Motivating Survey Responses Through Incentives

Alexandra Drossos, Anand Patel, Madeline Whitlow, Spencer Weston, Carlos Moreno
GitHub Repository: https://github.com/Spencer-Weston/W241_Final_Project.git

1: Introduction

A common challenge that researchers face in conducting survey studies is low response rates, which has implications in the form of reduced effective sample sizes and increased risk of nonresponse bias. Both scenarios limit the reliability of any conclusions drawn from survey data. There are a number of methods that have been adopted and tested to improve response rates, but for this study we will experiment on the effectiveness of offering monetary incentives for online survey completion. We will compare two treatment group response rates against the control group response rate to test the effectiveness of monetary incentives.

There have been various studies conducted to better understand what types of survey incentives best motivate respondents to participate. A paper by Harris et al.,(1) published in 2008 found that lottery incentives did not increase response rates on surveys while a different paper by Smith et al.(2) published in 2019 found that providing pre-incentives increased survey response rates. Additional research indicates that monetary incentives improve response rates and that different incentive schemes have different effects. A systematic review by Pit et al. surveyed different incentive schemes for inducing survey responses from general practitioners. They found that monetary incentives outperform non monetary incentives, upfront incentives outperform promised incentives, and large incentives outperform small incentives. These results are corroborated by Smith et al. which reports that the inclusion of $2 in a survey request letter more than triples the response rate. Pedersen et al. also reports a positive effect of incentives on survey response rates.

With various results and experiments out there we want to better establish which types of monetary incentives were most effective comparatively specifically among students in a Masters program. Knowing that the results of these experiments can vary based on the candidate pool of interest, we hypothesize that specific monetary incentives will have a causal impact on survey response rates for students in the MIDS (master in data science) program. This theory is based upon other established theories, particularly the incentive theory of motivation. This refers to a behavioral theory that suggests people are motivated by a drive for incentives and reinforcement. The incentive theory also proposes that people behave in a way they believe will result in a reward and avoid actions that may entail punishment. These established theories, alongside the previously mentioned studies investigating this concept have informed our hypotheses.

2: Experimental Details

As described above, our randomized control trial aims to test for a causal impact of a monetary incentive on survey response rate. To accomplish this we will be comparing the potential outcomes of our subjects against each other, in essence comparing two different states of the world. In the first state of the world, a subject receives an email requesting response to a survey and two weeks later we report if that subject has responded. In the second state of the world, a subject receives an email *with a monetary incentive* requesting response to a survey and two weeks later we report if that subject has responded. Then in accordance with the potential outcomes framework, the causal effect of the monetary incentive on the response rate to the survey is the difference in the response rate between these two states of the world. These two states of the world and our collection process is summarized below using an ROXO comparison.

0	R	X	O
1	R	X	O

Row 0 represents the state of the world in which a survey is sent out to the randomly assigned group “R” without any monetary incentive, whereas row 1 represents the state of the world in which a survey is sent out *with* some monetary incentive defined by the treatment “X”. The outcomes “O” are measured two weeks after treatment “X” is applied.

We gathered emails from all members of five UC Berkeley MIDS Slack channels corresponding to the members’ first semester in the UC Berkeley Masters in Data Science (MIDS) program. These channels include Summer and Fall in 2020 and Spring, Summer, and Fall in 2021 which corresponds to the subjects’ MIDS cohort. Table 8.1 (appendix) details the number of emails acquired by Slack channel.

Treatment groups were assigned with block randomization with blocks defined by the Slack channel where each subject’s email address was acquired. From every cohort, 72 subject’s email addresses were randomly selected for participation in the study for a total of 360 subjects. Each block of 72 subject email addresses was randomly assigned (1:1:1) into control, Treatment 1, and Treatment 2 groups. This results in 24 subjects per cohort within each block and 120 subjects per treatment group in the study as a whole. This results in group sizes of 120 individuals. (See details in Figure 2.1).

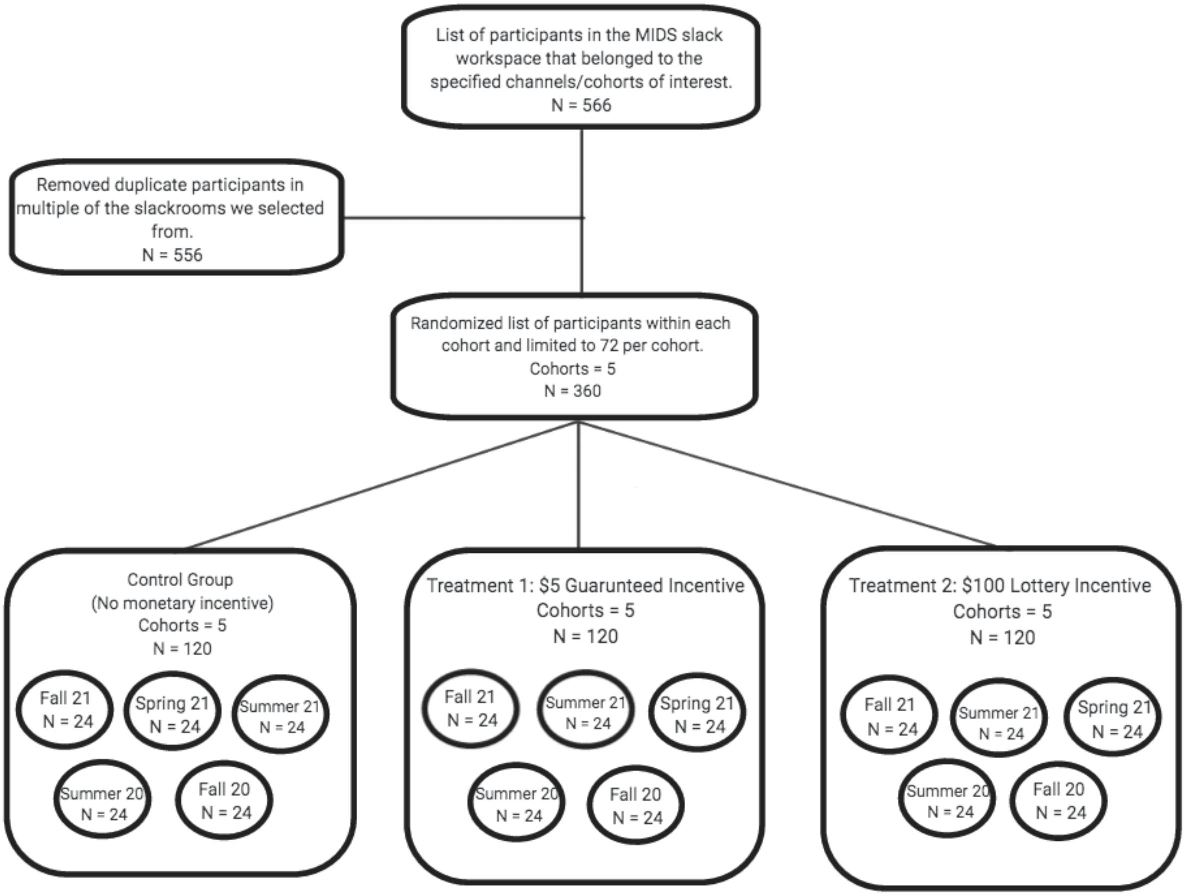


Fig 2.1: Detailed overview of how treatment groups were formed from the participant pool.

The abstract treatment for this randomized control trial involves providing monetary incentive to subjects to respond to a survey via email. However, prior research indicated that *different* incentive schemes have had different effects on response rate. To incorporate this, two treatment variants were created to then be compared against the control group of receiving no monetary incentive at all. For subjects in Treatment Group 0, the control group, subjects were sent an email requesting their response in a “MIDS Class Time Requirement” survey. For subjects in Treatment Group 1, they were sent an email guaranteeing a $5 Amazon Gift card for participating in the same survey. For subjects in Treatment Group 2, they were sent an email guaranteeing entry into a $100 Amazon Gift Card lottery for participating in the same survey. It’s important to note that the odds of winning this lottery were not disclosed as part of our email intervention, and this will be relevant to our analysis of our results. The content of each of these three emails was identical apart from the mention of a monetary incentive for Treatment groups 1 and 2 (see content in Appendix for reference). This initial response request email was the only intervention in this randomized control trial. We allotted two weeks for the subjects to respond to our survey, without any further correspondence such as a reminder email. Lastly, we used the

BCC feature to ensure that individuals receiving the email could not see who the other participants were.

Apart from the initial request email, the actual content of the sent survey is an important aspect of the treatment because survey response rates can depend on a combination of different factors apart from just monetary incentive - origin or sponsor of the survey, structure and length of the survey, assurance of confidentiality, and interest in focus of research - can all be identified as influential factors in increasing online survey response rates. For this experiment, these other variables were kept constant, but were chosen to give us the best results. Specifically, the intervention email came from a berkeley student's email address and contained an official Berkeley header, to show legitimacy and signal trust. In addition, the survey content asked students to provide their opinions on time requirements for certain classes and if there were any classes they wish MIDS provided (see full survey in Appendix). This subject matter was intended to peak the interest of subjects. And lastly, we kept the survey short to avoid subjects not wishing to complete it.

After the treatment was assigned and the intervention emails were sent out, subjects had two weeks to complete the survey before we began recording outcome measures. The primary outcome measure we collected was response rate, which was calculated as the number of completed surveys in treatment group X (0,1,2) divided by the number of surveys sent out to treatment group X.

