Mathematical Foundations of Data Science
Assignment 1

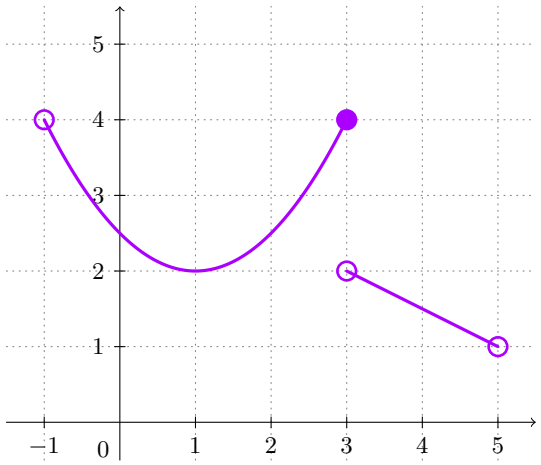
Trimester 2, 2024

1. Consider the following statement in set notation:

$$\forall y \in \mathbb{Q},\ \lfloor y \rfloor < y$$

Explain in words what this statement means, and state whether it is a true or false statement, giving justification.

2. Consider the function g whose graph is below.



   (a) What is the domain of g?

   (b) Is g a one-to-one function? Justify your answer.

   (c) Let f be the function obtained by restricting g to the domain $[1, 5)$. Then f is one-to-one. Sketch the graph of f^{-1}.

3. Let $A = \{0, \frac{1}{4}, \frac{3}{2}, 2, \pi\}$ and $B = \{x \in \mathbb{R} \mid -4 < 3x \leq 5\}$.

   (a) What is $A \setminus \mathbb{Z}$?

   (b) What is $A \cap B$?

4. Consider the function

$$h(x) = \frac{x+3}{2},\ x \geq 1.$$

Find $h^{-1}(x)$, clearly stating the domain of h^{-1} and showing all working.

5. You should complete this question using a Jupyter Notebook. All of the code you will need to complete this question can be taken directly or generalised from the week 1 practical, or will be given to you in the question.

   Download the file `swimming.csv`[1]. This file contains data about swimmers who competed in the men's 100m Freestyle at the 2020 Summer Olympics. Use `Python` to do the following:

   (a) Using `pandas`, read the data into a dataframe and print out its `head()`.

   (b) Create a histogram of all times (`Time`) in the dataset.
   Hint: Remember to add axis labels.

   (c) Calculate the mean time in the dataset. Print out a statement showing the mean time, rounded to the nearest 3 decimal places.
   Hint: In the week 1 practical, we saw how to print out a line containing text and numbers. The `round()` *function might also be useful here.*

   (d) Find the swimmer who swam the fastest time in the heats.
   Hint: Try creating a new dataframe that only contains the heat times for swimmers.

   (e) Find the swimmer who swam the slowest time in semifinal number 2.

   Present your answers as a full Jupyter Notebook. Your notebook must include code to find the results, and text answering the questions based on the output of your code. Download this notebook and convert to a PDF and submit with your assignment.

   Please note: You must include the code you used to find results. Each answer submitted without code will receive a mark of 0.

   Hint for submitting: You can "Download As PDF" in Jupyter, but that may not work on your computer. If it doesn't, you can download as HTML and convert that to a PDF. Make sure you join it to your assignment to make a single PDF when submitting! You might want to try googling things like "convert html to pdf" and "combine multiple pdfs". There is also a video in the Python Module on MyUni demonstrating how to save a Jupyter Notebook as a PDF.

[1] This data was partially sourced from Wikipedia.