

S

93201Q



NEW ZEALAND QUALIFICATIONS AUTHORITY
MANA TOHU MĀTAURANGA O AOTEAROA

QUALIFY FOR THE FUTURE WORLD
KIA NOHO TAKATŪ KI TŌ ĀMUA AO!

Scholarship 2021 Statistics

Time allowed: Three hours
Total score: 40

QUESTION BOOKLET

There are five questions in this booklet. Answer ALL FIVE questions.

Write your answers in Answer Booklet 93201A.

Pull out Formulae and Tables Booklet S–STATF from the centre of this booklet.

Show ALL working. Start your answer to each question on a new page. Carefully number each question.

Check that this booklet has pages 2–13 in the correct order and that none of these pages is blank.

YOU MAY KEEP THIS BOOKLET AT THE END OF THE EXAMINATION.

QUESTION ONE

- (a) Data about all Apple, Samsung, and Oppo branded mobile phones listed for sale was obtained from one New Zealand electronics retailer's website.

For each phone listing, its brand, current price (in \$NZ), and colour were recorded. Only phones that were black, blue, red, or white were used for exploratory data analysis.

Figures 1, 2, and 3 show three of the plots produced from the data.

Figure 1: Price by brand

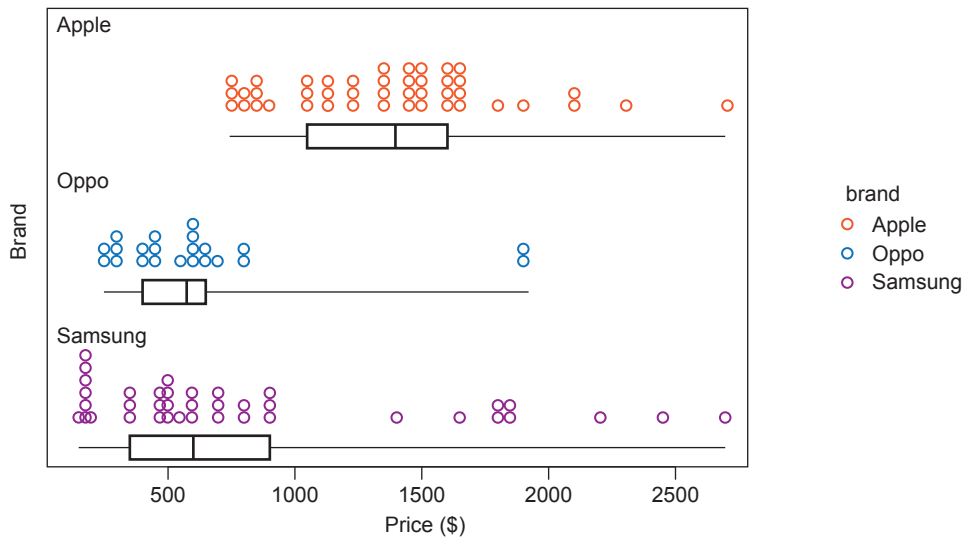


Figure 2: Colour by brand

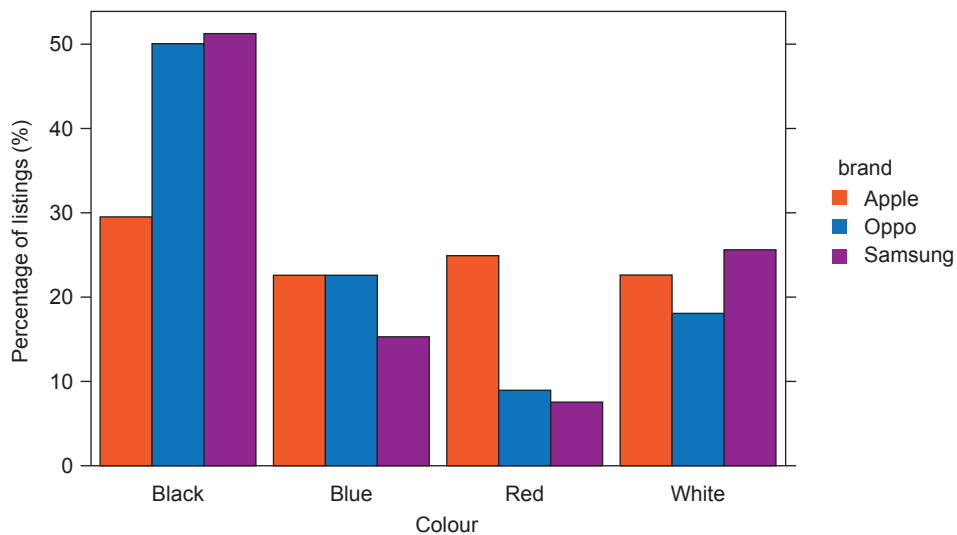
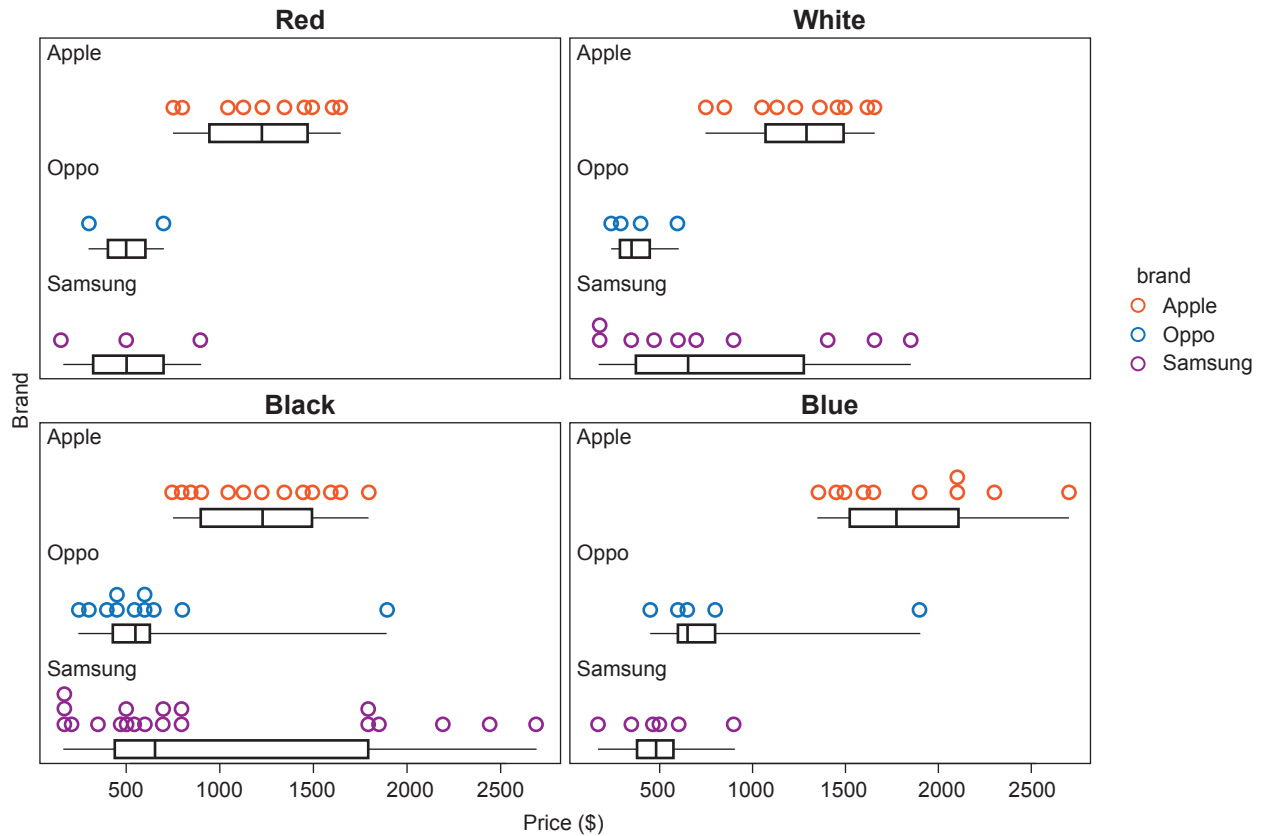


Figure 3: Price by brand subset by colour



- (i) Write a short paragraph describing the relationships between brand, colour, and price for these phone listings.
- (ii) Briefly discuss what additional information you would need to know before generalising your descriptions to all Apple, Samsung, and Oppo branded mobile phones listed for sale within New Zealand.

