INVENTORY REBALANCING AND DEMAND PLANNING

PRESENTED BY - TEAM 12

OUR TEAM

GARIMA CHOUHAN **GISHA GOPAL** **NIKET DALAL**

AGENDA

- Introduction
- Key Problems and challenges facing the company
- Data collection and preparation
- Full dataset vs. subset of data and Model outputs
- Comparing Accuracy measures
- How did our Analysis and findings help Management

INTRODUCTION

- Largest bike sharing system in the US.
- Named after lead sponsor Citigroup.
- Launched on 27th May 2013
- Operates in Bronx, Brooklyn, Manhattan, Queens.
- Also operates in JC and Hoboken in NJ (~80 stations)
- Acquired by Lyft in Q4 2018.

FACTS AND FIGURES

- In 2013, started with 332 Stations and an active bike inventory count of approximately 5000 bikes.
- Present day this has grown to 1500+ stations in NYC and an additional 80+ active stations in NJ with an active bike count of approximately 25000+ bikes.
- Ridership has grown from 5.8 million in 2013 to over 26 million in 2021.
- Annual membership has risen from 52k members to 160k in 2021.
- Revenue has grown from $14.3 million in 2013 to $80 million as of Oct 2021.

KEY PROBLEMS FACING THE COMPANY

- Inventory and network rebalancing and Demand Planning to overcome empty docks.
- Forecasting for accurate demand is extremely challenging as it requires us to consider the effects of intangibles (like weather, traffic pattern, Rider preferences, commuting patterns etc.) on the number of Trips and its duration.
- Volatility due to on ground conditions (constructions, Road closures, Pandemic etc.)
- Induced Demand in transportation continues to be a challenge.

WHAT ARE WE DOING AND HOW DOES IT MATTER?

- Our team will analyze Historic Trip Data for one specific station and run different forecasting models on it.
- We will then finalize a Forecasting model based on its accuracy and predict Trip counts for the next six months (2021)
- Accurate Forecasts will help Management to accurately plan and predict for future Demand at that station.
- It also helps Management make key decisions on factors like Inventory management, Inventory Rebalancing and Daily Operations.
- This helps the company mitigate Induced Demand and indirectly help with memberships and revenue increase.

WHAT DATA WE USED

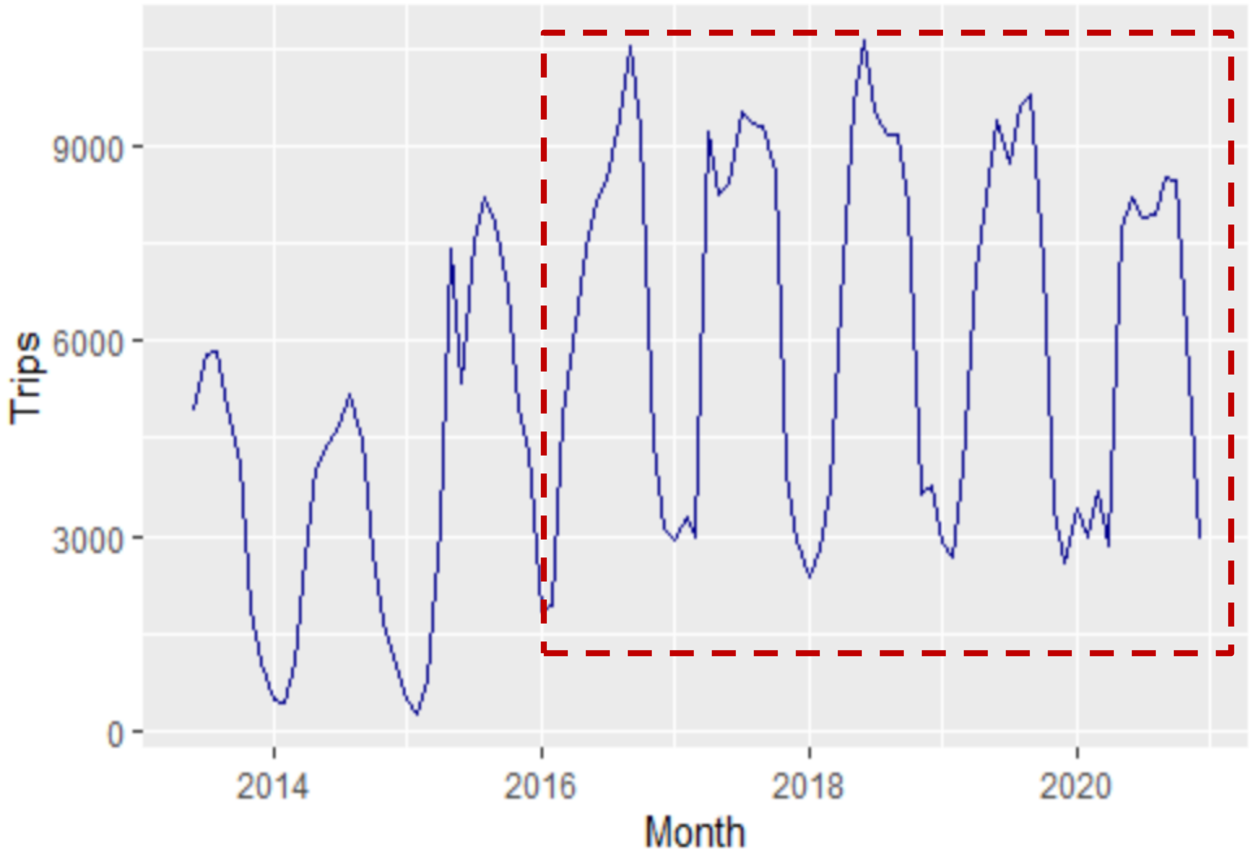
Data Source

- Citi Bike Trip data is published on their website every month
- Dataset is open for public use and is free to download.
- Files are segregated between NYC and NJ stations and broken down monthly.
- The Dataset grain is on a (single) Trip level.
- You can find further details here https://ride.citibikenyc.com/system-data

