



Tables, Charts and Graphs: Histograms

What is a Histogram?

A histogram is a type of bar graph that represents the distribution of numerical data. It shows how many data points fall into specific ranges (called class widths or intervals). Unlike a standard bar chart, where each bar represents a category, in a histogram, each bar represents a range of values, and the height of the bar shows how many data points fall within that range.

Why are Histograms used?

- Understanding Data Distribution:**
They give you a visual sense of how your data is spread — for example, whether it's skewed, symmetric, or has multiple peaks.
- Identifying Patterns:**
You can spot things like outliers, gaps, or clusters in the data.
- Comparing Groups:**
They help compare the frequency of values in different ranges, which is useful for analysis in stats, quality control, etc.

How to display your data on a Histogram

A psychologist was interested in how much time teenagers spend on social media each day. They asked 20 17 year old students to report how many hours they use social media per day.

The results of the working memory training group at the end of the study are shown in **Table 1**.

Number of hours a day (HRS) teenagers spend on Social Media	Frequency
0-2	3
3-5	10
6-8	5
9-11	3

(a) Using the data, draw a histogram to represent the number of students in each time interval for numbers of hours a day teenagers spend on social media. (4 marks)

Step 1: Calculate the size of each interval (class width). For example, $2-0=2$. If each interval is equal, there is no need to carry out any further calculations.

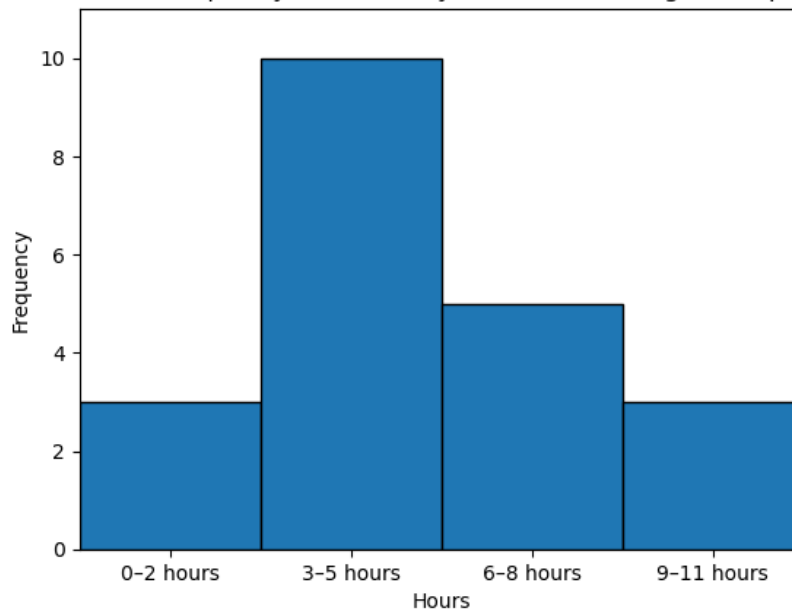


Step 2: Draw the Histogram

On the graph paper provided

- Title fully operationalised
- **X-axis** (horizontal) → your intervals (class widths)
- **Y-axis** (vertical) → frequency
- Draw on the frequencies, ensure that the bars are touching.

A Histogram to show the Frequency of how many hours students aged 17 spend on Social Media



Interpreting a Histogram

1. **Bars and Frequency:** Each bar represents a range of values. The height of the bar indicates the frequency data within that range. **For example, in the histogram, the first bar (0-2 hours) has a frequency of 3, meaning there are 3 instances of 17-year-olds spending between 0 and 2 hours on social media.**
2. **Shape of the Distribution:** The overall shape of the histogram can give insights into the distribution of the data. Common shapes include:
 - **Normal distribution:** Bell-shaped curve, indicating most data points are around the mean.
 - **Skewed distribution:** If the bars are higher on one side, the data is skewed. Positive skew means more data points are bunched to the left, and negative skew means more data points bunched to the right. **For example, the tail of the distribution extends to the right, showing that there are fewer instances of 17-year-olds spending a higher number of hours on social media.**
 - **Uniform distribution:** Bars are roughly the same height, indicating data is evenly spread across the range.
3. **Outliers:** Outliers are identified by looking for bars that stand out from the rest of the histogram. For example, if most of the data is clustered around the centre, but there is a single bar far to the right or left, that bar represents an outlier.



Figure 1 shows the lying frequency of the participants in the junior to senior Pinocchio study.

