Commercial Goals: Wifi Usage

Utilizing Synthetic Controls to Baseline Projected Performance

Insights and Analytics | Last Updated: 30 January, 2022

TangerOutlets

Table of Contents

Top Line

We built a forecast model to set a useful baseline for the business to build from. The business owner for this metric will ultimately set where were they want the business goal to land.

4 different forecasting models were applied, and are displayed below. These different models allow us to understand the potential impact of different scenarios.

- **Model 1:** Covid data included, and covid period counted as 3/1/2020 - 5/1/2021, includes growth trend
  - This model extracted a negative growth trend.
- **Model 2:** Covid data included, flattened trend.
  - This model produces the highest reasonable forecast, but is still below 2019 or 2021 values.
- **Model 3:** Covid data not included (then forecast 3 years out)
  - This model essentially is built with the belief that the trends pre-2020 will continue. Ultimately the business question here is whether shopper behavior will continue the trend that was started during to pre-pandemic behaviors in 2022.
- **Model 4:** Covid data not included (then forecast 3 years out), flattened trend.
  - This model essentially is built with the belief that the trends pre-2020 will continue, but with a flattened trend.

	2017	2018	2019	2020	2021	2022[1]	2022FC_M1	2022FC_M2	2022FC_M3	2022FC_M4	Business Goal
Jan	56.9K	27.8K	34.1K	102.6K	69.3K	58.7K	**77.3K**	**61.6K**	**632.8K**	**44.6K**	*TBD*
Feb	58.7K	34.3K	33.3K	98.0K	50.1K	-	**65.2K**	**53.8K**	**591.6K**	**42.7K**	*TBD*
Mar	74.0K	56.6K	58.1K	59.9K	104.0K	-	**89.4K**	**69.8K**	**688.0K**	**61.9K**	*TBD*
Apr	87.1K	45.5K	52.1K	3.3K	90.1K	-	**72.6K**	**57.0K**	**687.9K**	**63.6K**	*TBD*
May	73.6K	59.2K	51.4K	28.9K	97.2K	-	**62.4K**	**60.7K**	**726.3K**	**62.3K**	*TBD*
Jun	91.2K	57.3K	53.3K	80.0K	97.2K	-	**71.7K**	**75.0K**	**723.3K**	**65.6K**	*TBD*
Jul	108.9K	66.5K	85.4K	97.6K	123.2K	-	**88.8K**	**96.1K**	**780.6K**	**88.0K**	*TBD*
Aug	93.9K	56.3K	85.2K	107.9K	120.5K	-	**79.5K**	**91.5K**	**784.9K**	**78.4K**	*TBD*
Sep	23.8K	44.9K	77.8K	93.4K	82.8K	-	**55.5K**	**73.1K**	**756.4K**	**61.2K**	*TBD*
Oct	46.5K	49.4K	113.8K	41.2K	35.9K	-	**53.9K**	**69.8K**	**810.2K**	**74.2K**	*TBD*
Nov	66.4K	61.4K	175.8K	108.9K	40.8K	-	**78.5K**	**96.9K**	**818.6K**	**92.6K**	*TBD*
Dec	61.6K	73.7K	195.7K	138.3K	136.2K	-	**99.9K**	**126.3K**	**850.6K**	**108.6K**	*TBD*
total	842.6K	632.9K	1,016.2K	960.0K	1,047.2K	—	894.6K	931.6K	8,851.1K	843.7K	—

[1] Note that 2022 actuals represent the current traffic and is not a complete month

Reviewing Historical Data

Overview of Data

Data Generation

Wifi logins as captured within TangStats. We are looking for specific “promo-ids” (‘23610’, ‘24865’) as those are connected to when a user opts-in on the wifi landing page.

Data Pull from DB

For forecasting purposes we have pulled data for all centers from 2017 until now (2022.01.26).

Source: TangStats (server: appsql-prod.database.windows.net)

Tables: [dbo].[tblCustomerContacts]

