Assessment Schedule 2023

Mathematics and Statistics (Statistics): Evaluate statistically based reports (91584)

Evidence Statement

Q	Expected Coverage	Achievement (c)	Achievement with Merit (j)	Achievement with Excellence (i)
ONE (a)	47% ± 3.1% [43.9%, 50.1%] We're pretty sure that the population proportion of all **New Zealanders** who are in favour of **keeping the daylight saving system as it is**, is somewhere between 43.9% and 50.1%.	• CI correctly calculated.	• CI correctly calculated. AND Used to write inference statement in context (with population AND variable).	
(b)(i)	The question would have let respondents select as many reasons as apply, therefore lots of respondents would have their answers in multiple categories.	• Clearly indicates that the question would have been designed so that respondents can select multiple reasons.		
(ii)	It may not be appropriate as a respondent may have provided answers in both groups, so the two groups are not independent. Because the respondents could be in more than one group, the answers are not mutually exclusive, so a confidence interval is not appropriate. It would be wrong to do a two-group confidence interval for a difference between groups because the respondents are one group who are giving both answers. In this case, a respondent may have selected both the options that we are comparing, so it is not valid to create a confidence interval because one person may be in both groups. It would be wrong to do a one group confidence interval for a difference between answers because the answers are not mutually exclusive, and the same people could be included for both answers. In this case, a respondent may have selected both the options that we are comparing, so it is not valid to create a confidence interval because one person may have given both answers.	• Identifies that the answers are either not independent of each other, or mutually exclusive, with some context.	• Identifies one type of confidence interval for a difference, which would be inappropriate. OR Identifies that the answers are either not independent of each other, or mutually exclusive, with some context. AND Explains why a confidence interval is not appropriate without explicitly stating which type of confidence interval	• Identifies one type of confidence interval for a difference, which would be inappropriate. AND Explains why it is not appropriate.

(c)

	All year daylight saving	**Keep daylight saving as it is**
Poll % *enjoy daylight hours*	65%	57%
n	18% of 1010 = 182 respondents	47% of 1010 = 475 respondents
$\dfrac{1}{\sqrt{n}}$	7.41%	4.59%

Difference between groups: 65% – 57% = 8%

1.5 times average MoE = 9%

CI: 8% ± 9%

[–1%, 17%]

We're pretty sure that for **New Zealanders** who want all-year daylight saving, the proportion who gave the reason "**more time to enjoy daylight hours during summer**" is somewhere between 1% less and 17% more than for **New Zealanders who want to keep daylight saving as it is**.

Because this interval is both negative and positive, there is not sufficient evidence to support the claim that "the proportion of **New Zealanders** who want all-year daylight saving because it gives them "**more time to enjoy daylight hours during summer**" is larger than the proportion of **New Zealanders** who want to keep **daylight saving as it is** for the same reason."

- Average MoE calculated (or 1.5 × average MoE), i.e. 6% or 9% <u>with clear evidence of method.</u>

  OR

  Correct confidence interval without clear working.

  OR

  Correctly calculated an incorrect CI for difference between two groups (i.e. using wrong MoE method and / or wrong sample sizes) and both claim and interpretation are in context.

- Confidence interval correct.

  AND

  EITHER used to write inference statement in context (with population AND variable).

  OR

  Response to claim made in context (with population AND variable).

- Confidence interval correct used to write inference statement in context (with population variable).

  AND

  Response to claim made in context (with population AND variable).

NØ	N1	N2	A3	A4	M5	M6	E7	E8
No response; no relevant evidence.	Attempt at one part of the question.	1 of c	2 of c OR 1 of j	3 of c	1 of j AND 1 of c	2 of j	1 of i	2 of i