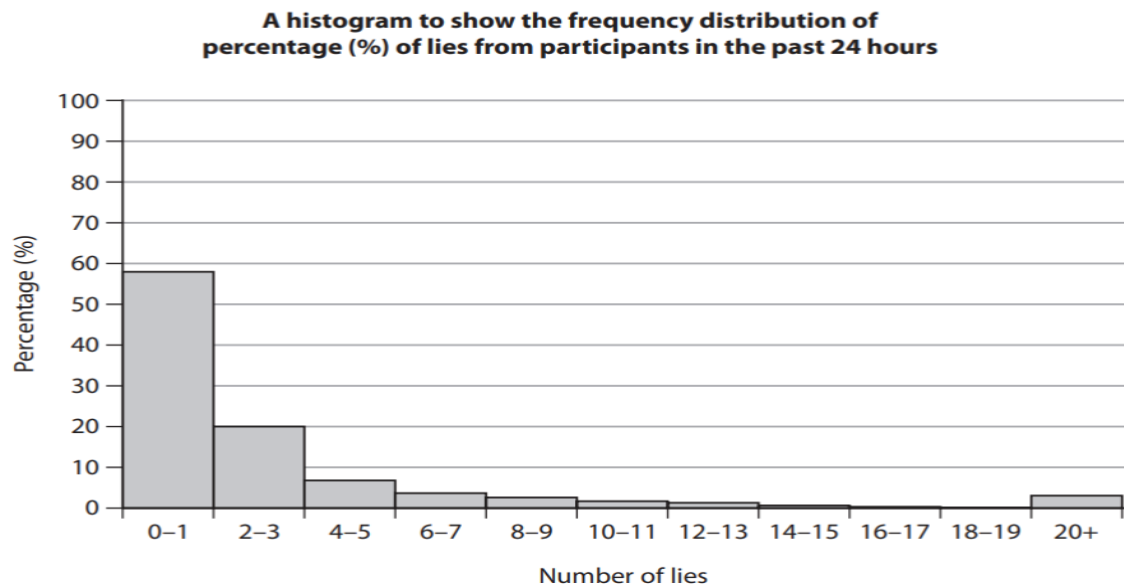


Figure 1

(e) State **one** conclusion that could be made from **Figure 1** with reference to the type of skew shown.

(1)

The majority of participants lied very infrequently in the past 24 hours which is shown by the positive skew in Figure 1 (1).



Practice Questions

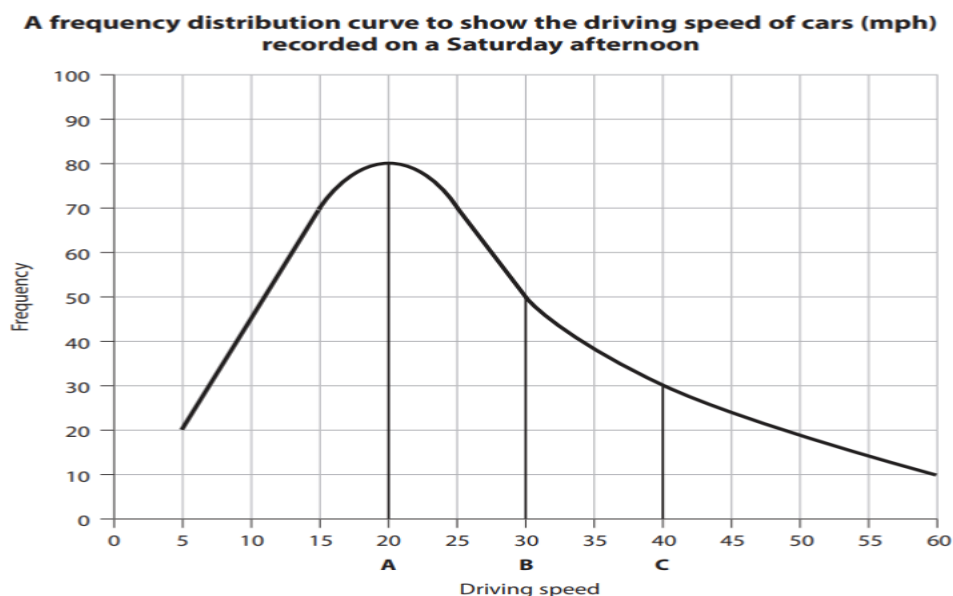
1. A psychologist was investigating how long (in minutes) it took a group of participants to complete a 20 piece jigsaw puzzle. They compared the data between for participants ages 10 years old and those aged 60. The results for 24 participants, age 10, are shown in the table below.

How long in minutes it took 10 year old ppts to complete a jigsaw puzzle	Frequency
4-6	2
7-9	5
10-12	7
13-15	4
16-18	6

Draw a histogram to represent the data in the table above, for 10 year olds.

(4 marks)

2 Jake wanted to find out if people obeyed the 30 miles per hour (mph) speed restriction in his local town. He recorded the driving speed of 200 cars on a Saturday afternoon and plotted the data on a frequency distribution curve.



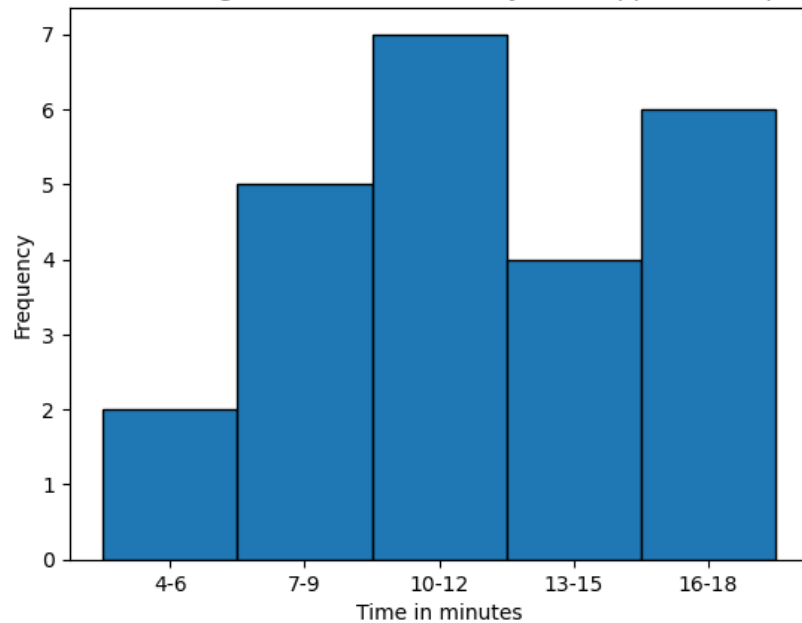
Interpret the data Jake gathered about driving speeds in his local town. (1)



Mark scheme

1.

A Histogram to show How long in minutes it took 10 year old ppts to complete a jigsaw puzzle



1 mark per correctly plotted bar.

2. Most drivers did not exceed the speed restrictions of 30mph in Jake's local town. (1)