Assessment Schedule – 2015

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.

QUESTION ONE

Tasks Q1 (a)
Evidence:
Figure 1

- Between 2010 and 2014 there was an upwards trend in the Auckland Region median price ($475K to almost $675K).
- The median prices peaked in the summer months and troughed in the winter months.
- The actual median prices fluctuated between 2010 and 2014 with small fluctuations of approx. $25K between 2010 and 2012 with the fluctuations increasing to approx. $50K between 2012 and 2014.

Figure 2:

- Between 2010 and 2014, the Auckland Region sales volumes showed a fluctuating trend with the seasonal component removed. There was an overall increase of a sales volume of 500 houses (2,000 to 2,500).
- Over the period 2010 to 2014, the sales volume peaked at approx. 3,400 houses in March 2013 **OR** over the period 2010 to 2014 the sales volume was least at approx. 1100 in the first quarter 2011.
- The fluctuations in actual sales volumes were fairly constant over the period (+/- 1,500 houses).
- The fluctuations in sales volumes troughed in January/February each year and peaked in March over the first quarter.

Note:

1. Can have a maximum of three bits when describing trend.
2. Allow flexibility in the specified location of peaks and troughs.
3. Must make at least one point from each Figure to gain 4 marks for (a).
4. The trend max volume was 2600 and trend min was 1600.

Tasks Q1 (b)
Evidence:
Table 1

- Both North Shore City and Waitakere City have the greatest increases in median sale prices at 17.6% and 17.0% respectively. Manukau's increase was 9.6%.
- Waitakere City showed the greatest volume increase at 16.7% with Manukau 6.1% and North Shore 3.7%.
- With total sales value estimated as median x sales volume; Waitakere City showed the greatest increase at 36.5% followed by North Shore City at 22.0% and last Manukau City at 16.3%.

Note:

1. One mark maximum for actuals.

Task Q1 (c)(i)
Evidence:

- Trend, Seasonal and Cyclical (or fluctuations in Trend).

Note:

1. Accept any two components for one mark.

Task Q1(c)(ii)
Evidence:

- Extrapolate the 12 month MA into 2016 and read off value for each month of 2016. Then obtain seasonal for each month by taking the difference between each month and the MA over the period 2010 to 2014. Average out the values over every corresponding month. Compute a forecast for each month in 2016 by taking the trend value for that month times the corresponding seasonal value. Add your 12 forecasts to get an overall forecast for 2016.

Note:

1. Marked on a 0, 1, 2 (complete description) basis.
2. A full answer must involve months not quarters.

Overall Judgement: for Question 1

Max 4 for (a), Max 3 for (b), Max 6 for (a) + (b) then Max 3 for (c) leading to max 8 overall.

QUESTION TWO

General Principle: Evidence provided as answers needs to be contextual.

Tasks Q2 (a)
Evidence:
- The national median price was up $3,000 excluding the impact of the Auckland Region**.**
- Sales volumes excluding Auckland were up 1.5% from March.

Note:
1. From the article evidence consisting of two distinct points is required of strength showing in regional markets.

Tasks Q2 (b)
Evidence:
- The mean prices are distorted by the sale houses at either the upper or lower end of the market.

Tasks Q2 (c)
Evidence:
- It allows for sales fluctuations that occur in different months/quarters of the year which are vastly different. The real trend vales can be observed. For example, sales before and after Christmas vary greatly so we wish to ascertain the trend.

Tasks Q2 (d)
Evidence:
- The demand for houses is substantially higher in Auckland so an auction was preferable by the vendor in order to get the highest possible selling price. Hence the percentage of 76% is substantially higher than in any other part of NZ.

Note:
1. Can omit last sentence and still score the mark.

Tasks Q2 (e)
Evidence:
To allow for variations between the regions that would affect the median price like demand for houses. It means that an overall index measuring the movement of house prices can be fairly weighted between the characteristics of each region.

Note:
1. The answer "adjusts for variations between regions" is too vague and unacceptable.

Tasks Q2 (f)
Evidence:
The percentage changes in the abridged extract are based on the complete data not a sample.

Tasks Q2 (g)
Evidence:
Sales volume measure sales activity which has a direct effect on (total value of sales) demand which is a driver of price. Using both price and volume estimates of sales value can be obtained along with a fairer comparison between regions. Two regions may have the same median sale price but the sales volumes could vary significantly.

Tasks Q2 (h)
Evidence:
- HPI for Auckland April 2015 = 1.189 x (612/450) x 100 = 161.7

Note:
1. One mark for getting 612/450.

Overall Judgement for Q2

Max 2 for (a), 1 mark for each of (b) to (f) inclusive. Overall (a) to (f) max 6. For (g) 1mark and (h) 2 marks max, overall max 8 marks.

