

PROJECT MODULE 3

Exploration of Populations and Earned Grades

Data Scientists In Training to Data Scientists In Training

Noah Sragow
MaryJo Zaborowski

DS 04222019
12 June 2019

THE PROJECT

Data Science Module 3 Project involves the assessment of a data set from the University of Wisconsin at Madison, including courses and grades covering years 2006 to 2017.

There are more than 9,000 courses in this dataset. There are nearly 200,000 course sections with grades, with 3 million grades reported in total. 18,000 instructors are also included in this dataset.

The objective of this study is statistical analysis and hypothesis testing. We pursue four hypotheses, follow proper experimental design principles, and report results.



TRANSCRIPTS

		GRADE	HOURS	MSH	CTP	MHP
Term Total	GPA:	A	3.00	3.00	3.00	1
Cumulative Total	GPA:		3.00	3.00	3.00	1
Fall 2006	Undergraduate Kinesiology					
EDUC 392	Educ Multicult Soc	B+	3.00	3.00	3.00	9.90
EDUC 402	Rdg Wrtg Content	A	3.00	3.00	3.00	12.00
PHYSED 305	Prac Elem Teach Meth	S	1.00	0.00	1.00	0.00
PHYSED 354	Meth of Teach K-5 PE	A+	3.00	3.00	3.00	12.00
PHYSED 373	Issue Hlth Wellness	A	3.00	3.00	3.00	12.00
Term Total	GPA:	3.825	13.00	12.00	13.00	45.90
Cumulative Total	GPA:	3.620	72.00	110.00	260.70	
Winter 2007	Undergraduate Kinesiology					
EDUC 307	Practicum	S	1.00	0.00	1.00	0.00
						12.00

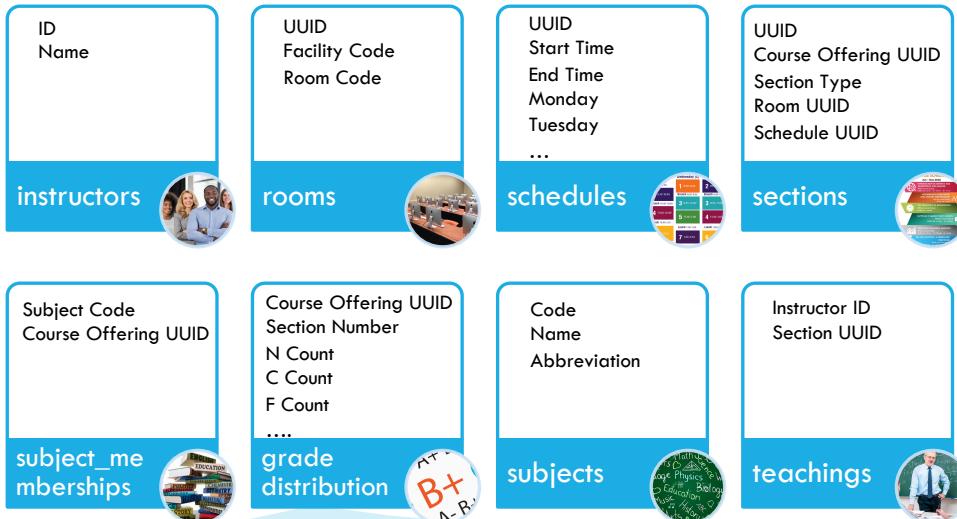
THE DATA

Tables are highly normalized

Normalization is a database design technique which organizes tables in a manner that reduces redundancy but increases dependency of data.

It divides larger tables to smaller tables and links them using relationships.

