National University of Computer and Emerging Sciences



Lab Exercise 05

CL461-Artificial Intelligence Lab

|  |  |
| --- | --- |
| Course Instructor | Dr. Kashif Zafar |
| Lab Instructor (s) | Saddam Khalil  Ahmad Abdullah |
| Section | D |
| Semester | Spring 2021 |

Department of Computer Science

FAST-NU, Lahore, Pakistan

# Exercise (20 Marks)

Attempt the exercise given below.

## Exploratory Data Analysis (20 Marks)

Perform Exploratory Data Analysis (EDA) using NumPy and Pandas on the dataset provided.

Read the **adult.data** csv file into your notebook and write the code to answer the following questions:

1. Count the number of men and women in this dataset?
2. Find the average age of women in this dataset?
3. How many United-States citizens are present in the dataset?
4. Find the standard deviation and mean value of the age of those who have salary more than 50K per year and those who receive less than 50K per year?
5. Find out (Yes/No) whether people who receive more than 50k salary have at least high school education? (High school education categories in the dataset- Bachelors, Prof-school, Assoc-acdm, Assoc-voc, Masters or Doctorate feature)
6. Using groupby() and describe(), display statistics of age for each race and each gender. Find the maximum age of men of Amer-Indian-Eskimo race.
7. Among whom the proportion of those who earn a lot(>50K) is more: among married or single men (marital-status feature)? Consider married those who have a marital-status starting with Married (Married-civ-spouse, Married-spouse-absent or Married-AF-spouse), the rest are considered bachelors.
8. What is the maximum number of hours a person works per week (hours-per-week feature)? How many people work such a number of hours and what is the percentage of those who earn a lot among them?
9. Count the average time of work (hours-per-week) those who earning a little and a lot (salary) for each country (native-country).

# Submission Instructions

Always read the submission instructions carefully.

* Rename your Jupyter notebook to your roll number and download the notebook as **.ipynb** extension.
* To download the required file, go to **File->Download .ipynb**
* Only submit the **.ipynb** file. DO NOT **zip** or **rar** your submission file
* Submit this file on Google Classroom under the relevant assignment.
* Late submissions will not be accepted