QUESTION THREE

Tasks Q3 (a)(i)
Evidence:
- In both cities 4 people per house is the most common, with the modal cluster between 2 and 4.
- The mean number of people per house in Auckland is likely to be between 0.41 and 0.99 more than the mean number of people per house in Christchurch.
- Number of people occupying a house in Christchurch is more consistent than in Auckland. (IQR 2 versus 3 or standard deviation 1.6 versus 2.5).
- Both graphs have a positive skew meaning that a smaller number of people occupy a majority of houses.

Note:
1. Must have value plus context.
2. Any two of a central measure comparison, a spread comparison or a bootstrap comment is acceptable for full marks.

Task Q3 (a)(ii)
Evidence:
- The distribution of the average number of people per house for the 413 suburbs is symmetric and clustered about 2.9, whereas the number of people per house has a right skewed distribution clustered about 3.
- Each value in Figure 5 is a mean value where in Table 2, each value is a number of people.

Note:
1. Central limit theorem is implied here.
2. One mark for description and one mark for reason. No marks for a lack of context.

Task Q3(a)(iii)
Evidence:
- The mean value of 3.947 is more likely to reflect the mean number of people per house in Auckland provided the survey was a random sample. A confidence interval calculation gives (3.7, 4.1) **OR**
- The mean value of 3.0 is more likely to reflect the mean number of people per house in Auckland as this is based on a census. However, there would be a different number of houses in each suburb, so the mean of these means is not likely to be the mean number of people in each house.

Note:
1. Mark is awarded for a sensible reason given for choice of mean.
2. Both dot points are required.

Tasks Q3 (b)(i)
Evidence:
- The asking price displays a weak positive relationship with the number of bedrooms **or** with the floor area.
- In both cases there is greater variation for larger x-values (number of bedroom and floor area).
- The asking price has strong, positive, linear relationship with the capital value.
- From CV $250000 to $1000000 the asking price is similar to the CV. For CV values above $1m, the graph shows a slighter steeper gradient, meaning that asking prices are above the CV
- There is no relationship between land area and asking price- most houses are on a small area of land.

Note:
1. Credit is given for only one dot point per graph.

Task Q3(b)(ii)
Evidence:
- I would use the CV to estimate the asking price. I could read an estimated value from the graph or add a piecewise linear trend line to the graph.

Task Q3(b)(iii)
Evidence:

- Categorical

Suburb/location	Some suburbs are seen to be more attractive than others. An attractive location would command a higher house price than a less desirable location.
Quality of the build	A house built with better quality products would command a higher price than a house built from cheaper products
Having a garage or not	A house with adequate garaging/parking might be more expensive than a house without garaging

- Numerical

Distance from CBD	As the distance from the CBD increases, the price would be expected to decrease
Age of the house	Older houses might be expected to be cheaper, although some old-style houses are seen to be 'chic' so would be more expensive

Note:

1. No marks awarded for just naming a variable without an explanation.
2. Number of bathrooms or demand not accepted as variables.

Overall Judgement for Q3
Max 2 for (a)(i), max 2 for (a)(ii), max 1 for (a)(iii), max 3 for (b)(i) leading to max 6 for (a) + (b)(i), max 1 for (b)(ii) and max 2 for (b) (iii) leading to max 8 overall.

QUESTION FOUR

Task Q4 (a)(i)
Evidence:

	Apartment	Free standing	
1–2 bed	2072	728	2800
3 bed	560	6192	6752
4+ bed	168	280	448
	2800	7200	10000

$$P(1\text{–}2 \mid \text{free standing}) = \frac{728}{7200} = 0.101$$

Note:
1. One mark for some indication of correct method.

Task Q4(a)(ii)
Evidence:

$$P(\text{apartment} \mid 1\text{–}2 \text{ bed}) = \frac{2072}{2800} = 0.74$$

$$P(\text{free standing} \mid 1\text{–}2 \text{ bed}) = \frac{728}{2800} = 0.26$$

$$\frac{0.74}{0.26} = 2.846 \left(\text{or } \frac{2072}{728}\right)$$

about 2.8 times as likely

Note:
1. One mark for some indication of correct method. Must have conditional probability.
2. No marks for having conditional probability the wrong way round.

Task Q4 (b)(i)
Evidence:

$$(0.5 \times 68) + 86 + 61 + (0.5 \times 23) = 192.5$$

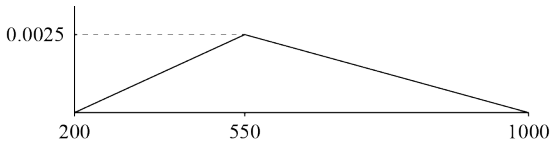
$$\frac{192.5}{325} = 0.592$$

About 59% of houses sold for between $450 000 and $750 000.

