Watching What You Watch:
An Analysis of the Predictive Value of Television Viewership Data in Targeted Online Advertising

by

Michael Christopher Eichert

An honors thesis submitted in partial fulfillment

of the requirements for the degree of

Bachelor of Science

Undergraduate College

Leonard N. Stern School of Business

New York University

May 2016

Professor Mart G. Subrahmanyam
Faculty Advisor

Professor Alvin Lieberman
Thesis Advisor

Abstract

The proliferation of 'big data' — defined here as the processes by which large amounts of information are recorded, processed, and stored by industrial computer systems — has enabled firms and marketers to employ myriad new techniques for identifying and targeting potential customers across interactive media platforms. With swaths of audience information available to them, firms are tasked with identifying which segments of data will be most useful to them in targeting ads to consumers. Often electing to purchase a portion of this data from third parties, the monetary value that media buyers, 'ad tech' firms, and their clients place on individual data segments, reflected to some degree in their market pricing, is directly proportional to the contribution of each in improving the accuracy of predictive modeling applications[1]. This paper utilizes methodology established in prior research literature to measure the predictive value of several data segments, highlighting in particular the broad category of television viewership and media consumption data. We find that these data segments, when integrated into a model that incorporates many other third-party data segment variables, indeed contribute to the model's ability to predict conversions, impacting the model's AUC score more significantly in some marketer industry cases than in others. Furthermore, we find that by adding the media segments to a model that relies exclusively on a base case of minimal demographic information, we in fact reduce the AUC score in the majority of our marketer test cases, while we improve the model's lift at a 10% threshold in three of the five cases.

[1] As outlined in *Bigger is Better, but at What Cost?,* "Naturally, data providers should be rewarded proportionately to the particular data's ability to effect positive change" (1).

Acknowledgements

I would like to thank Professor Lieberman for his influence and support throughout this project and for broadening my understanding of and appreciation for the entertainment industry throughout my four years at Stern. I would also like to thank Professor Subrahmanyam for giving me the opportunity to engage in this project and the Honors Program — I have learned so much over the course of this experience, and I appreciate his efforts to promote rigorous academic research at Stern. Finally, I would like to thank Claudia Perlich and Jessica Clark for introducing me to the world of data science and digital marketing. Without their guidance, support, and generosity, this project would not have been possible.

Table of Contents

Introduction

The technology press, guilty as it is perhaps of the occasional hyperbolic flourish, has taken to discussing the widespread proliferation of "big data"-enabled enterprise optimization applications in terms akin to other landmark discoveries made throughout the history of the human race. As a species that simply "can't seem to escape big data," Samuel Arbesman wrote in *Wired* in 2013, "We have more data inputs, storage, and computing resources than ever, so *Homo sapiens* naturally does what it has always done when given new tools: It goes even bigger, higher, and bolder[2]."

Indeed, the degree to which technology-enabled large-scale data collection and modeling have influenced and transformed business practices across industries has only grown since 2013[3], and perhaps nowhere has their impact been more conspicuous than in television and marketing. Data collection has long influenced the interplay between these two industries, but now, as new tools for audience measurement become available and the opportunities to incorporate data-informed decision-making enumerate, the need arrises to identify applications in which new data may be most useful in generating value for firms, and which specific data segments may be leveraged most productively to generate actionable insights in both industries.

Informing this discussion requires us to first examine the audience measurement methods currently employed by the television industry, and the industry use cases for this data. Understanding its function within the television industry, we can then analyze media

---

[2] Arbesman, Samuel. "Stop Hyping Big Data and Start Paying Attention to 'Long Data'." *Wired.com*. Conde Nast Digital, n.d. Web. 04 May 2016.

[3] An October, 2014 study conducted by Accenture and GE reported that "87% of enterprises believe Big Data analytics will redefine the competitive landscape of their industries within the next three years."

consumption data's usefulness to marketers, attempting to gain insight into how well modern advertising firms are able to leverage this data to target potential customers. Ultimately, we aim to uncover the degree to which this and other data segment categories may improve the predictive models utilized by ad firms, and how their usefulness may be reflected in the data's value.

The Current State of Television & Audience Measurement

The television industry is in the midst of a pivotal period of accelerated transformation, with firms new and old rethinking every aspect of the traditional business model from content programming and financing to distribution, marketing, and viewership. Industry leading content giants like Viacom face slumping ratings and a syphoning of ad dollars from traditional television networks to online video and social media platforms — Viacom themselves saw a domestic ad sales decline of 9% in Q3 2015.[4] Online video consumption, by comparison, only continues to grow more rapidly, with services like YouTube and Netflix commanding around 40% of all internet traffic in the UK at a given moment during the primetime television daypart.[5] In short, there are only so many hours in a day, and people seem to be spending fewer and fewer of them watching traditional cable television.

Many industry practitioners attribute these trends and market fragmentation to technology-enabled shifts in media consumption habits, and millennial "cord-cutting," replacing traditional cable subscriptions with other online streaming alternatives.[6] It is not a matter of people

[4] Viacom domestic ad sales tumbled for four consecutive quarters leading up to Nov. 2015, *Variety,* Nov. 2015

[5] The Guardian also reports that BBC's online iPlayer platform receives 7 million requests daily, *The Guardian*, Aug. 2014

[6] The *NYTimes* reports that 25% of millennials (ages 18-34) living without children do not have cable connected to their televisions.

watching less TV, but rather that they now have more content and more viewing options than ever before. Compounding their troubles, networks are now forced to compete with tech companies in Silicon Valley and New York for quality programming and advertising dollars — two areas in which they have historically had little need for technological innovation. For decades, the television industry has informed its programming decisions and set its advertising prices based on Nielsen ratings — the system established in the 1950s to measure household television viewership by installing meters in a sample of households that record what these households are watching.[7] But as advertising technology progressed with the advent of the internet and targeted online advertising, advertisers started to question the accuracy and bemoan the limitations of the Nielsen standard upon which the television ad sales business is built. A convergence of all these trends leaves the television industry in its current state of uncertainty, as it begins to look for answers in new places.

To combat declining revenues, reach new audiences, and ensure ad sales returns on increased investment in audience engagement, cable companies and television networks have started to adopt strategies employed by online streaming platforms. Netflix, often looked to by both tech firms as well as traditional media companies as the preeminent 21st century content company, has released extensive documentation about its unique approach to programming. Utilizing big data analytics and predictive modeling, Netflix is able to make programming decisions (including casting, plot point, and genre choices) based on tracked audience preferences and viewing behaviors such as drop-off rates and sticking points (the point in a season at which a

[7] "How Do Television Ratings Work?" *HowStuffWorks*. N.p., 28 July 2000. Web. 04 May 2016.

viewer decides to binge on the remainder of the season).[8] Similar audience measurement solutions are also available on other streaming platforms and internet-connected televisions, and here, the collected data is being used for ad sales. Firms like ComScore and Rentrak attempt to aggregate cross-platform viewership metrics and present them to advertisers seeking to purchase ad inventory 'programmatically,' trusting their media buying decisions to data-driven audience targeting algorithms.[9] Even the federal government seems poised to enable data collection firms to widen their nets, as impending FCC recommendations promote legislation that would force cable companies to unlock their content for streaming to any third-party devices using "open standards" — a move that would undoubtedly enable myriad new television audience measurement solutions.[10]

So as the television industry begins to rethink its approach to data collection and how networks use data both to make their own business decisions and to provide more detailed audience information to marketers, questions arise as to how marketers are in fact able to utilize this data in their own business models, and the degree to which they can derive value from it.

Overview of Targeted Online Advertising

In 2014, global online advertising revenues reached a record high of $135.42B, with that number forecasted to grow to $239.87B by 2019[11]. Accordingly, the number of players in the digital

[8] Person, and Adam Clark Estes. "Netflix Can Pinpoint Exactly When People Get Addicted to Shows." *Gizmodo*. Gawker Media, 24 Sept. 2015. Web. 04 May 2016.
Carr, David. "Giving Viewers What They Want." *The New York Times*. The New York Times, 24 Feb. 2013. Web. 04 May 2016.

[9] Steinberg, Brian. "Viacom Bets on Big Data to Boost Its Revenues." *Variety*. Variety, 03 Nov. 2015. Web. 04 May 2016.

[10] Newman, Jared. "The FCC Wants to Blow up the Cable Box. Here's What Its Proposal Will (and Won't) Do." *TechHive*. TechHive, n.d. Web. 04 May 2016.

