Targeting Segment Size Forecasting to Meet Higher Impression Targets

Giollamhir

July 8, 2017

Target Segment Size Forecasting to Meet New Impression Targets

Because not all website visitors register or log in consistently, a client company is using behavioral indicators to identify what they refer to as "lookalike segments" for targeting relevant advertising. This behavioral targeting (BT) has been a very successful program. For 2018, this website wants to increase their target numbers for ad impressions to these segments. Ad "impressions"" means views of individual ads. So if there are two advertisements visible on a page a visitor looks at, that is two ad "impressions".

The client company is currently targeting a stringently qualified subset of the lookalike segment., and the current segment population sizes would not afford the new target impression numbers. Based on the results of this analysis, I will be adjusting the logic that allows site visitors into the target segments to include more visitors while still using effective qualification methods.

The rules used in behavioral qualification program are out of scope of this discussion, but here are a few comments to provide context: Our confidence level for lookalike segment composition is monitored, as is the population size itself. Adjustments are made periodically to ensure both. Because of our need to maintain confidence levels and define lookalikes as stringently as possible, the accurate forecast of the necessary size of the segment population to be developed in this report is a critical consideration.

Components of the analysis to forecast segment size needed to support the new goals include:

- Real-time data on number of users from each segment actively on the site day to day
- How many actual advertisement impressions these current, active segment populations yield

By looking at how many users in a segment ("BT population" or "BTpop") result in n impressions, we can forecast how many users will be needed in future segments to guarantee future target impression numbers.

Complicating factors include:

- Impressions are served by an ad server making algorithmic ad display decisions on each page load. This algorithm prioritizes between many campaigns and many levels of target prioritization (BT is second or third priority, after various context-specific rules and various rules based on known logged-in user data).
  - It's not possible to gather observations without this algorithmic "interference".
  - We can assume the ad serving algorithms will continue to work consistently in 2018 as during the observation period.
- Observations can only be collected for campaigns that are live
  - Some segments were not running campaigns during months observed.
  - Several segments ran campaigns, but not at full capacity of maximum possible number of campaigns.
  - Some segments ran, but with configured serving constraints or timing issues by advertiser request
- Higher-level considerations such as annual seasonal patterns and total number of campaigns will be best analyzed when we have more longitudinal data. A few additional words on these are included at the end of the report.

April and prior months ran with target limits set into the ad server, imposing an artificial constraint on observations. For May and June, we ran these programs "uncapped" to collect more accurate observations in impression capacity of current BT segment population sizes. This report focuses on June because June has the best test segments to forecast from.

Business inputs

Proposed target numbers for the new, higher impression advertisement impression goals for 2018 are listed below.

```
##    Segment NewBTPP Campaigns NewTOT.Target
## 1        C   40000         5        200000
## 2        P   40000         4        160000
## 3        E   16000         3         48000
## 4       HO   32000         5        160000
## 5      ULM   16000         4         64000
## 6        I   24000         3         72000
## 7        R   15000         2         30000
## 8        G   30000         2         60000
## 9       SY   20000         2         40000
## 10     NPH   12000         2         24000
## 11     EUR   21000         2         42000
## 12      WH   20000         1         20000
## 13       D   19500         1         19500
## 14      EM    1500         1          1500
## 15     PDS   16000         2         32000
```

Segment Populations

The BT program involves writing cookies for the segments to be targeted. Since we only want to count users who are active, we can use an analytics report to determine daily counts of users who were on the site with a cookie indicator for each segment.

```
## 'data.frame':    30 obs. of  17 variables:
##  $ Dimension: chr  "Day" "Day" "Day" "Day" ...
##  $ Item     : Date, format: "2017-06-01" "2017-06-02" "2017-06-03" "2017-06-04" ...
##  $ C        : int  2254 1484 726 898 1787 2767 2415 2370 1547 698 ...
##  $ P        : int  7894 4715 2710 3147 5134 8411 8563 7800 4716 2632 ...
##  $ E        : int  1006 697 365 446 830 1166 1155 1153 785 464 ...
##  $ HO       : int  4371 2831 1609 2560 4406 5898 5793 5368 3365 1553 ...
##  $ ULM      : int  879 572 307 416 663 1203 1067 922 643 363 ...
##  $ I        : int  620 383 206 255 470 871 713 646 420 210 ...
##  $ G        : int  1339 831 446 565 931 1501 1429 1282 784 447 ...
##  $ SY       : int  662 488 245 315 485 723 736 597 459 275 ...
##  $ R        : int  586 415 273 304 557 858 734 682 438 255 ...
##  $ NPH      : int  1094 676 361 474 767 1114 1223 1092 705 366 ...
##  $ EUR      : int  2639 1642 841 1063 1973 2868 2889 2459 1693 851 ...
##  $ W        : int  1110 689 344 479 823 1259 1341 1188 790 369 ...
##  $ D        : int  900 555 301 366 591 838 995 1036 666 323 ...
##  $ EM       : int  1632 1044 677 734 1181 1901 1952 1868 1243 734 ...
##  $ PDS      : int  1755 1089 619 738 1216 2023 2005 1781 1083 563 ...
```

June 2017 Uncapped BT Ad Serving Test Results

June impressions were reported in the typical raw format that will need to be ingested for typical operational maintenance.

First, a summary of campaigns that ran through June. Aside from the overall considerations listed above, there are also sometimes special requests from the client on how we serve ads that impose their own constraints. The table below lists what campaigns actually ran (not all segments always run, and not necessarily to the maximum number of campaigns per segment).

- "CampNum" denotes the total campaigns that *actually ran* this month for that segment.
- The *Consider* column indicates any serving considerations (impressions will be lower due to constraints): SR means serving restrictions, and SP means Special Pacing.

```
##    Segment Campaign CampNum Consider
## 1        C        H       1     <NA>
## 2        P       GH       1       SR
## 3        E     L-Jn       1     <NA>
## 4       HO       AD       4     <NA>
## 5       HO       AM       4       SP
## 6       HO        K       4     <NA>
## 7       HO     N-Jn       4     <NA>
## 8        I        H       1       SR
## 9        R        H       1     <NA>
## 10       G       BS       2     <NA>
## 11       G       GH       2       SR
## 12      SY        N       1     <NA>
## 13     NPH      AMP       2     <NA>
## 14     NPH       FV       2     <NA>
## 15     EUR     A-Jn       2     <NA>
## 16     EUR     B-Jn       2     <NA>
```

June Campaign Data

Campaign impression logs are exported from the ad server in individual reports listing one campaign at a time, with impression numbers per ad position on the page (there are 2-4 possible positions depending on the template) per day. The following process ingests the campaign reports, consolidates all impressions per day, appends the relevant Behavioral Targeting segment population, and then binds the data frame for that campaign to an overall summary of all impression activity in June. I also add a "DoW" column noting the day of the week associated with that date, because day of week is critical to all calculations (more on that later).

The code is ingesting all segments that had any campaigns live during the observation period.

The structure of the assembled data we'll analyze is 224 rows (or "observations"), with 6 columns:

- **Segments:**: Labeled Seg
- **Campaign:**:Labeled Camp
- **Date**: Labeled Date
- **Impressions:**: Labeled Imps
- **Behavioral Targeting Segment Population:**: Labeled BTpop
- **Day of the Week (calculated from Date)**: Labeled DoW

```
## 'data.frame':    480 obs. of  6 variables:
##  $ Seg  : Factor w/ 10 levels "C","P","E","HO",..: 1 1 1 1 1 1 1 1 1 1 ...
##  $ Camp : Factor w/ 13 levels "H","GH","L-Jn",..: 1 1 1 1 1 1 1 1 1 1 ...
```

```
##  $ Date : Date, format: "2017-06-01" "2017-06-02" "2017-06-03" "2017-06-04" ...
##  $ Imps : int  1096 820 342 456 1119 1305 1414 1231 676 389 ...
##  $ BTpop: int  2254 1484 726 898 1787 2767 2415 2370 1547 698 ...
##  $ DoW  : Factor w/ 7 levels "Friday","Saturday",..: 7 1 2 3 4 5 6 7 1 2 ...
```

Total Impressions Per Campaign

Assembling the data is a first step to investigation. With the necessary data assembled into a usable structure, we can start to get a sense of:

- The relationship between BT segment population and ad impressions in a day.
- Differences in performance between segments – which operate at different magnitudes.
- Campaigns are individual ad campaigns, always within one segment. If they don't have imposed constraints, any significant variation between campaigns within a segment should be investigated.

```
## Source: local data frame [16 x 3]
## Groups: Seg [?]
##
##       Seg   Camp Impressions
##    <fctr> <fctr>       <int>
## 1       C      H       49771
## 2       P     GH       45184
## 3       E   L-Jn       19874
## 4      HO     AD       41050
## 5      HO     AM       25836
## 6      HO      K       32909
## 7      HO   N-Jn       33039
## 8       I      H       13093
## 9       R      H        8755
## 10      G     GH        7087
## 11      G     BS       18204
## 12     SY      N       15668
## 13    NPH    AMP       11982
## 14    NPH     FV       13568
## 15    EUR   A-Jn       21885
## 16    EUR   B-Jn       28299
```

Total Impressions Per Segment

This overview provides an at-a-glance summary of total impressions at the segment level during the observation period.

```
## # A tibble: 10 × 3
##       Seg campsRun TotImps
##    <fctr>    <int>   <int>
## 1       C        1   49771
## 2       P        1   45184
## 3       E        1   19874
## 4      HO        4  132834
## 5       I        1   13093
## 6       R        1    8755
## 7       G        2   25291
## 8      SY        1   15668
## 9     NPH        2   25550
```

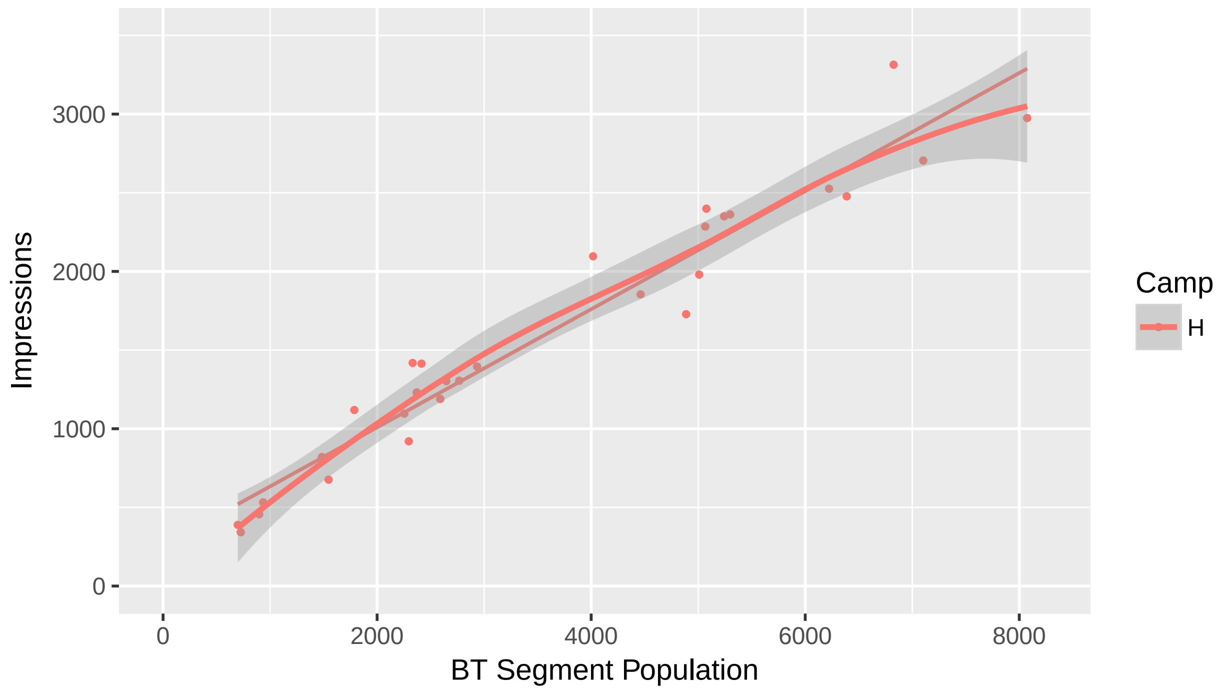
```
## 10    EUR          2   50184
```

What Does Impression per BT Population Look Like Graphed?

The following graphs depict two samples; one segment that ran one campaign, and one segment that ran four. The light, straight line for each campaign depicts the line of best fit for the campaign and segment data. Graphs for all segments are displayed in the appendix.