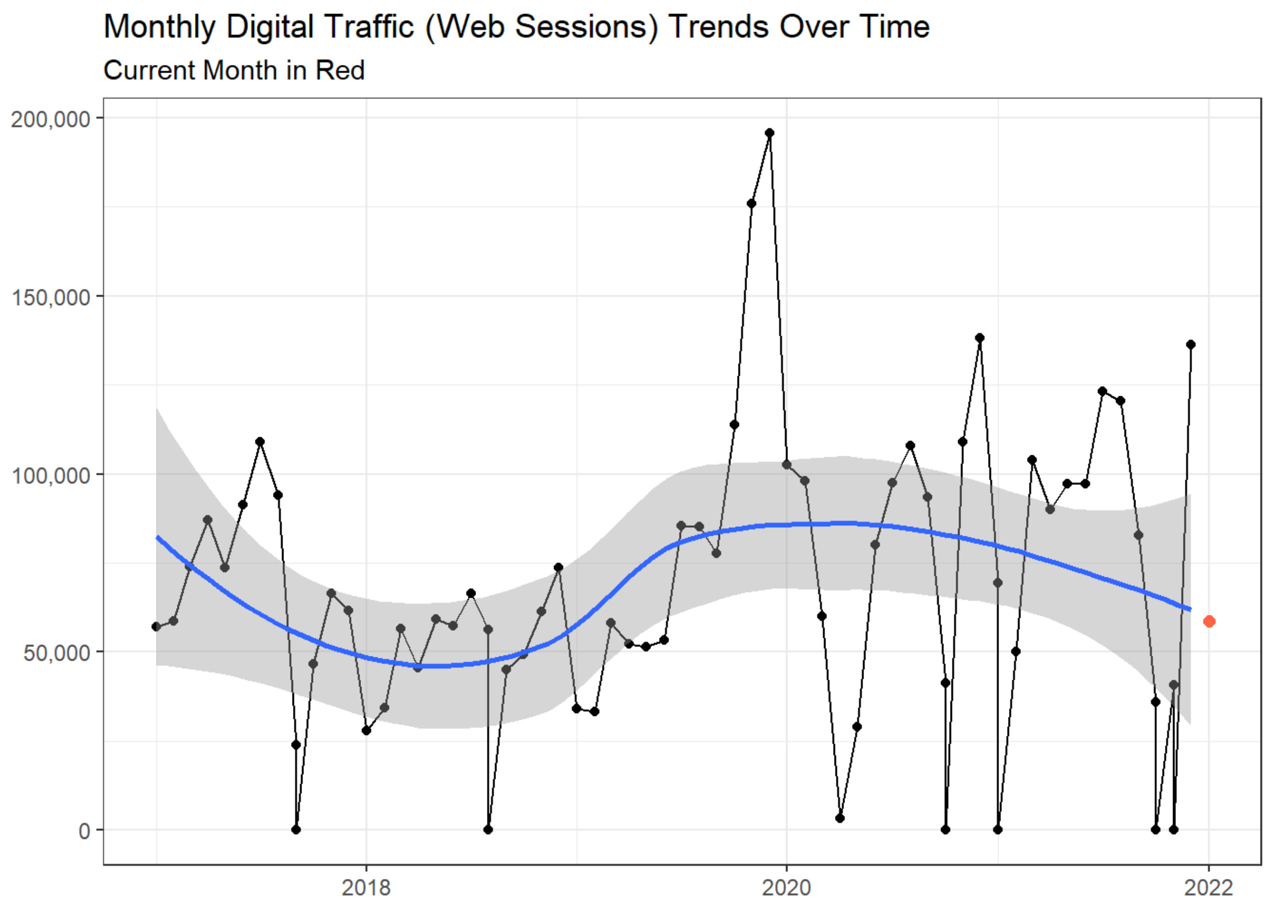
- The following centers were removed: “Bromont”, “Jeffersonville”, “Nags Head”, “Ocean City”, “Park City”, “Terrell”, “Williamsburg”, “Saint Sauveur”
    - This is because these centers are no longer in the portfolio
- Also if the center was identified as: “Corporate”,“No Center Assigned”, then those records were removed.
    - Records associated with these centers are assumed to be either created in error (No Center Assigned) or were part of an internal Dev test (Corporate)

Dealing with ZEROs/NULLS
In reviewing the data there are days where no wifi logins were captured. (ie near the end of year 2021). It is believed that this is an issue with the data generation process and not infact true zeros. For the forecasting model, the zeros have been removed and the forecast interpolates values based around “nearby” values.

View of Trends

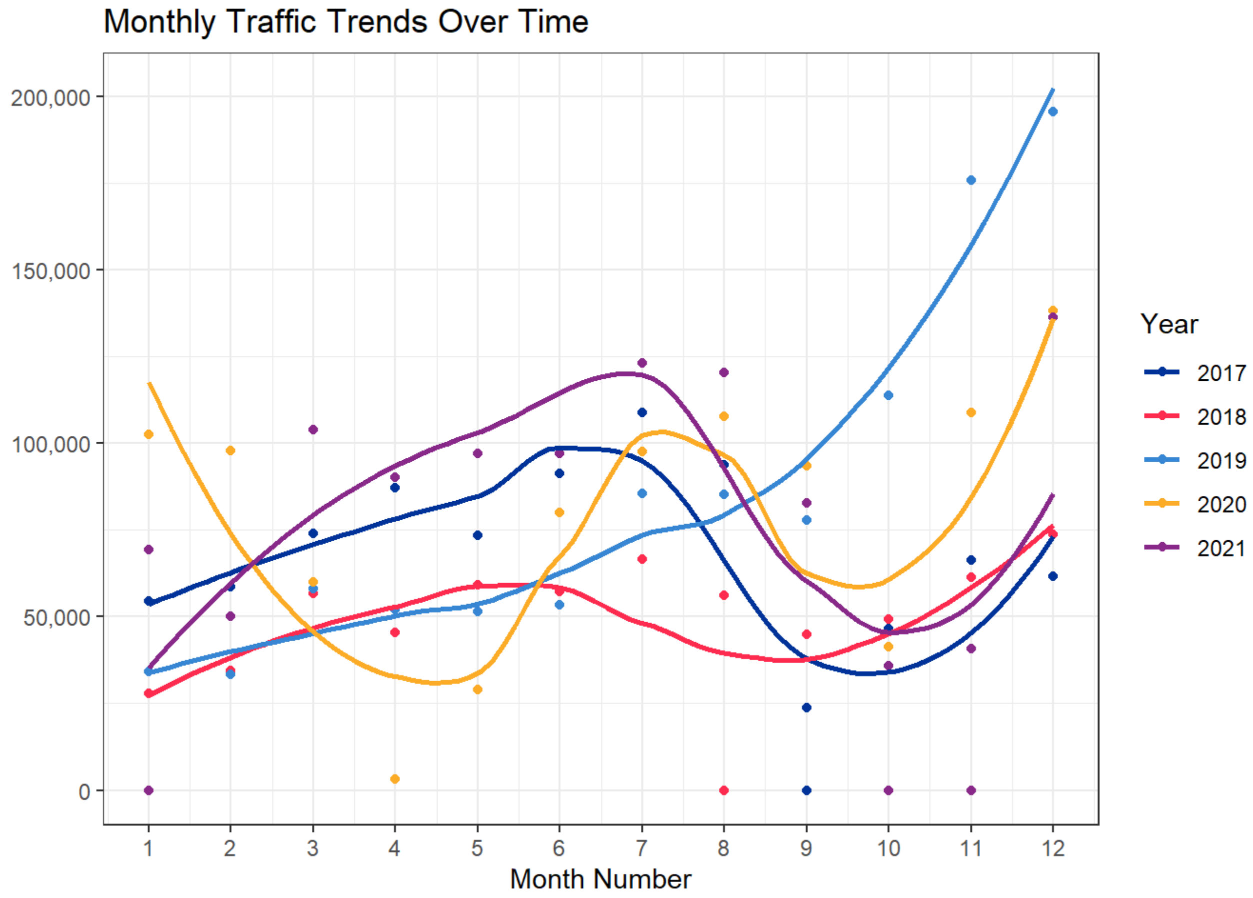
Before we begin to model this work, we want to visualize the general trends. This allows us to to flag for large outliers or anomalies within our data.

