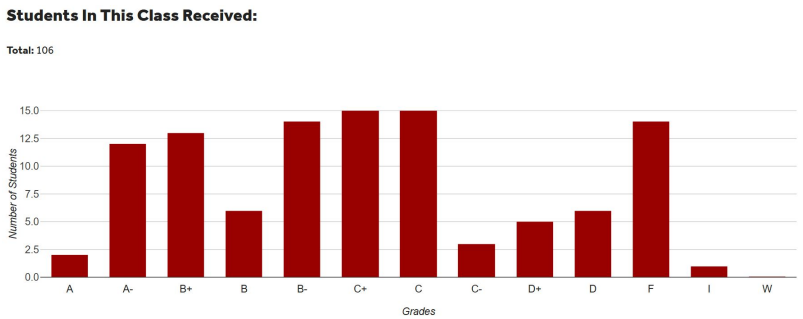
From Rate to Reality: a RateMyProfessor deepdive

Karen Cho, Christopher Lobato

Why look at class ratings?

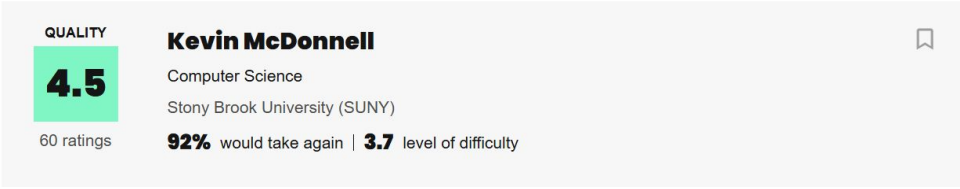
- CSE majors are now required to take 6 technical electives
- Advice from peers and advisors can often be anecdotal
- Data is scattered across multiple platforms RateMyProfessor, ClassieEvals, Reddit etc.

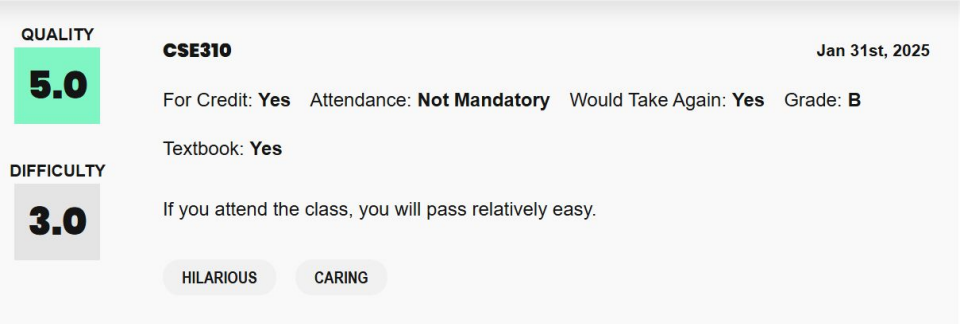


- CSE 355 - Computational Geometry
- CSE 356 - Cloud Computing
- CSE 357 - Statistical Methods for Data Science
- CSE 360 - Software Security
- CSE 361 - Web Security
- CSE 362 - Mobile Security
- CSE 363 - Offensive Security
- CSE 364 - Advanced Multimedia Techniques
- CSE 366 - Introduction to Virtual Reality
- CSE 370 - Wireless and Mobile Networking
- CSE 371 - Logic
- CSE 373 - Analysis of Algorithms
- CSE 376 - Advanced Systems Programming in UNIX/C
- CSE 377 - Introduction to Medical Imaging
- CSE 378 - Introduction to Robotics
- CSE 380 - 2D Game Programming
- CSE 381 - 3D Game Programming
- CSE 385 - Analysis of Algorithms: Honors
- CSE 390 - Special Topics in Computer Science

What we're looking for

- Simulate web browsing using selenium to retrieve information from each professor card
  - Name
  - Quality
  - Difficulty
- Aggregate data from the Psychology, Computer Science, Electrical/Computer Engineering department into a pandas dataframe





