(J**3)

Kaggle UW Madison Courses and Grades Dataset

Sqlite3 database:

- 9000 courses
- 200,000 course sections
- 3 million grades reported
- 18,000 instructors

Our task:

Ask questions of the data and answer them using statistical testing.

Stakeholders to whom our presentation is facing: *UW Madison Students*

Img src: https://morgridge.wiscweb.wisc.edu/wp-content/uploads/sites/4/2017/02/aerial_UW_17-35mm11_6861.jpg

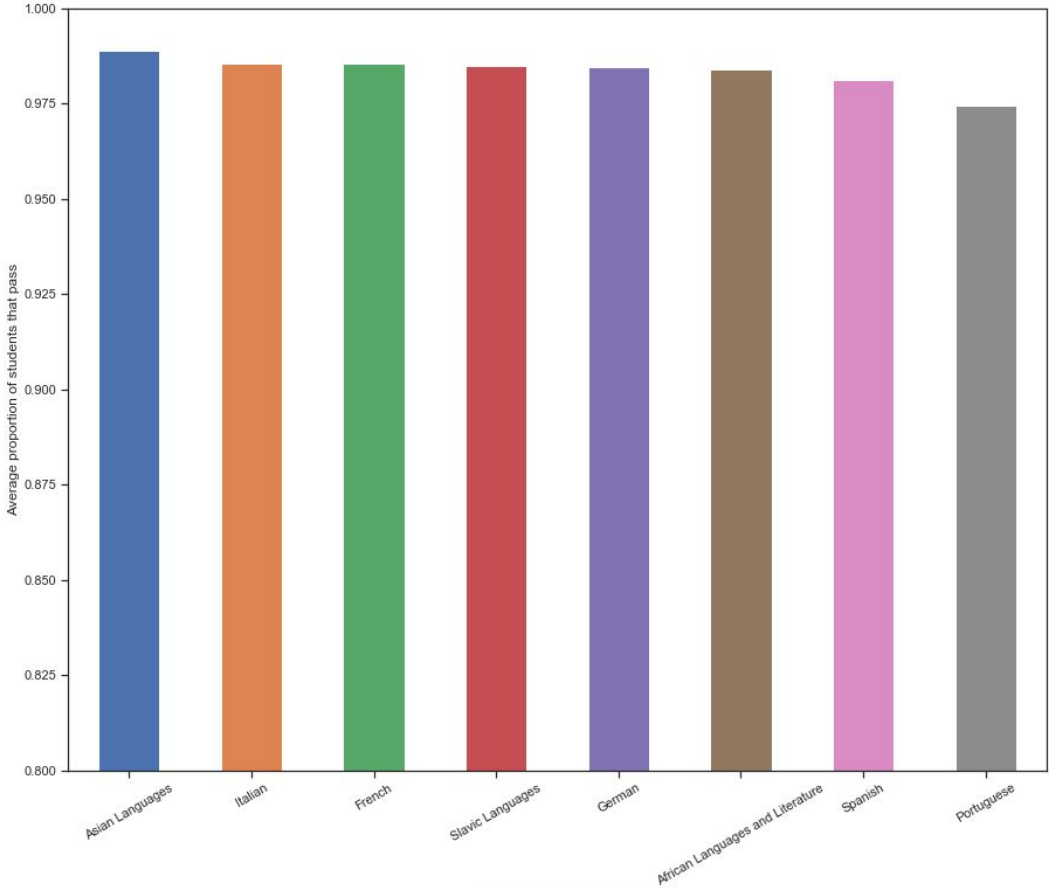
Question 1: Grade distribution by foreign language

Are higher grades awarded, on average, to classes in one foreign language versus the others?

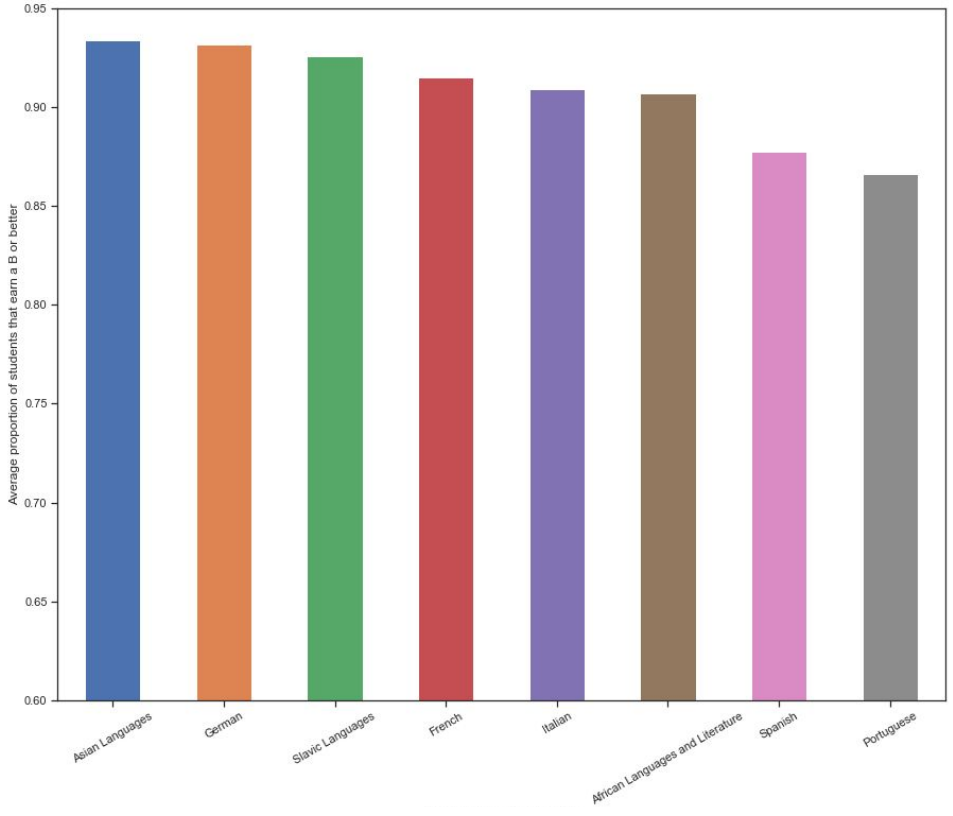
Approach:

- Kruskal-Wallis comparing proportion of students that pass in each foreign language
- Kruskal-Wallis comparing proportion of students who earn a B or better in each foreign language

Fewer students pass Portuguese classes than classes in other foreign languages



Students perform best in German and Asian language classes



Question 1 - Findings

- Fewer students pass Portuguese classes than classes in other languages
- Fewer students earn A’s and B’s in Spanish and Portuguese than in other languages
- More students earn A’s and B’s in German and Asian language classes than in other languages

Question 2: Is there a significant difference between failure rates from classes that start at 8:00-8:59 AM vs 1:00-1:59 PM.

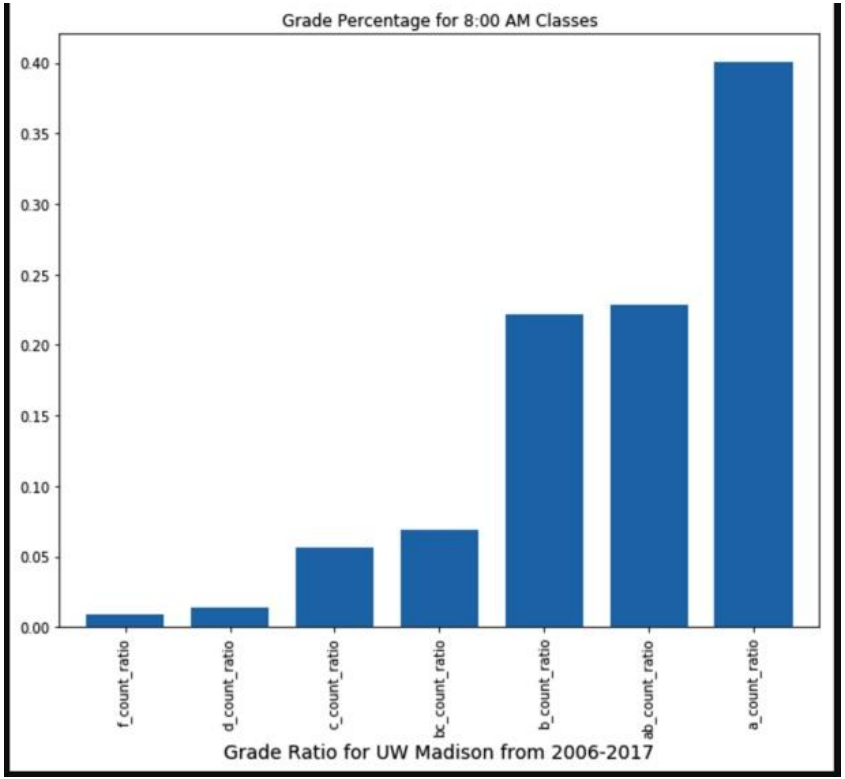
Grade Count for 8:00-8:59 AM Classes

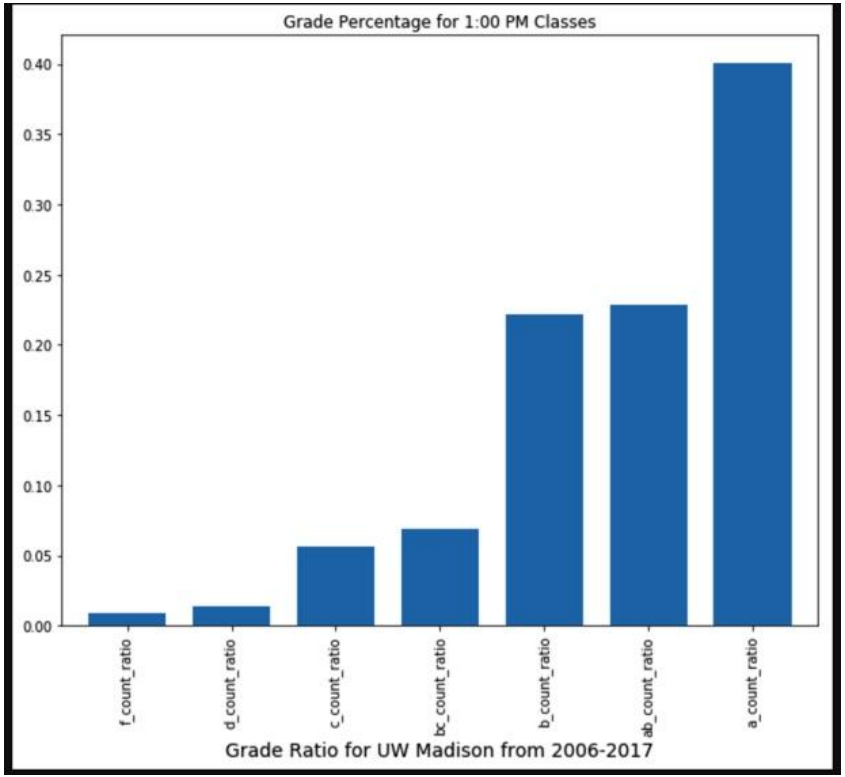
	f_count	d_count	c_count	bc_count	b_count	ab_count	a_count	start_time	total
count	88350.000000	88350.000000	88350.00000	88350.000000	88350.000000	88350.000000	88350.000000	88350.000000	88350.000000
mean	0.663113	1.423656	5.29193	4.921381	12.295020	9.361177	14.353537	8.612979	48.309813
std	2.152298	4.294526	13.53209	11.713782	24.392911	16.485655	21.185516	0.334209	81.743027
min	0.000000	0.000000	0.00000	0.000000	0.000000	0.000000	0.000000	8.000000	0.000000
25%	0.000000	0.000000	0.00000	0.000000	1.000000	1.000000	3.000000	8.500000	13.000000
50%	0.000000	0.000000	0.00000	1.000000	4.000000	4.000000	7.000000	8.833333	18.000000
75%	0.000000	0.000000	2.00000	3.000000	9.000000	8.000000	15.000000	8.833333	27.000000
max	37.000000	62.000000	132.00000	125.000000	219.000000	233.000000	347.000000	8.916667	572.000000

