Email Report

Justin Wininger

2024-06-13

# Introduction

This report analyzes several statistics of 645 email campaigns including over 1.5 million emails sent out by 59 nonprofit organizations represented by Endowment Development Services since 2010. When comparing two or more groups, this report will use the median instead of the mean or average. The median, being the point where there are equal amounts of observations larger and smaller, is more resistant to outliers than the mean and is therefore a much better representation of the “typical” observation. This report will mainly observe the open rate and click rate of these various email campaigns. Because it is physically impossible for either one to be negative, both will be skewed to the right. Because of this, using the mean would result in inflated and unrealistic claims. Using the median, therefore, is the best choice.

# Size

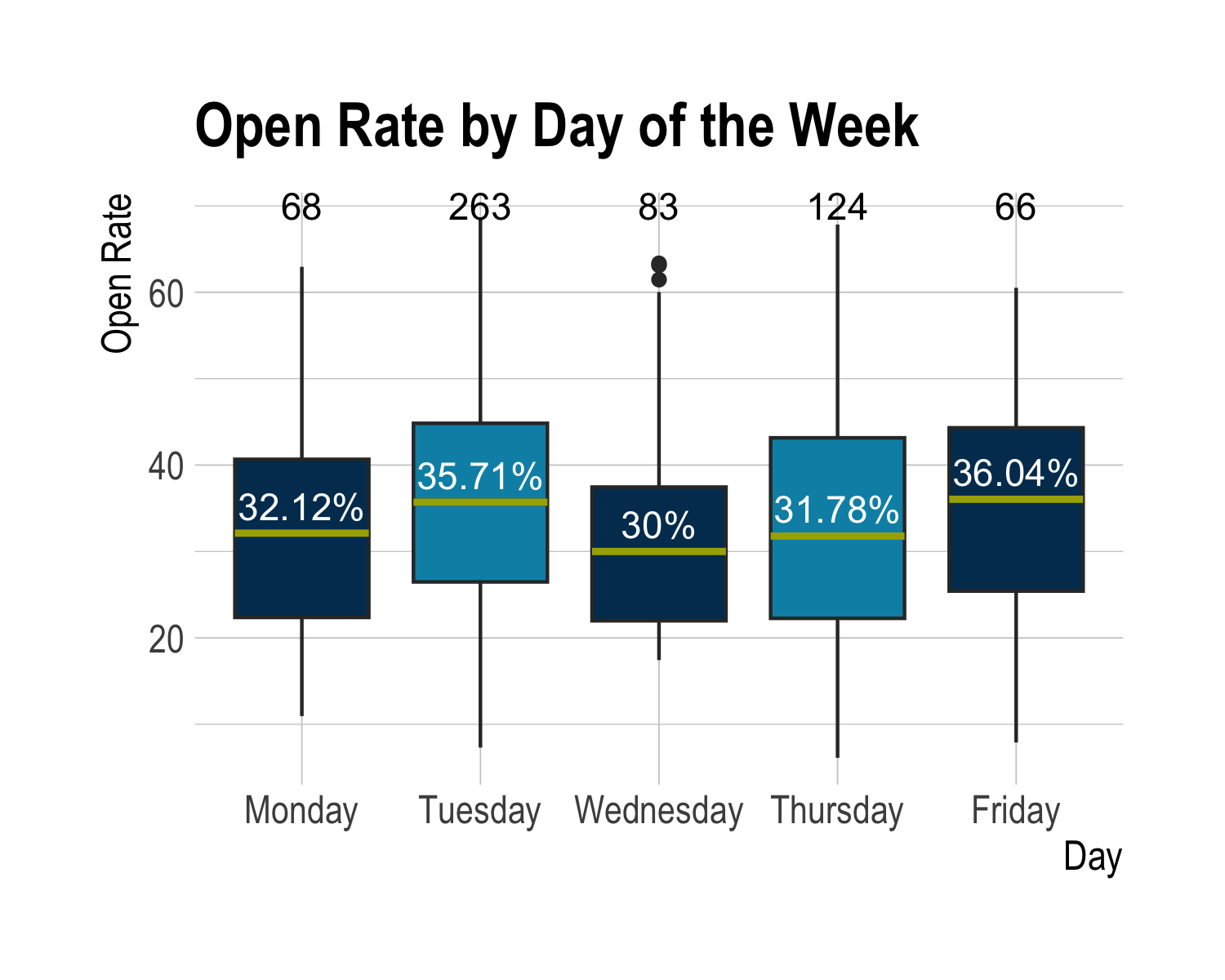
Nonprofit organizations sizes within these data vary wildly. Some have list sizes of just 50, while others have list sizes of almost 30,000. The average list size is only 2,473. Using 2,500 as an arbitrary benchmark between small and large, among the 59 nonprofit organizations represented, 27 of them are considered large and 32 of them are small. More statistics are in the table below.