For each individual subject, their outcome measure was binary: 1 = completed survey, 0 = did not complete survey. This yielded the count of completed surveys per treatment group. In collecting this outcome measure, we had to keep non-compliance in mind. Non-compliance in an experimental study arises when participants do not receive the treatment or intervention to which they were randomly allocated. In our case, this would occur if for some reason a participant did not receive or even see the email sent out to them inviting them to participate in the survey (this could have occurred if students don't have access to this email anymore or if the email went straight to spam). Our intervention process was designed specifically to try to avoid non-compliance - using a Berkeley.edu sender, only including subjects active in slack channels, etc. - but there's no guarantee that it was perfectly remedied, so we have taken that into consideration for our analysis.

3: Power Calculation

Power Calculation: Pre Experiment Implementation Calculation.

A simulation was implemented to estimate the group sizes required to achieve 80% power. “No Incentive” was considered the control group, and “$5 Incentive” was considered the main treatment group. The expected response rates were estimated as follows:

- 5% response rate for “No Incentive” group (no monetary incentive provided).
- 16% response rate for “$5 Incentive” group ($5 Amazon Gift card for participants answering the survey).

The response rates were estimated on the lower range to reflect the following:
- Participants were given only two weeks to respond (from October 31st to November 14th).
- One of the two weeks was a break at MIDS (Nov 8th to Nov 14th).
- No reminder was sent to participants.

The following graph presents the power calculation curve used to define group sizes to achieve 80% power. Based on the simulation, it was determined that groups of 120 participants are needed to achieve 80% power (indicated by the red line in the graph below).

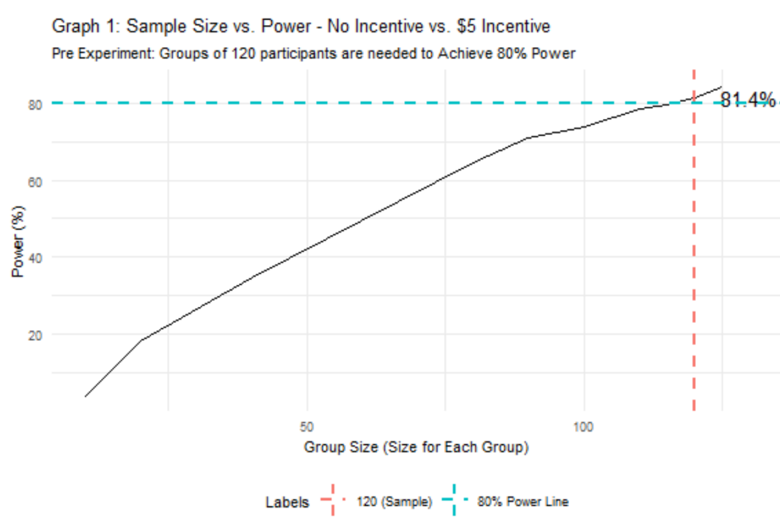


Fig 3.1: Pre-experiment power calculation curve used to define group sizes to achieve 80% power.

4: Data Collection Process

Response data was collected from Google Forms. The response data included the respondent's email and their response to the survey questions. Survey responses were received from 52 out of the 360 subjects for an overall response rate of 14.4%. To determine each respondent's block and treatment group, respondent email addresses were matched with respondents' email addresses by block and treatment group. No respondent emails matched treatment assignment emails, as explained below, so permutations were applied to email addresses to facilitate the matching process between the treatment and response email addresses.

Due to a peculiarity in how the MIDS program assigns email addresses, each student receives two email addresses. The first ends with "@berkeley.edu" and the second ends with "@ischool.berkeley.edu". Emails to the latter address are received in the former address's inbox. The "@ischool.berkeley.edu" address tends to be used to sign into campus resources whereas the "@berkeley.edu" is the standard email address for communications. MIDS students' Slack profiles provide the "@ischool.berkeley.edu" email address whereas all responses came from "@berkeley.edu" email addresses. When we account for this discrepancy, 41 of the 52 responses could be matched to their treatment group.

The 11 remaining response emails contained additional discrepancies. Consider someone named Jane Doe who has different treatment and response email addresses. Examples of their email could include: Jane.Doe@berkeley.edu, Jane-Doe@berkeley.edu, JDoe@berkeley.edu, JDoe123@berkeley.edu, JaneDoe@berkeley.edu, or JaneDoe123@berkeley.edu. Seven of the 11 remaining responses could be matched to treatment by considering these variations. For example, we may find a match by stripping digits from the treatment or response email address per subject. The 4 remaining unmatched response emails differed from their treatment emails in non-systemic ways such as using a nickname for the "@ischool.berkeley.edu" address and not the "@berkeley.edu" address or vice versa. Using the respondent's name and UC Berkeley's internal email address book, we were able to manually match these respondent's email addresses to their treatment email addresses.

5: Analysis

The data collected has columns representing variables of interest in our analysis, as summarized by Table 5.1. The outcome being measured in our analysis is the response rate of individuals. This is accomplished by using the `response` variable column since the mean of this binary variable gives the percent proportion of responses across a group. Treatment conditions are recorded by the `treatment_group_name` variable. The control group maps to individuals who did not receive any incentive for taking the survey, and have their `treatment_group_name` value as "No Incentive (Control)". The first treatment group, which incentivized survey responders with a $5 reward, maps to individuals who have their `treatment_group_name` value as “$5 Incentive”. The second treatment group, which incentivized survey responders with a chance to win $100 through a lottery, maps to individuals who have their `treatment_group_name` value as “$100 Lottery”.

Column Name	Description
full_name	Potential Survey Taker’s full name. Ignored in analysis.
email	Potential Survey Taker’s email on record. Ignored in analysis.
in_PST	Binary 1/0 if the individual is located in the Pacific time zone.
response	Binary 1/0 if the individual completed their survey.
treatment_group_name	Categorical Variable indicating which type of incentive treatment the individual received:; No Incentive (Control), $5 Incentive, $100 Lottery respectively.
cohort_name	Categorical variable indicating which MIDS starting semester cohort the individual belongs to. Represents how new the student is to this graduate program and where they are in the process of getting their MIDS degree. Values are: “Summer 2020”, “Fall 2020”, “Spring 2021”, “Summer 2021”, “Fall 2021”.

Table 5.1 A descriptive summary of variables in our data, represented by columns.

Section 5.1: Covariates

The important covariates used in our analysis are `cohort_name` and `in_PST`. We used `cohort_name` to conduct randomized blocking and ensured the five cohorts analyzed are equally represented across control and both treatments, with 24 randomly selected individuals from each cohort being included in each of our three treatment groups. Table 5.2 confirms that the size of each cohort is equal, regardless of which treatment was assigned.

Group Distribution and Response Rate by Cohort and Treatment Group:

Cohort	Treatment Group	Size	Response (N)	Response Rate (%)
Summer 2020	No Incentive (Control)	24	2	8.33
Summer 2020	$5 Incentive	24	6	25.00
Summer 2020	$100 Lottery	24	1	4.17
Fall 2020	No Incentive (Control)	24	0	0.00
Fall 2020	$5 Incentive	24	4	16.67
Fall 2020	$100 Lottery	24	4	16.67
Spring 2021	No Incentive (Control)	24	5	20.83
Spring 2021	$5 Incentive	24	3	12.50
Spring 2021	$100 Lottery	24	1	4.17
Summer 2021	No Incentive (Control)	24	4	16.67
Summer 2021	$5 Incentive	24	4	16.67
Summer 2021	$100 Lottery	24	0	0.00
Fall 2021	No Incentive (Control)	24	5	20.83
Fall 2021	$5 Incentive	24	8	33.33
Fall 2021	$100 Lottery	24	5	20.83

Note. Covariate analysis - checking for balance of cohorts across treatment groups.

Table 5.2 Group Distribution and Response Rate by Cohort and Treatment Group.

We control for students being in the Pacific Timezone by measuring `in_PST` and including it as a covariate in our analysis. Since we did not block by `in_PST`, the proportion of PST students may differ based on treatment group assignment if our randomization was poor. As a result, we examine how PST students are dispersed with treatment. If there is no relationship between assignment to a treatment group and the proportion of PST students in that group, then the covariate is balanced. Table 5.3 reveals only differences in the proportion of PST students in the “No Incentive (Control)” and “$5 Incentive” groups.

Group Distribution by Treatment and In_PST:

Treatment Group	In PST	Size	Response (N)	Response Rate (%)
No Incentive (Control)	1	54	8	14.81
No Incentive (Control)	0	66	8	12.12
$5 Incentive	0	68	15	22.06
$5 Incentive	1	52	10	19.23
$100 Lottery	1	54	8	14.81
$100 Lottery	0	66	3	4.55

Table 5.3: Group Distribution by Treatment and In_PST.

A covariate balance check on the statistical significance of the difference in proportion of PST students between control and $5 incentive is conducted by regressing `in_PST` on `treatment_group_name`, for this subset of data. The result of this covariate balance check is presented below in Table 5.4.

Covariate Balance Check on PST: No Incentive vs. 5 Per Survey Incentive

	Dependent variable:
	in_PST
5 Dollar Incentive	-0.017
	(0.065)
Baseline	0.450^{***}
	(0.046)
Observations	240
R^2	0.0003
Adjusted R^2	-0.004
Residual Std. Error	0.499 (df = 238)
F Statistic	0.067 (df = 1; 238)
Note:	*p<0.1; **p<0.05; ***p<0.01

Table 5.4: Covariate balance check on students in PST and $5 Incentive assignment.

