Interactive Graphics

Sateesh Chandra Babu, Akula

Course: MSc Data Science

Overview

- Understanding & Analyzing the given real-time dataset.
- Individual tasks on main project theme to draw final outcome.
- Visualizing the data using d3.js.
- Delivering conclusion on visualized graphs.

Project theme

- **Hypothesis**: Does attendance really effects the course outcome?
  - ✓ Collecting all the evidence's required to check whether the proposed theme makes sense.

- Individual Tasks
  - ✓ **Task-1**: Data analysis on Age different groups
  - ✓ **Task-2**: Stage of Study
  - ✓ **Task-3**: Outcome of the semester Overall
  - ✓ **Task-4**: Average Attendance of overall semester

Data Cleansing & Clustering

How it works:

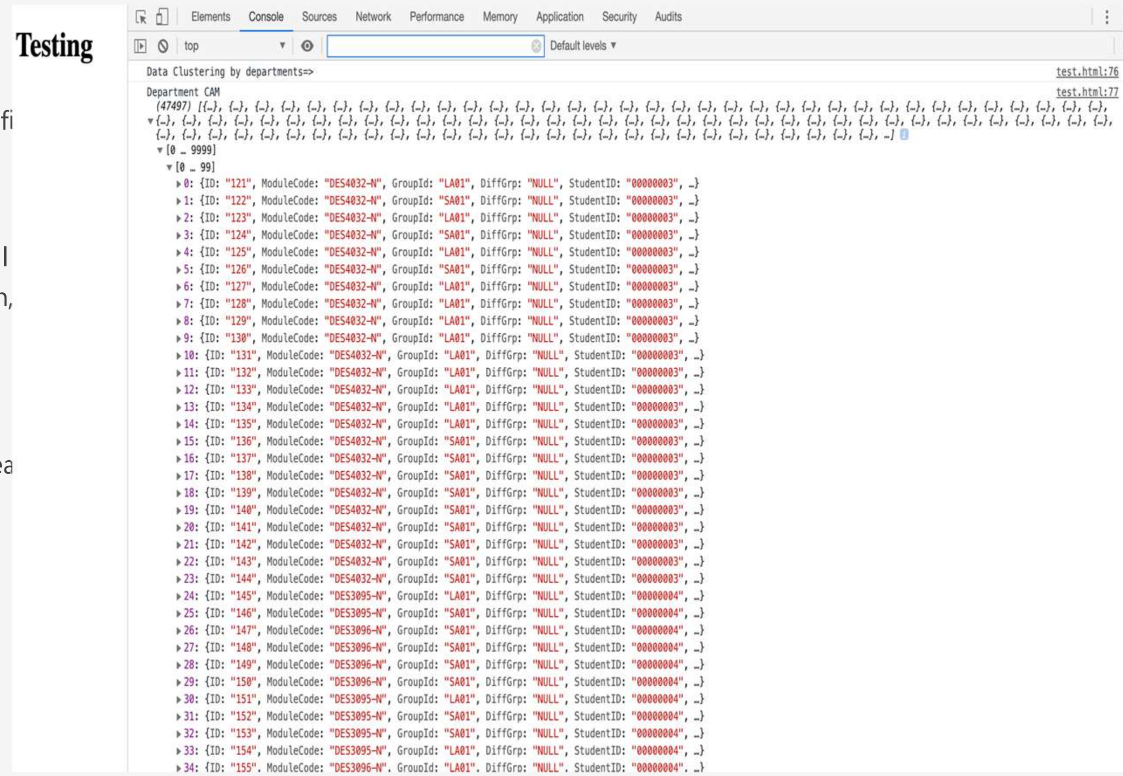
1. Identifying all the known departments from the given CSV fi

2. Clustering the data for each department. Then this data will clustered again according to Stage of study like Foundation, Undergraduate year.

3. With the above step, we will be able to access the data of ea module.

4. Resulting data: For each modulecode

ModuleCode: “MED0008-N”
StudentID: [“1”,”2”,”3”…..]
Status: [“P”, “ABS”….]
WeekNum: [“1”, “2”…..]

