Assessment Schedule – 2023
Scholarship Statistics (93201)
Evidence Statement
General Principles:

1 Ignore incorrect answers if alongside correct answers. The exception is contradictory statements.

2 Ignore minor copying errors.

3 When required in evidence, answers need to be contextual.

QUESTION ONE

Task Q1(a)
Evidence:

- While about 50% of energy used in the residential and commercial sector comes from electricity, this proportion is only about 25% for the industrial and agriculture sector.
- Diesel is the main source of energy in the agricultural sector (about 50%), while it plays only a minor role for the other sectors (residential <5%, commercial 10%, industrial 15%).
- Heating and cooling are the highest end use group for residential (about 65%), commercial (about 55%), industrial sector (about 65%), but only the second highest for the agricultural sector (about 25%).
- Mobile motive power is the highest end use group for the agricultural sector (about 55%), but it is only a minor group for the other sectors (residential 20%, commercial <5%, industrial 5%).
- Natural gas is the highest source of energy for industry (about 25%) but less important in other sectors (about 10% in residential and commercial, and 5% in agriculture).

Task Q1(b)(i)
Evidence:

- Difference: 0.6 – 0.39 = 0.21
- Estimated margin of error $= 2 \times \frac{1}{\sqrt{3039}} = 0.036.$
- CI for the difference of two proportions: (17.4%, 24.6%)
- With approximate 95% confidence, the proportion of households in NZ that use heat pumps is between 17.4 and 24.6 percentage points higher than those using electric plug-in heaters.

Note: Survey respondents could select more than one of the heating sources, hence the margin of error cannot be calculated assuming two independent groups

Task Q1(b)(ii)
Evidence:

- Electricity cost increased in most months after installation compared to the same month in the previous year (8 out of 12 months increased in cost).
- The general pattern of daily electricity cost remained the same with the majority being used between 5 p.m. and 10 p.m. This is true for all seasons.
- Both before and after the installation, the electricity cost was the highest on average in winter, regardless of the hour of day.
- The electricity cost in summer has increased after the installation by about $0.10 during peak hours.

Task Q1(b)(iii)
Evidence:

- It is hard to compare the mean cost for a specific hour of the day before and after the heat pump was installed, as these amounts are shown in different plots. To make comparisons requires estimating the mean costs on the two separate plots by reading the *y*-axis.
- A different way to visualise the data could be to further "subset" the data by season. For each season, the line for "before" and "after" could be displayed on the same plot (using different colours). In this way, you could more easily compare the mean electricity cost for each hour of the day within the same season.

Note: Accept other reasonable limitations of Figure 4 with respect to making comparisons of electricity costs before and after the heat pump was installed, and other reasonable suggestions for different visualisations.

Scholarship level	Outstanding level
Q1a • **3s marks** for making three comparison statements, with at least two about **fuel** and with at least two estimated **proportions** comparing features from two or more graphs. OR **2s marks** for two comparison statements with at least one estimated proportions from two or more graphs OR 1s mark for one comparison statement with at least one estimated proportion. OR **2s marks** for four comparison statements from two or more graphs (no proportions). **Q1bi** • **1s mark** for calculating an appropriate estimate for the margin of error. **Q1bii** • **3s marks** for describing three trends across **both** Figures 3 & 4 at least one with numerical evidence from the *y*-axis. OR **2s marks** for describing two trends from one or more figures. OR **1s mark** for describing one trend. **Q1biii** • **1s mark** for identifying a reasonable limitation of Figure 4 for making comparisons. *8 marks available* **Max 6 marks awarded**	**Q1a** • **1o mark** for making four comparison statements, with at least two about fuel and with at least three estimated proportions comparing features from two or more graphs **Q1bi** • **1o mark** for constructing and interpreting an appropriate confidence interval. **Q1biii** • **1o mark** suggesting an appropriate way to visualise the electricity data to support comparison. *3 marks available* **Max 2 marks**

QUESTION TWO

Task Q2(a)(i)
Evidence:

- The deviation is a measure of how different the actual production was compared to the predicted one. For example, in November, the actual amount was about 5300, while the predicted amount was about 4800. The difference of 500 was likely divided by the predicted, so that the deviation of this month would have been $\frac{500}{4800} = 10.4\%$. The deviation was likely an overall measure of this relative difference, by averaging the deviations for each month.