| Size | Count | Median List Size | Median Open Rate | Median Click Rate | Median Bounce Rate |
| --- | --- | --- | --- | --- | --- |
| All | 59 | 1403 | 33.27% | 0.97% | 4.73% |
| Large | 27 | 4219 | 33.28% | 0.87% | 3.88% |
| Small | 32 | 838 | 32.99% | 1.18% | 4.98% |

Looking at the table, the results do not appear to be significantly different. Performing a Chi-Square Test for Homogeneity confirms it. There is not a statistically significant difference between large and small list sizes in open rate, click rate, or bounce rate.

# Day of the Week

Once work on an email campaign is finished, perhaps the biggest question faced is when to send it. Below is a side-by-side box and whisker plot of the open rate plotted by day of the week. The lime colored line represents the median open rate for each day. The white text is that median.



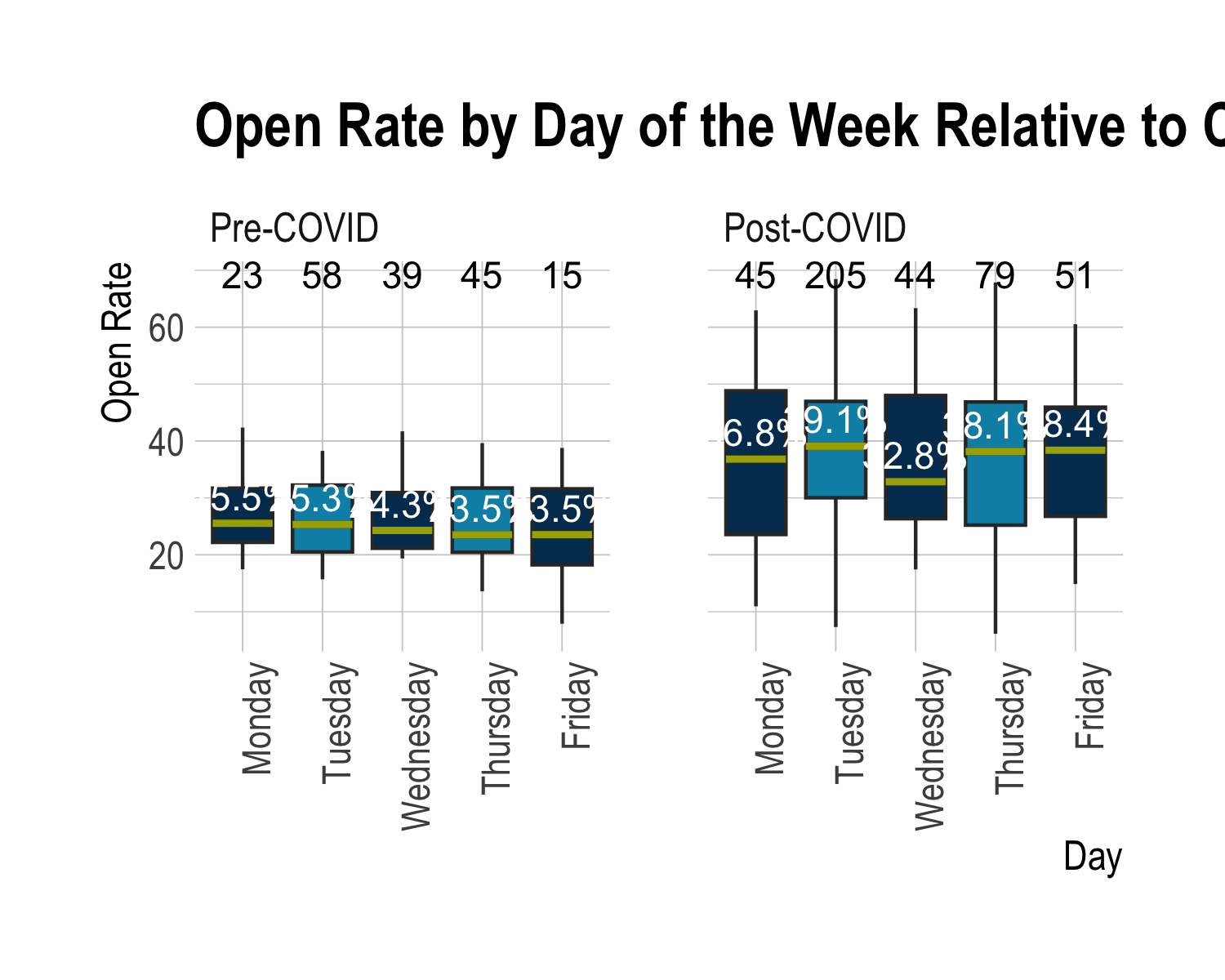
Open Rate by Day of the Week

This visualization reveals that not all days are created equal. Unfortunately, there was not enough data on Saturday or Sunday, so they were excluded. Friday and Tuesday seem to be the best days to send campaigns, and Wednesday and Thursday are the worst days. This relationship is statistically significant. Despite being statistically significant, there is always still the chance that any variation seen is due to random chance or lurking variables. For example, perhaps the nonprofits with the most loyal donors happen to have a policy to send campaigns on Fridays. This would skew the data. Sending campaigns on Tuesdays or Fridays, therefore, is not a guarantee the open rate will increase, but those days are the best bet to get a higher open rate.

