

Semester I

Subject Statistics for AIDS

Academic Year 2023-2024

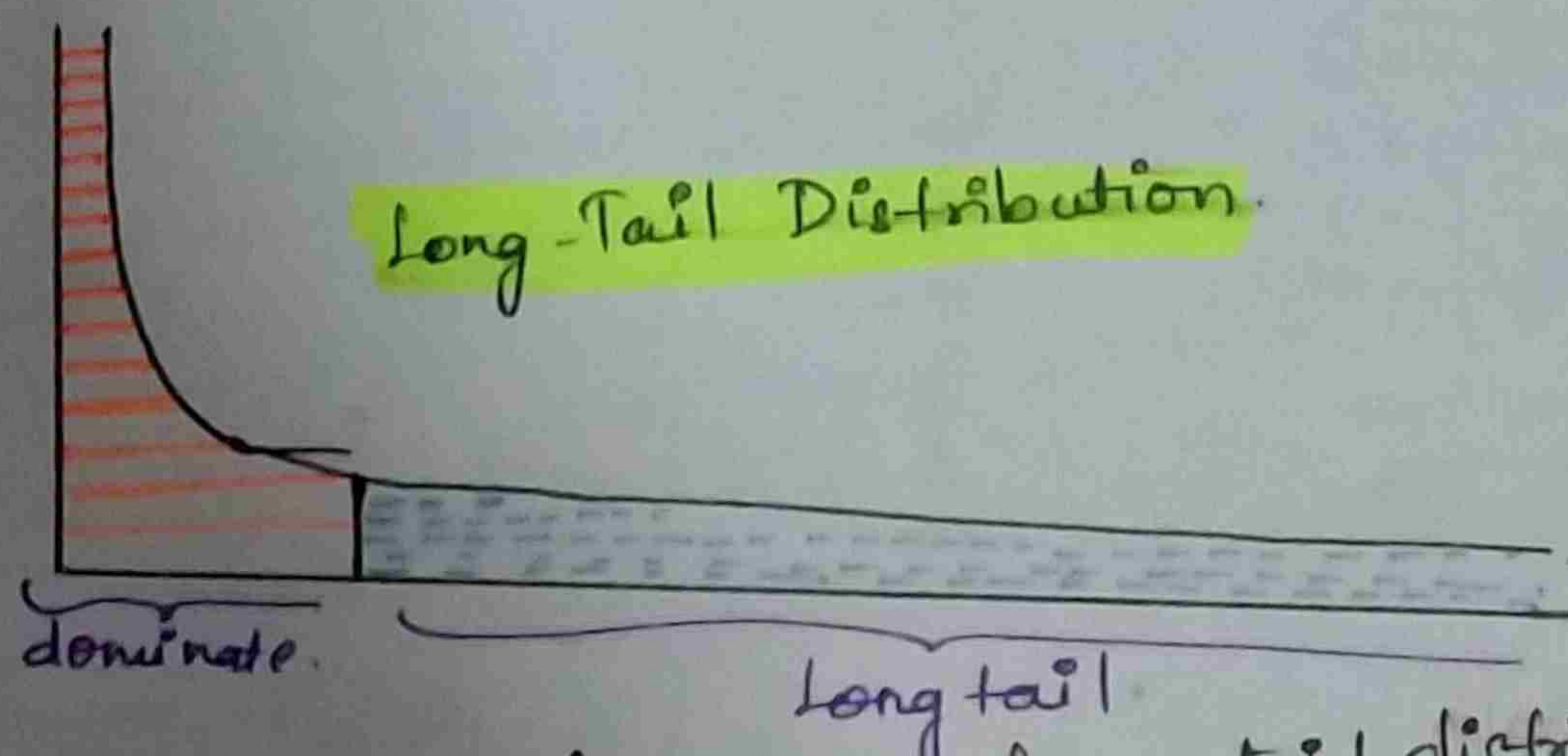
* The example 2 clearly describes how the QQ plot is generated.

