Below is a scatter plot for means of the 3 key metrics and over 4,000 companies. Bounds of the axis are limited so that outliers could be excluded.

We see that most of the observations are concentrated in the lower-left area, which is consistent with histograms in the notebook. A lot of observations are around x-axis, which means check-in counts stay at a low level even though the Facebook pages are popular. This also implies that many companies have very low conversion ratio.
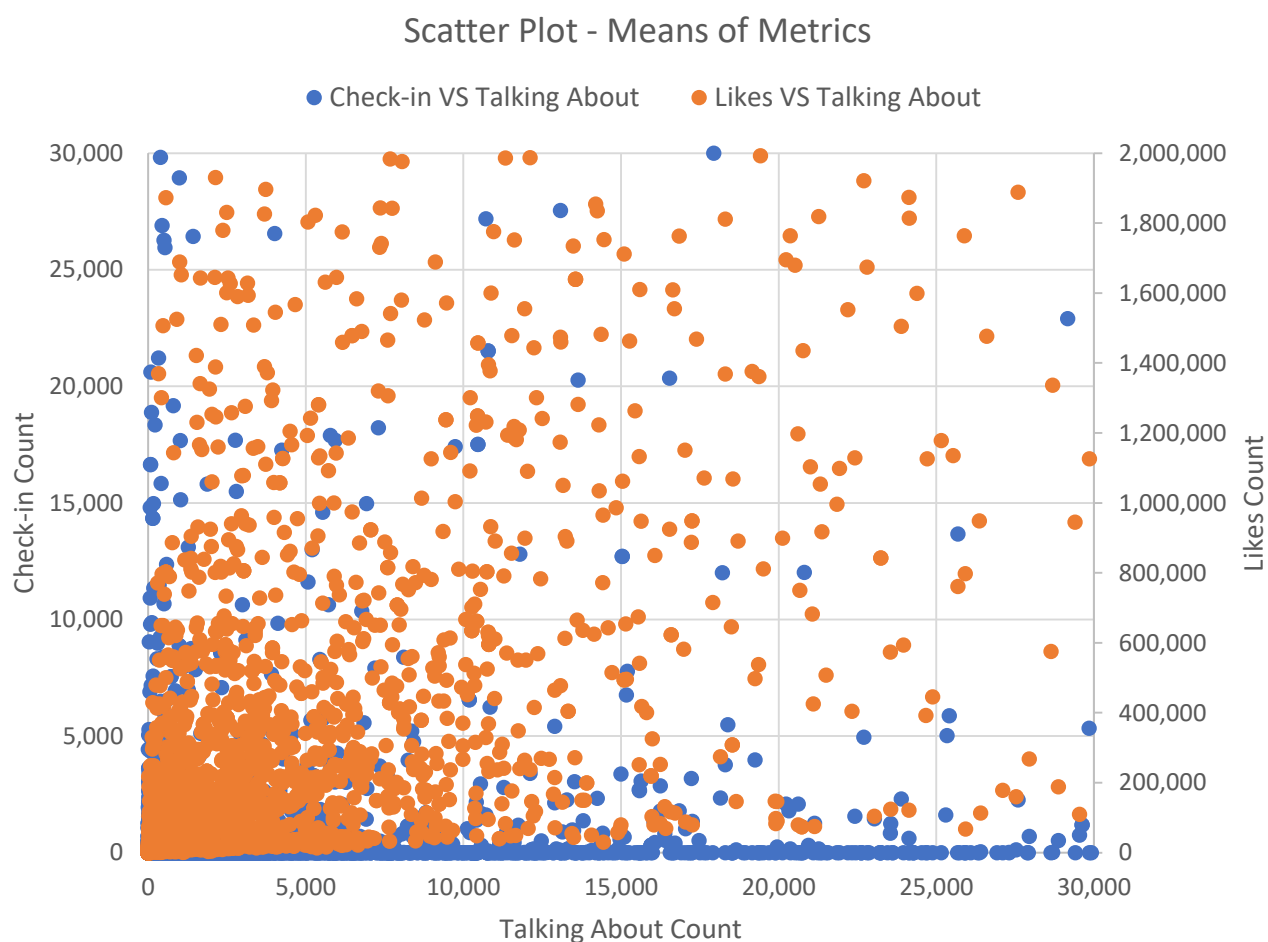

Though correlations in the notebook imply positive relationships among the 3 metrics, there is no linear relations observed in the scatter plot.