Data Wrangling and Preparation

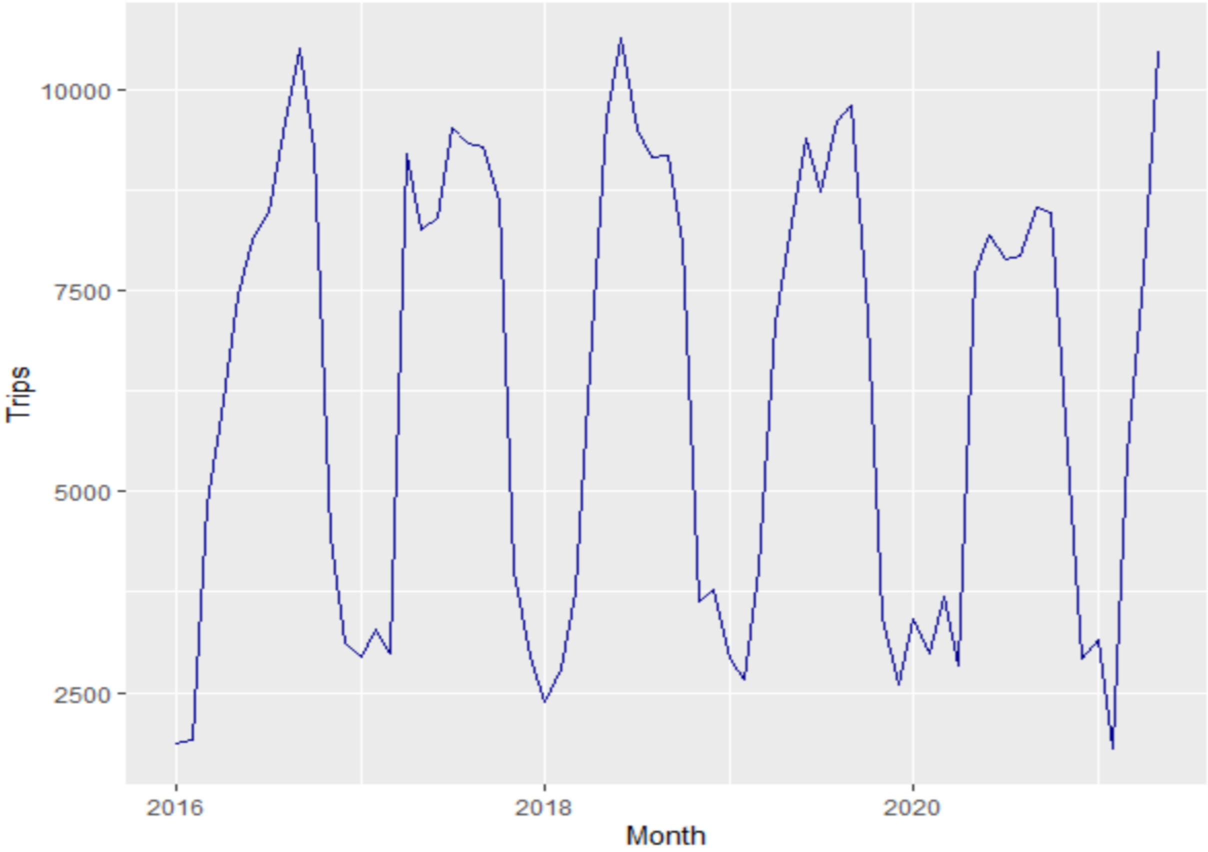
- In order to get past the limitations of the tools we are using we decided to focus on Trip data from one station.
- Station of choice [Central Park S. & 6th Ave]
- Download Monthly Trip Data and filter for Station ID 2006.
- Aggregate Trips originating from station ID = 2006 to Monthly counts
- Repeat this for all the months up to May 2021.
- Create .CSV file for consumption into R.

TRIP COUNT FROM START OF TIME



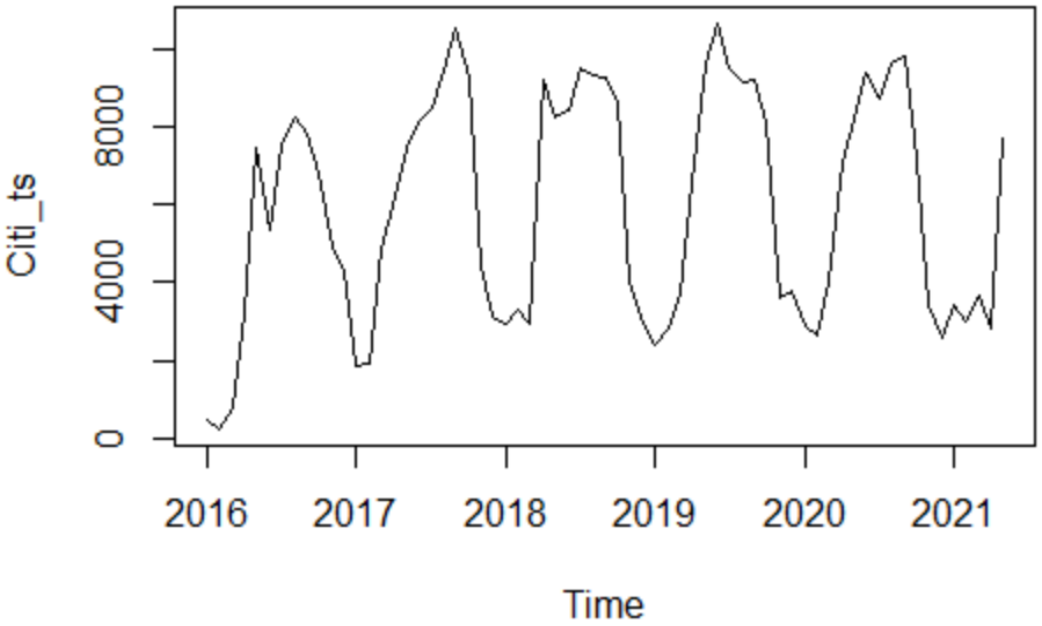
- Alongside graph represents Trip Counts from station ID = ‘2006’ which maps to location ‘Central Park S. & 6th Ave.’
- Monthly Trip Count data captured from 1st full month of operation [Jun-2013]
- As can be seen Trip Counts stabilize from start of 2016.

