Actionable Data Analysis: Harnessing the Full Potential of Your Data to Gain Insight

By Justin Gilbert

You've been in the meeting before. You're sitting down with a client or your executive team and presenting the great work you and your PR team have done. You show some examples and then rattle off some figures about the number of impressions or the estimated earned media value, and then you get asked the question: "So…what does that actually mean?"

> **The goal of any insight is always to understand what is performing and what is underperforming and then to shift resources (your time and effort) to the former.**

Sound familiar? The fact is that every day the volume of data available to marketers and communications professionals is expanding. The expansion is set to accelerate further as the social media footprint continues to change the way brands and consumers communicate. But, as we gain access to more data, some fundamental questions persist in challenging us:

- How do we move from reporting data to developing meaningful insights?
- How do we know what data to focus on and how to apply it?
- How do we calculate the impact of our efforts?

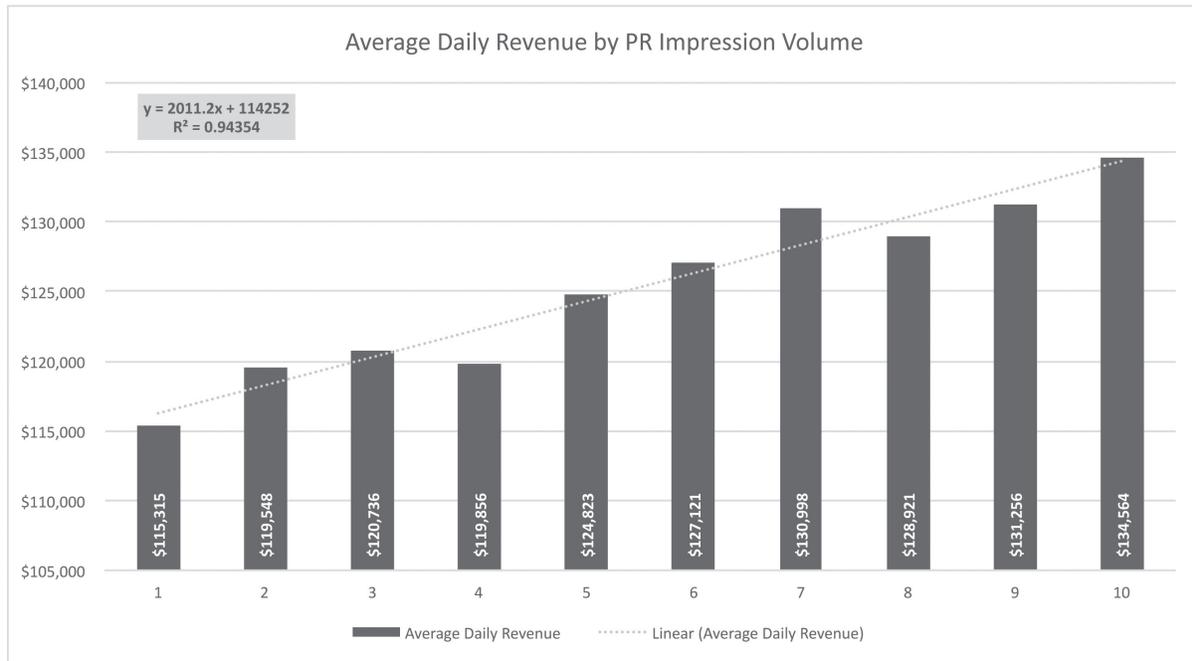The good news is that we can, in fact, answer these questions. Computer software (e.g. **Microsoft Excel**) now enables anyone to conduct complex analysis quickly, easily and without needing a degree in advanced statistics. So, let's work backwards through our list of questions and demonstrate how.

How Do We Calculate the Impact of Our Efforts?

In order to calculate the impact of your efforts, you need two things: (1) a metric that quantifies your effort over time (e.g. impressions) and (2) a key performance indicator (KPI) reflective of what you want to improve (e.g. revenue). Ideally, you'll have this information over a decent window of time (six months to a year are good), and you'll be able to get it at a granular level (daily is optimal).

So let's assume you have a year of earned media impressions and revenue by day. Now you'll need to apply two techniques: (1) structuring the data by decile and (2) performing basic linear regression. Those terms might sound intimidating, but the bark is worse than the bite.