Note: Prediction has been used as the "base" for the calculation of deviations, as we are attempting to evaluate the prediction model.

Task Q2(a)(ii)
Evidence:

- Model this situation using a Binomial distribution with $n = 12$, $p = 0.5$.
- Binomial distribution appropriate since we have a fixed number of trials (12 months), a fixed probability (assuming 50%), two outcomes (predicted production is higher than actual, or not higher than actual) and can assume independence (no reason to believe solar production from one month impacts solar production in another month).
- Seven of the 12 months the predicted production was higher than the actual production.
- $P(X \geq 7) = 1 - P(X \leq 6) = 0.3872$
- This probability is relatively large.
- No, we cannot conclude that more than half of the time, the predicted production of solar energy will be higher than the actual production of solar energy, as we do not have sufficient evidence to support this conclusion.

Task Q2(b)(i)
Evidence:

- The difference between the two means for alertness scores was 0.85, which means that, on average, students who completed the task in the office with incandescent lighting found more errors than those who completed the task in an office with LED lighting.
- The tail proportion is 0.178 (17.8%), which provides no evidence to support a claim that the light source used in an office affects the ability to process written information.

Task Q2(b)(ii)
Evidence:

- To control for differences in students' ability to process written information, a paired comparison design could be used in which each student does a similar but different task twice, once under LED lighting, and once under incandescent lighting.
- In this case, the response variable would be the difference in alertness score (based on number of spelling errors correctly identified) under the two lighting conditions.
- The order in which students are assigned the treatments LED / incandescent lighting should then be randomly allocated to control for other effects, such as fatigue and practice.

Scholarship level	Outstanding level
Q2ai • 1s mark for explaining how the deviation may have been calculated. MUST be the explanation for the deviation across all twelve months not just one month. • **1s mark** for providing one example of comparing predicted production and actual prediction using a relative approach e.g., percentage difference. **Q2aii** • **1s mark** for correctly identifying binomial distribution. • **1s mark** for identifying and discussing at least two key features of binomial distribution in context. • **1s mark** for correctly calculating an appropriate probability using a binomial distribution model. **Q2bi** • **1s mark** for correctly interpreting the difference in two means contextually. **Must include direction.** • **1s mark** for using the tail proportion to reject the claim (tail proportion of 5% or 10% not necessary) **Do not accept the position that it was caused by chance alone** (accept 'may have been due to chance' but not 'was due to chance'). **Q2bii** • **1s mark** for discussing an experiment with random allocation. OR 1s mark for discussing a paired comparison study. *8 marks available* **Max 6 marks**	**Q2aii** • **1o mark** for reasoning using and appropriate binomial distribution-based probability. **Q2bii** • **1o mark** for discussing the use of a paired comparison experiment **with** the order of treatment randomly allocated. • **1o mark** for discussion of factors that need to be controlled as part of an experiment. *3 marks available* **Max 2 marks**

QUESTION THREE

Task Q3(a)(i)
Evidence:

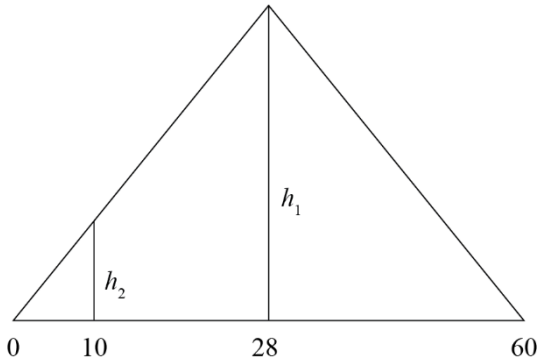
- It is likely that some of the greenness distributions for cities are skewed, with some streets with many trees.
- The median is a better overall measure for central tendency for skewed distributions, as it is not affected by extreme values.

