

Chapter 2

Frequency Distributions

Likan Zhan

Beijing Language and Culture University

2021-09-12

<https://likan.info>

zhanlikan@bncu.edu.cn

Table of Contents

1. Frequency Distributions
2. Grouped Frequency Distribution Tables
3. Frequency Distribution Graphs
4. Percentiles, Percentile Ranks, and Interpolation
5. Stem and Leaf Displays

Table of Contents

1. Frequency Distributions
2. Grouped Frequency Distribution Tables
3. Frequency Distribution Graphs
4. Percentiles, Percentile Ranks, and Interpolation
5. Stem and Leaf Displays

Frequency distribution

Frequency distribution

- The goal of descriptive statistics is to simplify the organization and presentation of data.

Frequency distribution

- The goal of descriptive statistics is to simplify the organization and presentation of data.
- One of the most common procedures for organizing a set of data is to place the scores in a frequency distribution.

Frequency distribution

- The goal of descriptive statistics is to simplify the organization and presentation of data.
- One of the most common procedures for organizing a set of data is to place the scores in a frequency distribution.
- A **frequency distribution** (频率分布) is an organized tabulation of the number of individuals located in each category on the scale of measurement.

Frequency distribution

Frequency distribution

- A frequency distribution can be structured either as a table or as a graph, but in either case, the distribution presents the same two elements:

Frequency distribution

- A frequency distribution can be structured either as a table or as a graph, but in either case, the distribution presents the same two elements:
 - a. The set of categories that make up the original measurement scale.

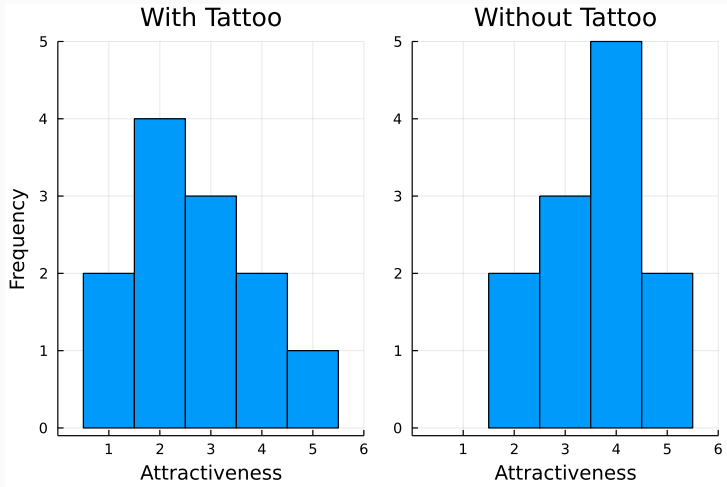
Frequency distribution

- A frequency distribution can be structured either as a table or as a graph, but in either case, the distribution presents the same two elements:
 - a. The set of categories that make up the original measurement scale.
 - b. A record of the frequency, or number of individuals in each category.

Tattoo and Attractiveness rating

```
# Attractiveness With Visible Tattoo  
x1 = [1, 2, 4, 3, 2, 2, 1, 3, 2, 5, 4, 3];  
  
# Attractiveness With No Visible Tattoo  
x2 = [2, 4, 4, 3, 5, 4, 2, 4, 4, 5, 3, 3];
```

Tattoo and Attractiveness rating



Tattoo and Attractiveness rating

```
x3 = [8, 9, 8, 7, 10, 9, 6, 4, 9, 8,  
      7, 8, 10, 9, 8, 6, 9, 7, 8, 8];
```

Tattoo and Attractiveness rating

```
x3 = [8, 9, 8, 7, 10, 9, 6, 4, 9, 8,  
      7, 8, 10, 9, 8, 6, 9, 7, 8, 8];
```

```
## 6x2 DataFrame  
##   Row   Score  Freq  
##           Int64  Int64  
##  
##    1     10     2  
##    2      9     5  
##    3      8     7  
##    4      7     3  
##    5      6     2  
##    6      4     1
```

Frequency Distribution Tables

Frequency Distribution Tables

- The following set of $N = 20$ scores was obtained from a 10-point statistics quiz.

##	[1]	8	9	8	7	10	9	6	4	9	8	7	8	10
##	[14]	9	8	6	9	7	8	8						

Frequency Distribution Tables

Scores	Frequency
10	2
9	5
8	7
7	3
6	2
5	0
4	1

Basic Computations

Basic Computations

- The total number of individuals, N

```
## [1] 20
```

Basic Computations

- The total number of individuals, N

```
## [1] 20
```

- The sum of individual scores, $\sum X$

```
## [1] 158
```

Basic Computations

- The total number of individuals, N

```
## [1] 20
```

- The sum of individual scores, $\sum X$

```
## [1] 158
```

- The sum of the squared individual scores, $\sum X^2$

```
## [1] 1288
```

Proportions and Percentages

Proportions and Percentages

- The table may include a proportion column showing the relative frequency for each category