- With this data we can see how significant Covid shutdowns had on traffic throughout 2020.
- We also see a fairly consistent month to month seasonal variation.



Monthly View

Each color line represents a different year's worth of traffic aggregated on a monthly basis.


Monthly Traffic Trends Over Time
200,000
150,000
100,000
50,000
0
1
2
3
4
5
6
7
8
9
10
11
12
Month Number
Year
2017
2018
2019
2020
2021

Forcasting model

Why Synthetic Controls?

This method can account for the effects of confounders changing over time, by weighting the control group to better match the treatment group before the intervention. Another advantage of the synthetic control method is that it allows researchers to systematically select comparison groups. It has been applied to the fields of political science, health policy, criminology, and economics.

In our case, we are using this methodology to help set a baseline of where the business would be if we continue with the same historical processes and procedures as before.

Step 1: Build Forecast Model

Building model for all entire portfolio.

- In this case, we are already into 2022, but we want to build the forecast model excluding the new year.
- In the future data frame, we go forwards 365 days (1 year) to be able to predict against.

```
Traffc <-
  CleanWifi %>%
  filter(Date < as.Date("2022-01-01")) %>%
  mutate(ds = Date) %>%
  group_by(ds) %>%
  summarise(y = sum(DistcCust, na.rm=T)) %>%
  ungroup()

# m <- prophet(Traffc, holidays = covid)
m <- prophet(holidays = covid)
m <- add_country_holidays(m, country_name = 'US')
m <- fit.prophet(m, Traffc)

future <- make_future_dataframe(m, periods = 365)
# tail(future)


forecast <- predict(m, future)
# tail(forecast[c('ds', 'yhat', 'yhat_lower', 'yhat_upper')])
```

Forecast Results

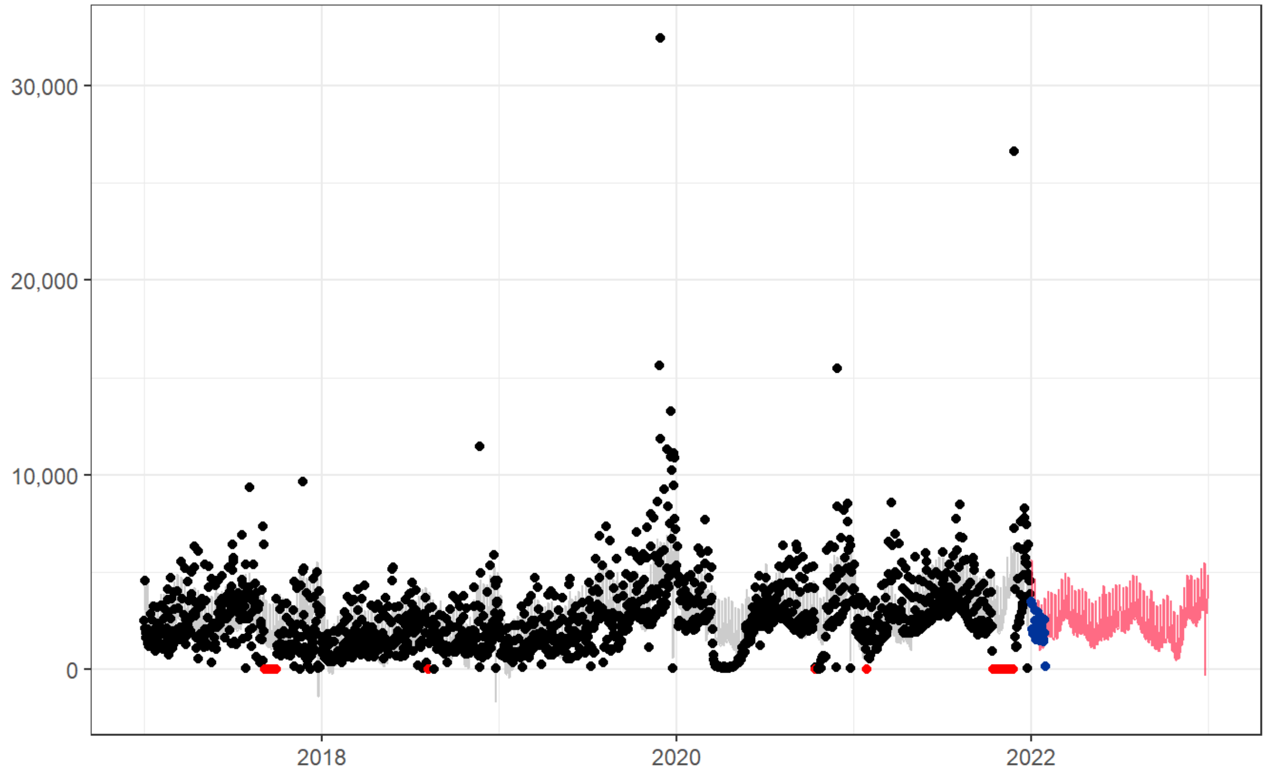
Below is the forecast model.

- The black dots represent to the total portfolio traffic by day.
- The blue line represents the forecasted values.

