DirectAvenue Project Overview

2019-12-06

Company Overview

DirectAvenue is a forward-thinking, performance-driven, tenacious team of media buying experts who are never satisfied with the status quo. We are motivated to work, both proactively and tirelessly, to pursue every potential opportunity on behalf of our clients. *Our mission is to maximize the performance of our clients' direct response media investments* by increasing clearance, achieving spending goals, driving down media costs, leveraging negotiations and delivering more media opportunities. Through our participation in this process we hope to engage with analysts who display a real knack and passion for our business and may ultimately become a part of our Carlsbad, CA-based team.

Project Description

Project Option 1: Spot's Level GRP and Monthly Level Reach Estimation

The goal of the project is to estimate the *Gross Rating Point (GRP)* of each *unrated TV spot* and then predict the reach based on the estimated GRPs rolled up into a monthly level.

To better help you understand *these* terminologies, please see the background information below.

What's a TV Spot? To make it simple, each TV Spot is an individual advertisement placed on any TV accessible channels (such as a Nike ad you see on ESPN or a 24-hour Fitness ad you see on Fox Sports West). When we buy a TV spot, we are buying the empty slot (10 seconds, 15s, 30s, 60s, 120s, etc.) within any TV program that's available for ad insertions. Since different TV stations have different audiences (viewer size, demographic, geographic, income, education level, etc.), the price of these empty slots varies.

In the advertising industry, we have a standard set of measurements to describe the relative competitiveness of these available slots as shown below (more information here).

Gross Rating Point (GRP): a term used in advertising to measure the size of an audience (or total amount of exposures) reached by any TV spots. For instance, if a TV spot (from a certain channel) has a GRP of 2.5 (% is omitted), it means the audience size of this TV channel at this time is about 2.5% of the overall US TV households. One thing to notice is that GRP may include duplicated audience. Therefore, there may be only 2% unique audience exposed to this spot with rest 0.5% as returning viewer. Thus, multiple TV spots' GRP can be rolled up to give us aggregated numbers.

Example 1: A 30 second spot should have doubled GRP compared to a 15s spot placed in the same station at the same time.

Example 2: If there are 30 spots in a week, we can add all 30 GRP values together to get the weekly total GRP.

Reach: Since GRP included duplicates, we use Reach to describe the unique % of audience. Taking the same example above, the reach of that TV spot would be 2 (% is also omitted) while its GRP is 2.5. Since reach does not count duplicates, its maximum value is 100 (%), which means the TV spot is delivered to all US households at that given time period. Neither can Reach be aggregated, as 2 TV spots might not be mutually exclusive.

Example 1: A 30 second spot should have less than doubled the Reach value compared to a 15s spot placed in the same station at the same time.

Example 2: If there are 2 spots aired in a day, the daily total reach should be equal or less than (depending on audience group overlapping) the sum of the 2 individual reach values.

The basic equation to calculate Reach is:

$$\text{Reach} = a + b - (ab)$$

$$\text{where} \quad a = \text{reach of the first medium (expressed as decimal)}$$
$$b = \text{reach of the second medium (expressed as a decimal)}$$
$$\text{Reach} = \text{decimal percent of the target exposed to either of the two media}$$

One essential assumption held here is that _the audience are randomly exposed to either of the 2 media (or random duplication)_. In many other cases, however, this is not true if the 2 media share many key metrics (content, demographic, geographic, etc.). In such cases, there would be a bigger portion of audience exposed to both media and we would have to adjust the equation to account for this (e.g. We would subtract 20 percent more than the random ab combination) as shown below.

$$a + b - 1.2(ab)$$

*Note that _the exact level of duplication has to come from a computer system. There is no way for a planner to estimate nonrandom duplication._

In general, the relationship between GRP and reach follows the below trend.
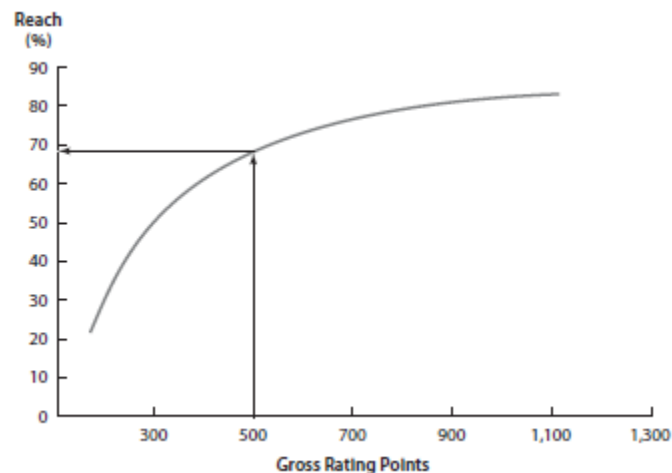
* Example shows relationship between GRPs and Reach in prime time (7pm – 10pm).

Frequency: This denotes, on average, how many times the TV spot is exposed to the same viewer. Frequency is simply calculated by using GRP divided by reach (GRP/Reach). In the above case, it would be 1.25 (2.5/2).

