Commercial Goals: New Members

Utilizing Synthetic Controls to Baseline Projected Performance

Insights and Analytics | Last Updated: 01 February, 2022

Table of Contents

Top Line

We built forecast models 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.

For the new member model acquisition model we forecasted new members from center and new members from web separately as those two different channels are managed differently. We also observe they are were affected by external factors differently (ie COVID). Therefore it is not reasonable to combine these two data sets together when forecasting.

3 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
- **Model 2:** Covid data included, utlizing a logistic curve
- **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.

Projected New Members from Center Teams

	2017	2018	2019	2020	2021	2022[1]	2022FC_M1C	2022FC_M2C	2022FC_M3C	Business Goal
Jan	5,424	7,891	6,825	5,132	1,574	1,788	**2,477**	**3,085**	**6,474**	*TBD*
Feb	6,765	8,389	6,391	9,612	1,651	-	**3,131**	**4,010**	**6,634**	*TBD*
Mar	10,677	10,124	10,508	4,899	5,904	-	**5,176**	**6,296**	**9,588**	*TBD*
Apr	12,095	8,390	10,895	1	4,740	-	**4,509**	**5,686**	**10,303**	*TBD*
May	11,429	9,269	10,748	1,466	4,813	-	**3,551**	**4,809**	**9,838**	*TBD*
Jun	8,022	13,776	11,852	3,078	4,542	-	**4,249**	**5,597**	**10,383**	*TBD*
Jul	12,600	12,832	16,234	6,524	8,885	-	**7,254**	**8,780**	**13,474**	*TBD*
Aug	12,958	14,310	19,023	8,270	8,590	-	**8,096**	**9,716**	**14,376**	*TBD*
Sep	8,077	9,248	8,261	5,781	5,865	-	**3,020**	**4,549**	**7,415**	*TBD*
Oct	7,399	5,934	6,250	3,874	5,634	-	**2,145**	**3,384**	**6,180**	*TBD*
Nov	13,418	11,116	11,007	6,012	6,612	-	**4,584**	**6,376**	**10,220**	*TBD*
Dec	11,257	11,639	9,620	4,842	5,626	-	**4,276**	**6,157**	**9,704**	*TBD*
total	120,121	122,918	127,614	59,491	64,436	—	52,467	68,446	114,590	—

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

Projected New Members from Web

	2017	2018	2019	2020	2021	2022[1]	2022FC_M1W	2022FC_M2W	2022FC_M3W	Business Goal
Jan	368	790	2,141	1,413	1,110	1,549	**3,291**	**2,821**	**2,890**	*TBD*
Feb	422	767	892	1,284	1,266	-	**2,825**	**2,351**	**2,301**	*TBD*
Mar	797	1,034	1,657	845	2,254	-	**3,584**	**3,054**	**2,885**	*TBD*
Apr	703	947	1,170	75	1,875	-	**3,262**	**2,696**	**2,706**	*TBD*
May	674	1,119	1,469	735	2,451	-	**3,519**	**2,826**	**2,935**	*TBD*
Jun	638	1,471	1,631	1,277	2,159	-	**3,632**	**2,905**	**2,947**	*TBD*
Jul	1,100	1,951	2,328	1,968	4,003	-	**4,613**	**3,806**	**3,692**	*TBD*
Aug	1,539	2,375	3,253	2,248	4,214	-	**5,012**	**4,146**	**4,140**	*TBD*
Sep	840	1,588	1,140	2,056	2,851	-	**3,980**	**3,087**	**2,967**	*TBD*
Oct	1,076	870	1,350	1,605	2,539	-	**3,927**	**2,945**	**2,964**	*TBD*
Nov	1,475	1,813	2,544	2,548	4,811	-	**4,878**	**3,869**	**3,626**	*TBD*
Dec	1,221	1,506	1,981	2,515	3,202	-	**4,599**	**3,510**	**3,335**	*TBD*
total	10,853	16,231	21,556	18,569	32,735	—	47,122	38,014	37,387	—

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

Reviewing Historical Data

Overview of Data

Data Generation

For this analysis we have focused on analyzing the count of new members by day. The data is then split on whether the source of the membership was the Web or Center team (as identified by the “FromWeb” field in the database.

Data Pull from DB

For forecasting purposes we have pulled data for all centers until now (2022.01.28). The forecast model excluded any data before 2017 as values were much lower.

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

Tables: [dbo].[tblTransactionDetail]

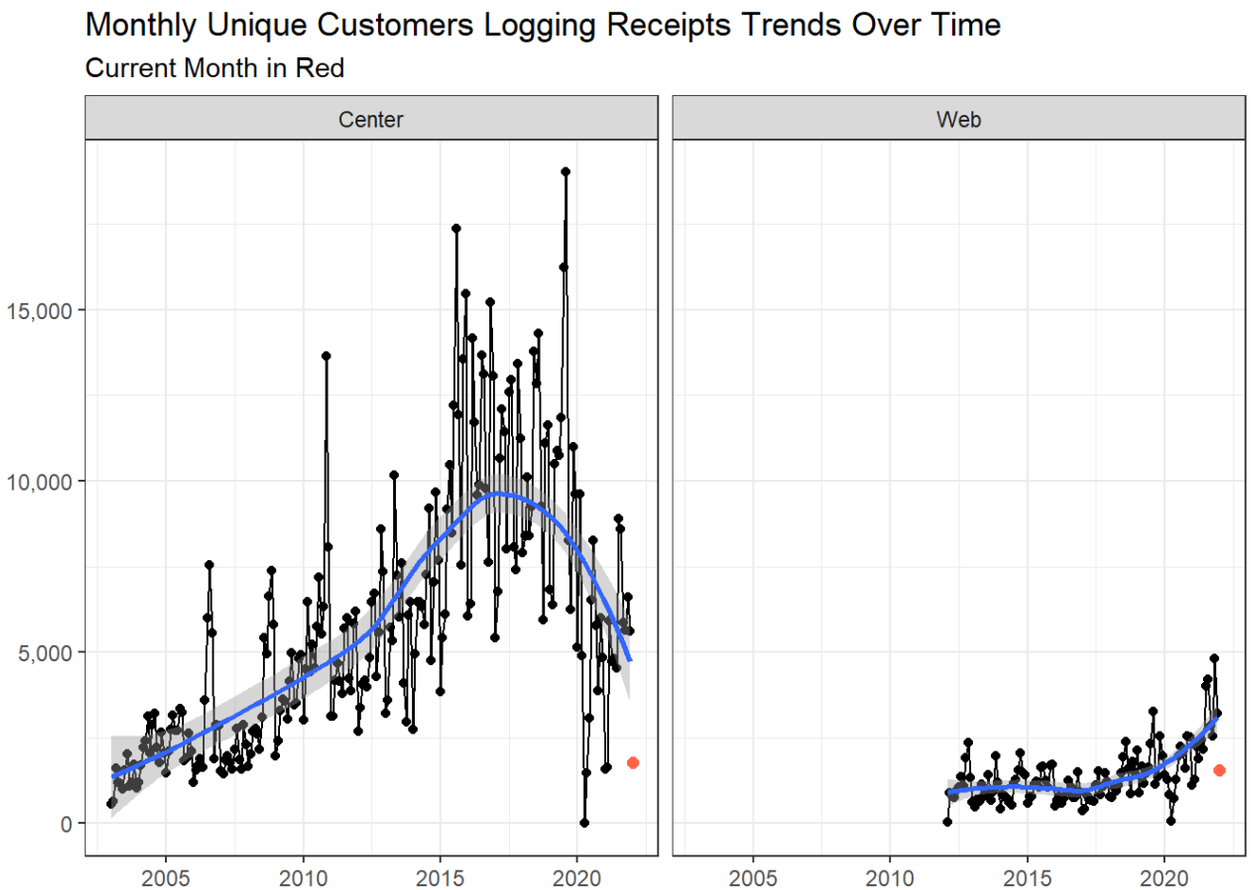
We were pull instances where the product id was “1” (New TC Membership)
* Removing records where there were status wasn’t either “A” or “N”
* Removing records where there was a deleted timestamp.