Any question requires at least three table joins



Data Sources

grade_distributions....	193k x 18
instructors.	18.7k x 2
rooms.	1350 x 3
schedules	4468 x 10
sections.	316k x 6
subject_membershi...	95.3k x 2
subjects	200 x 3
teachings.	315k x 2

course_offering_uuid	section_number	n_count	c_count	f_count	cr_count	p_count	nw_count	other_count	bc_count	u_count	d_count	ab_count	a_count	s_count	i_count	b_count	nr_count
344b3ebe-da7e-314c-83ed-9425269695fd	1	0	0	0	0	0	0	0	0	0	0	0	105	0	0	0	0
f718e6cd-33f0-3c14-a9a6-834d9c3610a8	1	0	0	0	0	0	0	0	0	0	0	0	158	0	1	0	0
ea3b717c-d66b-30dc-8b37-964d9688295f	1	0	3	0	0	0	0	0	0	0	0	12	139	0	0	2	0
075da420-5f49-3dd0-93df-13e3c152e1b1	1	0	0	0	0	0	0	0	0	0	0	0	87	0	1	0	0
2b4e216d-a728-3713-8c7c-19afffc6b2fd	1	0	0	0	0	0	0	0	0	0	0	0	70	0	1	0	0

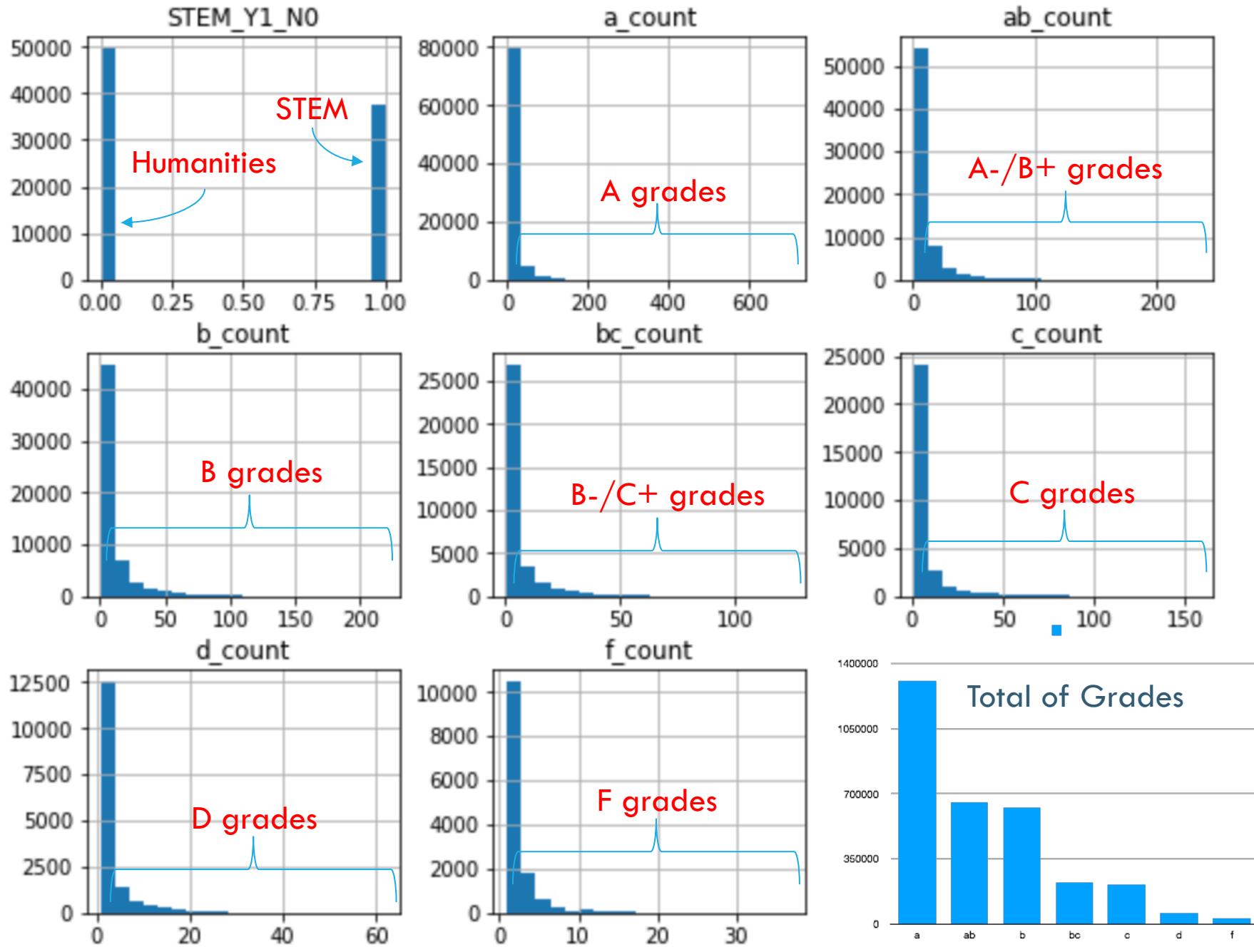
THE DATA

PART DEUX

This university awards many A's...

The real story is in the x-axes:

- The A grades x-axis extends to over 700, whereas others are much smaller
- There are 11 courses that 600-700 people each earned A's with an A- ratio of >94%
- What were some of these courses?
 - Surgery
 - Emergency Medicine
 - Engineering Mechanisms & Astronautics
 - Chem & Bio Engineering
 - Electrical & Computer Engineering



