Distribusi Frekuensi dan Grafik

EGI SAFITRI, S.MAT., M.SI

Table of Contents

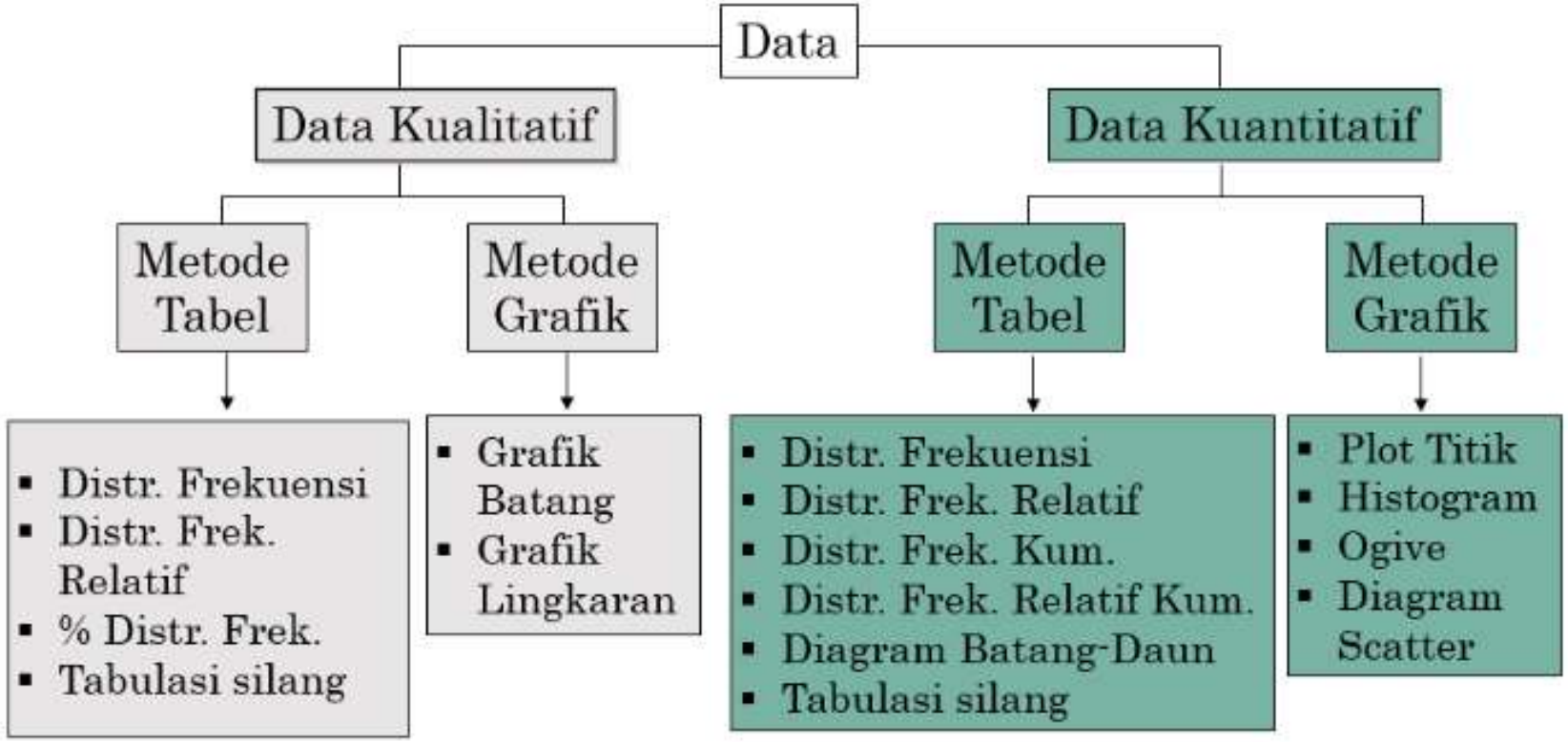
Definition

1. **Data Processing:** The process of obtaining data summaries based on raw data groups.
2. **Objective:** To obtain statistical data that can be used to analyze or answer problems collectively.
3. **Raw Data:** The recorded results of events or characteristics of elements obtained during the data collection phase.