$$\text{Proportion} = p = \frac{f}{N} \quad (1)$$

Proportions and Percentages

- The table may include a proportion column showing the relative frequency for each category

$$\text{Proportion} = p = \frac{f}{N} \quad (1)$$

- The table may include a percentage column showing the percentage associated with each X value

$$\text{Percentage} = p(100) = \frac{f}{N}(100) \quad (2)$$

Calculate Proportions and Percentages

Scores	Frequency
10	2
9	5
8	7
7	3
6	2
5	0
4	1

Calculate Proportions and Percentages

Scores	Frequency	Proportion	Percentage
10	2	0.10	10%
9	5	0.25	25%
8	7	0.35	35%
7	3	0.15	15%
6	2	0.10	10%
5	0	0.00	0%
4	1	0.05	5%

Frequency Distribution Table: A Practice

Frequency Distribution Table: A Practice

- Given the following frequency distribution table

<u>X</u>	<u>F</u>
5	1
4	2
3	3
2	3
1	1

Frequency Distribution Table: A Practice

- Given the following frequency distribution table

<u>X</u>	<u>F</u>
5	1
4	2
3	3
2	3
1	1

- $\sum X = 29$; $\sum X^2 = 97$

A Practice

A Practice

- Given the following frequency distribution table

<u>X</u>	<u>F</u>
5	1
4	2
3	3
2	3
1	1

A Practice: Proportion and Percentage

X	F	Proportion	Percentage
5	1	0.10	10%
4	2	0.20	20%
3	3	0.30	30%
2	3	0.30	30%
1	1	0.10	10%

Table of Contents

1. Frequency Distributions
2. Grouped Frequency Distribution Tables
3. Frequency Distribution Graphs
4. Percentiles, Percentile Ranks, and Interpolation
5. Stem and Leaf Displays

Grouped Frequency Distribution Tables

Grouped Frequency Distribution Tables

- It is recommended that a frequency distribution table have a maximum of 10–15 rows to keep it simple.

Grouped Frequency Distribution Tables

- It is recommended that a frequency distribution table have a maximum of 10–15 rows to keep it simple.
- If the scores cover a range that is wider than this suggested maximum, it is customary to divide the range into sections called **class intervals**.

Grouped Frequency Distribution Tables

- It is recommended that a frequency distribution table have a maximum of 10–15 rows to keep it simple.
- If the scores cover a range that is wider than this suggested maximum, it is customary to divide the range into sections called **class intervals**.
- These intervals are then listed in the frequency distribution table along with the frequency or number of individuals with scores in each interval.

Grouped Frequency Distribution Tables

- It is recommended that a frequency distribution table have a maximum of 10–15 rows to keep it simple.
- If the scores cover a range that is wider than this suggested maximum, it is customary to divide the range into sections called **class intervals**.
- These intervals are then listed in the frequency distribution table along with the frequency or number of individuals with scores in each interval.
- The result is called a **grouped frequency distribution** (分组频率分布).

The guidelines for constructing such a distribution

The guidelines for constructing such a distribution

- The grouped frequency distribution table should have about 10 class intervals.

The guidelines for constructing such a distribution

- The grouped frequency distribution table should have about 10 class intervals.
- The width of each interval should be a relatively simple number, such as 2, 5, 10, or 20.

The guidelines for constructing such a distribution

- The grouped frequency distribution table should have about 10 class intervals.
- The width of each interval should be a relatively simple number, such as 2, 5, 10, or 20.
- The bottom score in each class interval should be a multiple of the width.

The guidelines for constructing such a distribution

- The grouped frequency distribution table should have about 10 class intervals.
- The width of each interval should be a relatively simple number, such as 2, 5, 10, or 20.
- The bottom score in each class interval should be a multiple of the width.
- All intervals should be the same width.

Grouped Frequency Distribution Tables

Grouped Frequency Distribution Tables

- Suppose a set of $N = 25$ exam scores

```
## [1] 82 75 88 93 53 84 87 58 72 94 69 84 61
## [14] 91 64 87 84 70 76 89 75 80 73 78 60
```

Grouped Frequency Distribution Tables

- Suppose a set of $N = 25$ exam scores

```
## [1] 82 75 88 93 53 84 87 58 72 94 69 84 61  
## [14] 91 64 87 84 70 76 89 75 80 73 78 60
```

- The range of the data is

```
## [1] 53 94
```

Grouped frequency distribution table

Number	Group	Frequency
22	[92,94)	1
21	[90,92)	1
20	[88,90)	2
19	[86,88)	2
18	[84,86)	3
17	[82,84)	1
16	[80,82)	1
15	[78,80)	1
14	[76,78)	1
13	[74,76)	2
12	[72,74)	2
11	[70,72)	1
10	[68,70)	1
9	[66,68)	0
8	[64,66)	1
7	[62,64)	0
6	[60,62)	2
5	[58,60)	1
4	[56,58)	0
3	[54,56)	0
2	[52,54)	1
1	[50,52)	0

Grouped frequency distribution table

Number	Group	Frequency
4	[80,90)	9
3	[70,80)	7
2	[60,70)	4
1	[50,60)	2

Grouped frequency distribution table

Number	Group	Frequency
9	[90,95)	3
8	[85,90)	4
7	[80,85)	5
6	[75,80)	4
5	[70,75)	3
4	[65,70)	1
3	[60,65)	3
2	[55,60)	1
1	[50,55)	1

Real Limits and Frequency Distributions

Real Limits and Frequency Distributions

- When a continuous variable is measured, the resulting measurements correspond to intervals on the number line rather than single points.