Grade count for 1:00-1:59 PM Classes

	f_count	d_count	c_count	bc_count	b_count	ab_count	a_count	start_time	total
count	123270.000000	123270.00000	123270.000000	123270.000000	123270.000000	123270.000000	123270.000000	123270.000000	123270.000000
mean	0.672191	1.37357	5.066083	4.744561	12.446281	9.790549	15.200089	13.250758	49.293324
std	2.088344	4.11415	12.872938	10.723963	24.166786	16.560322	27.843329	0.155317	80.416239
min	0.000000	0.00000	0.000000	0.000000	0.000000	0.000000	0.000000	13.000000	0.000000
25%	0.000000	0.00000	0.000000	0.000000	1.000000	2.000000	3.000000	13.000000	13.000000
50%	0.000000	0.00000	0.000000	1.000000	4.000000	5.000000	7.000000	13.333333	19.000000
75%	0.000000	1.00000	2.000000	3.000000	9.000000	9.000000	16.000000	13.333333	35.000000
max	37.000000	62.00000	139.000000	125.000000	219.000000	234.000000	704.000000	13.916667	716.000000

Grade Ratio for 8:00 AM and 1:00 PM Classes





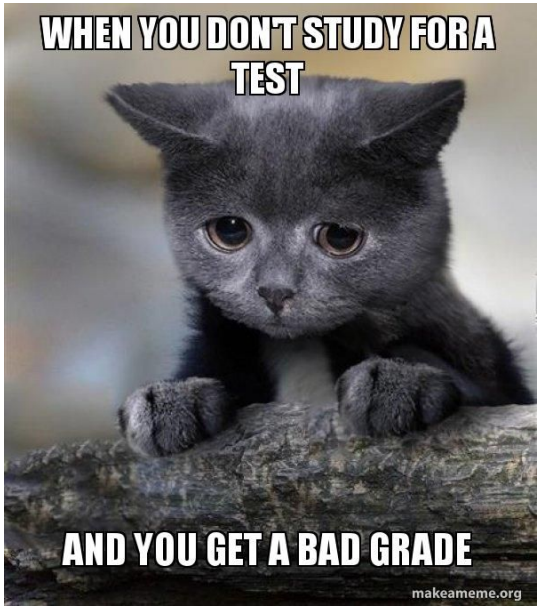
Fail Rates Conclusion

- Fail Rate for 8:00-8:59 AM Classes = 0.0226
- Fail Rate for 1:00-1:59 PM Classes = 0.0231
- Z-Test Results T-Stat = 0.9736 P-Value = 0.3302

With the results we can fail to reject the null hypothesis.

Question #3

When comparing the same course offering across different classrooms is there a significant difference in the percentage of A’s given?



Data Cleaning

- Convert grade count values to numeric
- Drop rows where there are no grades
- Drop rows where there are five or less grades
- Drop rows where the room code is ‘null’ (off campus)