The following centers were removed from the historical counts:
* “NE Prop/N Conway”, “McMinnville”, “Barstow”, “Boaz”, “Bourne”, “Branson”, “Bromont”, “Burlington”, “Casa Grande”, “Corporate”, “Dalton”, “Jeffersonville”, “Kittery”, “Lincoln City”, “Martinsburg”, “McMinnville”, “Nags Head”, “No Center Assigned”, “North Branch”, “Ocean City”, “Park City”, “Pigeon Forge”, “Saint Sauveur”, “Sanibel”, “Seymour”, “Stroud”, “Terrell”, “TEST CENTER”, “Tuscola”, “Vero Beach”, “West Branch”, “Westbrook”, “Williamsburg”, “Wisconsin Dells”

View of Trends

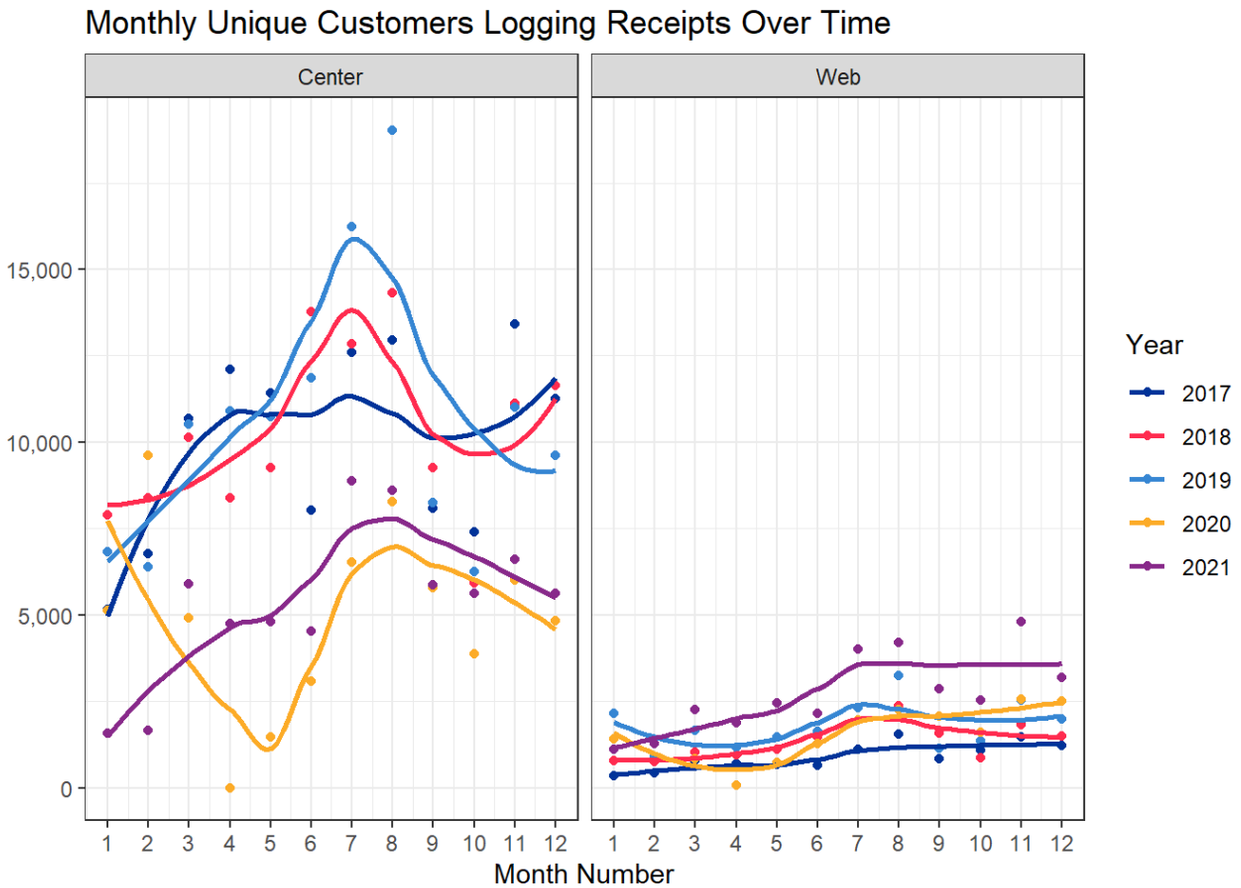
Before we begin to model this work, we want to visualize the general trends. This allows us to to flag for large outliers or anomalies within our data.

- With this data we can see how significant Covid shutdowns had on receipts 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.