Rated TV Spot: GRP and reach is measured by an advertising service agency called Nielsen. If we put ads on a Nielsen rated channel, we'll know its GRP and reach. In this case, these TV spots are "rated".

Unrated TV Spot: However, Nielsen doesn't have a measure for all TV channels. Therefore, the rest TV spots are "unrated".

In general, GRP and Reach are strongly correlated with the below factors:

- Market (national/regional/local)
- Spot length
- Spot cost
- TV channel
- Time of the airing (daypart)
- Day of the airing (DOW and holiday)

This leads us to believe there might be a way to train/validate/test a machine learning model based on rated TV spots and then generate a GRP and reach estimate for unrated spots.

The table below provides an example of available data attributes.

| Market	TV Station	Length of Spot
Date Aired (EST)	Time Aired (EST)	Spot Cost
Program	GRP	Reach (available at daily/weekly level)

Deliverables:

1. A machine learning framework capable of estimating GRP for unrated TV spot
2. Estimate monthly reach based on GRP obtained from step 1
3. Presentation of model performance & findings
4. Code & script sharing

Project Option 2: TV Adstock Exploration with Multi-stimulus

The goal of this project is to investigate and explore ways to capture _adstock effect_ (consumer response – a click, visit, download, registration, subscription, etc.) and reasonably attribute them to _multiple media stimulus_ that most likely caused it. More specifically, we're looking for solutions to measure _how long_ the adstock will last at individual TV spot airing (or station level) and how to _decompose the overall response trend_ so that each individual media stimulus will get a fair share of efforts. Currently, we've implemented a non-machine learning framework to capture and allocate valid TV driven responses as the below example shows. But we do believe there is a potential to develop a machine learning algorithm that can return us a more robust outcome.
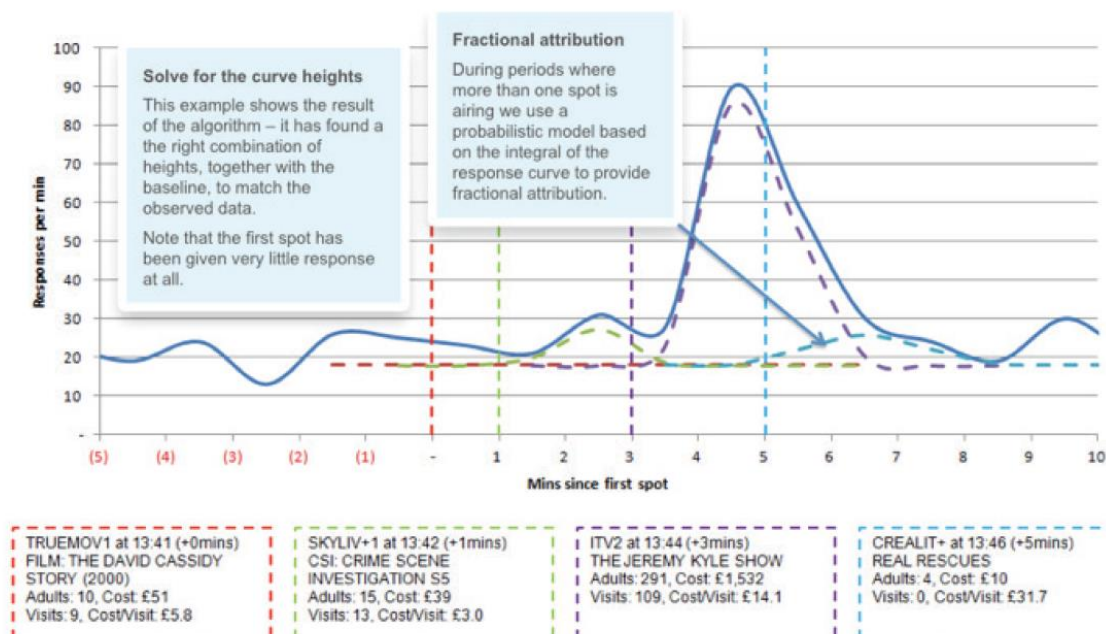
The tables below provide examples of data attributes.

TV Ad Data Example

Day of Week/WD vs. WE	Cost Per Spot	Spot Airing Time (HH:MM)
Station/Network	Access (Creative + Length)	Market (Local/Regional/National)
GRPs		

Consumer Response Data Example

Visits (Unique)/Users (New)	Registrations/Subscriptions	Geographic Info
Calls	App Downloads	Vanity URL
Orders	Revenue	Cost Per "X"

Deliverables:

1. A machine learning framework capable of capturing the true causal relationship between X spots and Y responses during period T (scalable). The output of this framework should answer the following questions:
    a. Which spots have stronger impact in driving responses? The spot/station impact should be measurable in a mathematical way
    b. What would be a good estimate of the adstock curve at the spot/station level and why?
2. Presentation of findings
3. Code & script sharing