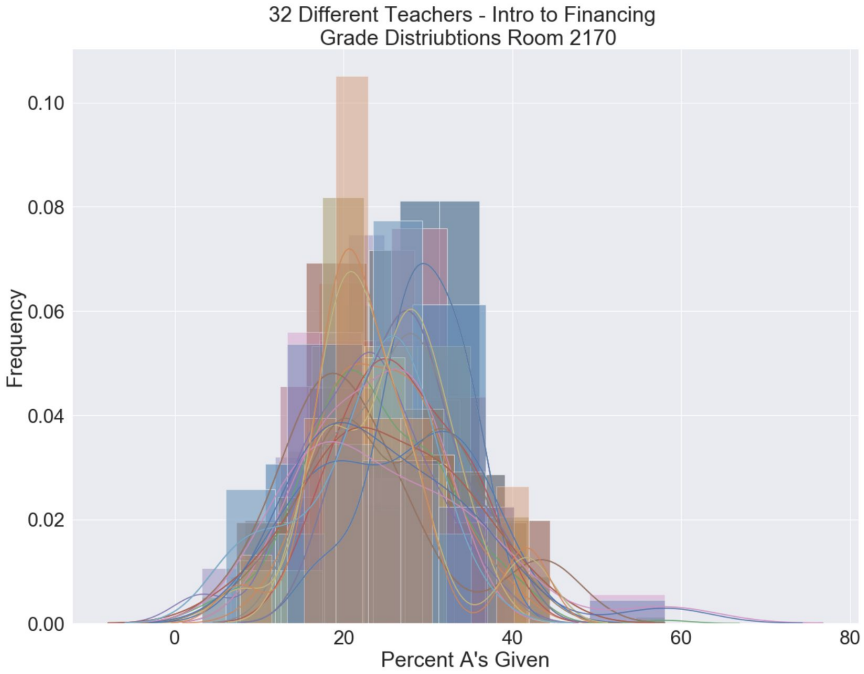
Choosing Classroom/Course Pairings

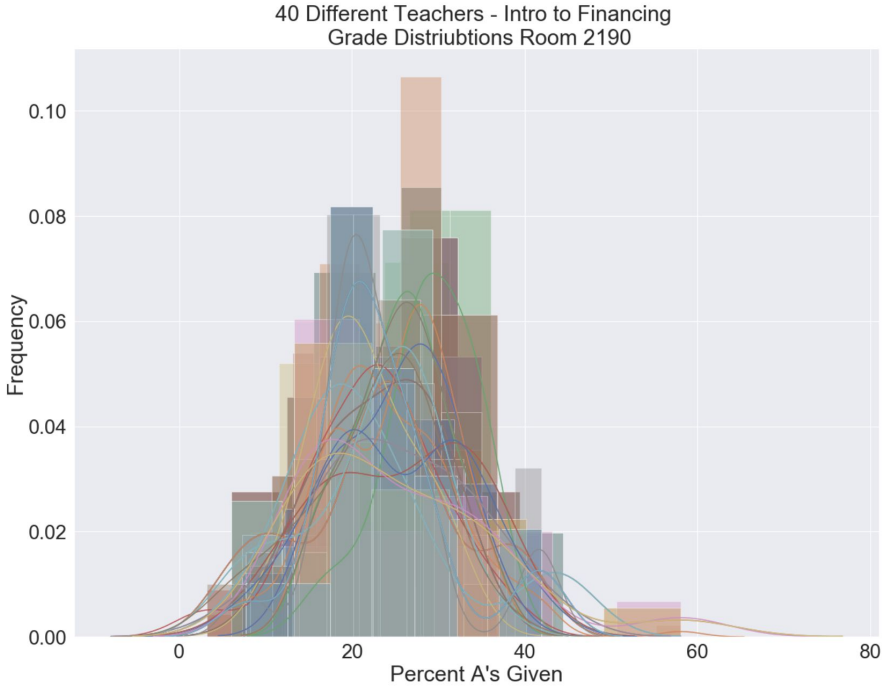
- Dropped bottom quartile of classrooms
  - Measure based on number of unique course offerings
- Also dropped bottom quartile of teachers
  - Measure based on number of unique classes taught
- Grouped by room, course, and instructor
  - About 3000 of such groupings - but we will only look at two
  - Converted grade counts in to ratios
  - Selected two classroom/course pairings to focus on

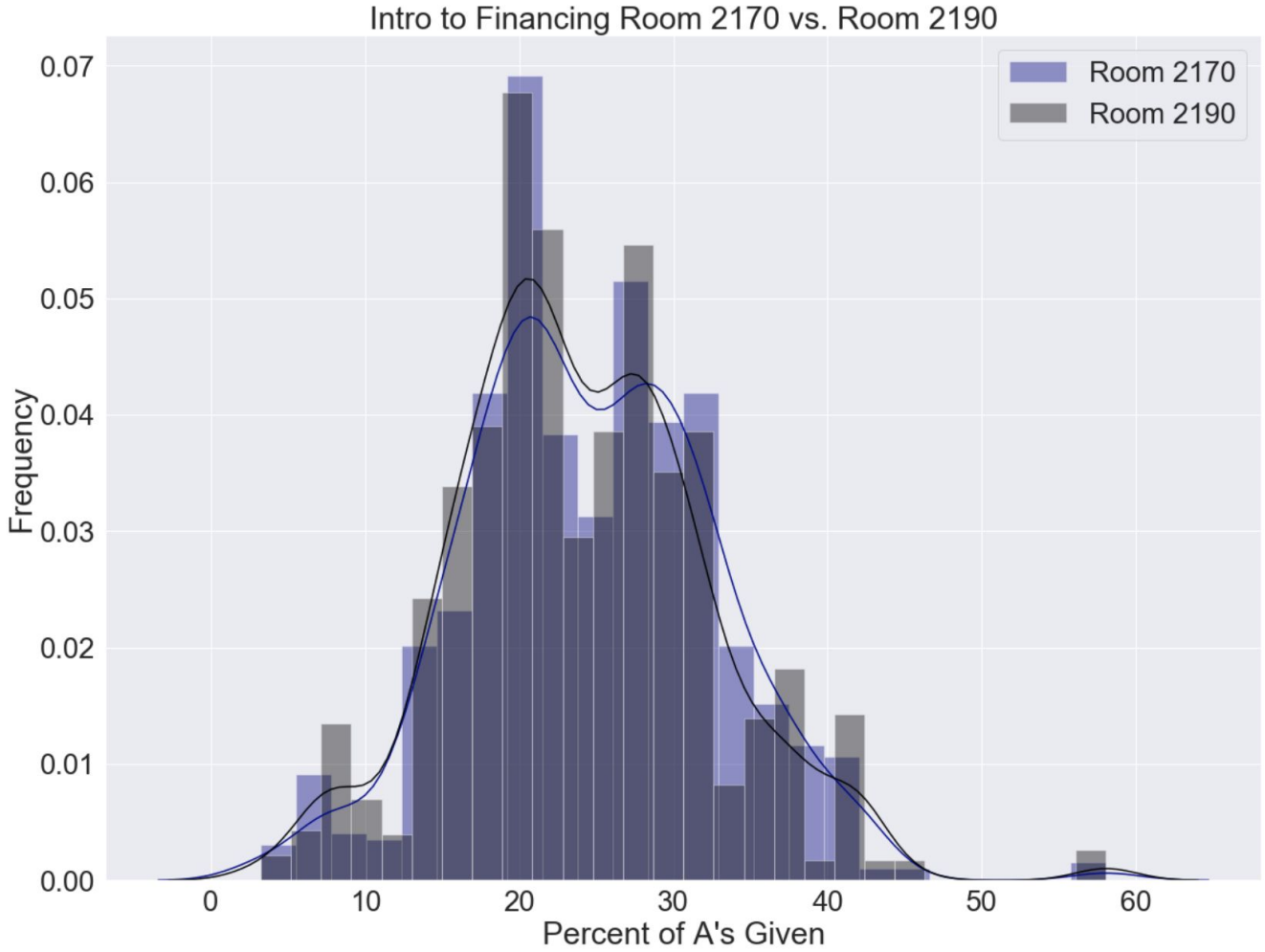
Example of room, course, instructor grouping:

[16]:

	room_code	course_name	instructor_name	A	AB	B	C	D	F	room_uuid
1356	2170	Intro Financial Accounting	DANIEL WANGERIN	0 0.212121 1 0.233333 2 0.250000 3...	0 0.000000 1 0.133333 2 0.125000 3...	0 0.363636 1 0.200000 2 0.406250 3...	0 0.242424 1 0.066667 2 0.156250 3...	0 0.151515 1 0.266667 2 0.031250 3...	0 0.030303 1 0.100000 2 0.031250 3...	026d4b38-9cc3-361d-af8d-ef162119f51e
1357	2170	Intro Financial Accounting	MATTHEW JUNEMANN	0 0.212121 1 0.233333 2 0.250000 3...	0 0.000000 1 0.133333 2 0.125000 3...	0 0.363636 1 0.200000 2 0.406250 3...	0 0.242424 1 0.066667 2 0.156250 3...	0 0.151515 1 0.266667 2 0.031250 3...	0 0.030303 1 0.100000 2 0.031250 3...	026d4b38-9cc3-361d-af8d-ef162119f51e
1358	2170	Intro Financial Accounting	TERRY WARFIELD	0 0.209302 1 0.268293 2 0.34285...	0 0.162791 1 0.048780 2 0.08571...	0 0.395349 1 0.439024 2 0.28571...	0 0.209302 1 0.097561 2 0.22857...	0 0.023256 1 0.121951 2 0.05714...	0 0.000000 1 0.024390 2 0.00000...	026d4b38-9cc3-361d-af8d-ef162119f51e
1359	2170	Intro Financial Accounting	PAUL NICHOLAS MICHAS	0 0.166667 1 0.233333 2 0.175000 3...	0 0.138889 1 0.066667 2 0.050000 3...	0 0.416667 1 0.433333 2 0.475000 3...	0 0.194444 1 0.266667 2 0.225000 3...	0 0.083333 1 0.000000 2 0.075000 3...	0 0.000000 1 0.000000 2 0.000000 3...	026d4b38-9cc3-361d-af8d-ef162119f51e


32 Different Teachers - Intro to Financing
Grade Distriubtions Room 2170
Frequency
0.00
0.02
0.04
0.06
0.08
0.10
0
20
40
60
80
Percent A's Given


40 Different Teachers - Intro to Financing
Grade Distriubtions Room 2190
Frequency
0.00
0.02
0.04
0.06
0.08
0.10
0
20
40
60
80
Percent A's Given



Are these the same distribution?

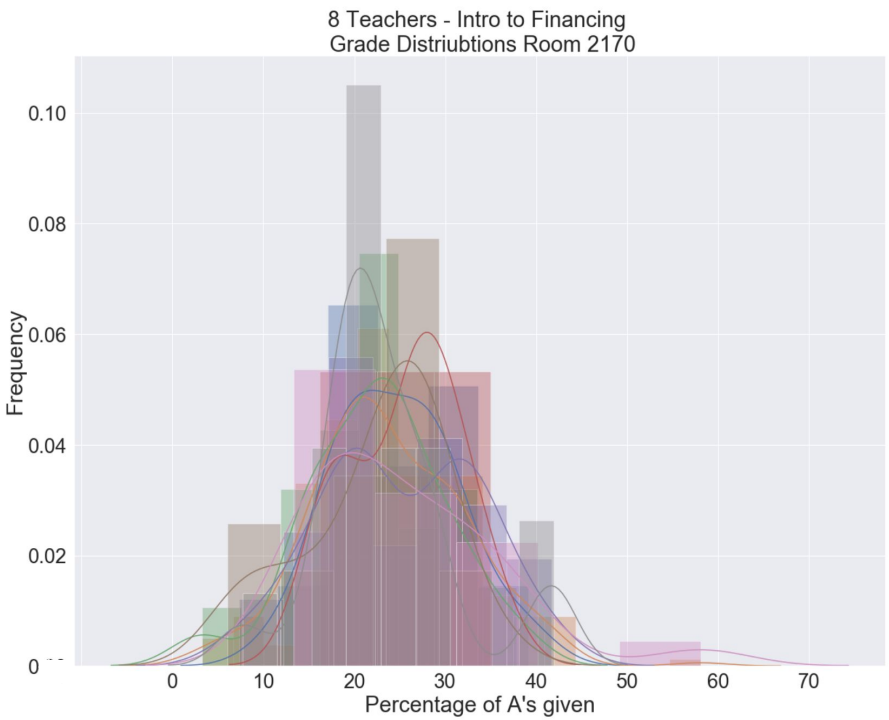
- Mean room 2170: 24.66% A's given
- Mean room 2190: 24.09% A's given

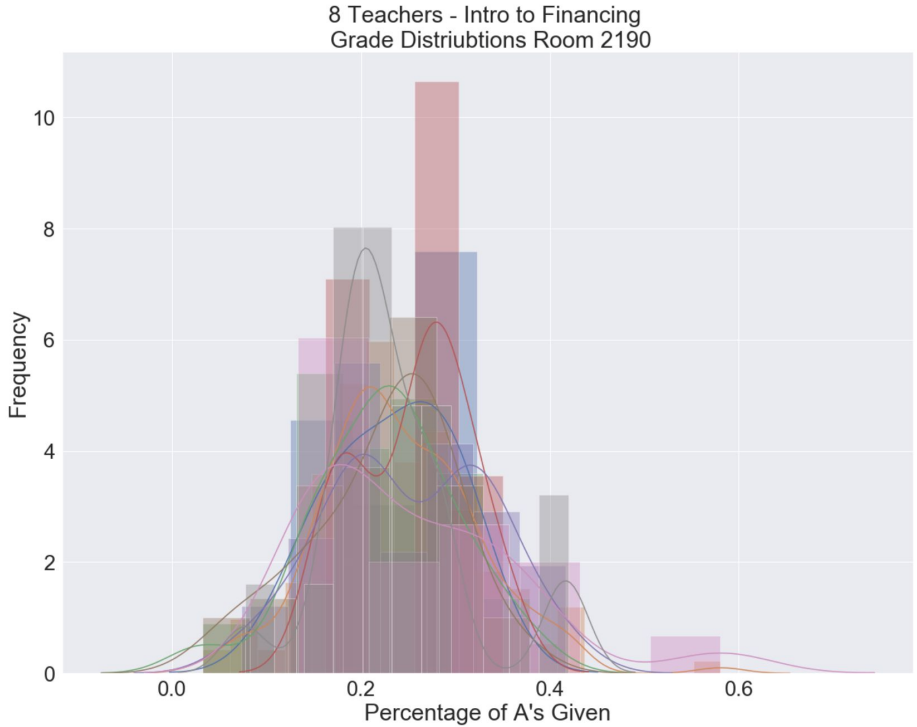
Test peformed: two side t-test.

