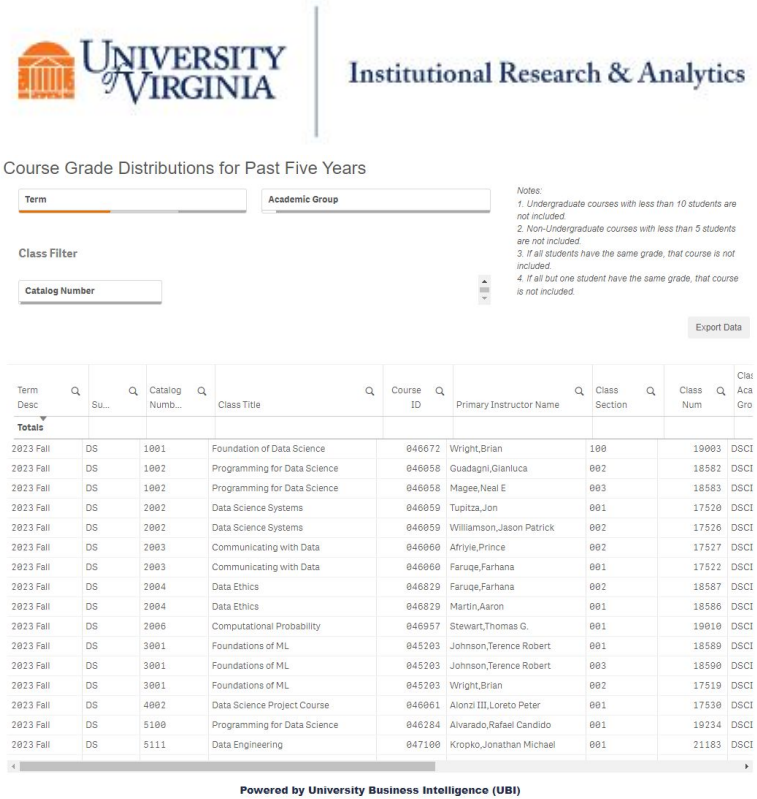
# Analysis of Undergraduate DS at UVA

Group 10 | Andrew Shifflett, Gabe Silverstein, Hanna Svenson, Masato Takedai

# The Dataset

- UVA IRA
- Filtering -> minimal cleaning
  - Term Desc
  - Subject
- Relevant variables:
  - Catalog Number
  - Class Title
  - Primary Instructor Name
  - Course GPA (A+, A)
  - # of Students



UNIVERSITY of VIRGINIA | Institutional Research & Analytics

Course Grade Distributions for Past Five Years

Term | Academic Group

Class Filter

Catalog Number

Notes:
1. Undergraduate courses with less than 10 students are not included.
2. Non-Undergraduate courses with less than 5 students are not included.
3. If all students have the same grade, that course is not included.
4. If all but one student have the same grade, that course is not included.