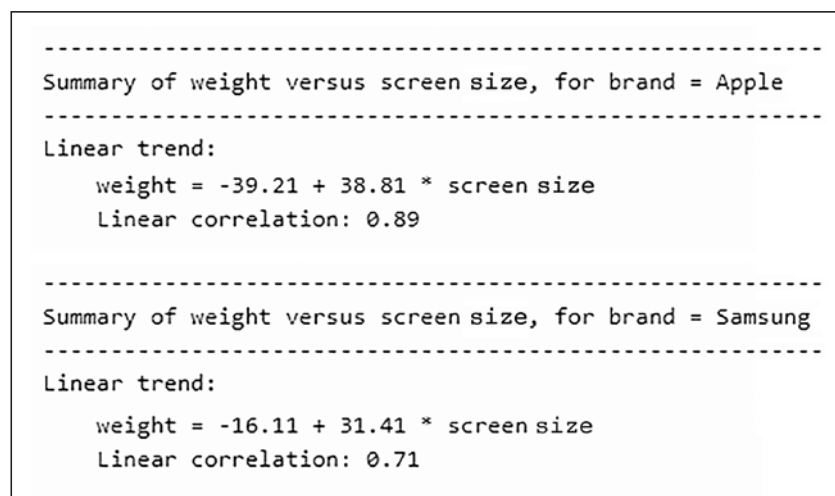
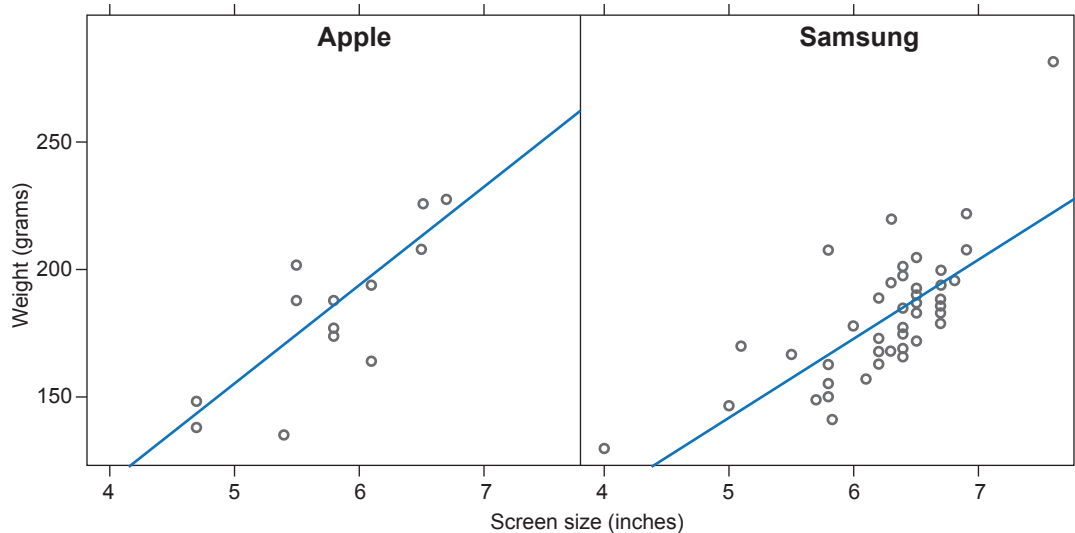
- (b) Data about Apple and Samsung branded mobile phones listed for sale was obtained from one United States electronics retailer's website.

For each phone listing, its brand, screen size (in inches), and weight (in grams) were recorded.

The relationship between the weight of a mobile phone and its screen size was then explored.

Figure 4 shows one of the plots produced, with the output showing the linear models fitted for each brand of mobile phone and the linear correlation values.

Figure 4: Weight by screen size subset by brand



- (i) Describe the relationship between weight and screen size, for these mobile phone listings.
- (ii) Suppose a new mobile phone will be released that has a screen size of 7.5 inches.

The weight of this new mobile phone is known to be 225 grams.

It is also known that the brand of the new phone is either Apple or Samsung.

Using the models given in Figure 4, calculate TWO different relevant predictions and use these to decide which brand is more likely for this new phone.