Data Processing

- To obtain statistical data, collected data from the investigated elements must be processed.
- **Data processing** means transforming raw data into summarized information in the form of numerical summaries.
- Collected raw data, when processed, presented, and analyzed, becomes more useful as a basis for decision-making.
- Data processing can be done manually or using electronic tools (calculators, computers).

Use of Graphs According to Data Type



Data Presentation

1. **Table**: A collection of numbers arranged by category to facilitate data analysis.
    - Example: Number of employees based on education and work experience.
2. **Graph**: Visual representations of data in numerical or symbolic form, usually derived from tables.

Descriptive Statistics

1. One function of descriptive statistics is to present data clearly and simply to provide an overview of the situation to relevant parties.
2. Collected data is initially raw or unorganized, making it difficult to interpret, necessitating data grouping.
3. Data arranged into specific interval classes is called a frequency distribution.

Definition of Distribution

- **Frequency distribution** is the arrangement of data into specific interval classes or categories, linked with their respective frequencies to provide a systematic and simplified summary of data.
- The purpose of creating frequency tables is to organize raw (ungrouped) data into a structured format without losing essential information.

Frequency Distribution

1. Data presented in a frequency distribution has both advantages and disadvantages.
2. **Advantages:** Provides a comprehensive summary of the available data.
3. **Disadvantages:** The initial details or raw information are lost, making grouped data appear less precise.

Example:
A table shows that three respondents are aged between **35-40 years**. This means the actual ages of these three respondents remain unknown.

Creating a Frequency Distribution Table

To create a frequency distribution list with equal class intervals, follow these steps:

1. Determine the **range** = largest data value - smallest data value.
2. Determine the number of **class intervals** needed using **Sturges' rule:**
    - Number of classes = 1 + (3.3) log n, (n = total data points).
3. Determine the **class width (interval length, P):**
    - $P = \text{range}/\text{number of classes.}$
4. Select the **lower boundary of the first class interval**, typically the smallest data value.

Example

Here are the exam results of 80 students

79	49	48	74	84	98	87	80
80	84	90	70	91	93	82	78
70	71	92	38	56	81	74	73
68	72	85	51	65	93	83	86
90	35	83	73	74	43	86	88
92	93	76	71	90	72	67	75
80	91	61	72	97	91	88	81
70	74	99	95	80	59	71	77
63	60	83	82	60	67	89	63
76	63	88	70	66	88	79	75

Work Steps

1. **Number of Classes:**
   - Number of classes = 1 + (3.3) log 80
   - = 1 + (3.3) log (1.9031)
   - ≈ 7.2802 ≈ 7.28 ≈ 7.3 ≈ 7
2. **Class Width:**
   - $P = \frac{64}{7}$
   - ≈ 9.1429 ≈ 9.143 ≈ 9.14 ≈ 9.1 ≈ 9
3. **Lower Boundary of the First Class Interval:**
   - 35

Result Table

From the data, an absolute frequency distribution table can be created: **Statistics exam scores for 80 students.**

Exam Score	Frequency
35-43	3
44-52	3
53-61	5
62-70	12
71-79	20
80-88	21
89-97	14
98-106	2
TOTAL	80

Terminology in Frequency Distribution

- **Class**: A group of data in the form of a-b.
  - Example: 35-39.
- **Class Limits:** The values at the boundaries of a class.
  - Example:
    - Class 35-39
    - Upper limit (ua): 39
    - Lower limit (ub): 35
- **Class Boundaries:** Depend on the data precision used:
  - **Formula:**
    - **Lower boundary (bb) = ub - 0.5**
    - **Upper boundary (ba) = ua + 0.5**
- **Example of Smallest Measurement Unit:**
  - If data includes 35, 34, 32, the smallest measurement unit is 1.
  - If data includes 35.1, 34.2, 32.5, the smallest measurement unit is 0.1.
- **Example of Class Boundaries:**
  - For class 35-39:
    - **Lower boundary (bb):** 35 - 0.5 = 34.5
    - **Upper boundary (ba):** 39 + 0.5 = 39.5

Relative and Cumulative Frequency Distribution

- **Relative frequency distribution** compares the frequency of each class with the total frequency, multiplied by 100%.
- **Cumulative frequency distribution** consists of two types:
  - **Less than cumulative frequency distribution**
  - **More than cumulative frequency distribution**