Export Data

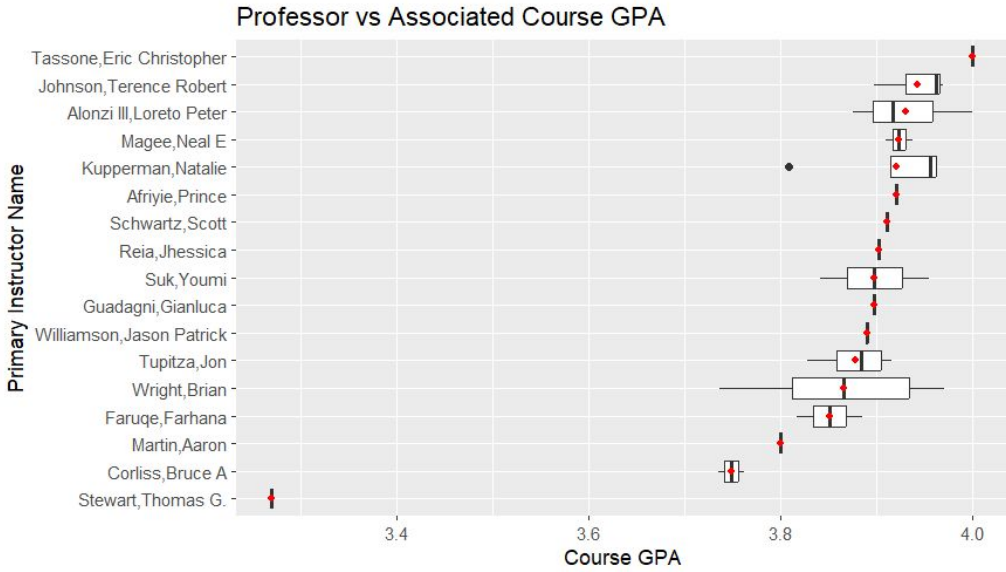
| Term Desc | Su... | Catalog Numb... | Class Title | Course ID | Primary Instructor Name | Class Section | Class Num | Cla Aca Gro |
|---|---|---|---|---|---|---|---|---|
| Totals | | | | | | | | |
| 2023 Fall | DS | 1001 | Foundation of Data Science | 046672 | Wright,Brian | 100 | 19003 | DSCI |
| 2023 Fall | DS | 1002 | Programming for Data Science | 046058 | Guadagni,Gianluca | 002 | 18582 | DSCI |
| 2023 Fall | DS | 1002 | Programming for Data Science | 046058 | Magee,Neal E | 003 | 18583 | DSCI |
| 2023 Fall | DS | 2002 | Data Science Systems | 046059 | Tupitza,Jon | 001 | 17520 | DSCI |
| 2023 Fall | DS | 2002 | Data Science Systems | 046059 | Williamson,Jason Patrick | 002 | 17526 | DSCI |
| 2023 Fall | DS | 2003 | Communicating with Data | 046060 | Afriyie,Prince | 002 | 17527 | DSCI |
| 2023 Fall | DS | 2003 | Communicating with Data | 046060 | Faruqe,Farhana | 001 | 17522 | DSCI |
| 2023 Fall | DS | 2004 | Data Ethics | 046829 | Faruqe,Farhana | 002 | 18587 | DSCI |
| 2023 Fall | DS | 2004 | Data Ethics | 046829 | Martin,Aaron | 001 | 18586 | DSCI |
| 2023 Fall | DS | 2006 | Computational Probability | 046957 | Stewart,Thomas G. | 001 | 19010 | DSCI |
| 2023 Fall | DS | 3001 | Foundations of ML | 045203 | Johnson,Terence Robert | 001 | 18589 | DSCI |
| 2023 Fall | DS | 3001 | Foundations of ML | 045203 | Johnson,Terence Robert | 003 | 18590 | DSCI |
| 2023 Fall | DS | 3001 | Foundations of ML | 045203 | Wright,Brian | 002 | 17519 | DSCI |
| 2023 Fall | DS | 4002 | Data Science Project Course | 046061 | Alonzi III,Loreto Peter | 001 | 17530 | DSCI |
| 2023 Fall | DS | 5100 | Programming for Data Science | 046284 | Alvarado,Rafael Candido | 001 | 19234 | DSCI |
| 2023 Fall | DS | 5111 | Data Engineering | 047100 | Kropko,Jonathan Michael | 001 | 21183 | DSCI |

Powered by University Business Intelligence (UBI)

# Question 1

Over the past three years, which five data science professors are associated with the highest average undergraduate course gpa?