Table of Contents

1. Frequency Distributions
2. Grouped Frequency Distribution Tables
3. Frequency Distribution Graphs
4. Percentiles, Percentile Ranks, and Interpolation
5. Stem and Leaf Displays

Frequency Distribution Graphs

Frequency Distribution Graphs

- All frequency distribution start with two perpendicular lines called axes.

Frequency Distribution Graphs

- All frequency distribution start with two perpendicular lines called axes.
- The horizontal line is the **X**-axis, or the abscissa (ab-SIS-uh). The vertical line is the **Y**-axis, or the ordinate.

Frequency Distribution Graphs

- All frequency distribution start with two perpendicular lines called axes.
- The horizontal line is the **X**-axis, or the abscissa (ab-SIS-uh). The vertical line is the **Y**-axis, or the ordinate.
- The measurement scale (set of X values) is listed along the X-axis with values increasing from left to right.

Frequency Distribution Graphs

- All frequency distribution start with two perpendicular lines called axes.
- The horizontal line is the **X**-axis, or the abscissa (ab-SIS-uh). The vertical line is the **Y**-axis, or the ordinate.
- The measurement scale (set of X values) is listed along the X-axis with values increasing from left to right.
- The frequencies are listed on the Y-axis with values increasing from bottom to top.

Frequency Distribution Graphs

Frequency Distribution Graphs

- As a general rule, the point where the two axes intersect should have a value of zero for both the scores and the frequencies.

Frequency Distribution Graphs

- As a general rule, the point where the two axes intersect should have a value of zero for both the scores and the frequencies.
- A final general rule is that the graph should be constructed so that its height (Y-axis) is approximately two-thirds to three-quarters of its length (X-axis).

Frequency Distribution Graphs

Frequency Distribution Graphs

- A frequency distribution graph lists scores on the horizontal axis and frequencies on the vertical axis. The type of graph used to display a distribution depends on the scale of measurement used.

Frequency Distribution Graphs

- A frequency distribution graph lists scores on the horizontal axis and frequencies on the vertical axis. The type of graph used to display a distribution depends on the scale of measurement used.
- For interval or ratio scales, you should use a **histogram** (直方图) or a **polygon** (多边形).

Frequency Distribution Graphs

- A frequency distribution graph lists scores on the horizontal axis and frequencies on the vertical axis. The type of graph used to display a distribution depends on the scale of measurement used.
- For interval or ratio scales, you should use a **histogram** (直方图) or a **polygon** (多边形).
- For nominal or ordinal scales, you should use a **bar graph** (柱形图).

Histograms

Histograms

- To construct a Histograms, you first list the numerical scores (the categories of measurement) along the X-axis. Then you draw a bar above each X value so that

Histograms

- To construct a Histograms, you first list the numerical scores (the categories of measurement) along the X-axis. Then you draw a bar above each X value so that
- a. The height of the bar corresponds to the frequency for that category.

Histograms

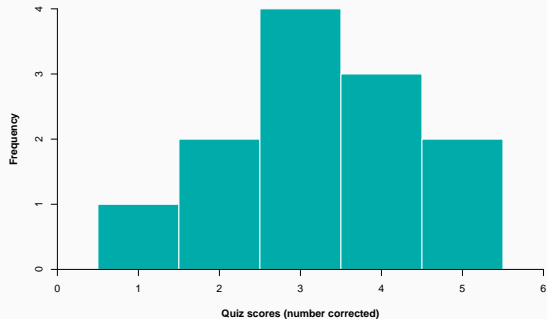
- To construct a Histograms, you first list the numerical scores (the categories of measurement) along the X-axis. Then you draw a bar above each X value so that
 - a. The height of the bar corresponds to the frequency for that category.
 - b. Each bar extends to the real limits of the score, so that adjacent bars touch.