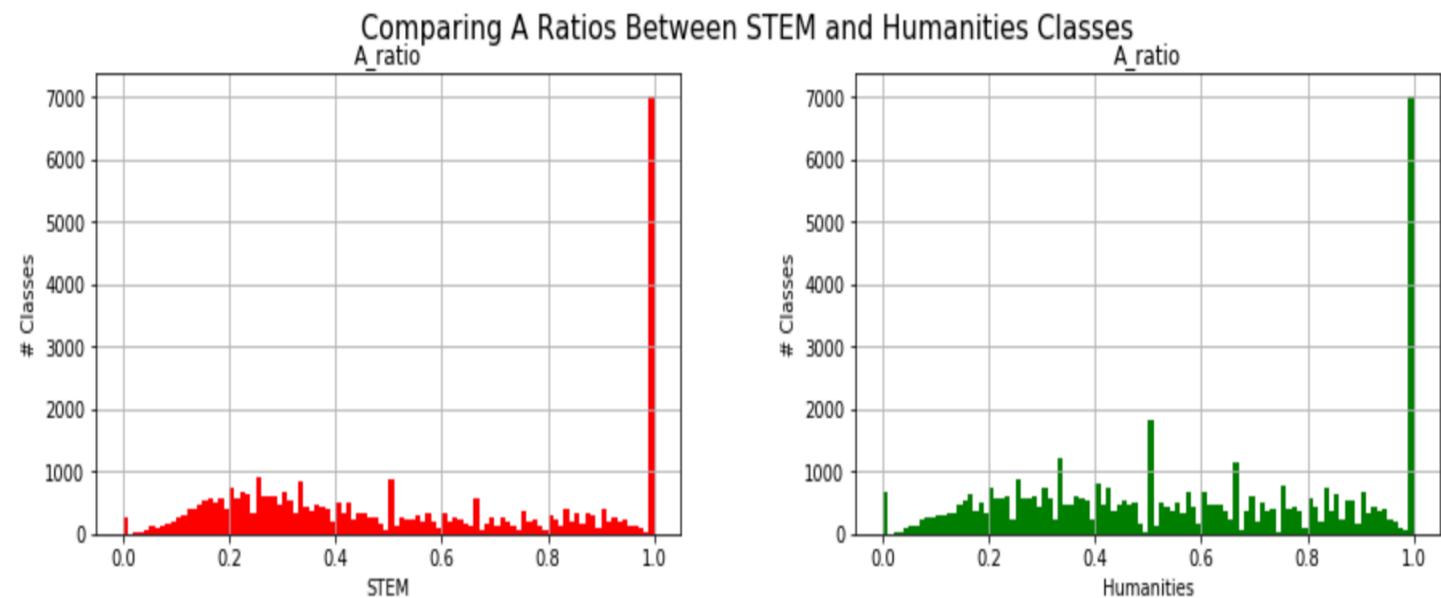
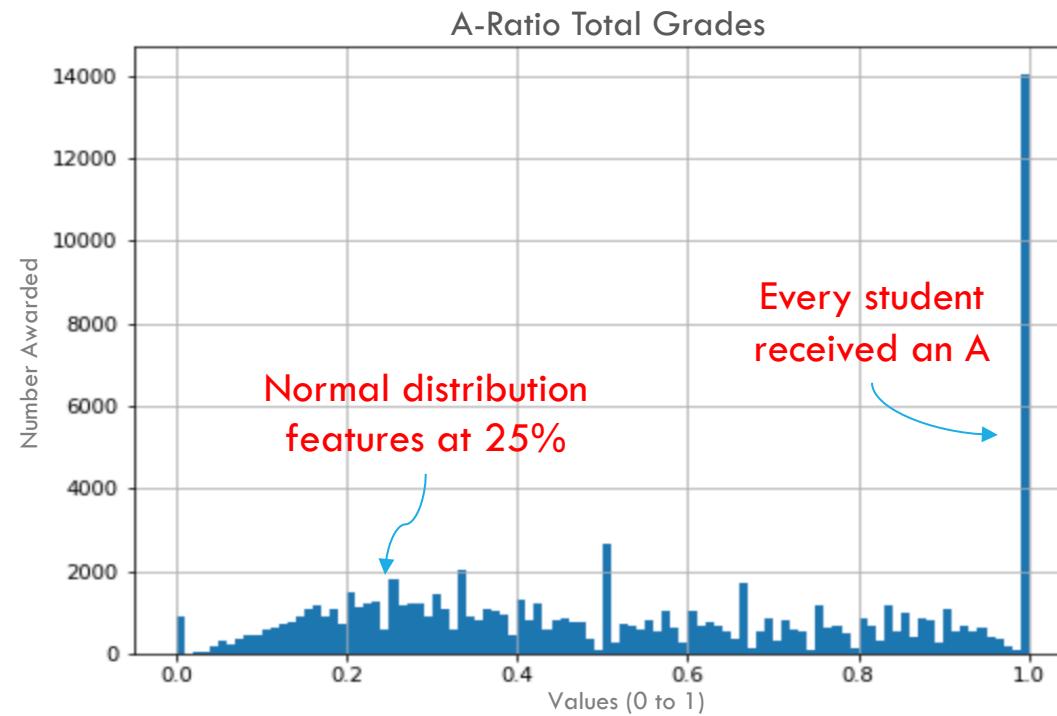
THE DATA

PART THREE

Because courses were graded without consistent normal distribution of data, the A data was standardized to improve comparability of assessment across classes.

An A-Ratio was calculated for every class:

$$\text{A Ratio} = \frac{\text{A's / Class}}{\text{Total Grades/Class}}$$



THE QUESTIONS

Q1: Do STEM fields have a statistically significant difference in the number of A's earned when compared to the humanities?