1. Eric Tassone
2. Terence Johnson
3. Loreto Alonzi III
4. Neal Magee
5. Natalie Kupperman

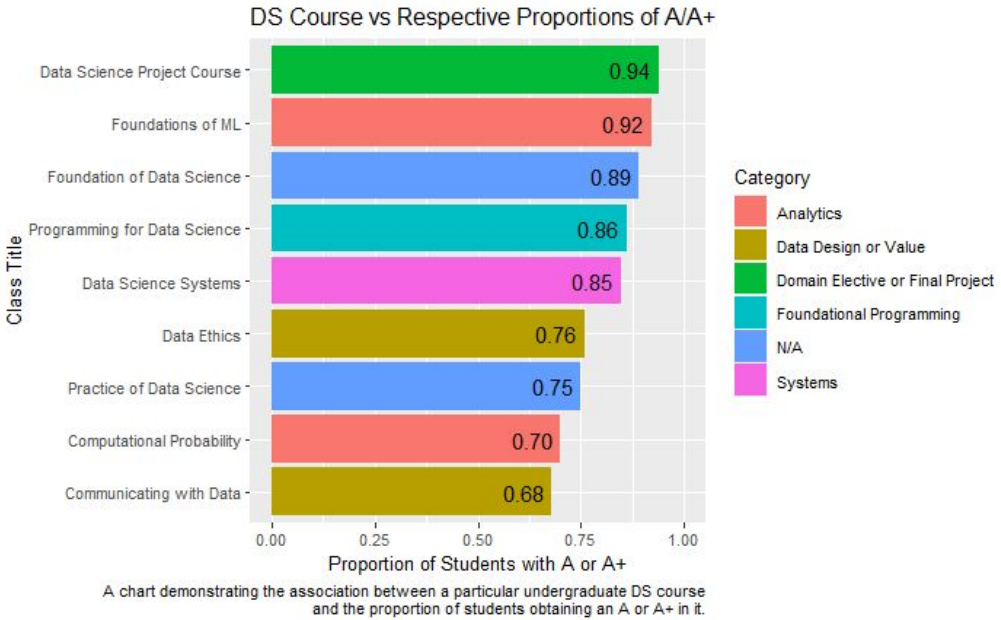



A comparison of DS Professors and their associated course gpa in the past 3 years. The red dots mark the average associated gpa for that professor.

# Question 2

Which five undergraduate data science courses have the highest proportion of students receiving an A or A+?

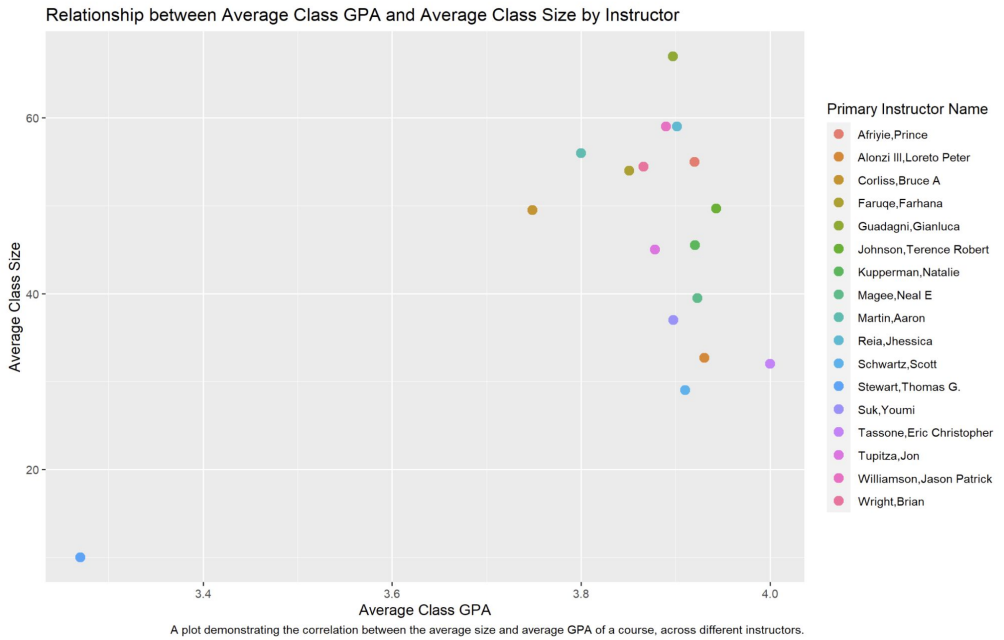
1. DS 4002: Data Science Project Course
2. DS 3001: Foundations of ML
3. DS 1001: Foundation of Data Science
4. DS 1002: Programming for Data Science
5. DS 2002: Data Science Systems



# Question 3

How does the GPA associated with a data science professor correlate to the size of the classes they have taught in the past three years?