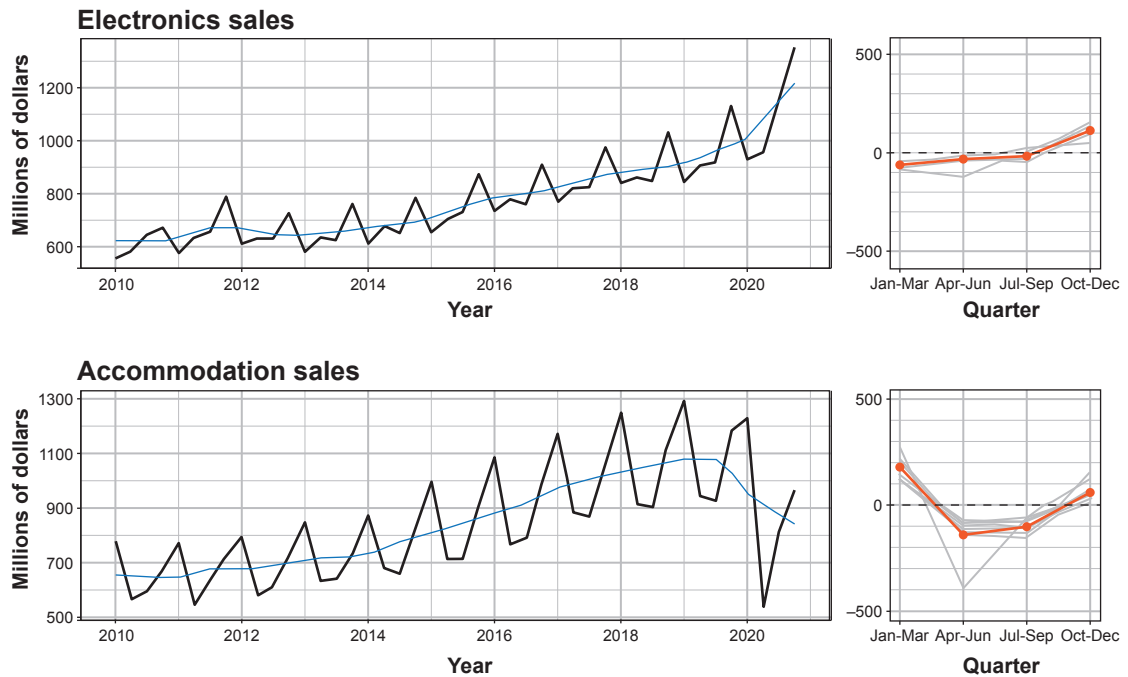
- (iii) Give TWO reservations you might have with using the data, models, and predictions from these models to decide the brand of this new phone.

QUESTION TWO

- (a) Data was obtained from Tauranga Aotearoa (Stats NZ) on the retail sales of electronics and accommodation per quarter, for the years 2010 to 2020. Sales were measured in millions of dollars.

Figure 5 displays the raw data for the retail sales of electronics and accommodation per quarter for the years 2010 to 2020, with a smoothed trend curve shown in blue. Figure 5 also displays the seasonal differences, with their average (mean) shown in red.

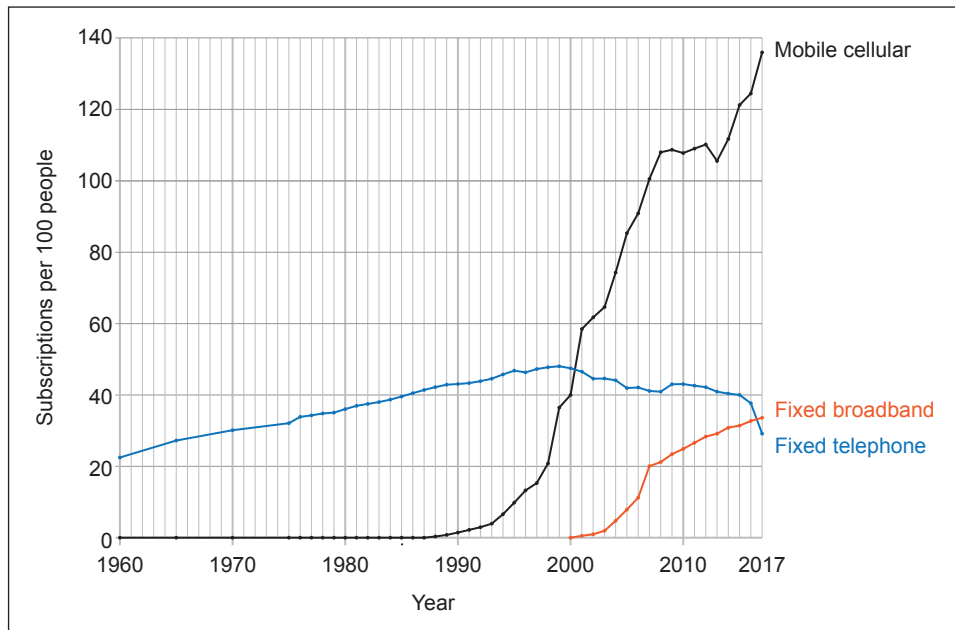
Figure 5: Retail sales of electronics and accommodation, 2010 – 2020



- (i) Write two short paragraphs comparing the features of the data for the retail sales of electronics and accommodation over the period 2010 to 2020.
- (ii) Discuss any issues with using the data shown in Figure 5 to make forecasts for the retail sales of electronics and accommodation during 2021.

(b) Figure 6 shows a graph provided by the website *ourworldindata.org*.

Figure 6: Mobile and fixed landline telephone subscriptions, New Zealand, 1960–2017



Discuss the different types of subscriptions for New Zealand over the period 1960 to 2017.

(c) Data was obtained from *databank.worldbank.org* on the number of mobile phone subscriptions in New Zealand, Fiji, and Samoa for the years 1997 to 2017. This data is shown in Table 1.