Recall that correlation between Talking About Count and Likes are much higher than that between Talking About Count and Check-in, which could also be observe from the plot. We see that as Talking About Count increases, most orange points, which represent Likes, are away from x-axis, compared to blue points, though this trend is not very clear.

About 20 companies are selected out of the 4,700, all of which are among the top 10 for either mean Check-in Count, mean Likes Count, or mean Talking About Count. The 2 charts below compared these companies' social media performance measured by the 3 metrics (companies are sorted by their ranks of Talking About Count):

Talking About Counts VS Likes

■ Mean Talking About Count ●— Mean Likes Count
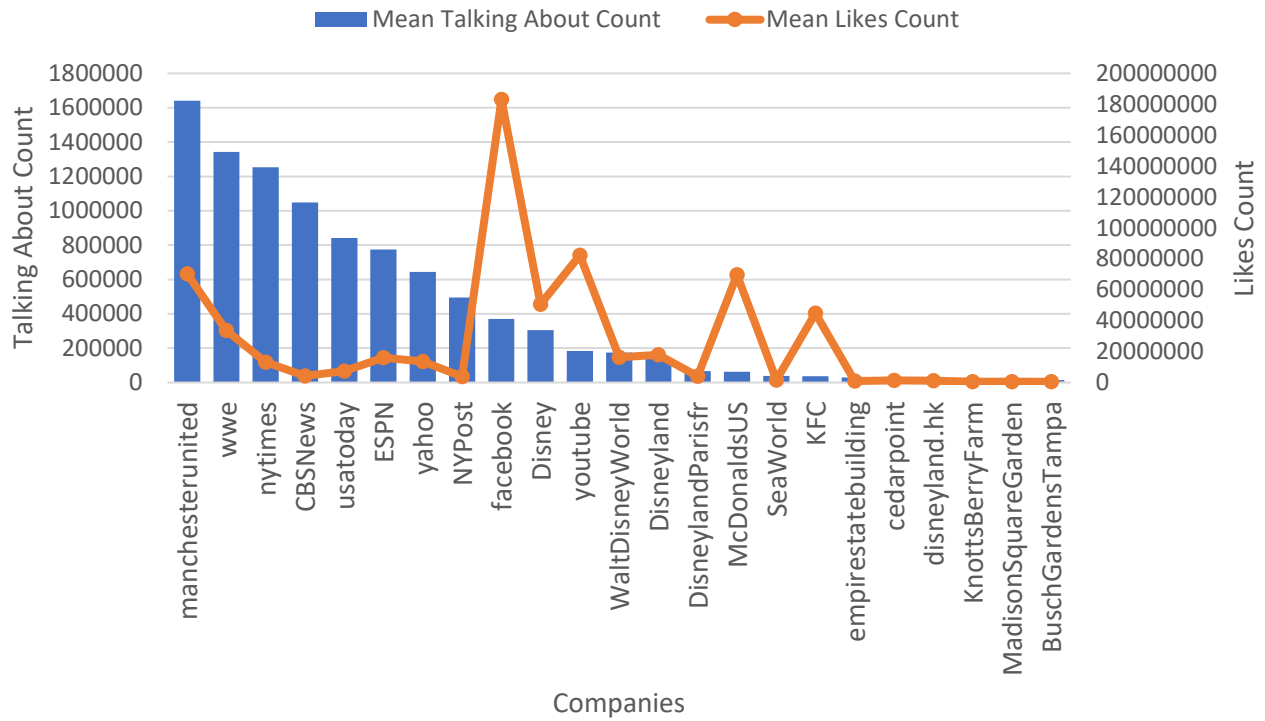
Talking About Counts VS Check-ins