Web Scraping Data

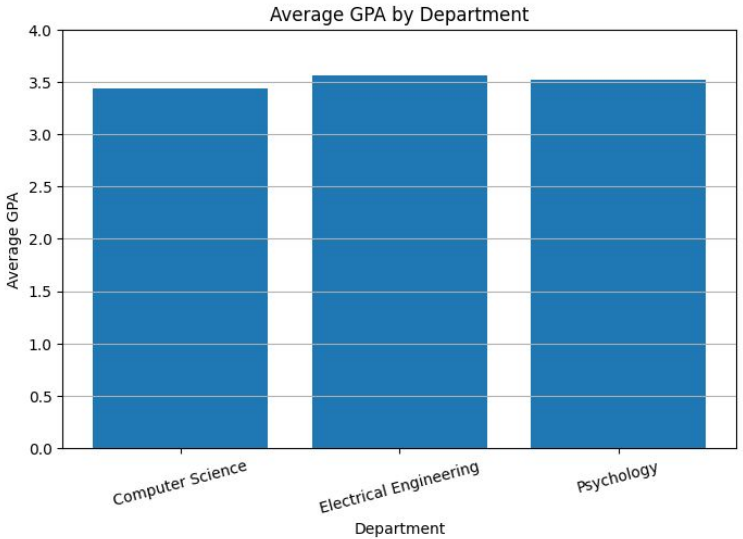

```
"UmF0aW5nLTQwNjI3ODMw": {
    "__id": "UmF0aW5nLTQwNjI3ODMw",
    "__typename": "Rating",
    "comment": "If you attend the class, you will pass relatively easy. ",
    "flagStatus": "UNFLAGGED",
    "createdByUser": false,
    "teacherNote": null,
    "legacyId": 40627830,
    "date": "2025-02-01 01:15:10 +0000 UTC",
    "class": "cse310",
    "helpfulRating": 5,
    "clarityRating": 5,
    "isForOnlineClass": false,
    "difficultyRating": 3,
    "attendanceMandatory": "non mandatory",
    "wouldTakeAgain": 1,
    "grade": "B",
    "textbookUse": 3,
    "isForCredit": true,
    "ratingTags": "Hilarious--Caring",
    "id": "UmF0aW5nLTQwNjI3ODMw",
    "adminReviewedAt": "2025-02-01 01:15:29 +0000 UTC",
    "thumbsUpTotal": 0,
    "thumbsDownTotal": 0,
    "thumbs": {
      "__refs": []
    }
  }
```


Final Rating Object

```
{
Professor: String,
class: String,
difficulty rating: int,
quality: int,
would take again: bool,
grade: String
attendance mandatory: String
}
```

Average Comparison by Department

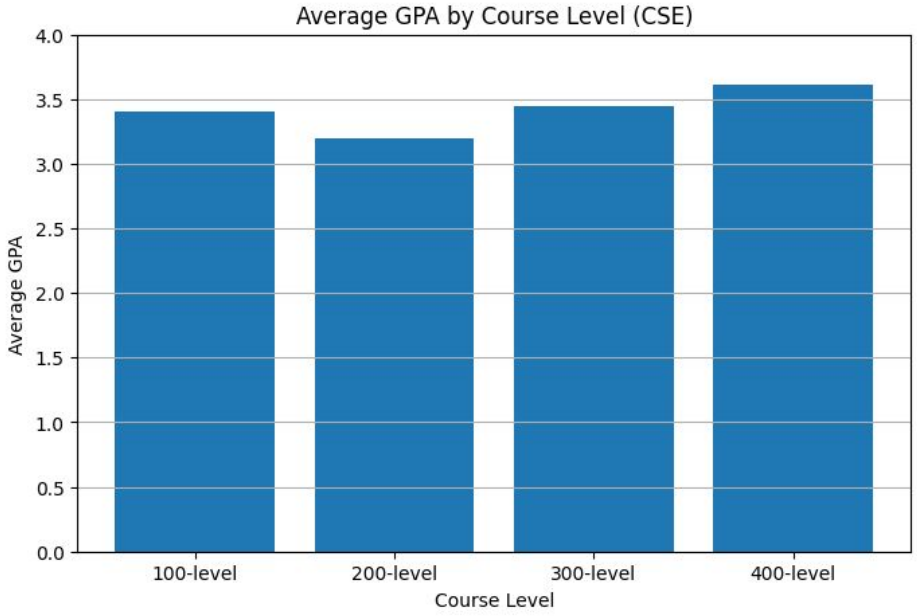
Question: Is certain major easier to earn higher grades?