We see a strong degree of seasonality

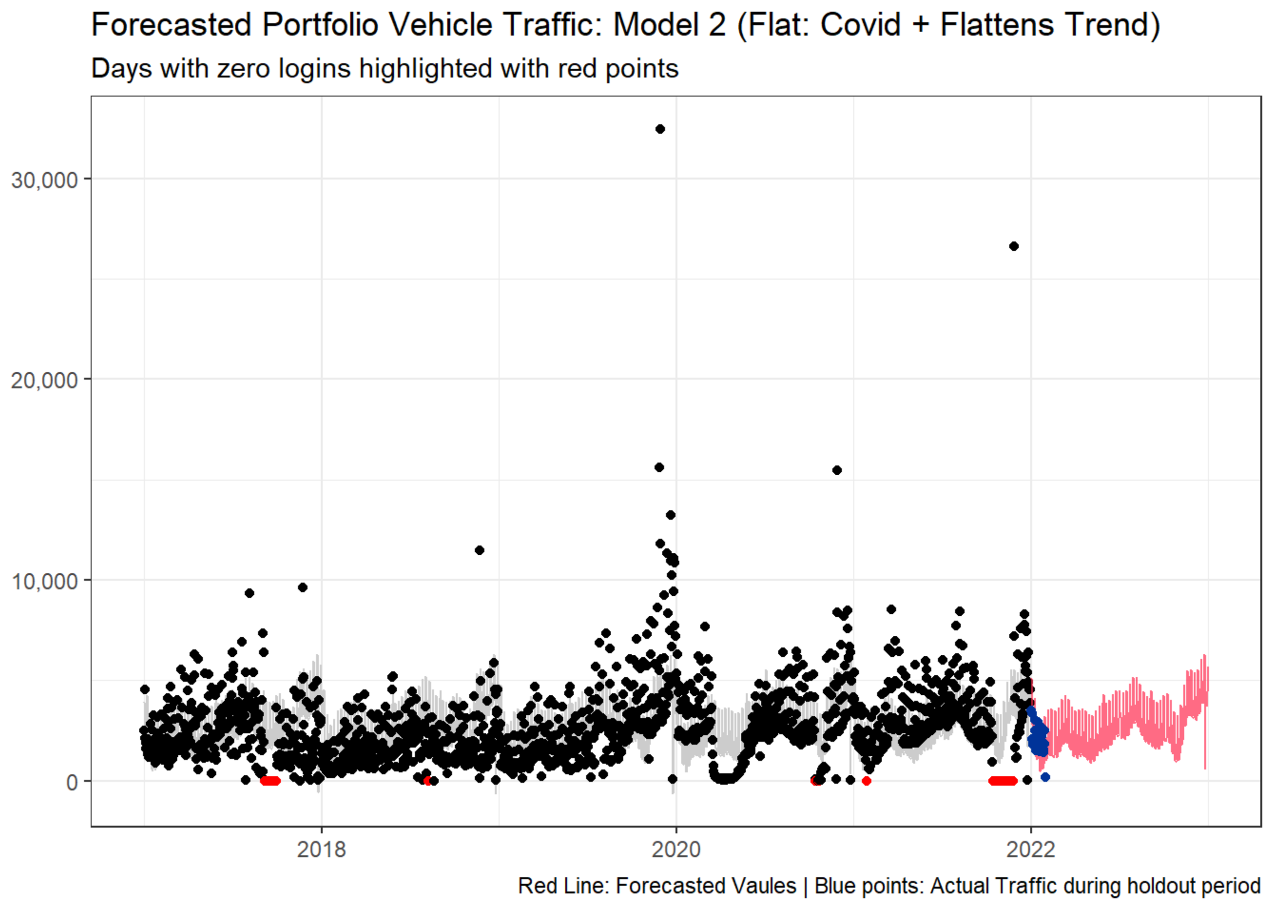
Forecast Model 1

Forecasted Portfolio Vehicle Traffic: Model 1 (Mid: Covid + Keeps Trend)
Days with zero logins highlighted with red points