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 <-
  CleanNewMem %>%
  filter(Date < as.Date("2022-01-01")) %>%
  mutate(ds = Date) %>%
  group_by(ds) %>%
  summarise(y = sum(NewMemCount, 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')])
```

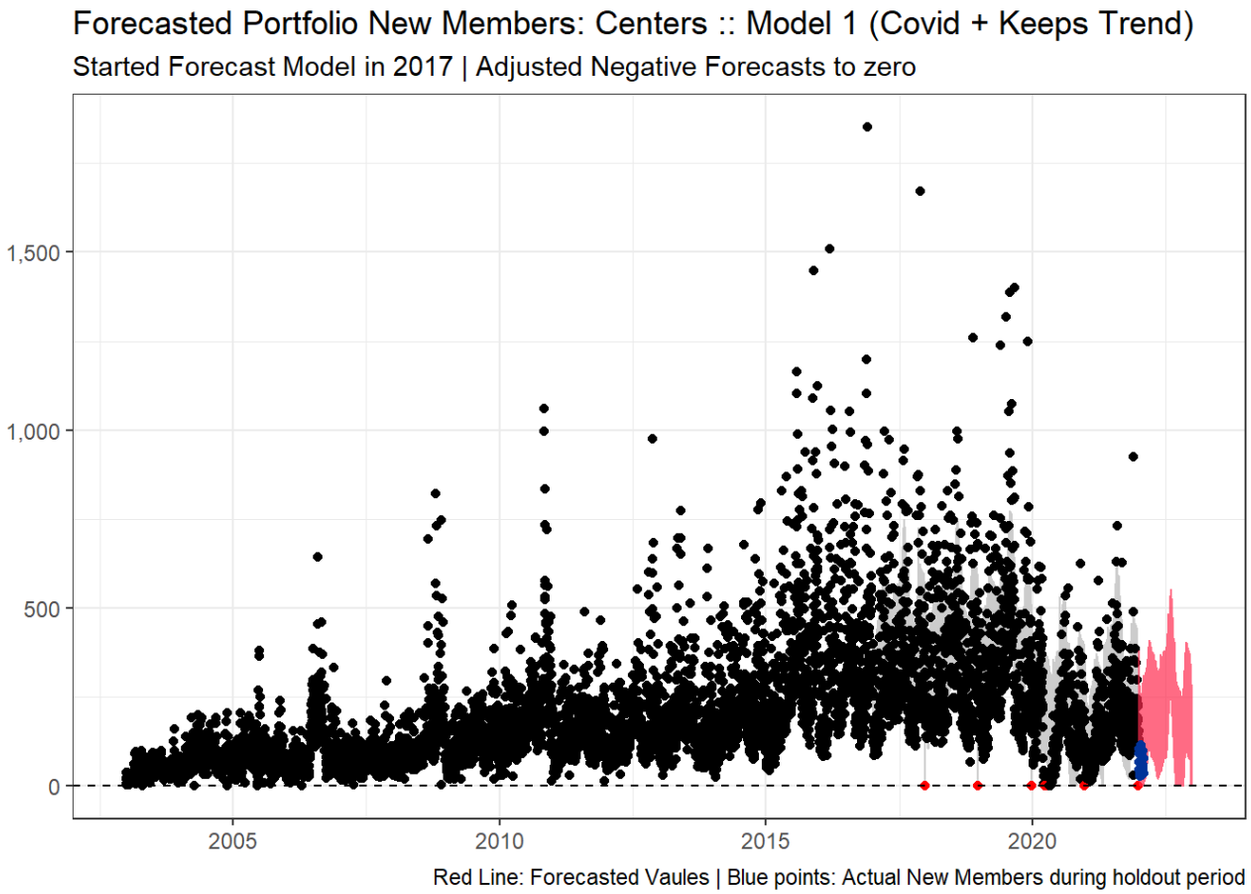
Forecast Results: Centers

Below is the forecast model.

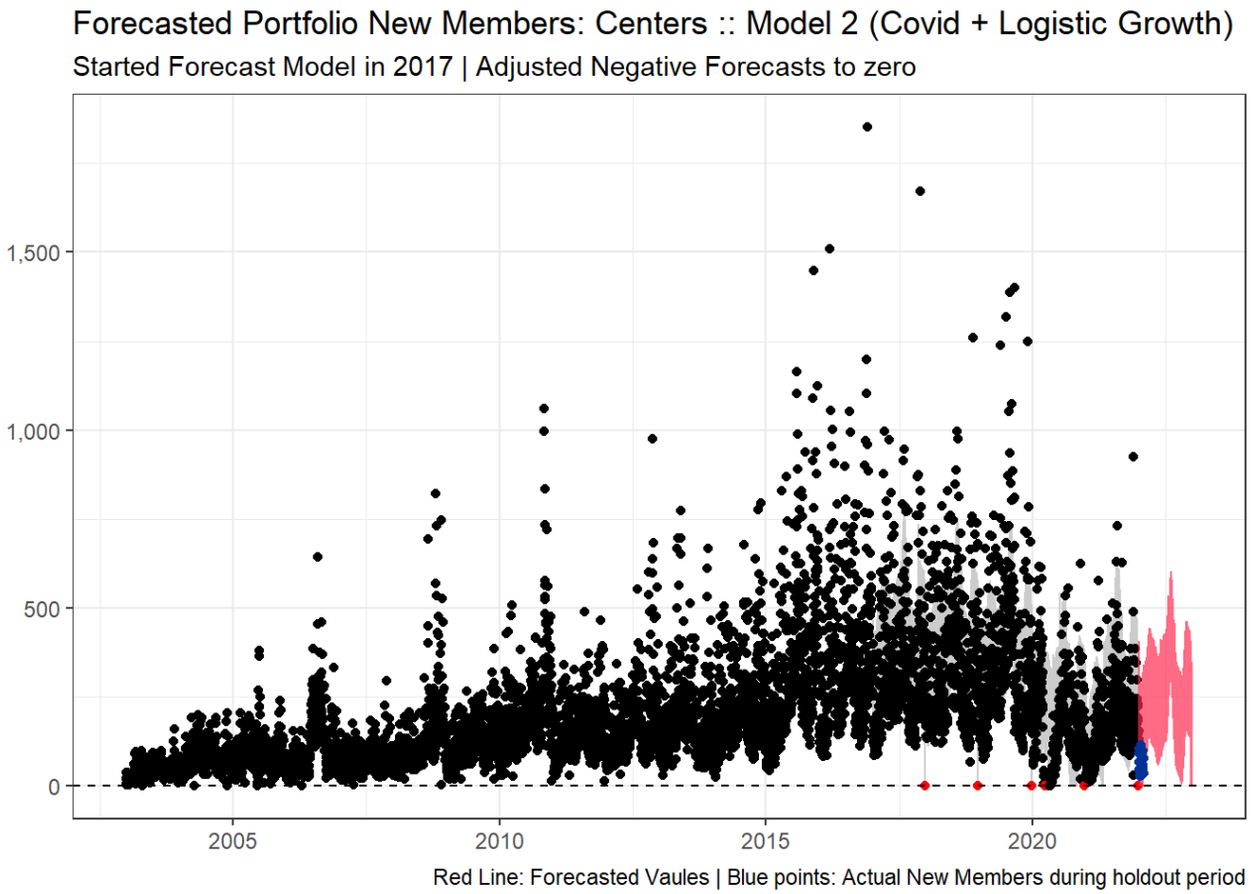
- The black dots represent to the total portfolio traffic by day.
- The blue line represents the forecasted values.

We see a strong degree of seasonality

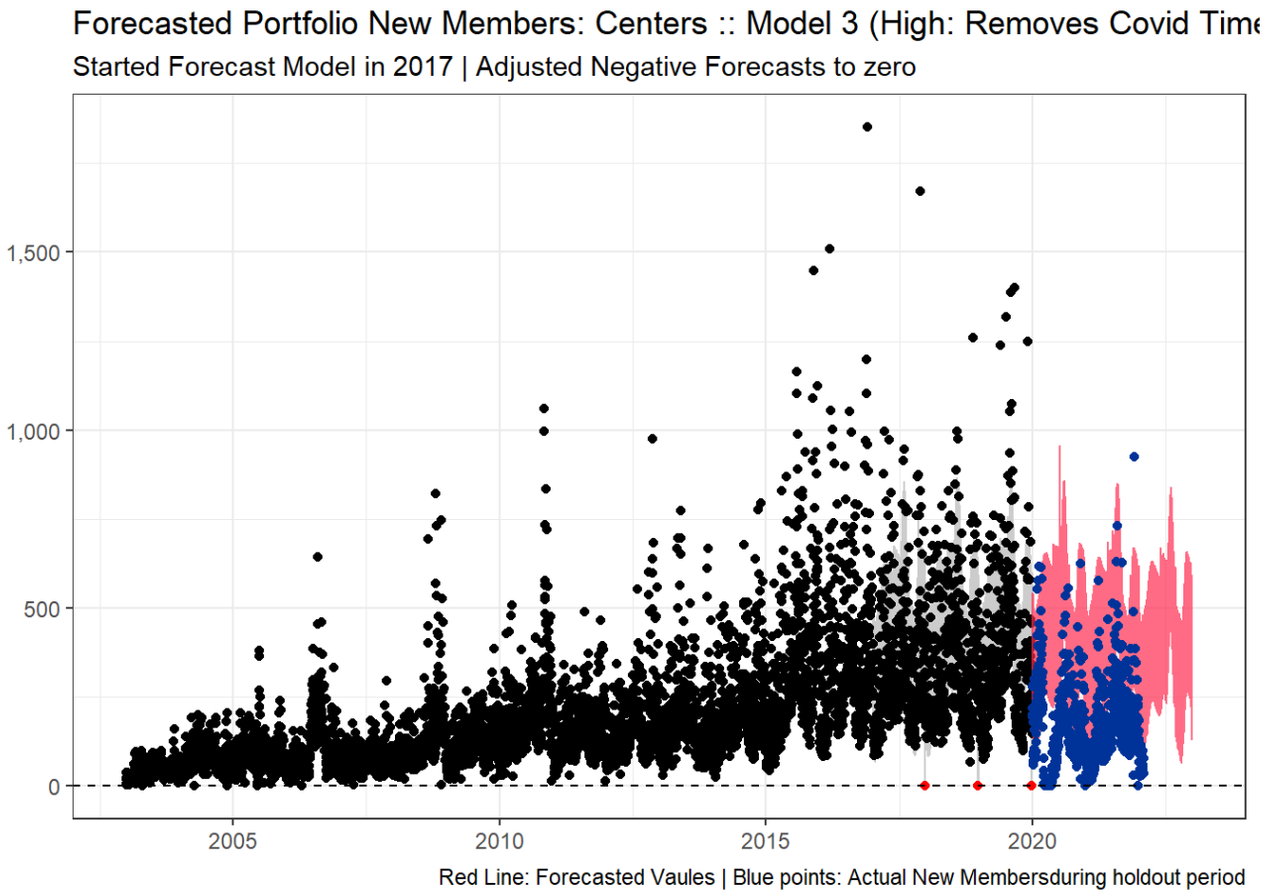
Forecast Model 1



Forecast Model 2



Forecast Model 3