Q	Expected Coverage	Achievement (c)	Achievement with Merit (j)	Achievement with Excellence (i)
TWO (a)	The number of towels used per room per day is a very clear, objective measurement that is easy for the hotel room staff to collect as it is easy to count the number of towels that need changing when servicing each room.	• Clearly states that towels are easy to count or measure (or similar).		
(b)	A potential confounding variable is the number of guests per room. If a hotel room has more guests, then they will likely use more towels. By randomly allocating guests to each treatment there should be about the same numbers of one person, couples, or families in a room for the three treatments. Other confounding variables that may be accepted: • Number of showers per day by guests: If a guest typically takes more than one shower a day for personal hygiene reasons, then they will likely use more towels. By randomly allocating guests to each treatment ,there should be about the same numbers of those guests that have no showers, one shower, or multiple showers per day in a room for the three treatments. • Gender: Discussion around the number of females/males in each group. If there are gender patterns of use (e.g. females showering more times than males per day) then they will likely use more towels. By randomly allocating guests by gender to each treatment, there should be about the same numbers of those gender groupings for the three treatments. *Accept other examples of relevant confounding variables.*	• Identifies a confounding variable AND how it impacts number of towels. OR Explains how random allocation will balance confounding in context.	• Identifies a confounding variable and how it impacts number of towels. AND Explains how random allocation will balance confounding in context.	
(c)(i) (ii)	**Hotel guests** were blinded, as they were not aware they were participants in the study and did not know that any other rooms had different signs (or similar). **Hotel room staff** were blinded to the researchers' hypothesis. They did not know the researchers were expecting that the guests with the message "75% of guests in this room" would use the least towels (or similar). It was important for hotel guests to not know their towel use was being monitored as this might change their behaviour. For example, they might use fewer towels if they knew the number of towels they were using was being counted. For hotel staff, it was important that they didn't know the researchers' hypothesis because they may not record the data accurately or they may encourage guests to reuse towels if they were in this treatment group.	• Describes what blinding involved for **hotel guests**. AND Describes what blinding involved for **hotel room staff**.	• Describes what blinding involved for **hotel guests** AND why it is important. AND Describes what blinding involved for **hotel room staff** AND why it is important. OR Describes blinding in BOTH situations AND ONE situation includes description of Why and How.	• Describes what blinding involved for **hotel guests** AND why it is important, including how it may impact the response variable. AND Describes what blinding involved for **hotel room staff** AND why it is important, including how it may impact the response variable.

| (d) | Issues include:

 • The study was conducted in 2013 – ten years ago. Attitudes to environmental behaviour have changed in this time, so hotel guests may be more conscious of towel use now compared to ten years ago.
 • The study was only conducted in two alpine ski resorts. Different types of hotels, e.g. lower priced or higher priced, will have different sorts of customers who might have different towel-use behaviour.
 • The study was only conducted in two Swiss and Austrian ski resorts. Hotels in different locations, e.g. beach resorts, will have different sorts of customers who might have different towel-use behaviour.

 Accept other valid issues.
 Don't accept "small sample size" or "differences in language" discussions. | • Describes TWO potential issues with extending the results. | • Describes TWO potential issues with extending the results.

 AND

 EITHER Discusses how ONE potential issue could limit extending the results by using specific features of the report / study.

 OR Discusses how ONE potential issue could limit extending the results. | • Describes TWO potential issues with extending the results.
 AND

 Discusses how BOTH potential issues could limit extending the results by using specific features of the report / study. |

NØ	N1	N2	A3	A4	M5	M6	E7	E8
No response; no relevant evidence.	Attempt at one part of the question.	1 of c	2 of c OR 1 of j	3 of c	1 of j AND 1 of c	2 of j	1 of i	2 of i

