

Features of Graphs and Statistics

The PPDAC Cycle

STOP AND CHECK (PAGE 5)

- For this external, the important stages are the analysis and the conclusion.

The Basics of Probability

STOP AND CHECK (PAGE 9)

- $\frac{13 + 9}{7 + 11 + 13 + 9} = \frac{22}{40}$ which is 0.55, or 55%

Special Features of Statistics

STOP AND CHECK (PAGE 9)

- Summary statistics:
 - Minimum = 2 (smallest number)
 - Maximum = 17 (biggest number)
 - Median = 6.5 (middle of 6 and 7)
 - Mean = 8.8 (add all number \div 20),
 - LQ = 5 (middle of first half of data),
 - UQ = 15 (middle of second half of data)

Features of Dot and Box-and-Whisker Plots

STOP AND CHECK (PAGE 11)

- 1, 3, 3, 4, 5, 6, 6, 6, 6, 6, 7, 7, 7, 8, 9, 9, 11, 11, 12, 13
- Summary statistics:

Minimum	LQ	Median	Mean	UQ	Maximum
1	5.5	6.5	7	9	13

Features of Scatter Plots and Time Series Graphs

STOP AND CHECK (PAGE 14)

- A dot plot would show the age and number of pizzas
- A time-series would show the number of slices and time (in months)

Bar Charts, Pie Graphs and Other Tasty Graphs

STOP AND CHECK (PAGE 17)

- Pie charts are useful when all groups add to 100%.
- When comparing two groups on a bar graph, it's best to use proportions to be able to as the data has been standardised and won't be changed by different sample sizes.
- There are 5 people in group B.

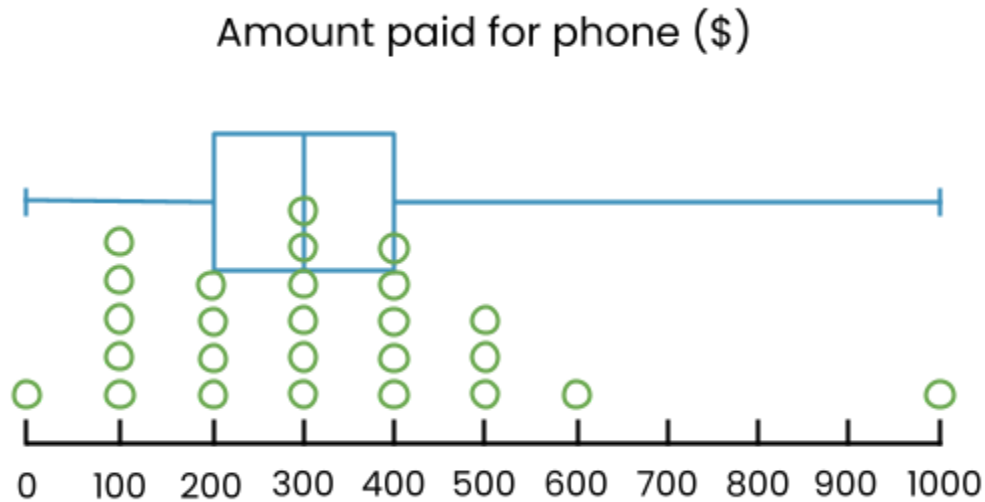
Features of Graphs and Statistics

QUICK QUESTIONS (PAGE 17)

- Values: 0, 100, 100, 100, 100, 100, 200, 200, 200, 200, 300, 300, 300, 300, 300, 300, 400, 400, 400, 400, 400, 500, 500, 500, 600, 1000.
 - With a total of 26 friends sampled.

Minimum	LQ	Median	Mean	UQ	Maximum
0	200	300	315	400	1000

- Box plot of prices:



- Adrian could not show the data on a time series as she has no information about time. If she wanted a time series, then she could collect data on a variable like the age of the phone model when it was purchased and then compare it to the price.

Features of Data

Outliers

STOP AND CHECK (PAGE 20)

- There is an outlier at the start of 2011 where there is a big peak and another outlier halfway through 2014, where there is a big decrease in the number of pizzas eaten. The first outlier might be because Isaac had his extended family over a lot so they had more pizza nights, and the decrease might have been because the extended family were all on a diet.