TRIP COUNT [2016] ONWARDS



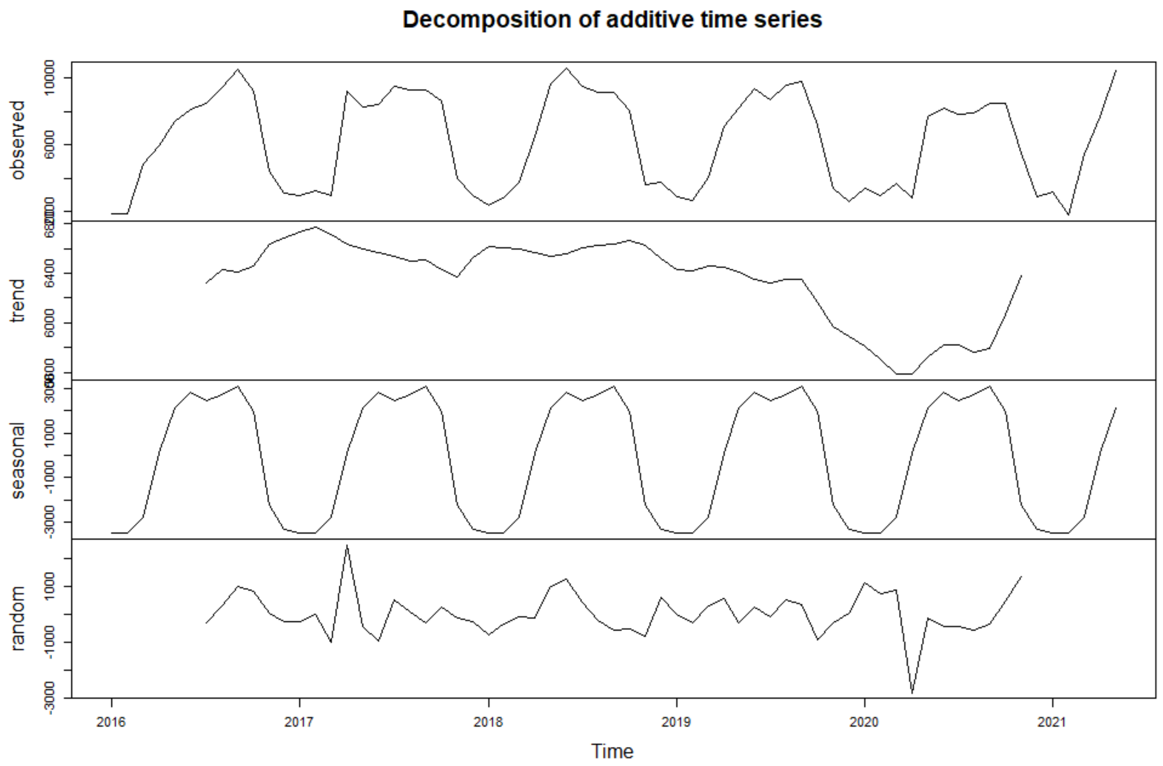
- Alongside graph represents Trip Counts from station ID = ‘2006’ which maps to location ‘Central Park S. & 6th Ave.’
- Excluded partial 2013 data along with 2014 and 2015 data to avoid any irregularities in forecasting.
- We will use this dataset to create our Time-Series.

TIME SERIES



- We plot the Time Series using data from 2016 until mid 2021.
- Our time series shows the following characteristics.
  - Trend: Slight overall year-over-year growth.
  - Seasonality: Users make more trips in the summer months and less in the winter.

DECOMPOSITION OF TIME SERIES

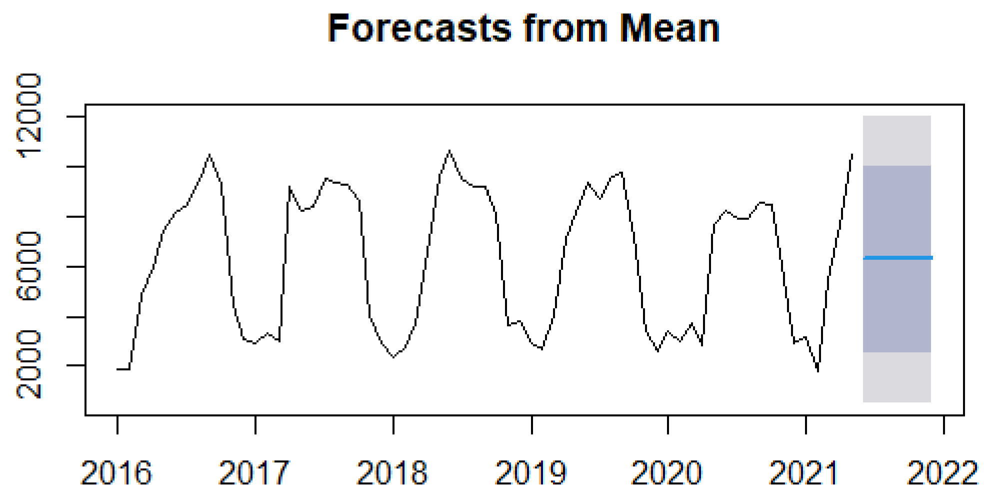


- We plot the Decomposition of our Time Series.
- Decomposition breakdown coincides with our observations from the previous slide.
- Slight overall year-over-year growth. Drop in around mid 2020 probably due to COVID but picked up quickly after mid 2020.
- Clear seasonal pattern detected.

MODELS

- Mean
- Naive
- Seasonal Naive
- Random walk
- Simple Moving Averages
- Exponential Smoothing
- Holt Winters
- Arima