[11] PwC. Global Entertainment and Media Outlook 2015-2019: Internet Advertising: Advertising Segment (n.d.): n. pag. PwC. Web.

advertising space has multiplied significantly over the course of the past few years, further complicating the advertising ecosystem itself, and lending tremendous complexity and sophistication to the means by which advertisers ultimately reach their audiences. Beyond sheer audience reach[12], internet display advertising is attractive to so many marketers due largely to the level of control, measurement, and efficiency afforded to them by targeting techniques, and the real-time bidding (RTB) process.

At this stage in the industry's development, it is commonplace for advertisers to hire third-parties to serve as media buying agents (we will refer to them as "ad agencies"), leveraging proprietary software and predictive models to target potential customers browsing online ("viewers"), and to engage in the bidding process on behalf of clients ("advertisers"). In a nutshell, the RTB process works as follows: In a market with multiple advertisers, each with an agency contracted to run certain campaigns for them, agencies bid in a "real-time auction" on behalf of their clients for advertising space on various websites ("publishers") across the internet (e.g. NBCNews.com, NYTimes.com, etc.).[13] Certain advertising inventory on these sites (the spatial areas where ads appear scattered throughout them) is purchased by third-party ad exchanges, which in turn host real-time auctions where each ad slot is subsequently sold to the agency that bid the highest on behalf of one of its advertisers. It is crucial to note that this process occurs instantaneously when a viewer navigates to a publisher's webpage where exchange-owned ad inventory is present and the page loads in the viewer's browser. Each ad slot may also be unique to the individual viewer who is viewing that web page at a given time. Agencies, which may represent several different

---

[12] In 2012, YouTube's homepage roadblock was estimated to average 70 million unique impressions daily. By means of comparison, around 114 million people watched the Super Bowl in 2015 (Hollywood Reporter).

[13] Here we'll reference the Aziz cookie paper

advertisers at once, will bid on behalf of whichever of its clients's ads will be most likely to convert into a purchase by the viewer viewing it — in other words, the agency will present the viewer with whichever ad it thinks the viewer will be most inclined to click on, and subsequently make a purchase. For this reason, the price an agency is willing to bid for a given ad impression depends on the viewer viewing the ad, and his or her likelihood to purchase the offer contained in the ad which the agency chose to display.

For this all to happen instantly when a viewer navigates to a webpage, agencies must be equipped with technology that rapidly evaluates the likelihood that a given impression will convert, and then calculate the price that it will bid for that impression. It goes without saying that firms want to target viewers with the highest probability of converting on an ad, and calculating this requires the agency to know something about the viewer. Exactly how each firm goes about calculating a conversion probability and bid price, and the specific viewer data that it uses to do so vary from firm to firm, but frequently agencies will use a combination of their own proprietary audience data and third-party data that they purchase from other firms. When an agency receives a bid request from an ad exchange, it also receives a unique cookie identifier for the viewer currently viewing the page.

Cookies, the common term for a few lines of code embedded in a website to track viewer behavior, may provide an ad agency with information ranging from a viewer's browsing and purchase history to his or her location and device model. The unique identifier also allows the agency to instantly associate the viewer with sizable amounts of proprietary and third-party audience data segments stored on the agency's databases. These segments can include behavioral information (i.e. a viewer's television viewership or travel preferences based on his browsing history), sociodemographic information (age, gender, household income, family size, education, etc.), geographic information about where the viewer lives or is located, technical information about his device or browser, and lifestyle profile information (amalgamations of several other data segments combined to present a more holistic understanding of the viewer).[14]

This data, though, comes at a price, and agencies and ad tech firms are challenged to maximize their investment in either collecting or purchasing audience information. As data scientists at New York-based ad tech firm Dstillery write, "most data vendors sell data at a fixed price and leave it to the buyers to determine if the data holds enough value to justify that price." [15]

It stands to reason that firms purchasing audience data would value that data proportionally to its usefulness to them in predicting impression conversions and pricing bids accordingly — the more a given data segment improves the accuracy of an agency's predictive models, the more valuable that data is to the firm. But exactly how third-party segments may work in conjunction with one another and with the firm's own proprietary data segments to improve these models predictive function is a topic of ongoing academic and practitioner research.

Literature Review

Methodologies for evaluating various data segments based on their effectiveness in improving predictive model accuracy have been presented. Concerned primarily with internet user privacy, Aziz and Telang (2015) use an area under the receiving operator characteristic cure (AUC)

---

[14] Oracle BlueKai. *Oracle Data Cloud Data Directory*. N.p.: Oracle BlueKai, 2016. Web.

[15] Dalessandro, Brian, Claudia Perlich, and Troy Raeder. "Bigger Is Better, but at What Cost?Estimating the Economic Value of Incremental Data Assets." *Big Data* 2.2 (2014): 87-96. Web.

analysis (described below) to show that incrementally adding increasingly "privacy-intrusive" data "increases the accuracy of prediction of purchases, but at a decreasing rate." [16] This study focuses almost exclusively on behavioral data tracked by cookies in online shopping platforms, evaluating the degree to which knowing a viewer's browsing behavior on such platforms aids the ad agency in predicting a purchase. It stops short of evaluating the effectiveness of other data segments in predicting outcomes for marketers in various different industries, and does not present an effective way to translate model improvement to a measure of the data's economic value.

Dalessandro, Perlich, and Raeder (2014) address both these issues by using a Precision/Lift analysis (described below) to measure the incremental value of targeting viewers in different segments across campaigns in 10 different marketer industries[17]. The lift metric, for reasons described below, allows for easily reflecting improved accuracy in monetary terms, enabling agencies to easily determine an optimal price for data they are purchasing. Our paper builds upon this body of research by adopting a similar methodology to evaluate the effectiveness of a broad category of third-party television viewership and media consumption behavioral data segments in improving predictive model accuracy in several marketer industry cases when compared with other third-party data segments.

[16] Aziz, Arslan, and Rahul Telang. *What Is a Cookie Worth?* Rep. Heinz College, Carnegie Mellon University, Mar. 2016. Web.

[17] Dalessandro, Brian, Claudia Perlich, and Troy Raeder. "Bigger Is Better, but at What Cost?Estimating the Economic Value of Incremental Data Assets." *Big Data* 2.2 (2014): 87-96. Web.

Data

Collaborating with a digital advertising technology ("ad tech") firm, we were given access to a dataset consisting of approximately 415,000 unique impressions over a single-day period, and another consisting of approximately 10.5 million ad conversions for various marketers and offers.

Relevant information contained within each row of the **impression dataset** consists of one unique **client_id** representing the unique, individual viewer, followed by a number of **segment_ids**, each representing a categorical piece of data about that viewer. Some of these data segments are proprietary to the ad tech firm, while others are tracked by third-parties (such as Oracle's Bluekai and others) and purchased by the firm to supplement their own data for use in predictive models.

The underlying data segments that the segment_ids proprietary to the ad tech firm represent are unknown to us. The remaining third-party data segments range from information on viewers' purchase habits, income information, and media consumption preferences to their travel destinations, transportation preferences, and likely entertainment activities. (Full list of relevant segments in **Figure 4** below)

Each row of the **conversions dataset** contains a **client_id**, a **marketer_id** corresponding to one of the ad firm's advertiser clients, and an offer_id corresponding to a particular campaign. For the purpose of this study, we ignored individual offer_ids and focused instead on the marketers themselves. For a chosen marketer_id, if the corresponding client_id also appears in the impression dataset, we conclude that the impression (with the attached data segments for that individual viewer) converted for that marketer's offer.

We aggregated this data according to marketer_id, first choosing a marketer, then taking a "balanced sample" of 5000 random impressions that did convert for this marketer, and 5000 random impressions that did not, giving us a conversion base rate of 50%. We collected samples on 5 marketers, each in a different industry, to use in 5 industry case studies.

Methodology, Analysis & Results

Our aim is to understand the degree to which television viewership and media consumption data (we will refer to it as "media data") either improves or worsens our model's ability to predict a binary outcome: Did an impression convert or not?

We prepared each of the 5 datasets for logistic regression modeling by structuring them as follows **(Figure 1)**: Each row, representing a single unique impression (10,000 in total) contains the unique client_id in Column 1, followed by a large number *c* columns (which varied with the particular impressions sampled) each with a data segment_id at the header and binary dummy variables in each row indicating the presence or not of that data segment in each impression. A final column denotes whether or not the impression converted, serving as our binary dependent variable.