Histograms

- A set of quiz scores

```
## [1] 5 5 4 4 4 3 3 3 3 2 2 1
```

Histograms



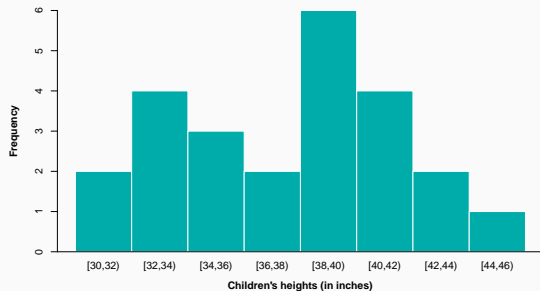
quize_scores	Freq
5	2
4	3
3	4
2	2
1	1

Histograms

- A set of height

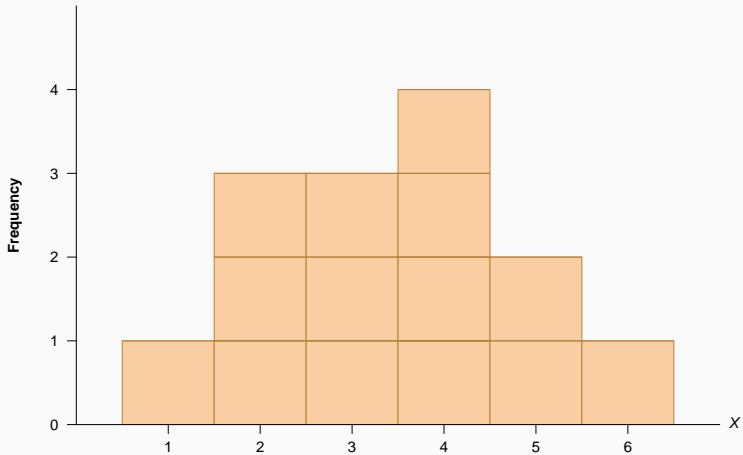
```
## [1] 44.5 42.2 42.8 40.1 40.2 40.3 40.4
## [8] 38.4 38.5 38.6 38.7 38.8 38.9 36.6
## [15] 36.8 34.2 34.3 34.5 33.0 32.1 32.2
## [22] 32.3 30.1 30.5
```


Histograms



	group	Freq
1	[44,46)	1
2	[42,44)	2
3	[40,42)	4
4	[38,40)	6
5	[36,38)	2
6	[34,36)	3
7	[32,34)	4
8	[30,32)	2

A Modified Histogram



Polygons

Polygons

- To construct a polygon Polygons, you begin by listing the numerical scores (the categories of measurement) along the X-axis. Then,

Polygons

- To construct a polygon Polygons, you begin by listing the numerical scores (the categories of measurement) along the X-axis. Then,
- a. A dot is centered above each score so that the vertical position of the dot corresponds to the frequency for the category.

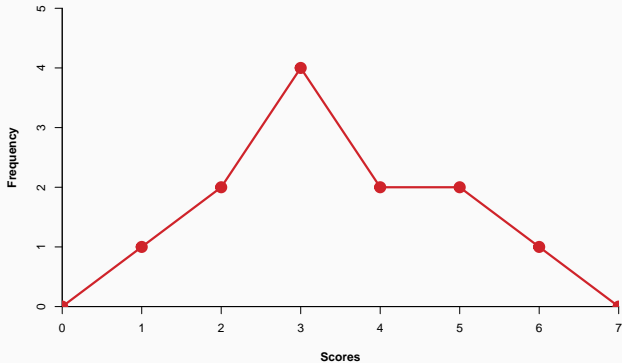
Polygons

- To construct a polygon Polygons, you begin by listing the numerical scores (the categories of measurement) along the X-axis. Then,
 - a. A dot is centered above each score so that the vertical position of the dot corresponds to the frequency for the category.
 - b. A continuous line is drawn from dot to dot to connect the series of dots.

Polygons

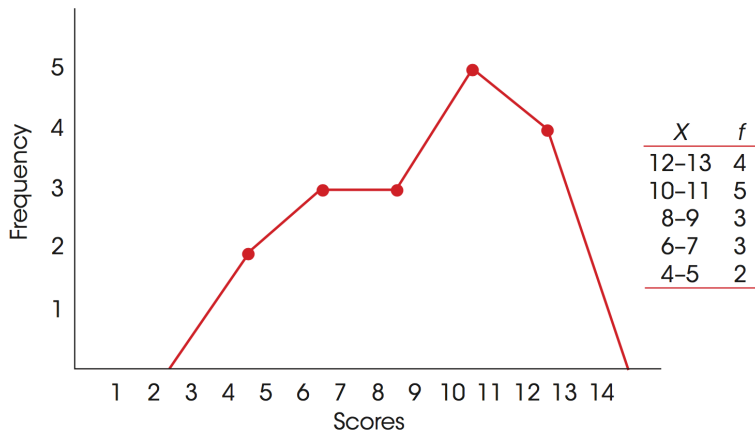
- To construct a polygon Polygons, you begin by listing the numerical scores (the categories of measurement) along the X-axis. Then,
 - a. A dot is centered above each score so that the vertical position of the dot corresponds to the frequency for the category.
 - b. A continuous line is drawn from dot to dot to connect the series of dots.
 - c. The graph is completed by drawing a line down to the X-axis (zero frequency) at each end of the range of scores.