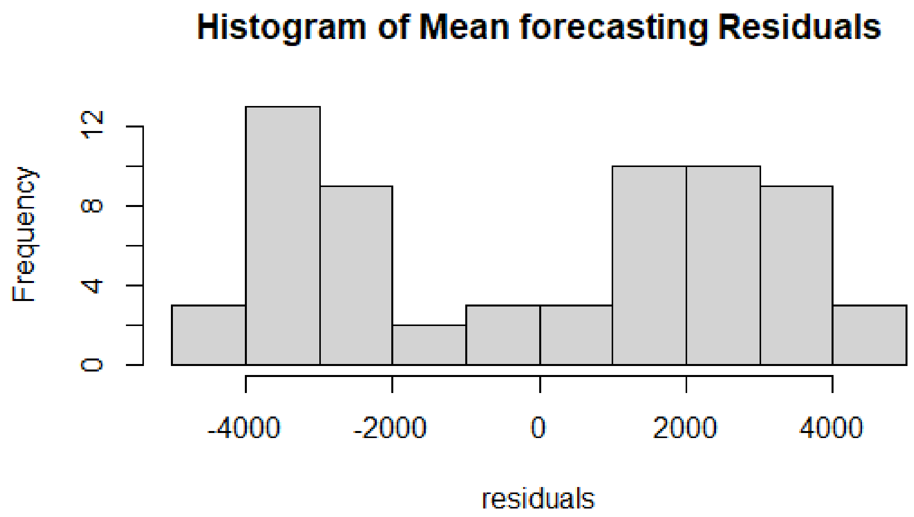
MEAN FORECASTING

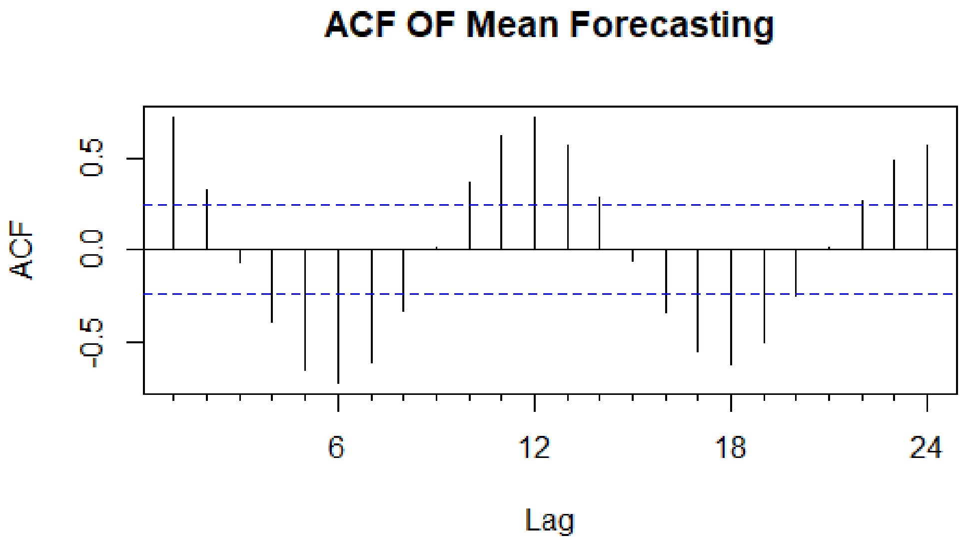


- ACCURACY VALUES

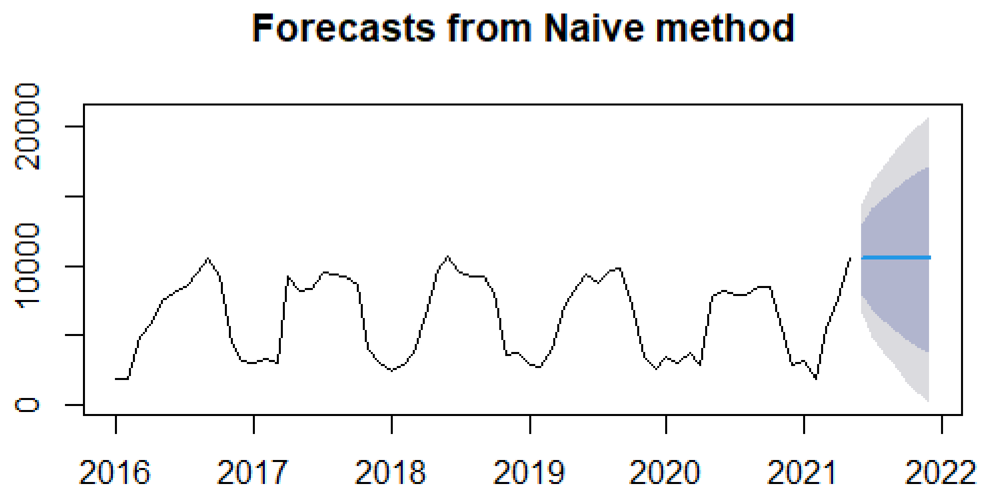
ME	RMSE	MAE	MPE	MAPE	MASE	ACF1
-2.794e-14	2831.015	2643.898	-31.73938	61.05208	2.498513	0.72066