Segment C − Sample Daily Impression Data June 2017

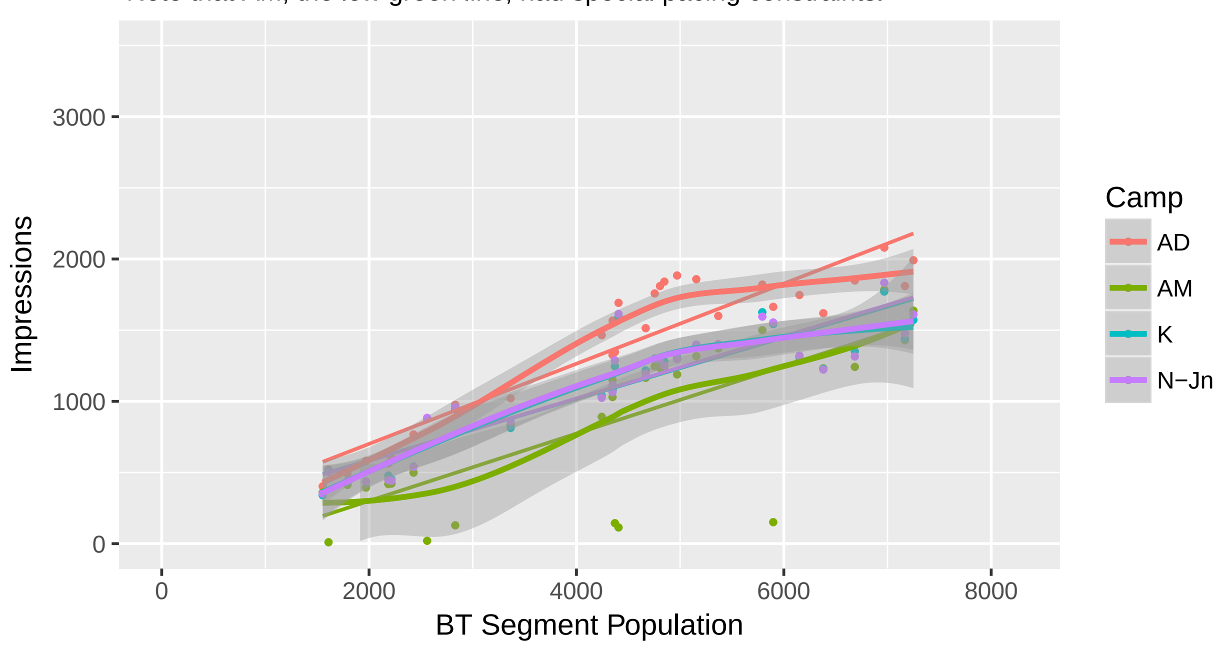
Gray shading shows confidence interval of the smoothing linear model



For each segment, this files loads impression numbers and BT population into a dataframe so they can be modeled into a line.

Daily Impressions Per BT Population,June 2017, Segment HO

Segment HO − Four HO Campaigns Were Live
Note that AM, the low green line, had special pacing constraints.



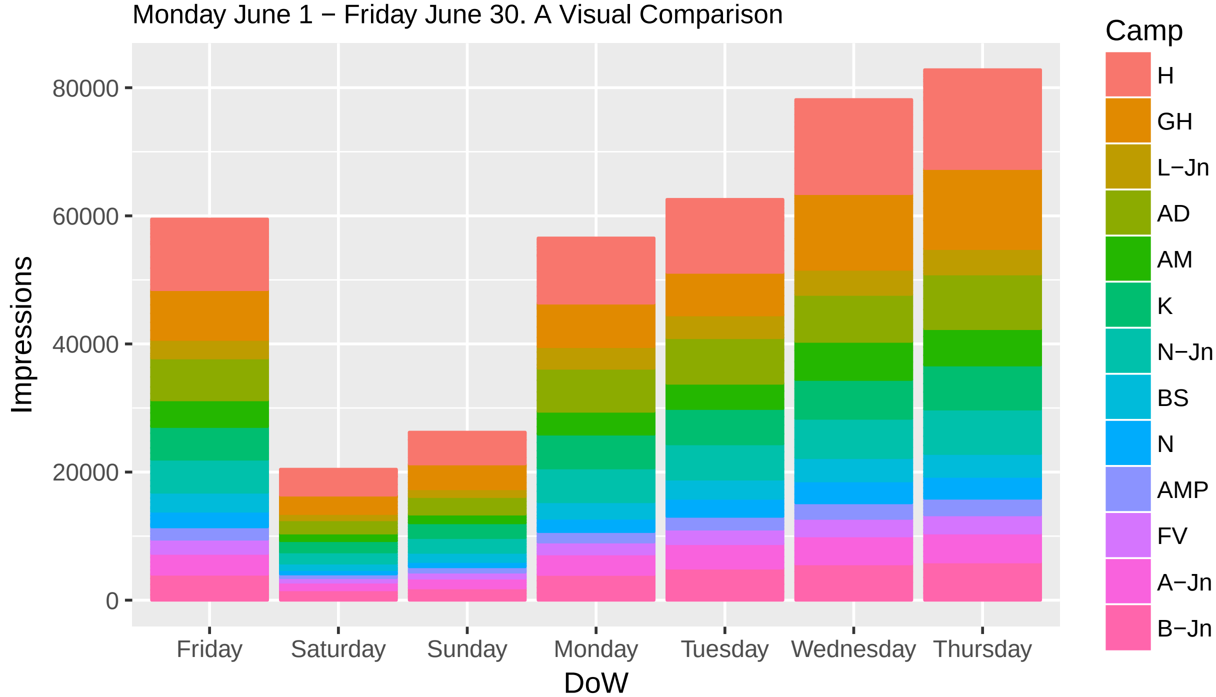
each segment, this process loads impression numbers and BT population into a dataframe.
Here is an example of a multi−campaign segment.
The other campaigns vary, but largely within a similar confidence range

Weekly Cycle

There is a pronounced weekly visit cycle. Some experimenting showed that taking into account impression performance per day of week was more accurate than calculating potential impressions from weekly and monthly averages. For a glimpse of the traffic variations, here are all the June impressions, mapped by day of the week.

Total Impressions June 2017 − All Campaigns − by Days of the Week

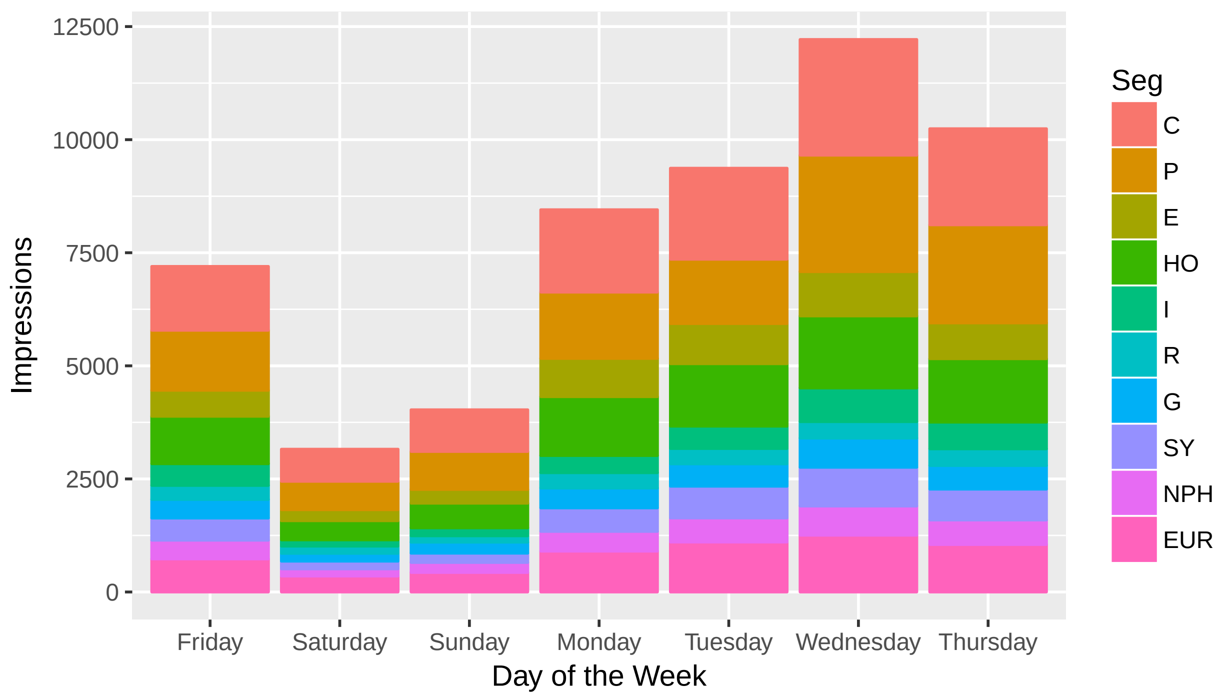
Monday June 1 − Friday June 30. A Visual Comparison



June began on a Thursday and ended on a Friday, so Thursday and Friday have one more day of impressions than other days.
Typically Wednesday and Thursday each yield 20% of all weekly impressions

Average Impressions per Segment by Day of Week, June 2017

Monday June 1 − Friday June 30. A Visual Comparison



Because this chart portrays averages per day of the week, note the reduced scale with 12,500 maximum, vs. over 80,000 total above.

Group Data for Predictive Analysis

Now that observations have been collected to one dataframe, I need to collapse the observation data per segment so that I can organize and process data by campaign, creating linear models and predictions campaign by campaign within each segment.

Our June activity dataframe is now significantly evolved from the six-variable list of observations we first collected. Rather than a dataframe of values, it is now adataframe of dataframes, so that it can contain more complex information such as the linear model for each campaign.

The structure of the assembled data we'll analyze is 224 rows (or "observations"), with 6 columns:

- **Segments:**: Labeled Seg
- **Campaign:**:Labeled Camp
- **data:**: BTPop / Imp/ Date / DoW data assembled earlier now stored in one cell for each campaign
- **linmodel:**: The linear model for the campaign impression/BTpop data
- **glance,tidy,rsq,augment**: Columns separately storing particular data points about the linear model
- **Consider**: Lists codes for any constraints applied to the campaign at the requst of the advertiser.

```
## # A tibble: 6 × 9
##      Seg   Camp              data linmodel                glance                tidy       rsq
##   <fctr> <fctr>            <list>   <list>                <list>              <list>     <dbl>
## 1      C      H <tibble [30 × 4]> <S3: lm> <data.frame [1 × 11]> <data.frame [2 × 5]> 0.9437481 <data
## 2      P     GH <tibble [30 × 4]> <S3: lm> <data.frame [1 × 11]> <data.frame [2 × 5]> 0.8518920 <data
## 3      E   L-Jn <tibble [30 × 4]> <S3: lm> <data.frame [1 × 11]> <data.frame [2 × 5]> 0.8645360 <data
## 4     HO     AD <tibble [30 × 4]> <S3: lm> <data.frame [1 × 11]> <data.frame [2 × 5]> 0.8653767 <data
## 5     HO     AM <tibble [30 × 4]> <S3: lm> <data.frame [1 × 11]> <data.frame [2 × 5]> 0.5887603 <data
## 6     HO      K <tibble [30 × 4]> <S3: lm> <data.frame [1 × 11]> <data.frame [2 × 5]> 0.8292500 <data
```

How Well Does Can we Model Impression Numbers from Segment Size?

Now that we've run the linear models, we can visualize statistical summary data about impressions relative to segment population for each campaign in a segment.