Figure 1: Data Structure Illustration

client_id	segment_1 (ex. "Frequent Traveler")	segment_2 (ex. "Past Purchases: Verizon Wireless")	…	segment_c-1 (ex. "Demographic, 18-25)	segment_c (ex. "TV Affinity: HBO")	converted?
1111	0	0		1	1	0
2222	0	1		1	0	1
3333	0	1		0	0	0
4444	1	1		0	1	0
5555	1	0		0	0	1

Figure 1: Table depicting the structure of the data frame used to test each of the five marketer datasets. The first column (client_id) shows the unique client_id associated with an individual impression (each row). The middle columns each indicate the presence (or lack) of individual third-party data segments (indicated by a 1 or 0 respectively). The last column indicates whether or not that impression converted for that marketer's ad. We made and tested 5 of these data frames in total, one for each marketer.

With the data structured appropriately, we conduct our analysis in three stages, running tests on all five marketer datasets in each. In the first stage, we estimate our model's accuracy with all the third-party data segments present, then measure the impact of removing the media data on our model's performance. In the second, we incrementally add and replace different categories of data segments to a base set of demographic data segments in order to test how our model reacts to each individually. Lastly, we repeat this test using a different classification metric in order to more easily translate our results for economically realistic applications (i.e. known budget constraints).

1. AUC Test - Removing Media Data

For our first test we use the area under the receiver operator curve (AUC) as our classification metric. The AUC score provides us with a statistically viable measurement of our model's accuracy, returning the probability with which our model is able to correctly rank a randomly chosen instance with a positive outcome above an instance with a negative outcome. Plotting our model's True Positive Rate (TPR) — the rate at which it predicts correctly that instances that do convert will convert — and False Positive Rate (FPR) — the rate at which it predicts incorrectly that instances that do not convert will convert (Type 1 Error) — as points on a coordinate plane for every possible threshold value *k* forms the receiver operator curve (ROC). This metric is useful in situations where a threshold is unknown or unknowable — in marketing applications, this could be when a budget has yet to be determined, and the size of the population in question is unknown.[18]

To improve our measurement ability, we first isolate only the third-party segments, removing the unknown proprietary segments from our model. We then take the AUC score of the model with all the available third-party segments present, followed by that of the model with only the media segments removed.

Figure 2: Model With and Without Media Segments AUC Scores

Marketer	All Third-Party Segments (AUC)	Third-Party - No Media	% Change	# of Segments (columns) Before	# of (columns) After	% of 3rd Party Segments Are Media
Travel/ Booking	0.53704	0.53532	0.3203%	826	760	8.0%
Streaming Video/TV	0.62313	0.62172	0.2263%	831	759	8.7%
Auto	0.55094	0.54199	1.6245%	808	742	8.2%
Specialty Retail	0.546194168	0.539422011	1.2399%	817	751	8.1%
Credit Card	0.540073272	0.547018508	-1.2860%	825	750	9.1%

Figure 2: Table depicting the results of our first AUC test. Column 1 (marketer) contains the five different marketer industries we chose from. Col 2. contains the AUC scores when all third-party segments are

[18] Dalessandro, Brian, Claudia Perlich, and Troy Raeder. "Bigger Is Better, but at What Cost?Estimating the Economic Value of Incremental Data Assets." *Big Data* 2.2 (2014): 87-96. Web.

present, while Col. 3 contains the AUC scores when the media segments are removed. Col. 4 shows us the change in AUC score between Cols. 2 and 3. Cols. 5, 6, and 7 contain respectively the number of total third-party segments (columns) in each marketer case, the number of segments remaining after the media segments were removed, and the percentage of third-party segments that are media segments.

We find first and foremost that removing the proprietary data segments reduces the AUC score of our base case 'third-party only' model down to near 0.5 (this being a model that would be equally likely to rank any randomly-chosen positive instance above any randomly-chosen negative instance as it would be to do the opposite — in other words, a model that tells us nothing). This is likely due to the scarcity of information contained within the third-party data segments alone, without including any of the behavioral information contained within the proprietary data segments. Examining the results of our test, we find that removing the media third-party segments caused a decline, albeit a minor one, in four out of the five industries tested. This may indicate that the media segments, when integrated into a model with many other third-party segments present, are indeed valuable to the marketer as they do slightly improve the likelihood that the model will correctly predict conversion outcomes.

2. AUC Test - Replacing Data Segment Categories

Our second test also uses the AUC score as a metric, this time measuring the predictive effectiveness of individual data segment categories when added to a base case of standard demographic and technical data segments. Our base case segments include: viewer age, gender, broad geography (country), and device OS. Onto this we add and then replace each of the following categories of data segments:

Figure 3: Data Segment Categories & Details

Media	Media consumption data including: cable providers, television network and show preferences, online streaming habits, etc.
Automotive	Automotive interest or past purchase data
Travel	Travel or travel interest data: flight destinations, departure dates, preferred hotels, preferred airlines, etc.
Financial	Financial information: Estimated HHI, credit card companies, investment preferences
Purchase	Past purchases or predicted purchase interests: cars, home goods, entertainment, ISPs, etc.

Figure 3: Table containing the five data segment categories and descriptions of the kinds of data contained within each.

Here we find that by adding the media segments to our base case, we in fact decrease the predictive accuracy of our model by a small amount in four out of the five marketer cases. The scores produced by this test are so small that additional testing such as a 10-fold cross-validation (which would produce 10 AUC estimates for each test case) may be useful in improving statistical significance.

Figure 4: % Change in AUC over Base Case

	Age/Profile (Base) - AUC	Add Media	% Change	Add Auto	% Change	Add Travel	% Change	Add Financial	% Change	Add Purchase History	% Change
Travel/Booking	0.507458864	0.50229276	-1.018%	0.50038966	-1.393%	0.52133803	2.735%	0.513054013	1.103%	0.536510779	5.725%
Streaming Video/TV	0.570275946	0.56466284	-0.984%	0.57220124	0.338%	0.58301438	2.234%	0.570850016	0.101%	0.586248369	2.801%
Auto	0.530989956	0.53453246	0.667%	0.53975665	1.651%	0.52789942	-0.582%	0.528672052	-0.437%	0.544282302	2.503%
Specialty Retail	0.541953665	0.54113509	-0.151%	0.53940175	-0.471%	0.53821597	-0.690%	0.538957277	-0.553%	0.540004988	-0.360%
Credit Card	0.52739552	0.51971672	-1.456%	0.52696809	-0.081%	0.53771925	1.957%	0.535343171	1.507%	0.542168159	2.801%

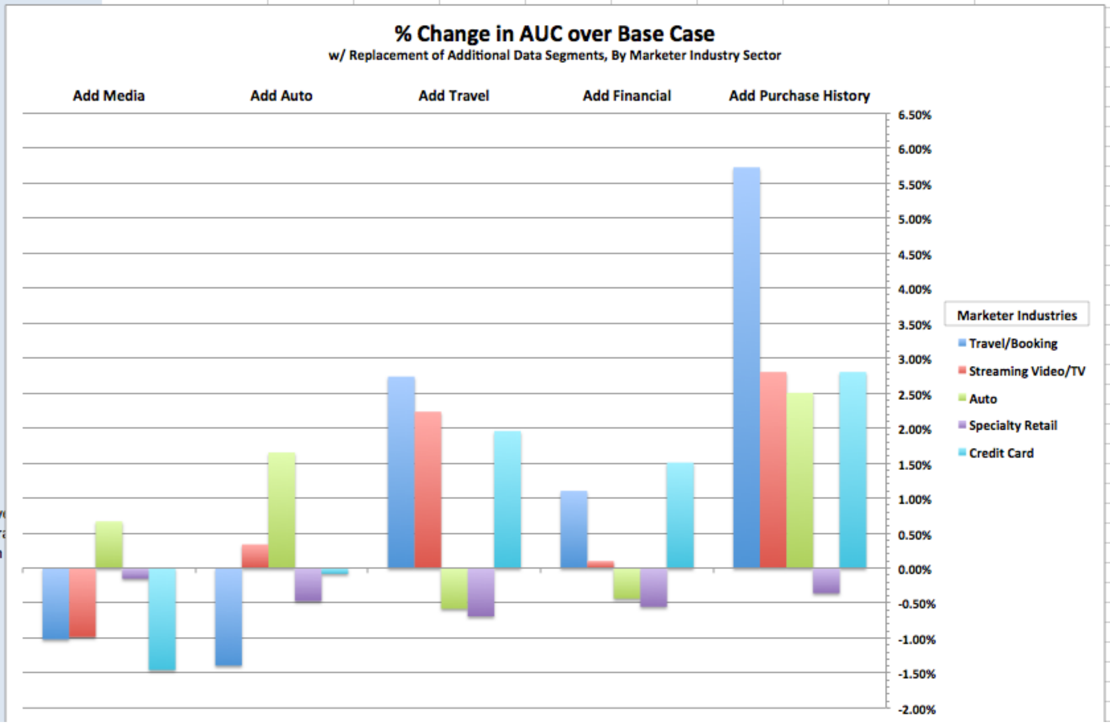


Figure 4 (larger picture in APPENDIX F): Results of the second AUC test, measuring AUC score before and after adding each data segment category. Column 2 (Age/Profile (Base) - AUC) contains the base case AUC scores for each marketer. Columns 3, 5, 7, 9, and 11 show the effect on AUC score in each case by adding that respective data segment category to the base case. Cols. 4, 6, 8, 10, and 12 contain the % change in AUC score achieved by adding each category. The bar chart compares the % changes in AUC score achieved for each marketer by adding each data segment.