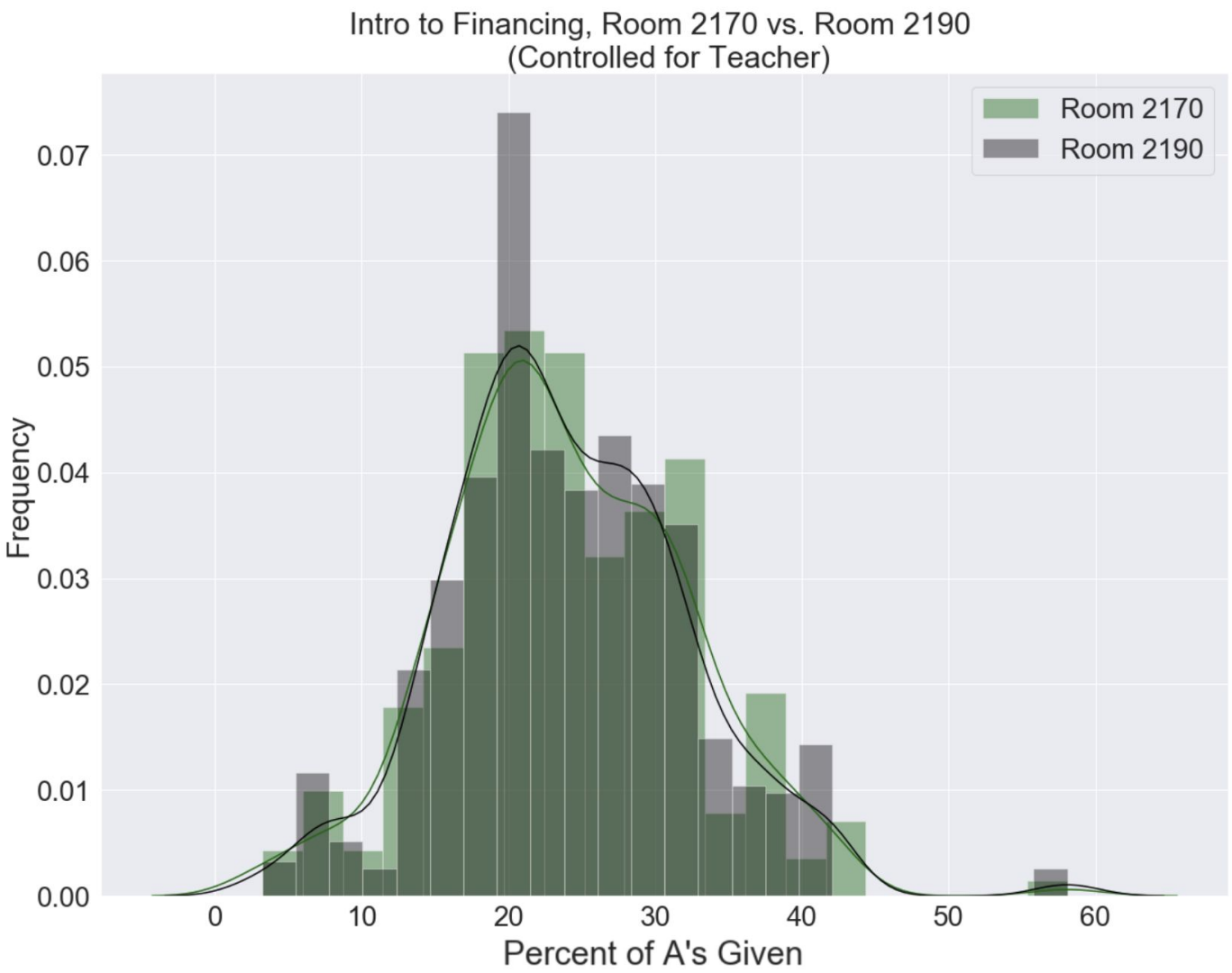
- P-value: 0.12
- Effect size: .57

Conclusion: Fails to reject that they are the same distribution at the .05 level of significance.

Now we take a look at only the same 8 teachers across the two rooms.







Are these the same distribution?

