

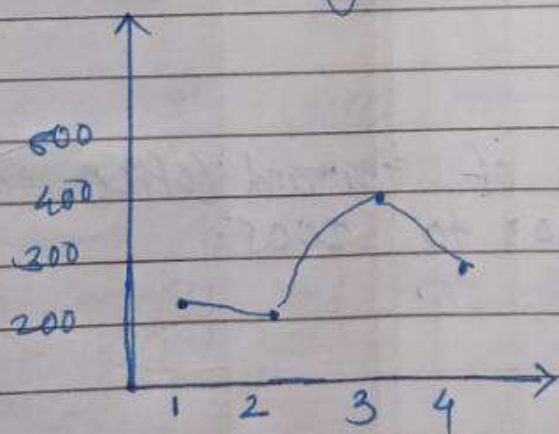
PRN : 2021291246048.

Q 3>

⇒

- Correlation:

- In Data visualization it's complicated to represent data using traditional way like draw charts and graphs.
- To simplify that using technology we represent that data in digital manner we use data visualization technique.
- To represent data using some kind of tools or programming like matlab in python we need variables like x, y.
- Those variables have relation or they maintain relation between them that relation known as Correlation.
- Consider following graph:



- In above eg. graph have scattered points which not understandable for simplifying this problem.
- We draw lines which connected to all points. to maintain Correlation.

- Preprocessing:

most data sets you find will be need some amount of preprocessing before they're even usable.

Return to the Seven Steps of Data Visualization accordingly we first take the raw data and then we will

- 1) parse

- 2) filter and

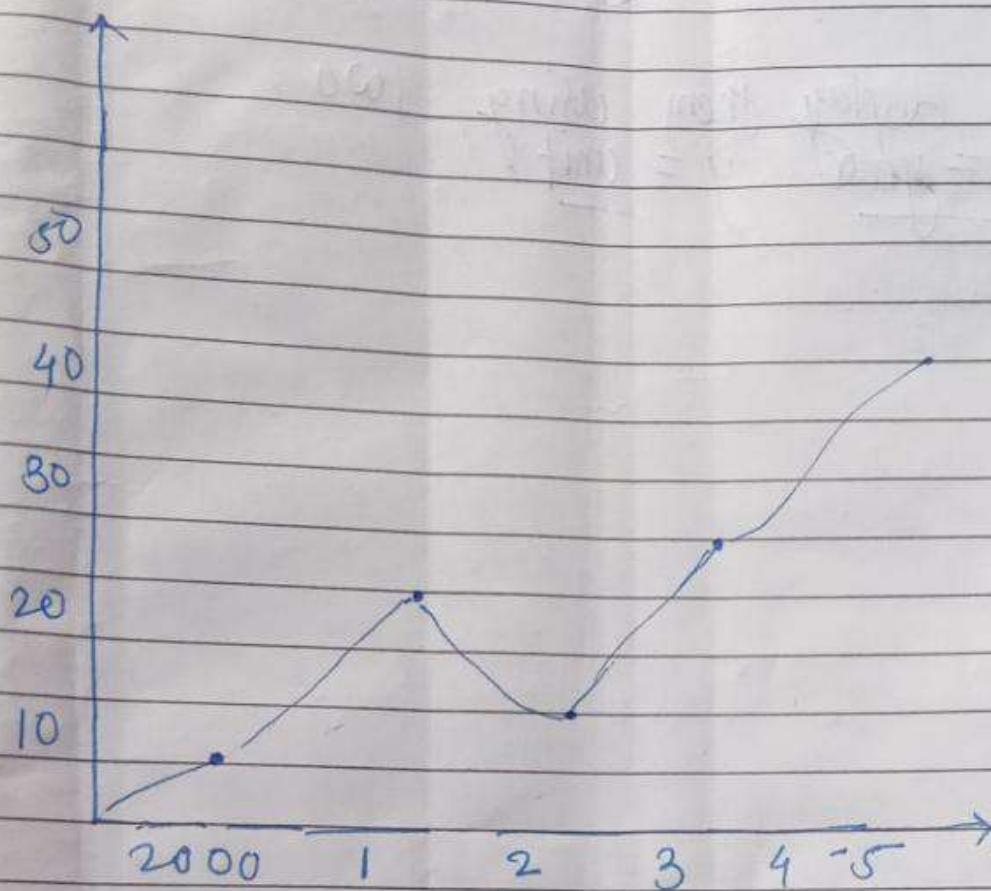
- 3) mine

them before the using of that data.

- Then again we clean that data and use it for specific purpose.

- After refining the data we will use that data to represent

- let be an example of tea and coffee consumed in india. from 2000 to 2005.



- here is an example to correlations and connections of data in data visualizations.
- firstly we collect that data of tea and coffee consumed by indians from different sources.
- like online and from different databases of shellers.
- It's common to know all data are in raw form
- we apply all steps of d.v. i.e. firstly we clean and filter them that means

- we work on pre processing before use it

- and then we display them using two variable $X = \text{year}$ $y = \text{unit}$