Week 1

LATEST SUBMISSION GRADE

100%

1. We have seen that the modern age is largely being driven by data. Which of the following is   **1 / 1 point**
    NOT a factor contributing to data being "the new oil"?

    ○ Modern technological advances allow data to be collected in a much larger scale than before.

    ○ Businesses are starting to realise that existing (and new) data can generate valuable insights to help them make
        better decisions.

    ⦿ Increasingly sophisticated statistical models and software are being developed to handle large volumes of data.

    ○ Modern technological advances, in particular advances in processing power, allow more sophisticated analyses of
        much larger datasets.

    ✓ **Correct**
        Correct. Advanced statistical models and software for handling data are more of a consequence than a cause for the
        importance of data.

2. Which of the following is NOT a reason for the boom in data-related jobs?   **1 / 1 point**

    ○ The idea of (and ability to) studying large amounts of data (in the context of this course) is still relatively new. The
        supply of data scientists has not kept up with demand.

    ○ The correct use of data can lead to highly profitable decisions.

    ⦿ Being able to work with data is a highly transferable skillset; the skills of an individual data scientist (say) are not
        restricted to any single industry.

    ○ Many businesses have accumulated large amounts of data that was previously unused.

    ✓ **Correct**
        Correct. The fluidity of the data science skillset would affect the supply of data scientists, not the demand for them.

3.
    What is involved in the process of "data cleaning"? Select the option that best describes this
    term.

    ○ If the data is missing some values, then you will either need to estimate those values or delete the entry
        corresponding to that missing value.

○ If the data available to you is not in a nicely structured format, then you will have to edit the data into a better format that will allow further analysis.

○ Your data may have genuinely erroneous entries, and you will need to either delete those entries or correct them.

⦿ All of the above.

✓ **Correct**
   Correct.Watch the video: Data value chain

4. Which of the following is NOT an example of collecting data, in the context of this course?   **1 / 1 point**

○ Supermarket reward programs that track what shoppers are buying.

○ Internet service providers (ISP's) collecting metadata on websites visited, duration, etc. by their customers.

⦿ A company using existing sales records and client data to advertise to existing customers.

○ A bank recording your spending on a credit card.

✓ **Correct**
   Because the company is using EXISTING records, no new data is being collected here.Watch the video: Data value chain

5. Which of the following is the clearest example of generating an insight from data that has been collected, in the context of this course?   **1 / 1 point**

○ A sales manager brought in from another country to restructure the sales department.

⦿ A social media site collecting data on the advertisements you click in order to decide which future advertisements to display.

○ An architect asking their client about their needs before drafting plans for a new building.

○ A large construction company keeping detailed records of their previous projects and clients.

✓ **Correct**
   Correct.

6. Which of the following is NOT a good reason for why a data analyst/scientist may generate visualisations from a dataset?   **1 / 1 point**

◯ "A picture is worth a thousand words"; a single visualisation can often convey an idea much clearer than a series of calculations and numbers.

◯ It allows people that may not have any technical expertise to still be able to understand a concept and be able to use the results of an analysis.

◉ Visualisations better allow the use of colour to make a presentation more interesting (say, to a potential investor).

◯ A visualisation can help the data analyst/scientist better understand the dataset, allowing more accurate methods to be chosen for the analysis/processing of the data and better conclusions to be drawn.

✓ **Correct**

Correct. The use of colour is at best a secondary benefit of using a visualisation.

7. What is the main difference between a data scientist and a data analyst?   `1 / 1 point`

◯ These titles are completely interchangeable.

◯ A data scientist typically works in a government role such as academia, whereas a data analyst typically works in private industry.

◉ A data analyst will typically look at a dataset and find trends, create charts and other visualisations to help with decision-making, whereas a data scientist typically works in more depth with data, often using more sophisticated methods for modelling and predicting.

◯ Data analysts generally have a different set of tertiary qualifications as compared to a data scientist; the former would typically have a statistics degree, whereas the latter would typically have a degree in computer science.

✓ **Correct**

Correct. Watch the video: Job roles available in the data insights market

8. For the practice task 1 involving the mtcars csv file, suppose we wanted to see if there was   `1 / 1 point`
any effect of the number of cylinders on a car's mileage. Which of the following would NOT
be an appropriate graph to use?

◯ A line chart

◯ A scatter plot.

◯ A bar chart

◉ A pie chart.

✓ **Correct**

A pie chart would be inappropriate as we are looking for the trend of mileage across the number of cylinders.