Relative Frequency Distribution Table

Class Interval	Class Boundaries	Midpoint	Frequency	Relative Frequency (%)
9-21	8.5-21.5	15	3	5
22-34	21.5-34.5	28	4	6.67
35-47	34.5-47.5	41	4	6.67
48-60	47.5-60.5	54	8	13.33
61-73	60.5-73.5	67	12	20
74-86	73.5-86.5	80	23	38.33
87-99	86.5-99.5	93	6	10
Total			**60**	**100**

Less Than Cumulative Frequency Distribution Table

Class Interval	Class Boundaries	Cumulative Frequency (Less Than)	Cumulative Percentage
9-21	less than 8.5	0	0
22-34	less than 21.5	3	5
35-47	less than 34.5	7	11.67
48-60	less than 47.5	11	18.34
61-73	less than 60.5	19	31.67
74-86	less than 73.5	31	51.67
87-99	less than 86.5	54	90
100+	less than 99.5	60	100

More Than Cumulative Frequency Distribution Table

Class Interval	Class Boundaries	Cumulative Frequency (More Than)	Cumulative Percentage
9-21	more than 8.5	60	100
22-34	more than 21.5	57	95
35-47	more than 34.5	53	88.33
48-60	more than 47.5	49	81.66
61-73	more than 60.5	41	68.33
74-86	more than 73.5	29	48.33
87-99	more than 86.5	6	10
100+	more than 99.5	0	0

Creating Graphs

Definition:

Creating frequency distributions in the form of visual representations such as bar charts, histograms, polygons, histograms, or ogives.

Key Terms in Data Presentation:

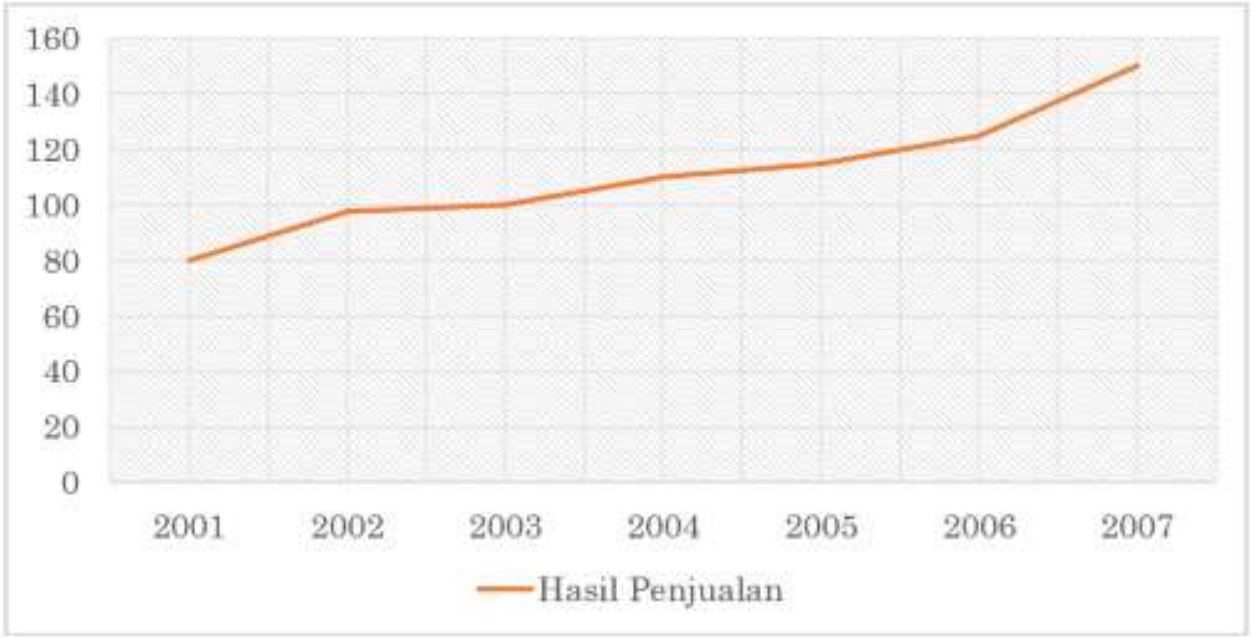
1. **Class Boundaries:** The lowest and highest values in a class.
2. **Class Midpoint:** The value located at the middle of the class.
3. **Class Edges:** Values that separate one class from another.
4. **Cumulative Frequency:** The sum of frequencies for each class, increasing or decreasing.

