***Instructions for handing in. Please...***

1. Attempt **ALL** questions.
2. Assignments may be typed in a word processor of your choice, or handwritten **neatly**.
3. Set answers out in order of the questions. Do **NOT** jump between questions.
4. There is no need to copy the questions out in your submission.
5. If answering a question using R, all commands that you enter should be provided as part of the answer, along with all relevant output. Include your ID number on output, where possible (e.g. as part of the title of a graph).

**Submission**: Assignments must be submitted using Canvas as a **single PDF** file, before the due date and time. Handwritten assignments will need to be scanned. **Prepare your assignments well in advance of the deadline in case of technical issues, as no extensions will be provided in this case.**

**Notes**:

- Summarising, analysing and communicating information is an important part of Operations Research. For this reason you will be expected to write answers which clearly communicate your thoughts. The mark you receive will be based on **your written English** as well as your technical work – review the relevant section in the Course Outline.
- **We encourage working together.** Discussing assignments and methods of solution with other students or getting help in understanding from staff and students is acceptable and encouraged. You must **write up your final assignment individually, in your own words.**
- By submitting this assignment, you confirm that you understand the University's policies on cheating, plagiarism and group work; that your submission is entirely your own work and you have not allowed access to any part of the assignment to any other person. **See the appropriate sections in the Course Outline for more details.**
- This assignment consists of **TWO** questions, and is marked out of **50 marks**.
  This assignment makes up **6%** of the final assessment for this course.

**1. Clear Blue Water Bottling (25 marks)**

Clear Blue is a company specialising in boutique water products. It recently signed a deal with the Waikato district council to enable the extraction of water from the local aquifers. Clear Blue is going to specialise in one particular product: the 1 litre bottle of mineral water, but needs to decide how large a bottling plant it should build to maximize its annualised profit. It is considering three plant sizes: 2 million bottles per year (`small`), 5 million bottles per year (`medium`), or 10 million bottles per year (`large`). The (annualised) plant costs are $4 million, $6 million and $8 million, respectively. The extraction costs are 10c per litre and there are 20c per litre of transportation, and bottling costs, regardless of the plant size.

The total international demand for New Zealand water is uncertain. However, Clear Blue predicts that there will either be `low`, `average`, or `high` demand. If Clear Blue builds a large plant then prices may be lower than with a larger plant (due to market saturation). These prices, which are given in the table below, are determined after the water has been bottled, and offered to the market.

| | Plant Capacity | | |
|---|---|---|---|
| | `small` | `medium` | `large` |
| `low` demand | $3.00 / litre | $1.50 / litre | $1.00 / litre |
| `average` demand | $3.50 / litre | $2.50 / litre | $1.50 / litre |
| `high` demand | $4.00 / litre | $3.00 / litre | $2.00 / litre |

(a) Construct a payoff matrix for this situation, with the rows being the decisions and the columns the states of nature.

(b) What is the best decision if Clear Blue is optimistic?

(c) What is the best decision if Clear Blue is pessimistic?

(d) What is the best decision if Clear Blue wishes to minimise their maximum regret?

Suppose that, based on preliminary market research by Clear Blue, it is estimated that there is a 0.2 probability that demand will be `low`, a 0.3 probability that demand will be `average`, and a 0.5 probability that demand will be `high`.

(e) Draw a decision tree for this situation, and compute the best decision, assuming that Clear Blue is risk-neutral.

(f) Compute the expected value of perfect information. Show all working.

Analytics experts approach Clear Blue and suggest that it's possible to conduct additional market research to get a better idea of the demand for these boutique water products. This market research can yield either a `positive` or `negative` result; the (conditional) probability distributions for demand are given in the table below.

| | Market Research | |
|---|---|---|
| | `negative` | `positive` |
| `low` demand | 0.400 | 0.150 |
| `average` demand | 0.600 | 0.225 |
| `high` demand | 0.000 | 0.625 |

(g) Compute the probability that the market research will be `positive`, in order for all the probabilities to be consistent.

