UN GOAL # 4: QUALITY EDUCATION

Aditya Melkote

Introduction

Goal

- Advance UN's 4th Goal "Ensure inclusive and equitable education and promote lifelong learning opportunities for all" in the context of 'Data Science for High School' course
- Identify course shortcomings
- Make recommendations to remedify the issues

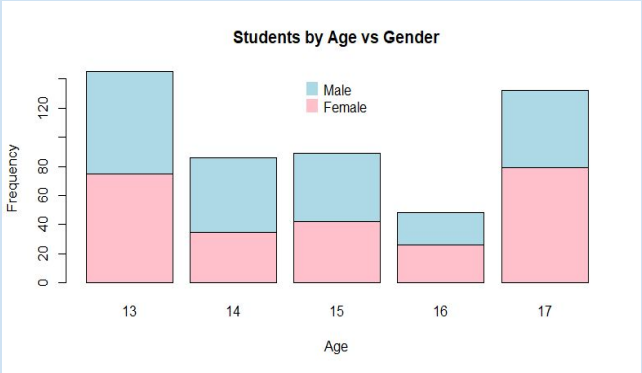
Problem Analysis and Solutioning Approach

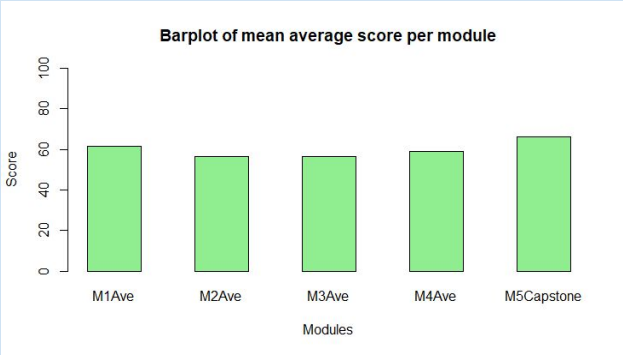
- Process the input data set to identify relevance
- Visualize the data set to make observations
  - Methods used:
    - Inherent R capabilities
    - Imported Libraries
    - Sentiment Analysis
- Perform problem root cause analysis
- Reason and make recommendations / propose solutions

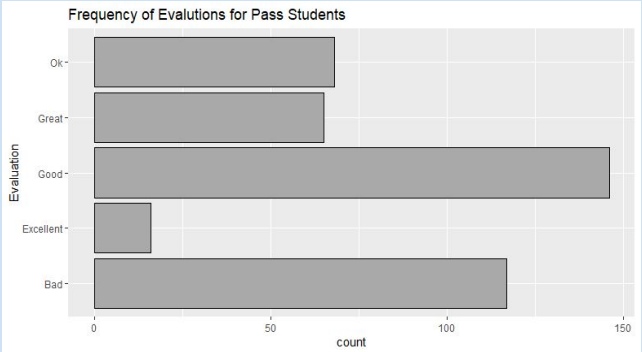
Data Analysis

Overall Statistics

- Equal number of males and females registered
- Distributed equally among all ages
- 13 and 17 year olds make up large majority
- Few 16 year olds
- General sentiment is mediocre
- Median Score is 57%
- Students struggle with all modules
  - Specifically module 2

