**MBG 6113 Midterm Exam (Take-home)**

**Name: Gözdem Çavdar**

**The exam papers should be sent to** [**ezgi.karaca@ibg.edu.tr**](mailto:ezgi.karaca@ibg.edu.tr)

**until Dec 14th 2022 midnight**

Please go to the github repo of the course book (<https://github.com/vsbuffalo/bds-files>).

**Q1:** Fork the repo to your own github account. Share the link of the cloned repo. (**10 points**)

[**https://github.com/gcavdar93/bds-files**](https://github.com/gcavdar93/bds-files)

Regarding the related chapter folders given in the repo:

**Q2:** chapter-00 (**10 points**):

Do you have Vim and Emacs installed on your computer? If not, please install them. Try each editor to enter a sample text into a new file. Describe with examples your experience.

* **Vim:**

**1-) I opened iterm2 and wrote:**

**which vim**

**The result is:**

**/usr/bin/vim**

**Then, I have vim installed on my computer.**

**2-) I wrote this command in order to create a file named Bioinformatics-Take-Home-Q2 in Downloads folder, and open this file:**

**vim Bioinformatics-Take-Home-Q2**

**3-) I opened insert mode by clicking ‘i’ letter on the keyboard.**

**I wrote the text, then saved the text and quitted from the file.**

**4-) I opened the file directly in the page with this code:**

**cat Bioinformatics-Take-Home-Q2**

**metin içeren bir resim

Açıklama otomatik olarak oluşturuldu**

* **Emacs:**

**1-) I have downloaded Emacs via** [**https://emacsformacosx.com**](https://emacsformacosx.com)**.**

**2-) I found Emacs in Applications folder and opened it via double clicking.**

**3-) A white page is appeared.**

**I wrote something to the page:**

**11.12.2022 Sunday**

**Izmir TR**

**4-) I clicked ‘File’ at the top and ‘Save as’ button.**

**5-) I saved the file in ‘emacs-q2’ name.**

**Q3:** chapter-01 (**20 points**):

Create a repo for a hypothetical tool that computes/predicts something that you currently need (for your daily research). Name the tool and write a README for it, including the hypothetical usage of the tool. While doing so, be mindful about the reproducibility of your research. Explain how you did so. Also share the relevant link.

[**https://github.com/gcavdar93/take-home-repo**](https://github.com/gcavdar93/take-home-repo)

**Q4:** python (**20 points**):

Please go to <https://www.learnpython.org/>. Finish the tutorial until the *Conditions* section. Summarize what you learned and what you liked the most about python in one paragraph.

**I learnt:**

* **how to write a command directly to the command line**
* **how to create a python file by ‘vi’ command and run it**
* **what are the variable types in python**
* **how to create lists in python**
* **how to use the arithmetic operators in python**
* **what are the basic string operations**
* **and how to use the conditions in python.**

**What I liked the most about python is its simplified syntax. That’s why, it is easier to learn compared to other languages. In addition, because it is free, everyone can use it and I can find a code I search easily via Google.**

**Q5:** chapter-06 (**20 points**):

Obtain reference IDs of 20 proteins from Uniprot.

