**Task 1: Types of Statistics**

**Descriptive Statistics** summarize and present data meaningfully. Example: "The average test score in a class is 85%." **Inferential Statistics** use sample data to make generalizations about a population. Example: "Based on a survey, 60% of people prefer online shopping."

**Task 2: Population vs. Sample**

**Scenario:** A researcher studies the average income of families in a city by surveying 100 households.

* **Population:** All families in the city.
* **Sample:** 100 surveyed households.

**Task 3: Sampling Techniques**

**Scenario:** Survey employee satisfaction in a company with departments of varying sizes.

* **Sampling Technique:** Stratified sampling.
* **Justification:** Ensures fair representation from each department based on size.

**Task 4: Types of Variables**

1. **Number of pets:** Quantitative (Discrete)
2. **Blood type:** Qualitative
3. **Weight in kg:** Quantitative (Continuous)
4. **Satisfaction level (low, medium, high):** Qualitative

**Task 5: Measures of Central Tendency**

**Dataset:** [3, 5, 5, 6, 7, 100]

* **Mean:** (3 + 5 + 5 + 6 + 7 + 100) / 6 = **21**
* **Median:** (5 + 6) / 2 = **5.5**
* **Mode:** **5**
* **Interpretation:** The mean is skewed by 100; the median and mode better represent the central trend.

**Task 6: Measures of Dispersion**

**Dataset:** [2, 4, 6, 8, 10]

* **Mean:** 6
* **Variance:** [(2-6)^2 + (4-6)^2 + (6-6)^2 + (8-6)^2 + (10-6)^2] / 5 = **8**
* **Standard Deviation:** √8 ≈ **2.828**
* **Significance:** Variance measures spread; standard deviation shows how far values deviate from the mean.

**Task 7: Z-Scores and Interpretation**

**Given:** Mean = 50, Standard deviation = 10

* **For x = 60:** z = (60 - 50) / 10 = **1**
* **For x = 40:** z = (40 - 50) / 10 = **-1**
* **Interpretation:** x = 60 is one standard deviation above the mean, and x = 40 is one below.

**Task 8: Gaussian Distribution**

The given data follows a normal distribution with a **mean of 100** and a **standard deviation of 15**. Based on the **68-95-99.7 rule**, we can estimate that:

* About **68%** of the data falls within **one standard deviation** of the mean.
* This means most values lie between **85 (100 - 15)** and **115 (100 + 15)**.

So, we can conclude that **approximately 68% of the data lies between 85 and 115** in this normal distribution.

**Bonus Task: Box Plot and Outlier Detection**

**Dataset:** [1, 2, 3, 4, 5, 6, 7, 8, 9, 20]

* **Q1 (25th percentile):** 2.5
* **Q3 (75th percentile):** 7.5
* **IQR:** 7.5 - 2.5 = **5**
* **Lower Fence:** Q1 - (1.5 × IQR) = 2.5 - 7.5 = **-5**
* **Upper Fence:** Q3 + (1.5 × IQR) = 7.5 + 7.5 = **15**
* **Outliers:** Values outside [-5, 15]. **Outlier: 20**