Interestingly, the days with the highest median click rate are Thursday and Wednesday, respectively. The two lowest days are Monday and Friday. Friday being the worst day for clicks is very surprising, especially considering that it is the best day for opens. This disparity makes a “best day” very hard to determine. Some days have high open rates, while other days have high click rates. Ultimately, it will be up to the individual to decide which is more valuable to him: opens or clicks. The only thing that can definitively be said is that Mondays have low open and click rates; do not send email campaigns on Mondays.

## COVID-19

Unfortunately, the practice of donating to nonprofits was not unaffected by the pandemic. It might have ended up, however, being a net positive. Open and click rates for every single day increased after COVID, as seen in the visualization below.



The observed shift makes sense. The pandemic forced people to digitize. A direct result of this digitization is that people spent more time working on their computers, which gives them more opportunities to check their emails. Median open rates increased, but so did the variability. This is represented by the size of the box. The bottom of the box is called the first quartile. It is the 25th percentile. The top of the box is called the third quartile and is the 75th percentile. Half of all observations, therefore, are contained within the box. Because the post-COVID boxes are taller, the data is more variable and has a wider range. Looking at only post-COVID, the best days to send campaigns for open rates are still Tuesday and Friday. However, Monday is now much better than it was, and Wednesday is much worse than it was.