Table 1: Number of mobile phone subscriptions in New Zealand, Fiji, and Samoa, 1997–2017

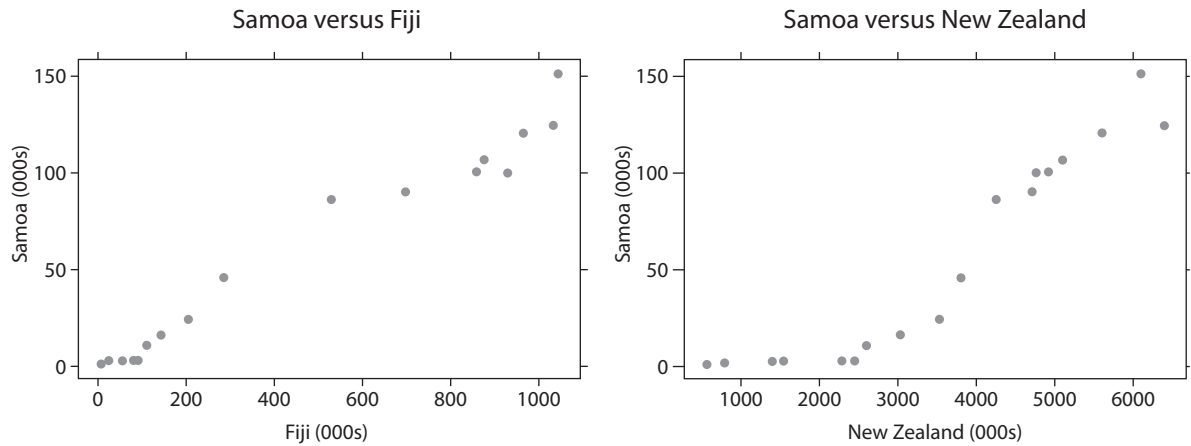
Year	New Zealand	Fiji	Samoa
1997	566 200	5 200	766
1998	790 000	8 000	1 480
1999	1 395 000	23 380	2 432
2000	1 542 000	55 057	2 500
2001	2 288 000	80 933	2 500
2002	2 449 000	89 900	2 700
2003	2 599 000	109 882	10 500
2004	3 027 000	142 190	16 000
2005	3 530 000	205 000	24 000
2006	3 802 290	284 661	45 500
2007	4 251 207	530 048	86 000
2008	4 620 000	600 000	
2009	4 700 000	640 000	
2010	4 710 000	697 920	90 000
2011	4 820 000	727 000	
2012	4 922 000	858 809	100 302
2013	4 766 000	930 406	99 887
2014	5 100 000	876 176	106 524
2015	5 600 000	965 950	120 517
2016	6 100 000	1 044 685	151 008
2017	6 400 000	1 033 915	124 211

(i) It was claimed that the number of mobile subscriptions per year for Fiji had increased at a faster rate than that for New Zealand over the years 2005 to 2008 inclusive.

Using the data supplied in Table 1, investigate if this claim is justified.

- (ii) Figure 7 shows two scatterplots produced using the data shown in Table 1. Note that for these scatterplots, the number of mobile phone subscriptions is shown in thousands.

Figure 7: Comparison of mobile phone subscriptions for different countries



Data on the number of mobile phone subscriptions is not available for Samoa for 2008, 2009, and 2011.

Use the information provided in Table 1 and Figure 7 to estimate the number of mobile phone subscriptions for Samoa for 2009.

Support your answer with statistical reasoning.