That said, it does give some directional indication of which data categories improve the model for certain industry marketers, many of which make sense intuitively. Automotive interests and purchase behavior is a comparatively strong indicator for conversions on ads from automakers, for example, while travel data tends to improve conversion predictability on ads for online travel and booking services. This analysis also points to estimated past purchase history and purchase interest data as the strongest indicator for four out of the five marketers by a substantial margin.

3. Precision/Lift Test - Replacing Data Segment Categories

In our third test, we use Precision and Lift metrics to evaluate model improvement under the same conditions we set in Test #2. Precision at a given threshold *k* is the percentage of predicted positives above this threshold that are in fact positives:

$$Pre_k = TP \cdot (TP + FP)^{-1}$$

Where:
TP is the number of positives above *k*
FP is the number of negatives above *k*

Lift at *k*, accordingly, is defined as Pre_k / TP_{Base} where TP_{Base} represents the base number of positives in the entire dataset.

This metric enables us to select a threshold *k* and measure the change in our rate of positives above *k.* Doing so is useful in marketing applications when the marketer is working under fixed budget constraints, and is forced to target only the top *k* percent of instances most likely to convert. We chose a threshold *k* of 0.1 or 10% to simulate a plausible scenario under which a marketer only wishes to target the top 10% of instances most likely to convert, and therefore cares only about how altering her model would impact its ability to predict positives above this threshold.

Figure 5: % Change in Lift over Base Case

	Threshold k	Age/Profile (Base) - Lift	Add Media	% Change	Add Auto	% Change	Add Travel	% Change	Add Financia	% Change	Add Purchas	% Change
Travel/Booking	10%	0.922667395	1.05821472	14.691%	0.95589539	3.601%	1.13739546	23.273%	0.96122946	4.179%	1.11495299	20.840%
Streaming Video/TV	10%	1.178686164	1.26017699	6.914%	1.15957374	-1.622%	1.13590126	-3.630%	1.16994189	-0.742%	1.25237417	6.252%
Auto	10%	0.947050601	1.03856749	9.663%	1.28139905	35.304%	0.91442308	-3.445%	1.03623628	9.417%	1.06964551	12.945%
Specialty Retail	10%	1.151289683	1.0632151	-7.650%	1.09086473	-5.248%	0.97280064	-15.503%	1.15128968	0.000%	1.1099292	-3.593%
Credit Card	10%	1.067652807	1.06765281	0.000%	1.45647586	36.418%	1.07450867	0.642%	1.27424062	19.350%	1.03703704	-2.868%

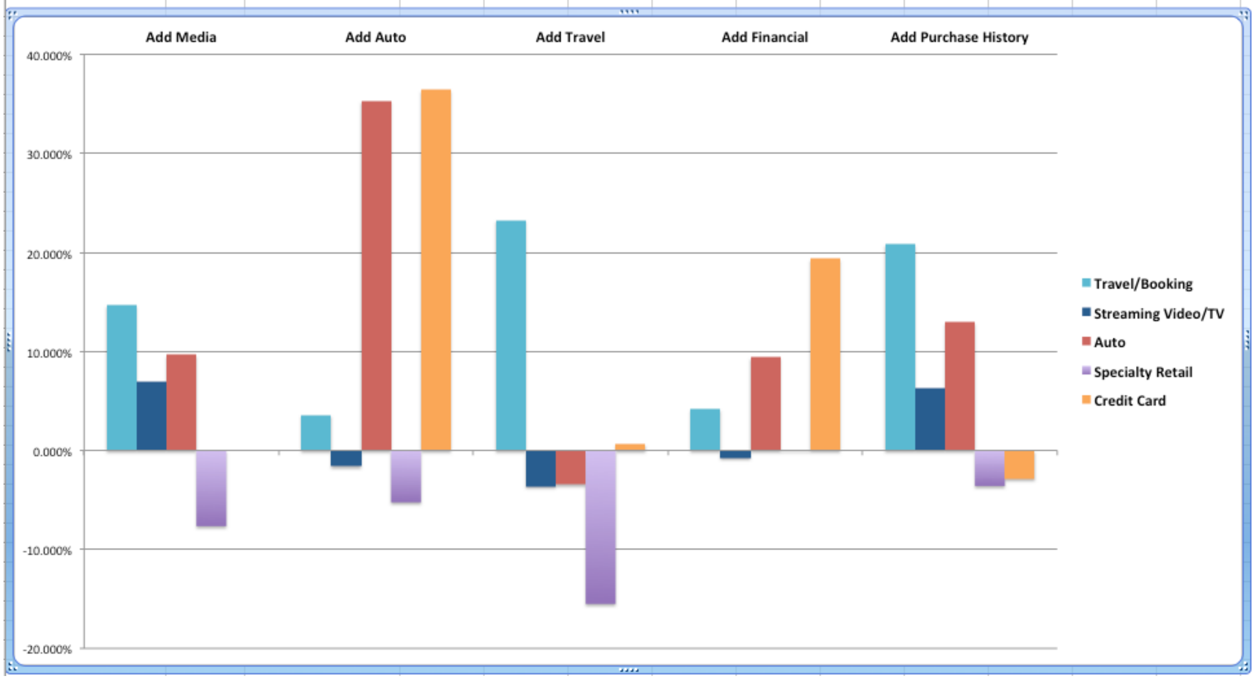


Figure 5 (larger picture in APPENDIX G): Results of the Precision/lift test, measuring lift before and after adding each data segment category. Column 2 shows the threshold k *chosen for each case - we used 10% for each. Column 3 (Age/Profile (Base) - Lift) contains the base case lifts for each marketer. Columns 4, 6, 8, 10, and 12 show the effect on lift in each case by adding that respective data segment category to the base case. Cols. 5, 7, 9, 11, and 13 contain the % change in lift achieved by adding each category. The bar chart compares the % changes in lift achieved for each marketer by adding each data segment.*

Here, our model produces results that are similar in direction and magnitude to those produced in Test #2 in some marketer/segment pairs, but not in others. Just as in the AUC test, we see that the presence of automotive segments and travel segments are relatively strong indicators of an impression converting on an automotive ad and a travel service ad respectively. Here, however, we also see the addition of purchase history data segments does not improve lift as much as it did AUC when compared to other segments.

Notable is the change in the media segments' impact on the model from Test #2. We observe here an improvement in the model's lift for three of the five marketer segments, while the Credit Card segment remains flat, and Specialty Retail declines.

Conclusions & Future Work

We have put forth three methodologies for evaluating the degree to which media consumption and television viewership data segments impact predictive models in online marketing applications in the context of other third-party data segments. From the results above, we can hypothesize that the media data, when integrated into a model containing other third-party segment data variables as well, does indeed contribute to the model's ability to predict conversions, acting more strongly as indicators in some marketer industry cases than in others.

That said, when the media data segments are incorporated into a model that relies on them and basic demographic data exclusively, our results are mixed. In the AUC test (Test #2), we find that adding the media segments to our base segments reduces the model's AUC score in four of the five marketer industries tested, while in our lift test (Test #3), the media segments improved the model's lift for the online travel service, the streaming website, and the automaker ads, while decreasing it for the specialty retailer ad and leaving the model's lift unchanged in the credit card scenario. Because the changes in AUC score produced in Test #2 are so small, conducting a more thorough AUC test such as the 10-fold cross-validation mentioned above may produce more statistically significant results, or even alter the results' directionality. That said, it is also possible that indeed the media segments improved the model's precision for instances with probabilities above a threshold of 10%, while impacting the model as a whole (at any threshold *k*) differently.