MEAN FORECASTING





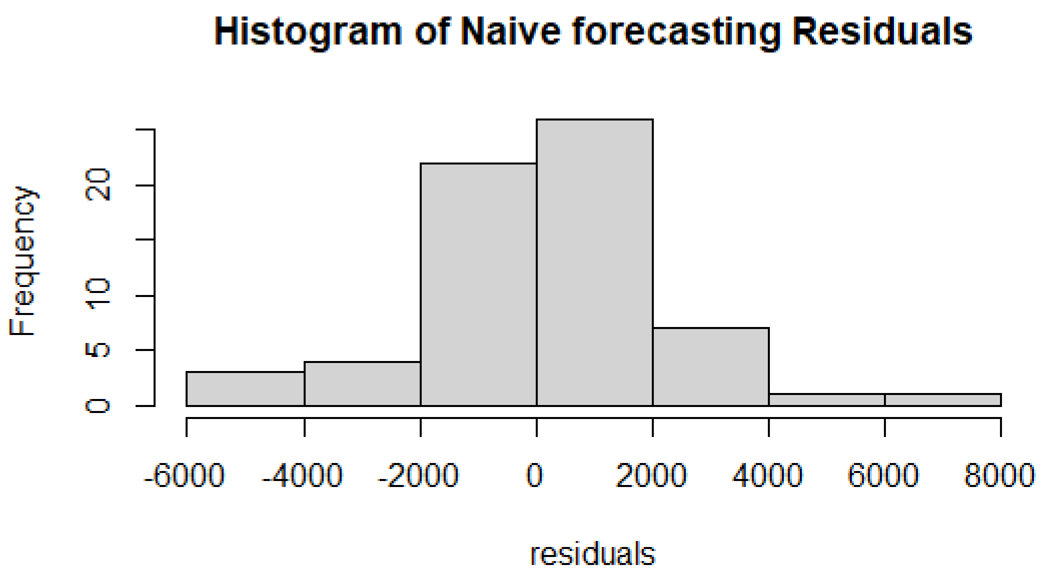
NAIVE FORECASTING

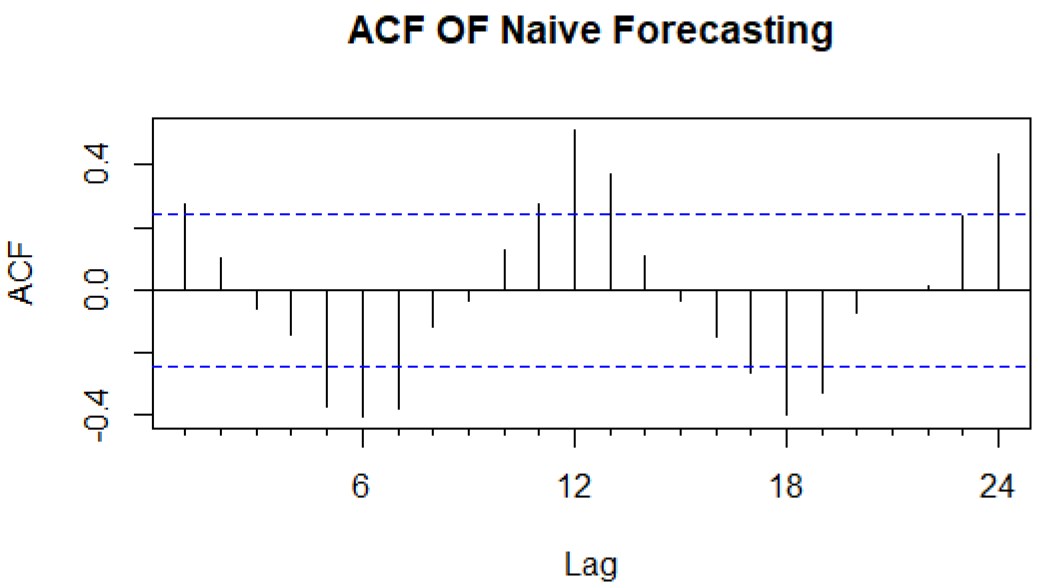


- ACCURACY VALUES

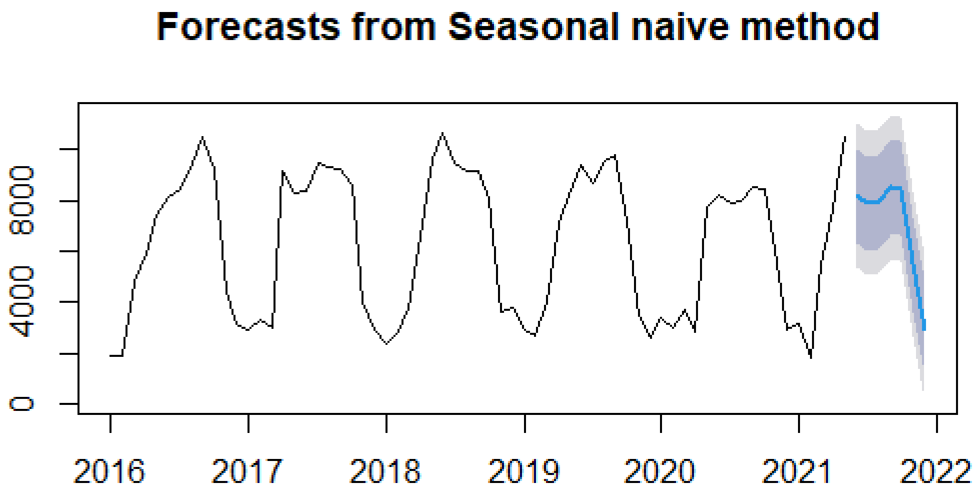
ME	RMSE	MAE	MPE	MAPE	MASE	ACF1
134.2656	1992.589	1399.453	-4.691763	27.14672	1.322499	0.2720531

NAIVE FORECASTING





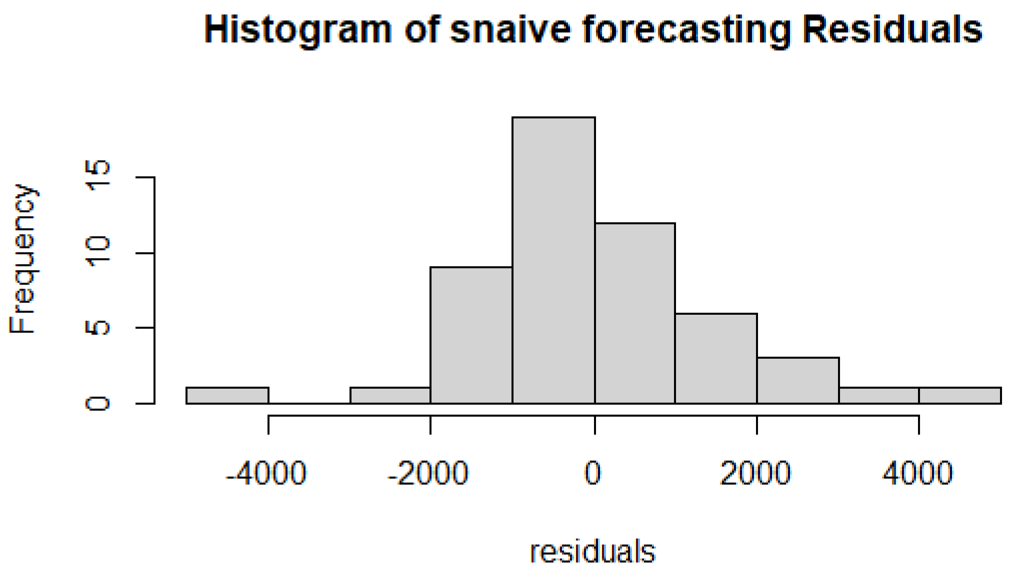
SEASONAL NAIVE FORECASTING

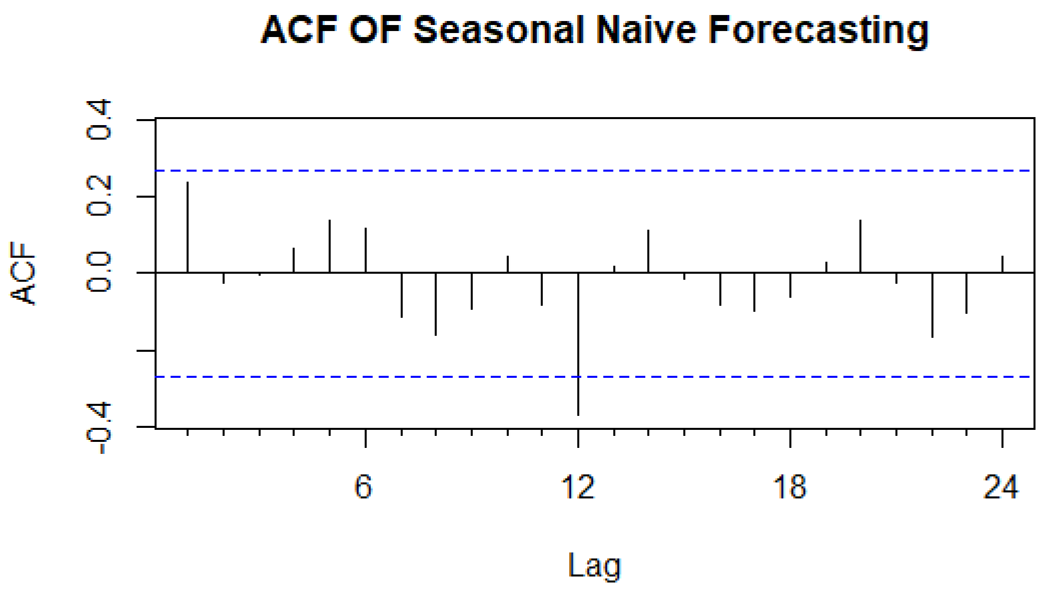


- ACCURACY VALUES

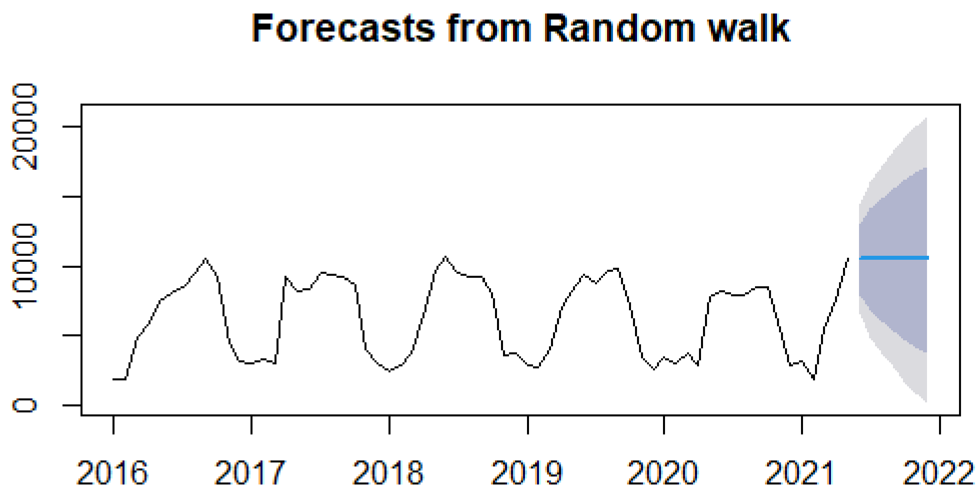
ME	RMSE	MAE	MPE	MAPE	MASE	ACF1
49.58491	1452.853	1058.189	-2.659102	20.42757	1	0.2370095