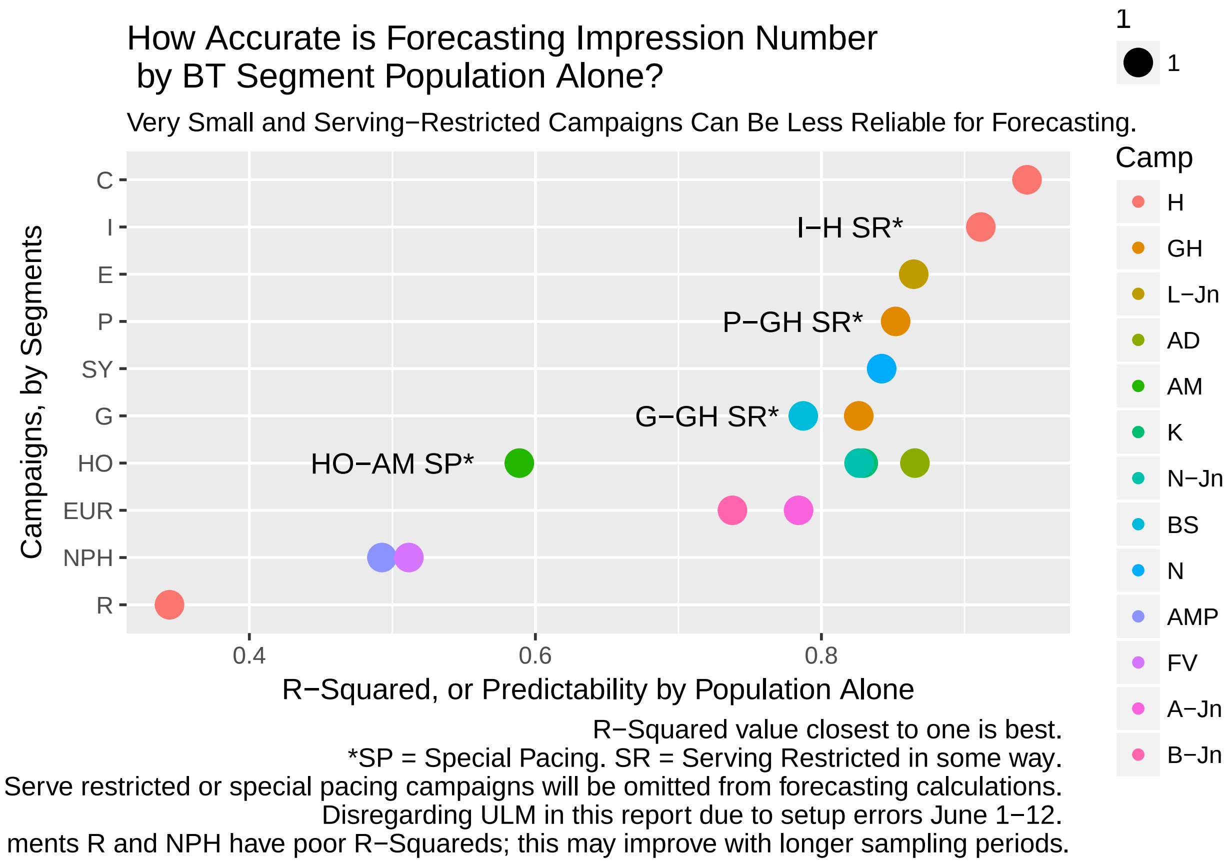
For Segments At Close to One in the Below Graph, Very Well!

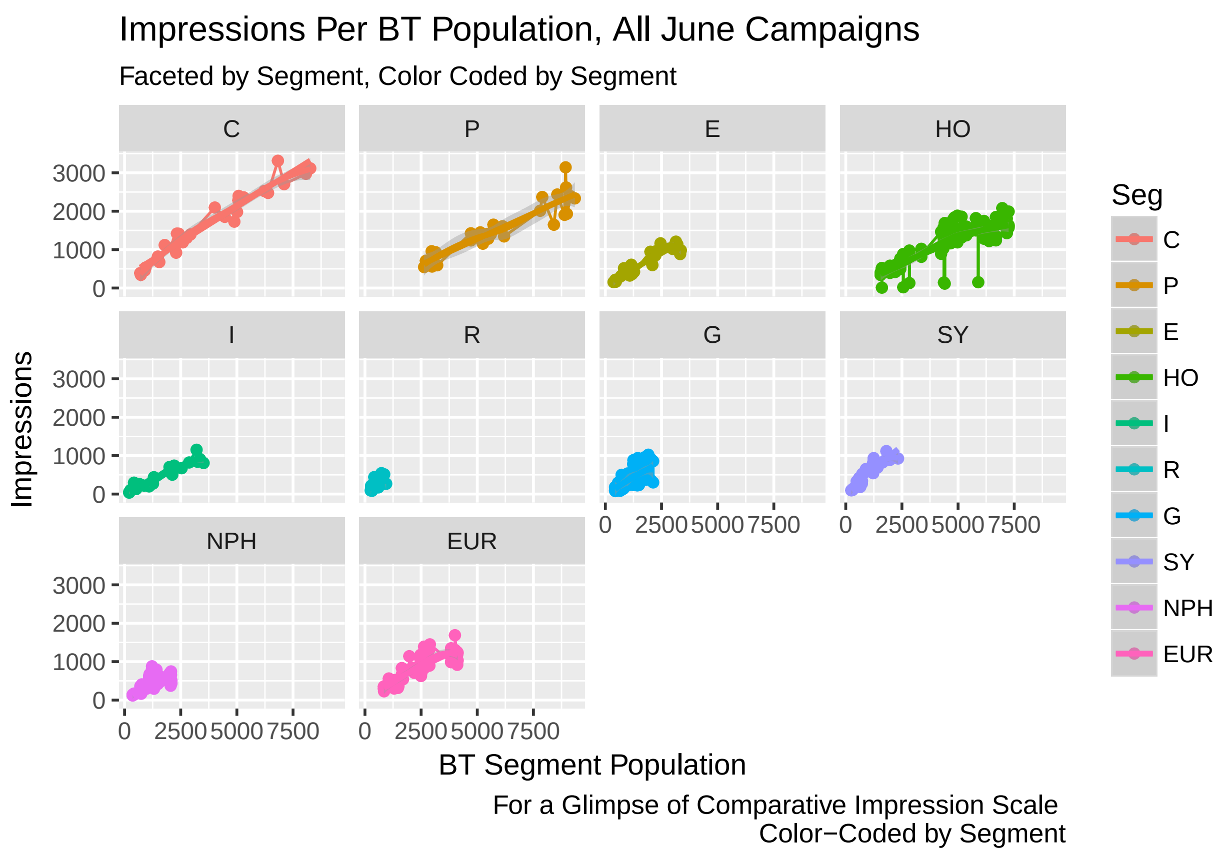
The below graphs show the "R Squared" of each campaign. R-squared is a statistic to indicate how well a linear model fits. Or, in our case, that means how well we can predict impression performance based on segment population alone, without other considerations. R squared is always between 0 and 1. The closer to 1 it is, the more perfectly the model fits. Where the R squared is low, more investigation is needed.

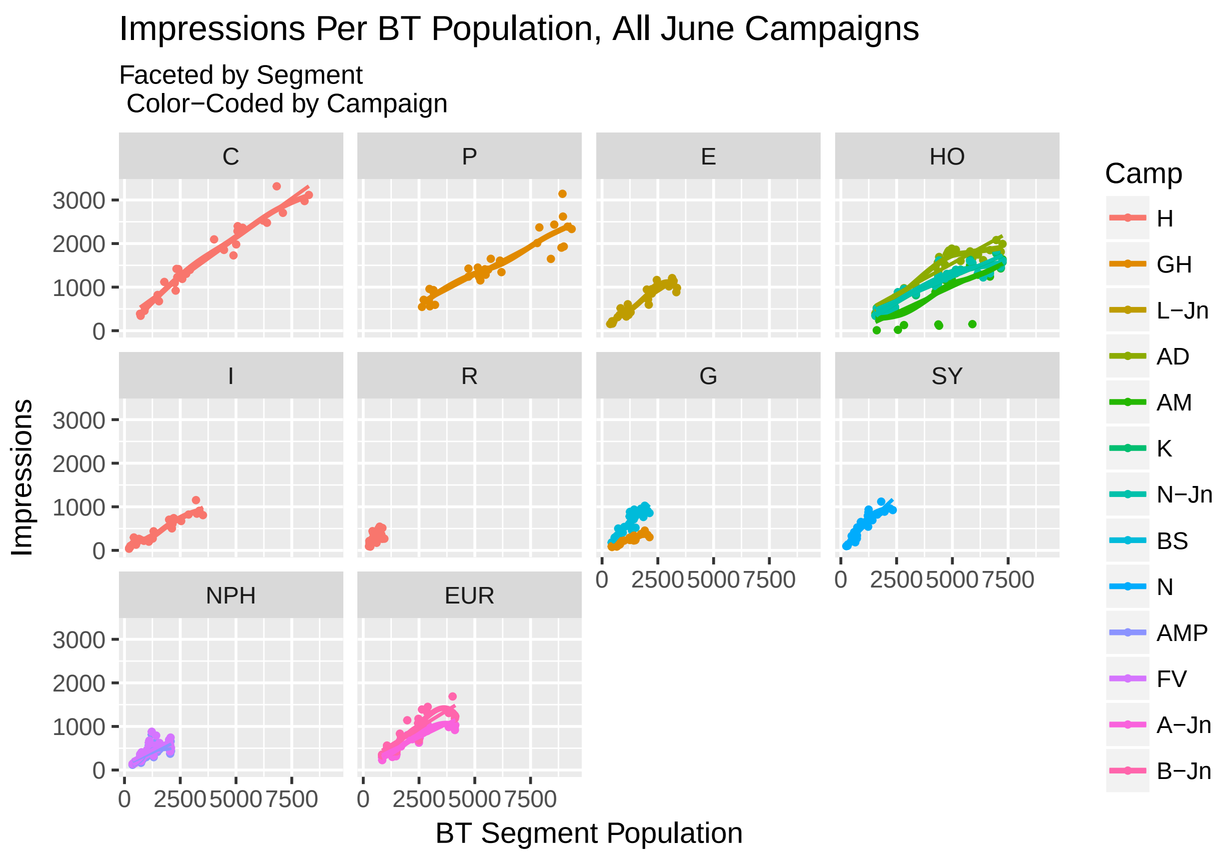
Segment-Campaign codes noted on the map indicate either serving restrictions or special pacing constraints.

The observation period was somewhat short, so fewer observations, especially in smaller segments, could also have an impact on poor fit. We can look to find other data to add to our modeling for smaller segments to make better predictions (for example, client company content publication patterns in relevant topics, or behavioral patterns more peculiar to that audience segment.)

Another question we can investigate is whether more or fewer total campaigns in a month impact predictability of individual campaign impression performance.


How Accurate is Forecasting Impression Number
by BT Segment Population Alone?
Very Small and Serving−Restricted Campaigns Can Be Less Reliable for Forecasting.
I−H SR*
P−GH SR*
G−GH SR*
HO−AM SP*
C
I
E
P
SY
G
HO
EUR
NPH
R
Campaigns, by Segments
0.4
0.6
0.8
R−Squared, or Predictability by Population Alone
1
1
Camp
H
GH
L−Jn
AD
AM
K
N−Jn
BS
N
AMP
FV
A−Jn
B−Jn
R−Squared value closest to one is best.
*SP = Special Pacing. SR = Serving Restricted in some way.
Serve restricted or special pacing campaigns will be omitted from forecasting calculations.
Disregarding ULM in this report due to setup errors June 1−12.
ments R and NPH have poor R−Squareds; this may improve with longer sampling periods.


Impressions Per BT Population, All June Campaigns
Faceted by Segment, Color Coded by Segment
C
P
E
HO
I
R
G
SY
NPH
EUR
3000
2000
1000
0
0
2500
5000
7500
Impressions
BT Segment Population
Seg
For a Glimpse of Comparative Impression Scale
Color−Coded by Segment



Test Segment Populations

Below I import segment populations of new test segments and use Set 1 and Set 2 as the inputs to predict impressions based on BT population of the test segment.

Predictions Based on First Set of Test Segments

Let's back up and get some context for the predictions and forecasts that will follow. "Segments" in this section refers to test segments defined in the segment manager tool, rather than the live targeting segments meant everywhere else in this report. There are hundreds of test segments defined in the segment manager, all being monitored for size, consistency/volatility, and confidence of accuracy.

- A new set of qualifying segments became the live, targeted set on June 13. This is "Set 1" in the following analysis.
- Several of the largest, most likely segment definition choices for Set 2 were created at noon on June 13. Where those are used, there is only input data starting from June 14 (volume for earlier dates in June are averaged from the second half of June).

The input data for these segments is from reports I've generated listing the daily "real-time", active populations of these segments. These reports are pulled from the segmentation tool, Adobe Audience Manager, rather than the traffic analytics tool (since Set 2 isn't live, there's no traffic data for it). Data from these two reports has been comparable historically. I'll run predictions and estimates against the populations for each of these sets.

CAUTIONARY NOTE: The predictions apply to a month running a the same number of campaigns listed here. Cases where all possible campaigns are running would reduce overall performance by an amount that will be predictable with a bit more data.

