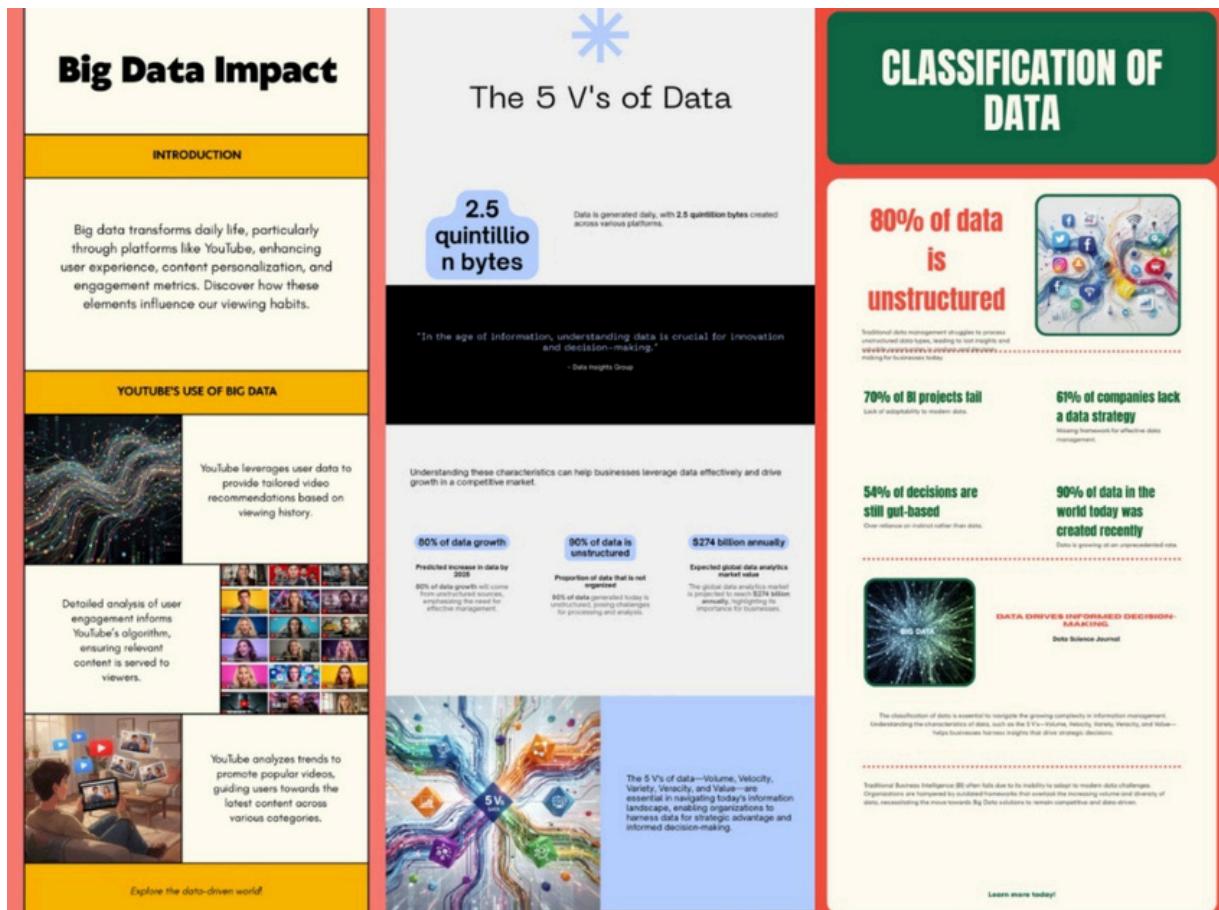


Fig 1.1 – infographic poster 1

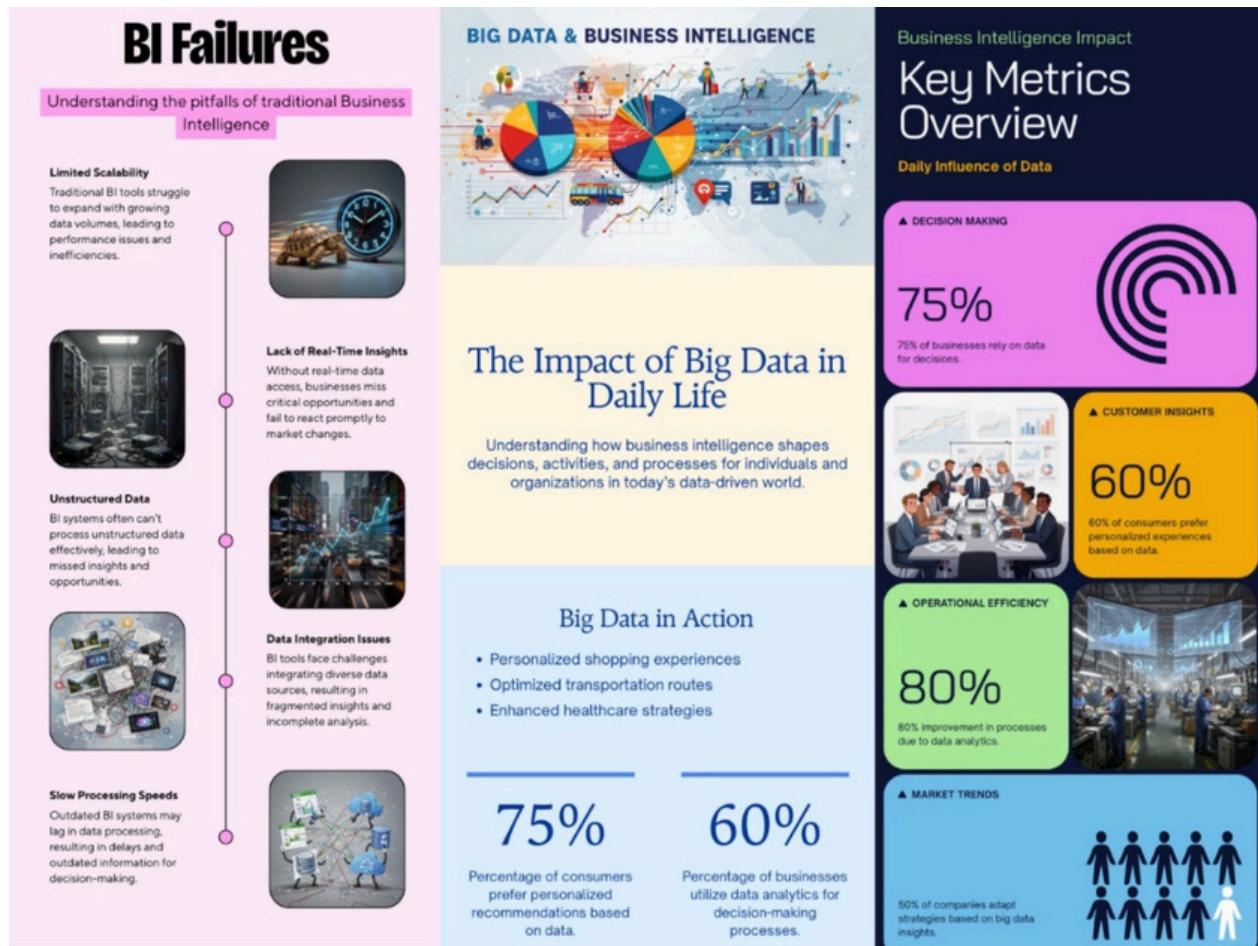


Fig 1.2 – Infographic Poster 2



### TASK 2 : BI vs Big Data Play

#### From Traditional BI to Big Data: A Technical Conversation

**Business Manager:**

Honestly, I don't understand all this Big Data hype. We already use Excel and SQL reports. That's enough for our decisions.

**Data Consultant:**

Excel and SQL are excellent for structured and moderate-size data. But modern businesses generate massive and diverse data like website clicks, app usage, and social media which traditional BI tools struggle to handle.

**Business Manager:**

But we store all our data in databases and run dashboards. What's the limitation?