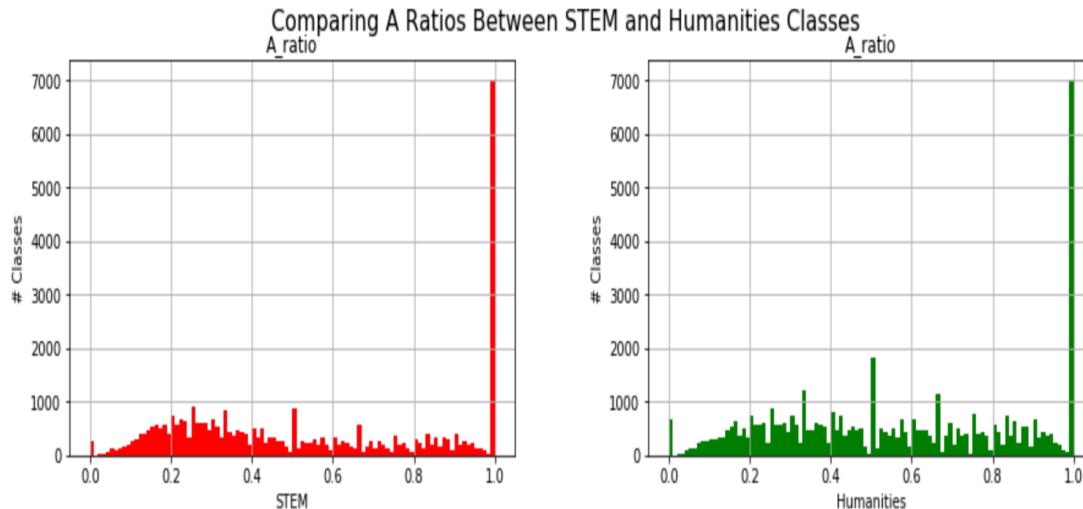
Q2: Does time of day have a statistically significant correlation with the number of A's earned in a course?

Q3: Do individual teachers' distribution of grades differ from the distribution of all other teachers?

Q4: Do individual class grade means follow the grade mean of all classes?

THE ANALYSIS Q1

Do STEM fields have a statistically significant difference in the number of A's earned when compared to the humanities?



To assess whether there is a difference in the number of A's earned in STEM courses vs Humanities courses, we tested the "A ratio," which was the number of A's divided by the number of total grades for each class. Our STEM courses included 37,390 classes and our Humanities courses included 49,819 classes.

Our Null Hypothesis (H0): There is no difference in the A's earned in STEM classes vs. Humanities (Not STEM) classes.

Our Alternative Hypothesis (HA): There is a difference in the A's earned in these populations.

Our analysis returned a t-statistic of -10.9 and a pvalue of 1.285e-27, given 87207 degrees of freedom.

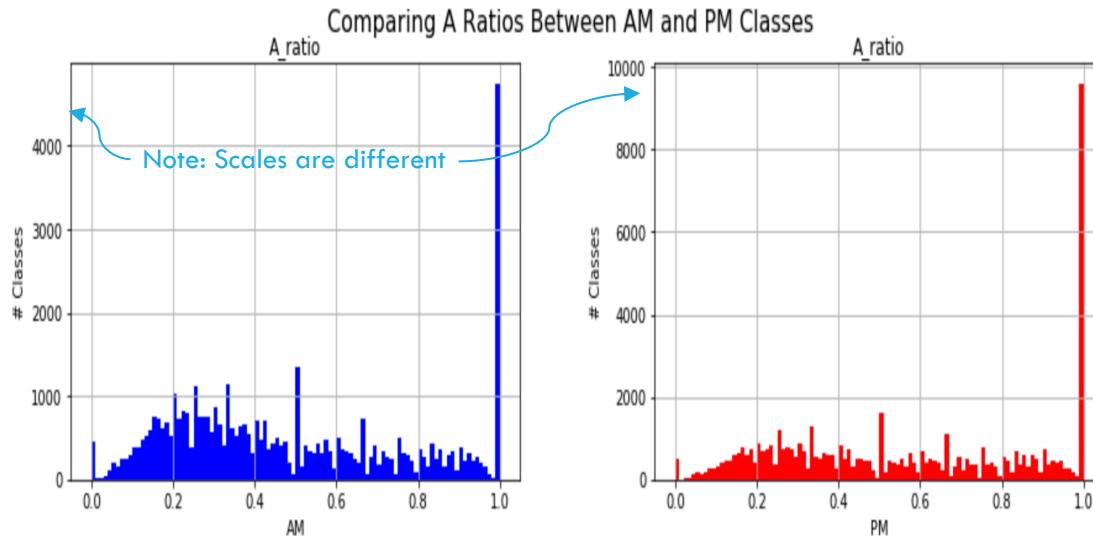
Negative t-value: The sign of a t-value tells us the direction of the difference in sample means. The mean of our STEM population was less than the mean of our Humanities population. In other words, on average, the A Ratio of STEM classes was lower than that of Humanities courses by 2.3%.