Department	Average grade
CSE	3.4417
ECE	3.5597
PSY	3.5189

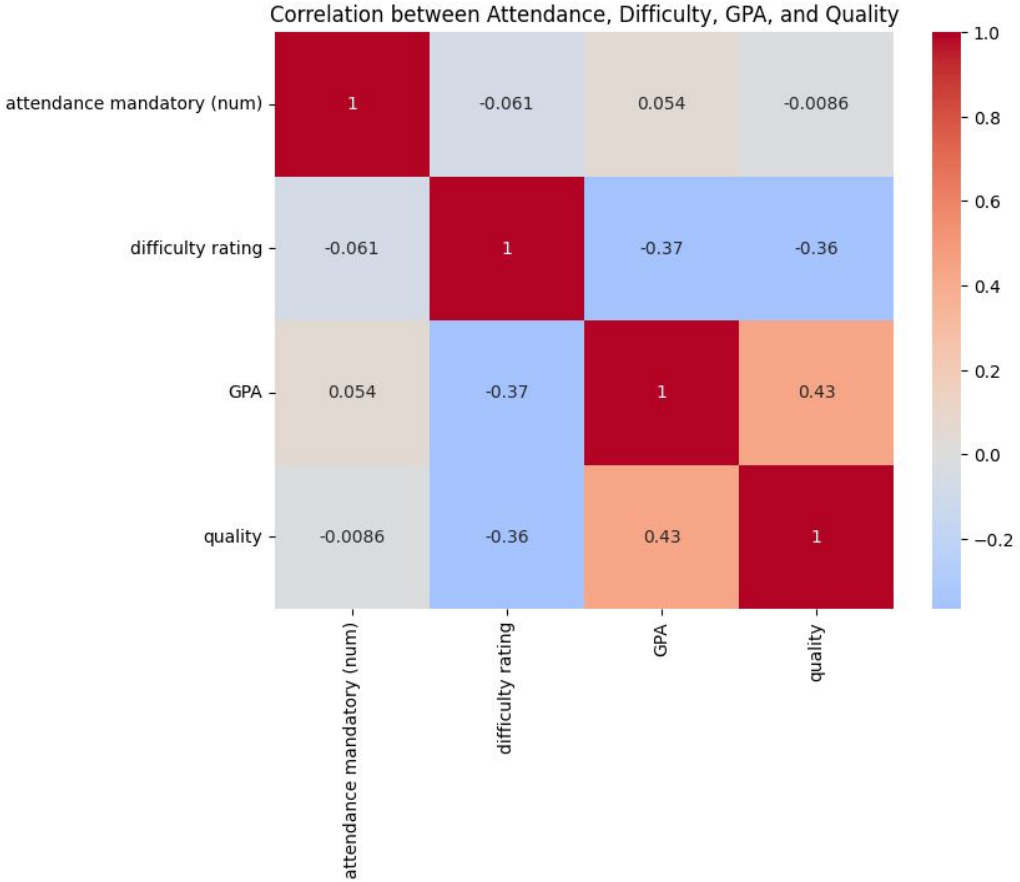
Average grades by course level

Question: As the course level goes up, is it hard to achieve higher grades?



Level	Average grade
100	3.4054
200	3.2019
300	3.4458
400	3.6178

Attendance Mandatory Relations



1.Is attendance mandatory because the course is difficult?

:No, attendance policies seems independent of difficulty

2.Does mandatory attendance help students achieve higher grades?