SEASONAL NAIVE FORECASTING





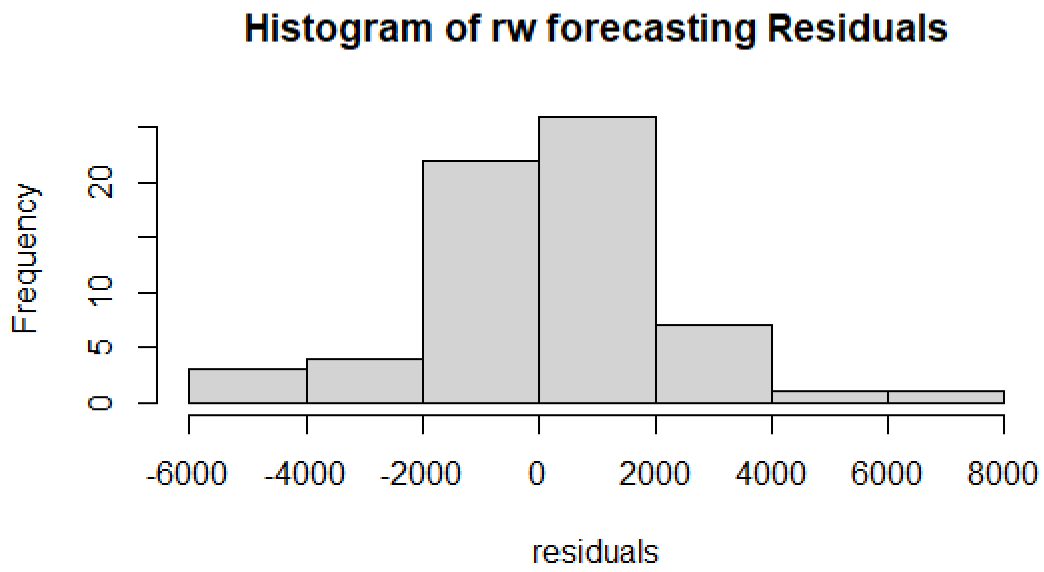
RANDOM WALK FORECASTING

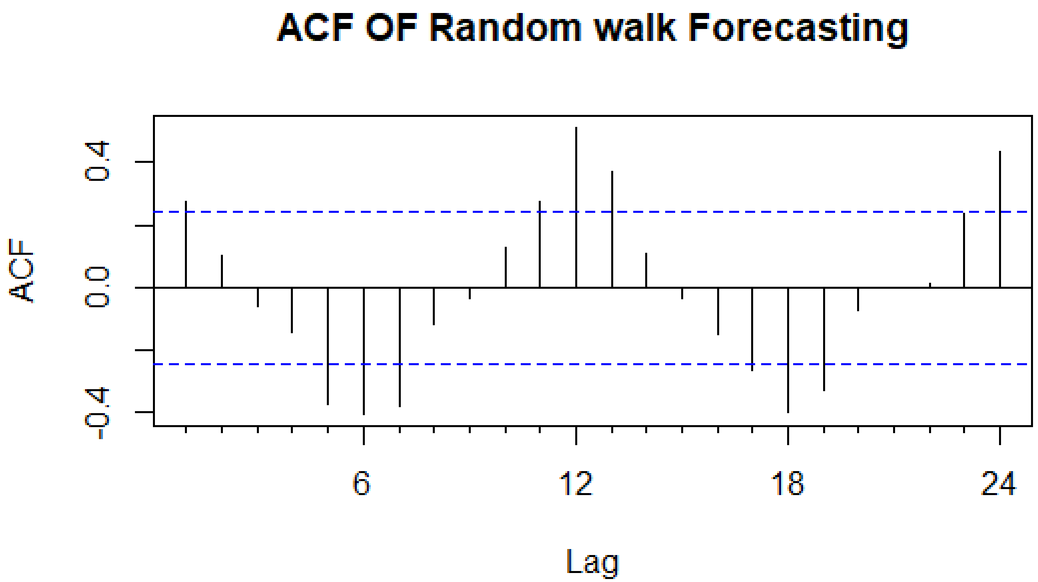


- ACCURACY VALUES

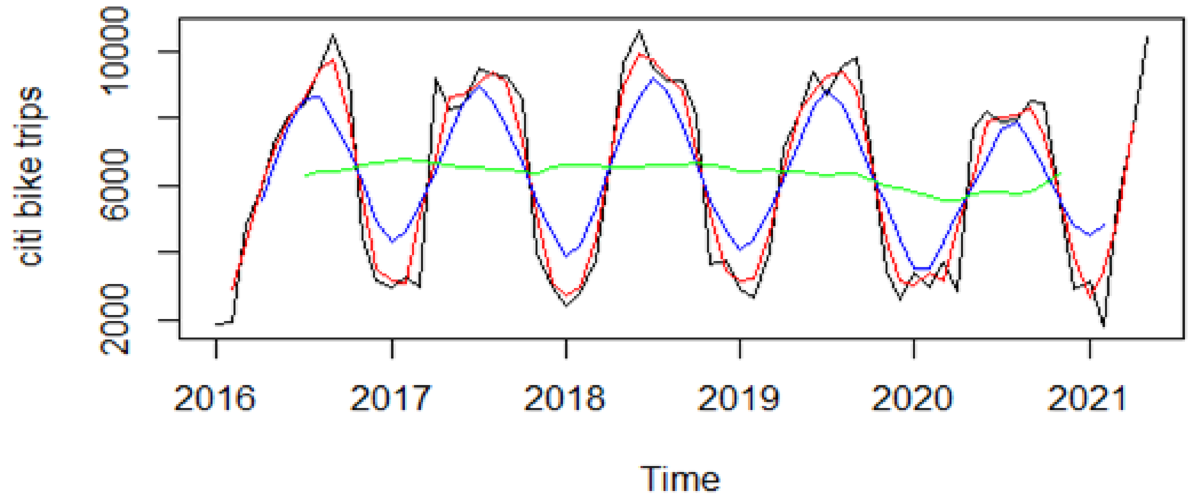
ME	RMSE	MAE	MPE	MAPE	MASE	ACF1
134.2656	1992.589	1399.453	-4.691763	27.14672	1.322499	0.2720531