The covariate balance check of `in_PST` on `treatment_group_name` in Table X.4 reveals that this difference between $5 Incentive and No Incentive groups is not statistically significant. Thus, `in_PST` can be assumed to be a balanced covariate.

Section 5.2: Exploratory Data Analysis

Visualizing the results of Table 5.2 in Fig 5.1 shows the heterogeneous treatment effects of incentives by cohort.

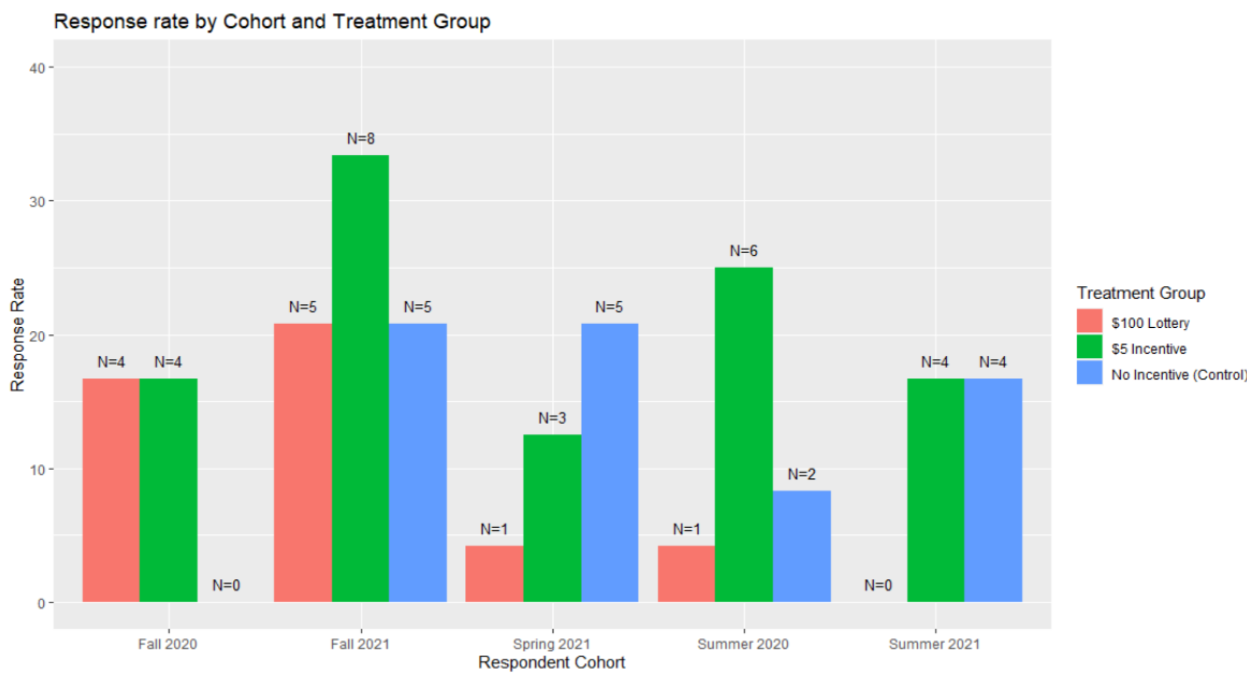


Fig 5.1: Response rate by cohort and treatment group assignment. Response rate is given as a percentage proportion and the number of responses is also written above for each cohort and treatment.

From Fig 5.1, we observe that there are more responses overall in the Fall 2021 cohort. This indicates that the newest students to the MIDS program may be more likely to answer surveys in their email. The response rate for the $100 Lottery Incentive treatment is the lowest among the 3 treatment conditions for every cohort except Fall 2020. The response rate for the $5 Incentive treatment yields the highest response rate in the Fall 2021 and Summer 2020 cohorts, is tied for highest response rate in two cohorts, and is second to the No Incentive control group for Spring 2020. Providing no incentive for survey completion only showed lower response rates than other treatment conditions for Fall 2020, possible due to more graduates being in that cohort and less motivation to give feedback on MIDS post-graduation without an incentive. It is important to note that the heterogeneous treatment effects portrayed in Fig 5.1 do not address if the differences are statistically significant. Given that the size of each subgroup portrayed in Fig 5.1 is only 26, the regression conducted to quantify the estimated heterogeneous treatment effects and their standard errors will most likely be severely underpowered due to low sample size. As a result, we leave the analysis of heterogeneous treatment effects of incentives on cohort assignment to a future study that accounts for this sample size problem.

Visualizing the results of Table 5.3 in Fig 5.2 shows the heterogeneous treatment effects of incentives by being in the Pacific Timezone..

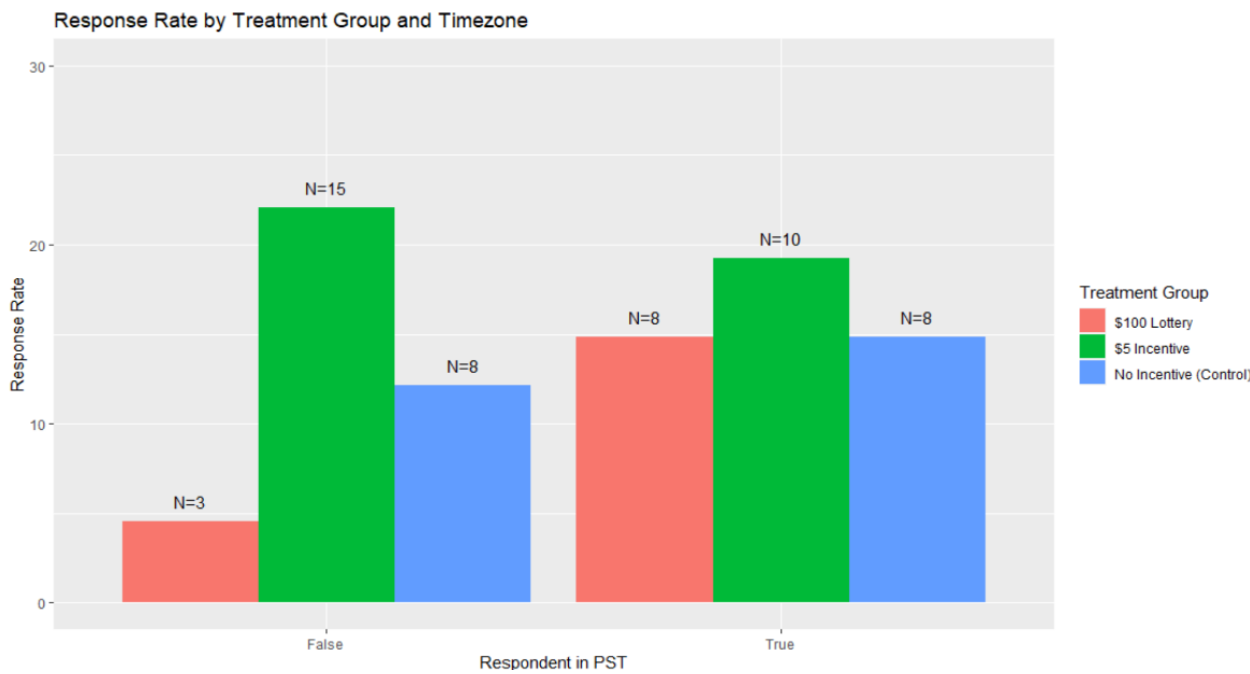


Fig 5.2: Response rate by Pacific Timezone indicator and treatment group assignment. Response rate is given as a percentage proportion and the number of responses is also written above for each in_PST condition and treatment.

From Fig 5.2, we observe that there are equal responses between students in PST and not in PST; both groups had 26 replies. However, it seems that students in PST's response rate did not change very much based on which incentive treatment they were assigned to. For students not in PST, the $5 Incentive yielded a greater response rate than both control, no incentive given, and the $100 lottery assignment. Given that the size of each subgroup portrayed in Fig 5.2 is between 50-70 samples, a sub-group analysis can be used to estimate heterogeneous treatment effects for in_PST and their standard errors.

Section 5.3: Summary Response Rate by Treatment Group

Table 5.5 summarizes the response rate for each treatment group. Each treatment group includes 120 participants, and the response rates for "No Incentive", for "$5 Incentive" and for "$100 Lottery" groups were 13.33%, 20.83%, and 9.17% respectively.

Table 1: Response Rate by Treatment.

Treatment	Group Size	Response (N)	Response (%)
No Incentive (Control)	120	16	13.33
$5 Incentive	120	25	20.83
$100 Lottery	120	11	9.17

Note. Summary response rate by treatment.

Table 5.5: Summary table giving the group size, response number, and response rate by the treatment condition.

Section 5.4: Statistical Power Calculation - Post Study Implementation

Given the observed response rates (see Table 5.5) in the experiment, using a power simulation comparing the “No Incentive” and “$5 Incentive” groups, the power for this experiment (120 participants per group) is around 32.8%. To achieve a power of 80%, it requires groups of around 395 participants.

In addition, using a power simulation comparing the “$5 Incentive” and “$100 Lottery” groups, given the observed response rates (see Table 5.5), the power for this experiment (120 participants per group) is around 73.7%. To achieve a power of 80%, it requires groups of around 145 participants.

