Let's have a look at the data generated per minute
on the internet

"2.1Million"   "3.8Million"   "1.0Million"   "4.5Million"
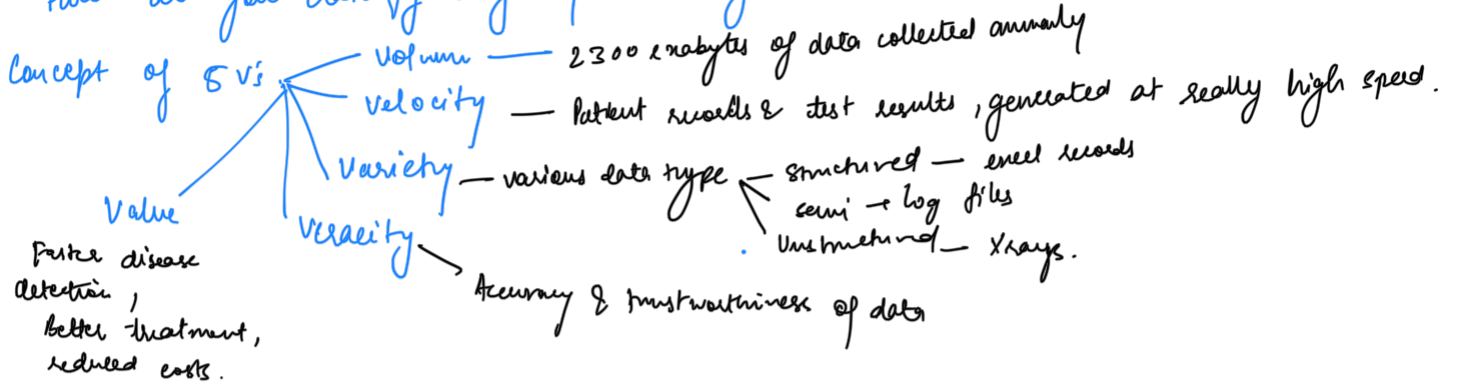
"188Million"   That's a lot of data

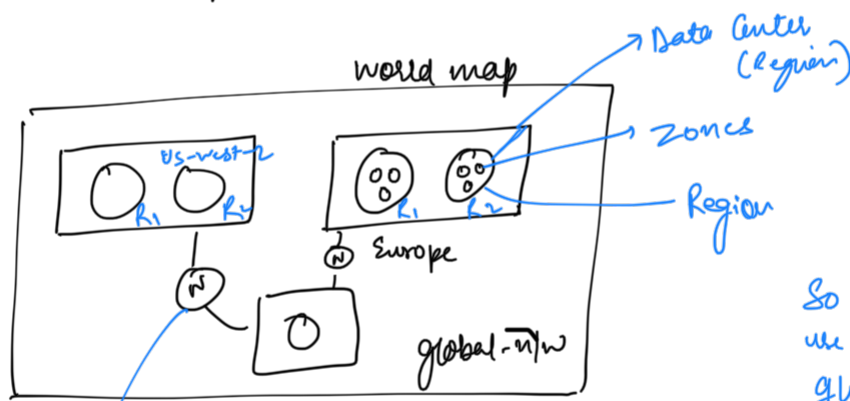# Big Data

How do you classify anything as big data.

Concept of 5 v's
- Volume —— 2300 exabytes of data collected annually
- Velocity —— Patient records & test results, generated at really high speed.
- Variety — various data type
  - structured — excel records
  - semi → log files
  - Unstructured — Xrays.
- Veracity → Accuracy & trustworthiness of data
- Value

Faster disease detection, Better treatment, reduced costs.

→ It can be incredibly hard to scalably ingest, store & analyze that data as it rapidly grows.

Bigquery : was designed to make large-scale data analysis accessible to everyone.
- → Storage → Perfect for BigData as BQ manage all storage & scaling ops
- → Ingestion → upload data from cloud storage / excel . . . . .
- → Querying → supports SQL

GCP

world map
- → Data Center (Region)
- → Zones
- — Region

us-west-2

R₁ R₂

Europe

global-n/w

google has 20 Regions
↓
61 Zones

So whatever service you use is either a
global resource,
regional resource,
zonal resource

① Network (global resource)

② Regional resource could be a
Sub-network —→ Subset of global network

when a global n/w comes & tries connecting to regional resources
↓
it will divide the particular n/w into multiple networks

breaks parent n/w into multiple child n/w

③ zonal resource

○   https://cloud.google.com/about/locations#regions

References:

BigQuery:

○   https://www.youtube.com/watch?v=d3MDxC_iuaw&ab_channel=GoogleCloudTech

○   https://www.youtube.com/watch?v=StEuT-pntZQ&ab_channel=GoogleCloudTech

○   https://www.youtube.com/watch?v=Abzj-Vyhi74&ab_channel=GoogleCloudTech

Service Account

→ Special account that can be used by an application or VM but not user

In google console

At user level → ○ email + admin → GCP Services

But at application level → [app] → should be able to use → Cloud storage / Cloud function

→ Your application uses the service account to call the Google API of a service, so that users aren't directly involved.

→ Service account belongs to the user rather than an individual end user