Structuring your data by decile simply means breaking it up by every 10th percentile. If you have a year's worth of impressions data by day, you want to know on which days the number of impressions fall below the number of impressions equal to the 10th percentile. This is the first decile. Those days where the number of impressions are greater than the 10th percentile but less than the 20th percentile equal the second decile, etc.

Average Daily Revenue by PR Impression Volume

$$y = 2011.2x + 114252$$
$$R^2 = 0.94354$$

Decile	Value
1	$115,315
2	$119,548
3	$120,736
4	$119,856
5	$124,823
6	$127,121
7	$130,998
8	$128,921
9	$131,256
10	$134,564

Average Daily Revenue .... Linear (Average Daily Revenue)

An example of decile range.

Once established, the next step is to determine the average daily revenue on the days that constitute each decile. To calculate this use the averageifs() function to tell Microsoft Excel to find all days that fall between each range of percentiles and then average the daily revenue from those same days.

Once that is complete, you should have a table that looks something like the table above. The first column represents the decile, the second column shows the number of impressions at the decile, and the third is the average daily revenue at that level.

A simple visual check can show whether or not revenue is increasing as impressions are increasing. If so, the next step is to perform linear regression to determine the strength of the correlation and the rate of that increase.

Linear regression may sound intimidating but, in a nutshell, it simply determines if an independent variable (impressions) changes as a dependent variable (revenue) changes.

Linear regression is quite simple to perform in Microsoft Excel with at right. First

Decile	Average Daily PR Impressions	Average Daily Revenue
1	1,151,815	$115,315
2	1,246,264	$119,548
3	1,348,458	$120,736
4	1,459,032	$119,856
5	1,578,673	$124,823
6	1,708,124	$127,121
7	1,848,190	$130,998
8	1,999,742	$128,921
9	2,163,721	$131,256
10	2,341,146	$134,564

An example of a linear regression chart.

highlight the 'Avg Daily Revenue' column and then insert a chart (line or column work fine, for this example column is used). Next click on the columns, right click and click 'Add Trendline'. Once the trendline is plotted, check the boxes for 'Display Equation' and 'Display R-Squared'.

The R^2 value is critical and demonstrates the strength of any correlation. The number will always range between 0 and 1—the closer the R^2 value is to one the stronger the correlation. The R^2 value should ideally be 0.9 or higher. Lower values can still be used in some cases but with greater skepticism. Here the R^2

Impact Analysis

Creating a decile:

With a data table sorted by date, use the percentile() function for all daily impressions to determine the number of impressions at 10 percent, 20 percent, etc.

Once established, use the averageifs() function to calculate the average daily revenue on days that fall within each decile—meaning the average revenue on days that fall within each decile based on impressions.

Linear regression:

Once the data is structured, this step determines whether there is a positive correlation and (1) if average daily revenue increases as impressions increase, (2) the strength of the correlation and (3) the rate at which it increases.

On a graph, plot the average daily revenue by decile. Then right click and add a trendline by default linear regression. Make sure to display both the equation and the R-Squared value.

If revenue increases and has an R^2 value of 0.9 or greater, there is a positive correlation between increases in impressions and revenue. The number before x on the equation is the rate revenue increases along the regression line.

How Do We Know What Data To Focus on and How to Apply It?

Before the analysis is conducted, it's important to know your KPIs. The easiest KPI to determine is the dependent variable, which is the variable you want to see change. This is often revenue, but it can be any quantifiable variable such as web traffic or search volume.

It's important to know realistic goals of the PR effort. While the analysis can correlate PR efforts to revenue, there will not always be a direct correlation. Often, the point of a PR effort is further up the funnel, generating awareness and interest. In this case a non-monetary KPI may be best. For example, if you know revenue increases with web traffic or search volume, it may be best to demonstrate a correlation between increasing PR impressions and those metrics.

This approach helps factor in causation. The fact that two variables correlate does not mean that the independent variable is actually a cause. It's best to use good judgment and common sense. Consider the PR effort: the goals, the placement, the quality of the impressions, calls to action, etc. It may often be best to run correlation tests on multiple KPIs.

