Data Science Notes by Sarowar Ahmed

📊 Chapter: Descriptive Statistics

📊 Topic: Percentiles

🔍 Hey there, GitHub fam! Ever heard of percentiles and wondered what they have to do with understanding data? Let's demystify this concept together in a way that's easy for everyone to grasp, regardless of age or background!

🌟 What Exactly Are Percentiles?

Think of percentiles as milestones along the data path, telling us how values are distributed in relation to each other. They help us

understand where a particular data point stands relative to the entire dataset.

📏 Formula for Calculating Percentiles:
For a dataset arranged in ascending order, the ρth percentile is the value below which ρ percent of the data fall.

▪ Mathematically, we find it using:
Percentileρ=(ρ/100)×(n+1)th value
Where n is the total number of values in the dataset.

📊 Example to Illustrate Percentiles:

Imagine you have the following set of test scores: 60, 65, 70, 75, 80, 85, 90, 95, 100.

▪ 50th Percentile (Median): Since there are 9 scores, the 50th percentile is the value at the (0.5 × 9)th position, which is the 5th value. So, the median score is 80.

▪ 25th Percentile: The value at the (0.25 × 9)th position is the 2.25th value, which is between the 2nd and 3rd scores (65 and 70). Since it's between two values, we typically take the average: (65+70)/2=67.5
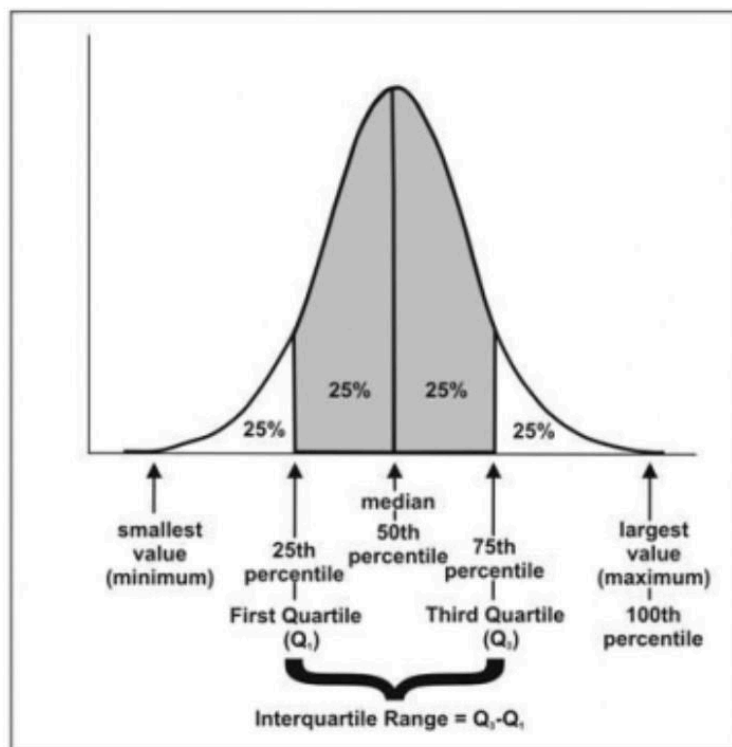
📈 Why Do Percentiles Matter?

Percentiles give us insight into the distribution of values within a dataset. They're invaluable when comparing individual data points to a larger group or when understanding where someone stands relative to their peers.

🌟 Visualization:

Whether you're analyzing student grades, customer satisfaction scores, or even employee performance metrics, understanding percentiles can provide deep insights into the distribution and relative positioning of data points.

Got any questions about Percentiles !? Feel free to ask me via Linkedin! Let's keep learning together.

My LinkedIn   Date: 05/04/2024