

Day 2 STEM Games Science Arena

Day 2 description

Task 1.

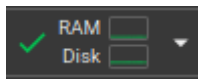
For day 2 you have gathered gene expression data from 5 Alopecia Areata patients and from 5 healthy control patients. Alopecia Areata is a disease that happens when the immune system attacks hair follicles and causes hair loss. There is no cure for this disease, but there are treatments that help with hair regrowth.

In order to answer first part of tasks for Day 2, you will have to run the code found on this link:

<https://colab.research.google.com/drive/12cIZDvAV-KXI0skpO8Xk2xrOePzVkssU?usp=sharing>

The code is written in the R, a programming language and software environment commonly used in life sciences research. It is an open-source, free-to-use software that is widely used for statistical analysis, data visualization, and modeling in a variety of fields, including genetics, genomics, proteomics, and bioinformatics.

In order to run the analysis firstly you will need to make a copy of the R notebook into your drive. To do so, select File and Save a copy in Drive. Next, press connect on the top right corner of the screen. This may take some time and when this appears:



you are ready to run the code.

Parts of the R notebook that contain code are called cells or chunks. You should run them one by one and in the order that they appear in the notebook. To run a cell select



the button on the left side of the cell which will appear when you hover your mouse over the cell.

After running a cell some output will appear underneath. Some of the outputs will be just log type messages which are not particularly important but some outputs will contain plots or other useful information. Additionally, some of the questions are general questions about statistics/data science which require you to do some research, and others just require you to look around in the notebook.

Questions from the R notebook.

?Q

Q1. What is GEO?

Q2. What can you tell about your samples (Platform, Extracted molecule, Characteristics)?

Q3. How many samples are there under the accession number GSE67492?

Q4. What is the difference in Experiment type between GSE229973, and GSE227774?
List two advantages of each method.

Q5. What is normalization of gene expression data? What would a boxplot of non-normalized data look like?

Q6. What is principal component analysis? Describe in one sentence what you can conclude from your PCA plot.

Q7. What is the main goal of differential expression analysis? What p-value was chosen for the analysis?

Q8. Interpret the results from the command: `head(temp)`.

Task 2.

Design a **biological experiment(s)** to test the importance of highlighted genes in differential analysis.

A proposed outline:

- Brief summary (max 300 characters without spaces)
- Introduction and overview of previous research (max 1000 characters without spaces)
- Research aim (max 250 characters without spaces)
- Research hypotheses (max 250 characters without spaces)
- Materials and methodology, focusing on the type of experimental model to be used, with brief justification of why it is appropriate and details of exactly how you would introduce/mutate/remove gene(s) in the proposed system (max 2000 characters without spaces)
- Research plan with diagram of the experimental protocol and details of how the data would be analyzed (max 2000 characters without spaces)
- Expected scientific contribution of the proposed research (max 300 characters without spaces)
- References (use AMA citation style)

Helpful references for inspiration:

- Fu Y, Zhao D, Zhou Y, Lu J, Kang L, Jiang X, Xu R, Ding Z, Zou Y. Identification of Differential Expression Genes between Volume and Pressure Overloaded Hearts Based on Bioinformatics Analysis. *Genes*. 2022; 13(7):1276.
<https://doi.org/10.3390/genes13071276>
- Chen X, Ding Q, Lin T, et al. An immune-related prognostic model predicts neoplasm-immunity interactions for metastatic nasopharyngeal carcinoma. *Front Immunol*. 2023;14:1109503. Published 2023 Mar 31.
doi:10.3389/fimmu.2023.1109503
- Yao Z, An W, Tuerdi M, Zhao J. Identification of novel prognostic indicators for oral squamous cell carcinoma based on proteomics and metabolomics [published online ahead of print, 2023 Apr 19]. *Transl Oncol*. 2023;33:101672.
doi:10.1016/j.tranon.2023.101672

Task 3.

Design a **project proposal** for a potential therapeutic approach based on the results from previous experiments.

Proposed outline:

- Introduction: A short disease background (pathophysiology, pathways, treatment), how your project goes beyond state of the art (max 5000 characters without spaces)
- Objectives of the project: what you plan on discovering/obtaining during the project (max 1000 characters without spaces)
- Methodology: the technologies and methodologies you plan on using and for which purposes, what is the type of data you will obtain with it; if possible, description of statistical and bioinformatics approaches are welcome (max 1500 characters without spaces)
- Impact: describe the possible impact success of your project could have on the improvement of science in the field, as well as wellbeing of patients suffering from the disease (max 1000 characters without spaces)
- Schedule: organize in the form of a Gantt chart the approximate deadlines for each stage of the project (subject recruitment, sampling, experimental work, dissemination – conferences, manuscript writing...)
- Possible pitfalls: list possible pitfalls of the project which could impede its progress (for example, sample contamination, etc) and explain how you would get back on track with the project in case of these situations (max 500-1000 characters without spaces)
- Budget: make a spreadsheet and allot a percentage of the total financial sum for each stage of the project mentioned in the schedule. Also try to allot additional percentages, signified in brackets next to the original percentage, e.g. 20 (+5)% in case of possible pitfalls.
- References: use AMA citation style

Helpful references for inspiration:

<https://scientific-publishing.webshop.elsevier.com/research-process/writing-scientific-research-project-proposal/>

<https://science.yalecollege.yale.edu/stem-fellowships/how-write-proposal>

<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2981889/>

<https://www.nature.com/articles/d41586-019-03914-5>