* Y axis takes the sample data and X-axis takes the normal Distribution Data

* They are together plotted in QQ-Plot. If it generates a straight line then it is normal distribution.

LONG TAILED DISTRIBUTION:

* In statistics, a long tail of some distributions of numbers is the portion of the distribution having many occurrence, and they are far from the "head" or central part of the distribution.



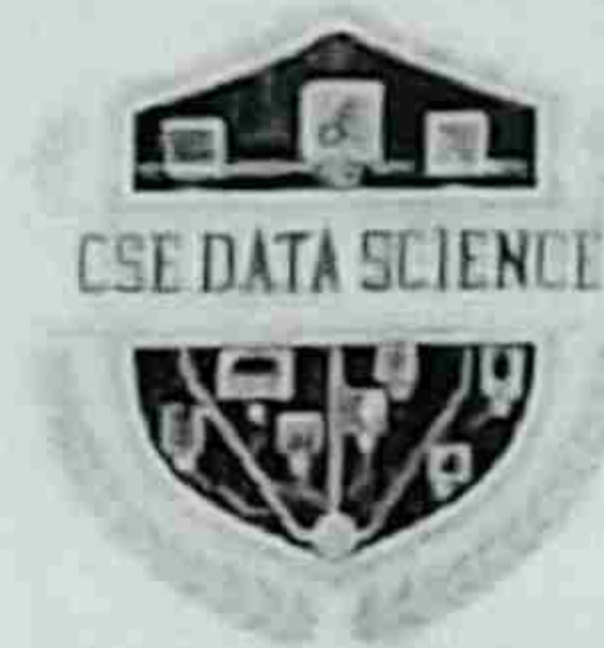
* The above graph shows long-tail distribution. To the right (dotted part) is the long tail and to the left (lined segment) are the few that dominate



PARSHWANATH CHARITABLE TRUST'S

A.P. SHAH INSTITUTE OF TECHNOLOGY

Department of Computer Science and Engineering
Data Science



Semester:



Subject: Statistics for AIDS

Academic Year: 2023-2024

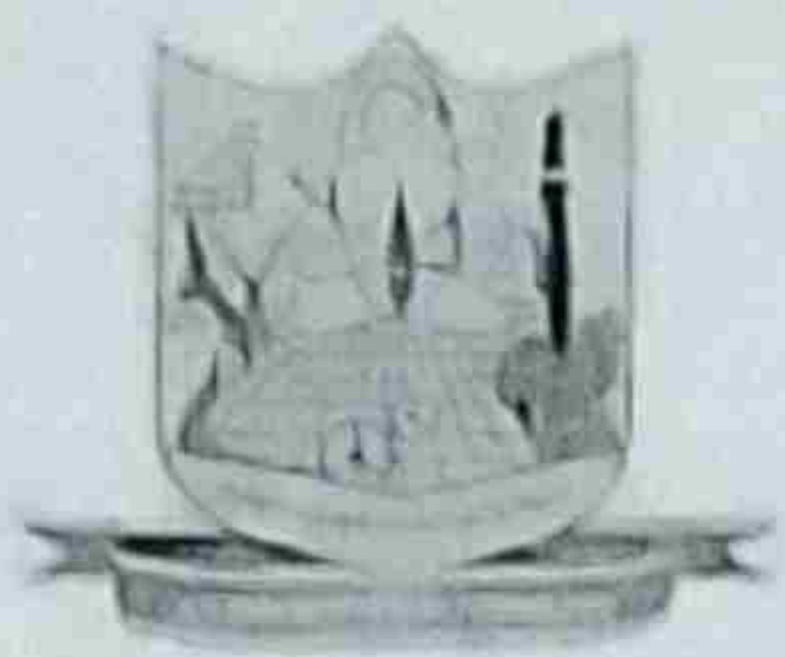
* In business, the long-tail is applied to rank-size distributions or rank-frequency distributions (of popularity), which often form power laws and are thus long-tailed distributions in the statistical sense.