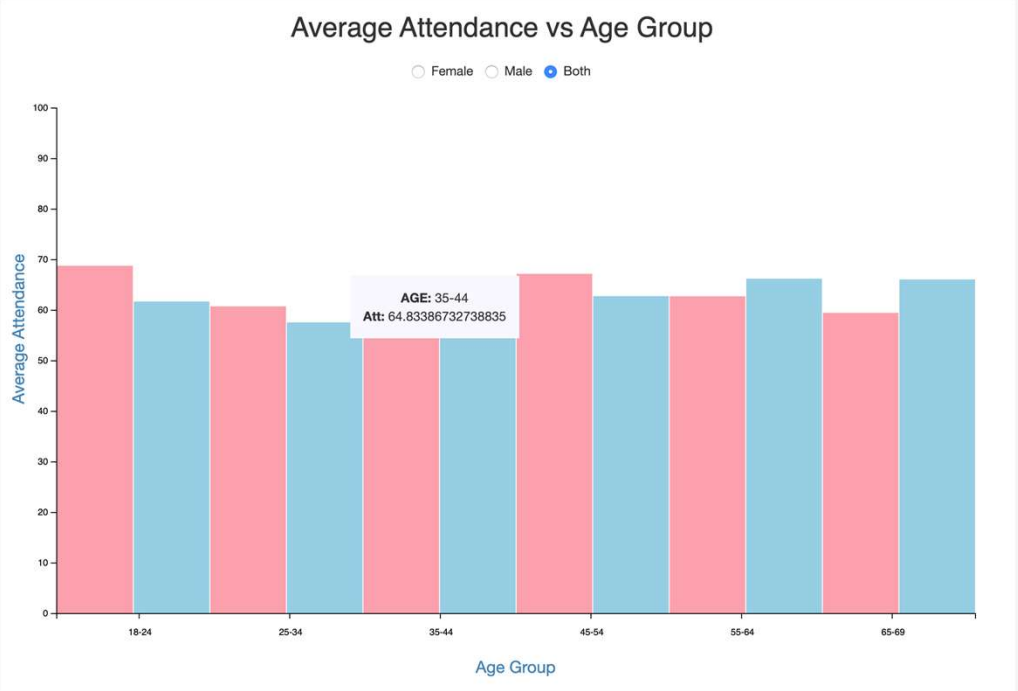


Finding

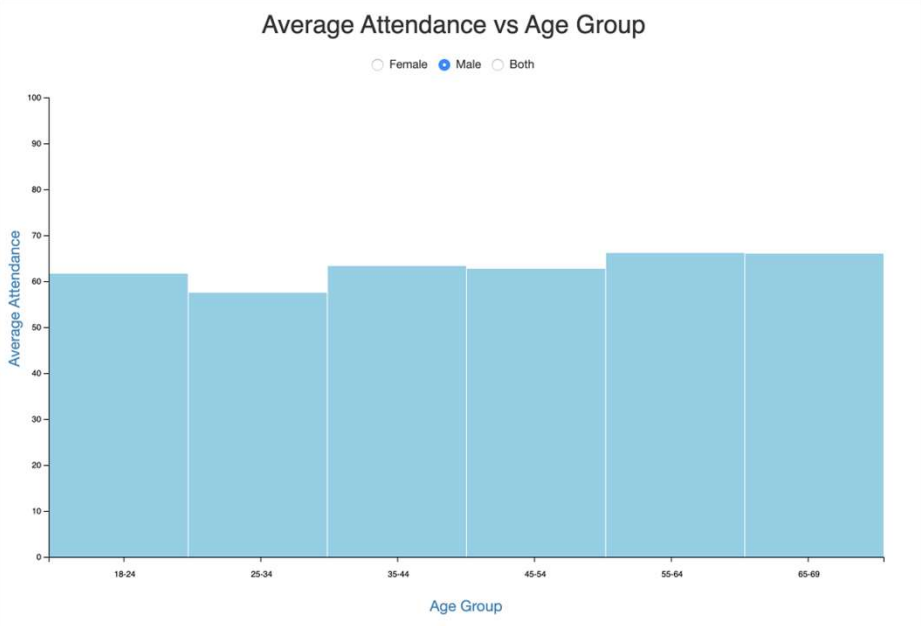
- From the given data, Total unique number of modules are 200.
- But when I segregated the data based on department ID's (GAV, CSIS, CAM),
- I have found that there has been a redundancy in no.of modules, which is 238.
- Due to similar modules in foundation year students & some other could be the elective or base foundation for there respective courses.