Given the p-value of 1.285e-27, and an alpha of 0.05, the data show we can reject the Null Hypothesis.

Conclusion: these data show there is a significant difference in the number of A's earned in STEM classes vs. Humanities classes.

THE ANALYSIS Q2

Do AM classes have a statistically significant difference in the number of A's earned when compared to the PM classes?



To assess whether there is a difference in the number of A's earned in morning (AM) courses vs afternoon/evening (PM) courses, we tested the "A ratio," which was the number of A's divided by the number of total grades for each class. Our AM courses included 45,281 classes and our PM courses included 55,811 classes.

Our Null Hypothesis (H0): There is no difference in the A's earned in AM classes (before 12:00) vs. PM (after 12:00) classes.

Our Alternative Hypothesis (HA): There is a difference in the A's earned in these populations.

Our analysis returned a t-statistic of -41.9 with a pvalue of 0, given 101,090 degrees of freedom.

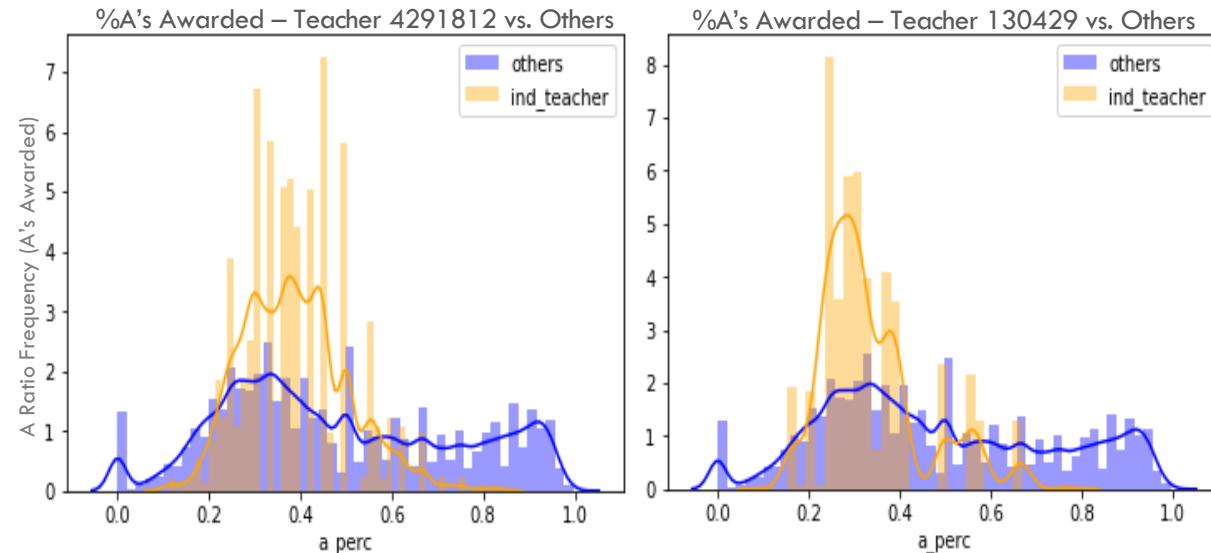
Negative t-value: The mean of our AM population was less than the mean of our PM population. In other words, on average, the A Ratio of AM classes was lower than that of PM courses by 7.9%.

Given the p-value of 0, and an alpha of 0.05, the data show we can reject the Null Hypothesis.

Conclusion: these data show there is a significant difference in the number of A's earned in AM classes vs. PM classes.