■ Mean Talking About Count ●— Mean Check-in Count
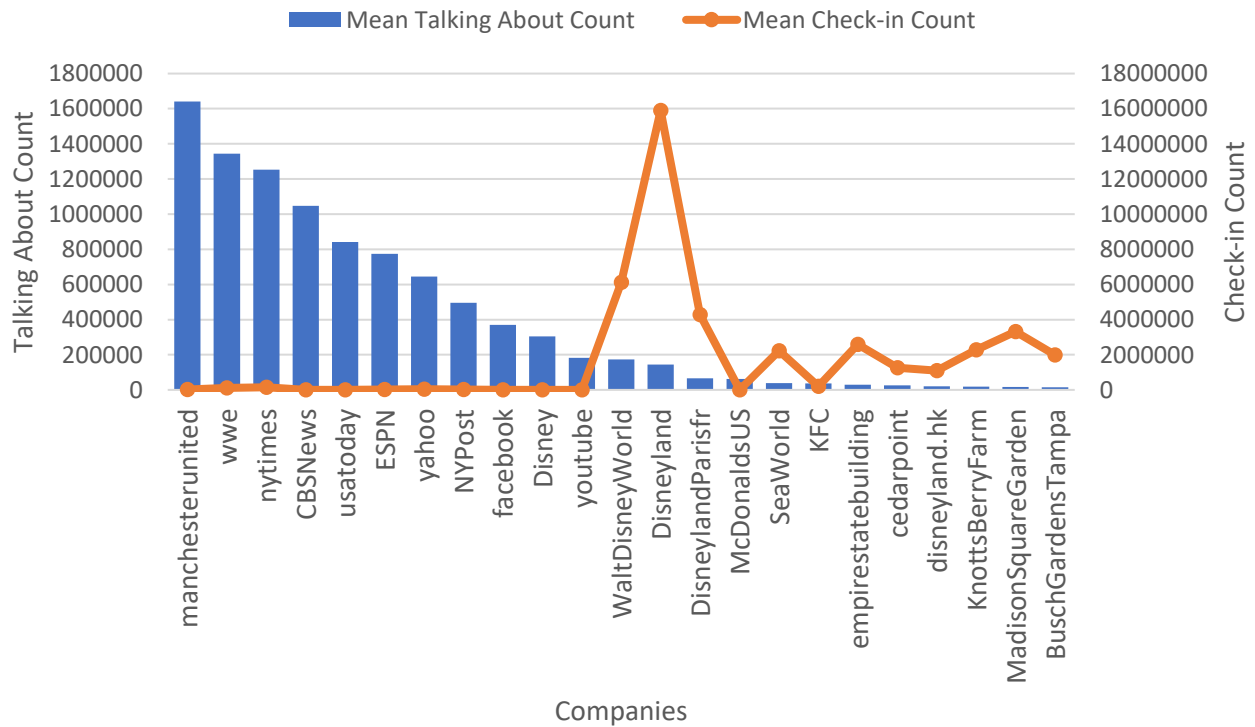
We see that all companies with the top mean Talking About counts do not have very high Likes or Check-in counts. It is the same with companies with very low Talking About counts.

However, the difference is that top popular companies tend to have moderate Likes counts but their Check-in counts are close to 0, which is because most of them do not serving people in their physical location, and foot traffic is not a performance metric for them; On the other hand, companies with low Talking About counts tend to also have poor Likes count but relatively good Check-in counts. KFC and McDonaldsUS are 2 exceptions in this case, their Likes counts are OK while Check-in counts are zero, which is because they are the official Facebook accounts for the 2 companies instead of any specific restaurants.

As for ratios, most these companies do not have good Engaging rates (above 5%), especially for those with lower Talking About counts. Not taking those 'no physical service' accounts into consideration, most of the companies' conversion rates are above 1, which implies that these top companies might not rely on social media traction.

