

Statistics

Lecture 08: | Frequency Table

Prof. Dr. Mostafa Elhosseini

<https://youtube.com/drmelhosseini>

| Agenda

- Frequency Table
- What - How - Why
- Proportion – Percentage
- Frequency Tables and Types of Data
- How to calculate Summary Statistics using Frequency Table

| Defect Counts in a Manufacturing Line

- Number of Defective Products per Day (over 50 days)

0, 1, 2, 0, 1, 3, 2, 2, 1, 0,
4, 3, 2, 1, 0, 2, 3, 4, 2, 3,
1, 0, 5, 3, 2, 4, 2, 1, 2, 3,
4, 3, 0, 2, 1, 2, 4, 5, 3, 2,
2, 1, 0, 2, 3, 4, 1, 0, 1, 2

- How often were there no defects at all?
- Which number of defects happened most frequently?
- How many days had 4 or more defects?
- Is the process mostly under control or not?

| Defect Counts in a Manufacturing Line

- Number of Defective Products per Day (over 50 days)

0, 1, 2, 0, 1, 3, 2, 2, 1, 0,
4, 3, 2, 1, 0, 2, 3, 4, 2, 3,
1, 0, 5, 3, 2, 4, 2, 1, 2, 3,
4, 3, 0, 2, 1, 2, 4, 5, 3, 2,
2, 1, 0, 2, 3, 4, 1, 0, 1, 2

Defects per Day	Frequency
0	8
1	10
2	15
3	9
4	6
5	2

✓ The Insights

- Zero defects: occurred on 8 days.
- Most frequent value: 2 defects/day (15 days).
- High-defect days (4 or more): only 8 days out of 50.
- The process shows signs of quality control—most days are low-defect.

| Delivery Times at a Logistics Company

■ Delivery Times (in hours) for 50 Packages

```
2, 3, 5, 7, 6, 4, 3, 5, 6, 8,  
10, 12, 11, 9, 8, 10, 11, 12, 13, 14,  
7, 6, 9, 10, 11, 12, 13, 14, 15, 16,  
5, 7, 6, 4, 6, 5, 7, 8, 9, 10,  
8, 7, 6, 5, 4, 3, 2, 2, 3, 4
```

- What is the most common delivery time?
- How many deliveries took more than 10 hours?
- Can you estimate how often deliveries took fewer than 5 hours?
- Is this data spread evenly or concentrated in certain values?

| Delivery Times at a Logistics Company

■ Delivery Times (in hours) for 50 Packages

2, 3, 5, 7, 6, 4, 3, 5, 6, 8,
10, 12, 11, 9, 8, 10, 11, 12, 13, 14,
7, 6, 9, 10, 11, 12, 13, 14, 15, 16,
5, 7, 6, 4, 6, 5, 7, 8, 9, 10,
8, 7, 6, 5, 4, 3, 2, 2, 3, 4



The Insights

- The most common delivery times are 6 and 7 hours.
- Deliveries over 10 hours occurred 12 times (values: 11–16).
- Deliveries under 5 hours = 11 deliveries (values: 2–4)

Delivery Time (hrs)	Frequency
2	3
3	4
4	4
5	5
6	6
7	5
8	4
9	3
10	4
11	3
12	3
13	2
14	2
15	1
16	1

| Problem 01

■ Server Response Time (ms) for 50 Web Requests

```
120, 200, 180, 150, 220, 190, 170, 160, 210, 200,  
230, 250, 270, 200, 180, 190, 220, 240, 250, 260,  
180, 190, 210, 170, 150, 140, 130, 200, 220, 230,  
260, 270, 280, 240, 230, 190, 200, 180, 160, 150,  
250, 260, 270, 280, 190, 210, 200, 150, 180, 170
```

- What is the most common server response time?
- How many responses were under 160 ms?
- Were there any slow responses (above 250 ms)? Is the server consistent in its performance?

| Problem 02

■ Number of Goals Scored by a Player in 50 Matches

0, 1, 2, 1, 0, 1, 2, 3, 1, 0,
2, 2, 3, 1, 0, 0, 1, 2, 1, 1,
0, 1, 2, 3, 0, 1, 1, 2, 2, 0,
0, 1, 2, 2, 3, 1, 0, 1, 2, 1,
1, 2, 3, 1, 0, 1, 2, 2, 1, 0

- How often did the player score no goals?
- What's the most frequent number of goals per match?
- How many matches included 2 or more goals?
- Can you quickly spot consistency or standout performances?