Figure 5.3’s Graph 2 and 3 (below) present the power calculation curve for “No Incentive” versus “$5 Incentive” groups, and for “$5 Incentive” versus “$100 Lottery” groups respectively. The red line indicates the power for the current experiment given the observed response rates for group sizes of 120 participants.

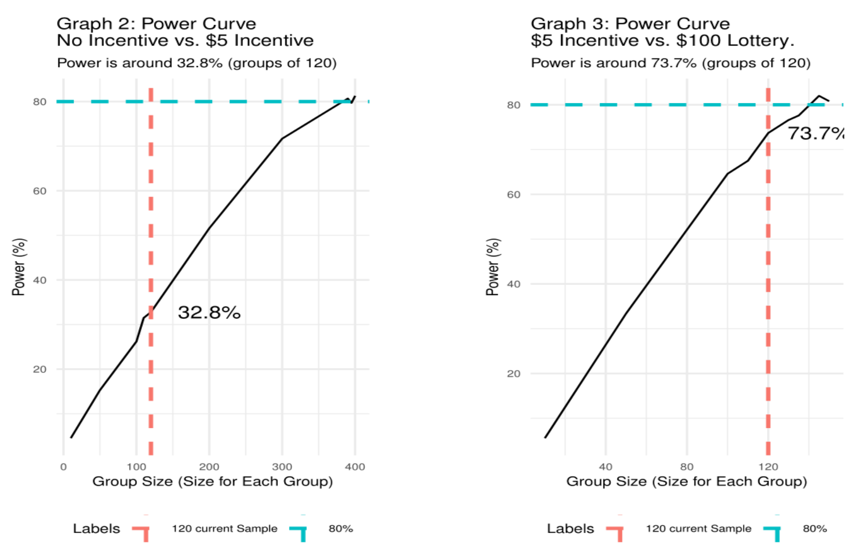


Fig 5.3: Power curves for the No Incentive vs $5 Incentive experiment (left) and the $5 Incentive vs. $100 Lottery Incentive (right). The blue line represents 80% power, the red line’s intersection with the curve gives our actual experimental power.

Section 5.5: T-Test Analysis

Based on a predefined statistical plan, a t-test analysis was performed to compare the following response rates:

1. “$5 Incentive” group (20.83%) vs. “No Incentive” group (13.33%).
2. “$100 Lottery” group (9.17%) vs. “No Incentive” group (13.33%).
3. “$100 Lottery” group (9.17%)vs. “$5 Incentive” group (20.83%).

The table below presents a summary of t-test results using un-adjusted α value of 0.05. In summary:

1. With a p-value of 0.124 (> 0.05), we fail to reject the null hypothesis of the average response rate being equal. Thus, there is not enough evidence supporting that the average response rate for “$5 Incentive” (20.83%) is statistically different from the average response rate for “No Incentive” (13.33%), at a 95% confidence.
2. With a p-value of 0.309 (> 0.05), we fail to reject the null hypothesis of the average response rate being equal. Thus, there is not enough evidence supporting that the average response rates for “$100 Lottery” (9.17%) is statistically different from the average response rates for “No Incentive” (13.33%), at a 95% confidence.
3. With a p-value of 0.011 (< 0.05), we reject the null hypothesis of the average response rate being equal. Thus, there is evidence suggesting the average response rates for “$100 Lottery” (9.17%) is statistically different from the average response rates for “$5 Incentive” (20.83%), at a 95% confidence.

Table 2: T-Test Summary Table:

Comparison	P-Value	CI_Lower	CI_Upper	Mean Diff	Mean Grp 1	Mean Grp 2
No Treatment vs. $5 Incentive	0.12	-2.07	17.07	7.50	13.33	20.83
No Treatment vs. $100 Lottery	0.31	-12.22	3.89	-4.17	13.33	9.17
$5 Incentive vs. $100 Lottery	0.01	-20.67	-2.66	-11.67	20.83	9.17

Note. Mean and Confidence interval in percentage (%). Uses Robust Standard Errors.

Table 5.6: Summary table of the T-Tests between mean response rates given treatment conditions and their lower and upper confidence intervals.

Given that three statistical tests were performed, we also evaluated the t-test results using an adjusted α of 0.017 (α=0.05/3 - Bonferroni Adjustment). Using the adjusted α, the above statistical conclusions are still valid. In the case of the third comparison, With a p-value of 0.011 (< 0.017), we reject the null hypothesis of the average response rate being equal. Thus, there is evidence suggesting the average response rates for “$100 Lottery” (9.17%) is statistically different from the average response rates for “$5 Incentive” (20.83%).

Though we fail to reject the null hypothesis of “No Incentive” response rate being equal to “$5 Incentive” response rate, the magnitude of the response rate difference is numerically large (7.5%), and the power analysis suggests the study may be underpowered to detect significant differences. Our power calculation indicates that a group size of 395 is needed to achieve 80% power, given the response rates observed.

The following graph presents the confidence interval for the differences in response rates for the t-test analysis. As it can be seen the confidence interval for the difference in average response rate between “$5 Incentive” and “$100 Lottery” does not include the value zero which aligns with the conclusion that the average response rates are statistically different. On the other hand, the other two confidence intervals include the value zero, indicating that the difference in average response rate may not be statistically different.

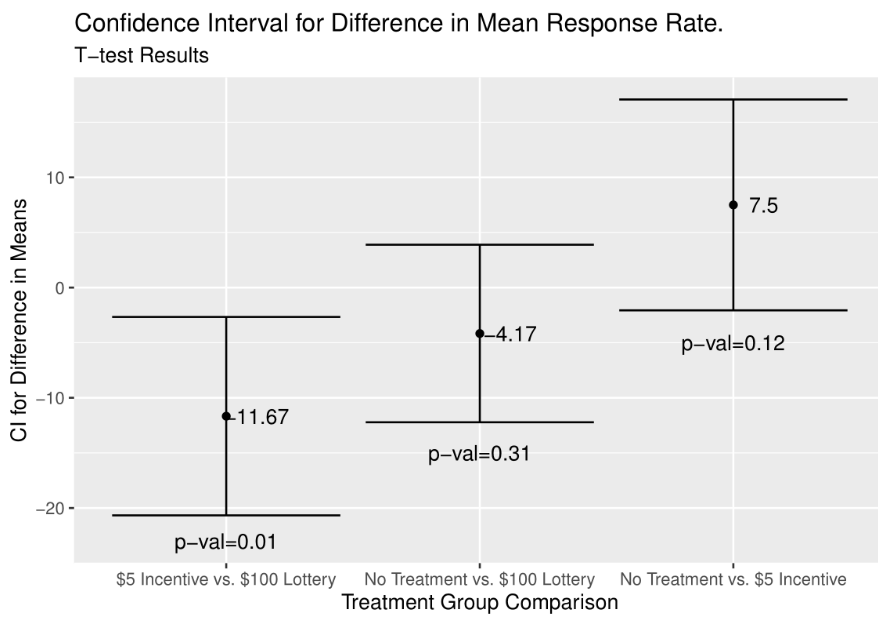


Fig 5.4: Confidence Intervals for Differences in Mean Response Rate.

In summary, the t-test analysis suggests that survey participants prefer to have an immediate incentive (in this case $5 incentive) to not having an incentive or being entered into a $100 lottery. The results were not statistically different when comparing “$5 incentive” and “No Incentive”, but the magnitude of the difference suggests that the average difference is substantial, and the study may have been underpowered to detect a statistically significant difference.

Section 5.6: Regression Analysis

Based on a predefined statistical plan, the following Linear Regression Analysis was performed:

- Comparison of “No Incentive” vs. “$5 Incentive” Linear Regression.
- Comparison of “No Incentive” vs. “$100 Lottery” Linear Regression.

- Comparison of “$5 Incentive”” vs. “$100 Lottery” Linear Regression.
- All Treatment (Incentives) Summary Linear Regression.
- Heterogeneous Treatment Effect Analysis Linear Regression.

The following sections discuss results from this linear regression analysis. Note that our estimated treatment effect is the Intent to Treat (ITT) effect, rather than the Average Treatment Effect (ATE), since we cannot confirm that every email sent was viewed and rejected when we do not get a response. There is a possibility that some individuals who did not respond to the survey might have missed the email in their inbox during the study’s duration.