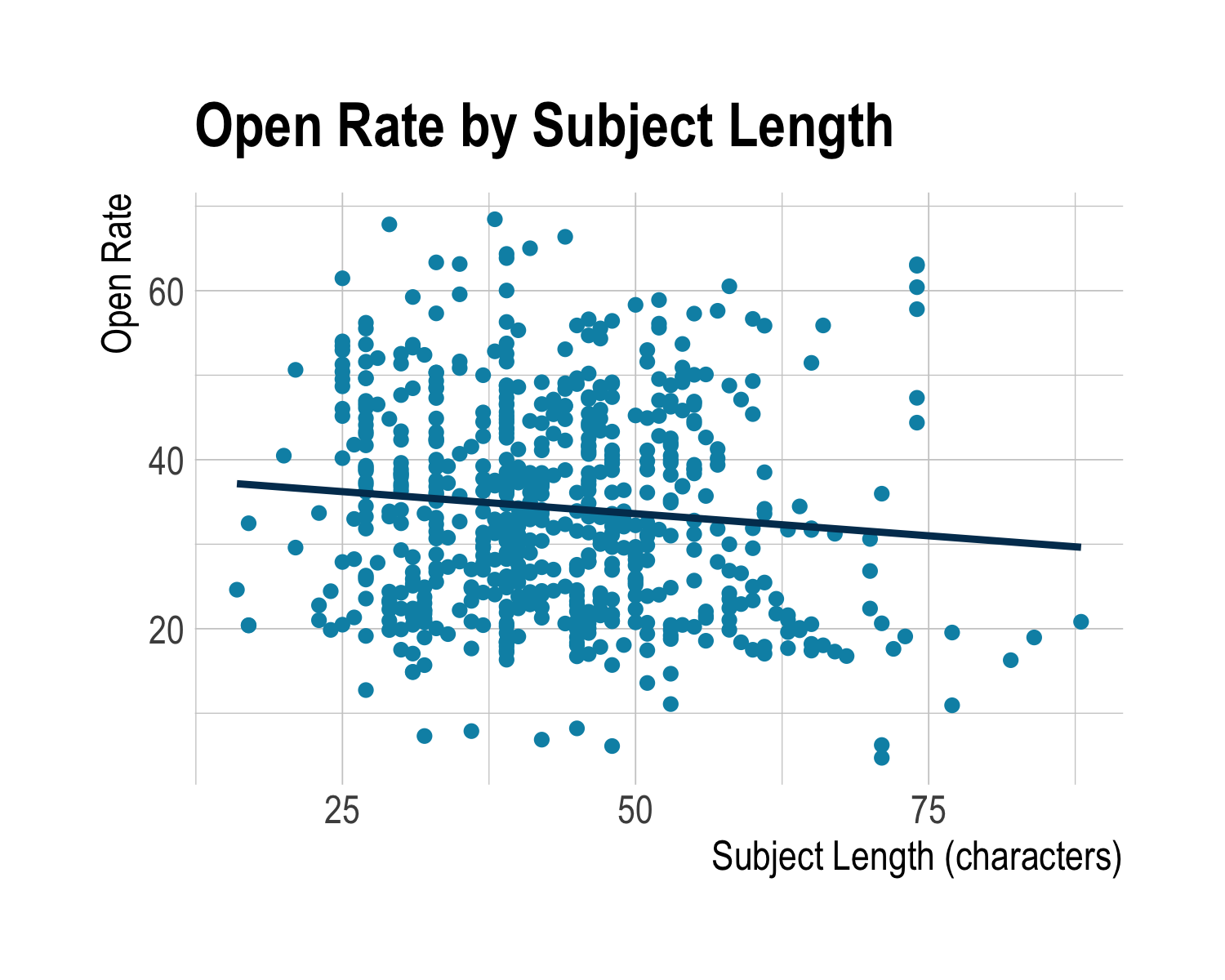
COVID made click rate decrease. The median click rate before the pandemic was 1.39%, but it is 0.86% after the pandemic. The days with the best click rates post-pandemic are Wednesday and Thursday, and the worst days are Monday and Friday. Once again, because the best days for open and click rate differ, it is up to the individual which is more important.

# Subject Line

The subject line is one of if not the most important parts of an email when it comes to engaging a potential donor. It is the first thing he will see. It is imperative, therefore, to maximize the effectiveness of the subject line.

## Subject Length

Are longer subjects more engaging? Below is a scatterplot that shows open rate by subject length in characters.



The pattern of the points shows no clear relationship, but the line of best fit has a definite negative slope. This should mean that as subject length increases, open rate decreases. This relationship is even statistically significant. is a measure of goodness-of-fit. It essentially measures how much the variation in open rate is caused by subject length. The value for this model is 1.01%. Only 1% of the variation in open rate is caused by subject length. So, even though the negative linear relationship is statistically significant, the impact of subject length on open rate is essentially 0. It is safe to conclude, therefore, that subject length does not realistically matter when observing open rate, as long as subject length is somewhere between 15 and 90 characters.