RANDOM WALK FORECASTING



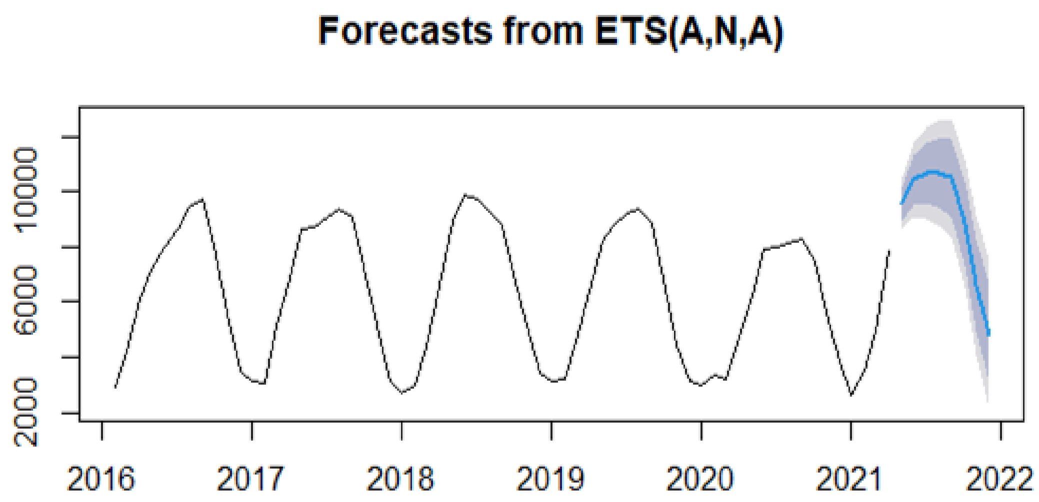


HISTORICAL SIMPLE MOVING AVERAGES



Red 3 order **Blue** 6 order **Green** 12 order

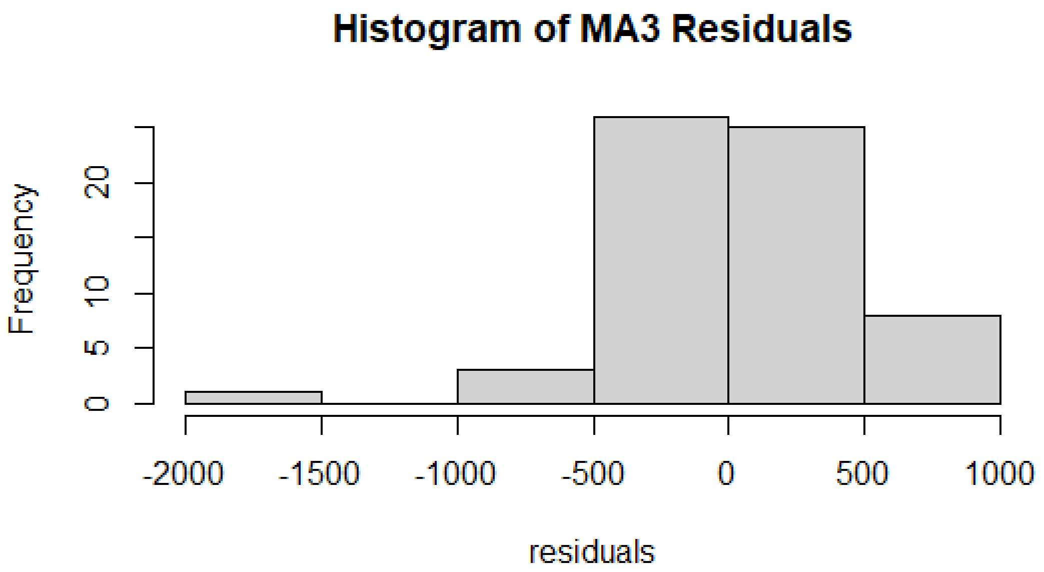
SIMPLE MOVING AVERAGES (ORDER 3)

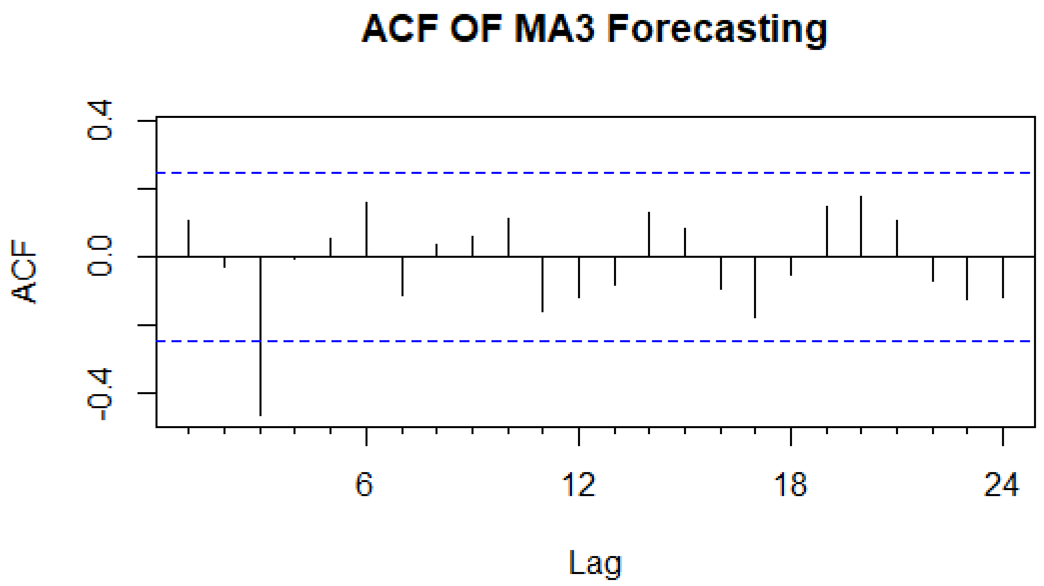


- ACCURACY VALUES

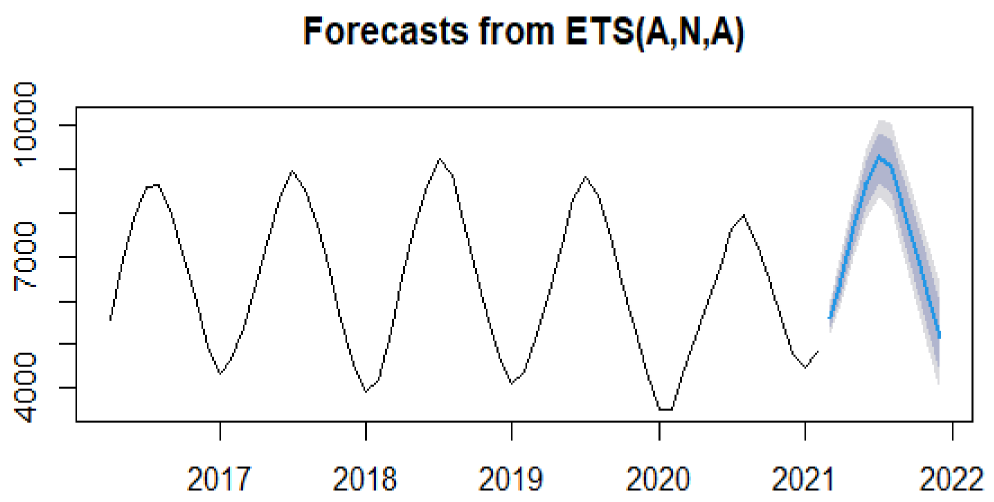
ME	RMSE	MAE	MPE	MAPE	MASE	ACF1
3.67512	429.1339	318.805	-0.214658	5.981897	0.484689	0.1088142