5.6.1: No Incentive vs. $5 Incentive

In this section, we evaluate the impact on Average Response Rate of having "$5 Incentive" versus "No Incentive". The table below presents the results from the different linear regressions comparing these treatments, using “No Incentive” as the reference level (baseline) for treatment. The linear regressions were developed using Robust Standard Errors. The four linear regression models are as follows:

- **Model 1:** Average Response Rate (ARR) on Treatment ($5).

$$ARR = b_0 + b_1 * Treatment_5$$

- **Model 2:** Average Response Rate (ARR) on Treatment ($5) with Cohort as a covariate.

$$ARR = b_0 + b_1 * Treat_5 + b_2 * Fall2020 + b_3 * Fall2021 + b_4 * Spring2021 + b_5 * Summer2021$$

- **Model 3:** Average Response Rate (ARR) on Treatment ($5) with PST (Pacific Time) as a covariate.

$$ARR = b_0 + b_1 * Treatment_5 + b_2 * InPST$$

- **Model 4:** Average Response Rate (ARR) on Treatment ($5) with Cohort and PST (Pacific Time) as covariates.

$$ARR = b_0 + b_1 * Treat_5 + b_2 * Fall2020 + b_3 * Fall2021 + b_4 * Spring2021 + b_5 * Summer2021 + b_6 * InPST$$

Table 1: Linear Regression - No Incentive vs. 5 Per Survey Incentive

	Dependent variable:			
	Response Rate			
	Simple	Cohort Included	PST included	Cohort and PST included
	(1)	(2)	(3)	(4)
5 Dollar Incentive	0.075	0.075	0.075	0.075
	(0.049)	(0.049)	(0.049)	(0.049)
Fall 2020		-0.083		-0.083
		(0.067)		(0.068)
Fall 2021		0.104		0.104
		(0.085)		(0.085)
Spring 2021		0.000		-0.0002
		(0.078)		(0.078)
Summer 2021		0.000		-0.0003
		(0.078)		(0.078)
In PST			-0.001	0.002
			(0.049)	(0.049)
Baseline	0.133^{***}	0.129^{**}	0.134^{***}	0.129^{**}
	(0.031)	(0.055)	(0.037)	(0.057)
Cohort fixed effects	No	Yes	No	Yes
PST fixed effects	No	No	Yes	Yes
Baseline is:	No Incentive (Control)	No Incentive (Control)	No Incentive (Control)	No Incentive (Control)
Observations	240	240	240	240
R^2	0.010	0.035	0.010	0.035
Adjusted R^2	0.006	0.014	0.002	0.010
Residual Std. Error	0.376 (df = 238)	0.374 (df = 234)	0.377 (df = 237)	0.375 (df = 233)
F Statistic	2.386 (df = 1; 238)	1.694 (df = 5; 234)	1.188 (df = 2; 237)	1.406 (df = 6; 233)
Note:				*p<0.1; **p<0.05; ***p<0.01
				Note: Uses Robust Standard Errors.

Table 5.7: Linear regression analysis to estimate No Incentive vs $5 Guaranteed Incentive's treatment effect on Average Response Rate.

From model 1, we conclude that the “No Incentive” treatment group had an Average Response Rate (ARR) of 13.33%, and the “$5 Incentive” treatment group had an increase in Average Response Rate, versus “No Incentive”, of 7.5%. Using Robust Standard Errors, with a p-value of 0.124 (>0.05), model 1 suggests that the increase in response rate for “$5 Incentive” is not statistically significant. Thus we fail to reject the null hypothesis of “$5 Incentive” coefficient being equal to zero. In other words, even though the magnitude of the increase in ARR is relatively large for “$5 Incentive” versus “No Incentive”, it is not statistically different from zero.

As covariates for Cohort were added in model 2, the estimated increase in ARR for “$5 Incentive” versus “No Incentive” is maintained at 7.5%. This indicates that the random block assignment using the different cohorts was successful in breaking any potential correlation between cohort and ARR. In addition, with a p-value of 0.122 (>0.05), this coefficient is not statistically different from zero. However, the addition of the cohort as covariate did not help in improving the precision for the coefficient, as the coefficient's Robust Standard Error is the same as in model 1.

In model 3, a covariate of Pacific Time (PST) was added to model 1. Similarly to what was observed in model 1 and 2, the estimated increase in ARR for “$5 Incentive” versus “No Incentive” is maintained at 7.5%. Though we did not perform a random block assignment using PST, this covariate may not be correlated with ARR, as the estimate for ARR for “$5 Incentive” is maintained. As in previous models, with a p-value of 0.125 (>0.05), this coefficient is not statistically different from zero, and the inclusion of PST as covariate did not improve the precision of the ARR coefficient, as the coefficient’s Robust Standard Error is the same as in model 1 and 2.

Finally, in model 4, covariates for Cohort and for Pacific Time (PST) were added. Similarly to what was observed in model 1, 2 and 3, the estimated increase in ARR for “$5 Incentive” versus “No Incentive” is maintained at 7.5%. As in previous models, with a p-value of 0.123 (>0.05), this coefficient is not statistically different from zero, and the inclusion of these covariates did not improve the precision of the ARR coefficient.

In summary, the increase in ARR for “$5 Incentive” is estimated at 7.5%. Even though this increase is relatively large when compared to the ARR for “No Incentive”, the coefficient is not statistically different from zero (using Robust Standard Errors), as p-value for all models is greater than 0.05. The treatment effect is equal in all models as covariate for Cohort and Pacific Time (PST) are added, indicating that the random assignment of subjects was successful in breaking any potential correlation of the covariates and the ARR. However, none of the covariates helped to increase precision as the coefficient’s Robust Standard Error for the main treatment effect did not change across the models. Overall, this analysis demonstrates that, for increasing survey response rate, a $5 guaranteed incentive is not significantly better than giving no incentive.

5.6.2: No Incentive vs. $100 Lottery Incentive

In this section, we evaluate the impact on Average Response Rate of entering a $100 lottery for answering the survey (“$100 Lottery”) versus “No Incentive”. The table below presents the results from the different linear regressions comparing these treatments, using “No Incentive” as the reference level (baseline) for treatment. The linear regressions were developed using Robust Standard Errors. The four linear regression models are as follows:

- **Model 1:** Average Response Rate (ARR) on Treatment (Lottery).

$$ARR = b_0 + b_1 * Treatment_L$$

- **Model 2:** Average Response Rate (ARR) on Treatment (Lottery) plus Cohort as covariate.

$$ARR = b_0 + b_1 * Treat_L + b_2 * Fall2020 + b_3 * Fall2021 + b_4 * Spring2021 + b_5 * Summer2021$$

- **Model 3:** Average Response Rate (ARR) on Treatment (Lottery) plus PST (Pacific Time) as covariate.

$$ARR = b_0 + b_1 * Treatment_L + b_2 * InPST$$

- **Model 4:** Average Response Rate (ARR) on Treatment (Lottery) plus Cohort and PST (Pacific Time) as covariate.

$$ARR = b_0 + b_1 * Treat_L + b_2 * Fall2020 + b_3 * Fall2021 + b_4 * Spring2021 + b_5 * Summer2021 + b_6 * InPST$$

Table 2: Linear Regression - No Incentive vs. 100 Lottery Win

	Dependent variable:			
	Response Rate			
	Simple	Cohort Included	PST included	Cohort and PST Included
	(1)	(2)	(3)	(4)
100 Dollar Lottery Incentive	-0.042	-0.042	-0.042	-0.042
	(0.041)	(0.041)	(0.041)	(0.041)
Fall 2020		0.021		0.021
		(0.055)		(0.055)
Fall 2021		0.146^{**}		0.147^{**}
		(0.070)		(0.070)
Spring 2021		0.062		0.053
		(0.060)		(0.059)
Summer 2021		0.021		0.018
		(0.054)		(0.054)
In PST			0.065	0.067
			(0.042)	(0.042)
Baseline	0.133^{***}	0.083^{*}	0.104^{***}	0.055
	(0.031)	(0.043)	(0.036)	(0.049)
Cohort fixed effects	No	Yes	No	Yes
PST fixed effects	No	No	Yes	Yes
Baseline is:	No Incentive (Control)	No Incentive (Control)	No Incentive (Control)	No Incentive (Control)
Observations	240	240	240	240
R^2	0.004	0.031	0.015	0.042
Adjusted R^2	0.0002	0.011	0.006	0.018
Residual Std. Error	0.317 (df = 238)	0.315 (df = 234)	0.316 (df = 237)	0.314 (df = 233)
F Statistic	1.039 (df = 1; 238)	1.521 (df = 5; 234)	1.775 (df = 2; 237)	1.720 (df = 6; 233)
Note:				*p<0.1; **p<0.05; ***p<0.01
				Note: Uses Robust Standard Error.

Table 5.8: Linear regression analysis to estimate No Incentive vs $100 Lottery Entry Incentive treatment effect on Average Response Rate.

From model 1, we conclude that the “No Incentive” treatment group had an Average Response Rate (ARR) of 13.33%, and the “$100 Lottery” treatment group had a decrease in Average Response Rate, versus “No Treatment”, of -4.17%. Using Robust Standard Errors, with a p-value of 0.309 (>0.05), model 1 suggests that the reduced ARR for “$100 Lottery” is not statistically significant. Thus we fail to reject the null hypothesis of the “$100 Lottery” coefficient being equal to zero. Thus, the decrease in ARR for “$100 Lottery” versus “No Incentive” is not statistically different from zero. Offering participants the opportunity to win $100 for completing the survey did not significantly change the response rate from being offered No Incentive for completing the survey.

As covariates for Cohort were added in model 2, the estimated decreases in ARR for "$100 Lottery" versus "No Incentive" is maintained at -4.17%, which indicates that the random block assignment using the different cohorts was successful in breaking any potential correlation between cohort and ARR. In addition, with a p-value of 0.307 (>0.05), this coefficient is not statistically different from zero, and the addition of the cohort as covariate did not help in improving the precision for the coefficient, as the coefficient's Robust Standard Error is the same as in model 1.

In model 3, a covariate of Pacific Time (PST) was added to model 1. Similarly to what was observed in model 1 and 2, the estimated decrease in ARR for "$100 Lottery" versus "No Incentive" is maintained at -4.17%. Though we did not perform a random block assignment using PST, this covariate may not be correlated with ARR, as the estimate for ARR for "$100 Lottery" is maintained. As in previous models, with a p-value of 0.308 (>0.05), this coefficient is not statistically different from zero, and the inclusion of PST as covariate did not improve the precision of the ARR coefficient.

