***CMPE230 / HOMEWORK-2***

**Problem :**

In this project, we developed a Python program called bucourses.py that will crawl Bogazici University’s OBIKAS registration pages and extract course offering information. Program should print all departments for asked semesters.

We had to handle some problems such as:

* Some departments doesn’t exist in some semesters
* Duplicate course names are not allowed but one course can have multiple section, so we tried to clean section parts while recording the distinct instructors in the courses

**Solution :**

Pandas is an open source, BSD-licensed library providing high-performance, easy-to-use data structures and data analysis tools for the [Python](https://www.python.org/) programming language. We used Python’s Pandas library to crawl all courses. We describe two main methods as **singleterm** and **processdepartment.**

* **Singleterm** is used to firstly processing semesters, we regulated the semesters given as input (2018-Fall 2019-Spring) to be appropriate for the website link (2018/2019-1, 2018/2019-2) which is used to read html code of the schedules and transform them to data frames.

In this method, the total instructor set is also calculated and the graduate and undergraduate courses are counted in the decided semester for one department.

* **Processdepartment** takes initial and final semesters, code of the department and long name of the department as parameters.

We loop through the initial semester to final semester and add them to the termlist if the link is legal (in case of an closed department in some semester ), it is checked by try except block.

At the end of the **processdepartment** we called **singleterm** with wanted department and semester list, then for all the all semesters in the cleantermlist we call **singleterm** and merge all semesters.

Finally, we call **processdepartment** for all the departments in the deptlist and merge them.

For instructors, in each semester we create an instructor set consists of distinct instructors then we merge them for all semesters

To calculate semester number in total offerings while traversing in termlist we append all courses in a different table (for all semesters) Then calculate which course appeared how many times and record it in total offerings.