Task Q3(a)(ii)
Evidence:

- While both distributions have most of the values between 0% and around 60%, the majority of the values in the Sydney distribution are closer to the centre.
- As the standard deviation measures the average distance of the values from the mean, the standard deviation of the Sydney distribution is smaller.

Task Q3(a)(iii)
Evidence:

- Sydney: Normal distribution with mean = 26 and standard deviation = 9.
- $P(X < 10) = 0.0377$
- Vancouver: Triangular distribution with min = 0, mode = 28, max = 60



- $\frac{1}{2} \times 60 \times h_1 = 1$
- $h_1 = 0.0333$
- $h_2 = \frac{10}{28} \times h_1 = 0.0119$
- $P(X < 10) = \frac{1}{2} \times 10 \times 0.0119 = 0.0595$

Note: estimates for distribution parameters will vary, so accept reasonable values.

Task Q3(b)
Evidence:

- For an observational study, existing houses within Wellington would need to be selected. Whether each house had a light-coloured or dark-coloured roof would need to be recorded, as well as indoor temperatures. Any conclusions made from this study would be limited in terms of causation, as there are other factors that could be related to both roof colour and indoor temperatures. If a random sample of houses within Wellington was used, the conclusions could be generalised to all houses in Wellington.
- For an experiment, houses would need to be constructed, or modified, to have either a light-coloured or a dark-coloured roof, specifically for the purposes of the study. Whether each house had a light-coloured or a dark-coloured roof should be randomly allocated, and all other features of the houses should be as similar as possible. Indoor temperatures would also need to be recorded. Any conclusions made from this study could support claims of causation; however, these conclusions could be limited to the types of houses constructed or modified for the experiment.

Scholarship level	Outstanding level
Q3ai • **1s mark** for demonstrating sound understanding of measures of central tendency (mean vs median). **Q3aii** • **1s mark** for reasoning with visual features of probability distributions to compare standard deviation. **Q3aiii** • **1s mark** for providing suitable parameters for a normal distribution model. Accept mean 24–28 and std dev 7–10. • **1s mark** for correctly calculating a probability from a normal distribution model. **can use unsuitable parameters.** • **1s mark** for providing suitable parameters for a triangular distribution model. Accept h2 of 25.9-32 • **1s mark** for correctly calculating a probability from a triangular distribution model. **can use unsuitable parameters.** OR **1s mark** for correctly calculating a probability for Vancouver from a normal distribution model with suitable parameters and a standard deviation that is larger than Sydney's standard deviation. **Q3b** • **1s mark** for demonstrating understanding of the need for random allocation in an experiment. • **1s mark** for demonstrating understanding that an observational study cannot make a causal claim. OR **1s mark** for discussing confounding variables in observational studies. *7 marks available* **Max 6 marks awarded**	**Q3a** • **1o mark** for demonstrating sound understanding of measures of central tendency (mean vs median) **and** variation (standard deviation) **in context**. **Q3b** • **1o mark** for discussing limitations of causal claims made from an experiment. • **1o mark** for discussing random sampling and the extension of results for an observational study. *3 marks available* **Max 2 marks awarded**

QUESTION FOUR

Task Q4(a)
Evidence:

- Three sampling strategies are involved.
- Of all the suburbs of Auckland and Christchurch (sampling frame), 100 are chosen at random. As Auckland has more suburbs, it is likely that there are more Auckland suburbs in the sample, but as we are comparing across the two cities, this imbalance within the sample is not necessarily an issue.
- Just one set of co-ordinates was used to represent each suburb and around these a square of size 150 metres by 150 metres was generated. This may be an issue, as depending on the size and layout of the suburb, this sampling method may not select a representative area of the suburb's greenness.
- For each of the squares generated for each suburb, a random sample of 1000 pixels was taken from the Google map satellite image. A potential issue is that using the green pixels to measure greenness could capture roofs.