Finally, in model 4, covariates for Cohort and for Pacific Time (PST) were added. Similarly to what was observed in model 1, 2 and 3, the estimated decrease in ARR for "$100 Lottery" versus "No Incentive" is maintained at -4.17%. As in previous models, with a p-value of 0.305 (>0.05), this coefficient is not statistically different from zero, and the inclusion of these covariates did not improve the precision of the ARR coefficient.

In summary, the decrease in ARR for "$100 Lottery" is estimated at -4.17%. The coefficient is not statistically different from zero (using Robust Standard Errors), as p-value for all models is greater than 0.05. The treatment effect is equal in all models as covariate for Cohort and Pacific Time (PST) are added, indicating that the random assignment of subjects was successful in breaking any potential correlation of the covariates and the ARR. However, none of the covariates helped to increase precision as the coefficient's Robust Standard Error for the main treatment effect did not change across the models. Overall, this analysis demonstrates that, for increasing survey response rate, a $100 lottery entry incentive is not significantly different than giving no incentive.

5.6.3: $5 Incentive vs $100 Lottery Incentive

In this session, we evaluate the impact on Average Response Rate of offering a $5 Incentive for answering the survey ("$5 Incentive") versus entering participants in a $100 Lottery for answering the survey ("$100 Lottery"). The table below presents the results from the different linear regressions comparing these treatments, using "$5 Incentive" as the reference level (baseline) for treatment. The linear regressions were developed using Robust Standard Errors. The four linear regression models are as follows:

- **Model 1:** Average Response Rate (ARR) on Treatment (Lottery).

$$ARR = b_0 + b_1 * Treatment_L$$

- **Model 2:** Average Response Rate (ARR) on Treatment (Lottery) plus Cohort as covariate.

$$ARR = b_0 + b_1 * Treat_L + b_2 * Fall2020 + b_3 * Fall2021 + b_4 * Spring2021 + b_5 * Summer2021$$

- **Model 3:** Average Response Rate (ARR) on Treatment (Lottery) plus PST (Pacific Time) as covariate.

$$ARR = b_0 + b_1 * Treatment_L + b_2 * InPST$$

- **Model 4:** Average Response Rate (ARR) on Treatment (Lottery) plus Cohort and PST (Pacific Time) as covariate.

$$ARR = b_0 + b_1 * Treat_L + b_2 * Fall2020 + b_3 * Fall2021 + b_4 * Spring2021 + b_5 * Summer2021 + b_6 * InPST$$

Table 3: Linear Regression - 5 Per Survey Incentive vs. 100 Lottery Win

	Dependent variable:			
	Response Rate			
	Simple	Cohort Included	PST included	Cohort and PST Included
	(1)	(2)	(3)	(4)
100 Dollar Lottery Incentive	-0.117^{**}	-0.117^{**}	-0.117^{**}	-0.117^{**}
	(0.046)	(0.046)	(0.046)	(0.046)
Fall 2020		0.021		0.022
		(0.075)		(0.076)
Fall 2021		0.125		0.122
		(0.082)		(0.082)
Spring 2021		-0.062		-0.065
		(0.065)		(0.065)
Summer 2021		-0.062		-0.068
		(0.064)		(0.063)
In PST			0.037	0.042
			(0.047)	(0.046)
Baseline	0.208^{***}	0.204^{***}	0.192^{***}	0.188^{***}
	(0.037)	(0.061)	(0.043)	(0.066)
Cohort fixed effects	No	Yes	No	Yes
PST fixed effects	No	No	Yes	Yes
Baseline is:	5 dollar Incentive	5 dollar Incentive	5 dollar Incentive	5 dollar Incentive
Observations	240	240	240	240
R^2	0.027	0.064	0.029	0.067
Adjusted R^2	0.023	0.044	0.021	0.043
Residual Std. Error	0.354 (df = 238)	0.350 (df = 234)	0.354 (df = 237)	0.350 (df = 233)
F Statistic	6.526^{**} (df = 1; 238)	3.200^{***} (df = 5; 234)	3.590^{**} (df = 2; 237)	2.807^{**} (df = 6; 233)
Note:				*p<0.1; **p<0.05; ***p<0.01
				Note: Uses Robust Standard Error.

Table 5.9: Linear regression analysis to estimate $5 Guaranteed Incentive vs $100 Lottery Entry Incentive treatment effect on Average Response Rate.

From model 1, we conclude that “$5 Incentive” treatment group had an Average Response Rate (ARR) of 20.83%, and “$100 Lottery” treatment group had a decrease in Average Response Rate, versus “$5 Incentive”, of -11.67%. Using Robust Standard Errors, with a p-value of 0.011 (<0.05), model 1 suggests that the reduced ARR for “$100 Lottery” is statistically significant. Thus we reject the null hypothesis of the “$100 Lottery” coefficient being equal to zero. As we performed three statisticals tests, we also evaluated the results using an adjusted α of 0.017 (α = 0.05/3, Bonferroni Adjustment). Using the adjusted α, with a p-value of 0.011 (<0.017), we reject the null hypothesis of “$100 Lottery” coefficient being equal to zero. Thus, the decrease in ARR for “$100 Lottery” versus “$5 Incentive” is statistically different from zero. Offering a $5 Guaranteed Incentive instead of a $100 Lottery Entry Incentive will increase the survey response rate.

As covariates for Cohort were added in model 2, the estimated decreases in ARR for "$100 Lottery" versus "$5 Incentive" is maintained at -11.67%, which indicates that the random block assignment using the different cohorts was successful in breaking any potential correlation between cohort and ARR. In addition, with a p-value of 0.01 (>0.05), this coefficient is statistically different from zero, and the addition of the cohort as covariate did not help in improving the precision for the coefficient, as the coefficient's Robust Standard Error is the same as in model 1. Using the adjusted α (0.017, Bonferroni Adjustment), with a p-value of 0.01 (<0.017), we reject the null hypothesis of "$100 Lottery" coefficient being equal to zero. Thus, the decrease in ARR for "$100 Lottery" versus "$5 Incentive" is statistically different from zero (using adjusted α).

In model 3, a covariate of Pacific Time (PST) was added to model 1. Similarly to what was observed in model 1 and 2, the estimated decrease in ARR for "$100 Lottery" versus "No Incentive" is maintained at -11.73%. Though we did not perform a random block assignment using PST, this covariate may not be correlated with ARR, as the estimate for ARR for "$100 Lottery" is maintained. As in previous models, with a p-value of 0.011 (>0.05), this coefficient is statistically different from zero, and the inclusion of PST as covariate did not improve the precision of the ARR coefficient. This result holds when using adjusted α (0.017, Bonferroni Adjustment).

Finally, in model 4, covariates for Cohort and for Pacific Time (PST) were added. Similarly to what was observed in model 1, 2 and 3, the estimated decrease in ARR for "$100 Lottery" versus "No Incentive" is maintained at -11.74%. As in previous models, with a p-value of 0.01 (>0.05), this coefficient is statistically different from zero, and the inclusion of these covariates did not improve the precision of the ARR coefficient. This result also holds when using adjusted α (0.017, Bonferroni Adjustment).

In summary, the decrease in ARR for "$100 Lottery" is estimated at -11.67%. The coefficient is statistically different from zero (using Robust Standard Errors), as p-value for all models is greater than 0.05, and this result holds when using adjusted α (0.017, Bonferroni Adjustment). The treatment effect is equal in all models as covariate for Cohort and Pacific Time (PST) are added, indicating that the random assignment of subjects was successful in breaking any potential correlation of the covariates and the ARR. However, none of the covariates helped to increase precision as the coefficient's Robust Standard Error for the main treatment effect did not change across the models. The results of this analysis show that people are more likely to respond to a survey when offered guaranteed compensation than when offered a higher compensation amount in expectation. Perhaps a lottery's chance for zero compensation for 15 minutes of work outweighs the potentially larger compensation awarded if chosen, so fewer people responded to the survey than those who were offered a relatively small $5 compensation for their efforts.

5.6.4: Incentive Summary Table

The table below summarizes the regression results of the previous three tables by regressing the ARR on the three treatment conditions: “$100 Lottery Incentive", “No Incentive”, and “$5 Incentive”. The table uses “$5 Incentive” as the reference level (baseline) for treatment, so the coefficients of “100 Dollar Lottery Incentive” and “No Incentive” quantify how much the response rate would change if this incentive was offered instead of $5 guaranteed. The linear regressions were developed using Robust Standard Errors. The table includes 4 regressions corresponding, again, to the inclusion of Cohort, if a student was in PST, and both Cohort and if a student was in PST as covariates to increase the precision of our estimated treatment effects.

