

Q1.A. What are the features of Big data?
Also enlist drivers of any two features.

Q1.B. Justify the significance of data cleaning with example.

Q2.A. Consider the following five application of data science and mention what kind of analysis can be done in that application (to the point analysis).

- i) E-commerce applications.
- ii) Product Manufacturing company.
- iii) Finance sector.
- (iv) Pharmaceutical company designing drugs.
- (v) Academic Institute.

Q2.B. Consider the following data and perform all four types of data analysis

Sr. No.	Name	Age	Salary in LPA	Job Profile	City
1	Rakhi	34	12	IT	Pune
2	Yash	21	8	Teaching	Nagpur
3	Anvee	38	16	IT	Mumbai
4	Arnav	48	18	IT	Mumbai
5	Athang	45	21	IT	Mumbai
6	Sakshi	25	5	Pharmacy	Nagpur
7	Rajesh	22	6	Pharmacy	Nagpur
8	Himanshu	24	8	Pharmacy	Hyderabad
9	Saurabh	31	10	Pharmacy	Delhi
10	Ritika	40	15	IT	Nagpur

Q3.A. Saffy Markets compares prices charged for identical item in all of the food stores. Here are the prices charged by each store for a pound of bacon last week.

\$1.08	\$0.98	\$1.09	\$1.24	\$1.33	\$1.14	\$1.55	\$1.08	\$1.22	\$1.05
--------	--------	--------	--------	--------	--------	--------	--------	--------	--------

Calculate the median price per pound and mean price per pound. Which value is better measure of the central tendency of these data?

Q3.B. The Birch company a manufacturer of electrical circuit boards has manufactured the following number of unit over the last 5 years.

1992	1993	1994	1995	1996
12500	13250	14300	15741	17630

Calculate the average percentage increase in unit produced over this time period and use this to estimate the production for ~~1999~~ 1999.