Assignments for students

Group Essay

- Write an essay on how next-generation sequencing has been used to research a specific cancer of your choice. What has the impact been? (2000 word limit)

Group Presentation

- Your group will be assigned a cancer database to analyse. Find more about this task in 3 slides.

The multi-omic methods covered

- Bulk DNA analysis
- Bulk RNA analysis
- Proteomics (e.g. western blotting)
- Single-cell and single-nucleus sequencing
- Epigenomic (e.g. ATAC-seq and ChIP-seq)

Essential criteria for assignments

- Essays and presentations must be adequately sourced, and a bibliography should be used.
- References should follow a Nature referencing style:

Author(s). Title of article. *Title of Journal* **Volume number**, Page range (Year).

- Essays and presentations should answer the question/problem.
- Essays and presentations should be formatted (Times New Roman, size 12, justified, 1.5 spacing)
- Images and figures should be sourced and labelled.
- For presentations, all members of the group need to speak.

Essay Requirement

- Provide an introduction(or abstract), main body and conclusion
- Give essay a title.
- The main body should contain subheading where needed.

Data Analysis Presentation Groups

- Calista and Audrey: Alzheimer's dataset
- https://r-packages.io/datasets/alzheimer
- Andy and Elaine: Pima Indians Diabetes Database
- https://r-packages.io/datasets/PimaIndiansDiabetes
- Kate, Iris, Leco: BreastCancer
- https://r-packages.io/datasets/BreastCancer

Data Analysis Presentation Requirements

- Your group have been assigned a dataset to analyse.
- I would like you to conduct some research on the data you have been assigned and add this to your presentation:
- Provide background information on the data set (e.g. what the data shows, why it's useful, etc).
- Explain each step of your data analysis

Data Analysis Presentation Requirements

- The data analysis steps I would like you to present:

1. Upload data set (read data into R)
2. Clean data (i.e. remove and NAs)
3. Create a new column(s) based on a loop and/or conditional statements
4. Select columns and assign to a new data frame.
5. Filter data and assign to a new data frame.
6. Write one of these new data frames into a csv file.
7. You must create at least the following plots of the data using ggplot2 (you can present any part of the filtered data):
   - Box plot
   - Scatter plot with a line of best fit
   - Histogram

Data Analysis Presentation Requirements

- You are welcome to make more plots if you like.
- Plots must contain axis names, titles and colours.
- At least 1 plot must be faceted.
- Use head() where relevant to show changes made to data frame (if showing the whole data frame is impossible).

Important dates

- **Sunday 21st July:** The group essay must be handed in by the evening.
- **Monday 22nd July:** Group presentations are held on the final day of the summit.

Plagiarism and use of AI

- Plagiarism is strictly prohibited and any students found making use of use of sources without proper citation will not be eligible for awards.
- The use of AI tools such as chatbots and virtual assistants is strictly prohibited, and any students found making use of them will not be eligible for awards.

Assignment assessment criteria

- Understanding of the subject
- Engagement with literature
- Satisfying question asked
- Originality of thought

Assessment awards

- 1 Best Group Presentation
- 1 Best Group Essay
- 1 Outstanding Student
- 1 Academic Excellence