Table 4: Linear Regression - Including All Treatments

	Dependent variable:			
	Response Rate			
	Simple	Cohort Included	PST Included	Cohort and PST Included
	(1)	(2)	(3)	(4)
100 Dollar Lottery Incentive	-0.117^{**}	-0.117^{**}	-0.117^{**}	-0.117^{**}
	(0.046)	(0.046)	(0.046)	(0.046)
No Incentive	-0.075	-0.075	-0.076	-0.076
	(0.049)	(0.049)	(0.049)	(0.049)
Fall 2020		-0.014		-0.014
		(0.054)		(0.054)
Fall 2021		0.125^{*}		0.124^{*}
		(0.064)		(0.065)
Spring 2021		-0.000		-0.004
		(0.055)		(0.055)
Summer 2021		-0.014		-0.019
		(0.054)		(0.053)
In PST			0.034	0.037
			(0.038)	(0.037)
Baseline	0.208^{***}	0.189^{***}	0.194^{***}	0.175^{***}
	(0.037)	(0.053)	(0.041)	(0.056)
Cohort fixed effects	No	Yes	No	Yes
PST fixed effects	No	No	Yes	Yes
Baseline is:	5 dollar Incentive	5 dollar Incentive	5 dollar Incentive	5 dollar Incentive
Observations	360	360	360	360
R^2	0.019	0.042	0.021	0.044
Adjusted R^2	0.013	0.025	0.013	0.025
Residual Std. Error	0.350 (df = 357)	0.348 (df = 353)	0.350 (df = 356)	0.348 (df = 352)
F Statistic	3.431^{**} (df = 2; 357)	2.561^{**} (df = 6; 353)	2.565^{*} (df = 3; 356)	2.341^{**} (df = 7; 352)
Note:				$^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$
				Note: Uses Robust Standard Error.

Table 5.10: Summary Regression Table comparing the treatment effect of "$100 Lottery Incentive", "No Incentive", and "$5 Incentive" on the Average Response Rate.

From the table above, we see that offering a $100 Lottery Entry Incentive instead of $5 guaranteed incentive leads to a statistically significant decrease in response rate of -11.7%, with a standard error of 4.6%. This coefficient and robust standard error is consistent across all regressions in Table 5.10 and matches our earlier results in Table 5.9. If a survey designer must compensate survey responders, and their options are a small guaranteed reward versus using a large reward lottery entry, our study supports the use of guaranteed incentives since they yield a significant increase in survey response rate over lottery incentives.

We also see that offering no Incentive instead of $5 guaranteed incentive leads to a statistically insignificant decrease in response rate of -7.5%, with a standard error of 4.9%. This coefficient and robust standard error is consistent across all regressions in Table 5.10 and matches our earlier results in Table 5.7. If a survey designer does not need to compensate survey responders, and their primary goal is to have a high response rate, our study does not support that offering even guaranteed incentives significantly improves response rate over no incentive. Our 95% confidence interval for the change in survey response for offering no incentive instead of a $5 incentive is -17.3% to 2.3%. Offering no incentive may decrease survey responses by up to 17.3%, but it could also increase survey responses by 2.3%. If a survey designer's goal is to increase survey response by all means, they can take their chances and provide a guaranteed $5 incentive. However, the data does not support that an increase in response rate is statistically significant, so a money conscious survey designer might be motivated to provide no incentive without significant change in survey response rate.

5.6.5: Heterogeneous Treatment Effect on being in PST

In Fig 5.2, we observe that students in PST's response rate did not change very much based on which incentive treatment they were assigned to. For students not in PST, the $5 Incentive yielded a greater response rate than both no incentive and the $100 lottery incentive. In Table 5.11, a sub-group analysis is used to estimate heterogeneous treatment effects for in_PST and their robust standard errors. Columns 1 and 2, for students outside of PST and in PST respectively, provide the heterogeneous treatment effects of $100 Lottery Incentive and No Incentive over the reference level ($5 Incentive Baseline). Columns 3 and 4, for students outside of PST and in PST respectively, estimate the same heterogeneous treatment effects but more precisely through the use of Cohort as covariates.

Table 5: Linear Regression - Heterogeneous Treatment Effects for being in PST

	Dependent variable:			
	Response Rate			
	Not PST	PST	Not PST, Cohorts	PST, Cohorts
	(1)	(2)	(3)	(4)
100 Dollar Lottery Incentive	-0.175^{***}	-0.044	-0.172^{***}	-0.040
	(0.057)	(0.074)	(0.057)	(0.074)
No Incentive	-0.099	-0.044	-0.103	-0.031
	(0.065)	(0.074)	(0.065)	(0.073)
Fall 2020			-0.015	-0.005
			(0.071)	(0.088)
Fall 2021			0.084	0.190^{*}
			(0.082)	(0.107)
Spring 2021			-0.030	0.033
			(0.074)	(0.085)
Summer 2021			-0.061	0.050
			(0.067)	(0.088)
Baseline	0.221^{***}	0.192^{***}	0.222^{***}	0.133
	(0.051)	(0.056)	(0.070)	(0.085)
Cohort fixed effects	No	No	Yes	Yes
Baseline is:	5 dollar Incentive	5 dollar Incentive	5 dollar Incentive	5 dollar Incentive
Observations	200	160	200	160
R^2	0.046	0.003	0.067	0.038
Adjusted R^2	0.036	-0.010	0.038	-0.0001
Residual Std. Error	0.331 (df = 197)	0.372 (df = 157)	0.331 (df = 193)	0.370 (df = 153)
F Statistic	4.723^{***} (df = 2; 197)	0.248 (df = 2; 157)	2.293^{**} (df = 6; 193)	0.998 (df = 6; 153)
Note:				$^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$
				Note: Uses Robust Standard Error.

Table 5.11: In_PST Sub-group analysis regression table comparing the heterogeneous treatment effect, for students in PST, of "\$100 Lottery Incentive", "No Incentive", and "\$5 Incentive" on the Average Response Rate.

Column 1 of Table 5.11 shows that students not in PST had a large, statistically significant, decrease in response rate of -17.5% when offered a \$100 Lottery Incentive over the baseline \$5 Incentive. Curiously, this decrease in response rate is statistically insignificant and small, -4.4%, for students in PST as shown in Column 2. This result indicates that students outside of the Pacific Standard Time Zone are significantly more likely to respond to the survey when offered a \$5 guaranteed incentive over the \$100 Lottery incentive than students who are located in the Pacific Standard Time Zone. Perhaps \$5 goes a long way when not located in the West Coast U.S., and maybe even more so for internationally attending students. Perhaps students located on the West U.S. Coast, in PST, in the UC Berkeley area feel closer to the college community or identity such that the presence of incentives does not additionally motivate them to take surveys on the topic of improving their school. Neither students outside of PST or in PST respond significantly differently to the survey if offered no incentive instead of $5, since the coefficients in the No Incentive row are not statistically significant for any column.

After introducing Cohort Covariates in Columns 3 and 4, our estimates slightly change but the robust standard errors do not decrease by much. The students not in PST had a large, statistically significant, decrease in response rate of -17.2% (Column 3) when offered a $100 Lottery Incentive over the baseline $5 Incentive. The students in PST had a small, not statistically significant, decrease in response rate of -4.0% (Column 4) when offered a $100 Lottery Incentive over the baseline $5 Incentive. This means that the change in survey response for students in PST still was insignificant in the regression with Cohort covariates. Curiously, for students in PST with the Cohort covariate (Column 4) the only statistically significant result for a non-zero change in response rate was for students in the Fall 2021 Cohort of 19.0% with a 0.1 p-value threshold. Please note, however, that Cohort assignment is not randomly assigned, so this coefficient can not be interpreted as causal.

Overall, it appears that the significant increased response rate for $5 Guaranteed Incentive over $100 Lottery Entry Incentive is present for students outside of the Pacific Standard Time Zone and not for students in it.

6: Conclusion

Section 6.1: Concerns & Limitations

Estimation of causal effects requires generating a model that contains less detail than reality itself. Any such generalization requires careful consideration of how our model may differ from the true state of causal effects in the world. While we use statistically rigorous methods to ensure our model reflects reality in expectation, our model's limitations require careful consideration when interpreting results.

Our first limitation is statistical power, our ability to detect an effect if it truly exists. For the $5 incentive, our power is approximately 32% indicating that we have the statistical power to capture ~32% of true statistically significant effects were we to repeat the experiment many times. Increasing the power to 80% would require nearly 400 subjects. If we consider the possibility that our point estimate of the $5 incentive is the true effect size, we would need more than triple the number of subjects to find this effect statistically significant 80% of the time.

Compare this to the 73% power of the $5 versus lottery incentive where we observe a significant effect of the $5 incentive. The increased power of this estimate indicates not only helps identify a statistically significant effect but increases the likelihood that the observed effect is not a false positive. For these reasons, increased power would increase the rigor of attempts to replicate our results.

The negative effect of the $100 lottery incentive is also worth discussing. If we chose an unrealistic incentive of one trillion dollars, a negative effect would be shocking. Reasoning from this extreme scenario, we might also expect a positive effect from the lower dose $100 incentive. It's possible that we observed a negative effect by chance, but we also consider how our experimental design may have contributed to this result. Subjects were not informed of the expected value of the incentive. If subjects assumed that the $100 incentive email was sent to thousands of people, the assumed expected value of the incentive approaches zero, and the $100 lottery treatment only differs from the control in the wording of the email received. We do not know if subjects made such an implicit assumption. However, contrast the lottery incentive with the $5 incentive where the expected value is $5 for every subject and every subject is aware of this fact. Therefore, the $100 lottery differs not only in that it is a lottery but also in that the subjects do not have enough information to determine the expected value of the incentive when they decide to respond or not respond to the survey. Were we to repeat the experiment, the number of subjects eligible for the lottery would be communicated to the subjects so that the implicit expected value is available to subjects.