Types of Graphs

- Line Chart
- Bar Chart/Histogram
- Pie Chart
- Pictogram
- Cartogram (Map-Based Graph)

Line Chart

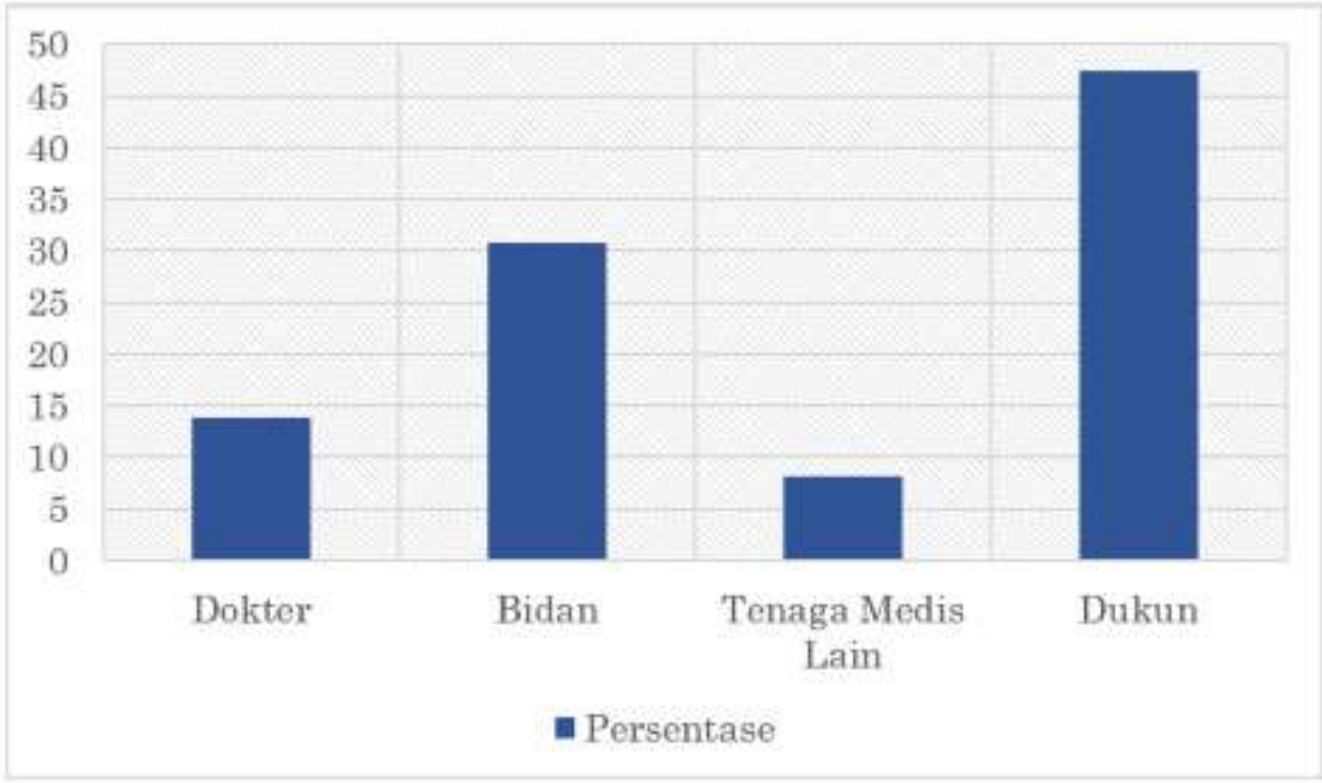
- A graph consisting of a single line to represent trends or developments of a characteristic over time.
- **Example:** Sales of "Harapan Kita" Company from 2001-2007 in million Rupiahs.