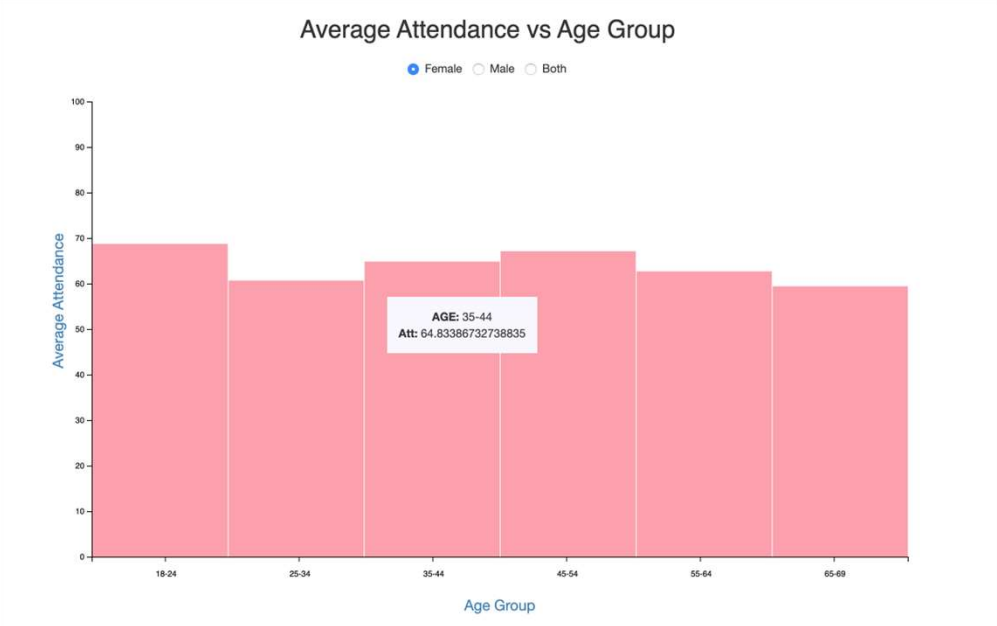
Macroview – Age Group & Average Attendance

- This graph represents an overview of different age groups and their Average attendances.
- **I chose to draw a relation between these two because there is a huge diversity in the age group of students, throughout the University**
- Identified age groups range from 18 to 69.
- To visualize this information, I have choose to create a custom age groups [18-24, 25-34, 35-44, 45-54, 55-64, 65-69]
- And I have added a gender filter, to check the stats based on gender.
- On X-axis: Age groups is defined.
- On y-axis: Average attendance percentage is defined.
- Each bar represents the average attendance of the particular age group.