Red Line: Forecasted Vaules | Blue points: Actual Traffic during holdout period

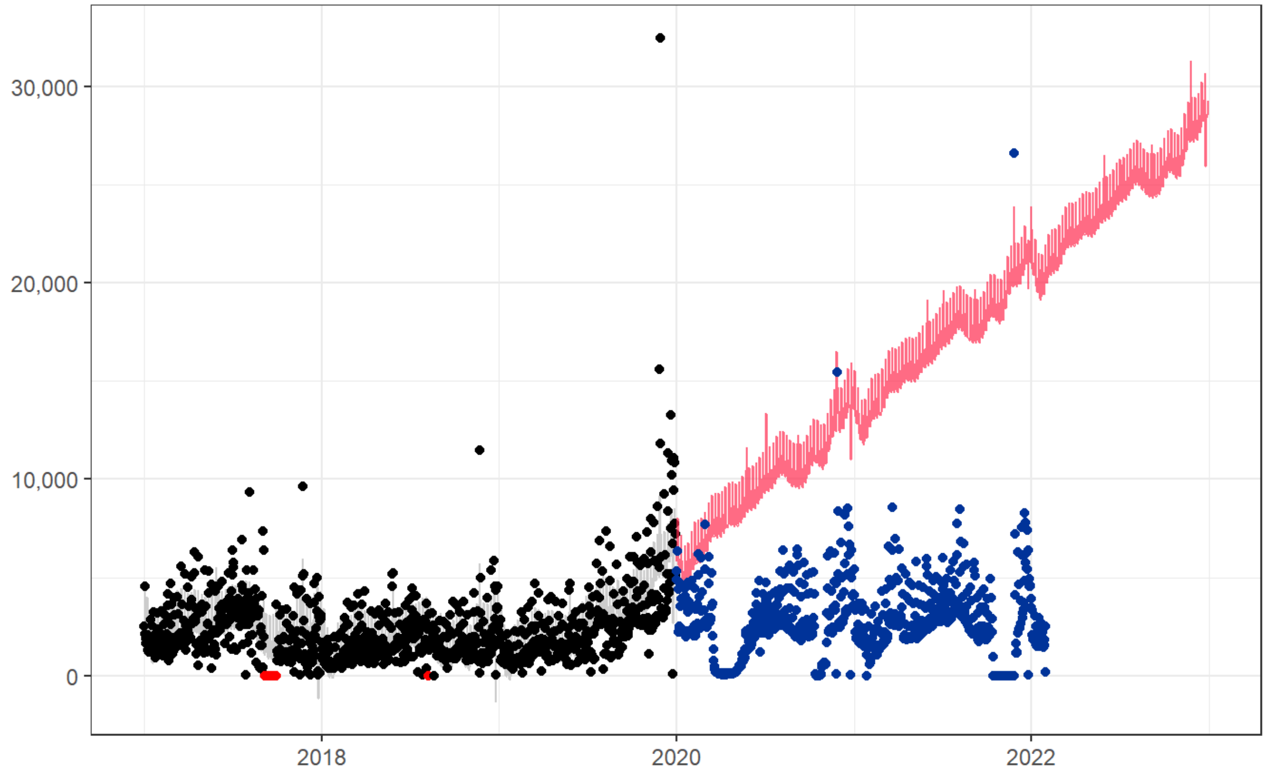
Forecast Model 2



Forecast Model 3

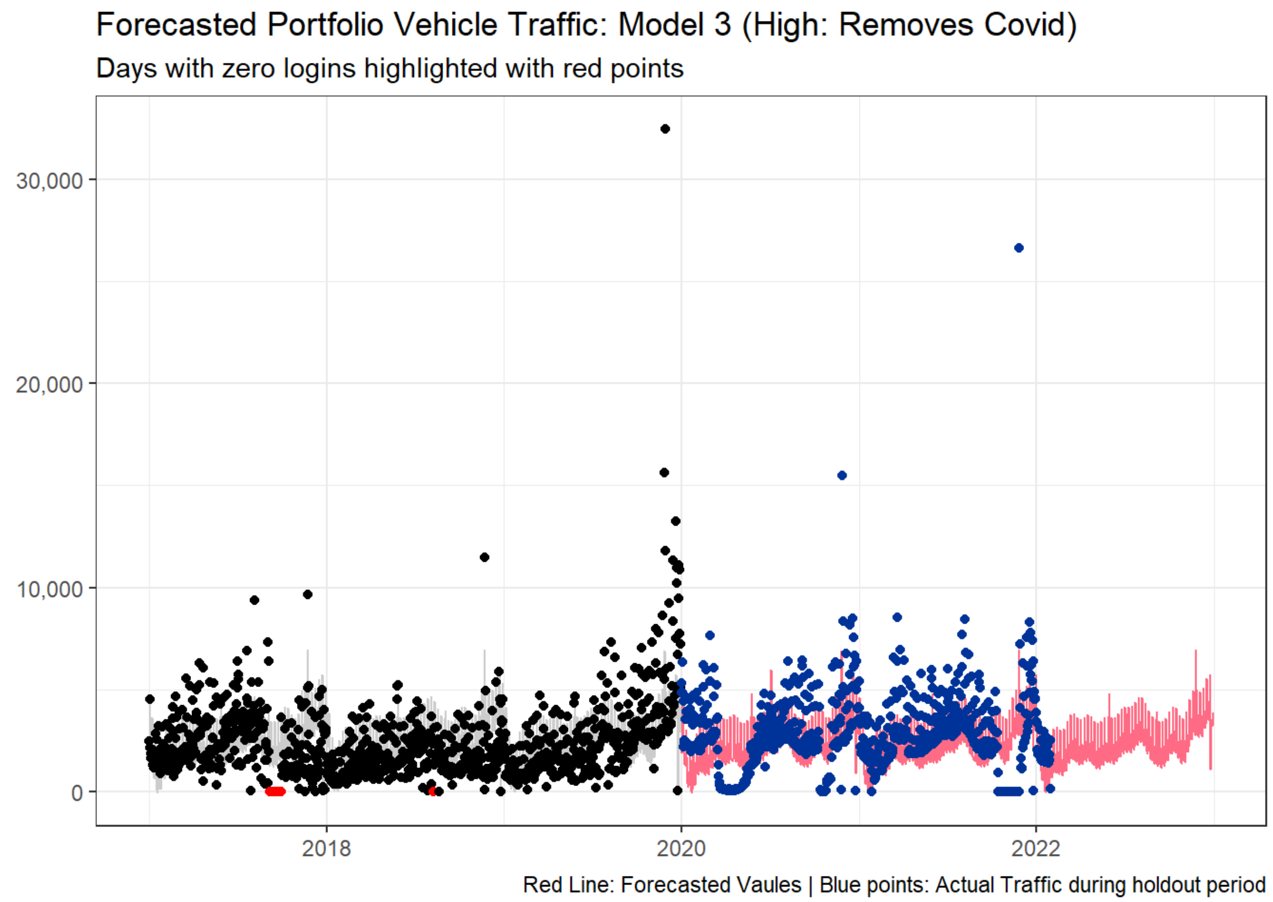
Forecasted Portfolio Vehicle Traffic: Model 3 (High: Removes Covid)

Days with zero logins highlighted with red points



Red Line: Forecasted Vaules | Blue points: Actual Traffic during holdout period

Forecast Model 4



Forecast Components

A forecast model can be broken up into the different elements that add (or multiply) together. By splitting them out and visualizing them, we then can view the index values and assess the impact of different items on the results