It is quite possible, due in part to the limitations of the data provided to us and in part to the complexity of the model (the number of variables), that during our tests, the model was subject to overfitting — reacting badly to noise that may have impacted our results. In order to prevent overfitting and any variance error, in future work, the model should be subjected to more rigorous testing including a 10-fold cross-validation, and our methodology should be applied to larger datasets.

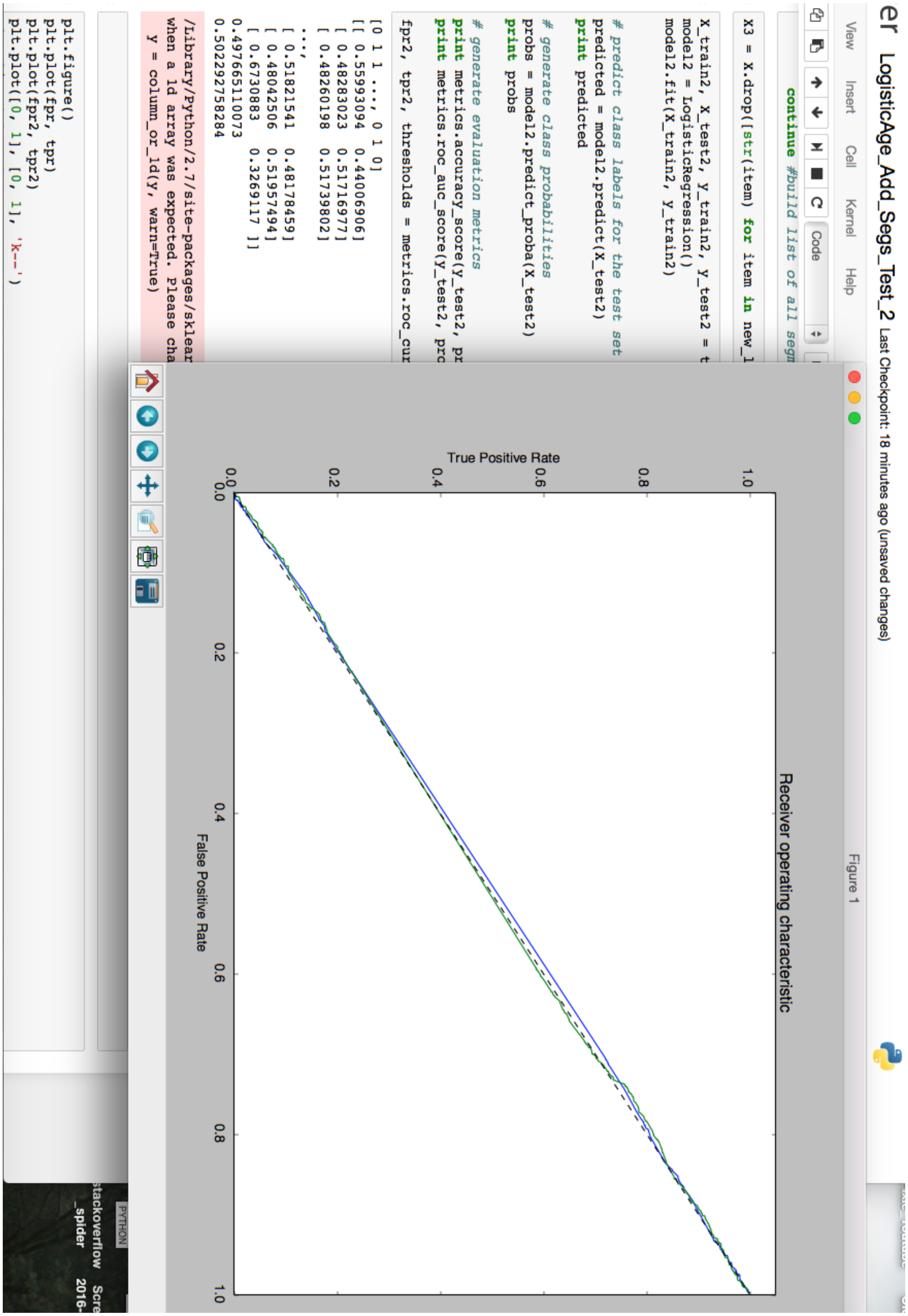
Improving upon this work further, future studies may average the results of models run on campaigns for many advertisers in each industry, rather than taking one for each. This would give a more accurate representation of how the model performs for a given *industry* rather than a given advertiser in that industry. The same result could be accomplished by averaging many campaigns rather than advertisers, sampling the test dataset according to offer_id. Furthermore, future models may control for media segments as a percentage of total third-party segments present.

Implications

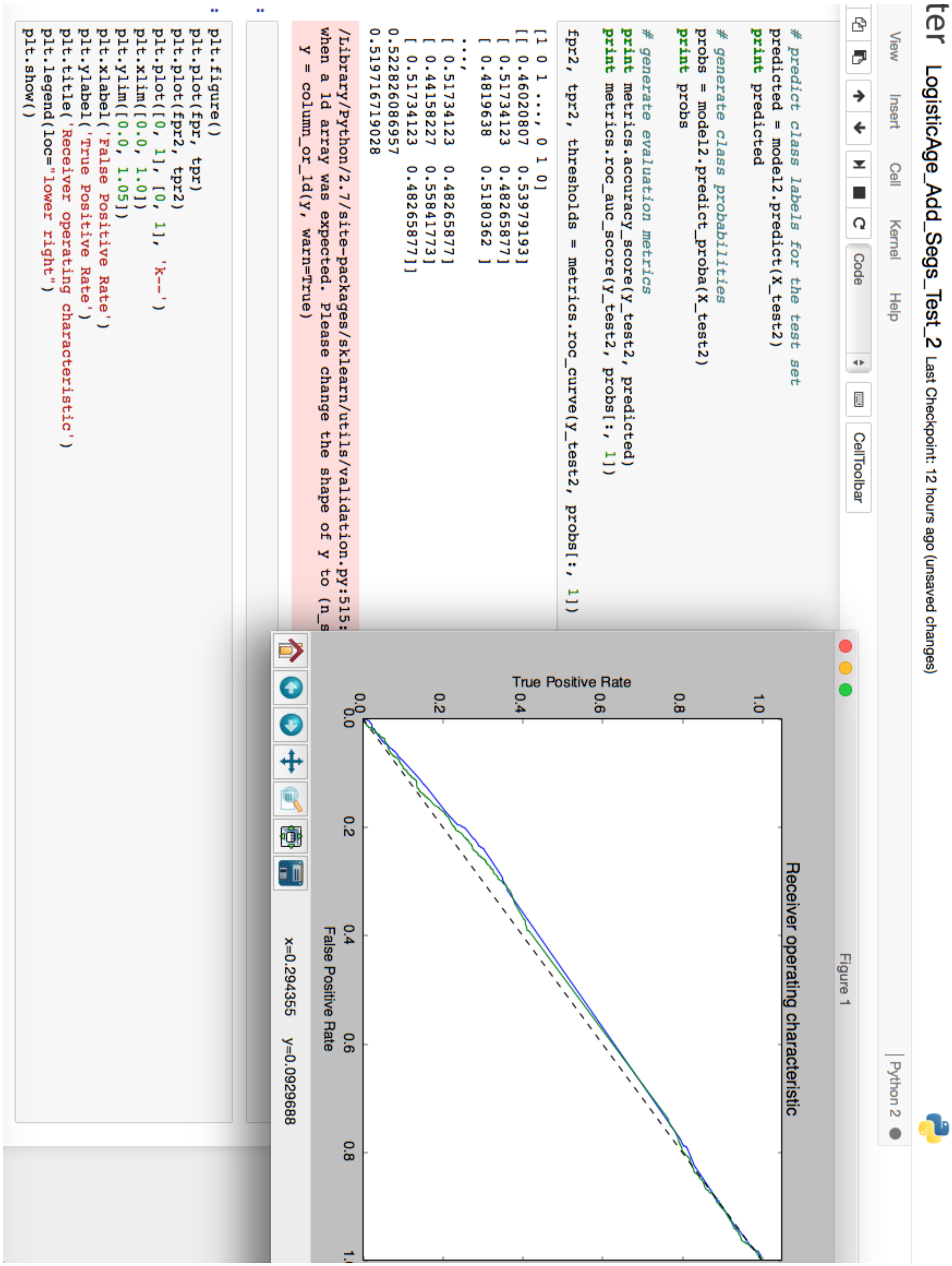
With this study, our aim was to inform the current industry discussion surrounding the predictive value of media consumption data, and to contribute to the growing body of academic research examining targeted digital advertising. We hope that marketers and those engaging in audience measurement are able to look to our methodology as a means by which to better understand how leveraging certain data segments impacts predictive models, and that they may use this knowledge in order to make more informed pricing and investment decisions.