Note:
1. One mark awarded for getting 192.5 or 0.59 with no working.
2. One mark for indication of correct method.

Task Q4(b)(ii)
Evidence:
Triangular distribution

min = 200, max = 1000, modal value at 500, height of triangle at $x = 550$ is 0.0025.



$$H_1 = \frac{2(450-200)}{(1000-200)(550-200)} = \frac{500}{280\,000} = \frac{1}{560}$$

$$H_2 = \frac{2(1000-750)}{(1000-200)(1000-550)} = \frac{500}{360\,000} = \frac{1}{720}$$

$$P(450 < X < 750) = 0.5 \times \left(\frac{1}{560} + 0.0025\right) \times 100 + 0.5 \times \left(\frac{1}{720} + 0.0025\right) \times 200$$

$$= 0.603$$

Note:

1. Min, Max and height at highest point is insufficient.
2. Must have three points to describe distribution: min, max and modal value.

Normal distribution

Estimate mean, std dev from graph:

Mean = 550 Std dev = $\frac{800}{6}$ = 133.3

P (450 < X < 750) = 0.7068

OR calculating from table mean = 510, std dev = 150

P (450 < X < 750) = 0.601

OR other sensible approach to finding mean, std dev

- Triangular distribution fits well on left side where there is a linear increase, but not so well on right side. Normal distribution might fit right side better where sales are clustered closer to centre.
- Triangular distribution might be more appropriate having a definite smallest price and largest price sale. Normal model assumes that there is no upper and lower limit.
- Graph is slightly skewed so triangular distribution can model this better. Normal distribution assumes symmetry about the mean.
- Estimates of probability – consistent comment for their model eg both models give similar probability value and this value is close to proportion calculated in (i).

Note:

1. Marks are scored as ½ + ½ + ½ + ½ for model, parameters, probability and discussion about each distribution leading to a maximum of 4 with an overall ½ rounded down.
2. Probabilities can be conditionally correct.

Overall Judgement for Q4

Max 2 for (a)(i), max 2 for (a)(ii), max 2 for (b)(i), max 4 for (b)(ii) leading to max 8 overall.

QUESTION FIVE

General Principle: Evidence provided as answers needs to be contextual.

Tasks Q5 (a)

Evidence:

- The median number score increase from those in Course X was 1.0 higher than that from those in Course Y (Mean difference was 3.15).
- The increases in scores from the participant that did Course Y were slightly more varied than from those who did Course X. The IQR differed by 3.75; the standard deviation differed by 1.37.
- The distribution of score differences were skewed differently; Course X positively and Course Y negatively.
- The lower quartile of the differences in scores from the Course X participants was greater than the median of the differences from the Course Y participants.

Note:

1. One mark for each of central measure, spread and distribution comparisons. The 4th mark is for anything sensible for comparison. Can use either 3.15 or 3.25 for the mean difference.
2. Must have value with context (penalty minus one mark).

Tasks Q5 (b) (i)

Evidence:

- To remove possible effects (even out characteristics) to the experimental outcomes like selling areas, branches.

Tasks Q5 (b) (ii)

Evidence:

- All the agents in the experiment are likely to be operating at different levels of selling effectiveness and we wish to allow for these effects from the experiment.
- We are comparing changes due to the courses.

Note:

1. One of these dot points is sufficient.

Tasks Q5 (b) (iii)

Evidence:

- At the end of the experiment, the mean score was 6.5 for Course X increases and 3.35 for Course Y increases giving a difference of 3.15 (refer Table 6) or 3.25 (refer Figure 12).
- The re-randomisation distribution for the mean of the difference between (After – Before) for Course X and (After – Before) for Course Y has a tail proportion of 0,017. A difference of 3.15 or more could occur by chance alone with probability of 1.7 % **so** this experiment gives sufficient evidence (1.7% is less than a 5% threshold) to claim that Course X is improving the selling effectiveness to the agents more effectively than Course Y.

Note:

1. A conclusion referring to a 10% tail is acceptable.
2. The second part is marked on a 0, 1, 2 basis.
3. Overall mark is 1+2 for a maximum of 3.
4. Can have 3.25 or 3.15 for the mean difference.

Tasks Q5 (b) (iv)

Evidence:

- Claim isn't necessarily correct as it assumes a uniform increase in scores from all the participants after either Course X or Course Y.

Note:

1. Any suitable explanation is acceptable.

Overall Judgement for Q5

Max 4 for (a), max 1 for (b)(i), max 2 for (b)(ii) leading to max 6 so far. Max 3 for (b)(iii), max 1 for (b)(iv) leading to max 8 overall.