Forecast Components: Centers

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

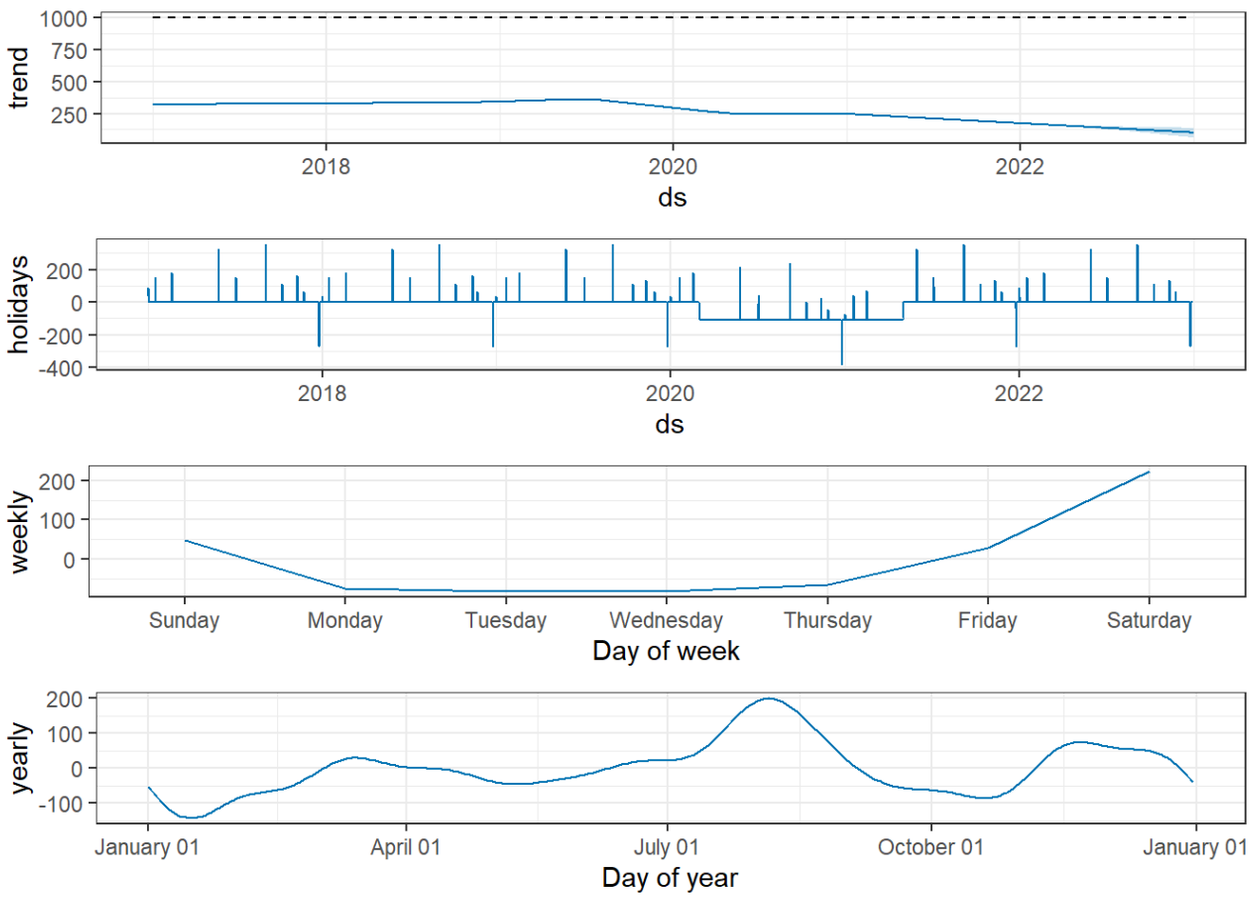
By breaking down the forecast into it's component parts we see a couple of things:

- 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.

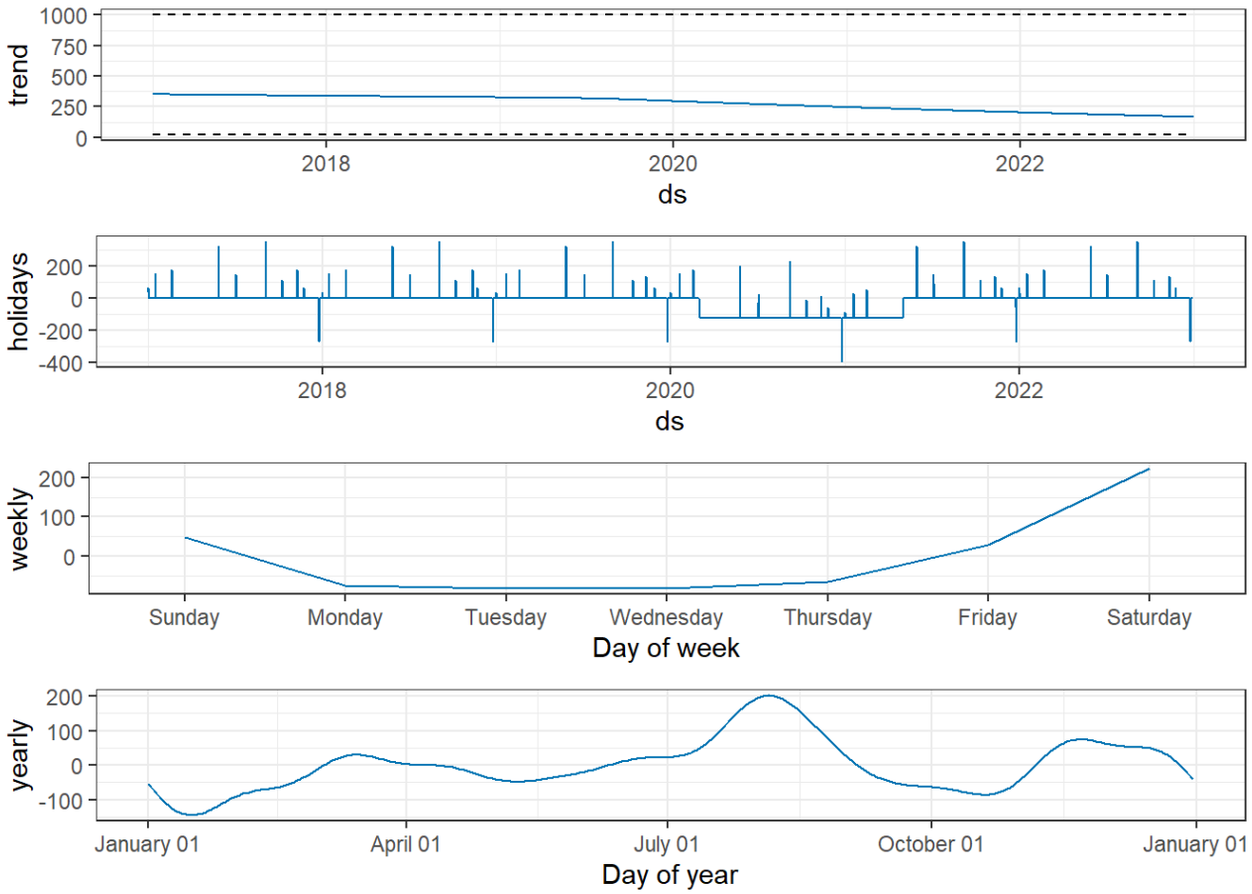
Note about negative values: These 4 items below are added together to create the forecast model. This is why there are some items that are negative and some that are positive. A negative value does not mean that we would forecast negative vehicles, but instead that at those instances the traffic would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