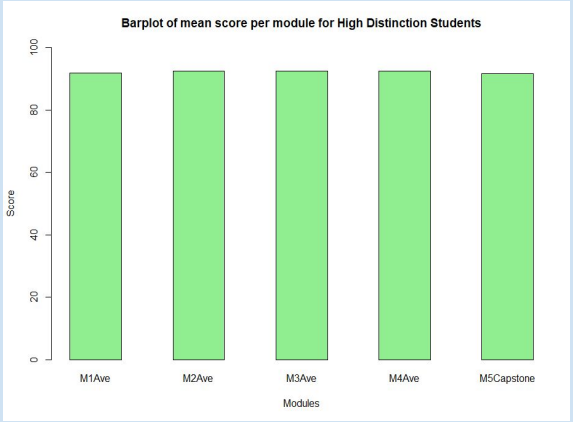


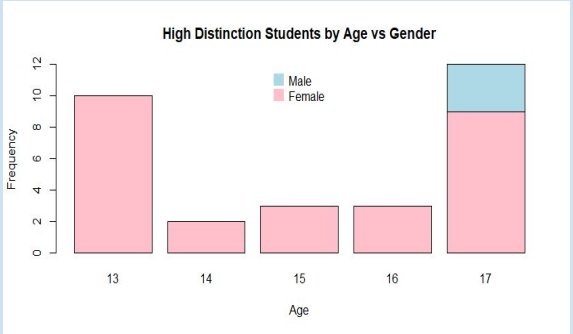




High Distinction Students

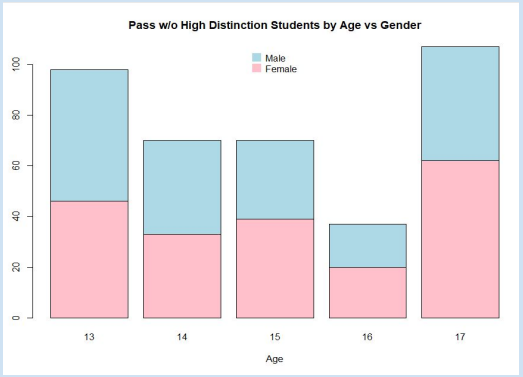
- Only 30 students pass with High Distinction
  - 6% of total students
- Female to Male ratio is 27:3
- Overwhelmingly composed of 13 and 17 year olds
  - Proportionate to number registered
- Excel in all modules (~90%)
  - Average of 92.3%
- Course rating is very positive
  - Only Excellent and Great

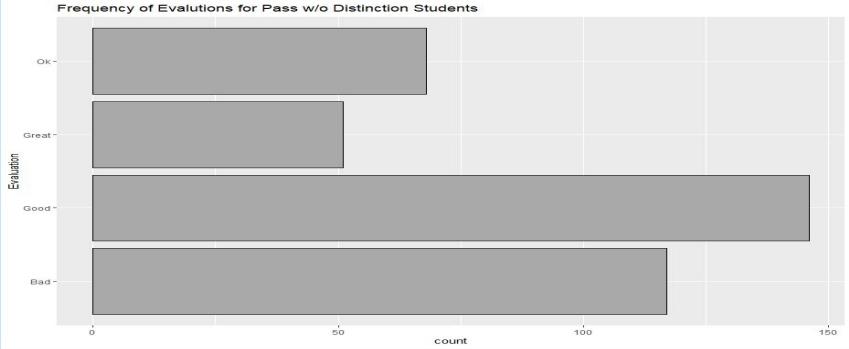


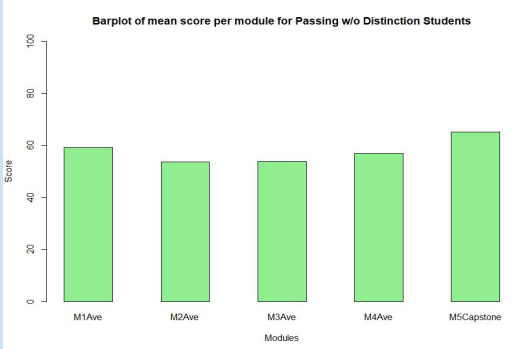


Passing Students w/o Distinction

- Perform significantly worse in module 2
  - Suggest module 2 is too difficult
- Equal distribution of males and females
- Pass with Barely Passing Grades
  - 76.25% pass
  - Average of 57.6%
  - Suggests course is too difficult
- Mediocre ratings
  - Generally dissatisfied