**Data Consultant:**

Traditional BI mainly works with structured tables sales, finance, inventory. It cannot efficiently process unstructured or semi-structured data such as images, logs, videos, or customer behavior streams.

**Business Manager:**

Do we really need that kind of messy data?

**Data Consultant:**

Yes. That "messy" data reveals real-time customer behavior what users browse, like, or abandon. It helps companies personalize marketing and predict demand.

**Business Manager:**

Still, our SQL queries run fine. Why change?

**Data Consultant:**

They run fine because your data fits on a single server. But today data grows to terabytes or petabytes. Traditional databases scale vertically (bigger machines), which becomes costly and slow.

**Business Manager:**

So what triggered the Big Data evolution?

**Data Consultant:**

Three factors: huge data volume, high speed of data generation, and many data formats. These are known as the 3Vs Volume, Velocity, and Variety.



**Business Manager:**

And Big Data technologies solve this?

**Data Consultant:**

Exactly. Technologies like Hadoop store and process massive datasets across clusters of inexpensive computers instead of one powerful server.

**Business Manager:**

But our relational database is reliable and structured.

**Data Consultant:**

Relational databases require fixed schemas. Big Data often has flexible or changing structure. That's why NoSQL databases such as MongoDB and Cassandra are used they scale easily and handle diverse data types.

**Business Manager:**

So are you saying BI is outdated now?

**Data Consultant:**

Not at all. BI is still essential for historical reporting, KPIs, and dashboards. Big Data complements BI by supplying richer and larger datasets.

**Business Manager:**

So Big Data is about prediction and real-time insight?

**Data Consultant:**

Exactly. BI answers "What happened?" Big Data answers "What is happening now?" and "What will happen next?"

**Business Manager:**

I see Excel and SQL analyze the past, Big Data drives future decisions.

**Data Consultant:**

Right. Companies using Big Data gain faster insights, personalization, and competitive advantage.

**Business Manager:**

Alright, I understand now. We should explore Big Data along with our BI tools.