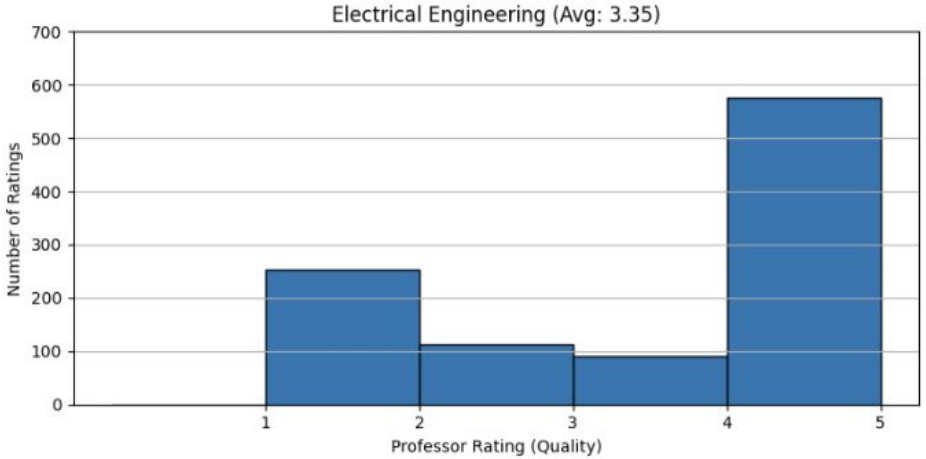
:No, mandatory attendance does not significantly boost student grades.

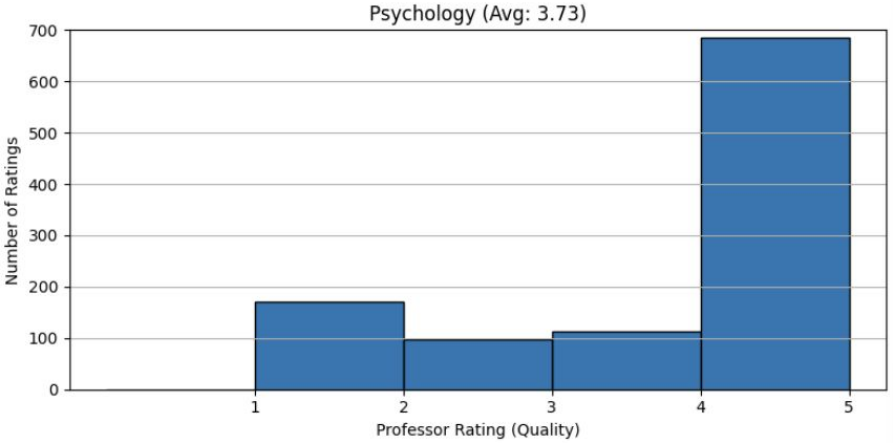
3.Do students think the quality is higher/lower if attendance is mandatory?

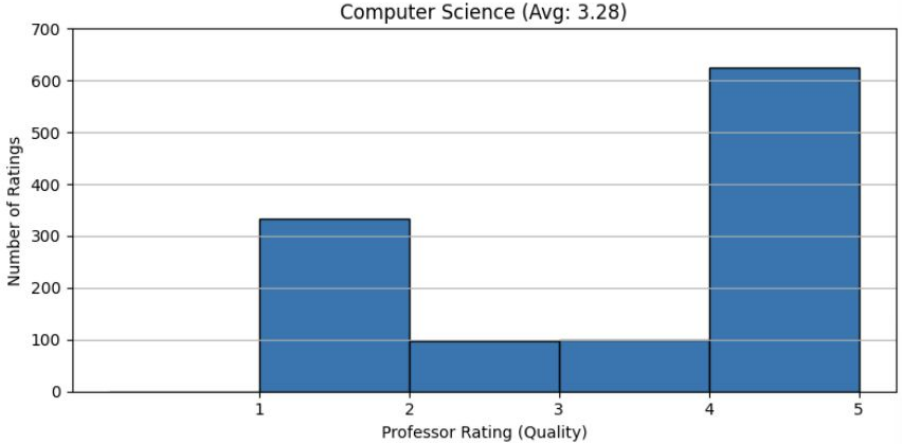
:No, students don't think class quality is better (or worse) because of mandatory attendance.