How Graphs Shift and Overlap

STOP AND CHECK (PAGE 22)

- Children have a bigger spread in the amount of pizza eaten but adults have a bigger interquartile range and a higher median. The adults' middle 50% completely overlap the children's middle 50%, with their upper 50% the same size as the upper 25% of the children. Both have the same mean. In general, even though the median number of slices is bigger for adults than children, you cannot confidently say that there is a significant difference

The Shapes of Graphs

STOP AND CHECK (PAGE 24)

- There is a right skew and there is clustering around the low values such as 0 - 4 slices of pizza, which means the younger kids tend to eat roughly the same number of pizzas
- The scatter plot shows an increased range in the variance of slices as people get older. This means that children are more likely to eat less, but an adult may eat between 4 - 9 slices.

How Graphs Follow Trends

STOP AND CHECK (PAGE 29)

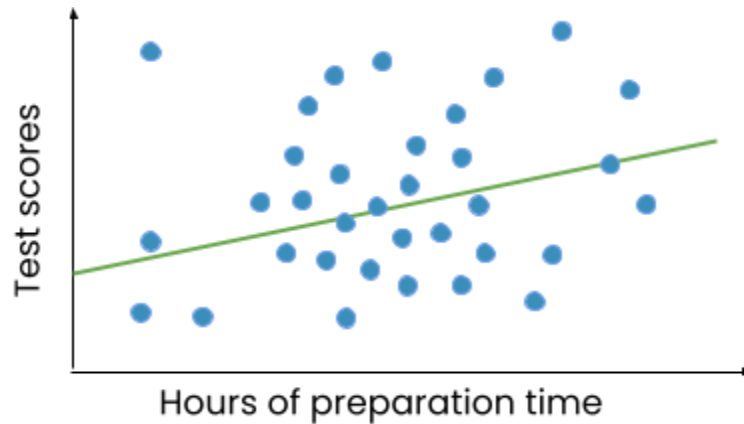
- Overall the number of pizzas eaten is decreasing, with a decrease of about two pizzas per year. The highest point is in early 2011 at 20 pizzas and the lowest point is partway through 2014 at 4 pizzas. There is a strong seasonal effect, with fewer pizzas eaten in the middle of the year than at the start of the year.

Features of Data

QUICK QUESTIONS (PAGE 30)

- Station 1 has a left skew and station 2 has a right skew, so station 1 tends to have high results and station 2 tends to have low results. Station 2 has a

bigger spread than station 1 so a bigger range of scores, but station 1 has a larger interquartile range, so the middle 50% of scores are more variable. Station 1 is shifted to the right of station 2, with the top 75% of scores the same as the top 25% of scores at station 2. Station 2 also has an outlier where one person got a score that was much higher than others at that station. In general, people score better at station 1 than at station 2.



- The trend line shows a positive relationship between hours of preparation and test scores, so the more time you spend preparing the better you are predicted to score in the test. However, this is a relatively weak relationship, so the amount of preparation time might not be the only thing affecting test scores.

Analysing Chance and Data

How to Calculate Conditional Probability

STOP AND CHECK (PAGE 25)

- The probability that a randomly selected person is an adult who wins is $\frac{25}{100}$. As a proportion, this would be 0.25, or 25% as a percentage.

Bias and Unfair Surveys

STOP AND CHECK (PAGE 35)

- There are a lot more girls than boys sampled, with the sample size of the boys quite small at only 30, so it may not be representative. People were also not randomly sampled as she took the sample from a library, and most people there might read more books than the general population.

Comparing Two Sets of Data

STOP AND CHECK (PAGE 36)

- You need which group is higher up in the graph, differences in long term trends and differences in seasonal trends.

Analysing Validity

STOP AND CHECK (PAGE 39)

- Compare the two medians, talk about the graphs (differences in shift, skew, overlap, trend, and outliers), and say if the conclusion is valid or not.

Analysing Chance and Data

STOP AND CHECK (PAGE 39)

- Children have a bigger spread than adults, meaning that there is more variation in the number of slices they eat. Adults have a larger interquartile range, meaning they have more variation in their middle 50%. At 8 slices compared to 5.5 slices, the median number of slices of pizza an adult eats at his family reunion is more than the children's median. The medians are probably not different enough to draw a confident conclusion, particularly as the question does not say the sample size. Both means, at about 6, are the same, so on average children and adults eat the same number of slices. The adults show a slight left skew while the children are more symmetrical in shape, meaning that the number of slices a child eats is probably more evenly distributed than the adults. The median number of slices for an adult is the

same as the upper quartile for the children and the middle 50% of adults overlap the middle 50% of children. There might be some outliers for children in the number of slices they eat, as some eat 0 and others eat 12. From this evidence, Isaac's claim that children eat more slices than adults is not supported. The medians are too similar and the middle 50% of children completely overlapped by 25% of the adults.