* This is used to describe the retailing strategy of selling many unique items with relatively small quantities sold of each (the "long tail") - in addition to selling many fewer popular items in large quantities (the "head").

The long-tail concept is founded on application, research and experimentation. It is used in online business, mass media, microfinance, user-driven innovation, knowledge management and social network mechanisms.

Example:-

Let us consider there is a huge book store. The book store consists of different types of books. The book store has a huge collection of books. Assume 10% of Books are best and popular ones, 90% of Books are not popular. Here the question arises why the book store keeps all the books, why can't they keep only the best ones. The



Semester

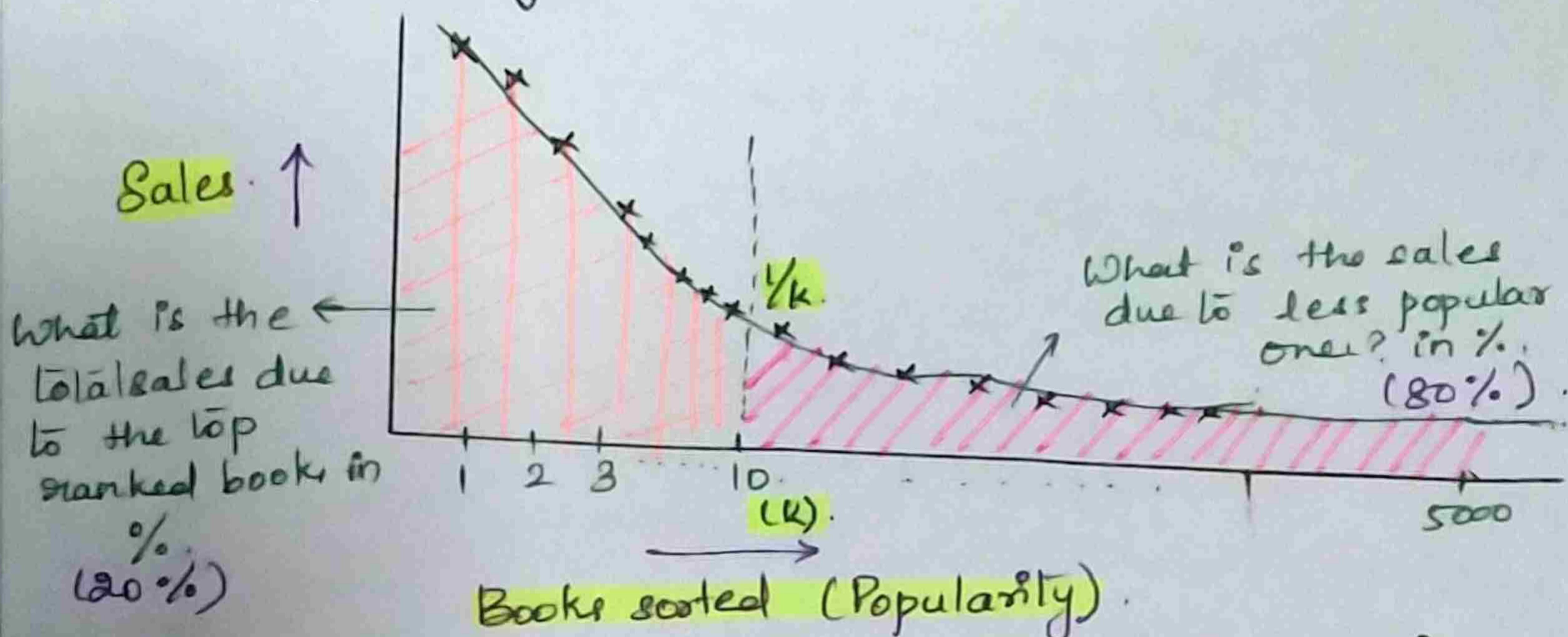


Subject Statistics for AIDS

Academic Year 2023-2024

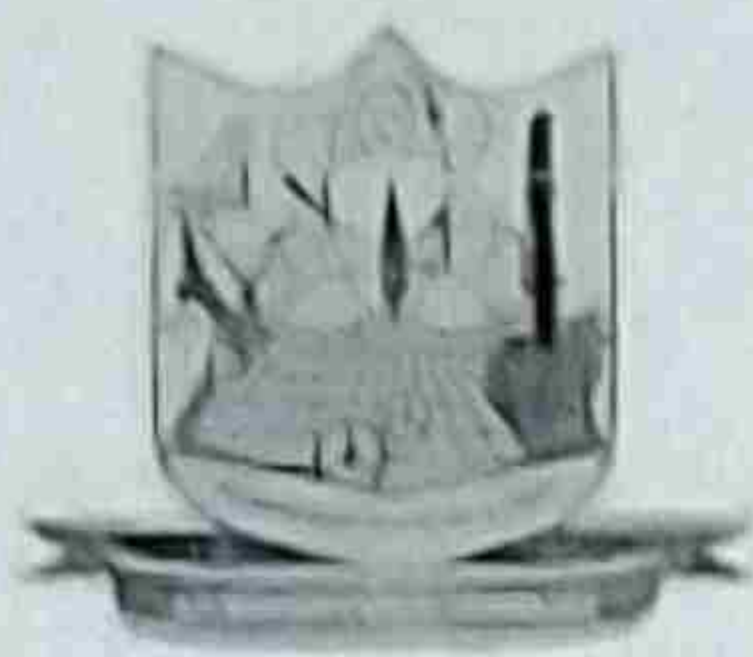
long tail distribution given the answer for this question.