| Frequency Table

- **Frequency table** is a tool used to organize and summarize data by showing **how often** each value or range of values (class) occurs in a dataset.
- Displays the **distribution** of a variable numerically
 - listing of **possible values** for a variable, together with the **number of observations** for each value
 - For a **categorical** variable, a frequency table lists the categories and the number of times each category was observed

| Proportion

- A **proportion** is a number that represents a part of a whole.
- It is usually expressed as a decimal or a fraction.
- Formula:

$$\text{Proportion} = \frac{\text{Part}}{\text{Total}}$$

Example:

- If 8 out of 40 students read 2 books per week:

$$\text{Proportion} = \frac{8}{40} = 0.20$$

| Percentage

- A **percentage** is a proportion multiplied by 100 to express the value out of 100.
- It is usually followed by the % symbol.
- Formula:

$$\text{Percentage} = \text{Proportion} \times 100$$

Example:

- If 8 out of 40 students read 2 books per week:

$$\text{Proportion} = \frac{8}{40} = 0.20$$

$$\text{Percentage} = 0.20 \times 100 = 20\%$$

| Uses of a Frequency Table

- Simplifies large data sets for better understanding.
- Helps in constructing histograms or bar charts.
- Supports statistical analysis like mode identification or median estimation.
- Essential for calculating descriptive statistics (mean, variance, etc.).

| Example: Library Book Genres Survey

- A librarian surveyed 40 visitors about their favorite genre of books. The results are summarized in the table below:

Genre	Frequency
Fiction	12
Mystery	8
Science	6
History	7
Biography	4
Poetry	3

- Complete the table by calculating the proportion and percentage for each genre.
- Which genre was most popular and least popular?
- What proportion and percentage of visitors preferred non-fiction genres (Science, History, Biography, Poetry)?

| Example: Library Book Genres Survey

- Frequency Table with Proportion and Percentage

Genre	Frequency (f)	Proportion (f/40)	Percentage (%)
Fiction	12	0.30	30.0%
Mystery	8	0.20	20.0%
Science	6	0.15	15.0%
History	7	0.175	17.5%
Biography	4	0.10	10.0%
Poetry	3	0.075	7.5%
Total	40	1.000	100%

- Most popular: Fiction (12 responses → 30%)
- Least popular: Poetry (3 responses → 7.5%)

- Genres: Science, History, Biography, Poetry
- Frequencies: $6 + 7 + 4 + 3 = 20$
- 50% of visitors preferred non-fiction genres.



Example

- An electronics technician tested **30 electronic devices** and recorded the **number of capacitors** found in each device. The recorded values (number of capacitors per device) are:

```
3, 4, 4, 5, 5, 5, 6, 6, 6, 6,  
7, 7, 7, 8, 8, 8, 8, 8, 9, 9,  
9, 10, 10, 11, 11, 11, 12, 12, 13, 14
```

Construct a frequency table



Answer: Frequency Table

3, 4, 4, 5, 5, 5, 6, 6, 6, 6,
7, 7, 7, 8, 8, 8, 8, 8, 9, 9,
9, 10, 10, 11, 11, 11, 12, 12, 13, 14

Capacitors per Device	Frequency	Proportion	Percentage
3	1	0.033	3.3%
4	2	0.067	6.7%
5	3	0.100	10.0%
6	4	0.133	13.3%
7	3	0.100	10.0%
8	5	0.167	16.7%
9	3	0.100	10.0%
10	2	0.067	6.7%
11	3	0.100	10.0%
12	2	0.067	6.7%
13	1	0.033	3.3%
14	1	0.033	3.3%
Total	30	1.000	100%



Example

- The following are the test scores (out of 100) of **30** students in a statistics exam:

12, 18, 27, 31, 35, 39, 43, 45, 48, 52,
54, 55, 58, 61, 63, 66, 68, 71, 73, 75,
78, 80, 83, 85, 88, 91, 93, 95, 97, 99

Create a frequency table using class intervals of width 20



Answer: Frequency Table

12, 18, 27, 31, 35, 39, 43, 45, 48, 52,
54, 55, 58, 61, 63, 66, 68, 71, 73, 75,
78, 80, 83, 85, 88, 91, 93, 95, 97, 99