A frequency distribution polygon



X	f
0	0
1	1
2	2
3	4
4	2
5	2
6	1
7	0

A frequency distribution polygon for grouped data



Bar Graphs

Bar Graphs

- A bar graph is essentially the same as a histogram, except that spaces are left between adjacent bars.

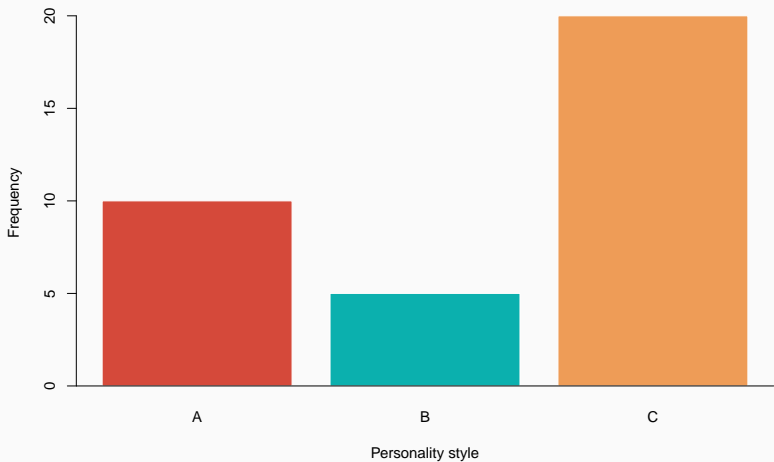
Bar Graphs

- A bar graph is essentially the same as a histogram, except that spaces are left between adjacent bars.
- For a nominal scale, the space between bars emphasizes that the scale consists of separate, distinct categories.

Bar Graphs

- A bar graph is essentially the same as a histogram, except that spaces are left between adjacent bars.
- For a nominal scale, the space between bars emphasizes that the scale consists of separate, distinct categories.
- For ordinal scales, separate bars are used because you cannot assume that the categories are all the same size.

A bar graph



Graphs for Population Distributions

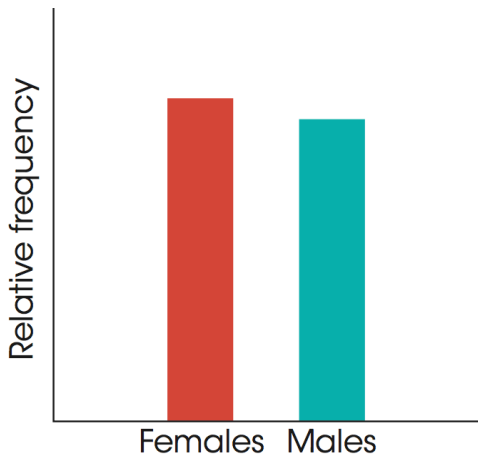
Graphs for Population Distributions

- Although it is still possible to construct graphs showing frequency distributions for extremely large populations, the graphs usually involve two special features: **relative frequencies** (相对频率) and **smooth curves** (平滑曲线).

Graphs for Population Distributions

- Although it is still possible to construct graphs showing frequency distributions for extremely large populations, the graphs usually involve two special features: **relative frequencies** (相对频率) and **smooth curves** (平滑曲线).
- Although you usually cannot find the absolute frequency for each score in a population, you very often can obtain relative frequencies.

Relative frequencies



Smooth curves

Smooth curves

- When a population consists of numerical scores from an interval or a ratio scale, it is customary to draw the distribution with a smooth curve instead of the jagged, step-wise shapes that occur with histograms and polygons.

Smooth curves

- When a population consists of numerical scores from an interval or a ratio scale, it is customary to draw the distribution with a smooth curve instead of the jagged, step-wise shapes that occur with histograms and polygons.
- The smooth curve indicates that you are not connecting a series of dots (real frequencies) but instead are showing the relative changes that occur from one score to the next.

Smooth curves

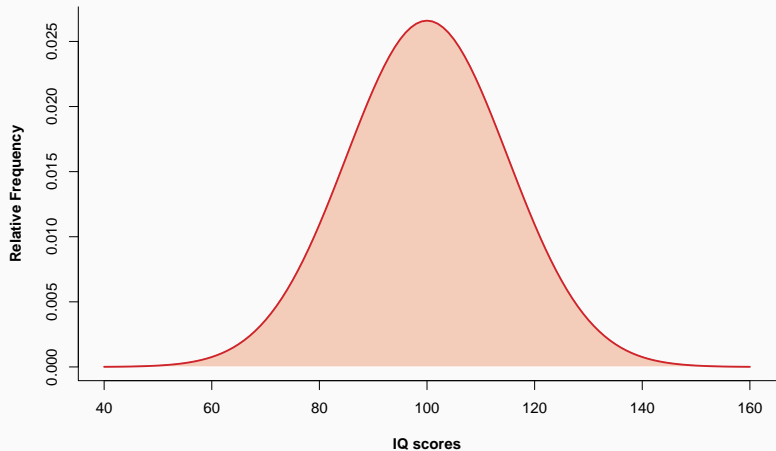
Smooth curves

- One commonly occurring population distribution is the **normal curve**.