Certainly, as content companies and television networks continue to innovate, the use cases for media consumption data will only multiply, and as the Dstillery scientists point out in their paper, data itself "has no intrinsic value, and the estimate of its value is only as good as the abilities of the modeler undertaking this exercise." That said, for many in the television industry, it pays tremendously to be able to understand the ways in which marketers derive value from audience data, and the tools used to translate predictive value to pricing decisions. At a minimum, the better informed each party involved in the audience data value chain is, the more efficient this market becomes, and the utility derived from the data may be maximized. Looking forward, we hope that this research and the body of knowledge of which it is a part mark the initial steps in a more general rethinking of the relationship between content companies (both digital and not) and their audiences — one that sees both parties forthrightly acknowledge the value that is extracted from information about the latter, and increased transparency about the process as a whole.

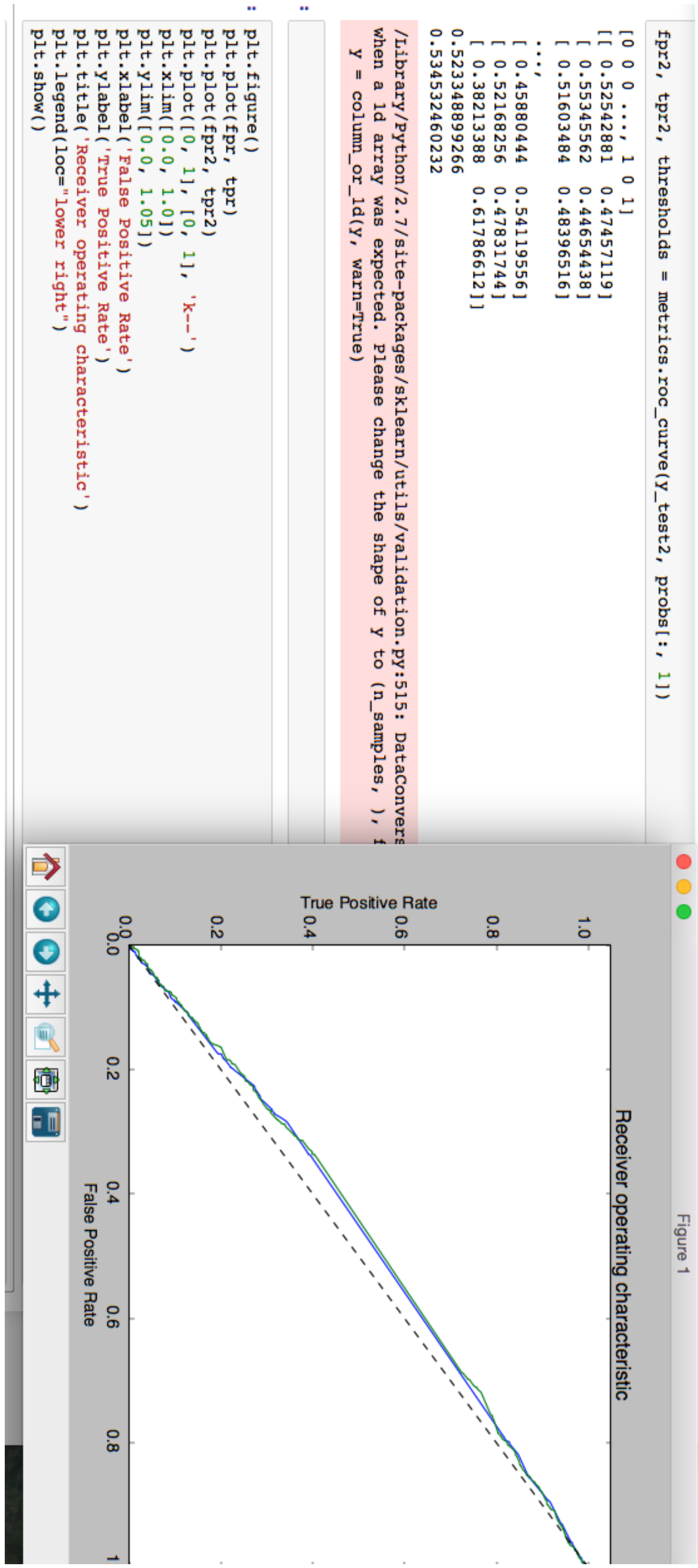
APPENDIX A - Test #2 ROC Curves & AUC Score, Travel Service, Base + Media Segments



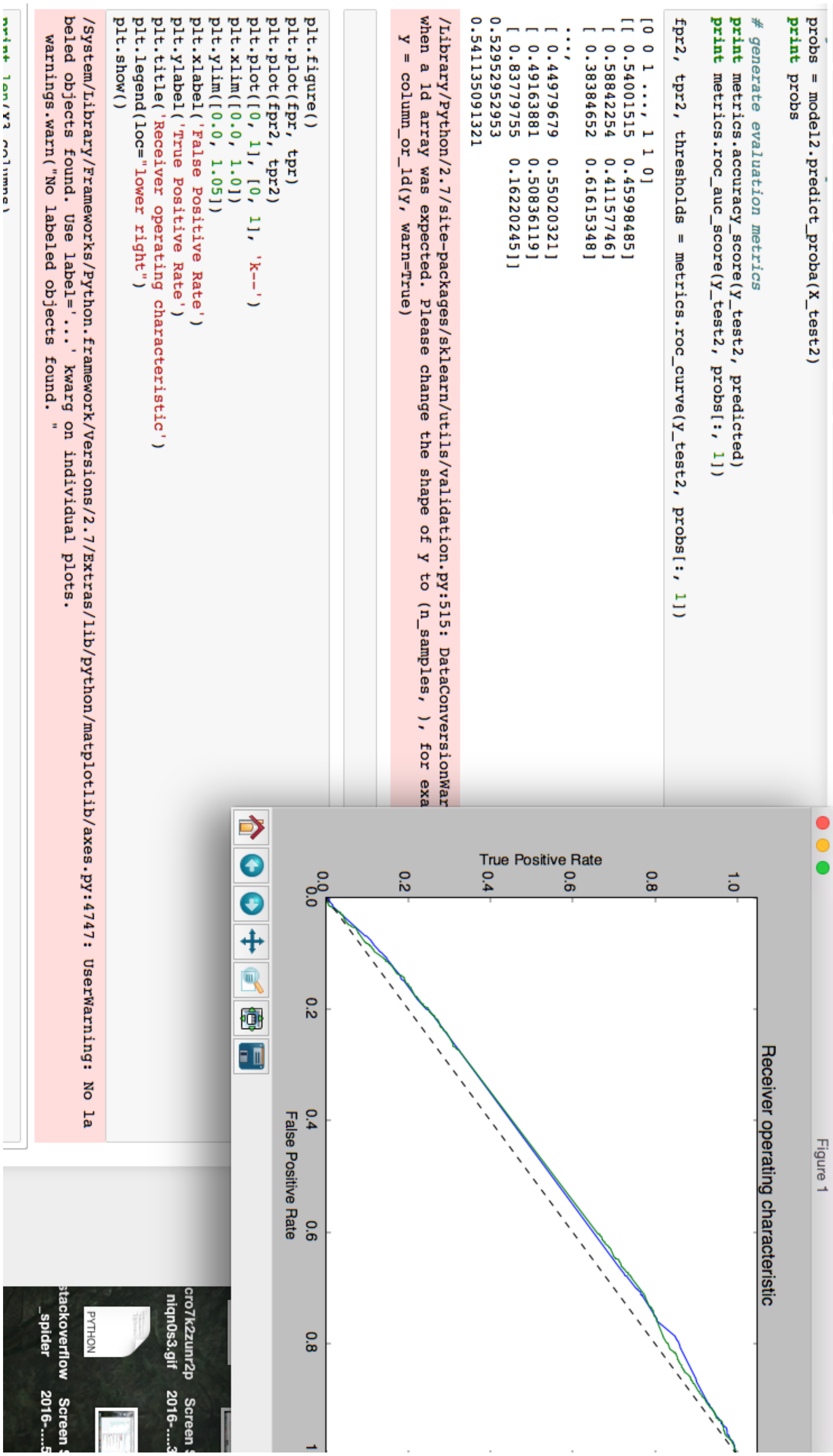
APPENDIX B - Test #2 ROC Curves & AUC Score, Credit Card, Base + Media Segments