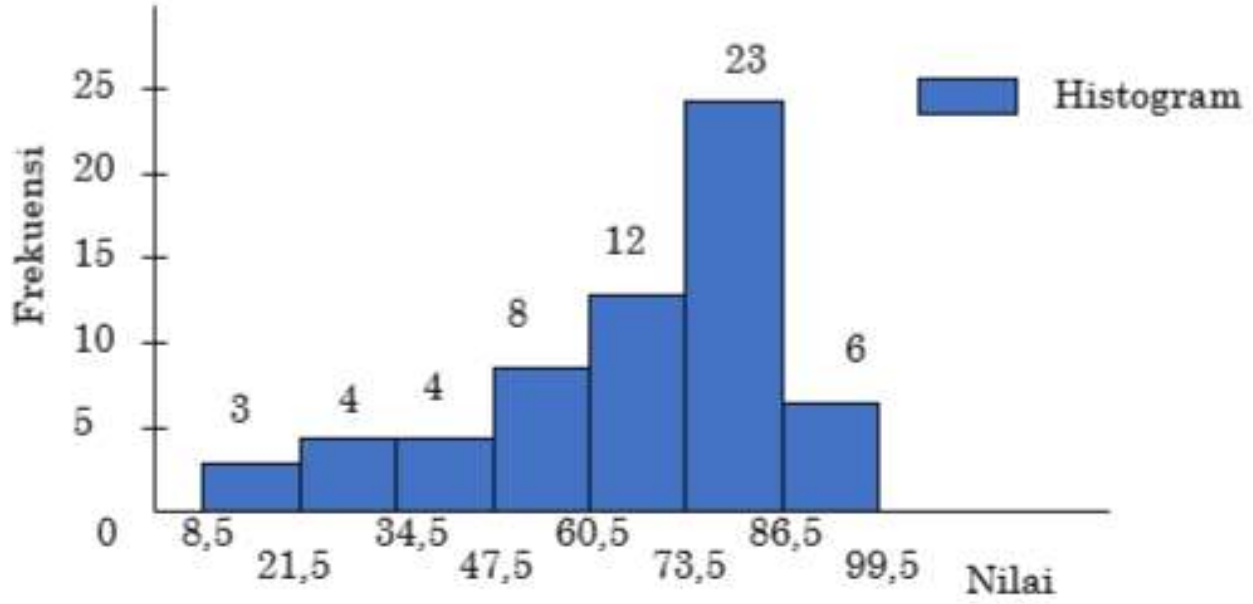
Gambar: Sumber : Supranto J. , M.A. 2008. Statistik Teori dan Aplikasi. Erlangga. Jakarta

Bar Chart

- A bar chart consists of one (single) bar to illustrate the development of an event or occurrence.
- Example: The percentage of infant births by birth attendants in Indonesia in 1999.