There is little correlation between a professor's associated gpa and class size.

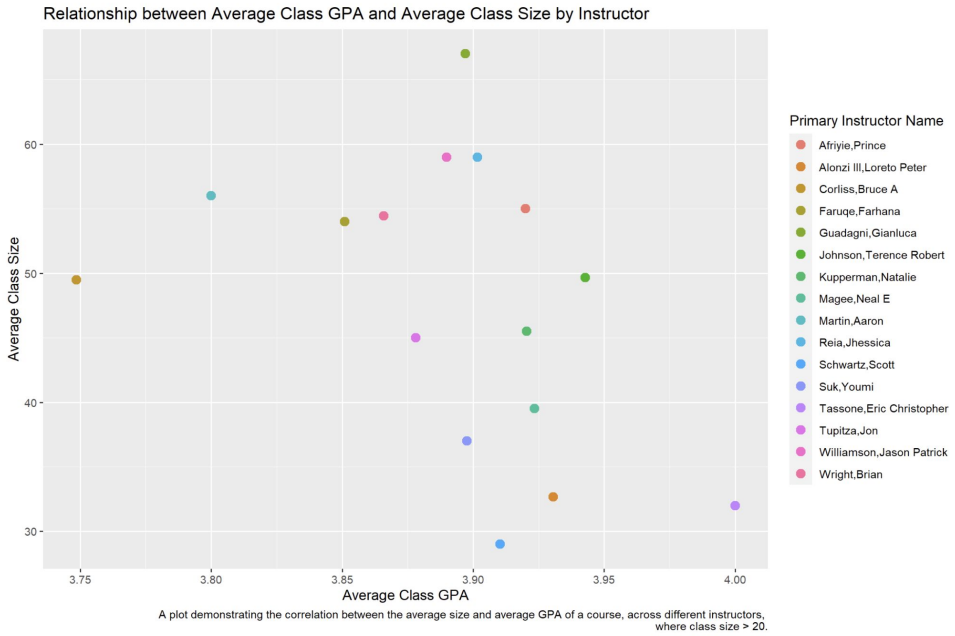


A plot demonstrating the correlation between the average size and average GPA of a course, across different instructors.

# Question 3 cont.

After adding a filter for class size > 20, we see a slight negative correlation between a professor's average GPA and class size.

Conclusion: Smaller class sizes are somewhat correlated with higher average GPAs, except in cases where class size ≤ 20.



A plot demonstrating the correlation between the average size and average GPA of a course, across different instructors, where class size > 20.

# Question 4

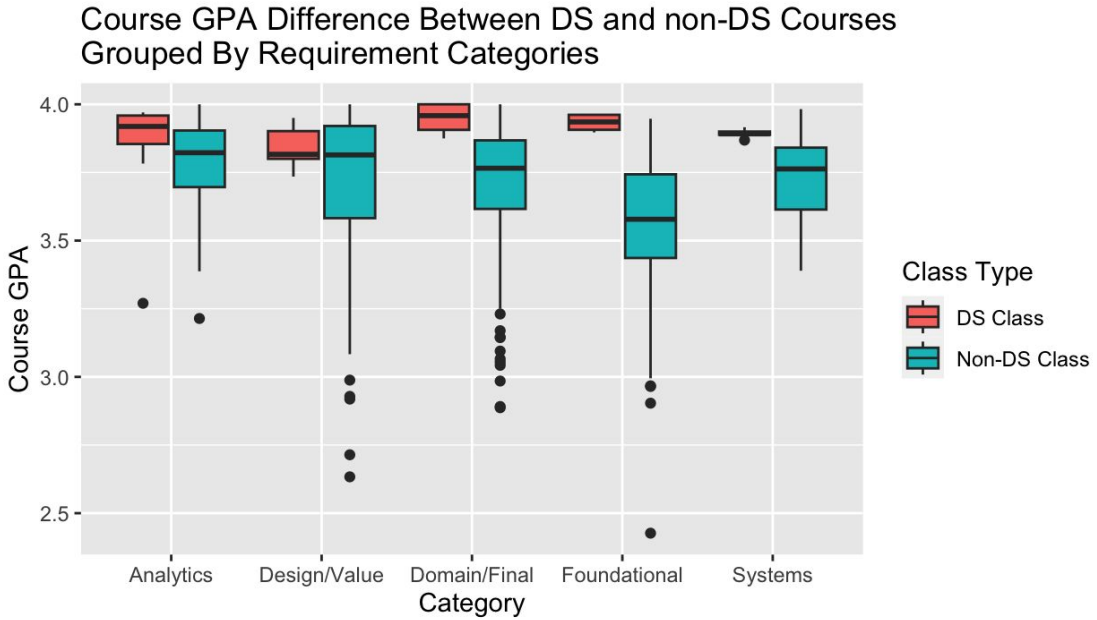
Over the past 5 years, how do the GPA of equivalent non-DS courses compare to the DS courses?