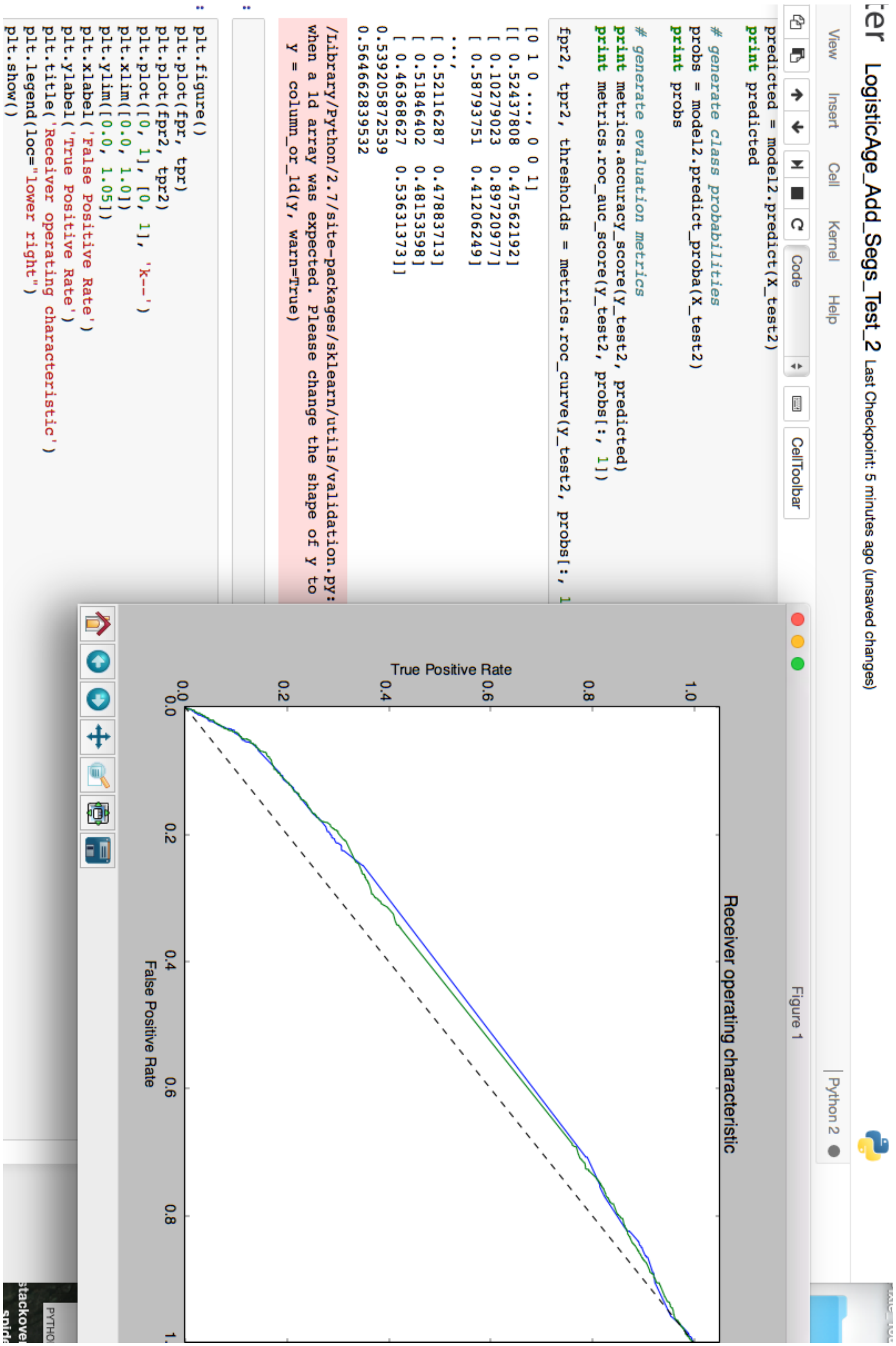
APPENDIX C - Test #2 ROC Curves & AUC Score, Automobile, Base + Media Segments