Task Q4(b)
Evidence:

- There are more Auckland suburbs in the sample than Christchurch suburbs.
- The mean greenness of Auckland suburbs in the sample is 43.3 higher than that of Christchurch suburbs.
- Based on the bootstrap confidence interval, it is a fairly safe bet that the mean greenness of all Auckland suburbs is somewhere between 51.7 lower than and 135.1 higher than that of all Christchurch suburbs.
- Based on this confidence interval, there is not enough evidence to conclude that the mean greenness of suburbs is higher for either of the two cities.

Task Q4(c)
Evidence:

- Overall, there is a moderate positive relationship between the distance from the city and the greenness of a suburb. In general, the further a suburb is from the city centre, the greener it is.
- Christchurch suburbs are generally greener when close to the city centre than Auckland suburbs are.
- All but one of the Christchurch suburbs are at most 12 km from the city centre, whereas the Auckland suburbs can be up to 65 km from the centre.
- If a separate linear model was fitted for each city individually, the slope would be steeper for Christchurch suburbs, which suggests that greenness increases faster with distance from the city.
- There is less scatter for Christchurch compared to Auckland, and the variation in greenness score for the Auckland suburbs increases as the distance from the city centre increases.

Task Q4(d)
Evidence:

- As all but one of the Christchurch suburbs are at most 12 km from the city centre, the sampling process could be changed so that only suburbs within 12 km of the city centre are sampled.
- The Christchurch sample size of 20 is very small, which is of concern because small samples may have more sampling error due to sampling variation. Therefore, the sampling process could be changed to increase the number of suburbs from Christchurch that are sampled.

Scholarship level	Outstanding level
Q4a • **2s marks** for identifying the sampling strategies related to suburbs, co-ordinates and pixels. OR **1s mark** for identifying TWO of the above strategies. **Q4b** • **1s mark** for comparing the sample means or another relevant feature in context. • **1s mark** for interpreting the bootstrap confidence interval in context. **Must include reference to the mean.** • **1s mark** for identifying that there is not enough evidence to conclude that the mean greenness of suburbs is higher for either of the two cities. **Q4c** • **1s mark** for describing the overall relationship between distances to city centre and greenness scores. **Must include direction and strength (weak to moderate).** OR 1s mark for describing the relationships of Christchurch and Auckland individually. **Must include direction and strength, Auckland strength must be weaker than Christchurch.** • **2s marks** for making TWO valid statements comparing the relationship for the two cities. **Comparisons must be direct, not just implied.** OR 1s mark for making ONE valid statement comparing the relationship for the two cities. *8 marks available* **Max 6 marks awarded**	**Q4a** • **2o marks** for discussing potential issues with TWO of the sampling strategies, including use of suburb co-ordinates / area. OR 1o mark for discussing potential issues with ONE of the sampling strategies. **Q4d** • **1o mark** for discussing how to change the sampling process to make comparisons fairer. **Must make a link to Figure 10. Do not accept stratified sampling without further, reasonable, explanation.** *3 marks available* **Max 2 marks awarded**

Sufficiency Statement

For each question:

Score 1 – 4 No award	5 – 6 Scholarship level	7 – 8 Outstanding Scholarship level
Shows understanding of relevant statistical and probability concepts and methods, and some progress towards applying this in context.	Application of high-level statistical analysis and critical thinking, knowledge and skills, to complex situations. Shows logical development, precision, and clarity of ideas.	In addition to the requirements of Scholarship, demonstration of perception and insight, sophisticated integration and abstraction of ideas, independent reflection and extrapolation, and convincing communication.

Cut Scores

Scholarship	Outstanding Scholarship
20 – 27	28 – 32