Cleaning:

- Required more preparation
- Vectors with lists of classes for each requirement were used to filter classes

Conclusion:

- DS Classes on average have a higher GPA than outside classes
- Outside classes tend to vary much more in GPA



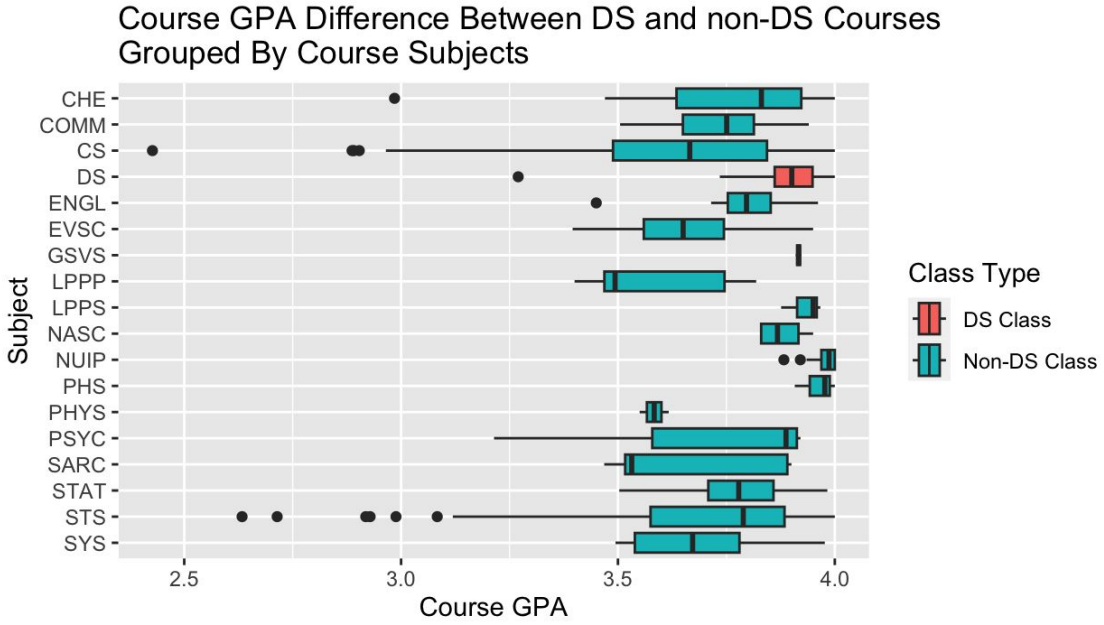

A box plot demonstrating the difference in grade distribution for DS and equivalent non-DS classes for each requirement category

# Question 4 cont.

Over the past 5 years, how do the GPA of equivalent non-DS courses compare to the DS courses?

Conclusion:

- CS and STS classes tend to have lower outliers in GPA
- NUIP and PHS classes have a higher GPA than DS classes




A box plot demonstrating the difference in grade distribution for DS and equivalent non-DS for each subject

# Conclusions

- Most DS professors have high associated GPAs regardless of taught courses, but there is a slight association between smaller classes and higher GPAs
- High probability a student can get an A/A+ in most DS courses
- DS classes tend to have higher GPAs than equivalent non-DS classes in nearly all requirement categories