Note: Most Test Set 2 segments have only existed from 6/14. So daily segment sizes input for 6/1 - 6/13 are an average of the sizes from the two-week period 6/14 - 6/27.

```
## # A tibble: 16 × 6
##       Seg   Camp       rsq Consider ImpFcstSet1 ImpFcstSet2
##    <fctr> <fctr>     <dbl>   <fctr>       <dbl>       <dbl>
## 1       C      H 0.9437481       NA      113888      113888
## 2       P     GH 0.8518920       SR       31217      100013
## 3       E  L-Jn 0.8645360        NA       23680       28385
## 4      HO     AD 0.8653767       NA       80568       67307
## 5      HO     AM 0.5887603       SP       59044       47901
## 6      HO      K 0.8292500       NA       63224       53052
## 7      HO   N-Jn 0.8263334       NA       63732       53433
## 8       I      H 0.9114516       SR       23299       42653
## 9       R      H 0.3441589       NA        7632       14189
## 10      G     BS 0.7873161       NA       16922       52764
## 11      G     GH 0.8261066       SR        6596       20332
## 12     SY      N 0.8421271       NA       22109       34591
## 13    NPH    AMP 0.4926972       NA       10664       27833
## 14    NPH     FV 0.5115040       NA       12062       31690
## 15    EUR   A-Jn 0.7840176       NA       22017       23917
## 16    EUR   B-Jn 0.7377760       NA       28468       30895
```

Estimates for Four Segments Where We Lack Observations

Four segments were not live during our observation period:

- D
- EM
- WH
- PDS

I will use the observations of other segment impression performance to incorporate estimations of potential impression numbers for these segments. All these have a maximum or one or two campaigns, so we will use the coefficients for observations that ran 1-2 campaigns. This describes nearly all of our segments. After omitting coefficients for segment HO and for campaigns that had serving constraints, I'll average coefficients for "normal" 1-2-campaign segments. I'll use the mean coefficient to estimate potential performance of test segments for the four segments on which we have no impression data.

I will now use the estimate associated only for segments running 1-2 campaigns, with no considerations, (so, eliminating campaigns P - GH, HO-AM, I-H, and G-GH) for the estimates.

Finally, I will append those estimates to a summary table presenting the Forecasts where a prediction was calculated and Estimates where a prediction was estimated.

```
## # A tibble: 9 × 11
##      Seg   Camp       rsq Consider ImpFcstSet1 ImpFcstSet2  term  estimate  std.error statistic
##   <fctr> <fctr>     <dbl>   <fctr>       <dbl>       <dbl> <chr>     <dbl>      <dbl>     <dbl>
## 1      C      H 0.9437481       NA      113888      113888 BTpop 0.3678802 0.01697336 21.673973 4.87
## 2      E   L-Jn 0.8645360       NA       23680       28385 BTpop 0.3278809 0.02452774 13.367759 1.12
## 3      R      H 0.3441589       NA        7632       14189 BTpop 0.3615514 0.09432153  3.833180 6.56
## 4      G     BS 0.7873161       NA       16922       52764 BTpop 0.4830291 0.04744463 10.180901 6.46
## 5     SY      N 0.8421271       NA       22109       34591 BTpop 0.4872341 0.03986793 12.221205 9.66
## 6    NPH    AMP 0.4926972       NA       10664       27833 BTpop 0.2323360 0.04455342  5.214774 1.54
```

```
## 7    NPH      FV 0.5115040       NA       12062       31690 BTpop 0.2656207 0.04905563  5.414683 8.93
## 8    EUR    A-Jn 0.7840176       NA       22017       23917 BTpop 0.2467318 0.02447329 10.081678 8.04
## 9    EUR    B-Jn 0.7377760       NA       28468       30895 BTpop 0.3151940 0.03551182  8.875749 1.25
```

```
##     Seg Camp       rsq Consider ImpFcstSet1 ImpFcstSet2  Est1  Est2  term estimate std.error CampMax
## 1     C    H 0.9437481     <NA>      113888      113888    NA    NA BTpop    0.368     0.017       5
## 2     P   GH 0.8518920       SR       31217      100013    NA    NA BTpop    0.268     0.021       4
## 3     E L-Jn 0.8645360     <NA>       23680       28385    NA    NA BTpop    0.328     0.025       3
## 4    HO   AD 0.8653767     <NA>       80568       67307    NA    NA BTpop    0.282     0.021       5
## 5    HO   AM 0.5887603       SP       59044       47901    NA    NA BTpop    0.237     0.037       5
## 6    HO    K 0.8292500     <NA>       63224       53052    NA    NA BTpop    0.216     0.019       5
## 7    HO N-Jn 0.8263334     <NA>       63732       53433    NA    NA BTpop    0.219     0.019       5
## 8     I    H 0.9114516       SR       23299       42653    NA    NA BTpop    0.275     0.016       3
## 9     R    H 0.3441589     <NA>        7632       14189    NA    NA BTpop    0.362     0.094       2
## 10    G   BS 0.7873161     <NA>       16922       52764    NA    NA BTpop    0.483     0.047       2
## 11    G   GH 0.8261066       SR        6596       20332    NA    NA BTpop    0.185     0.016       2
## 12   SY    N 0.8421271     <NA>       22109       34591    NA    NA BTpop    0.487     0.040       2
## 13  NPH  AMP 0.4926972     <NA>       10664       27833    NA    NA BTpop    0.232     0.045       2
## 14  NPH   FV 0.5115040     <NA>       12062       31690    NA    NA BTpop    0.266     0.049       2
## 15  EUR A-Jn 0.7840176     <NA>       22017       23917    NA    NA BTpop    0.247     0.024       2
## 16  EUR B-Jn 0.7377760     <NA>       28468       30895    NA    NA BTpop    0.315     0.036       2
## 17   WH  EST        NA      EST          NA          NA  7441 19808  <NA>    0.343     0.042       1
## 18    D  EST        NA      EST          NA          NA 10421 30332  <NA>    0.343     0.042       1
## 19   EM  EST        NA      EST          NA          NA 25375 29432  <NA>    0.343     0.042       1
## 20  PDS  EST        NA      EST          NA          NA 16860 31889  <NA>    0.343     0.042       2
```

Note that the average linear model estimate for segments running 1-2 campaigns is .343, and the average estimate for HO, running four campaigns, is .239. This may generalize to a reduction of roughly .1 when more campaigns are running than there are ad positions on ad-bearing pages. We'll return to this in the future opportunities section.

Create Output

To easily share with other departments, I will write a subset of the data containing the forecast and estimate summaries for each campaign to a csv file.

Any number in either the prediction or estimate column of this output indicates how many impressions per month we can anticipate for the coming year. These projections can be used in marketing materials, sales discussions, and so on. Bear in mind the caveat that we haven't seen the actual impact of 5 campaigns.

If all campaigns are full, *we made the rough estimate for Segment C that impressions would be closer to 50,000 impressions per campaign at 5 campaigns.* A similar estimate would apply to Segment P * HO had four campaigns running. We need to find out what happens when HO runs five campaigns live. Our rough estimate would put the fifth campaign at around 8,000 impressions

Future Opportunities

The data collected so far is most informative for segments with a simultaneous campaign maximum of 1-2, since that's the number of unrestricted campaigns we were able to collect reliable data on for forecasting or estimating. We have one sample with four HO campaigns and no samples with 5 simultaneous campaigns. Because there are typically 2-3 ad positions per page, running 5 simultaneous campaigns will require more users (with more page views). It's a third tier of forecasting we don't have any data on yet. It will be important to collect more data over time for accurate forecasting and estimating with 5-campaign segments.

As we collect and analyze data with this method going forward, we will be able to add observations and predictive analysis not just month to month impression performance to BT segment population, but on the impact of:

1. Performance of number of campaigns live in a segment in a given month. I'm calling this "Tiers": 1-2, vs. 3-4, vs. 5 campaigns live
2. How performance varies depending on total overall number of campaigns running in a given month.
3. Seasonal patterns.

As an example of campaign tiering it item one, let's do some gross estimation based on very limited observation. For simplicity let's call it running four campaigns when there are only 2 ad positions available. If we say this added tier reduces impression performance by roughly .1, we might guess that going to a third tier and running 5 campaigns might reduce impression performance by .2 for that segment.

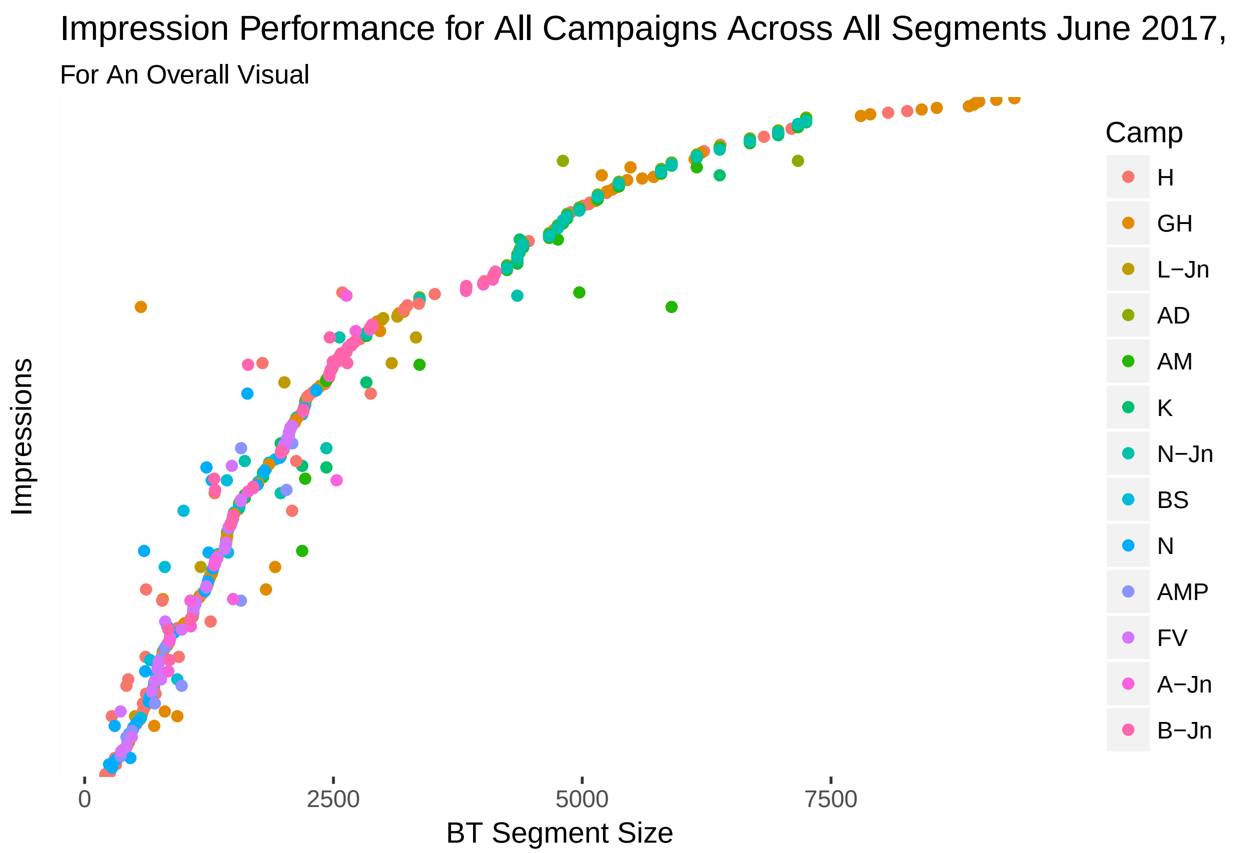
For example, the C segment forecast above is 113,888 with segment C running one campaign. Estimated performance for C with one campaign is .368 impressions per each user in the active BT segment population. If C were running five campaigns live, with the same segment population, that estimate goes down to about .168. This would make the per campaign forecast *while running five campaigns* as low as 51,992. This is why it will be important to continue observing and measuring performance, especially when there are more simultaneous campaigns than we've been able to test so far.