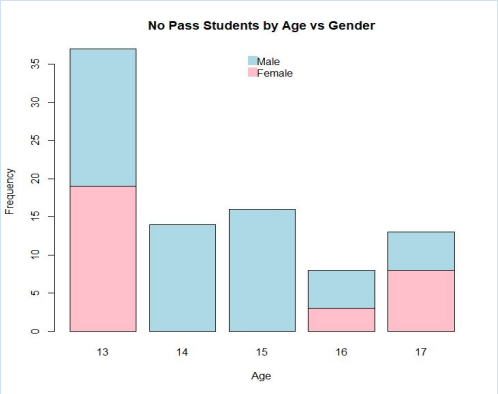


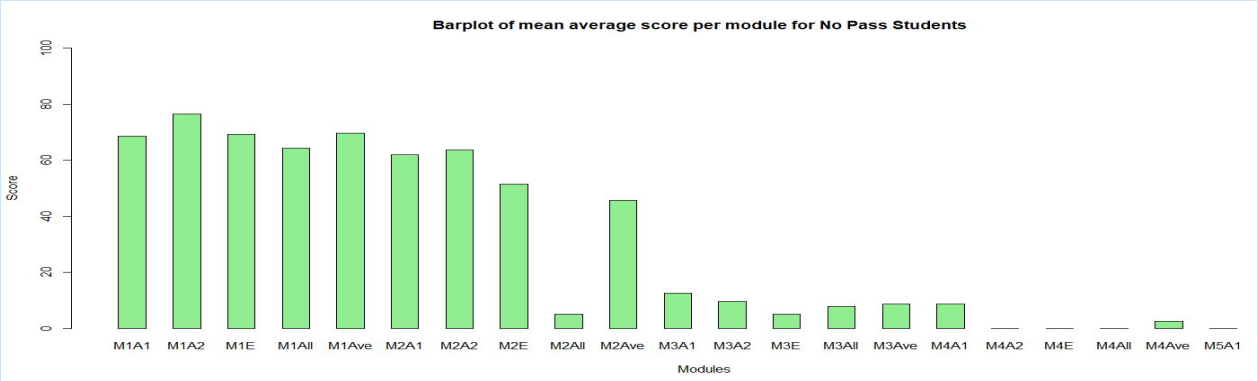




Failing Students

- 17.63% of registered students fail
- Males fail twice as compared to females
- Failure seen largely with 13 year olds
- High scores in earlier modules
  - Drop after Module 2
  - Problem seen with Module 3





Sentiment Analysis

- Running sentiment analysis on various Data Science related terms:
  - Data Management: 0.53
  - Data Science Course: 1.22
  - R Data Science: 1.07
  - Learn Data Science: 0.82
  - Intro to Statistics: 0.23
  - Data Representation: 0.45
  - Data Visualization: 0.95
  - Machine Learning: 0.58
  - Big Data: 0.1
  - Sentiment Analysis: 0.35
- Generally positive sentiment on Data Science related topics and terms

Problem Synthesis

Problems

- Topics introduced are at surface level and require consulting outside resources
- Code is hard to replicate
  - Explanations are provided for what the code does, but not why the code works
    - Code in later units get extremely complicated very fast
  - Very little information on how R works and coding practices past very basics
    - Reference cards lack in-depth explanations
- Unit assignment directions are vague and require assumptions to be made
  - Difficult to clarify with teacher all the time due to time zone differences
- Modules contain dead hyperlinks
- Example code is out of date at times
- Assignments lack a clear connection to rest of course at times Ex: Module 3
  - Results in lack of interest
  - Potential reason for why drop rate is so high
- Course material in its current form is too difficult for students
  - Interested students potentially drop the course due to difficulty in understanding course material

Disparity between genders

- No clear reason as to why males are more unsuccessful
  - No data suggests males are less capable
  - Not enough information about students' background to make any informed conclusions
  - Too many underlying factors
  - Possibility of being coincidental

Proposed Solutions

Recommendations

- Introduce Video Lectures and/or notes
  - Demonstrate exactly what each function does
- Introduce coding exercises
  - Build code from scratch
- Clean up dead hyperlinks
  - Host on VHS servers
- Clear explanations on how unit assignments relate to the bigger picture
  - Unit 3 Final Assignment
- Optional sections introduced to give more in-depth explanations of the concepts learned
  - Ex. Statistical Concepts
  - Extra support for students who need them
    - Specifically for 13 year olds
    - Specifically Unit 2

Conclusion and Reflection

Summary

Problems:

- Course is too difficult for majority of students in its current state
- Students are too unsuccessful
  - Students barely pass
  - Nearly a 5th of registered students fail
- Students lose interest
- Certain demographics perform worse
  - 13 year olds
  - Males
- Student's are not gaining a positive experience
  - Majority of students reporting mediocre or negative responses
    - Decreases likelihood of pursuing career in this field

Proposed Solutions:

- Increase course clarity
  - Introduce interactive resources and exercises
  - Clean up and modernize course links and code
  - Explain example code in greater detail
  - Explain connections between certain assignments and overall theme better
- Provide optional extra support for students who require more help

Conclusion

- Proposed solutions will advance UN's 4th goal
  - Solutions will increase student retention rates
  - Solutions will improve the quality of education delivered to students
  - Solutions will promote equity in education
- Students will have a more positive experience
  - Increase the likelihood of students pursuing higher education in Data Science
  - Students will acquire better grasp of foundational concepts to succeed in higher education

Reflection

- Coding was difficult
  - Was very tough to figure out what code accomplished the goal I had in mind
  - Optimizing code
- Hard to provide solutions with very little background information of the students
  - Only given age and gender
    - No link between gender and success
- Hard to perform sentiment analysis
  - Course specific keywords had 0 results
  - Had to broaden search terms which gave results that were less relevant to course