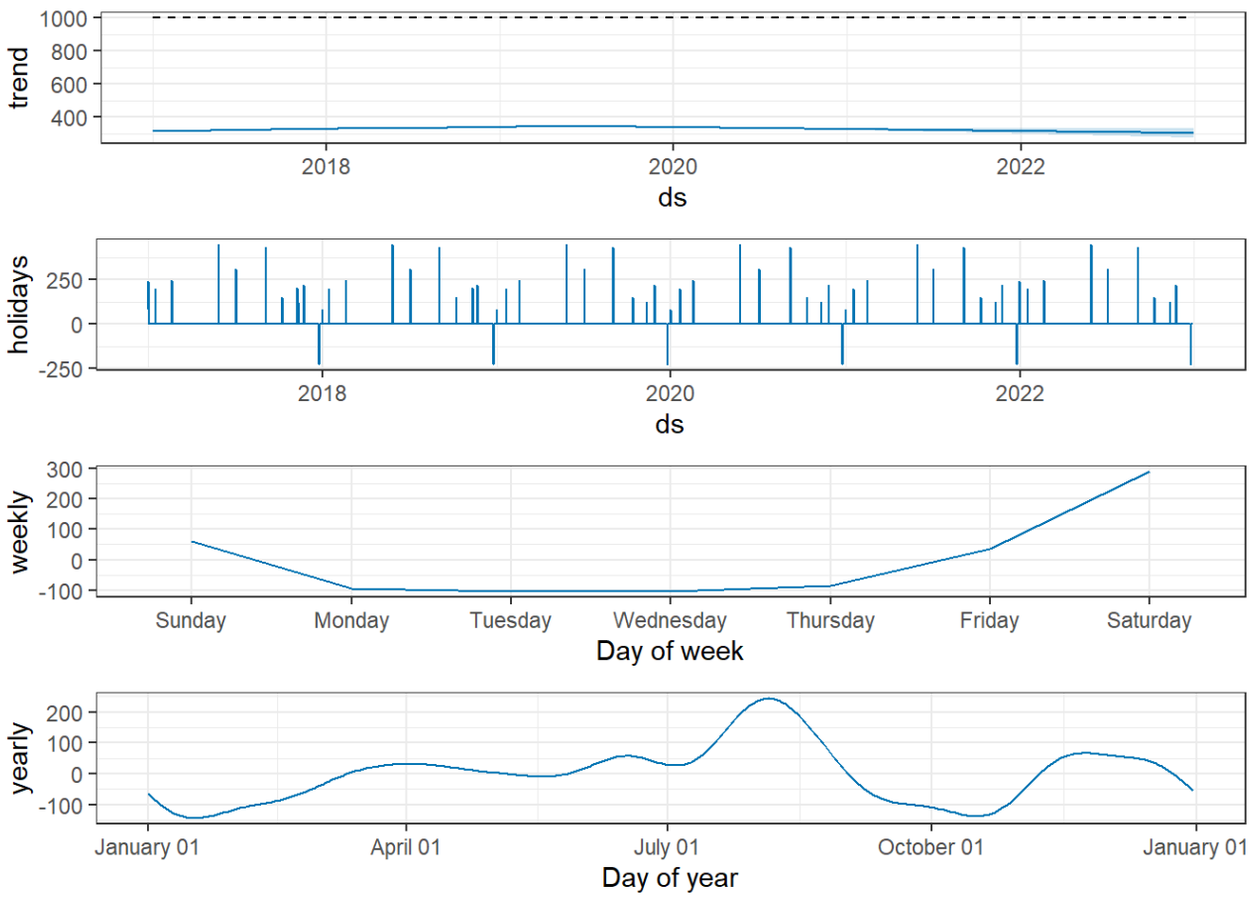
Model 1



Model 2


trend
1000
750
500
250
0
2018
2020
2022
ds
holidays
200
0
-200
-400
2018
2020
2022
ds
weekly
200
100
0
Sunday
Monday
Tuesday
Wednesday
Thursday
Friday
Saturday
Day of week
yearly
200
100
0
-100
January 01
April 01
July 01
October 01
January 01
Day of year

Model 3



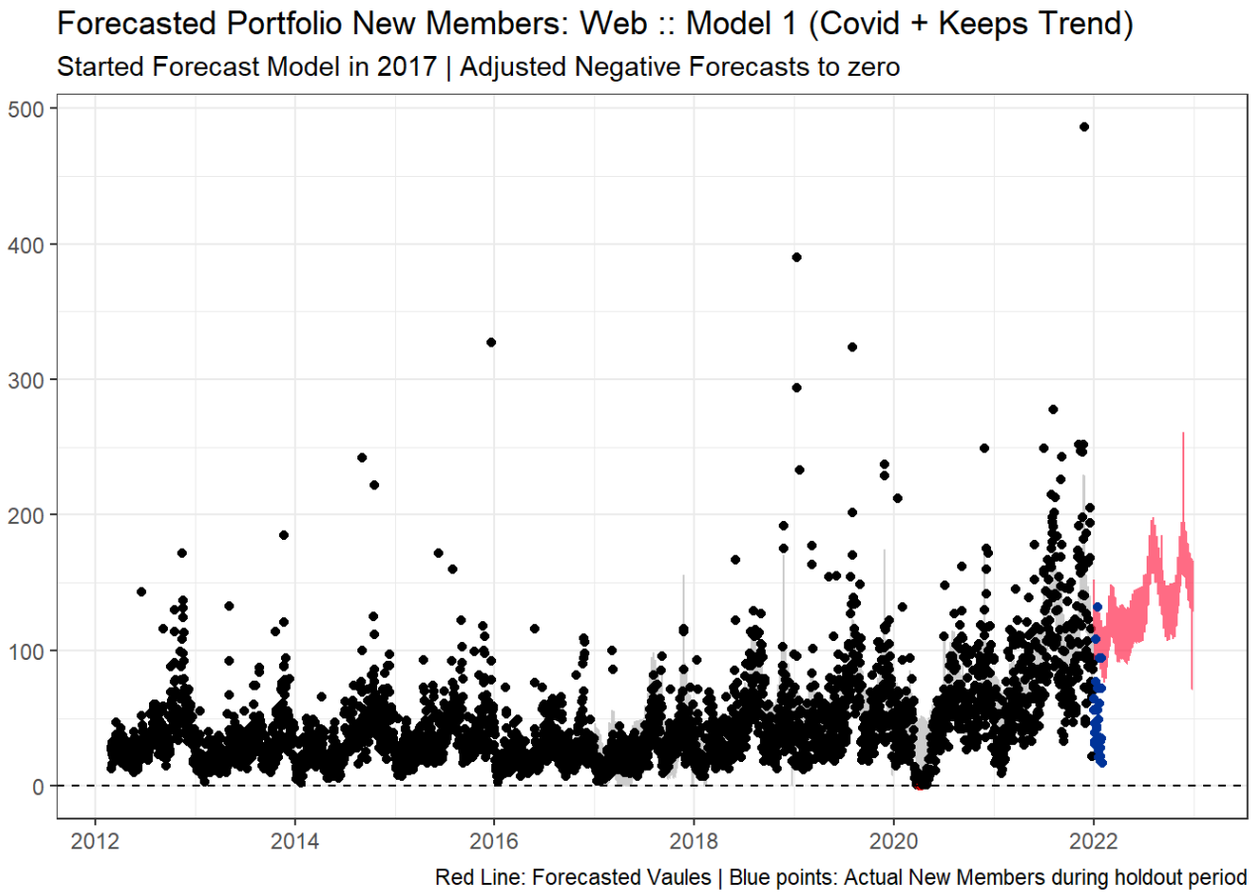
Forecast Results: Web

Below is the forecast model.

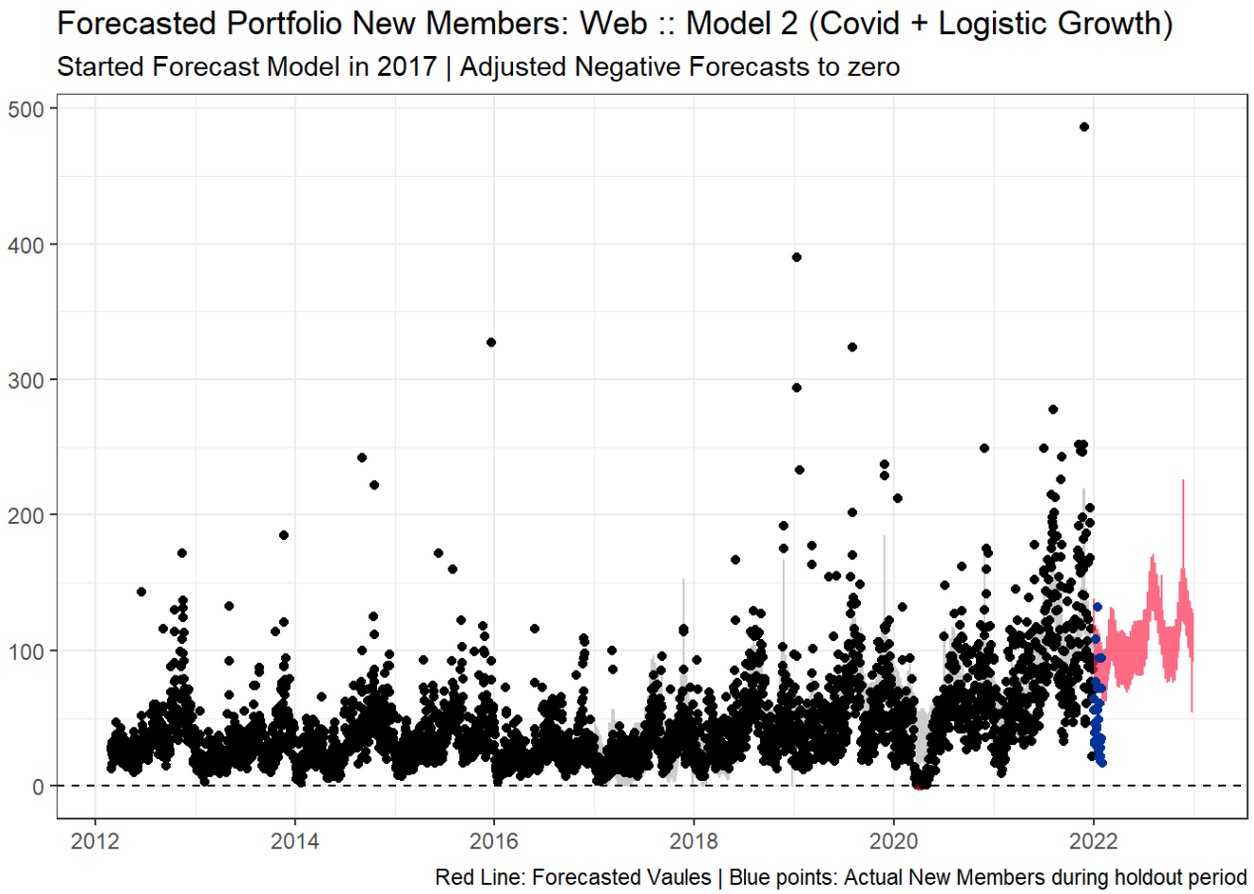
- The black dots represent to the total portfolio traffic by day.
- The blue line represents the forecasted values.

We see a strong degree of seasonality