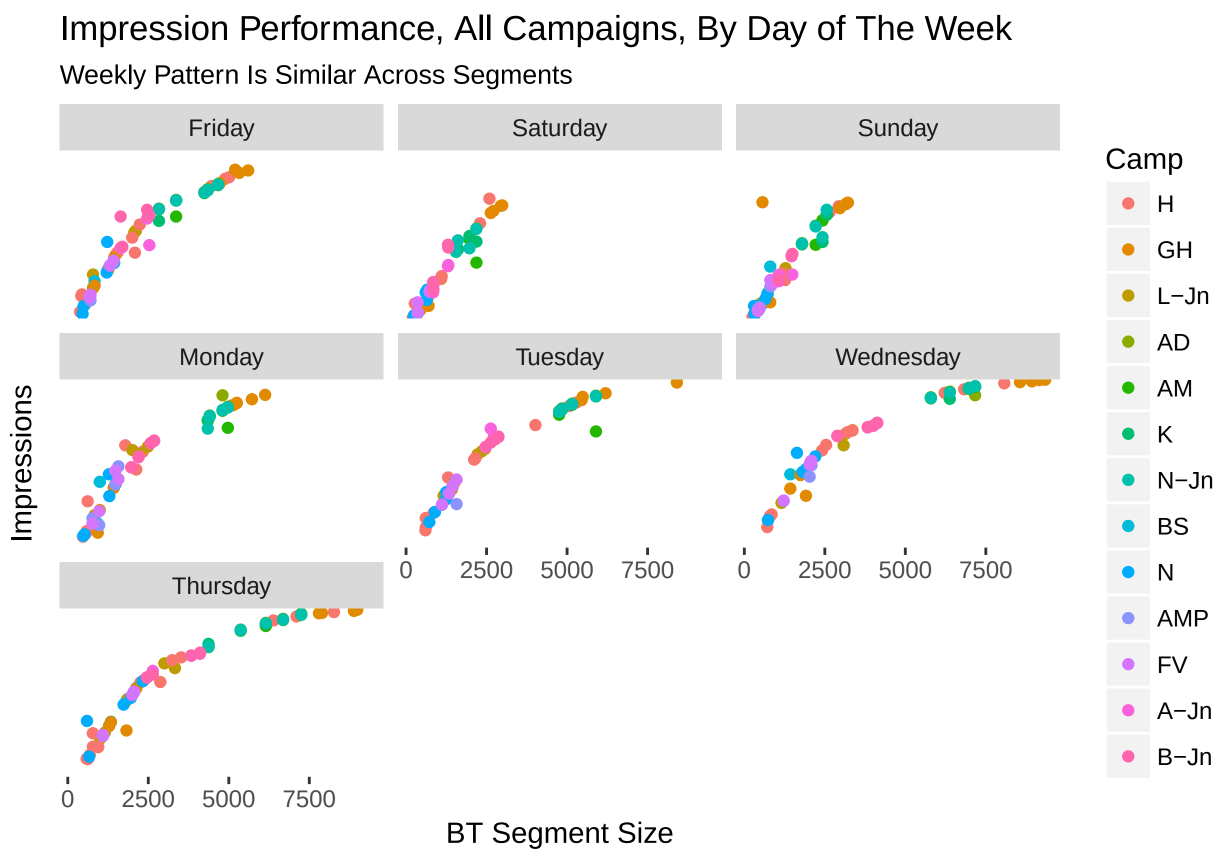
APPENDIX: A Deeper Look at Data for Each Segment

Now I'll break the data back out to make it easier to look a little deeper at performance for each segment.

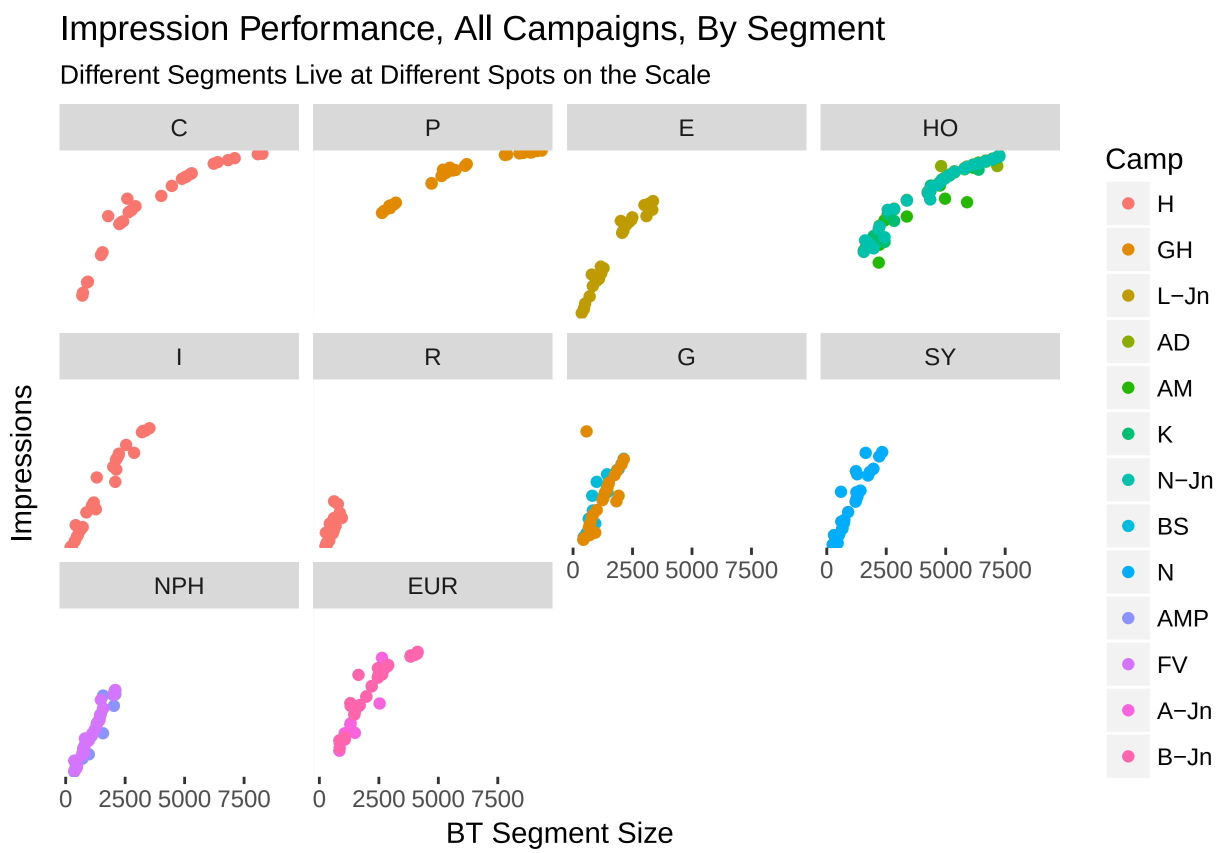
The effects of special pacing and serving restrictions on P-GH, HO-AM, I-H,and G-GH are visible here.