THE ANALYSIS Q3

Do individual teachers' distribution of grades differ from the distribution of all other teachers?



To assess whether there is a difference in the number of A's awarded by the top 100 teachers with most classes, we tested the "A ratio" for each teacher against the rest of the teachers. The A ratio is the number of A's divided by the number of total grades for each class. Classes where only A's were awarded were removed from this analysis. The resulting dataset included 17,199 teachers.

Our Null Hypothesis (H0): There is no difference in the A's awarded by individual teachers vs. all other teachers.

Our Alternative Hypothesis (HA): There is a difference in the A's awarded in these populations.

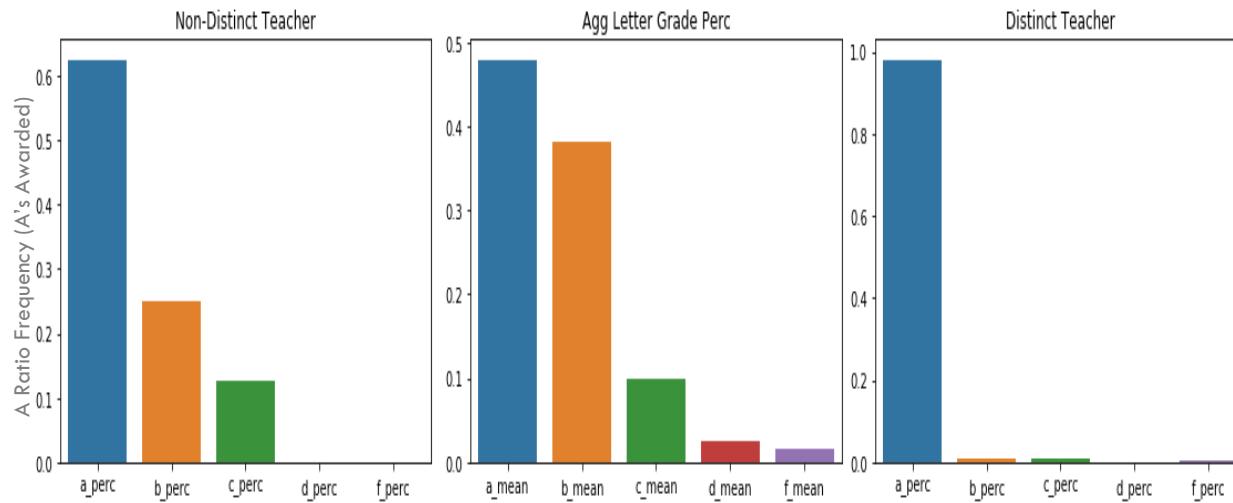
Using Welch's two sample T-test, 90% of the top most prolific teachers had a pvalue <0.05. 10% had a pvalue >0.05.

For the 90% set, given the p-values and an alpha of 0.05, the data show we can reject the Null Hypothesis.

Conclusion: these data show there is a significant difference in the number of A's awarded by 90% of the most prolific teachers vs. all other teachers.

THE ANALYSIS Q4

Do individual class grade means follow the grade mean of all classes?



To assess whether a class gave out grades based on the school's average grade distribution. Classes where only A's were awarded were removed from this analysis. The resulting dataset included 608,321 classes.

Our Null Hypothesis (H0): The grades follow the school's grade distribution.

Our Alternative Hypothesis (HA): There is a difference in the grade distributions across these populations.

A Pearson's Chi Squared test was applied to each class, with the expected values calculated based on the mean percents of each grade.

Approximately 40% of the classes returned a p-value less than an alpha of 0.05. For these classes, we reject the Null Hypothesis.

The remaining approximate 60% of the classes showed a p-value \geq an alpha of 0.05. For these, we are unable to reject the Null Hypothesis.

Conclusion: 40% of the classes showed a significant difference in the grading distribution than the overall grading distribution.