Engaging VS Conversion
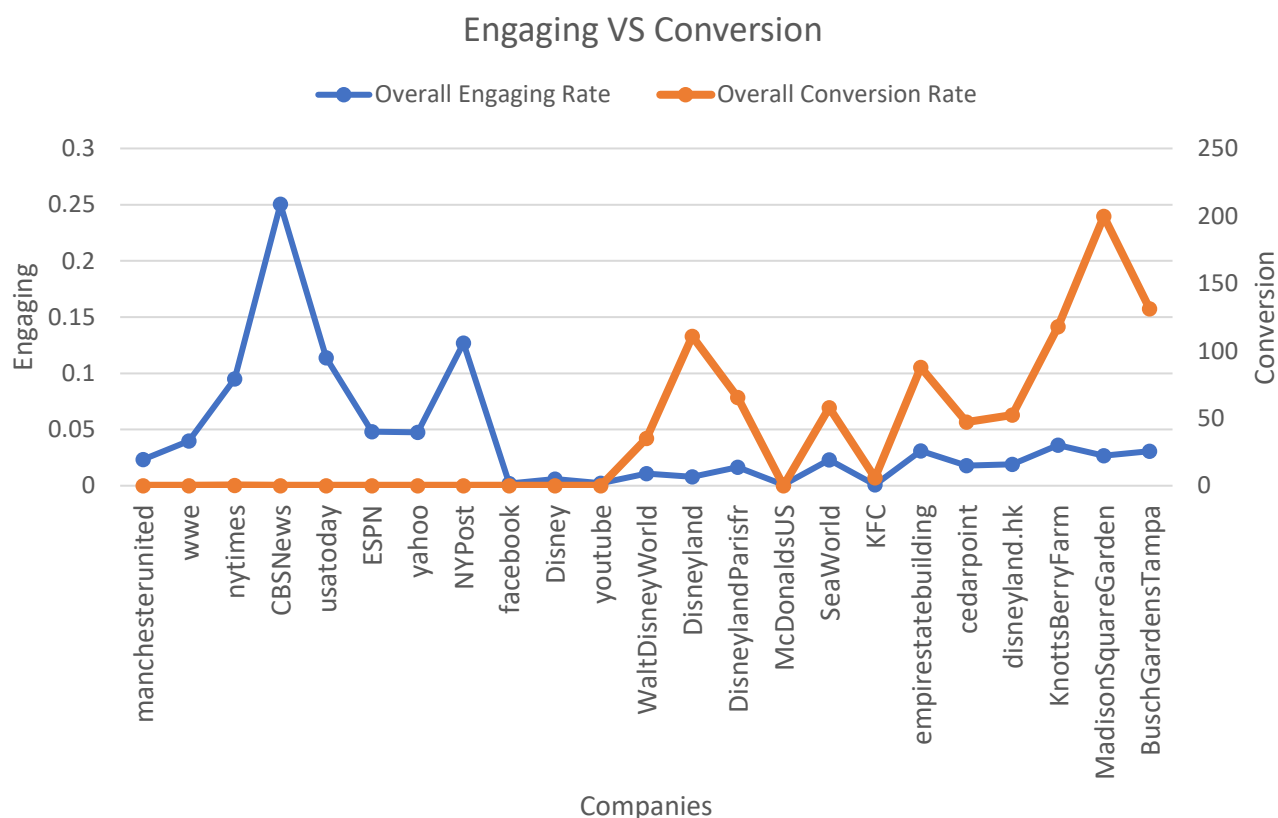
This could be an issue in future modeling processes, since purpose of this project is looking quantitative relations between those metrics, time, and any important events or campaigns, as well as trying to predict future foot traffic. If an account / company does not offer physical service, or do not rely on social media traction, the project would be meaningless for it. A possible solution could be simply exclude such companies from future discussion, or to focus on several specific types of companies.