Determining an independent variable can be slightly trickier. The best approach is to apply a combination of logic and multiple tests. For example, impressions are generally a good baseline KPI for an independent variable. However, as any marketing professional will tell you, not all impressions are equal. Other KPIs like estimated earned media value may better factor the value of impressions by medium and placement.

Applying the data requires a degree of intuition in understanding what you can effectively control in your PR efforts. Do impressions generated via certain outlets result in a stronger increase in revenue? If so, then focusing efforts on generating PR in those and similar outlets is certainly an application.

value is 0.9435, which shows a strong correlation between increasing impressions and average daily revenue.

The rate at which average daily revenue is generally increasing is also on the chart. The number in front of the x in the equation is how much average daily revenue is increasing along the regression line: in this case about $2,011. You can also calculate this using the linest() function on the revenue column in the table above. If you also apply this function to the impression column, it shows that the increase in impressions from one decile to the next is 131,391 impressions. It's now possible to say that in general an increase in approximately 131,000 PR impressions correlates to a $2,011 increase in average daily revenue.

If Nothing Else…

To understand the impact of efforts, you have to structure the data properly. Creating a decile range is one approach, but there are many other ways to do this as well. A good rule of thumb is that however you structure your data you should be able to glean some basic information by looking at the table before doing the in-depth analysis.

Correlation is not causation. Understand the goal of your PR effort and look for correlations to metrics that reflect those goals. It often won't be revenue, so don't box yourself in.

Measurement for the sake of measurement misses the point. Every bit of analysis you conduct should be rooted in a question. Those questions should be based on what you can change, depending on the answers.

Sometimes you simply do not have the data you need. If that's the case, make a plan to get the data you need.

The most readily available data (e.g. Google Analytics) may often not be the best data to use.

Don't always go past the mark. Know what's expected of your efforts and conduct analysis to those points. At the end of the day, your efforts are part of a whole. You don't want to project findings that fail to recognize that.

Analysis is not gospel. It's an attempt to better understand data. Take everything with a grain of salt and understand that there are few absolutes.

Build your analytical skills over time. Spend a little bit of time each week practicing functions in Excel or researching analysis techniques online. As with anything, the more you practice the more proficient you will get.

Never be afraid to ask for help. As a data nerd myself, I can say we often love helping on these types of projects.

The same can be said for medium, timing and topic or subject matter. Applying the data is about peeling back layers to understand the impact across dimensions where you have some control.

How Do We Move from Reporting Data to Developing Meaningful Insights?

This is the most critical aspect of any measurement or data analysis. The most beautiful chart or complex analysis has little or no relevance without being able to answer the simple question: "What can I do with this?"

As previously mentioned, applying data analysis and measurement involves peeling back layers to determine impact in areas you can somewhat control. Look for how KPIs change in relation to these dimensions.

Is there a stronger correlation between increases in impressions and revenue with outlets that fit a certain category versus overall? The insight is then to focus efforts on increasing exposure in such outlets and developing better relationships with journalists who work in the field.

Do certain impressions related to certain subjects perform better? Are they general subjects or are they topical? If certain subjects perform better, the insight is to pitch stories along those lines. If more topical subject matter performs better, the insight is to then scout for other areas of opportunity on the calendar and build a messaging calendar around those opportunities.

The same approach can be applied to any dimension: day of week, time of day, medium, etc. The goal of any insight is always to understand what is performing and what is underperforming and then to shift resources (your time and effort) to the former.

Final Thoughts

In many ways, explaining a process for measurement in an article can be harder than the analysis itself. At first, this approach may seem like it might as well be written in Ancient Greek. But take time to work through the process and it will become clear. More important, it will help demonstrate the value

of your effort and show a path to improving it. Below are some final pointers on more effectively incorporating measurement into your efforts:

- **Math and/or Excel is not everyone's strong suit.** The same can be said for PR. If the steps here are just too difficult, find someone in your organization who can help with that part. You'll be able to help them understand what you're looking for from the analysis and they'll be able to help deliver.

- **Never underestimate the power of the help resources in a product like Micro soft Excel.** The formulas may seem overwhelming at first, but this tool provides great examples in easy-to-understand terminology. If you need more assistance, there are great resources on the internet such as MrExcel.com or Lynda.com.

- **Analytics always starts with questions.** To get the most out of analytics, know the questions before you start. **PRN**

Justin Gilbert is marketing analytics director at R&R Partners.