Score Interval	Frequency	Proportion	Percentage
0 - 20	2	0.067	6.7%
20 - 40	4	0.133	13.3%
40 - 60	7	0.233	23.3%
60 - 80	7	0.233	23.3%
80 - 100	10	0.334	33.4%
Total	30	1.000	100%



Example

- A logistics company measured the weights (in kilograms) of 30 delivered packages:

1.2, 1.5, 2.3, 2.8, 2.9, 3.1, 3.4, 3.6, 3.9, 4.0,
4.2, 4.3, 4.5, 4.7, 4.9, 5.1, 5.3, 5.5, 5.6, 5.8,
6.0, 6.1, 6.4, 6.7, 6.8, 7.0, 7.2, 7.5, 7.8, 8.0

Construct a frequency table using equal class intervals of width 2.0 kg



Answer: Frequency Table

1.2, 1.5, 2.3, 2.8, 2.9, 3.1, 3.4, 3.6, 3.9, 4.0,
4.2, 4.3, 4.5, 4.7, 4.9, 5.1, 5.3, 5.5, 5.6, 5.8,
6.0, 6.1, 6.4, 6.7, 6.8, 7.0, 7.2, 7.5, 7.8, 8.0

Weight Interval (kg)	Frequency	Proportion	Percentage
1.0 - 3.0	5	0.167	16.7%
3.0 - 5.0	10	0.333	33.3%
5.0 - 7.0	9	0.300	30.0%
7.0 - 9.0	6	0.200	20.0%
Total	30	1.000	100%



Example

- A survey was conducted among 30 computer science students to identify their preferred programming language.

```
Python, Java, Python, C++, JavaScript, Python, C++, Java, Python, Python,  
Java, C++, C++, JavaScript, Python, Java, Python, C#, Java, C++,  
Python, C#, JavaScript, Java, JavaScript, Python, C#, Java, Python, Python
```

Construct a frequency table for the nominal data



Answer: Frequency Table (Nominal)

Python, Java, Python, C++, JavaScript, Python, C++, Java, Python, Python,
Java, C++, C++, JavaScript, Python, Java, Python, C#, Java, C++,
Python, C#, JavaScript, Java, JavaScript, Python, C#, Java, Python, Python

Programming Language	Frequency	Proportion	Percentage
Python	11	0.367	36.7%
Java	7	0.233	23.3%
C++	5	0.167	16.7%
JavaScript	4	0.133	13.3%
C#	3	0.100	10.0%
Total	30	1.000	100%



Example

- A quality control manager in a manufacturing plant asked 25 workers to rate machine maintenance quality on a 5-point ordinal scale:
Poor, Fair, Good, Very Good, Excellent.

Good, Fair, Good, Very Good, Excellent, Good, Good, Fair, Excellent, Good,
Very Good, Poor, Good, Fair, Good, Excellent, Fair, Good, Good, Very Good,
Fair, Poor, Excellent, Good, Very Good

Construct a frequency table for the ordinal data



Answer: Frequency Table (Ordinal)

Good, Fair, Good, Very Good, Excellent, Good, Good, Fair, Excellent, Good,
Very Good, Poor, Good, Fair, Good, Excellent, Fair, Good, Good, Very Good,
Fair, Poor, Excellent, Good, Very Good

Maintenance Rating	Frequency	Proportion	Percentage
Poor	2	0.080	8.0%
Fair	5	0.200	20.0%
Good	10	0.400	40.0%
Very Good	5	0.200	20.0%
Excellent	3	0.120	12.0%
Total	25	1.000	100%

| Summary statistics

- How to calculate summary statistics — mean, median, and mode — from a frequency table.

| Summary statistics

■ Number of Books Read per Week.

Total number of people (N):

$$N = 4 + 8 + 10 + 6 + 2 = 30$$

Books per Week (x)	Frequency (f)
0	4
1	8
2	10
3	6
4	2

1. Mean (Average)

Use the formula:

$$\begin{aligned}\bar{x} &= \frac{\sum(x \cdot f)}{N} \\ &= \frac{(0 \cdot 4) + (1 \cdot 8) + (2 \cdot 10) + (3 \cdot 6) + (4 \cdot 2)}{30} = \frac{0 + 8 + 20 + 18 + 8}{30} = \frac{54}{30} = 1.8\end{aligned}$$

Mean = 1.8 books/week

| Summary statistics

- Number of Books Read per Week.