## Punctuation

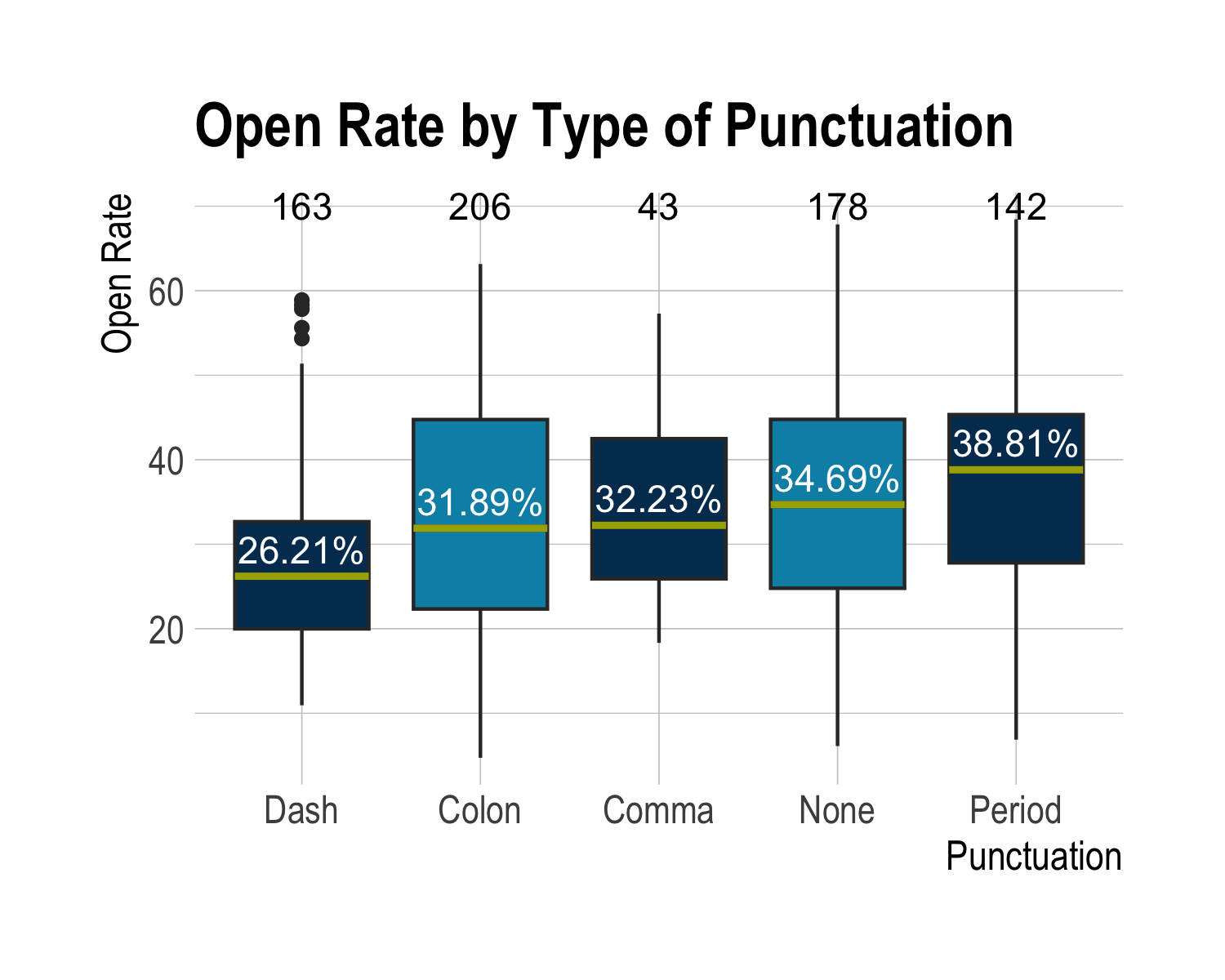
Below is a table comparing subjects with punctuation versus subjects without punctuation.

| Punctuation | Count | Median Open Rate | Median Click Rate |
| --- | --- | --- | --- |
| No | 178 | 34.69% | 0.73% |
| Yes | 432 | 32.84% | 1.26% |

Like before, the presence of punctuation does not significantly affect open or click rate. However, just like the days of the week, one must choose whether he wants a higher open or click rate. Subject lines with no punctuation have a higher open rate, but subject lines with punctuation have a higher click rate.