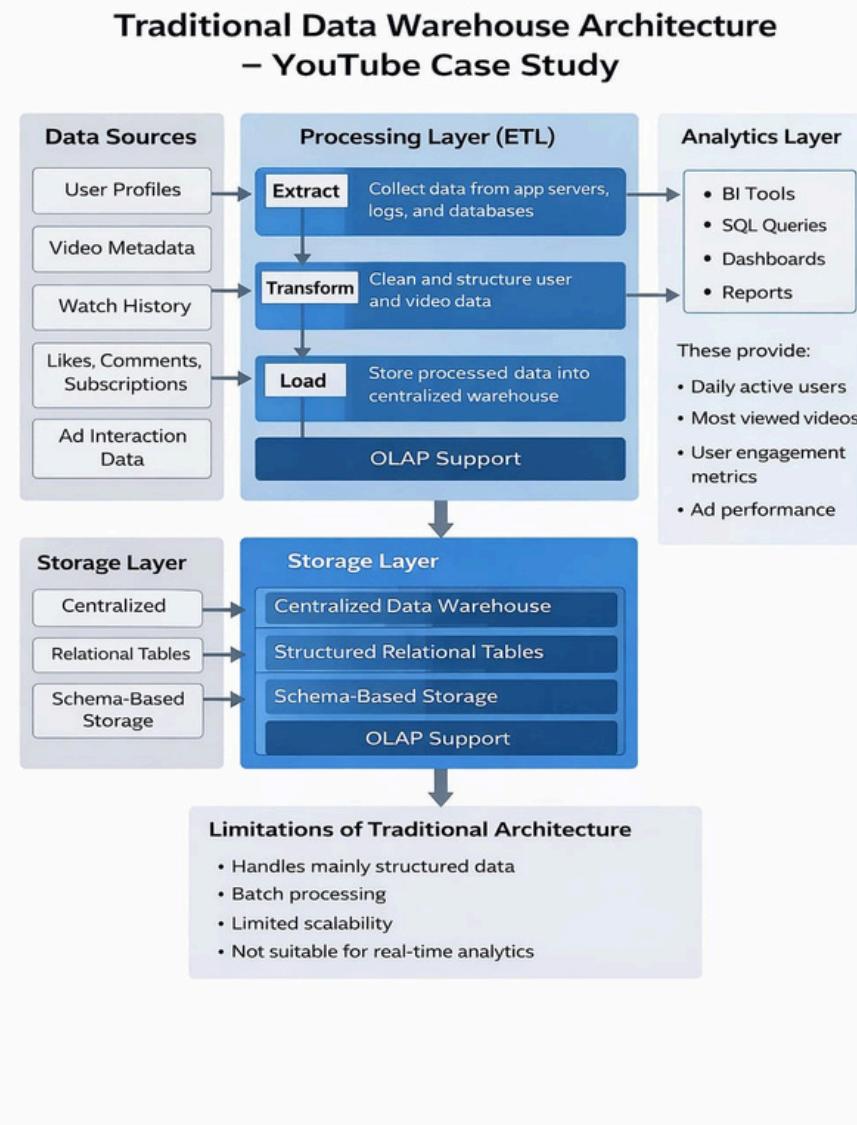
**Data Consultant:**

That's the best approach BI for structured reporting, Big Data for scale, speed, and advanced intelligence.

### TASK 3 : Architecture Design Challenge

3.1 Traditional Data Warehouse Architecture In traditional systems, data is stored in relational databases and processed using structured query methods for reporting and analysis in YouTube

#### Traditional Data Warehouse:



### 3.2 Hadoop-Based Big Data Architecture

Hadoop-based big data architecture of YouTube uses distributed storage (HDFS) and parallel processing (MapReduce/Spark) to store, process, and analyze massive video and user-interaction data at scale





### TASK 4 : Analytics and Tool Match

#### Descriptive Analytics

**Answers:** What happened?

- Summarizes historical data using reports and dashboards.
- Identifies trends and patterns from past data.

#### Diagnostic Analytics

**Answers:** Why did it happen?

- Analyses data to find the root cause of a problem or trend.
- Uses comparison and correlation analysis to explain changes.

#### Predictive Analytics

**Answers:** What will happen next?

- Uses historical data and machine learning to forecast future outcomes.
- Identifies patterns to predict trends and user behaviour.

#### Prescriptive Analytics

**Answers:** What action should be taken?

- Recommends the best action based on predictive results. Uses
- optimization and AI models to improve decision-making.



# ACHARYA INSTITUTE OF TECHNOLOGY

Affiliated to Visvesvaraya Technological University, Belagavi, Govt. of Karnataka.

Approved by AICTE, New Delhi

## Computer Science & Engineering (Data Science)

Business Question	Analytics Type	Tool Used
What happened?	Descriptive Analytics	Hadoop + BigQuery + Tableau
Why did it happen?	Diagnostic Analytics	Spark + Data Mining
What will happen next?	Predictive Analytics	Spark MLlib / Machine Learning
What action should be taken?	Prescriptive Analytics	AI Recommendation Engine + NoSQL

Table 4.1 – Analytics and Tools Table



# ACHARYA INSTITUTE OF TECHNOLOGY

Affiliated to Visvesvaraya Technological University, Belagavi, Govt. of Karnataka.

Approved by AICTE, New Delhi

## Computer Science & Engineering (Data Science)

### BONUS CHALLENGE : Explain Big Data to a 10-Year-Old

Big Data can be explained to a 10-year-old by imagining a huge toy box that keeps filling up every day with new toys, puzzle pieces, and LEGO blocks. After some time, the box becomes so full that it is hard to count or organize everything by hand. You may want to know which toy you play with the most, which one is your favorite, or which one you lost, but there are simply too many items to check one by one. So you use a smart robot that can quickly sort all the toys, find patterns, and give answers. Big Data is like that giant toy box full of information, and Big Data tools are like the smart robot that helps understand it. In real life, apps like YouTube collect huge amounts of data when millions of people watch videos, like, comment, and search every day. This creates more data than normal computers can easily handle. Big Data technology helps YouTube analyze all this information to recommend videos, find trending content, and personalize what each user sees. In simple words, Big Data means extremely large amounts of information that require special tools to store, process, and learn from so that computers can make smart decisions.