

## Q. What is Big Data?

→ (A) Intro.

(1) →  $40 \times$  bytes of data → generated every day by Smartphone  
[Text, audio, Video]  
month

58 Vsub  $\rightarrow 40 \times 5$

## (B) Data generated per minute:

Snapchat  
Snap

→ 2.1 Millions Snap Sent.

Google

3.8M

→ 3.8 Millions Searches.

Facebook

→ 1.1 Million

▶

→ 4.5 Million Videos Watched on Youtube

@

→ 118M Emails Sent.

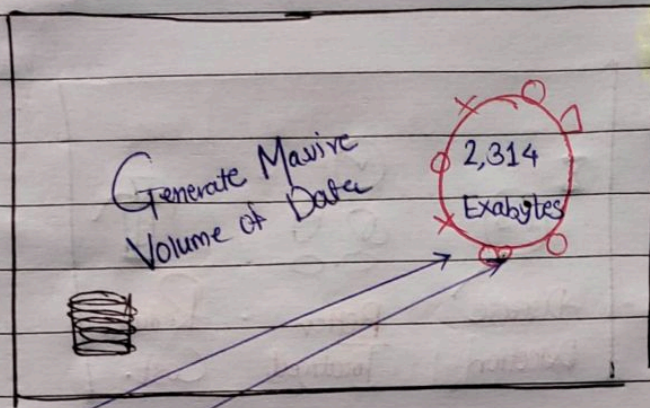
## (C) Classification of Big Data :

Concept of 5 V's

- Volume
- Velocity
- Variety
- Veracity
- Value.