Macroview – Age Group & Average Attendance

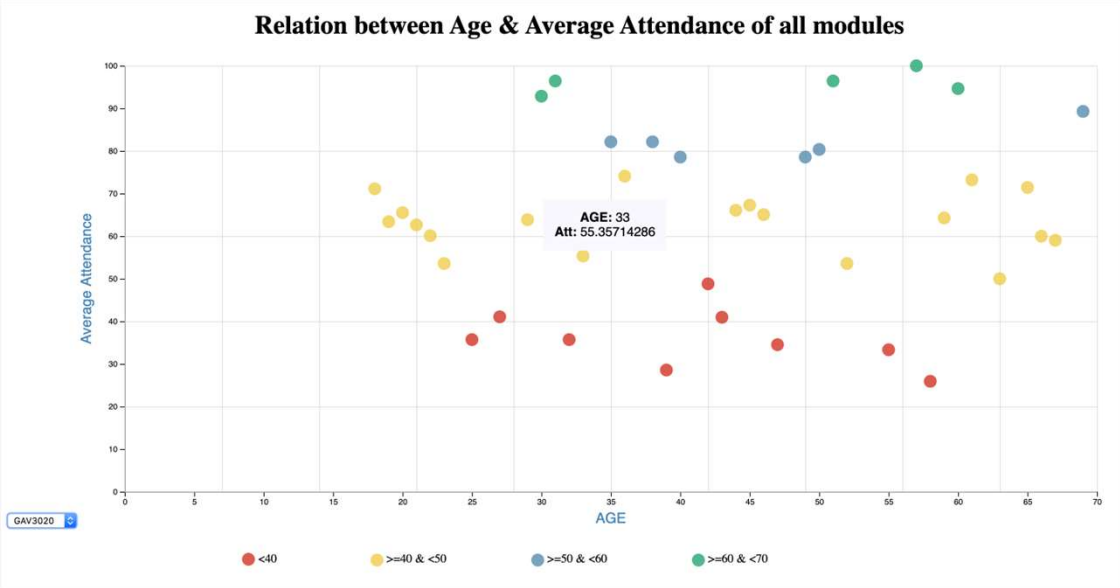




- Furthermore, I have calculated & plotted average attendance for male & female in all age groups separately.
- I have given user flexibility to choose different options (Female, Male, Both), using radio buttons.
- I have used tooltip to display each bar's (x, y) values, when mouseover events are called.

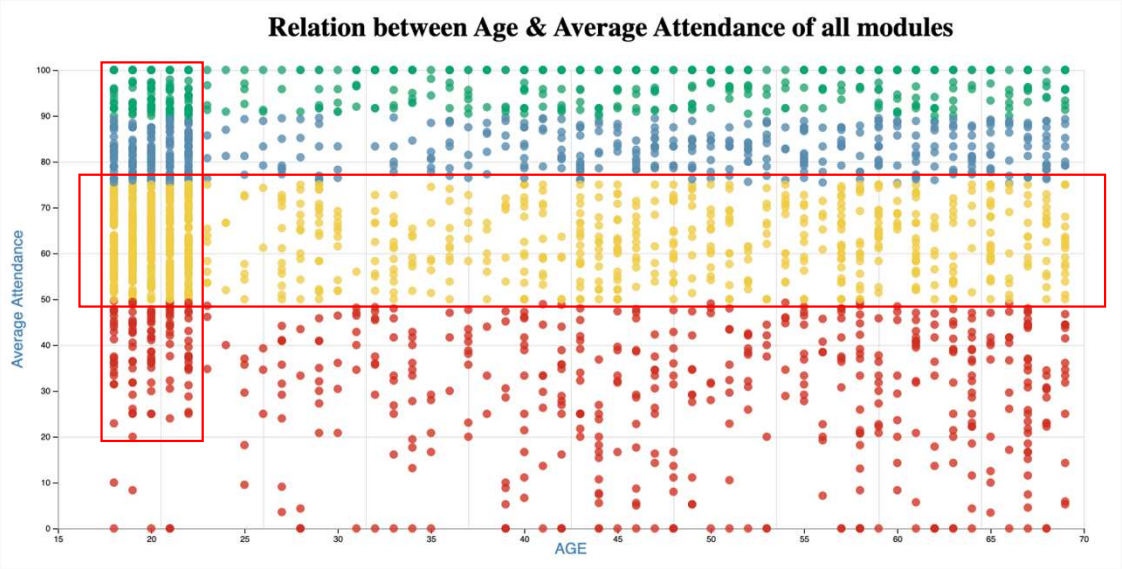
Microview – Age vs Average Attendance in each module

- **This graph draws a relation between student's in a particular age and their average attendance percentage in every module from the given dataset.**
- I have chosen scatterplot because of it's effectiveness in depicting large datasets.
- On X-axis: Individual age range
- On Y-axis: Average attendance percentage
- Color filter's based on attendance percentages.
- For example : Red represents attendance less than 50%.
- Drop Down menu for modules filtering.
- By examining the graph, There isn't any pattern to conclude the trend.
- And in fact, there isn't any correlation to suggest that age has some effect on attendance.



Microview – Age vs Average Attendance

- **Upon further analysis, I choose to eliminate the module filter to check the following trend.**
- Due to different course types & modules, I haven't found any correlation between Age & Attendance.
- But the graph suggests the increase in number of dots in a range of 74% to 50% attendance.
- From the Age 18 to 24, a high number of continuous dots been observed. [Which is in between 30% to 100% attendance]
- This shows that, Students from 18 to 24 Age groups are more active than other age groups.
- These may suggest that, this particular age group is showing more participation in different courses & degrees.
- Though I may not conclude the correlation between age & attendance. But I have noticed that 18to24 age groups has showed more interest in pursuing higher education.



Students with Average attendance (>50 & <75) are higher

Conclusion

- From the above visualization, We can see that age has a little effect on total attendance. But It varied from course to course.
- Due to certain circumstances, We are not able to visualize the average attendance for modules.
- This limited us to conclude our project theme with enough evidence.

Thank You!