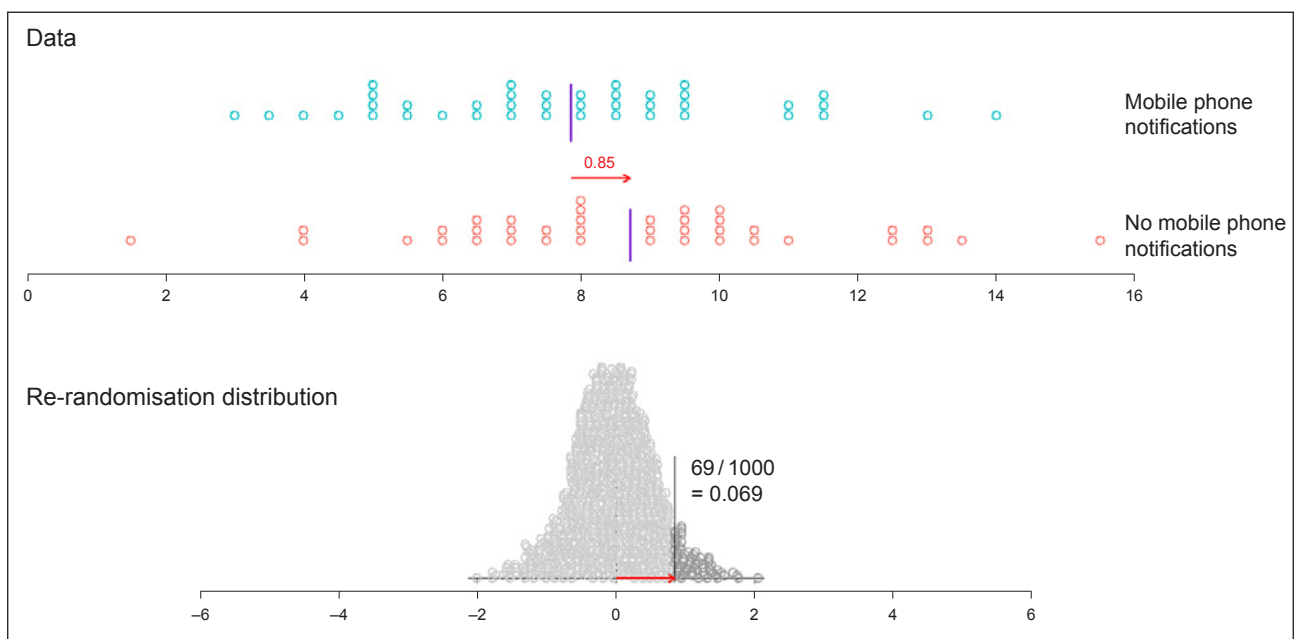
QUESTION THREE

A study was carried out to investigate whether mobile phone notifications affected memory. A group of university students was given different lists of words to memorise, and studied the word lists under one of two levels of distraction: (1) no mobile phone notifications and (2) mobile phone notifications with instructions to read the notifications before dismissing them. The students were randomly allocated to one of the levels of distraction and the number of words correctly recalled for each student was recorded.

- Write a short paragraph that identifies the key elements of the study design **using appropriate statistical terminology**.
- Explain why the students were randomly allocated to one of the levels of distraction.
- A randomisation test was carried out using the difference between the mean number of words correctly recalled by students for each level of distraction.

Figure 8 gives some output from this test.

Figure 8: Randomisation test output



Interpret the randomisation test output.

- Discuss how a similar study could be designed so that each student's memory ability could be taken into account.

- (e) After completing the word recall activity for the study, the students completed questionnaires about their video-gaming, musical, and texting experiences.

Suppose another randomisation test was carried out comparing the mean number of words correctly recalled by students who were classified as **gamers** or **non-gamers**.

For this randomisation test, the gamers correctly recalled 1.2 words more, on average, than the non-gamers and the tail proportion for the randomisation test was 0.032.

Discuss whether the following claim could be supported by both the design of this study and by this randomisation test: *Playing video games significantly affects your memory for the better!*

QUESTION FOUR

TataurangaKiTeKura Aotearoa (CensusAtSchool New Zealand) has conducted an online survey of Year 3 to Year 13 students every two years since 2005.

Three of the variables for which data has been collected across all surveys include:

- year level (e.g. Year 3, Year 4, etc.)
- region (e.g. Wellington, Canterbury, etc.)
- mobile phone ownership (Yes, No).

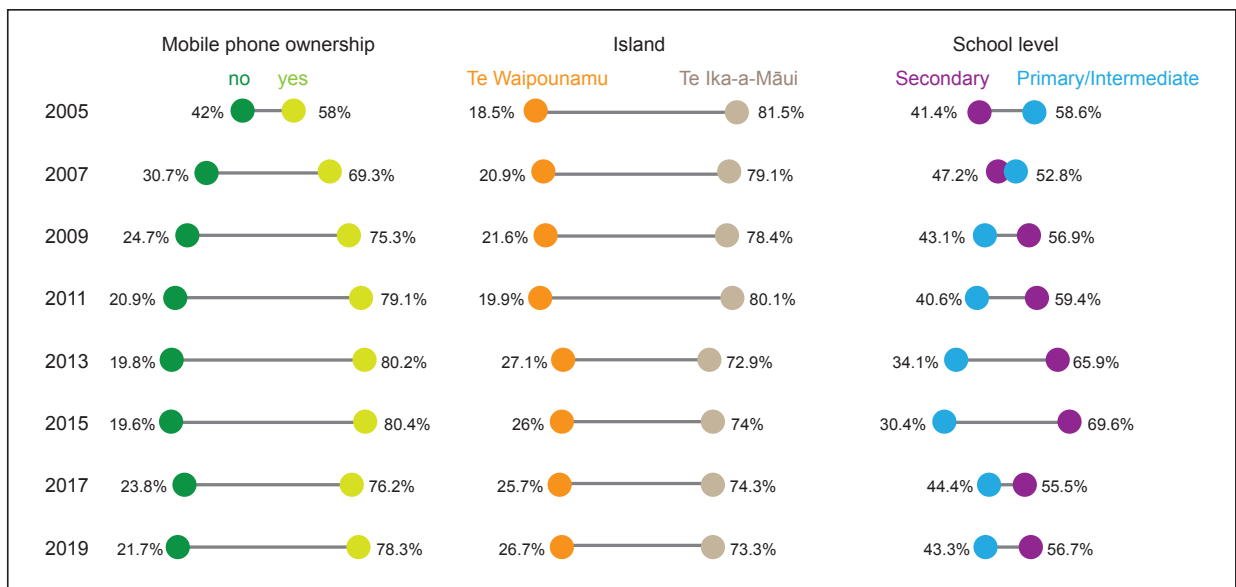
Random samples of 1000 students from each of the biennial surveys from 2005 to 2019 were obtained and combined to create a new data set. Five rows from this data set are shown in Table 2.