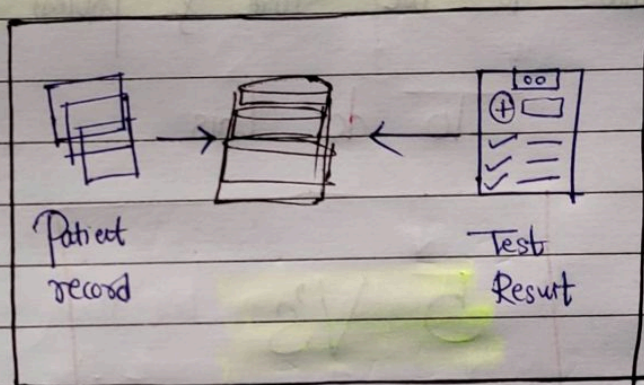


## ① Volume :

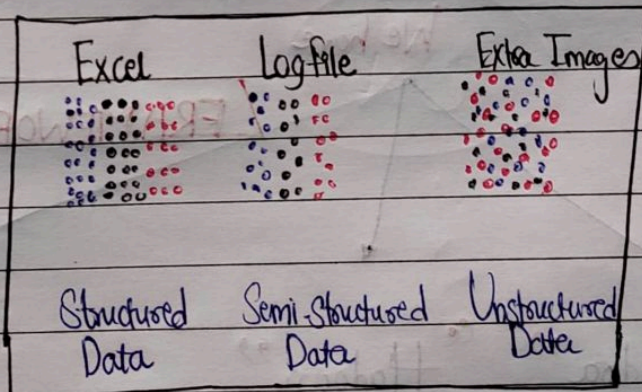


Clinic & Hospital generate lot of Data

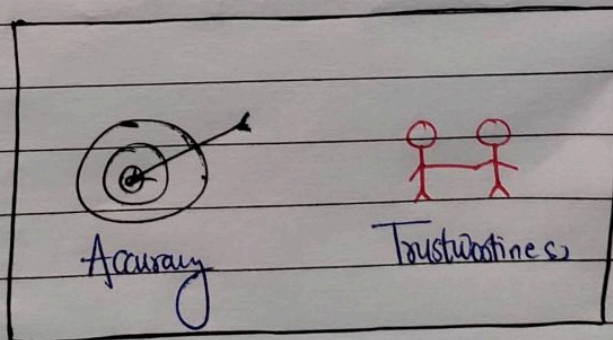
## ② Velocity



## ③ Variety

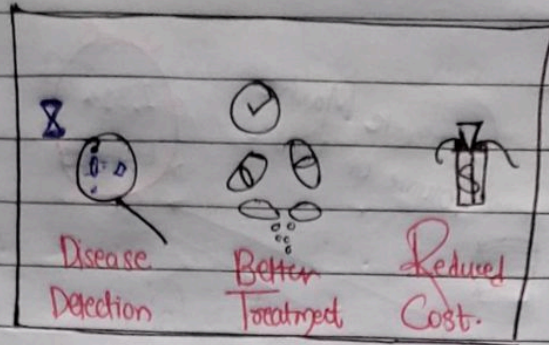


## ④ Veracity





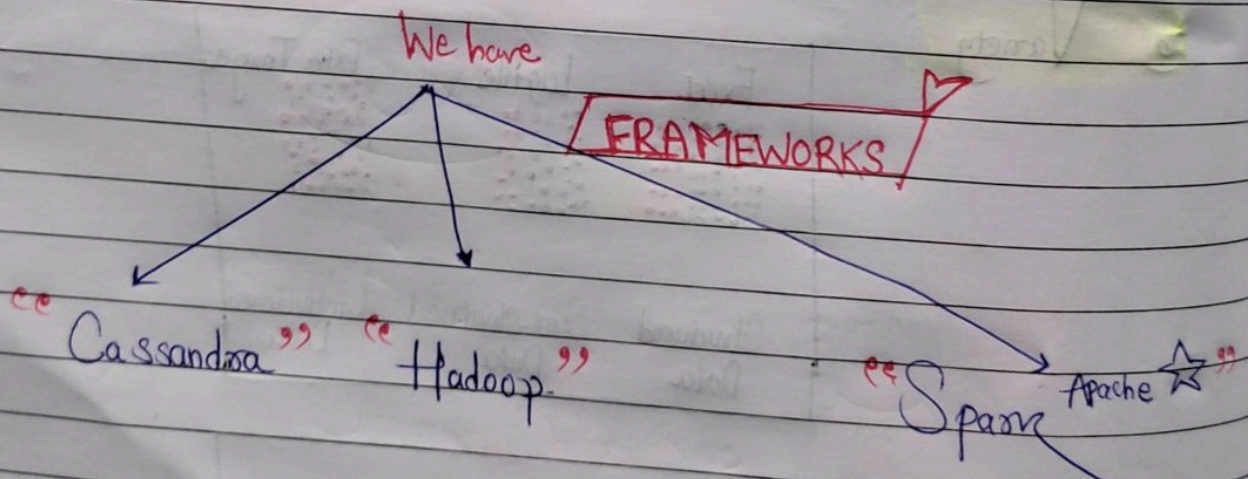
## ⑤ Value



# But how do we store & process this data?

To do this

5 V's

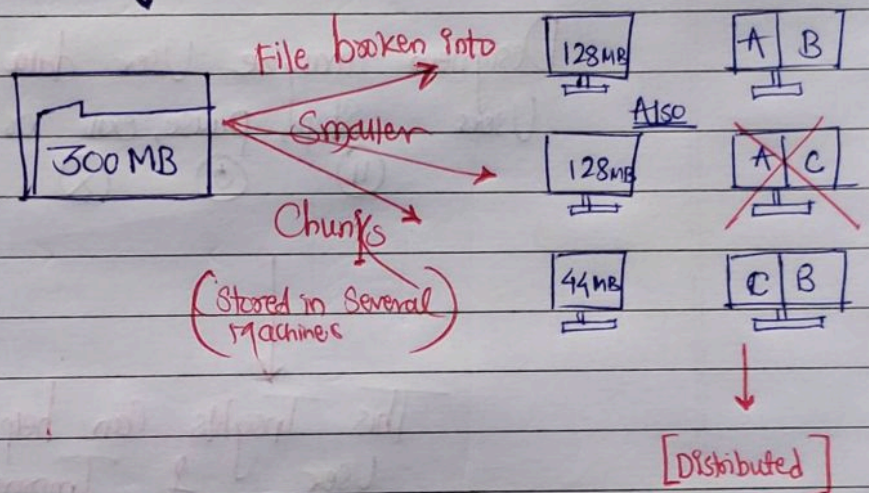




Let's See How it Stores & processes the Data!  
 Eg: Hadoop

↳ It Uses distributed file system known as "Hadoop distributed File system"

To store Big Data.

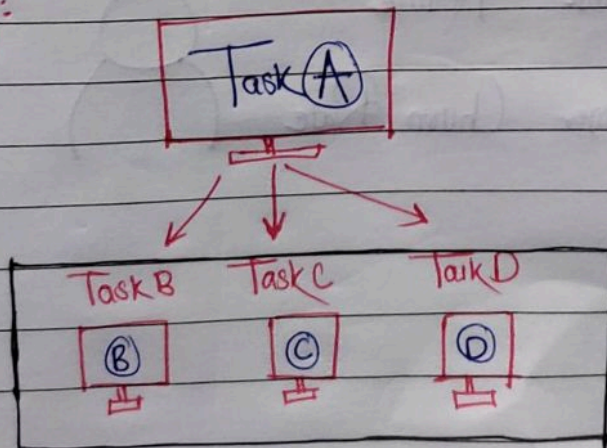


↳ Even if one Machine fails other Machine is Safe... ✓

Eg: ~~A/C~~ Fail But ~~A/B~~ ✓ ~~C/B~~ ✓ Works

# Map Reduce technique is Used to process Big Data.

Eg:



Results

• Parallel Processing •

↳ Fast Processing Easy ✓



Now We Stored & processed Big Data

↳ We Can Analyze Data for Various Application.

## (E) Application of Big Data

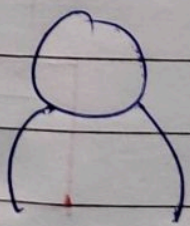
