Metric Choice

Choosing Invariant Metrics

Check each metric you would use as an invariant metric.

- **Number of cookies:** That is, number of unique cookies to view the course overview page.
- **Number of user-ids:** That is, number of users who enroll in the free trial.
- **Number of clicks:** That is, number of unique cookies to click the "Start free trial" button (which happens before the free trial screener is trigger).
- **Click-through-probability:** That is, number of unique cookies to click the "Start free trial" button divided by number of unique cookies to view the course overview page.
- **Gross conversion:** That is, number of user-ids to complete checkout and enroll in the free trial divided by number of unique cookies to click the "Start free trial" button.
- **Retention:** That is, number of user-ids to remain enrolled past the 14-day boundary (and thus make at least one payment) divided by number of user-ids to complete checkout.
- **Net conversion:** That is, number of user-ids to remain enrolled past the 14-day boundary (and thus make at least one payment) divided by the number of unique cookies to click the "Start free trial" button.

Invariant Metrics: number of cookies, number of clicks

Choosing Evaluation Metrics

Check each metric you would use as an evaluation metric.

- **Number of cookies:** That is, number of unique cookies to view the course overview page.
- **Number of user-ids:** That is, number of users who enroll in the free trial.
- **Number of clicks:** That is, number of unique cookies to click the "Start free trial" button (which happens before the free trial screener is trigger).
- **Click-through-probability:** That is, number of unique cookies to click the "Start free trial" button divided by number of unique cookies to view the course overview page.
- **Gross conversion:** That is, number of user-ids to complete checkout and enroll in the free trial divided by number of unique cookies to click the "Start free trial" button.
- **Retention:** That is, number of user-ids to remain enrolled past the 14-day boundary (and thus make at least one payment) divided by number of user-ids to complete checkout.
- **Net conversion:** That is, number of user-ids to remain enrolled past the 14-day boundary (and thus make at least one payment) divided by the number of unique cookies to click the "Start free trial" button.

Evaluation Metrics: Gross conversion, (Retention), Net conversion

Calculating standard deviation

For each metric you selected as an evaluation metric, make an analytic estimate of its standard deviation, given a sample size of 5000 cookies visiting the course overview page. Enter each estimate in the appropriate box to 4 decimal places.

- Number of cookies
- Number of user-ids
- Number of clicks on "Start free trial"
- Click-through-probability on "Start free trial"
- Gross conversion: 0.0202
- Retention: 0.0549
- Net conversion: 0.0156

Will you use the Bonferroni correction in your analysis phase?

- ○ Yes
- ◉ No

Which evaluation metrics did you select?

- ☐ Number of cookies
- ☐ Number of user-ids
- ☐ Number of clicks on "Start free trial"
- ☐ Click-through-probability on "Start free trial"
- ☑ Gross conversion
- ☐ Retention
- ☑ Net conversion

How many pageviews will you need?

Use alpha = 0.05 and beta = 0.2. Round your answer to the nearest integer, if necessary. 685325

As we now have more than one hypothesis, the chance to get false positives increases. However, our metrics are not fully independent which is why the true probability for false positives will still be lower than 9.75% (that's the case for independent metrics). We could then use family-wise error rate such as Bonferroni or false discovery rate methods to account for the multiple hypotheses problem. However, they have flaws as well (e.g. we could easily end up with more false negatives; see here and here). Hence, given that the chance to get more false positives is only slightly increased in this case, we won't control for multiple hypothesis here.

Choosing duration and exposure

Number of pageviews

How many pageviews are required? (Enter your answer from the last exercise.) 685325

Fraction of traffic exposed

What fraction of Udacity's traffic would you divert to this experiment? Enter your answer as a number between 0 and 1. 1

Length of experiment

Given this, how many days will Udacity need to run the experiment? Enter an integer number of days.

18

Sanity checks

For each metric that you chose as an invariant metric, compute a 95% confidence interval for the value you expect to observe. Enter the upper and lower bounds, and the observed value, all to 4 decimal places. Check the box if the metric passes your sanity check.

	Lower bound	Upper bound	Observed	Passes
• Number of cookies	0.4988	0.5011	0.5006	☑
• Number of user-ids				☐
• Number of clicks on "Start free trial"	0.4959	0.5041	0.5004	☑
• Click-through-probability on "Start free trial"				☐
• Gross conversion				☐
• Retention				☐
• Net conversion				☐

Effect Size Tests

For each of your evaluation metrics, compute a confidence interval around the difference.

Did you use the Bonferroni correction? ○ Yes ◉ No

	Lower bound	Upper bound	Statistical significance	Practical significance
• Number of cookies			☐	☐
• Number of user-ids			☐	☐
• Number of clicks on "Start free trial"			☐	☐
• Click-through-probability on "Start free trial"			☐	☐
• Gross conversion	-0.0291	-0.0120	☑	☑
• Retention	0.0081	0.0541	☑	☐
• Net conversion	-0.0116	0.0019	☐	☐

Sign Tests

Run a sign test on each of your evaluation metrics using the day-by-data data. Enter each p-value, and indicate whether each result is statistically significant.

Did you use the Bonferroni correction? ○ Yes ◉ No

	p-value	Statistical significance
• Number of cookies		☐
• Number of user-ids		☐
• Number of clicks on "Start free trial"		☐
• Click-through-probability on "Start free trial"		☐
• Gross conversion	0.0026	☑
• Retention	0.6776	☐
• Net conversion	0.6776	☐