1. Download fasta amino acid sequences of the selected proteins using wget or curl.

2. Obtain checksum numbers of fasta files that you downloaded.

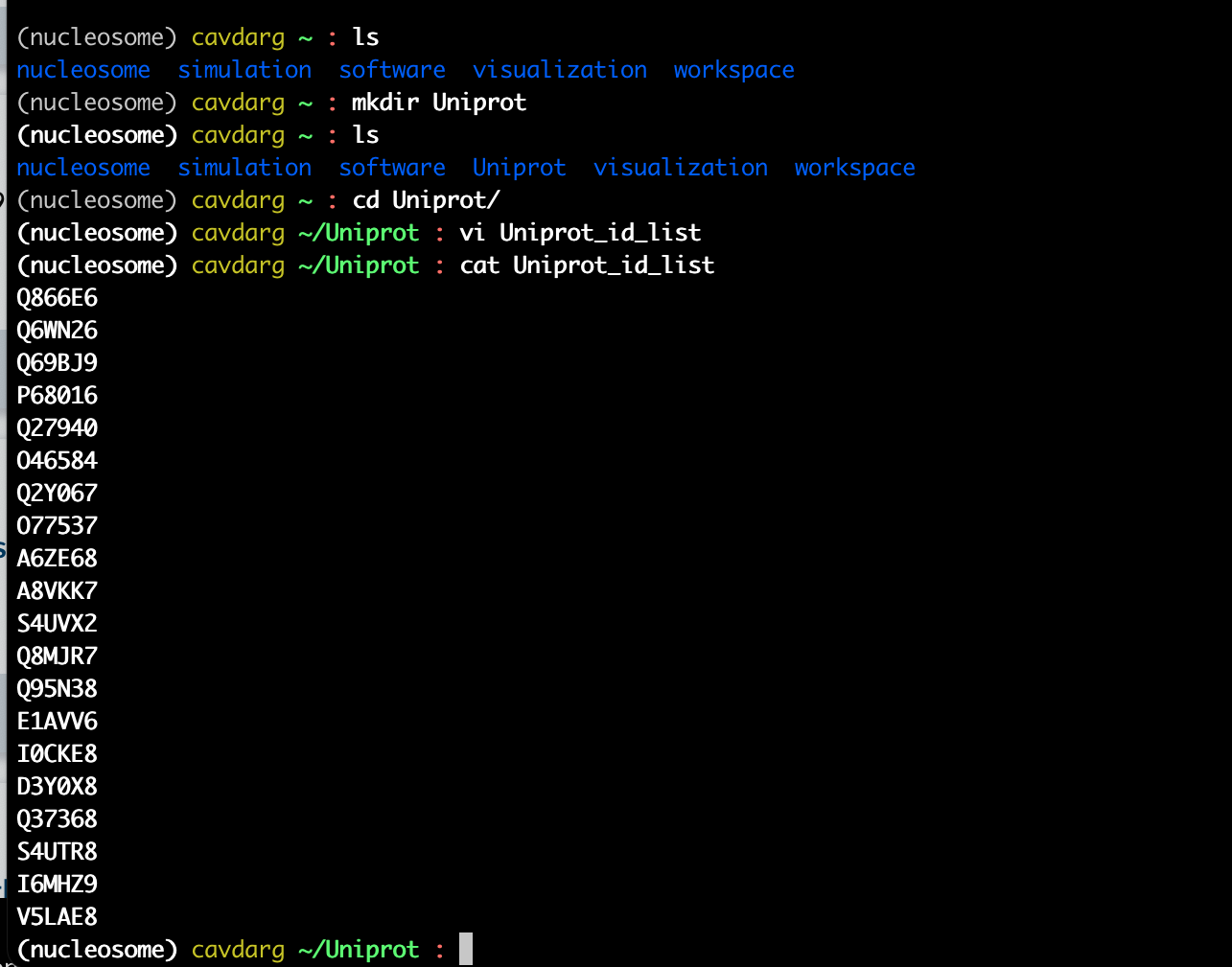
3. Append your name/surname, your lab name to one of the fasta files (keep the

original file) as the first line. Compare the checksums of the new files with the previous ones.

4. Illustrate the differences in the fasta files that were changed with diff.

Include all the comments you used to answer the above-given questions and the relevant screenshots

**1-)**



metin içeren bir resim

Açıklama otomatik olarak oluşturuldu

metin içeren bir resim

Açıklama otomatik olarak oluşturuldu

metin içeren bir resim

Açıklama otomatik olarak oluşturuldu

metin içeren bir resim

Açıklama otomatik olarak oluşturuldu

metin içeren bir resim

Açıklama otomatik olarak oluşturuldu

**2-)**

metin içeren bir resim

Açıklama otomatik olarak oluşturuldu

**3-)**

**metin içeren bir resim

Açıklama otomatik olarak oluşturuldu**

**metin içeren bir resim

Açıklama otomatik olarak oluşturuldu**

**4-)**

**metin içeren bir resim

Açıklama otomatik olarak oluşturuldu**

**Q6:** chapter-07 (**20 points**):

Get the *Mus\_musculus.GRCm38.75\_chr1\_genes.txt* file from chapter07’s github page

**I directly copied the text via cmd-C and pasted it to an Emacs file. Then, I saved it by giving it Mus\_musculus.GRCm38.75\_chr1\_genes.txt name.**

1. What are number of words, lines, and characters of the given file?

metin içeren bir resim

Açıklama otomatik olarak oluşturuldu

1. What is the size of the file in kilobytes?

metin içeren bir resim

Açıklama otomatik olarak oluşturuldu

**32100 bytes = 32.1 kilobytes**

1. What are the first 4 and last 4 lines in the given file?

**First 4 lines:**

metin içeren bir resim

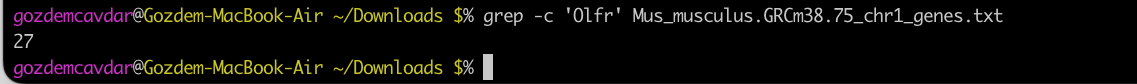
Açıklama otomatik olarak oluşturuldu

**Last 4 lines:**

**metin içeren bir resim

Açıklama otomatik olarak oluşturuldu**

1. Find out how many lines there are containing “Olfr”



5. Make an *Olfr.txt*  file, expanding on the results you obtained for the previous question. The final *Olfr.txt*  file should only contain the gene names, not the gene id’s.

metin içeren bir resim

Açıklama otomatik olarak oluşturuldu