Overal Visualizations

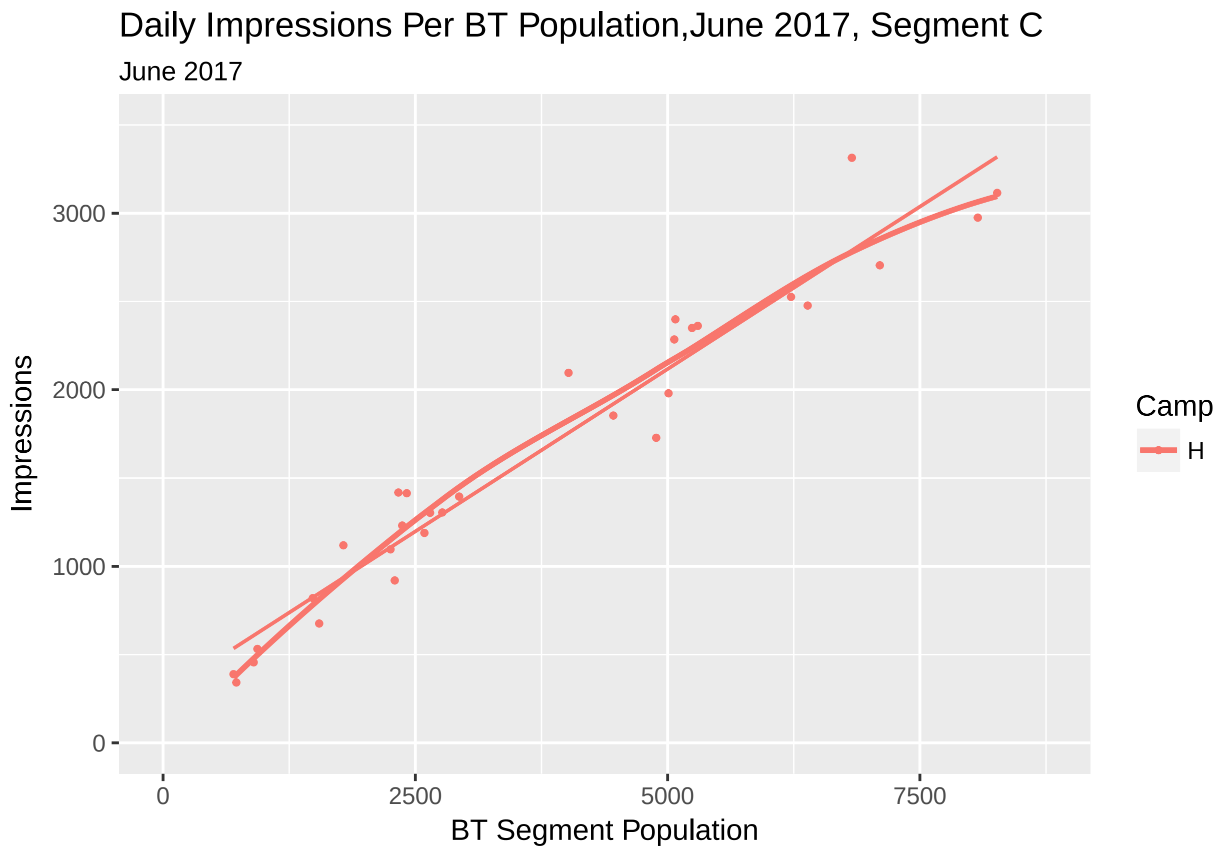


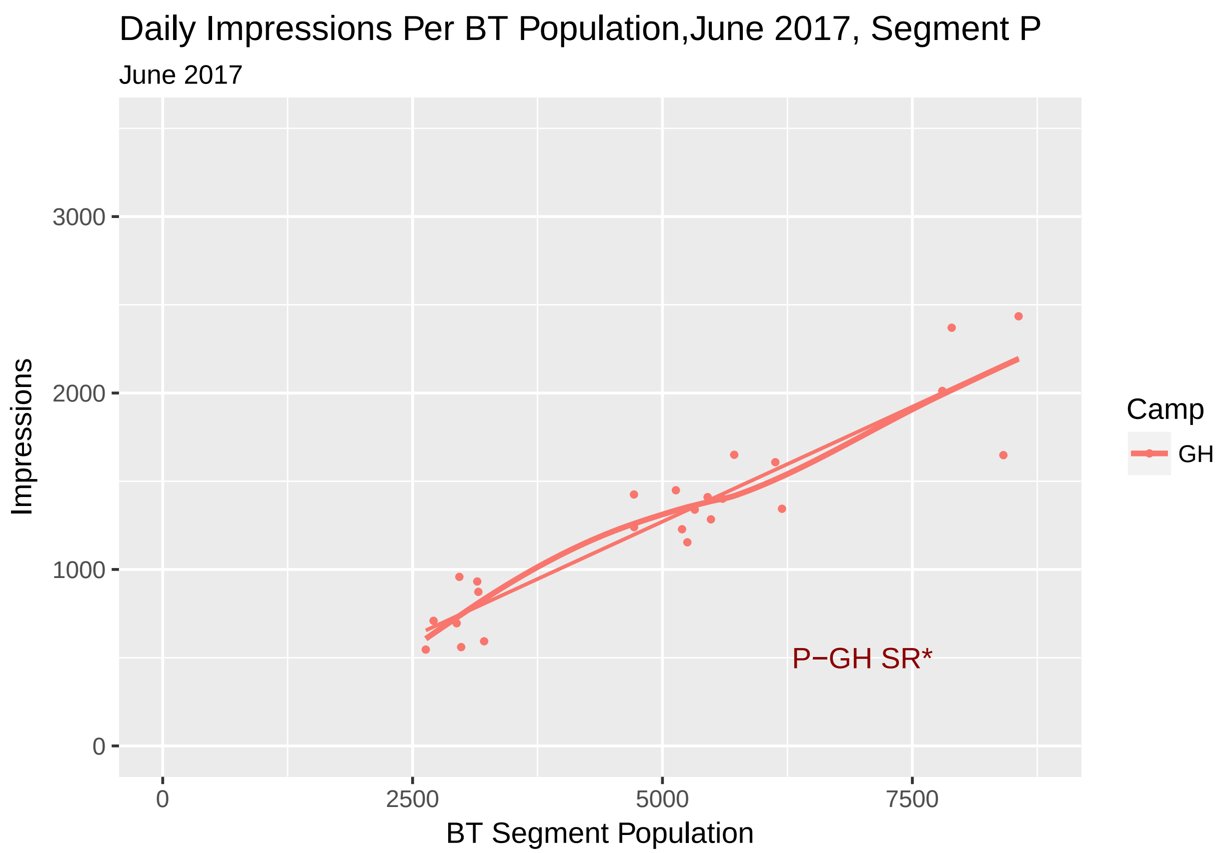


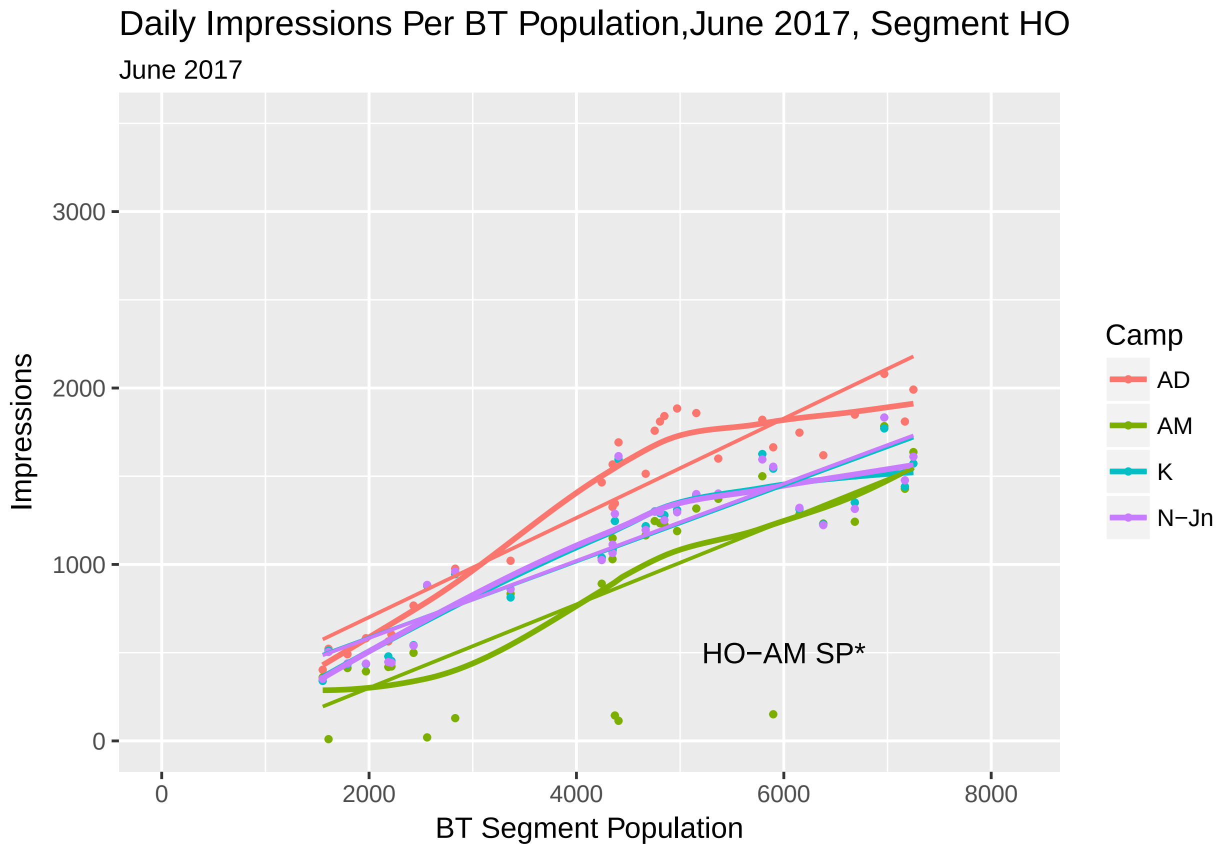
Impression Performance, All Campaigns, By Day of The Week
Weekly Pattern Is Similar Across Segments
Friday
Saturday
Sunday
Monday
Tuesday
Wednesday
Thursday
Impressions
BT Segment Size
0
2500
5000
7500
Camp
H
GH
L−Jn
AD
AM
K
N−Jn
BS
N
AMP
FV
A−Jn
B−Jn


Impression Performance, All Campaigns, By Segment
Different Segments Live at Different Spots on the Scale
C
P
E
HO
I
R
G
SY
NPH
EUR
Impressions
0 2500 5000 7500
BT Segment Size
Camp
H
GH
L−Jn
AD
AM
K
N−Jn
BS
N
AMP
FV
A−Jn
B−Jn