Consider, the x axis resembles Books sorted according to popularity and y axis is the sales of those books.



While calculating we understand that 20% of sales is from top ranked books and 80% sales is from the less popular ones. In this case we have to concentrate on the pink shaded graph. The percentage could be different. One should always check in the business is popular product % vs not popular product.

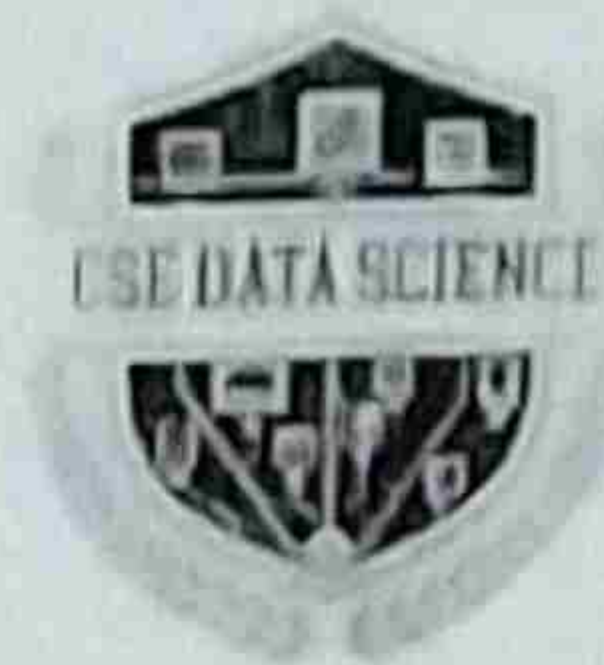
	Popular	Vs	Not Popular
It can be	20%		80%
	80%		70%
	50%		50%



PARSHWANATH CHARITABLE TRUST'S

A.P. SHAH INSTITUTE OF TECHNOLOGY

Department of Computer Science and Engineering
Data Science



Semester: I

Subject: Statistics for AIDS Academic Year: 20 23 20 24

* It is historically observed that, one cannot run the business with only the popular ones. It requires both. In this case the long-tail distribution is used.

* In this distribution if x-axis is plotted at k , then y-axis is plotted at $1/k$.