

## Final Project

- **Due: 12/17 18:00**
- Team: Random assignment after survey.
  - 3 or 4 team members
- Topic: Any topic and dataset you want.
  - **You should choose an appropriate dataset that can be applied to Classification, Regression, and Clustering well with one dataset.**
  - Each problem type (Classification, Regression, and Clustering) must be derived from an appropriate problem statement and hypothesis from the dataset.
- Assignment: A report of about 10 pages + Code
  - 1. Problem statement (20)
    - Dataset explanation
    - Hypothesis constructing for classification, regression, and clustering
  - 2. EDA (Data analysis) (5)
    - Type of attributes
    - Visualization
    - Summary Statistics
  - 3. Modeling
    - Classification (20)
      - Pre-processing (cleaning data)
      - Reasons for choosing each model
      - Model selection: Candidate model comparison, Hyper-parameter
      - Evaluation: plots, tables
      - Explanations of the results
    - Regression (20)
      - Pre-processing (cleaning data)
      - Reasons for choosing each model
      - Model selection: Candidate model comparison, Hyper-parameter
      - Evaluation: plots, tables
      - Explanations of the results
    - Clustering (20)
      - Pre-processing (cleaning data)
      - Reasons for choosing each model
      - Model selection: Candidate model comparison, Hyper-parameter
      - Evaluation: plots, tables
      - Explanations of the results
  - 4. Discussion & Conclusion (15)
- There is no example-code for your final project.
- After submitting: Peer-review
  - Every student will write down the 1-2 people who contributed the most.
  - Additional points will be awarded to the person with the highest score in each team. (1-2 points)

### **Notice**

Going forward, we will not be accepting any questions via email other than those posted on Classum. If you have any questions about the class, please post your questions in Classum's questions tag. Please email only for personal reasons.

1	라수련	20211094
	박수민	20211113
	김용민	20211054
	최성훈	20181381
2	김예진	20191055
	심재운	20211173
	고현서	20201015
3	Yerlen Saifulla	20202033
	Beksultan Abdikarimov	20212004
	Rizabek Alpamys	20222002
4	이나영	20191201
	한승찬	20221404
	정재건	20191285
5	홍대의	20201325
	김보민	20211368
	김영진	20211053
	김현서	20211081
6	권기완	20191015
	허찬범	20191327
	조용진	20191293
7	김민겸	20201037
	이희윤	20201252
	이금용	20191364
	이제호	20191369
8	유용근	20191185
	배민옥	20171312
	이희창	20191257
9	고준호	20211010
	김시우	20231071
	김재옥	20181310
	김주영	20201072
10	김진원	20171372
	배병남	20191355
	손수연	20221223
11	이상재	20191365
	이유진	20171194
	장준석	20211288
12	Amirbek Abdirasulov	20222001
	Daniyar Beishekeev	20222003
	Mansur Mukimbekov	20212009
	Temuge Munkhbat	20202024
13	Duc Minh Nguyen	20202026
	Thu Phuong Nguyen	20202027
	Mohammed Osman	20192023
14	Marlen Raushanov	20212010
	Zhaniya Smagulova	20202049
	Yihalem Yimolal Tiruneh	20202039