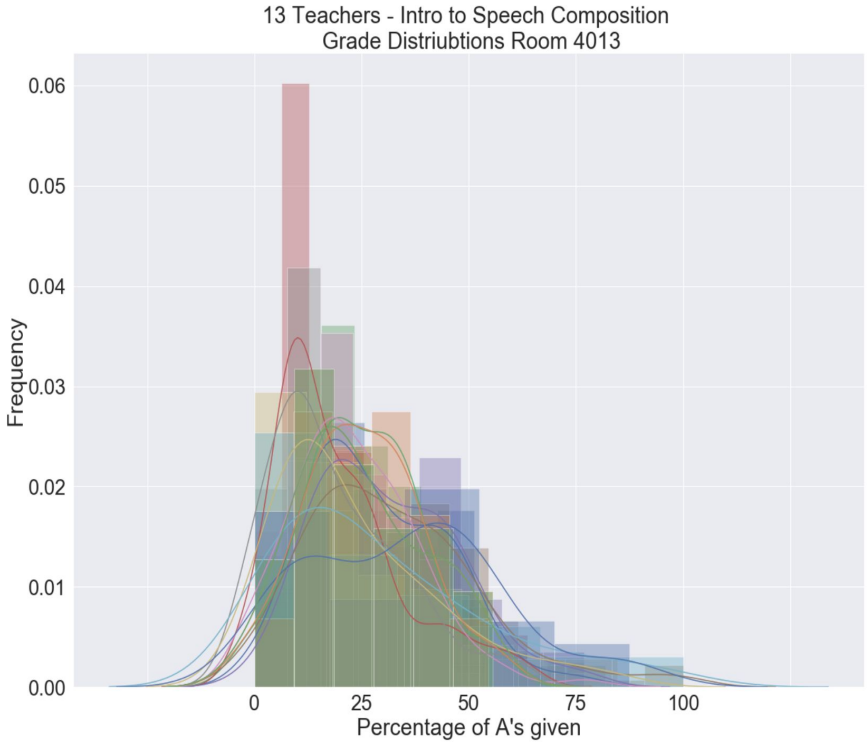
- Mean room 2170: 23.98% A's given
- Mean room 2190: 24.01% A's given

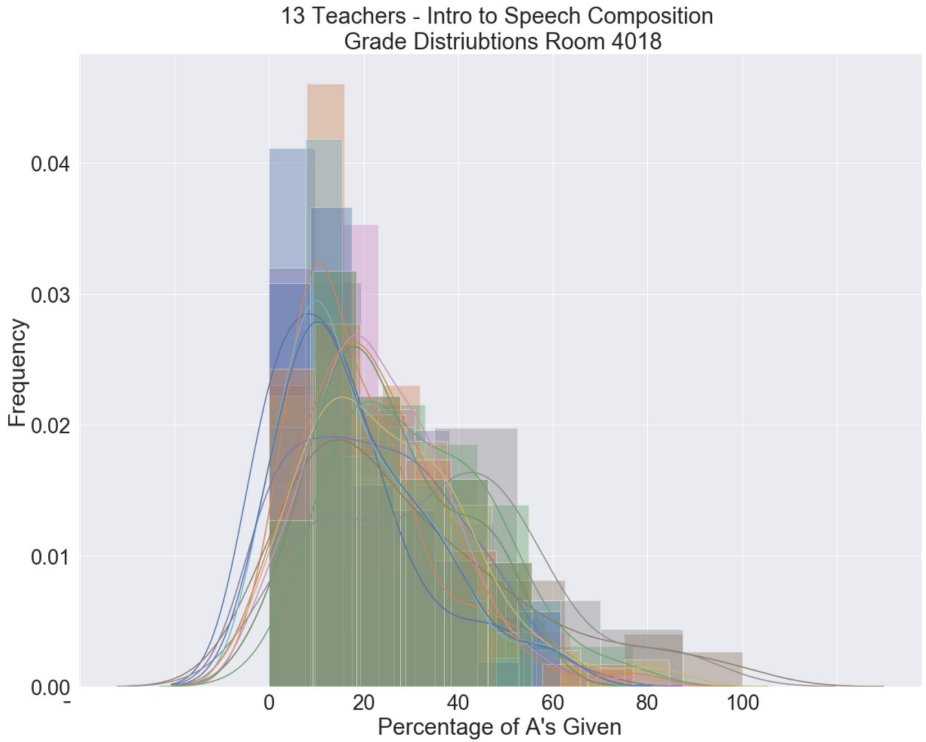
Test peformed: two side t-test.

