that are late because of formatting issues or because a version is incomplete, will not be accepted.

Assignment Tasks and Questions

Exercise 1 (34 marks = 4 + 18 + 4 + 8)

One of the issues that usually come up in the US presidential elections is how to deal with a sluggish economy. Specifically, should the government *cut* spending, *raise* taxes, *inflate* the economy by printing more money, or do none of the above and *let* the deficit rise? And as with most other issues, politicians need to know which parts of the electorate support these options.

Suppose that a random sample of 1,000 people was asked which option they support and their political affiliations. The possible responses to the question about the preferred economic option were the four options mentioned above (i.e., *Cut*, *Raise*, *Inflate*, and *Let*), and to the question about political affiliation the respondents could answer *Democrat*, *Republican*, or *Independent* (which included a variety of political persuasions). The responses are summarised in the table below.

	Affiliation		
Option	Democrat	Republican	Independent
Cut	101	282	61
Raise	38	67	25
Inflate	131	88	31
Let	61	90	25

Do these data allow us to conclude at the 1% significance level that political affiliation affects support for the economic options?

(a) Describe the variables of interest? Are they qualitative or quantitative? If they are qualitative, are they ranked or unranked? If they are quantitative, are they discrete or continuous? What are the levels of measurement? Explain your answers.

(b) Answer the research question by performing all necessary calculations manually. State the hypotheses, the decision rule, calculate the test statistic showing the details of your calculations, make a statistical decision with reference to the proper critical value and draw your conclusion. Based on the relevant critical value table, what can you tell about the *p*-value of your test? Explain your answer.

(c) Repeat the test you performed in part (b) with *R*. Evaluate the test using the reported *p*-value this time.

(d) List all requirements that must be met to validate the test in parts (b) and (c). Are they likely satisfied this time? Explain your answers.

Exercise 2 (31 marks: 6 + 4 + 8 + 13)

Consumer Research is an independent agency that conducts research on consumer attitudes and behaviours for a variety of firms. In one study, a client asked for an investigation on consumer characteristics that can be used to predict the amount charged by credit card users. Data were collected on annual income (*Income*, $1000), household size (*Size*), and annual amount of credit charged to the credit card (*Credit*, $) for a random sample of 50 customers. These data are saved in the *a3e2.xlsx* file.

(a) For each of the three variables, answer the following questions. Is the variable qualitative or quantitative? If it is qualitative, is it ranked or unranked? If it is quantitative, is it discrete or continuous? What is its level of measurement? Explain your answers.

(b) Consider the following two pairs of variables: *Credit* and *Income*, and *Credit* and *Size*. In each case, do you expect the variables to be related to each other? If yes, do you expect the relationship to be positive or negative. Explain your answers.

(c) Using *R*, calculate the Pearson or Spearman correlation coefficient, whichever is more appropriate, for the two pairs of variables in part (b). In each case, briefly explain your choice between the Pearson and Spearman correlation coefficients and comment on the direction and relative strength of the relationship as implied by the point estimate.

(d) Based on your answers in part (b), perform an appropriate test with *R* at the 5% significance level on each pair of variables to determine whether there is a linear, or at least monotonic, relationship between the variables in the expected direction. In each case, show the hypotheses and state the statistical decision and the conclusion.

Exercise 3 (35 marks: 4 + 6 + 8 + 6 + 11)

Refer to Exercise 2.

(a) Estimate a multiple linear regression model with *R* for the annual credit card charges using annual income and household size as independent variables. Write out the sample regression equation using the actual variable names.

(b) Interpret the adjusted coefficient of determination. Is the model likely to be useful in predicting the annual credit card charges? Explain your answer.

(c) Test the overall significance of the model at the 1% significance level. What are the hypotheses, the observed test statistic, the statistical decision, and the conclusion? Be precise.

(d) Based on your expectations in part (b) of Exercise 2, test appropriate hypotheses concerning the slope coefficients using *t*-tests at the 1% significance level. What are your decisions and conclusions?

(e) Interpret the slope coefficients that you deemed significant in the expected direction. Does the *y*-intercept estimate have a logical interpretation this time? If yes, what is it? If not, why not?