Finally, by using MIDS students as subjects, we run the risk that students may know the authors and respond at a different rate as a result. This applies especially to the author who sent out the survey requests. If a subject knows the author, they may be more or less inclined to respond. While we cannot compute the risk of this effect, we note that there could be heterogeneous treatment effects between those who know the authors and those who do not, and this may bias our results.

Section 6.2: Lessons Learned

First, given the observed effects and power calculations, we did not have enough power to consistently detect a significant effect of the $5 incentive versus the control even if our point estimate is accurate. A follow up experiment under similar conditions with a larger sample size would help delineate whether we failed to observe a significant effect due to being underpowered versus a state of the world where the null hypothesis is true.

Second, we would recommend collecting additional information in the survey to ensure we can make a correct match between survey responses and assignment to treatment. We were only able to match all responses to treatment due to our privilege as Berkeley students to associate emails with names. If we did not sample Berkeley students and ran into similar issues, we would have difficulty attributing responses to their treatment group.

Finally, identification of more informative covariates, such as work hours per week or salary, would be beneficial in any follow up experiment. Our covariates did not reduce the standard errors of any of the treatments. While an increased sample size would reduce our standard errors, more informative covariates may serve the same purpose and may be cheaper and/or easier to obtain than a larger sample size.

Section 6.3 Final Thoughts

Neither the guaranteed $5 incentive or the $100 lottery incentive rejected the null hypothesis of equal response rates between the control and treatment group. The guaranteed $5 incentive had a statistically significant positive effect relative to the $100 lottery incentive. This may be an indication that guaranteed incentives outperform lottery incentives when attempting to generate survey responses. This effect was present for the sub-group of students who were outside of the Pacific Timezone, and insignificant for those in the Pacific Timezone. If a survey designer seeks to compensate respondents, then the guaranteed incentive is recommended over the lottery incentive due to the larger response rate.

Our experience indicates several improvements that may be made in future work to replicate or extend this research. Power analysis indicates we were underpowered to detect significant effects between both treatment groups and the control group. A larger sample size and more informative covariates would increase power. We suggest including information to clarify the expected value of lottery incentives. Finally, exclusion of subjects who may know the researchers likely requires collecting subjects without connections to the researchers' institution.

7: References

1. Harris IA, Khoo OK, Young JM, Solomon MJ, Rae H. Lottery incentives did not improve response rate to a mailed survey: a randomized controlled trial. J Clin Epidemiol. 2008 Jun;61(6):609-10. doi: 10.1016/j.jclinepi.2007.07.012. Epub 2008 Feb 14. PMID: 18471665.
2. Smith, M. G., Witte, M., Rocha, S., & Basner, M. (2019, December 5). *Effectiveness of incentives and follow-up on increasing survey response rates and participation in field studies*. BMC Medical Research Methodology. Effectiveness of incentives and follow-up on increasing survey response rates and participation in field studies
3. Examining Factors Impacting Online Survey Response Rates in Educational Research: Perceptions of Graduate Students. Amany Saleh (Arkansas State University), Krishna Bista (Morgan State University). Journal of MultiDisciplinary Evaluation Volume 13, Issue 29, 2017. Examining factors impacting online survey response rates in educational research: Perceptions of graduate students.
4. Pit, S.W., Vo, Tham, Pyakurel, Sagun (2014, June 6). The Effectiveness of Recruitment Strategies on General Practitioner's Survey Response Rates - a systematic review. BMC Medical Research Methodology. The effectiveness of recruitment strategies on general practitioner's survey response rates – a systematic review
5. Pedersen, M. J., & Nielsen, C. V. (2014, December 17). *Improving Survey Response Rates in Online Panels: Effects of Low-Cost Incentives and Cost-Free Text Appeal Interventions*. SAGE Journals. Improving Survey Response Rates in Online Panels: Effects of Low-Cost Incentives and Cost-Free Text Appeal Interventions

8: Appendix

Treatment Group 1 Email Content - $5 Incentive

Response Requested: $5 Guaranteed to Complete MIDS Course Time Survey Inbox ×

Madeline Whitlow <madeline_94@berkeley.edu> Sun, Oct 31, 2:16 PM

Berkeley
UNIVERSITY OF CALIFORNIA

Dear MIDS Student,

To help better inform MIDS students on which classes they should take, we're launching a survey for *you* to provide perspective on the time requirement that certain program classes take. The survey should take no longer than 10 minutes total to complete. Individual feedback will remain anonymous. However, aggregated data will be made available to current students looking for class selection guidance. To show our gratitude, upon completion of the survey, a $5 Amazon gift card will be sent to your Berkeley email address provided on the survey. We will be collecting data until November 14th, please submit a response by then if you would like to participate. Please follow this link here to complete. Thanks!

Best,
Madeline Whitlow
(MIDS Spring '21 Cohort)

Treatment Group 2 Email Content - $100 Incentive

Response Requested: $100 Lottery Chance to Complete Course Time Survey Inbox ×

Madeline Whitlow <madeline_94@berkeley.edu> Sun, Oct 31, 2:16 PM

Berkeley
UNIVERSITY OF CALIFORNIA

Dear MIDS Student,

To help better inform MIDS students on which classes they should take, we're launching a survey for *you* to provide perspective on the time requirement that certain program classes take. The survey should take no longer than 10 minutes total to complete. Individual feedback will remain anonymous. However, aggregated data will be made available to current students looking for class selection guidance. To show our gratitude, upon completion of the survey, you will be entered into a lottery with the chance to win a $100 Amazon gift card. We will be collecting data until November 14th, please submit a response by then if you would like to participate. Please follow this link here to complete. Thanks!

Best,
Madeline Whitlow
(MIDS Spring '21 Cohort)

Treatment Group 0 Email Content - No Incentive

Response Requested: Complete MIDS Course Time Requirement Survey

Inbox ×

Madeline Whitlow <madeline_94@berkeley.edu> Sun, Oct 31, 2:14 PM

Berkeley
UNIVERSITY OF CALIFORNIA

Dear MIDS Student,

To help better inform MIDS students on which classes they should take, we're launching a survey for *you* to provide perspective on the time requirement that certain program classes take. The survey should take no longer than 10 minutes total to complete. Individual feedback will remain anonymous. However, aggregated data will be made available to current students looking for class selection guidance. We will be collecting data until November 14th, please submit a response by then if you would like to participate. Please follow this link here to complete. Thanks!

Best,
Madeline Whitlow
(MIDS Spring '21 Cohort)

Survey Content

Of the MIDS courses you've taken please indicate how many hours a week each course took you (this includes asynchronous material, studying, homework and office hours).

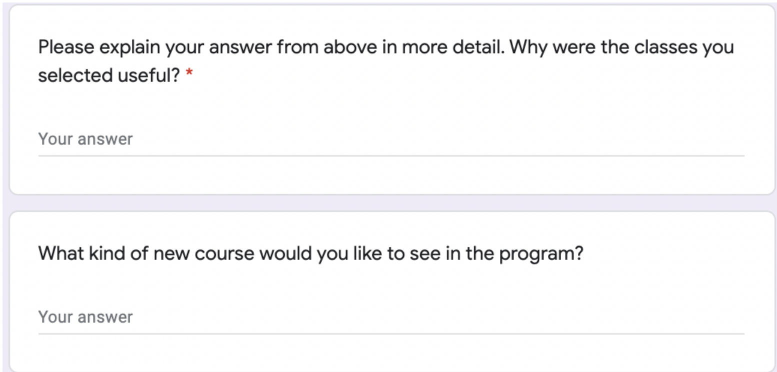
	0-10	10-20	20-30	30+
W200	☐	☐	☐	☐
W201	☐	☐	☐	☐
W203	☐	☐	☐	☐
W205	☐	☐	☐	☐
W207	☐	☐	☐	☐
W209	☐	☐	☐	☐
W210	☐	☐	☐	☐
W231	☐	☐	☐	☐
W233	☐	☐	☐	☐
W241	☐	☐	☐	☐
W251	☐	☐	☐	☐
W261	☐	☐	☐	☐
W266	☐	☐	☐	☐
W271	☐	☐	☐	☐

For the classes that took the most time, what requirement was the most time consuming? *

- ○ Homework
- ○ Async Material
- ○ Studying
- ○ Other: ____________

Which courses did you find the most useful? *

- ☐ W200: Introduction to Data Science Programming
- ☐ W201: Research Design and Applications for Data and Analysis
- ☐ W203: Statistics for Data Science
- ☐ W205: Fundamentals of Data Engineering
- ☐ W207: Applied Machine Learning
- ☐ W209: Data Visualization
- ☐ W210: Capstone
- ☐ W231: Behind the Data: Humans and Values
- ☐ W233: Privacy Engineering
- ☐ W241: Experiments and Causal Inference
- ☐ W251: Deep Learning in the Cloud and at the Edge
- ☐ W261: Machine Learning at Scale
- ☐ W266: Natural Language Processing with Deep Learning
- ☐ W271: Statistical Methods for Discrete Response, Time Series, and Panel Data



Please explain your answer from above in more detail. Why were the classes you selected useful? *

Your answer

What kind of new course would you like to see in the program?

Your answer

Table 8.1: Number of Emails per Cohort

Cohort	Num. Emails Gathered per Slack Channel
Summer 2020	79
Fall 2020	122
Spring 2021	146
Summer 2021	117
Fall 2021	92