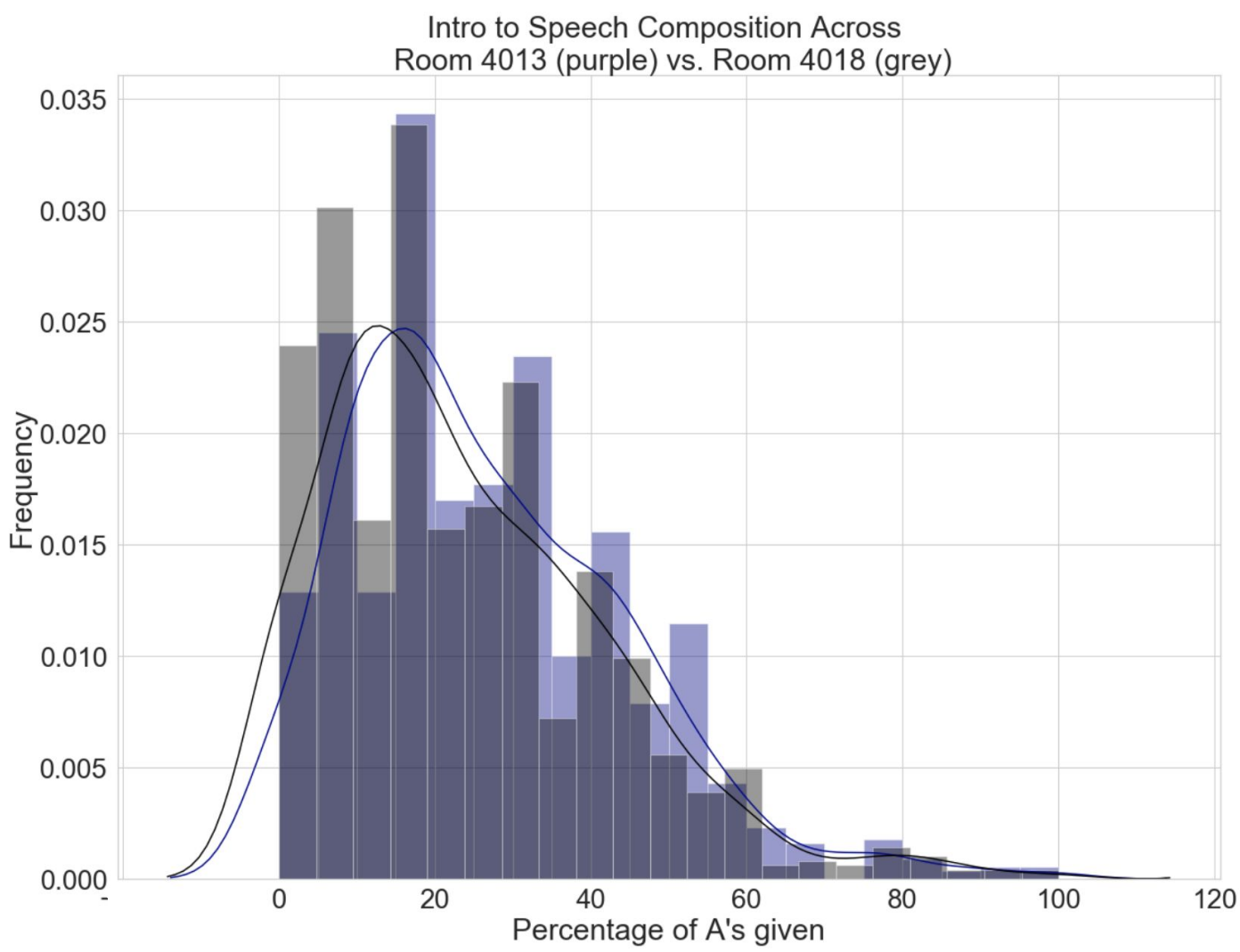
- P-value: 0.95
- Effect size: .03

Conclusion: Fails to reject that they are the same distribution at the .05 level of significance.

Now we take a look at another two classrooms across a different course - controlling again for only teachers that taught that course in both rooms.







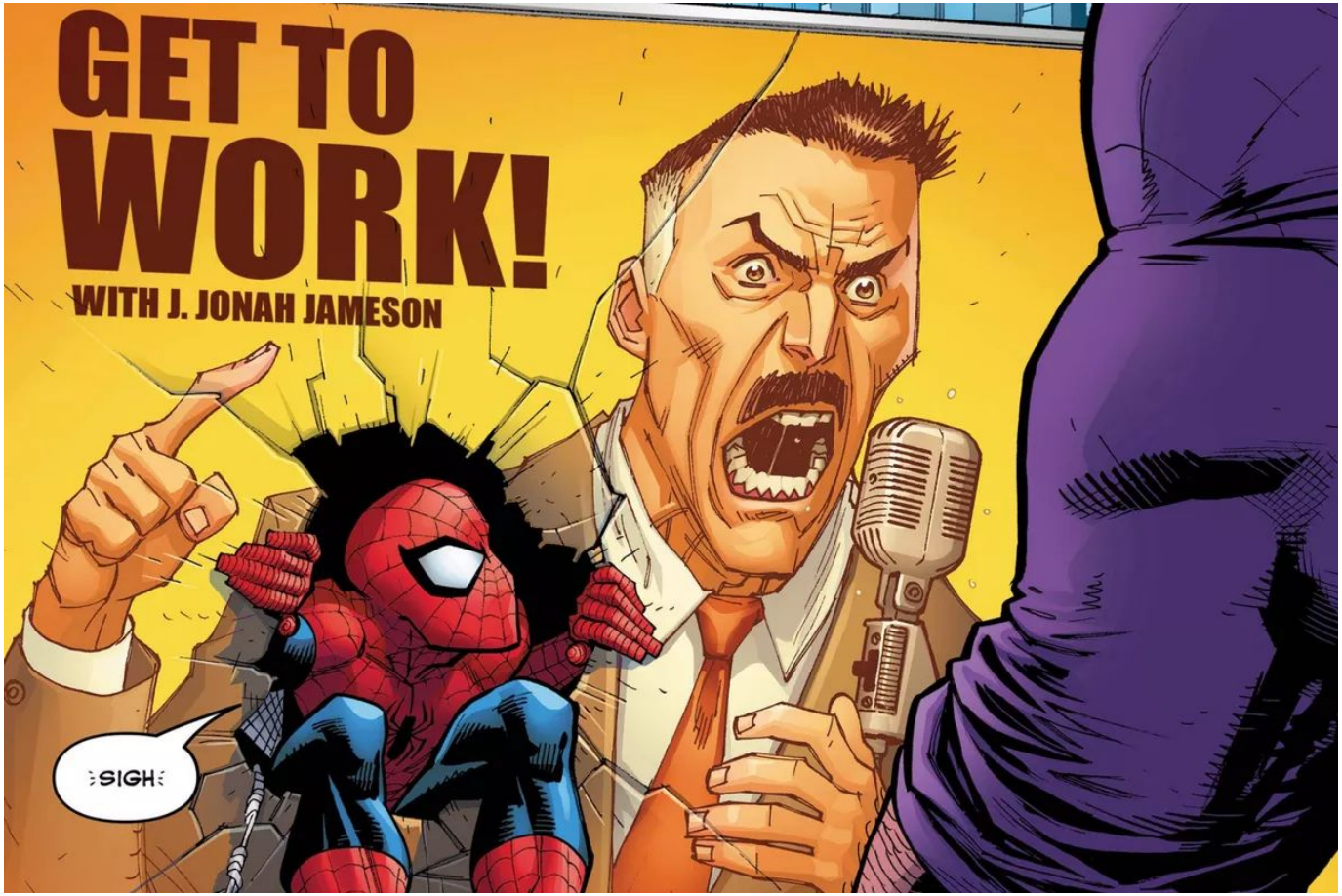
Are these the same distribution?

- Mean room 4013: 26.48% A's given
- Mean room 2190: 23.8% A's given

Test peformed: two side t-test.

- P-value: 0.00058
- Effect size: .02678

Conclusion: This test rejects these are the same distribution at the .05 level of significance.


GET TO WORK!
WITH J. JONAH JAMESON
SIGH