### Types of Punctuation

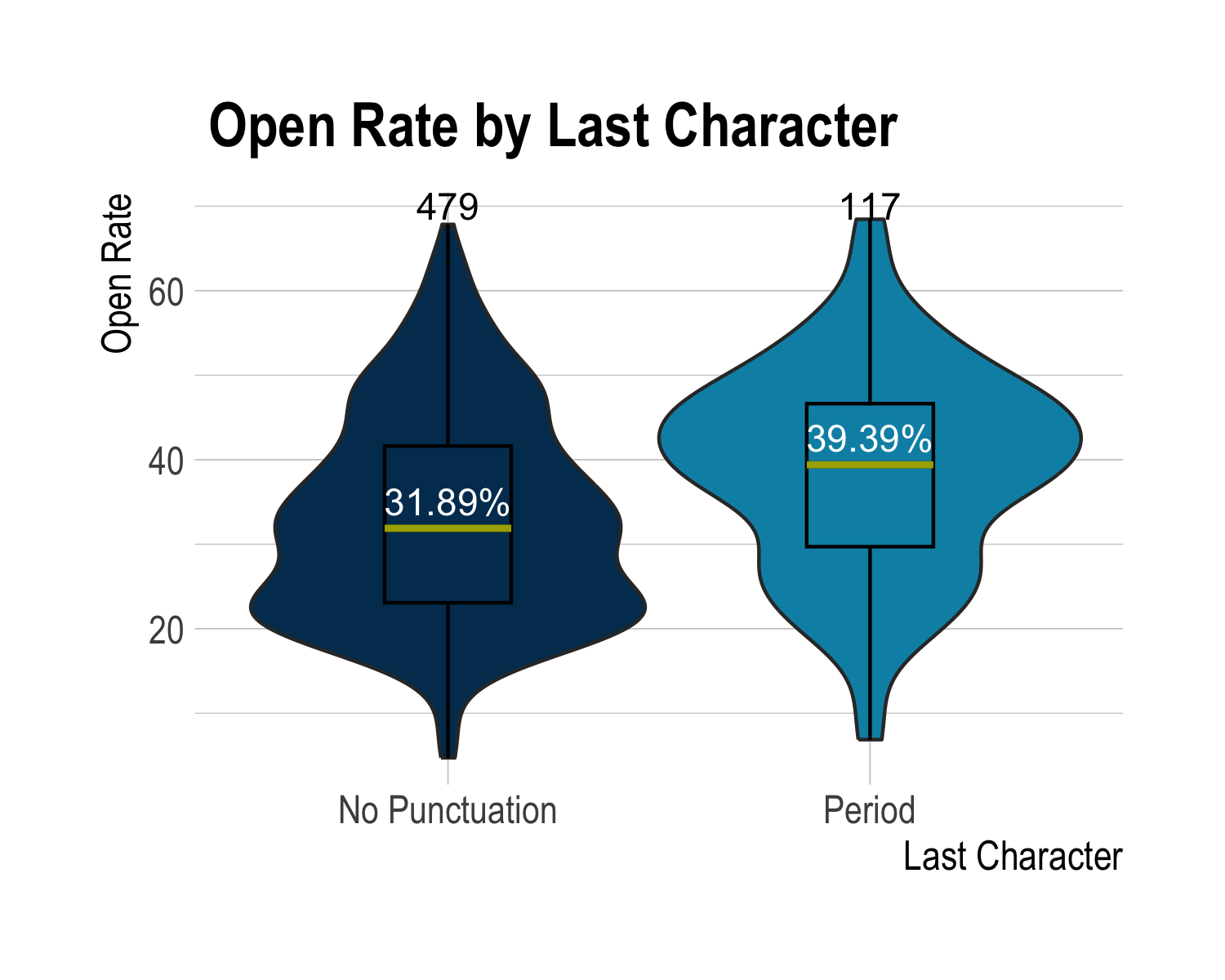
Punctuation alone is not enough information. There are tens of different types of punctuation. Below is a plot showing all types of punctuation with at least 15 observations.



This visualization is unusual in that is contains some observations multiple times. For example, a subject line with a colon, dash, and period would be included three times, once in each box. Based on the plot, dashes do not perform well and periods perform the best. As expected, click rates are the exact opposite. Dashes have the highest median click rate, and periods have the lowest. Subject lines, however, are designed to persuade people to open the email, not necessarily to click within it. When discussing subject lines, therefore, catering to open rates is much more important.

### Last Character

Within a subject line, the last character may be among the most important. Below is a violin plot comparing the difference between ending a subject with a period versus ending it with no punctuation at all. This violin plot contains the all-too-familiar boxplot, but also includes the distribution of the groups.



The boxplots alone reveal that ending with a period seems to be significantly better. The violin plot tells the same story, but to an even greater extent. The violin shows the distribution of observations. Where the violin is wider, there are more observations. It is very evident from this plot that ending the subject line with a period is significantly better. Plots alone are not enough, though. Luckily, Analysis of Variance (ANOVA) agrees with the plots and returns a statistically significant result. Naturally, however, click rates show the opposite relationship. Like previously stated, thugh, click rates do not matter