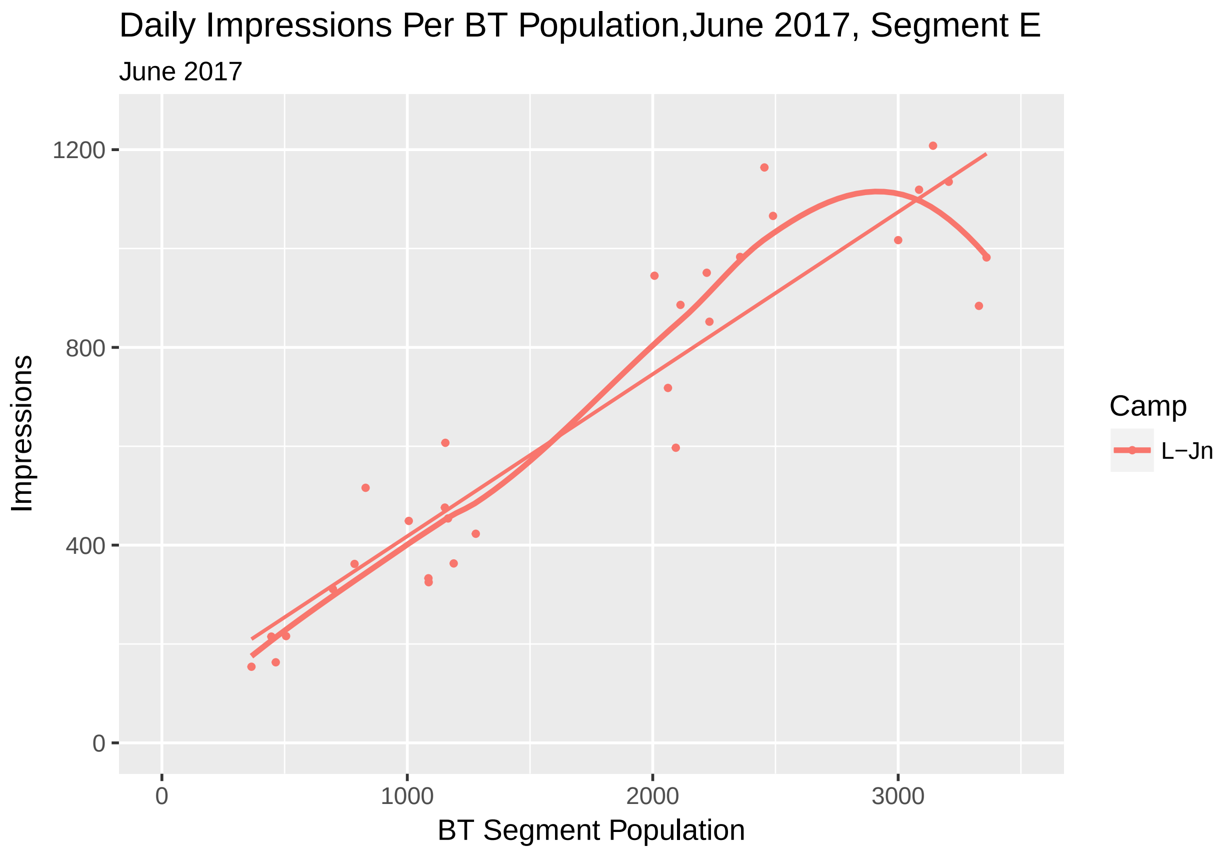
Larger Scale Segments

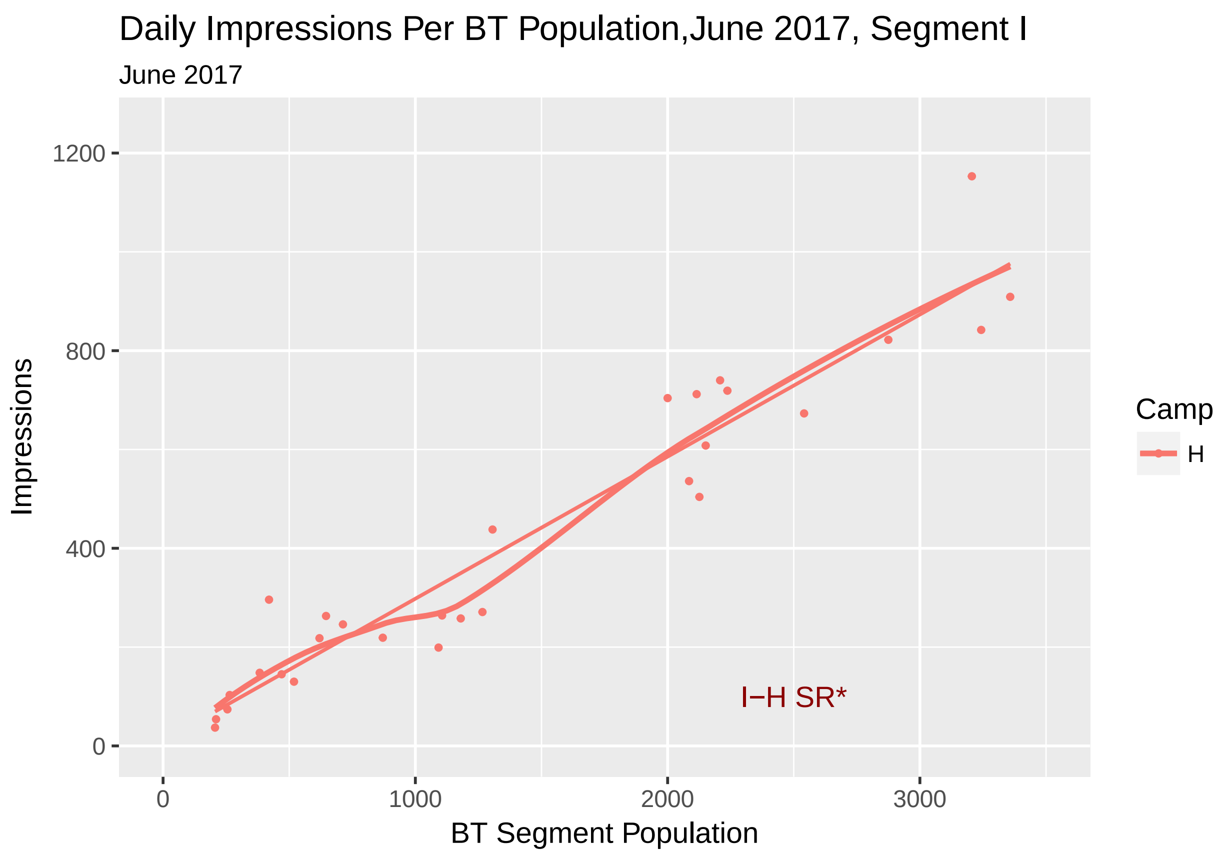


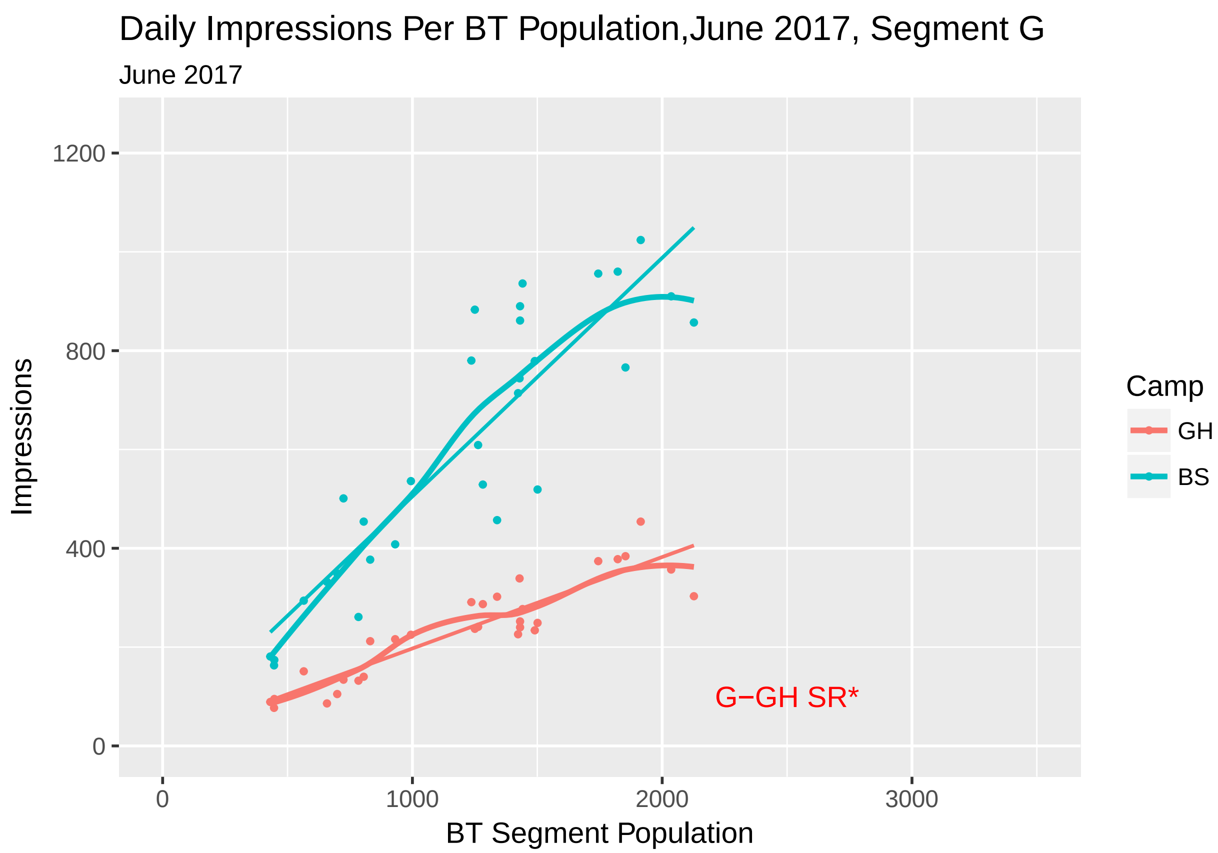

Daily Impressions Per BT Population,June 2017, Segment P
June 2017
Impressions
3000
2000
1000
0
0
2500
5000
7500
BT Segment Population
Camp
GH
P−GH SR*


Daily Impressions Per BT Population,June 2017, Segment HO
June 2017
Impressions
3000
2000
1000
0
0
2000
4000
6000
8000
BT Segment Population
HO−AM SP*
Camp
AD
AM
K
N−Jn

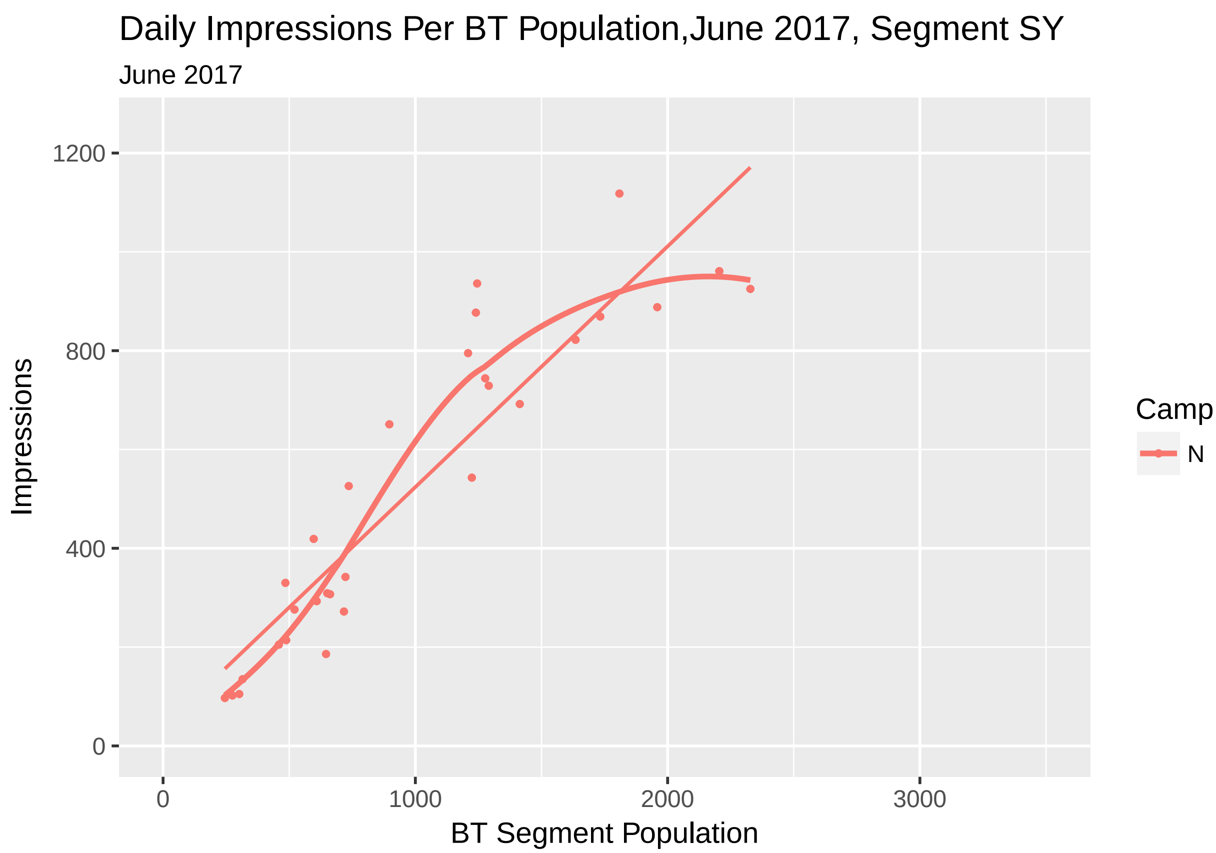
Smaller Scale Segments

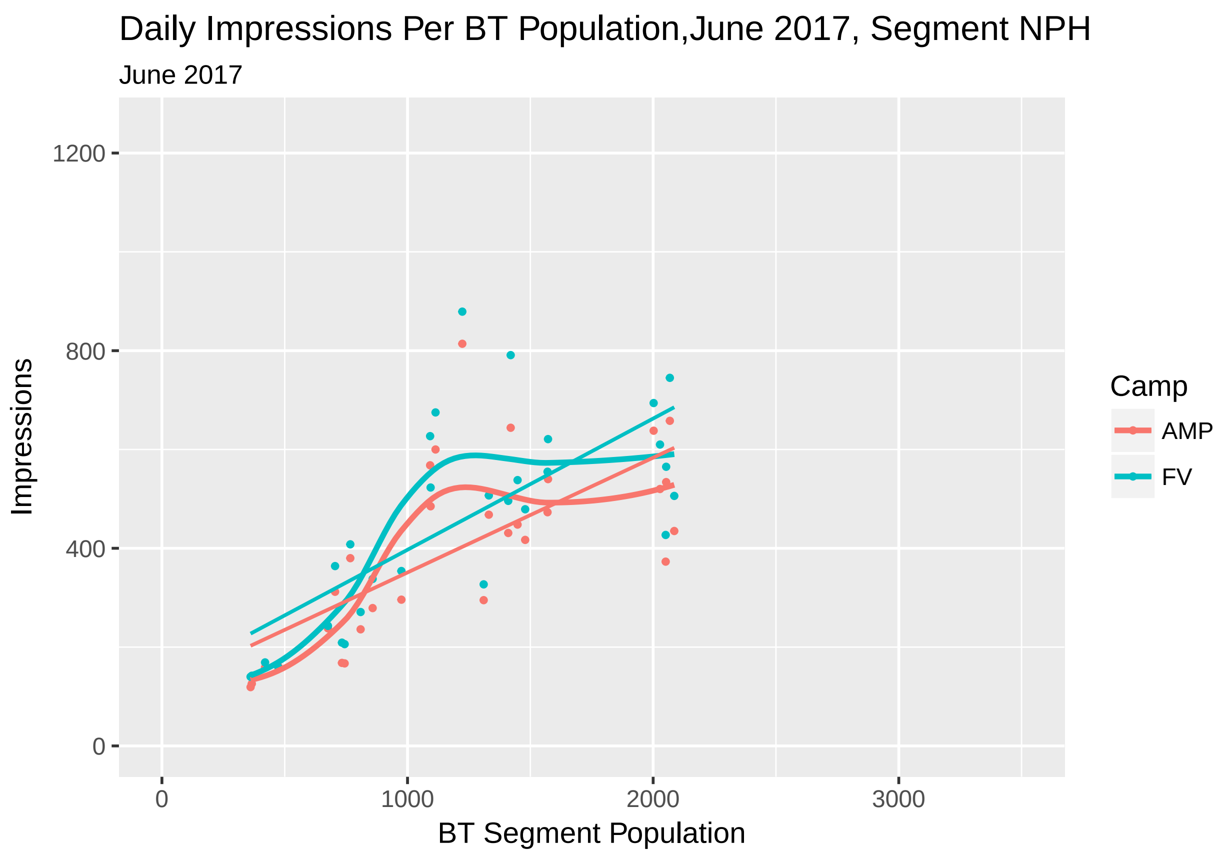



Daily Impressions Per BT Population,June 2017, Segment I
June 2017
Impressions
BT Segment Population
1200
800
400
0
0
1000
2000
3000
Camp
H
I−H SR*



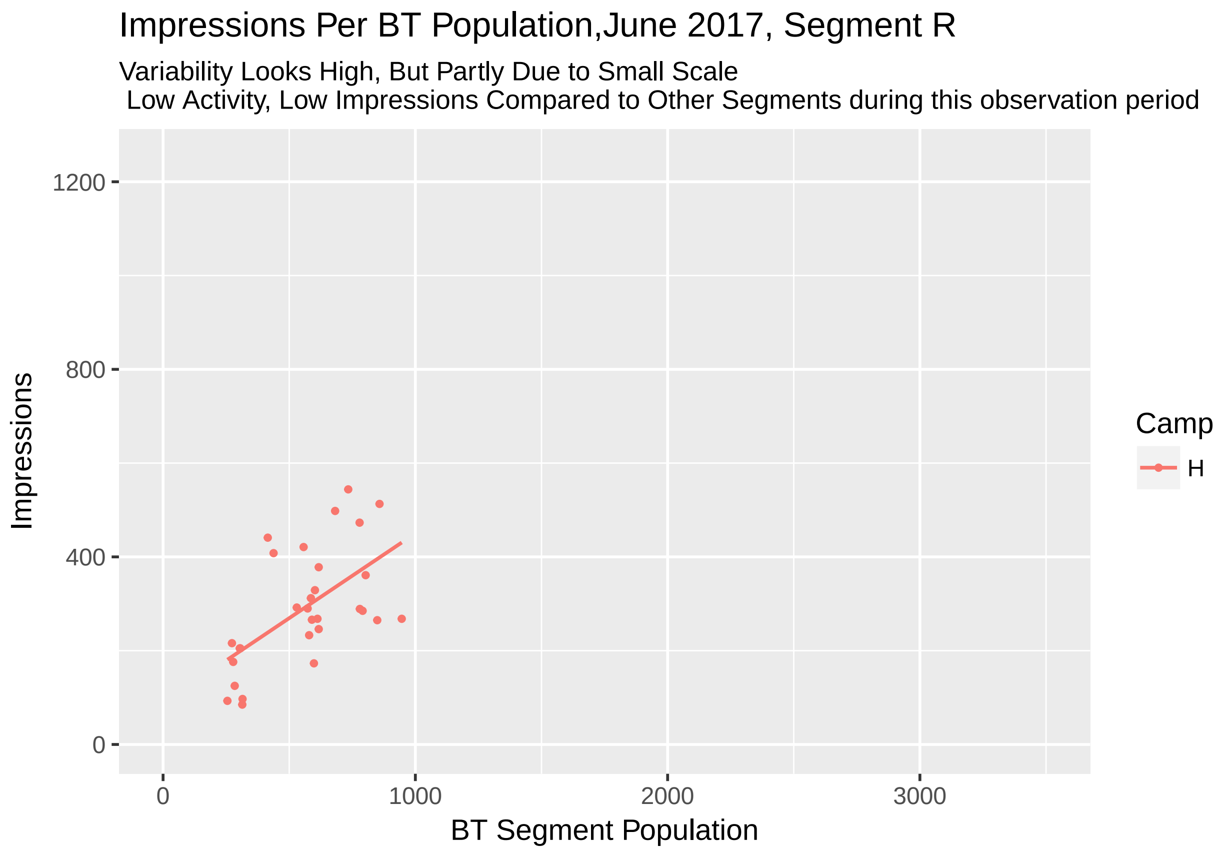
Daily Impressions Per BT Population,June 2017, Segment G
June 2017
Impressions
1200
800
400
0
0
1000
2000
3000
BT Segment Population
G−GH SR*
Camp
GH
BS