Table 2

census_year	year_level	school_level	Region	Island	mobile_phone
cas2019	Year 9	Secondary	Auckland Region	Te Ika-a-Māui (North Island)	yes
cas2019	Year 10	Secondary	Canterbury Region	Te Waipounamu (South Island)	no
cas2015	Year 9	Secondary	Northland Region	Te Ika-a-Māui (North Island)	yes
cas2009	Year 7	Primary or Intermediate	Auckland Region	Te Ika-a-Māui (North Island)	yes
cas2019	Year 7	Primary or Intermediate	Canterbury Region	Te Waipounamu (South Island)	yes

- (a) Identify TWO new variables that have been created for the data set shown in Table 2 using the three variables listed in the bullet points above, and describe how they were created.
- (b) Data from the combined sample data set of 8000 students was used to create the visualisation shown in Figure 9.

Figure 9



- (i) Write a short paragraph describing what the visualisation shown in Figure 9 reveals about students who participate in TaurangaKiTeKura Aotearoa (CensusAtSchool New Zealand).
- (ii) Discuss TWO of the graphical techniques used for the visualisation shown in Figure 9.
- (iii) Suppose a claim was made that the overall proportion of New Zealand Year 3 to Year 13 school students who own mobile phones has increased between 2005 and 2019.

Discuss TWO reservations you would have with using the sample percentages as they are presented in Figure 9 to make inferences about all New Zealand Year 3 to Year 13 school students who own mobile phones.

- (c) In the 2019 survey, the following new question was used:

Do you check your phone for messages or notifications as soon as you wake up?

This question was shown only to students who said they owned a mobile phone earlier in the survey.

Students had to select one of five options for their answer: always, often, sometimes, rarely, or never.

The numbers of students who selected each option within the sample data described earlier are shown in Table 3.

Table 3

Always	Often	Sometimes	Rarely	Never
190	156	184	138	98

Investigate the claim that most New Zealand Year 3 to Year 13 students who own mobile phones either sometimes, often, or always check their phone for messages or notifications as soon as they wake up.

Assume that this random sample is representative of all New Zealand Year 3 to Year 13 students who own mobile phones.