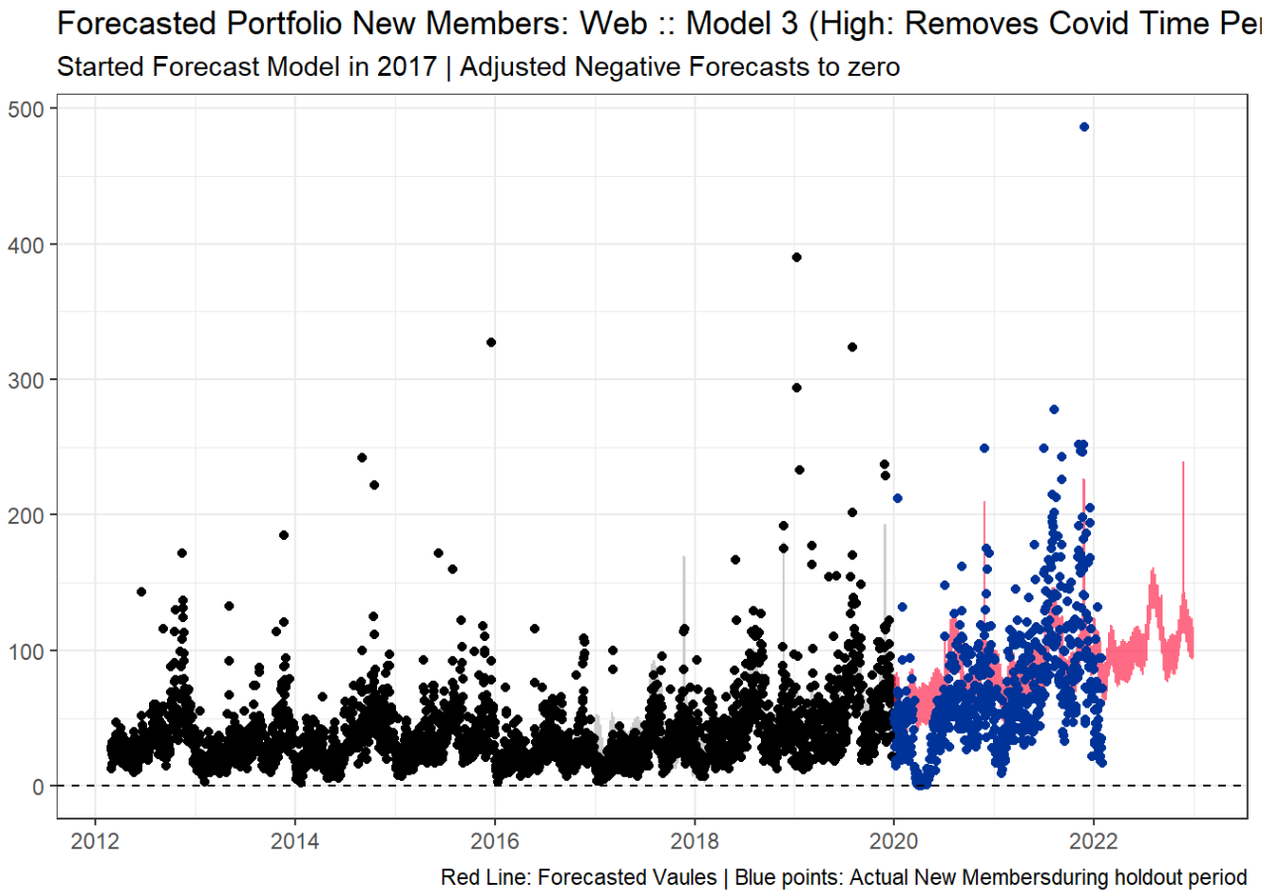
Forecast Model 1



Forecast Model 2

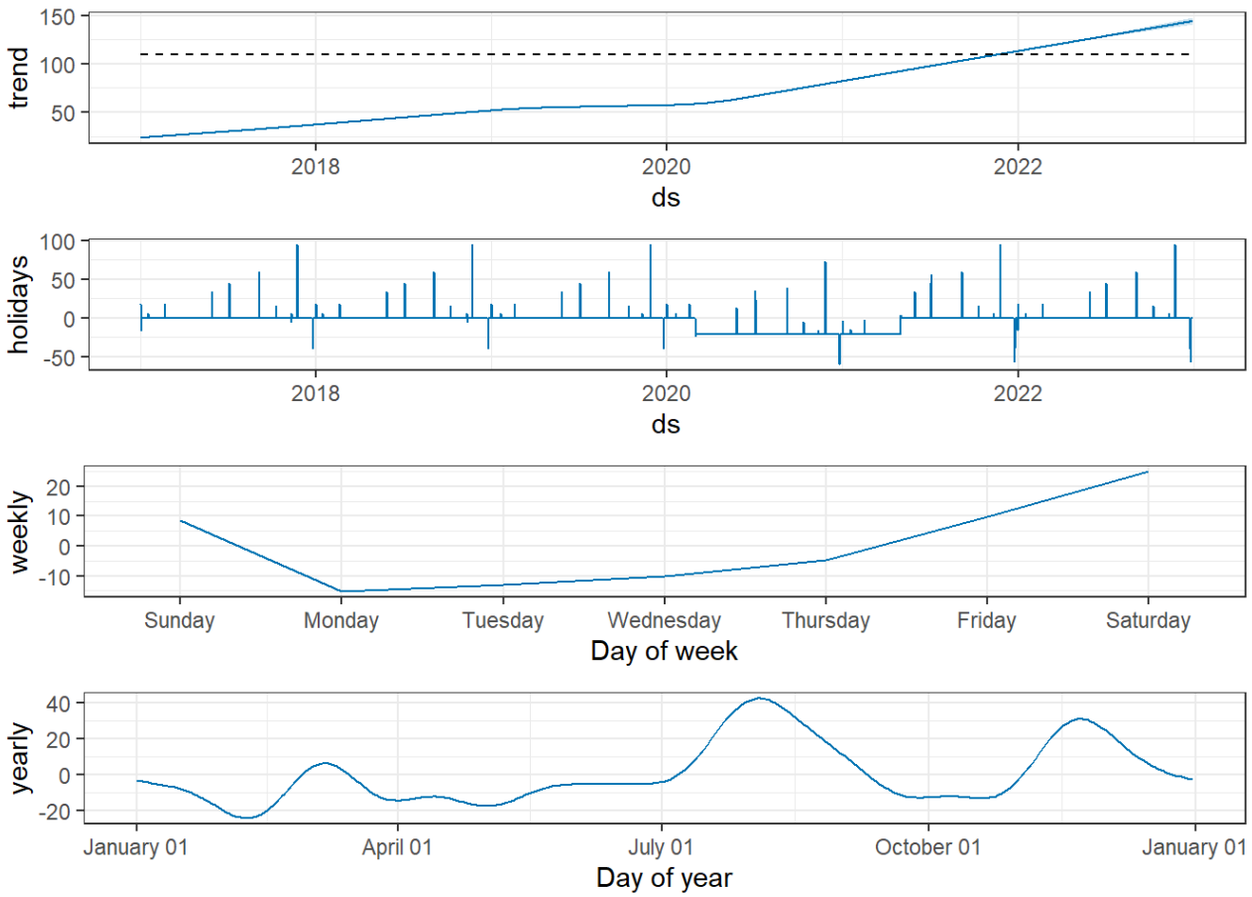


Forecast Model 3

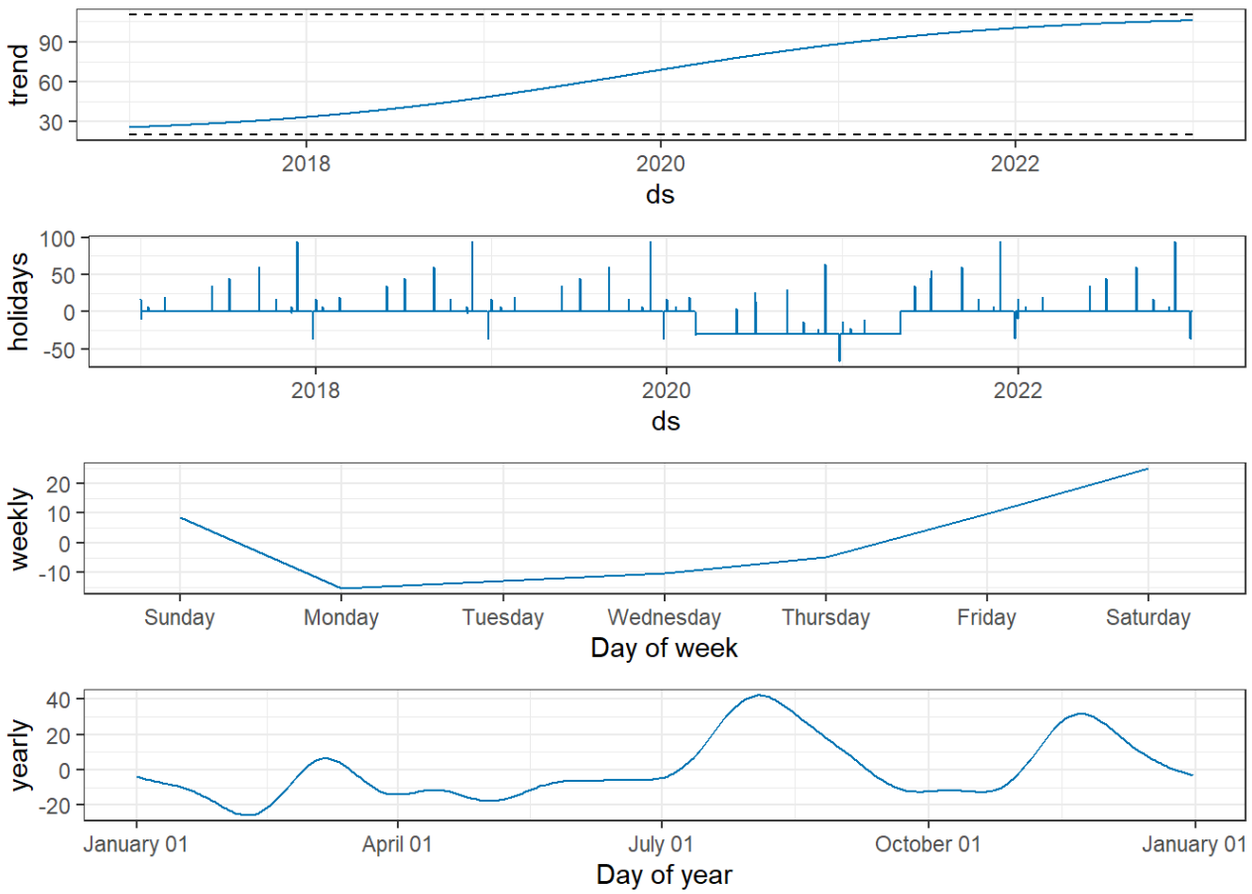


Forecast Components: Web

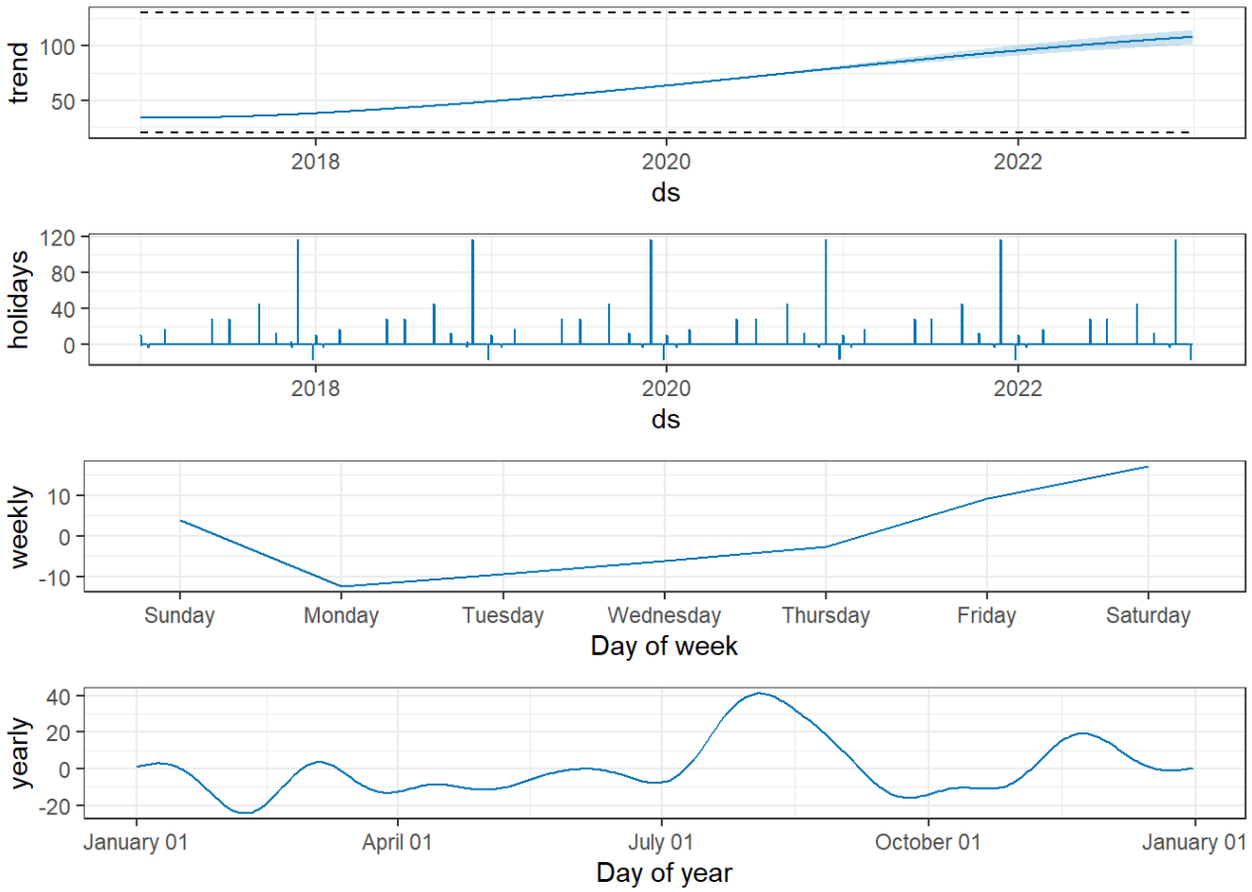
Model 1


trend
150
100
50
2018
2020
2022
ds
holidays
100
50
0
-50
2018
2020
2022
ds
weekly
20
10
0
-10
Sunday
Monday
Tuesday
Wednesday
Thursday
Friday
Saturday
Day of week
yearly
40
20
0
-20
January 01
April 01
July 01
October 01
January 01
Day of year

Model 2


trend
90
60
30
2018
2020
2022
ds
holidays
100
50
0
-50
2018
2020
2022
ds
weekly
20
10
0
-10
Sunday
Monday
Tuesday
Wednesday
Thursday