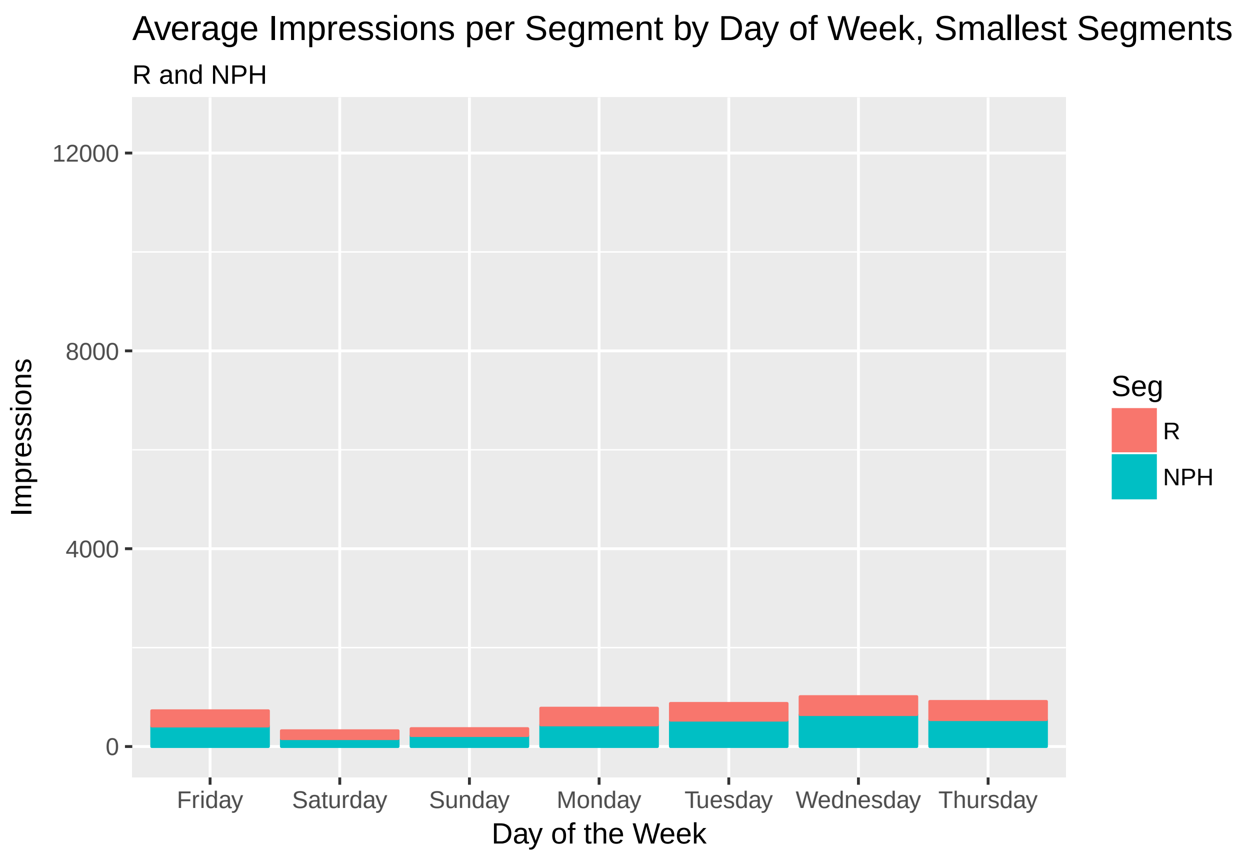

Daily Impressions Per BT Population,June 2017, Segment SY
June 2017
Impressions
1200
800
400
0
0
1000
2000
3000
BT Segment Population
Camp
N

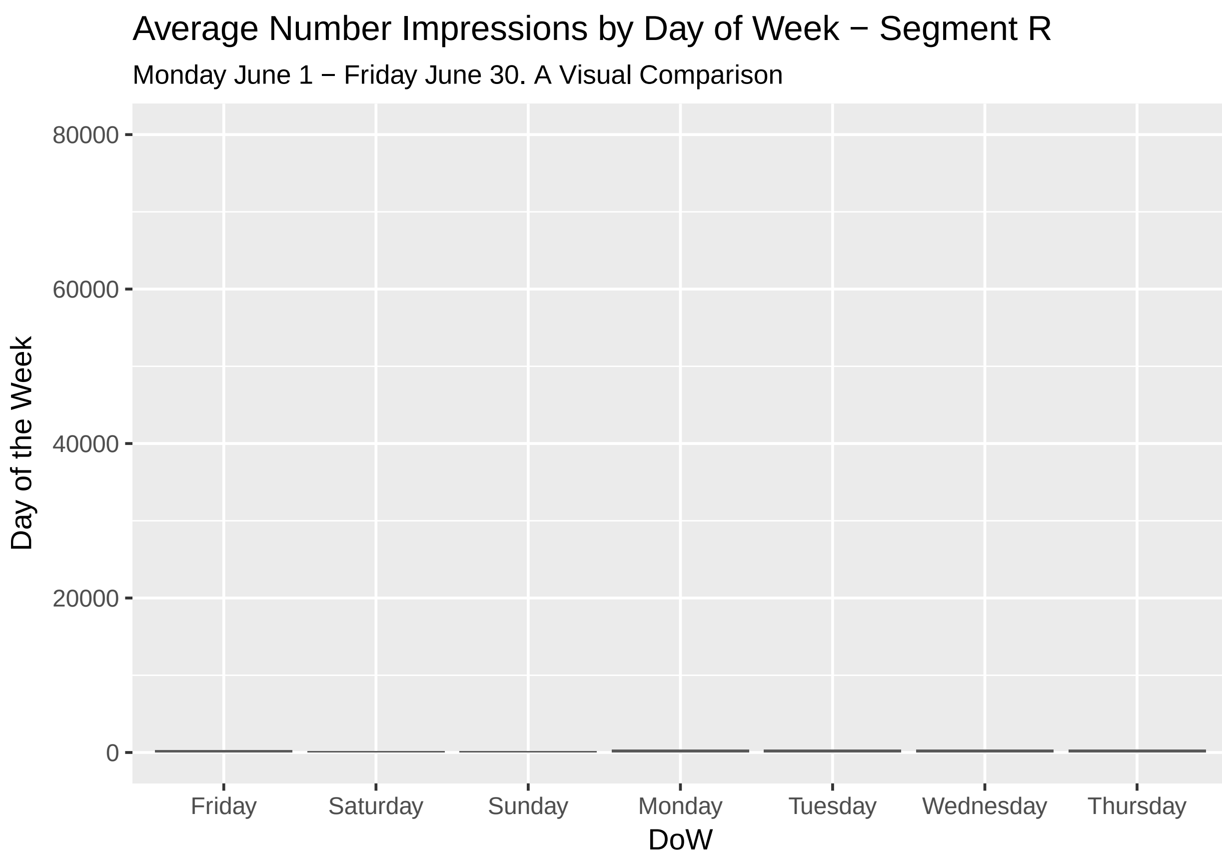

Daily Impressions Per BT Population,June 2017, Segment NPH
June 2017
Impressions
1200
800
400
0
0
1000
2000
3000
BT Segment Population
Camp
AMP
FV

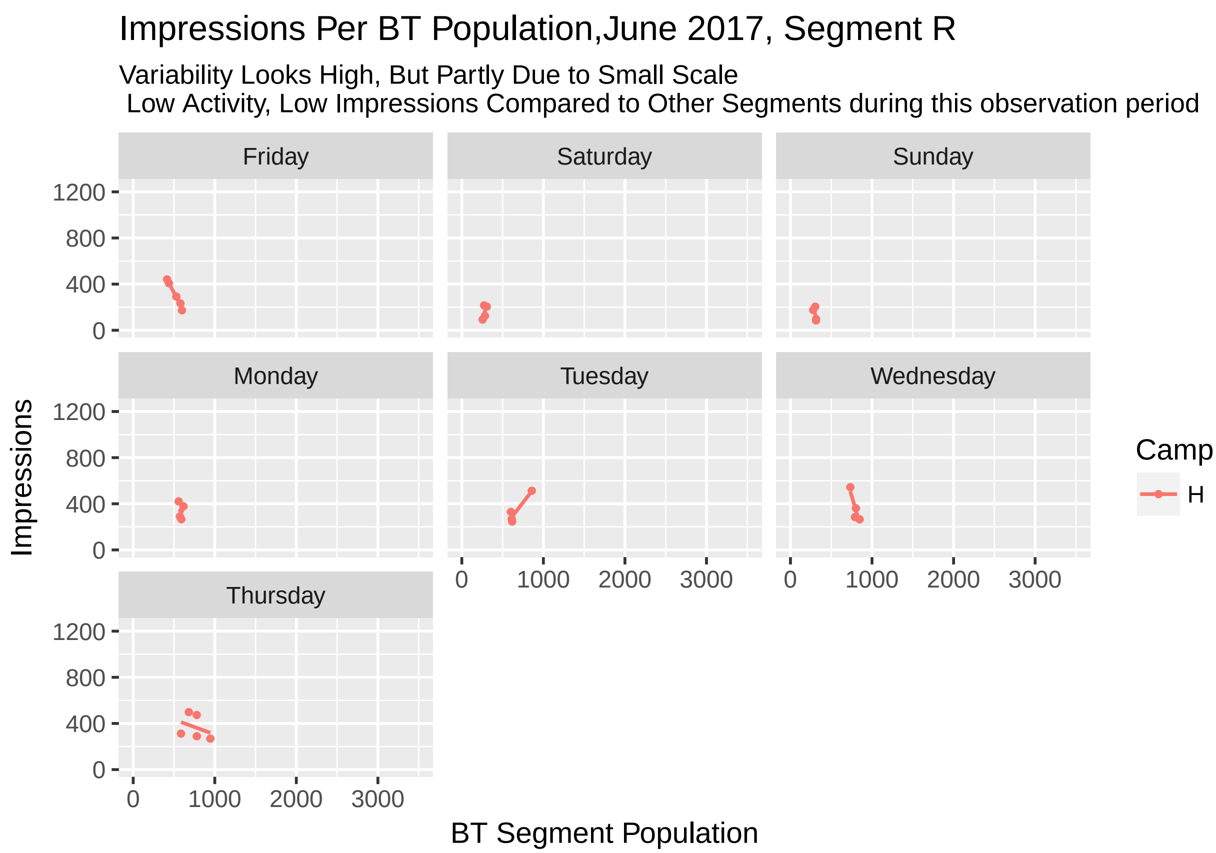
Daily Impressions Per BT Population,June 2017, Segment EUR
June 2017
Impressions
1200
800
400
0
0
1000
2000
3000
BT Segment Population
Camp
A−Jn
B−Jn

The Smallest Segments: R Investigation




Average Impressions per Segment by Day of Week, Smallest Segments
R and NPH
12000
8000
4000
0
Impressions
Friday
Saturday
Sunday
Monday
Tuesday
Wednesday
Thursday
Day of the Week
Seg
R
NPH


Average Number Impressions by Day of Week − Segment R
Monday June 1 − Friday June 30. A Visual Comparison
Day of the Week
80000
60000
40000
20000
0
Friday
Saturday
Sunday
Monday
Tuesday
Wednesday
Thursday
DoW


Impressions Per BT Population,June 2017, Segment R
Variability Looks High, But Partly Due to Small Scale
Low Activity, Low Impressions Compared to Other Segments during this observation period
Friday
Saturday
Sunday
Monday
Tuesday
Wednesday
Thursday
Impressions
1200
800
400
0
0
1000
2000
3000
BT Segment Population
Camp
H