

Recap and Outline

Chapter 3

Presenting Data

- Stem and leaf plots, Bar charts and Histograms.
- Sometimes it better to look at relative frequencies.
- Other kinds of data.
- MINITAB

Percentage Relative Frequency Histograms

| Service time | Frequency | Relative Frequency (%) |
|-----------------------|-----------|------------------------|
| $175 \leq time < 180$ | 1 | 2 |
| $180 \leq time < 185$ | 3 | 6 |
| $185 \leq time < 190$ | 3 | 6 |
| $190 \leq time < 195$ | 6 | 12 |
| $195 \leq time < 200$ | 10 | 20 |
| $200 \leq time < 205$ | 12 | 24 |
| $205 \leq time < 210$ | 8 | 16 |
| $210 \leq time < 215$ | 3 | 6 |
| $215 \leq time < 220$ | 3 | 6 |
| $220 \leq time < 225$ | 1 | 2 |
| Totals | 50 | 100 |

Percentage Relative Frequency Histograms

- Allows comparison between different sized data sets.
- Percentages are perhaps more easily understood than frequencies.

Relative Frequency Polygons

| Class Interval | Mid Point | % Relative Frequency |
|------------------|-----------|----------------------|
| $0 \leq x < 10$ | 5 | 10 |
| $10 \leq x < 20$ | 15 | 20 |
| $20 \leq x < 30$ | 25 | 35 |
| $30 \leq x < 40$ | 35 | 25 |
| $40 \leq x < 50$ | 45 | 10 |

Another example:

| Weekly Income (£) | West Road (%) | Jesmond Road (%) |
|---------------------------------|---------------|------------------|
| $0 \leq \text{income} < 100$ | 9.3 | 0.0 |
| $100 \leq \text{income} < 200$ | 26.2 | 0.0 |
| $200 \leq \text{income} < 300$ | 21.3 | 4.5 |
| $300 \leq \text{income} < 400$ | 17.3 | 16.0 |
| $400 \leq \text{income} < 500$ | 11.3 | 29.7 |
| $500 \leq \text{income} < 600$ | 6.0 | 22.9 |
| $600 \leq \text{income} < 700$ | 4.0 | 17.7 |
| $700 \leq \text{income} < 800$ | 3.3 | 4.6 |
| $800 \leq \text{income} < 900$ | 1.3 | 2.3 |
| $900 \leq \text{income} < 1000$ | 0.0 | 2.3 |

Cumulative Frequency Polygons (Ogive)

| Class Interval | % Relative Frequency | Cumulative % Relative Freq |
|------------------|----------------------|----------------------------|
| $0 \leq x < 10$ | 10 | 10 |
| $10 \leq x < 20$ | 20 | 30 |
| $20 \leq x < 30$ | 35 | 65 |
| $30 \leq x < 40$ | 25 | 90 |
| $40 \leq x < 50$ | 10 | 100 |

Pie Charts

$$angle = \frac{\text{Number in category}}{\text{Total number in sample}(n)} \times 360.$$

| Paper | Frequency | Degrees |
|---------------------|-----------|---------|
| The Times | 140 | 77.5 |
| The Sun | 200 | 110.8 |
| The Sport | 50 | 27.7 |
| The Guardian | 120 | 66.5 |
| The Financial Times | 20 | 11.1 |
| The Mirror | 80 | 44.3 |
| The Daily Mail | 10 | 5.5 |
| The Independent | 30 | 16.6 |
| Totals | 650 | 360.0 |

Time Series Plots

| Quarter | Units Sold |
|---------|------------|
| Q1 2000 | 86.7 |
| Q2 2000 | 94.9 |
| Q3 2000 | 94.2 |
| Q4 2000 | 106.5 |
| Q1 2001 | 105.9 |
| Q2 2001 | 102.4 |
| Q3 2001 | 103.1 |
| Q4 2001 | 115.2 |
| Q1 2002 | 113.7 |
| Q2 2002 | 108.0 |
| Q3 2002 | 113.5 |
| Q4 2002 | 132.9 |
| Q1 2003 | 126.3 |
| Q2 2003 | 119.4 |
| Q3 2003 | 128.9 |
| Q4 2003 | 142.3 |
| Q1 2004 | 136.4 |
| Q2 2004 | 124.6 |
| Q3 2004 | 127.9 |

Scatter Plots

| Total costs (£) | Monthly Output |
|-----------------|----------------|
| 10300 | 2400 |
| 12000 | 3900 |
| 12000 | 3100 |
| 13500 | 4500 |
| 12200 | 4100 |
| 14200 | 5400 |
| 10800 | 1100 |
| 18200 | 7800 |
| 16200 | 7200 |
| 19500 | 9500 |
| 17100 | 6400 |
| 19200 | 8300 |

If you were interested in the relationship between the cost of production and the number of units produced you could easily plot this by hand.

1. The “response” variable is placed on the y -axis. Here we are trying to understand how total costs relate to monthly output and so the response variable is “total costs”.
2. The variable that is used to try to explain the response variable (here, monthly output) is placed on the x -axis.
3. Plot the pairs of points on the graph.