Friday
Saturday
Day of week
yearly
40
20
0
-20
January 01
April 01
July 01
October 01
January 01
Day of year

Model 3


trend
100
50
2018
2020
2022
ds
holidays
120
80
40
0
2018
2020
2022
ds
weekly
10
0
-10
Sunday
Monday
Tuesday
Wednesday
Thursday
Friday
Saturday
Day of week
yearly
40
20
0
-20
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: New Members

```
With tc_new_members (TC_CustID, CreatedBy, CenterID,CenterName, FromWeb,
        TransactionDetailID, TransactionID, ProductID, ForCenterID,
Quantity, Price, Status, CreatedOn, ModifiedOn, TransactionDate) as
(
    Select t.CustomerID, t.CreatedBy, t.CenterID,c.CenterName, t.FromWeb,
        td.TransactionDetailID, td.TransactionID, td.ProductID,
     td.ForCenterID, td.Quantity, td.Price, td.Status, cast(td.CreatedOn as
        date), td.ModifiedOn, cast(t.TransactionDate as date)
    From tblTransactionDetail td
    Inner Join tblTransactions t
        On td.TransactionID = t.TransactionID
    left join tblcenters c ON t.centerid = c.centerid
    Where td.ProductID = 1  --New TC Membership
        And td.DeletedOn is null
        And td.Status in ('A', 'N')
        And t.Status in ('A', 'N')
)

select *
from tc_new_members
```

Forecast Model

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

TangerOutlets