Game: ↳ Call of Duty.  
↳ Halo 3

Designers analyze User data at which place/time  
Users Stop, pause, exit or Restart  
① ② ③ ④

↓  
This Insights can help them to Understand  
User & Improve the  
User Experience

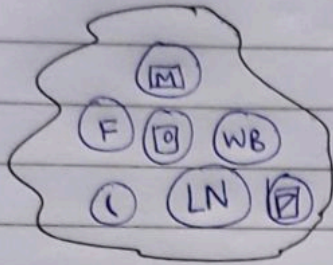
UX

↓  
All these Reduce  
Customer Churn Rate

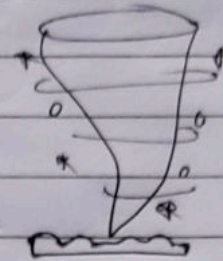




## (F) Disaster Management:



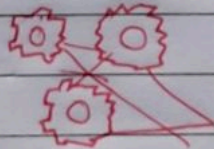
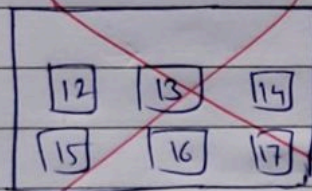
"Big Data"



"Hurricane  
Hurricane Sandy  
2012"

Necessary Measure we taken

It could predict hurricane's landfall five days in advance  
Which wasn't possible earlier



"Processed & Analyzed"



Q. What's Big Data?

(A) Intro.

→ ①  $40^x$  bytes of data → generated ~~every day~~ monthly  
[Text, audio, Video]

50 users  $\rightarrow 40 \times 5$

(B) Data generated per minute:

Snapchat  
Snap

→ 2.1 Millions

Google

3.8M

→ 3.8 Millions

Facebook

→ 1.1 Million

D

→ 4.5 Million