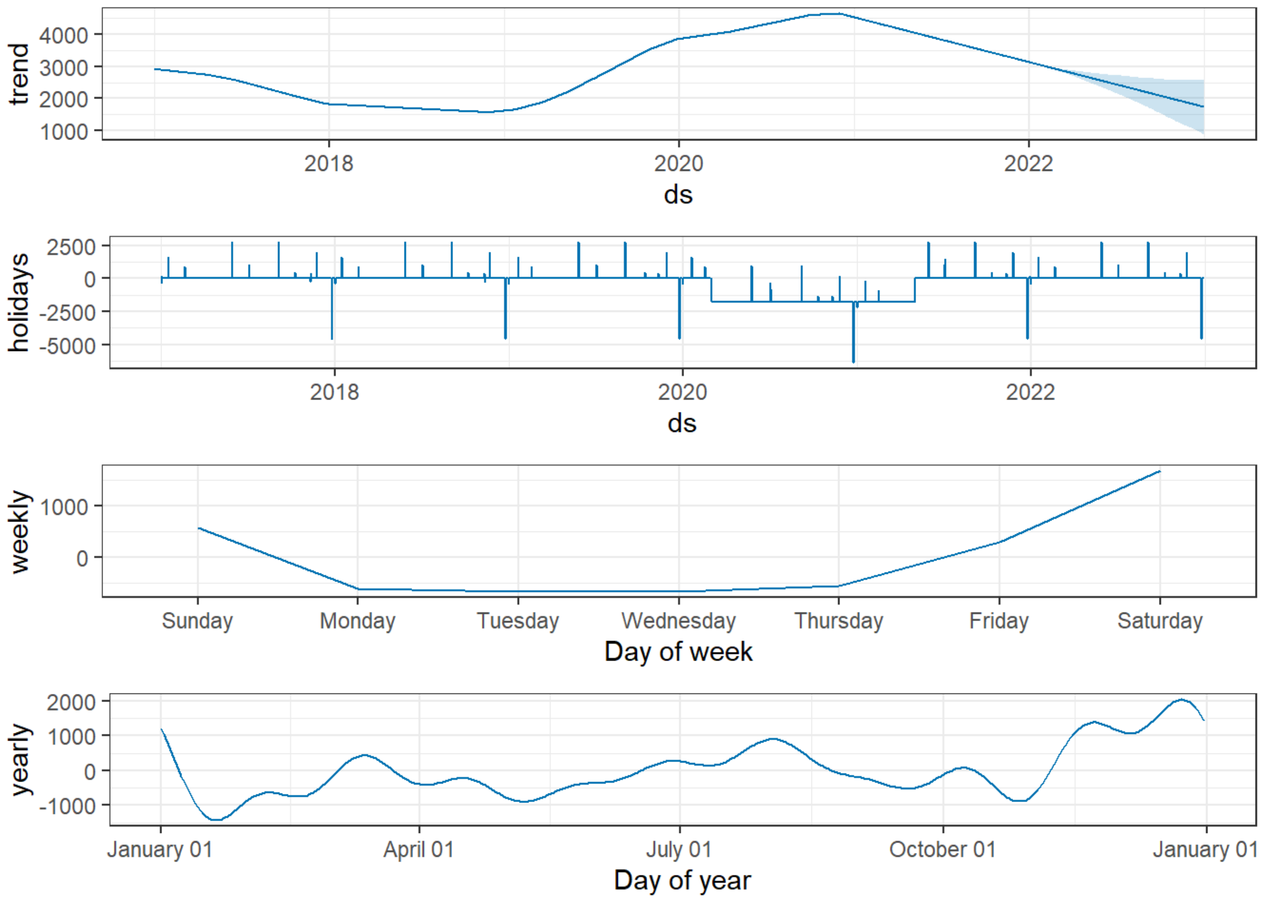
By breaking down the forecast into it's component parts we see a couple of things:

- Trend: This element shows what the general trend is year over year.
    - In this case, there is a forecasted to be a downwards trend for future years.
- Holidays: This elements flags both the US holidays as well as the date range from 3/1/2020 - 5/1/2021 as the main effect of Covid shutdowns.
- Weekly: This element shows how traffic varies within the week.
    - The outlets being a mainly weekend business really shows up in this view, with the positive indices being for Friday, Saturday, Sunday.
- Yearly: This is the element of the within year seasonality.
    - In general we see that there is a general ramp up throughout the first part of the year, a small lull in October, and then our biggest traffic volumes seen during the end of year.

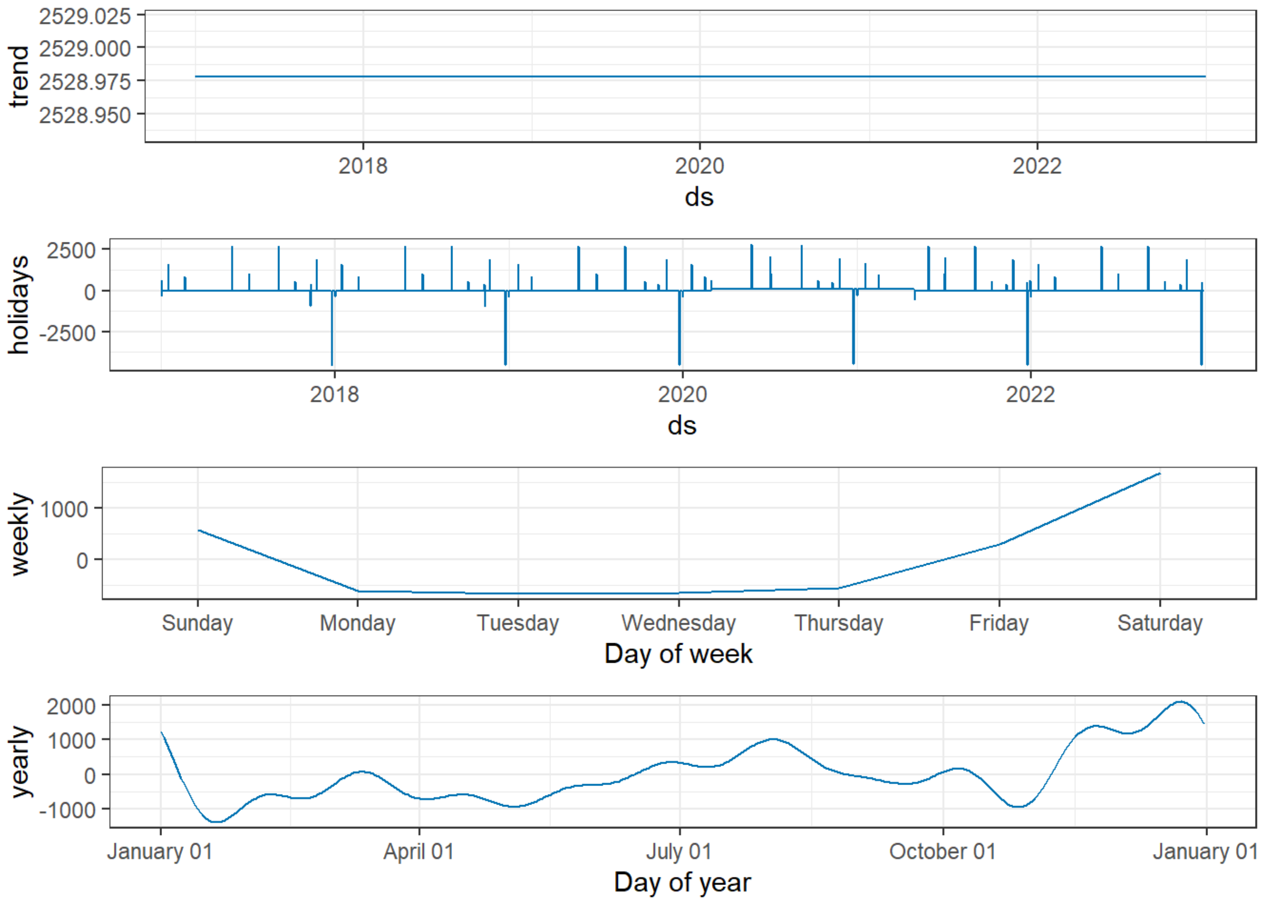
Note about negative values: These 4 items below are added together to create the forecast model. This is why there are some items that are negative and some that are positive. A negative value does not mean that we would forecast negative vehicles, but instead that at those instances the traffic would decrease from the trend.