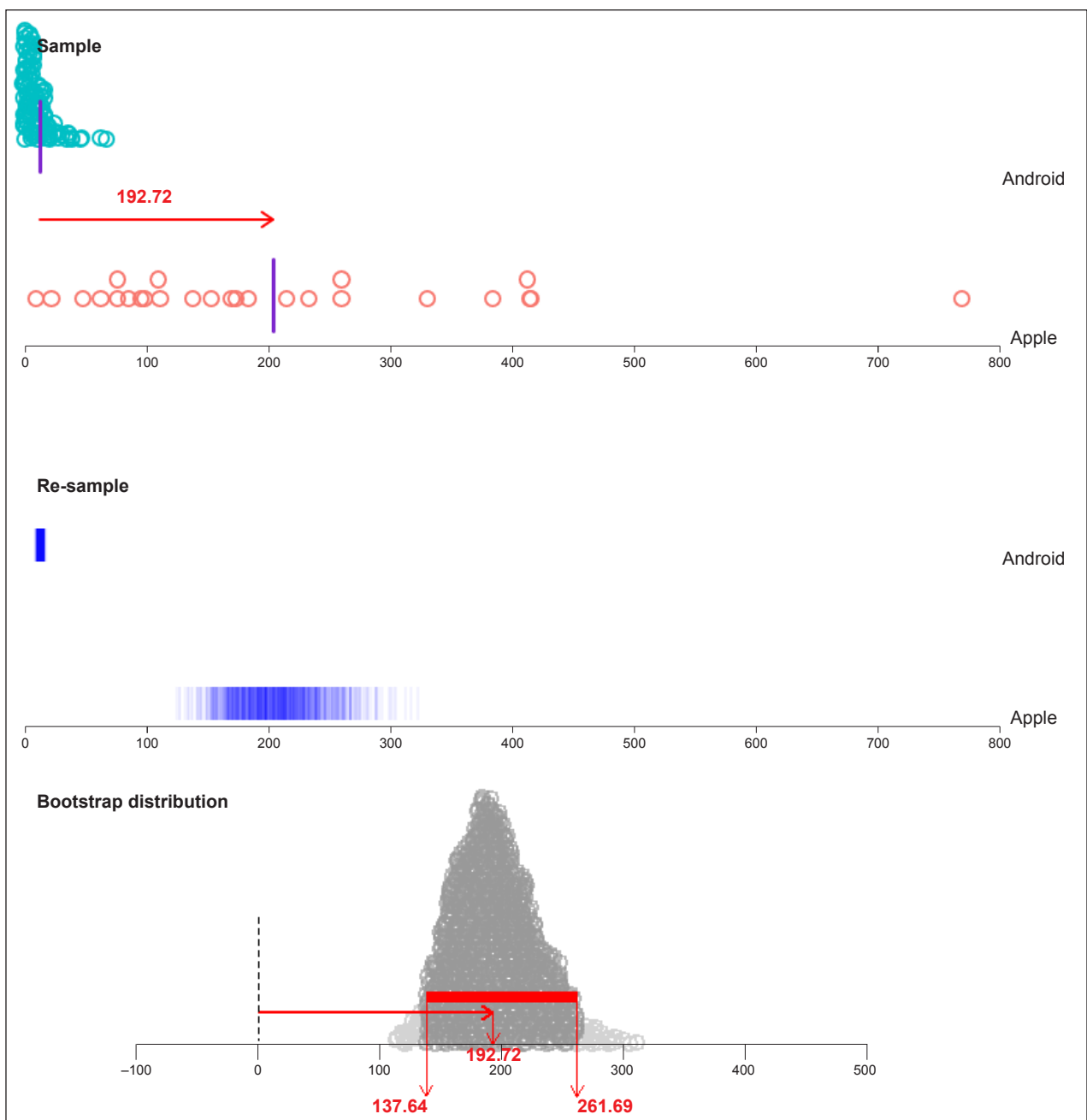
Support your answer with statistical reasoning and appropriate calculations.

QUESTION FIVE

- (a) Common advice for mobile phone app developers is to keep the size of their apps small to avoid issues with users downloading or installing them. The size of each new app released during one randomly selected hour on one randomly selected day during 2021 was recorded for mobile phone apps sold on both the Google Play store (Android) and the iOS App store (Apple). The size of the app was measured in megabytes (MB).
- (i) These samples were used to construct a bootstrap confidence interval for the difference between the mean size of Android mobile phone apps and the mean size of Apple mobile phone apps.

The output from this analysis is shown in Figure 10.

Figure 10: Bootstrap confidence interval output



Discuss what can be concluded from both the features of the sample data distributions and the confidence interval constructed using the sample data.

- (ii) A student wants to find out the average size of the largest mobile phone app students at their school have installed on their phones, and whether the type of phone (Apple or Android) is related to this.

Apply the steps of the statistical enquiry cycle to this situation, giving a short description of what each step would involve.

- (b) A study was conducted to explore how people charge their mobile phones. The study took place over a period of four weeks. During this time, anonymous battery information was collected from Android mobile phones.

- (i) At 8 p.m. on one day during the study, the mean battery charge percentage for participants who were not currently charging their phones was 52.4% and the standard deviation was 10.6%.

Assuming the distribution of battery charge percentages can be modelled by a normal distribution, calculate an estimate for the probability that the battery charge percentage of an Android phone not currently being charged at 8 p.m. is over 60%, given that the battery charge percentage is under 70%.

- (ii) Discuss how another factor could be taken into account when modelling the probability for part (i).

- (iii) Participants in the study charged their phones using their computers 39% of the time. The rest of the time they plugged their phones into power outlets to charge them.

In 23% of charging instances, the phone was unplugged from the computer or power outlet within the first 30 minutes after the battery had been fully charged.

Of these charging instances when the phone was unplugged from the computer or power outlet within the first 30 minutes after the battery had been charged, 60% of the phones were being charged using a computer.

Use this information and relevant calculations to compare how participants in this study charged their mobile phones.

93201Q