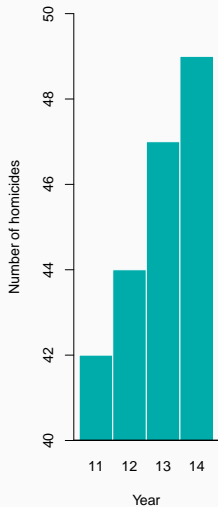
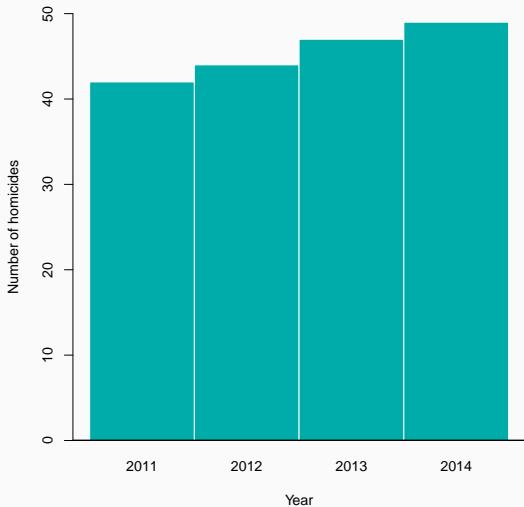
Smooth curves

- One commonly occurring population distribution is the **normal curve**.
- In the future, we will be referring to **distributions** of scores. Whenever the term distribution appears, you should conjure up an image of a frequency distribution graph.

The normal curve



The Use and Misuse Of Graphs



The Shape of a Frequency Distribution

The Shape of a Frequency Distribution

- Rather than drawing a complete frequency distribution graph, researchers often simply describe a distribution by listing its characteristics.

The Shape of a Frequency Distribution

- Rather than drawing a complete frequency distribution graph, researchers often simply describe a distribution by listing its characteristics.
- There are three characteristics that completely describe any distribution: A **shape** (形状), A **central tendency** (集中趋势), and A **variability** (变异性).

The Shape of a Frequency Distribution

The Shape of a Frequency Distribution

- In simple terms, **central tendency** measures where the center of the distribution is located and **variability** measures the degree to which the scores are spread over a wide range or are clustered together.

The Shape of a Frequency Distribution

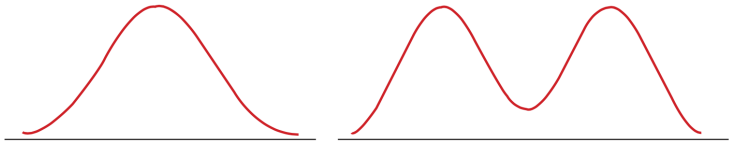
- In simple terms, **central tendency** measures where the center of the distribution is located and **variability** measures the degree to which the scores are spread over a wide range or are clustered together.
- Technically, the shape of a distribution is defined by an equation that prescribes the exact relationship between each X and Y value on the graph.

The Shape of a Frequency Distribution

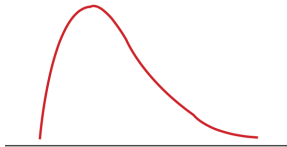
- In simple terms, **central tendency** measures where the center of the distribution is located and **variability** measures the degree to which the scores are spread over a wide range or are clustered together.
- Technically, the shape of a distribution is defined by an equation that prescribes the exact relationship between each X and Y value on the graph.
- However, we will rely on a few less-precise terms that serve to describe the shape of most distributions.

The Shape of a Frequency Distribution

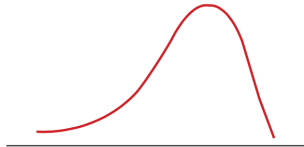
Symmetrical distributions



Skewed distributions



Positive skew



Negative skew

The Shape of a Frequency Distribution

The Shape of a Frequency Distribution

- In a **symmetrical distribution** (对称分布), it is possible to draw a vertical line through the middle so that one side of the distribution is a mirror image of the other.

The Shape of a Frequency Distribution

- In a **symmetrical distribution** (对称分布), it is possible to draw a vertical line through the middle so that one side of the distribution is a mirror image of the other.
- In a **skewed distribution** (偏态分布), the scores tend to pile up toward one end of the scale and taper off gradually at the other end.

The Shape of a Frequency Distribution

The Shape of a Frequency Distribution

- The section where the scores taper off toward one end of a distribution is called the **tail of the distribution** (分布的尾巴).

