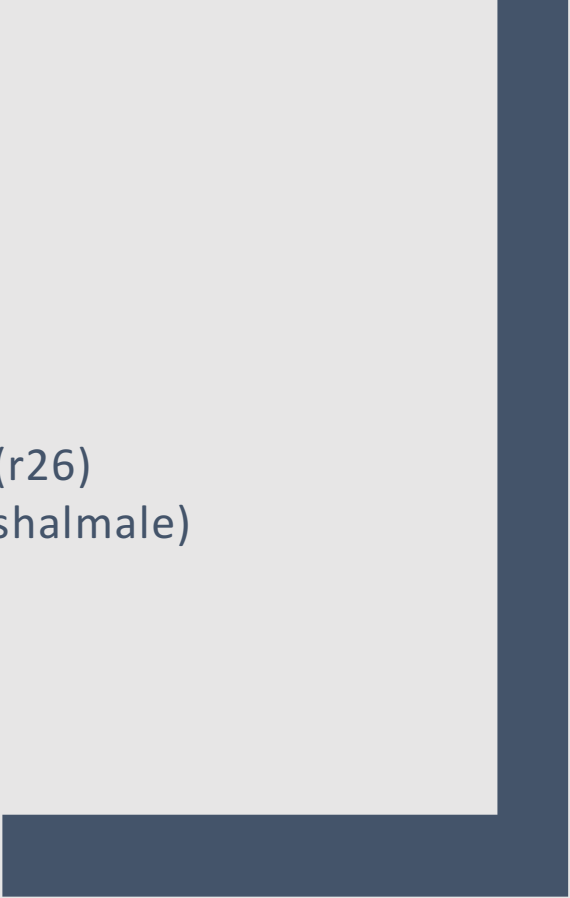




Project 2 Part 2

DESIGN OF QUESTIONS

Submitted by:
Ranganath Sundar (r26)
Shalmalee Vishwasrao (shalmale)



Questions

CourseSchedule1.java

- 1) What is the total enrollment for all subjects per year?
- 2) How has the enrollment trend been over the years? By what number has enrollment increased or decreased between consecutive years?
- 3) What is the maximum amount by which enrollment has increased between 2 consecutive years? What is the corresponding period?
- 4) What is the maximum amount by which enrollment has decreased between 2 consecutive years? What is the corresponding period?

CourseSchedule1.java uses 2 mappers and 2 reducers. While the first MapReduce computes the total enrollment, the second one is used for determining how the enrollment trend has changed over the years. From the output of the second MapReduce, we can also see the period for which the change in trend has been the maximum for both increase in enrollment and decrease in enrollment.

CourseSchedule2.java

- 5) How many students have enrolled for subjects from each department every year?
- 6) Has there been a change in the number of students enrolled for each department per year? How has the trend been?
- 7) Which department has had the highest growth in enrollments? In which period did this change take place?
- 8) Which department has most suffered in number of enrollments? In which period did this drop take place?

CourseSchedule2.java uses 2 mappers and 2 reducers. MapReduce 1 computes the sum of enrollments for all subjects offered by each department for each year. MapReduce 2 determines the trend i.e. year wise change in this enrollment. It is possible to interpret from the analysis performed by the second MapReduce, the time period between which there has been the maximum increase and decrease in the enrollments for each department.

CourseSchedule3.java

- 9) What is the total capacity of each hall over all time slots for each year?
- 10) Has there been any change in the capacity of any of the halls over the years? Has the capacity increased or decreased? If yes, then by how much? During what period did the change happen?
- 11) When has there been maximum increase in the capacity of hall? How much was the increase?
- 12) When has there been maximum decrease in the capacity of hall? How much was the decrease?

CourseSchedule3.java uses 2 mappers and 2 reducers. It deals with the capacity of all the lecture halls in the university and how it has changed over the years. In the first phase of MapReduce, the total capacity each lecture hall offers each year has been deduced from the data. This output when subjected to the second phase of MapReduce gives the change in the total capacities of the lecture halls in consecutive years. The highest change both positive and negative can be interpreted from the results along with the exact period in which it occurred.

CourseSchedule4.java

- 13) How many students are enrolled to attend classes in any of the halls for a given time slot?
- 14) How many of the halls are empty for a particular time slot, and how many times?

CourseSchedule4.java uses 2 mappers and 2 reducers. The first MapReduce job calculates the total number of students of all departments together attending class in each of the lecture halls at a particular time of the day over all the years. From the result it can be observed that there are no students in some hall at some time of the day. The second MapReduce job calculates the number of such cases, that is, the number of lecture halls which have always been empty at a certain time of the day.

CourseSchedule5.java

- 15)What is the enrollment for all subjects, the subsequent seating capacity of each of them in all the Halls for every term across all years?
- 16)What is the utilization percentage of each hall over all the years?
- 17)Each year, what percent of the total availability of halls has actually been utilized?
- 18)What is the percent utilization of each hall in each year?
- 19)What is the increase or decrease in percent utilization of each hall in consecutive years?

CourseSchedule5.java has 5 mappers and 5 reducers. Except for the first pair, the rest all are independent of each other and are only used to answer individual questions. They all however, make use of the first pair of Mapper and Reducer for their computations.

- *The first MapReduce task computes total enrollment and total capacity of each lecture hall in each term.*
- *The second MapReduce task calculates the year wise total utilization of all lecture halls, where utilization is determined by the formula 'total enrollment / total capacity'.*
- *The third MapReduce task calculates the hall wise total utilization over all the years.*
- *The fourth MapReduce task combines the functionality of the above two and gives a more detailed result. It gives the year wise total utilization for each lecture hall.*
- *The fifth MapReduce task gives the trend in utilization in terms of increase or decrease and the corresponding difference, among subsequent years for all the lecture halls.*

CourseSchedule6.java

- 20)What is the number of times each hall has been under-utilized over all the years?

CourseSchedule6.java uses a single MapReduce job. It considers only those records for which the number of enrollments is significantly lesser than the capacity of the course. In such scenarios, the space is being inefficiently allocated leading to under-utilization. This MapReduce job computes the total number of times such a situation has been recorded for each hall over the years.

CourseSchedule7.java

21)What is the number of times each hall has been over-utilized through the years?

CourseSchedule7.java uses a single MapReduce job. It considers only those records for which the number of enrollments is greater than the capacity of the course. Since in such situations, a lecture hall accommodates more students than the capacity it offers, it is a situation of over-utilization. This MapReduce job computes for each hall, the total number of times such a situation has been recorded.

Conclusion:

Thus, we have implemented 7 different MapReduce programs, with multiple mappers and reducers, to answer several questions which allow us to analyze the data about the class schedules over the years at UB. The outputs of these MapReduce programs enable us to gain insights regarding the utilization of classrooms and the enrollment trends.