- An example here would be the impact of Christmas day on digital traffic, where with the centers closed it's less likely people are logging into the website.
- We do see a yearly spike around Thanksgiving/Black Friday time period.

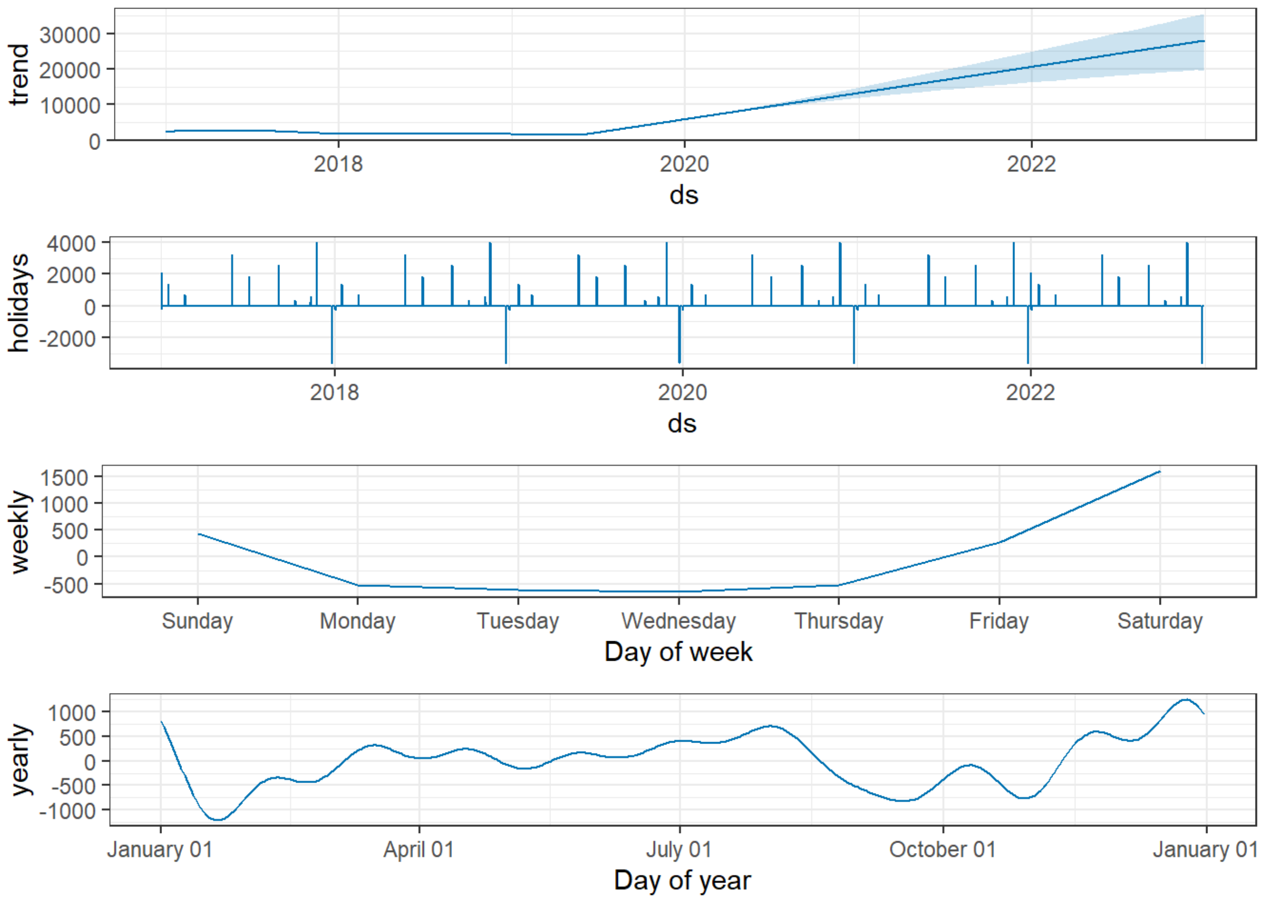
Model 1



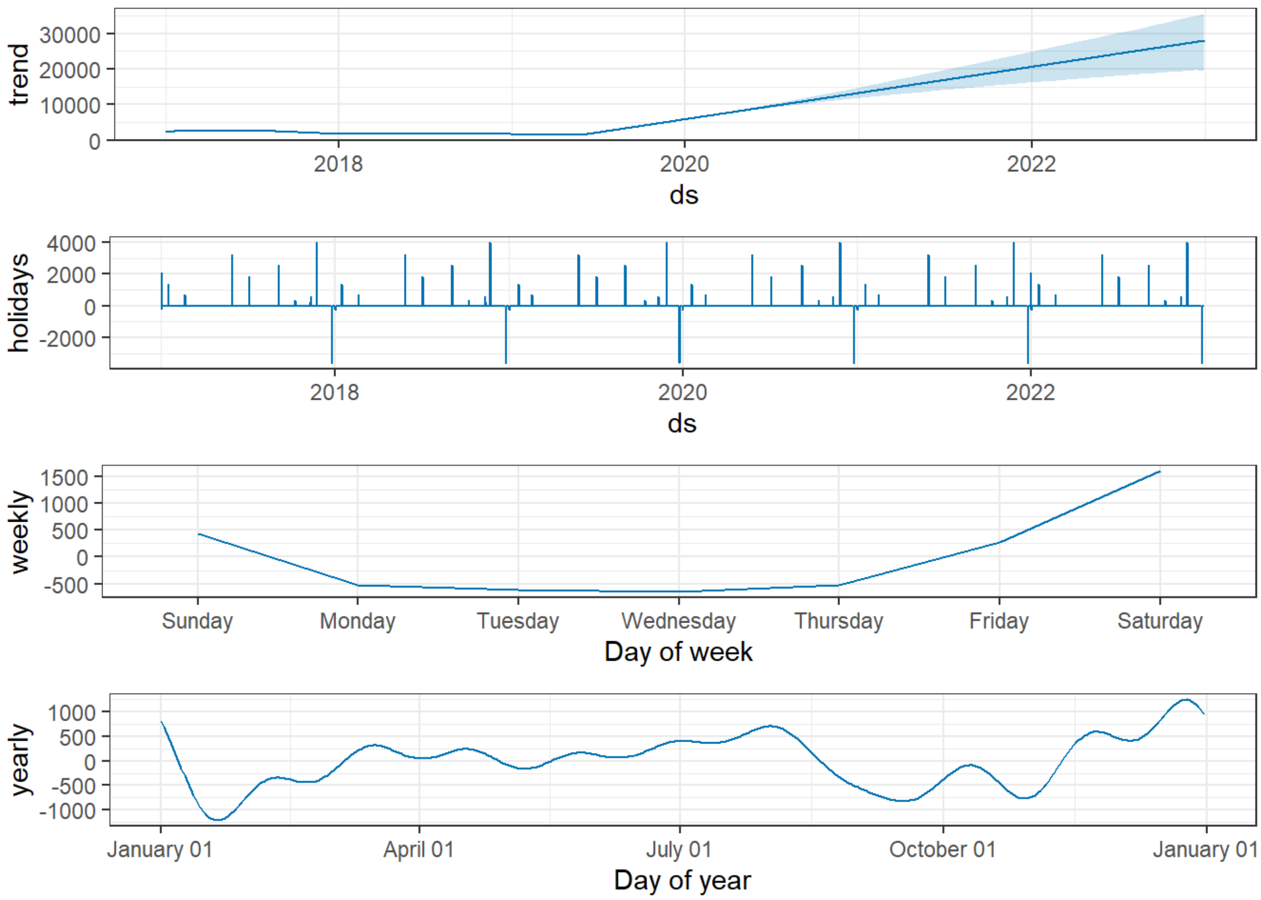
Model 2


2529.025
2529.000
2528.975
2528.950
trend
2018
2020
2022
ds
2500
0
-2500
holidays
2018
2020
2022
ds
1000
0
weekly
Sunday
Monday
Tuesday
Wednesday
Thursday
Friday
Saturday
Day of week
2000
1000
0
-1000
yearly
January 01
April 01
July 01
October 01
January 01
Day of year

Model 3


trend
30000
20000
10000
0
2018
2020
2022
ds
holidays
4000
2000
0
-2000
2018
2020
2022
ds
weekly
1500
1000
500
0
-500
Sunday
Monday
Tuesday
Wednesday
Thursday
Friday
Saturday
Day of week
yearly
1000
500
0
-500
-1000
January 01
April 01
July 01
October 01
January 01
Day of year

Model 4


trend
30000
20000
10000
0
2018
2020
2022
ds
holidays
4000
2000
0
-2000
2018
2020
2022
ds
weekly
1500
1000
500
0
-500
Sunday
Monday
Tuesday
Wednesday
Thursday
Friday
Saturday
Day of week
yearly
1000
500
0
-500
-1000
January 01
April 01
July 01
October 01
January 01
Day of year

Next Steps

The “Top Line” section presented both the actuals traffic by month for prior years with the forecasted values for 2022 (bolded).