The Shape of a Frequency Distribution

- The section where the scores taper off toward one end of a distribution is called the **tail of the distribution** (分布的尾巴).
- A skewed distribution with the tail on the right-hand side is **positively skewed** (正偏态分布) because the tail points toward the positive (above-zero) end of the X-axis. If the tail points to the left, the distribution is **negatively skewed** (负偏态分布).

Table of Contents

1. Frequency Distributions
2. Grouped Frequency Distribution Tables
3. Frequency Distribution Graphs
4. Percentiles, Percentile Ranks, and Interpolation
5. Stem and Leaf Displays

Percentile and percentile rank

Percentile and percentile rank

- Percentile and percentile rank are used to describe the position of individual scores within a distribution.

Percentile and percentile rank

- Percentile and percentile rank are used to describe the position of individual scores within a distribution.
- The rank or **percentile rank** (百分比等级) of a particular score is defined as the percentage of individuals in the distribution with scores at or below the particular value.

Percentile and percentile rank

- Percentile and percentile rank are used to describe the position of individual scores within a distribution.
- The rank or **percentile rank** (百分比等级) of a particular score is defined as the percentage of individuals in the distribution with scores at or below the particular value.
- When a score is identified by its percentile rank, the score is called a **percentile** (百分位数).

Percentile and percentile rank

- Percentile and percentile rank are used to describe the position of individual scores within a distribution.
- The rank or **percentile rank** (百分比等级) of a particular score is defined as the percentage of individuals in the distribution with scores at or below the particular value.
- When a score is identified by its percentile rank, the score is called a **percentile** (百分位数).
- Notice that percentile rank refers to a percentage and that percentile refers to a score. Also notice that your rank or percentile describes your exact position within the distribution.

Cumulative Frequency and Cumulative Percentage

Cumulative Frequency and Cumulative Percentage

- To determine percentiles or percentile ranks, the first step is to find the number of individuals who are located at or below each point in the distribution.

Cumulative Frequency and Cumulative Percentage

- To determine percentiles or percentile ranks, the first step is to find the number of individuals who are located at or below each point in the distribution.
- This can be done most easily with a frequency distribution table by simply counting the number of scores that are in or below each category on the scale.

Cumulative Frequency and Cumulative Percentage

- To determine percentiles or percentile ranks, the first step is to find the number of individuals who are located at or below each point in the distribution.
- This can be done most easily with a frequency distribution table by simply counting the number of scores that are in or below each category on the scale.
- The resulting values are called **cumulative frequencies** because they represent the accumulation of individuals as you move up the scale.

Cumulative Frequency and Cumulative Percentage

Cumulative Frequency and Cumulative Percentage

- The cumulative frequencies show the number of individuals located at or below each score.

Cumulative Frequency and Cumulative Percentage

- The cumulative frequencies show the number of individuals located at or below each score.
- To find percentiles, we must convert these frequencies into percentages. The resulting values are called **cumulative percentages** because they show the percentage of individuals who are accumulated as you move up the scale.

Cumulative Frequency and Cumulative Percentage

- The cumulative frequencies show the number of individuals located at or below each score.
- To find percentiles, we must convert these frequencies into percentages. The resulting values are called **cumulative percentages** because they show the percentage of individuals who are accumulated as you move up the scale.
- The cumulative percentage values are associated with the upper real limits of the corresponding scores or intervals.

Cumulative Frequency and Cumulative Percentage

X	F
5	1
4	5
3	8
2	4
1	2

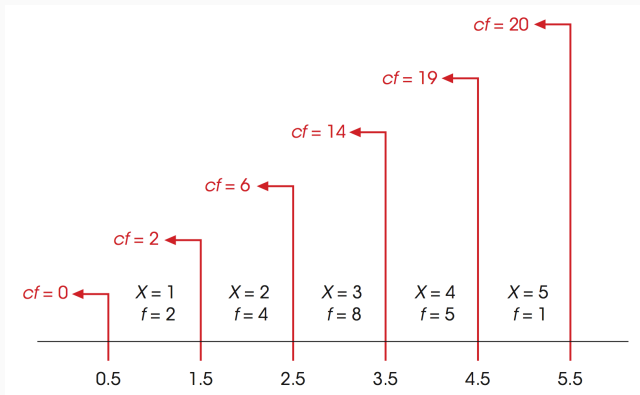
Cumulative Frequency and Cumulative Percentage

X	F	cf
5	1	20
4	5	19
3	8	14
2	4	6
1	2	2

Cumulative Frequency and Cumulative Percentage

X	F	cf	c%
5	1	20	100%
4	5	19	95%
3	8	14	70%
2	4	6	30%
1	2	2	10%

Cumulative frequencies and upper real limits



Cumulative Frequency and Cumulative Percentage

Cumulative Frequency and Cumulative Percentage

- What is the 95th percentile?

Cumulative Frequency and Cumulative Percentage

- What is the 95th percentile? (Answer: $X = 4.5$.)