(h) How much would Clear Blue be willing to pay for this market research to be conducted?

Suppose that Clear Blue is risk averse, with utility function: $U(\pi) = \sqrt{\pi + 2}$, where $\pi$ is the payoff in $millions.

(i) If making the investment decision, without the additional market research, compute Clear Blue's expected utility for each plant capacity. Which should Clear Blue choose, and why?

(j) What is the CEV associated with the decision determined in (i)?

(k) Given this utility function, how much would Clear Blue be willing pay for the additional market research, described above? This question can only be computed numerically. **(2 bonus marks)**

2. **New Zealand Electricity Sector (25 marks)**

The New Zealand electricity sector comprises is a multi-nodal market, with generators, retailers, transmission and distribution. In this question we will be looking at the generation of power in 2018.

Generation data is collected on a half-hourly basis from 64 generation sites across New Zealand. We wish to use this data to investigate how different generation technologies contribute to New Zealand's electrical energy mix.

For this question you need to first install RStudio, and download `generation.csv` and `plants.csv` from Canvas. Remember to first load the `tidyverse` library, and then import `generation.csv` and `plants.csv` into R.

(a) Use the match function to create `Fuel`, `Region` and `Renewable` attributes in the `generation` data set, by matching the nodal information with the `plants` data set.
**Hint:** for `Fuel` you will need to complete this command:
`> generation$Fuel=as.factor(plants$Fuel[match(_______,_______)])`

(b) Filter the data in the `generation` data set to create a new data set called `windgen` containing only data about wind generators.

(c) Create a `pairs` plot that compares wind generation in different trading periods each day. Since 48 half-hour trading periods are too many to show on a single plot, select 6 trading periods to include in your pairs plot. Comment on any observations.

(d) Use the `gather` function to reshape the `generation` data set to have a single generation output attribute `MWh` for all trading periods, you should set the `key` to be `TP`).
**Hint:** you can use the `starts_with(___)` function, instead of listing all columns to 'gather'.

(e) Use the `group_by()` and `summarise(______,____(____,na.rm=TRUE))` functions to display the total generation by `Fuel` type.

(f) Create a stacked bar chart comparing the generation output for each of the `Regions` (the `fill` colour should be set to the `Fuel` type). Comment on any observations.
**Hint:** you will need to use the `group_by()` and `summarise()` functions again.

Use the following command to convert the values in the `TP` attribute to integers:
`> generation.gather$TP=as.integer(gsub("TP","",generation.gather$TP))`

For the following questions, first import the `lubridate` library into R and use the command:
`> generation.gather$Date=parse_date_time(generation.gather$Date,"d/m/Y")`
in order to parse the `Date` attribute.

(g) Create a line graph showing the half-hourly average generation of trading periods 1–48 with different coloured lines for each month. Comment on any observations.
**Hint:** first create a new attribute in the data set for the month (set as a `factor`), and then use the `filter()`, `group_by()`, `summarise()` functions to prepare the data before plotting.

For the following questions, load the `leaflet` package, and import the `regions.csv` file.

(h) Using `leaflet` create a map of New Zealand, with a circle located at the coordinates for each region (as given in the `regions.csv` file).

(i) Use the `group_by()` and `summarise()` functions to compute the total renewable and fossil fuel generation in each region.

(j) Use the `spread()` command to separate the renewable and fossil fuel generation values into separate attributes, storing the resulting tibble as `totalgen.spread` and replace `NA` values with `0` using the command `totalgen.spread[is.na(totalgen.spread)]=0`.

(k) Modify the map created in (h) so that the ***areas*** of the circles are *proportional* to the total generation in that region, and the colour is based on the proportion of generation in that region that is renewable; show a legend on the map. (Additional attributes will need to be created in `totalgen.spread` to complete this, and the `match()` function will need to be used.)