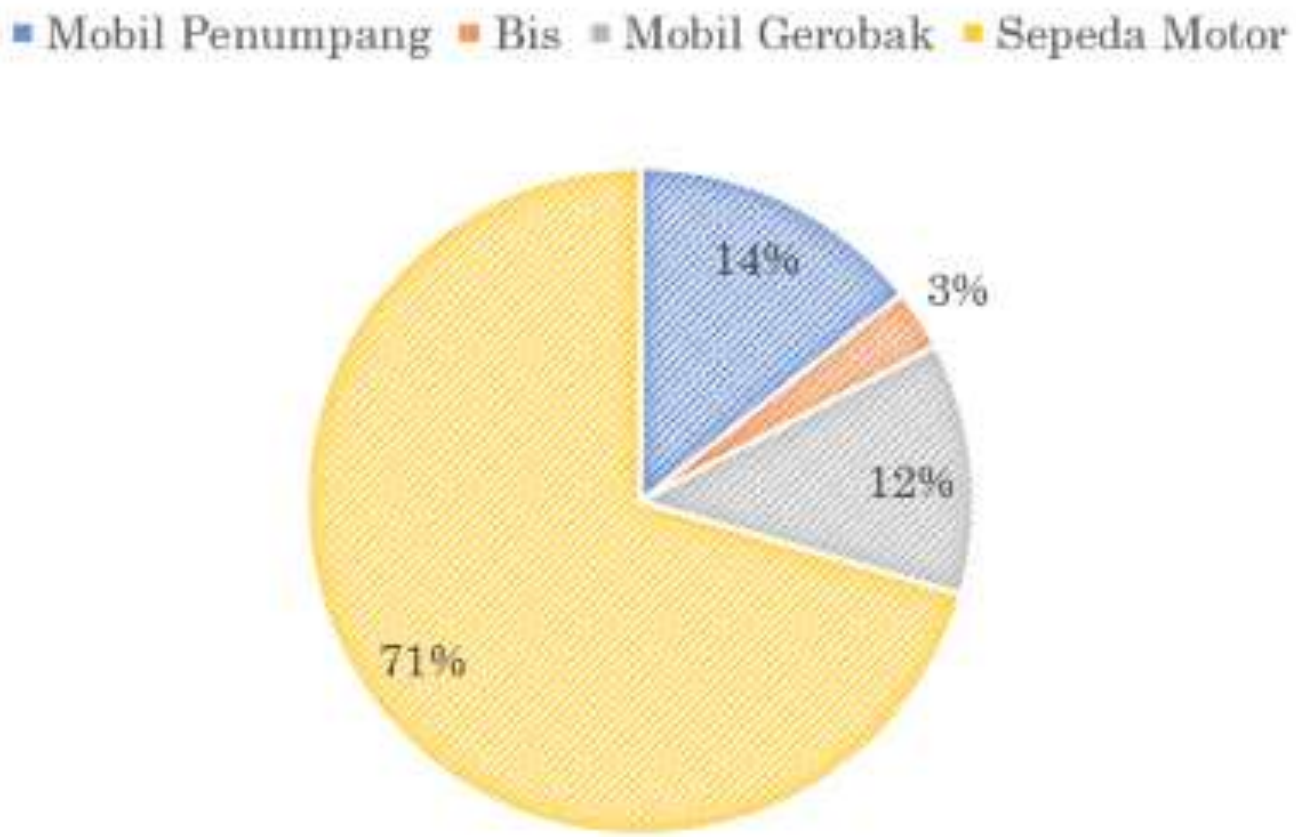


Histogram


Frekuensi
25
20
15
10
5
0
3
4
4
8
12
23
6
Histogram
8,5
21,5
34,5
47,5
60,5
73,5
86,5
99,5
Nilai

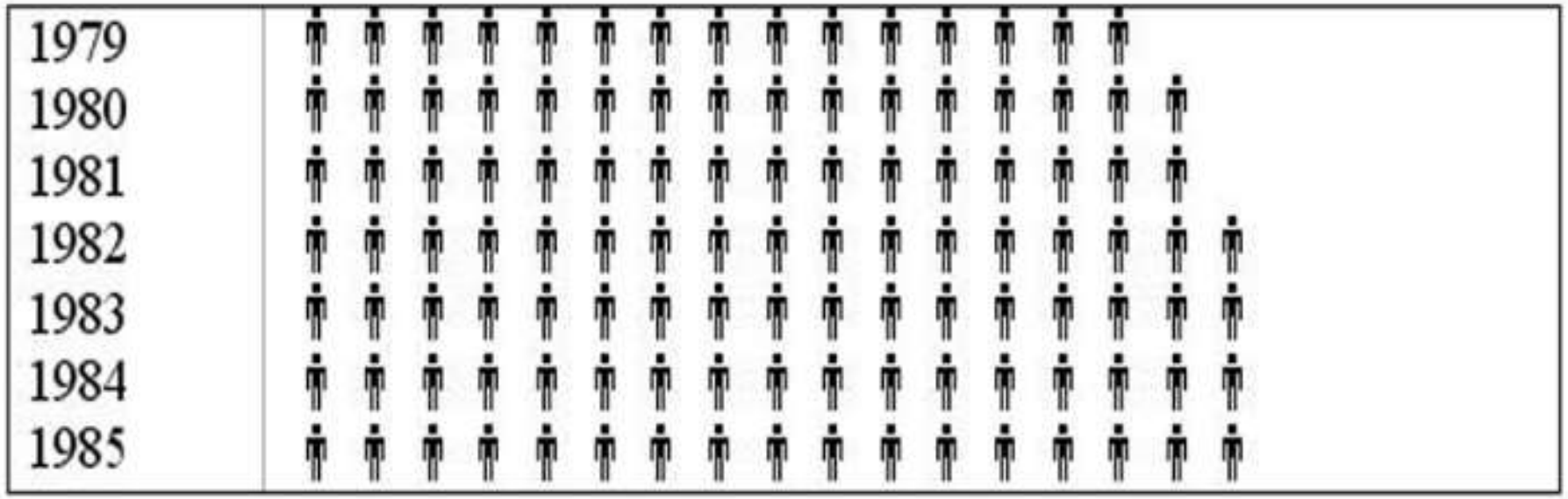
Pie Chart

- A pie chart consists of a single circle.
- Example: The number of motorized vehicles in Indonesia in 1999.



Pictogram

- A chart that is presented in the form of images.
- Example: The population growth represented with human figure symbols for different years.




THANK
YOU