Cumulative Frequency and Cumulative Percentage

- What is the 95th percentile? (Answer: $X = 4.5$.)
- What is the percentile rank for $X = 3.5$?

Cumulative Frequency and Cumulative Percentage

- What is the 95th percentile? (Answer: $X = 4.5$.)
- What is the percentile rank for $X = 3.5$? (Answer: 70%.)

Cumulative Frequency and Cumulative Percentage

- What is the 95th percentile? (Answer: $X = 4.5$.)
- What is the percentile rank for $X = 3.5$? (Answer: 70%.)
- What is the 50th percentile?

Cumulative Frequency and Cumulative Percentage

- What is the 95th percentile? (Answer: $X = 4.5$.)
- What is the percentile rank for $X = 3.5$? (Answer: 70%.)
- What is the 50th percentile?
- What is the percentile rank for $X = 4$?

Interpolation

- When a desired percentile or percentile rank is located between two known values, it is possible to estimate the desired value using the process of interpolation.

Interpolation

- When a desired percentile or percentile rank is located between two known values, it is possible to estimate the desired value using the process of interpolation.
- Interpolation assumes a regular linear change between the two known values.

Interpolation

Interpolation

- Find the percentile rank corresponding to $X = 7.0$:

X	f	cf	$c\%$
10	2	25	100%
9	8	23	92%
8	4	15	60%
7	6	11	44%
6	4	5	20%
5	1	1	4%

Interpolation

- These values are shown in the following table:

	Scores (X)	Percentages
Top	7.5	44%
Intermediate value →	7.0	?
Bottom	6.5	20%

Interpolation

- These values are shown in the following table:

	Scores (X)	Percentages
Top	7.5	44%
Intermediate value →	7.0	?
Bottom	6.5	20%

- It can be calculated as

$$\frac{7.5 - 7.0}{7.5 - 6.5} = \frac{44\% - ?}{44\% - 20\%}$$

Interpolation

- These values are shown in the following table:

	Scores (X)	Percentages
Top	7.5	44%
Intermediate value →	7.0	?
Bottom	6.5	20%

- It can be calculated as

$$\frac{7.5 - 7.0}{7.5 - 6.5} = \frac{44\% - ?}{44\% - 20\%}$$

- The result is 32%.

Interpolation

Interpolation

- Find the 40th percentile:

X	f	cf	c%
			NA%
			NA%
			NA%
			NA%
			NA%

Interpolation

- These values are shown in the following table:

	Scores (X)	Percentages	
Top	9.5	60%	
	?	40%	← Intermediate value
Bottom	4.5	10%	

Interpolation

- These values are shown in the following table:

	Scores (X)	Percentages	
Top	9.5	60%	
	?	40%	← Intermediate value
Bottom	4.5	10%	

- It can be calculated as

$$\frac{9.5 - ?}{9.5 - 4.5} = \frac{60\% - 40\%}{60\% - 10\%}$$

Interpolation

- These values are shown in the following table:

	Scores (X)	Percentages	
Top	9.5	60%	
	?	40%	← Intermediate value
Bottom	4.5	10%	

- It can be calculated as

$$\frac{9.5 - ?}{9.5 - 4.5} = \frac{60\% - 40\%}{60\% - 10\%}$$

- The 40th percentile is $X = 7.5$.

Table of Contents

1. Frequency Distributions
2. Grouped Frequency Distribution Tables
3. Frequency Distribution Graphs
4. Percentiles, Percentile Ranks, and Interpolation
5. Stem and Leaf Displays

Stem and Leaf Displays

- A stem and leaf display is an alternative procedure for organizing data. Each score is separated into a stem (the first digit or digits) and a leaf (the last digit).

Stem and Leaf Displays

- A stem and leaf display is an alternative procedure for organizing data. Each score is separated into a stem (the first digit or digits) and a leaf (the last digit).
- The display consists of the stems listed in a column with the leaf for each score written beside its stem.

Stem and Leaf Displays

- A stem and leaf display is an alternative procedure for organizing data. Each score is separated into a stem (the first digit or digits) and a leaf (the last digit).
- The display consists of the stems listed in a column with the leaf for each score written beside its stem.
- A stem and leaf display is similar to a grouped frequency distribution table, however the stem and leaf display identifies the exact value of each score and the grouped frequency distribution does not.

Stem and Leaf Displays

Stem and Leaf Displays

- The raw data of a set of $N = 24$ scores

```
## [1] 32 33 42 46 52 56 57 59 62 63 68 71  
## [13] 73 74 74 76 76 78 81 82 83 85 93 97
```

Stem and Leaf Displays

```
## 1 | 2: represents 12
```

```
## leaf unit: 1
```

```
## n: 24
```

```
## 2 3 | 23
```

```
## 4 4 | 26
```

```
## 8 5 | 2679
```

```
## 11 6 | 238
```

```
## (7) 7 | 1344668
```

```
## 6 8 | 1235
```

```
## 2 9 | 37
```

Questions?