The forecasted values could represent what would be considered as the baseline for goal setting by the business. Upon which the business leaders responsible for this metric would review and adjust the values according to where they believe they can made an impact. It will be up to the business to determine what strategies they will implement and adjust the goals accordingly.

This analysis body of work is set up to allow for the adjustment of goals based upon the values decided upon by the business owners.

Appendix

SQL: Wifi Logins

```
SELECT cc.CustomerID,  - cc.CustomerContactID AS TransactionID,
    - cc.CustomerContactID AS TransactionDetailID,
    cast(cc.ContactDate AS DATE) as TransactionDate,c.CenterName,
    'A' AS Status, coalesce(p.Description, '')  AS DetailDesc,
    'Customer Contact' AS TransactionHistoryType,
    cc.PromoID
    --, cc.Description as promo_desc
FROM tblCustomerContacts cc
INNER JOIN tblCustomerContactTypes cct ON cc.ContactTypeID =
        cct.CustomerContactTypeID
INNER JOIN tblPromos p ON cc.PromoID = p.PromoID
LEFT JOIN tblcenters c ON cc.centerid = c.centerid

WHERE year(cc.contactDate) >= 2019
    and cc.PromoID <> 0
and cc.promoid in ('23610', '24865') --wifi landing page opt-in
```

Forecast Model

The forecast model code is available within the I&A OneDrive.

Tanger®Outlets