Q	Expected Coverage	Achievement (c)	Achievement with Merit (j)	Achievement with Excellence (i)
THREE (a)	As media consumption is so frequent and includes so many different platforms, it can be difficult to remember exactly what media you have consumed reasonably recently (e.g. did you listen to the radio or a podcast or music on your way home a week ago?). By asking respondents about what they consumed "yesterday", this eliminates as much of these memory recall issues as possible and gets responses that are fresh in the respondents' minds.	• Identifies memory recall as an issue with some link to media consumption.	• Detailed comment about memory recall issues clearly linked to media consumption.	
(b)	Regional sample stratification is the process of dividing a larger population into distinct geographical areas or regions before selecting a sample for research or survey purposes. This ensures representation from different locations, and can account for regions that are under or over represented in a sample. Regional sample stratification means that the proportion of respondents selected from each region will be proportional to the population of that region. For example, there would be more survey responses from Auckland than Dunedin, as Auckland has a much larger population than Dunedin. Stratification is important if the media consumption is different in different regions.	• Describes concept of regional sample stratification, potentially by example.	Describes concept of regional sample stratification, potentially by example and be specific to this context).	• Describes concept of regional sample stratification, potentially by example (should be specific to this context). AND Why stratification is important in this context.
(c)	One advantage of using interviews is that complex questions can be asked about media consumption and the responder can seek clarification about anything as the interview progresses. This would mean that the data collected would be as accurate as possible. One disadvantage of using interviews is that responders may be embarrassed by some of their media consumption habits (e.g. watching inappropriate material) and therefore not be truthful in their responses. This could mean that platforms with more easily accessible inappropriate material (e.g. YouTube) could be viewed more than what is reported. *Accept other valid comments related to interviews.* *Do not accept comments related to advantages or disadvantages of using telephone or online platforms.*	• ONE general advantage AND ONE general disadvantage of interviews given. OR EITHER ONE advantage OR ONE disadvantage described specifically in relation to interviews to collect data on respondents' media consumption.	• ONE advantage AND ONE disadvantage of interviews described specifically in relation to interviews to collect data on respondents' media consumption. OR EITHER ONE advantage OR ONE disadvantage described specifically in relation to interviews to collect data on respondents' media consumption AND its impact on the results of the survey discussed.	• ONE advantage AND ONE disadvantage of interviews described specifically in relation to interviews to collect data on respondents' media consumption. AND Discussion of the impact on the results of the survey discussed for at least one.

| (d) | **Landline**: Respondents who still have a landline might do so because they are in a more rural area where the internet may also be slow or unreliable. Therefore, these respondents viewing habits might be more reliant of traditional platforms of live TV or radio. TV and radio consumption for respondents who still have a landline might be higher as a result. It is important that results combining landlines with other groups are presented in a way that ensures the proportions of each group matches that in the population. If the sample contained more people with landlines than the population, the consumption of traditional platforms reported may be higher than in the actual population.

Gender: Media consumption on the different platforms may be different for different genders – e.g. males are less likely to consume podcasts. Therefore, it is important that results combining gender are presented in a way that ensures the proportion of male responses matches that in the population – if, for example, the sample contained 1/3 males. yet the population is ½ males, then the consumption of podcasts overall may be reported as lower than in the actual population.

Age: Media consumption is different for different age groups, with respondents aged 60+ more likely to watch live TV and radio. Therefore, it is important that results combining ages are presented in a way that ensures the proportions of each age group matches that in the population – if, for example, the sample had more respondents aged 60+ than the population, then the consumption of live TV and radio overall may be reported as higher than in the actual population.

Ethnicity: Different ethnic groups may have different trends in media consumption. For example, Pasifika people are more likely to watch online videos. Therefore, it is important that results combining all ethnicities are presented in a way that ensures the proportions of each ethnicity closely matches that in the population – if, for example, the sample had less Pasifika people than the population, then the consumption of online videos may be reported as lower than in the actual population.

Accept other valid comments related to factors of landline, gender, age, or ethnicity.

Do not accept comments related to other factors. | • TWO factors described in context.

OR

ONE factor described in context and linked to differences in media consumption. | • TWO factors described in context AND both factors linked to differences in media consumption.

OR

ONE factor described in context, linked to differences in media consumption AND how the results of the survey could be impacted are discussed. | • TWO factors described in context

AND both factors linked to differences in media consumption.

AND

How the results of the survey could be impacted are discussed for both factors. |

NØ	N1	N2	A3	A4	M5	M6	E7	E8
No response; no relevant evidence.	Attempt at one part of the question.	1 of c	2 of c OR 1 of j	3 of c	1 of j AND 1 of c	2 of j	1 of i	2 of i

Cut Scores

Not Achieved	Achievement	Achievement with Merit	Achievement with Excellence
0 – 7	8 – 13	14 – 18	19 – 24