Summary: Mandatory attendance is basically random

Professor Ratings Comparisons

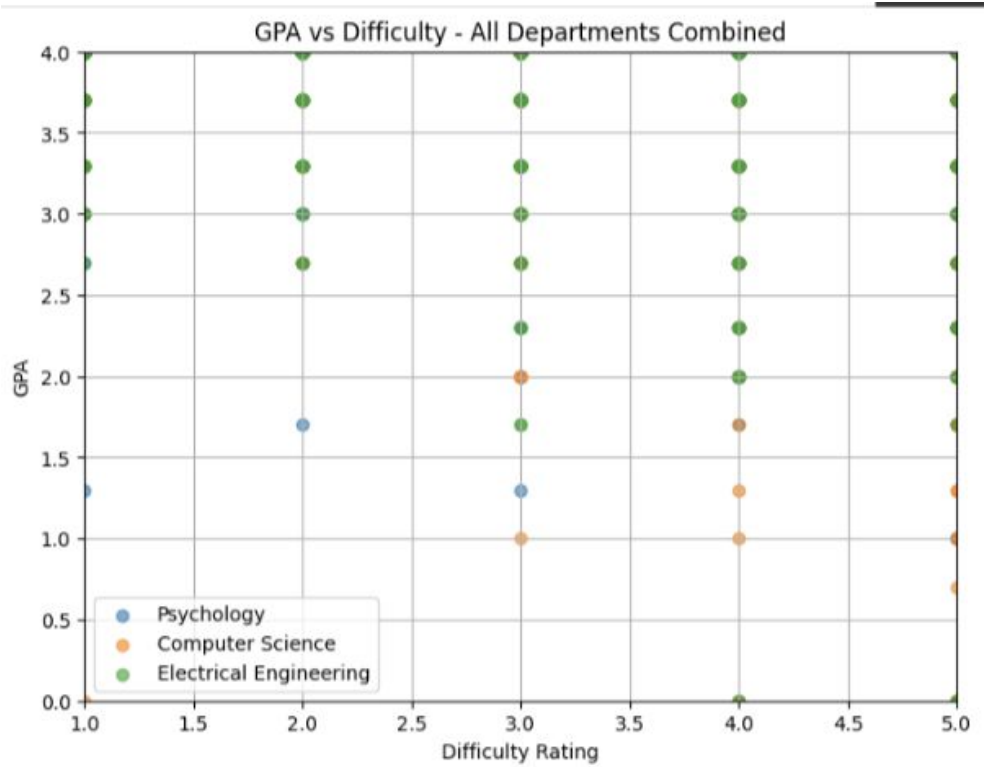






Easiness & Grade Relations

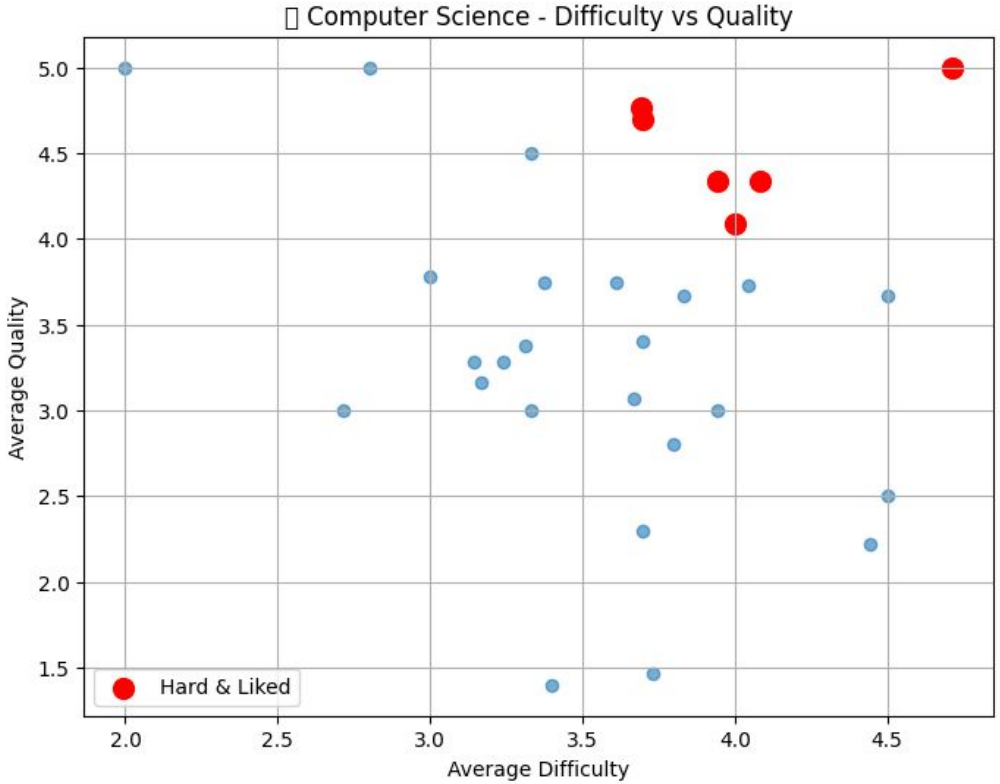
Question: Do easy going classes really end up with high grade?



Yes, in Psychology and Computer Science, easier classes often lead to higher grades.

However, no clear pattern in Electrical Engineering – grade is more spread out regardless of difficulty.

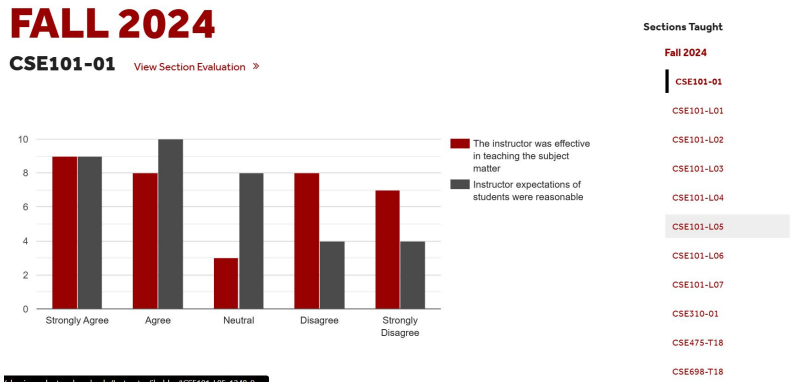
Hard Course That is worth taking?

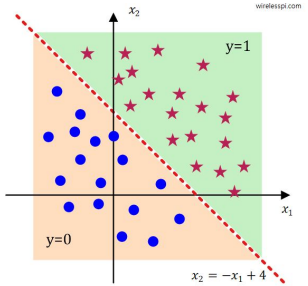


Class	Difficulty	Quality
CSE320	4.0833	4.3333
CSE331	3.7000	4.7000
CSE356	4.7142	5.0000
CSE373	4.0000	4.0909
CSE506	3.6923	4.7692
CSE548	3.9444	4.3333

Where to go from here?

- Scrape and aggregate data from classie evals to find similarities or discrepancies between quality ratings
- Impute data for missing optional fields on RateMyProfessor
- Use student input data with a logistic regression model to classify if they would like a class or not





Logistic Regression
Binary classifier

Resource and Google Colabs

Data Analysis Colab:

https://colab.research.google.com/drive/1pMhZbP_V2k7COGfLxRUEfZjKBviP8oxq?authuser=0#scrollTo=0Gkg8D4HHdb3

Scraping Colab:
https://colab.research.google.com/drive/1m0FziomUFKMj_hT0vHceA2DGkpMg9L3Y?usp=sharing

Inspiration for web scraping: SBUScheduler by MadR7 and the Gordon