APPENDIX D - Test #2 ROC Curves & AUC Score, Specialty Retail, Base + Media

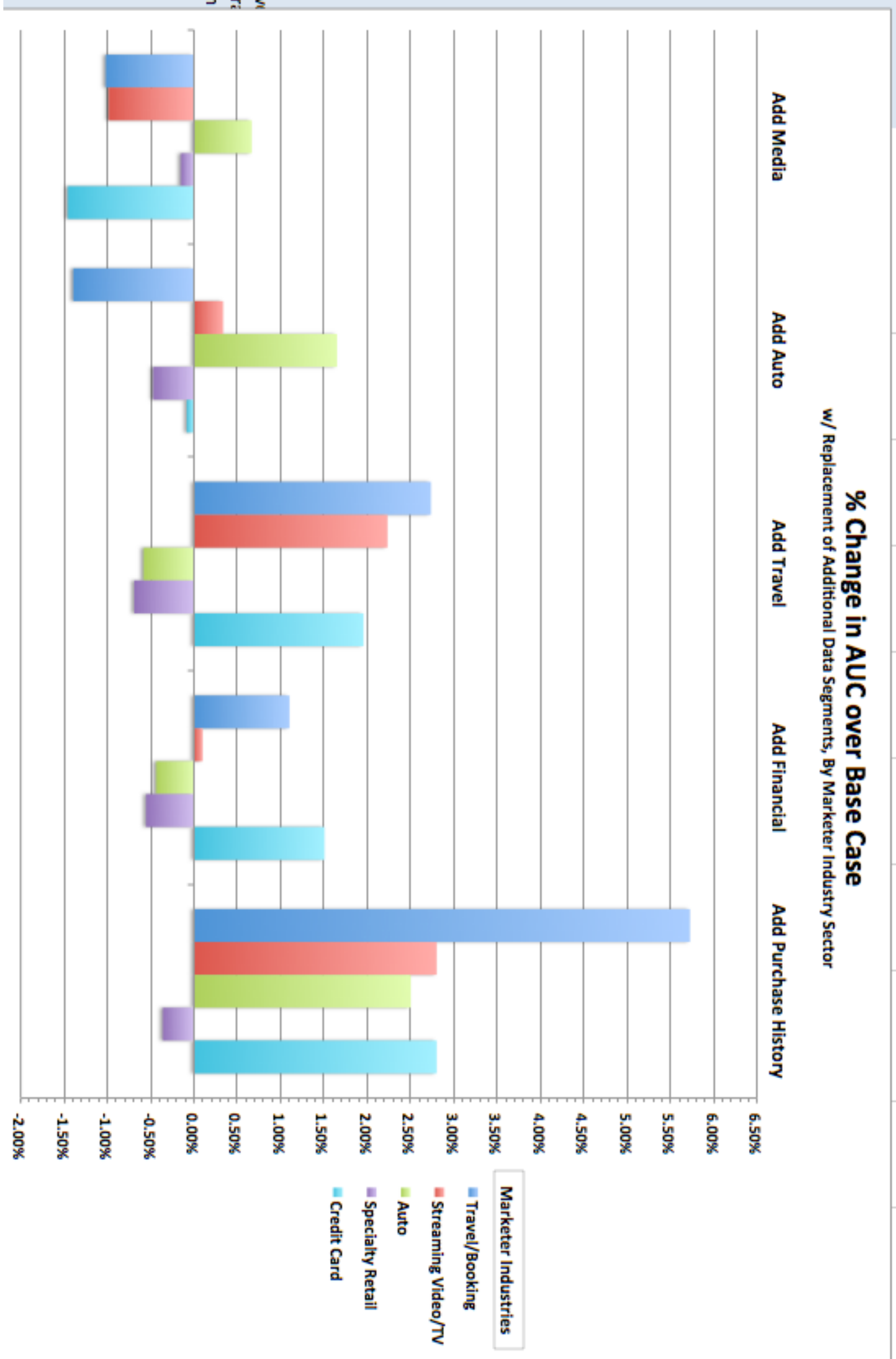


APPENDIX E - Test #2 ROC Curves & AUC Score, Streaming Service, Base + Media



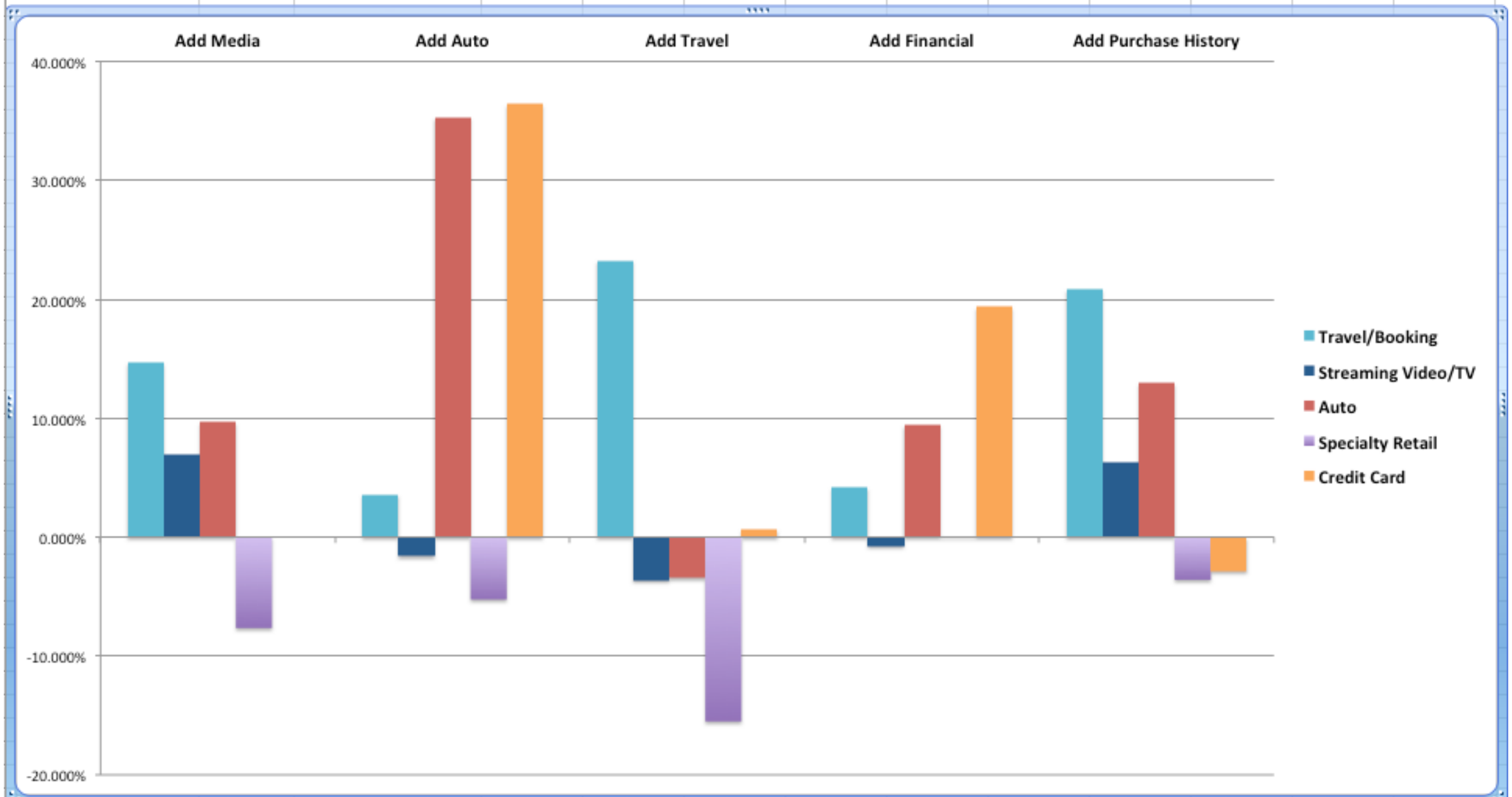
APPENDIX F - Test #2 Results (larger picture)

	Age/Profile (Base) - AUC	Add Media	% Change	Add Auto	% Change	Add Travel	% Change	Add Financial	% Change	Add Purchase History	% Change
Travel/Booking	0.507458864	0.50229276	-1.018%	0.50038966	-1.393%	0.52133803	2.735%	0.513054013	1.103%	0.536510779	5.725%
Streaming Video/TV	0.570275946	0.56466284	-0.984%	0.57220124	0.338%	0.58301438	2.234%	0.570850016	0.101%	0.586248369	2.801%
Auto	0.530989956	0.53453246	0.667%	0.53975665	1.651%	0.52789942	-0.582%	0.528672052	-0.437%	0.544282302	2.503%
Specialty Retail	0.541953665	0.54113509	-0.151%	0.53940175	-0.471%	0.53821597	-0.690%	0.538957277	-0.553%	0.540004988	-0.360%
Credit Card	0.52739552	0.51971672	-1.456%	0.52696809	-0.081%	0.53771925	1.957%	0.535343171	1.507%	0.542168159	2.801%



APPENDIX G - Test #3 Results (larger picture)

	Threshold k	Age/Profile (Base) - Lift	Add Media	% Change	Add Auto	% Change	Add Travel	% Change	Add Financia	% Change	Add Purchas	% Change
Travel/Booking	10%	0.922667395	1.05821472	14.691%	0.95589539	3.601%	1.13739546	23.273%	0.96122946	4.179%	1.11495299	20.840%
Streaming Video/TV	10%	1.178686164	1.26017699	6.914%	1.15957374	-1.622%	1.13590126	-3.630%	1.16994189	-0.742%	1.25237417	6.252%
Auto	10%	0.947050601	1.03856749	9.663%	1.28139905	35.304%	0.91442308	-3.445%	1.03623628	9.417%	1.06964551	12.945%
Specialty Retail	10%	1.151289683	1.0632151	-7.650%	1.09086473	-5.248%	0.97280064	-15.503%	1.15128968	0.000%	1.1099292	-3.593%
Credit Card	10%	1.067652807	1.06765281	0.000%	1.45647586	36.418%	1.07450867	0.642%	1.27424062	19.350%	1.03703704	-2.868%



Bibliography

Arbesman, Samuel. "Stop Hyping Big Data and Start Paying Attention to 'Long Data'." *Wired.com*. Conde Nast Digital, n.d. Web. 04 May 2016.

Aziz, Arslan, and Rahul Telang. *What Is a Cookie Worth?* Rep. Heinz College, Carnegie Mellon University, Mar. 2016. Web.

Carr, David. "Giving Viewers What They Want." *The New York Times*. The New York Times, 24 Feb. 2013. Web. 04 May 2016.

Columbus, Louis. "84% Of Enterprises See Big Data Analytics Changing Their Industries' Competitive Landscapes In The Next Year." *Forbes*. Forbes Magazine, n.d. Web. 04 May 2016.

Dalessandro, Brian, Claudia Perlich, and Troy Raeder. "Bigger Is Better, but at What Cost? Estimating the Economic Value of Incremental Data Assets." *Big Data* 2.2 (2014): 87-96. Web.

Newman, Jared. "The FCC Wants to Blow up the Cable Box. Here's What Its Proposal Will (and Won't) Do." *TechHive*. TechHive, n.d. Web. 04 May 2016.

O'Connell, Michael. "TV Ratings: Super Bowl 50 Falls Shy of Record With 111.9 Million Viewers." *The Hollywood Reporter*. The Hollywood Reporter, n.d. Web. 04 May 2016.

Oracle BlueKai. *Oracle Data Cloud Data Directory*. N.p.: Oracle BlueKai, 2016. Print.

Person, and Adam Clark Estes. "Netflix Can Pinpoint Exactly When People Get Addicted to Shows." *Gizmodo*. Gawker Media, 24 Sept. 2015. Web. 04 May 2016.

PwC. *Global Entertainment and Media Outlook 2015-2019: Internet Advertising: Advertising Segment* (n.d.): n. pag. PwC. Web.

Steel, Emily, and Bill Marsh. "Millennials and Cutting the Cord." *The New York Times*. The New York Times, 02 Oct. 2015. Web. 04 May 2016.

Steinberg, Brian. "Viacom Bets on Big Data to Boost Its Revenues." *Variety*. Variety, 03 Nov. 2015. Web. 04 May 2016.

Tsuchiya, Aki. "How Can TV Ride the Big Data Wave?" *The Guardian*. Guardian News and Media, 29 Aug. 2014. Web. 04 May 2016.

"What Is Real Time Bidding and Why Is It More Effective than Direct Bidding Methods?" *Executive Digital*. N.p., n.d. Web. 04 May 2016.