Median

- Find the **cumulative frequency**:

Books per Week (x)	Frequency (f)
0	4
1	8
2	10
3	6
4	2

Books per Week (x)	Frequency (f)	Cumulative Frequency
0	4	4
1	8	12
2	10	22
3	6	28
4	2	30

| Summary statistics

- Number of Books Read per Week.

Median

- Find the **cumulative frequency**:

Books per Week (x)	Frequency (f)	Cumulative Frequency
0	4	4
1	8	12
2	10	22
3	6	28
4	2	30

Books per Week (x)	Frequency (f)
0	4
1	8
2	10
3	6
4	2

- The median position is:

$$\frac{N + 1}{2} = \frac{30 + 1}{2} = 15.5$$

So, the 15th and 16th values fall in the cumulative range of 13–22, which corresponds to $x = 2$. Median = **2 books/week**

| Summary statistics

- Number of Books Read per Week.

Books per Week (x)	Frequency (f)	Cumulative Frequency
0	4	4
1	8	12
2	10	22
3	6	28
4	2	30

Books per Week (x)	Frequency (f)
0	4
1	8
2	10
3	6
4	2

- The median position is:

$$\frac{N + 1}{2} = \frac{30 + 1}{2} = 15.5$$

So, the 15th and 16th values fall in the cumulative range of 13–22, which corresponds to $x = 2$. Median = **2 books/week**

| Summary statistics

Time Spent on a Website.

- A company recorded how long users spent on their website across 60 sessions.

Time Interval (minutes)	Frequency (f)
0 - 10	6
10 - 20	10
20 - 30	16
30 - 40	14
40 - 50	8
50 - 60	6

| Summary statistics

✓ Mean (Estimated)

Interval	Midpoint (x)	Frequency (f)	$x \cdot f$
0 - 10	5	6	30
10 - 20	15	10	150
20 - 30	25	16	400
30 - 40	35	14	490
40 - 50	45	8	360
50 - 60	55	6	330
Total		60	1760

$$\bar{x} = \frac{\sum(x \cdot f)}{N} = \frac{1760}{60} = 29.33$$

Estimated Mean = 29.33 minutes

| Summary statistics

✓ Median (Estimated)

Interval	Frequency (f)	Cumulative Frequency
0 - 10	6	6
10 - 20	10	16
20 - 30	16	32
30 - 40	14	46
40 - 50	8	54
50 - 60	6	60

- Median position = $N/2 = 30$
- The 30th value is in the interval 20 - 30 (since cumulative freq = 32).
- Use the formula:

$$\text{Median} = L + \left(\frac{\frac{N}{2} - F}{f_m} \right) \cdot h$$

| Summary statistics

✓ Median (Estimated)

Interval	Frequency (f)	Cumulative Frequency
0 - 10	6	6
10 - 20	10	16
20 - 30	16	32
30 - 40	14	46
40 - 50	8	54
50 - 60	6	60

$$\text{Median} = 20 + \left(\frac{30 - 16}{16} \right) \cdot 10 = 20 + (0.875 \cdot 10) = 20 + 8.75 = 28.75$$

$$\text{Median} = L + \left(\frac{\frac{N}{2} - F}{f_m} \right) \cdot h$$

- L = lower boundary of median class = 20
- F = cumulative frequency before median class = 16
- f_m = frequency of median class = 16
- h = class width = 10

| Summary statistics

Mode (Estimated)

- Find the modal class
 - The class with the highest frequency = 20 - 30 ($f = 16$)
- Use the formula:

$$\text{Mode} = L + \left(\frac{f_1 - f_0}{(2f_1 - f_0 - f_2)} \right) \cdot h$$

| Summary statistics

✓ Mode (Estimated)

$$\text{Mode} = L + \left(\frac{f_1 - f_0}{(2f_1 - f_0 - f_2)} \right) \cdot h$$

- L = lower boundary of modal class = 20
- f_1 = frequency of modal class = 16
- f_0 = frequency of class before modal = 10
- f_2 = frequency of class after modal = 14
- h = class width = 10

Interval	Frequency (f)	Cumulative Frequency
0 - 10	6	6
10 - 20	10	16
20 - 30	16	32
30 - 40	14	46
40 - 50	8	54
50 - 60	6	60

$$\text{Mode} = 20 + \left(\frac{16 - 10}{2 \cdot 16 - 10 - 14} \right) \cdot 10 = 20 + \left(\frac{6}{32 - 24} \right) \cdot 10 = 20 + \left(\frac{6}{8} \right) \cdot 10 = 20 + 7.5 = 27.5$$