SIMPLE MOVING AVERAGES (ORDER 3)





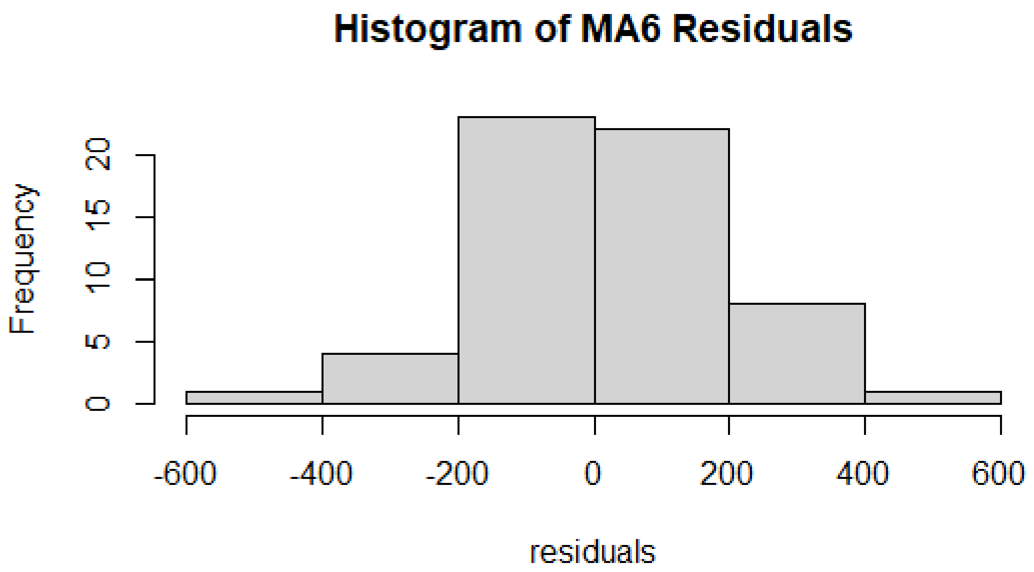
SIMPLE MOVING AVERAGES (ORDER 6)

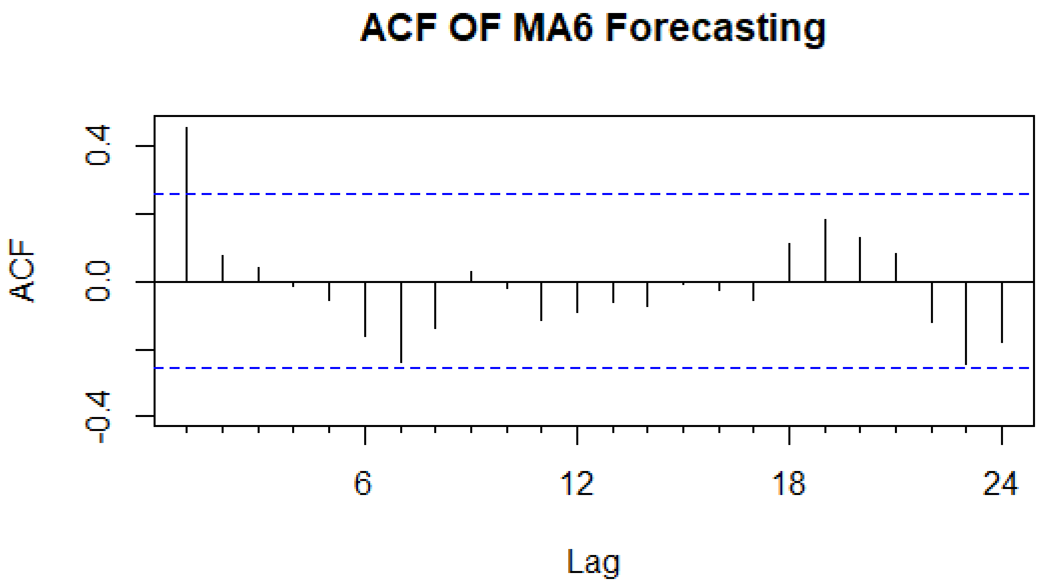


- ACCURACY VALUES

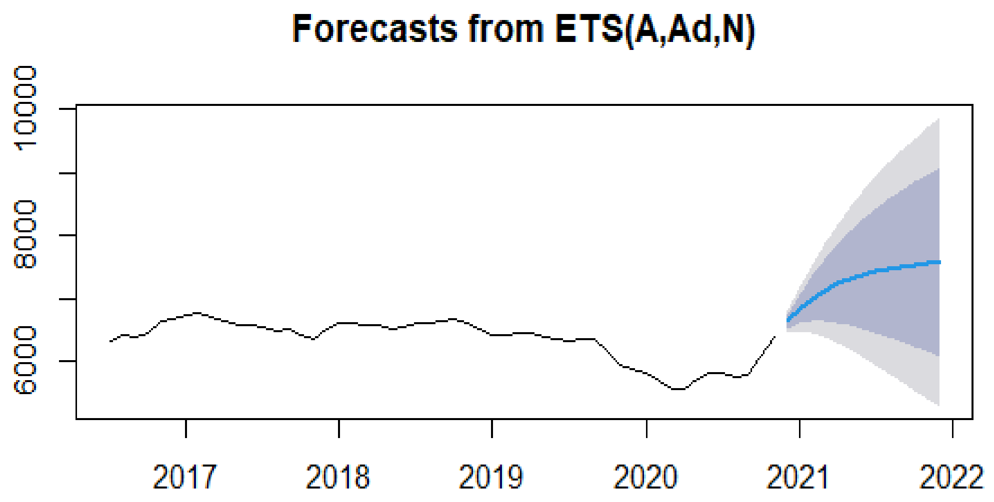
ME	RMSE	MAE	MPE	MAPE	MASE	ACF1
7.185922	177.134	134.4685	-0.0677174	2.219086	0.3065491	0.4510698

SIMPLE MOVING AVERAGES (ORDER 6)





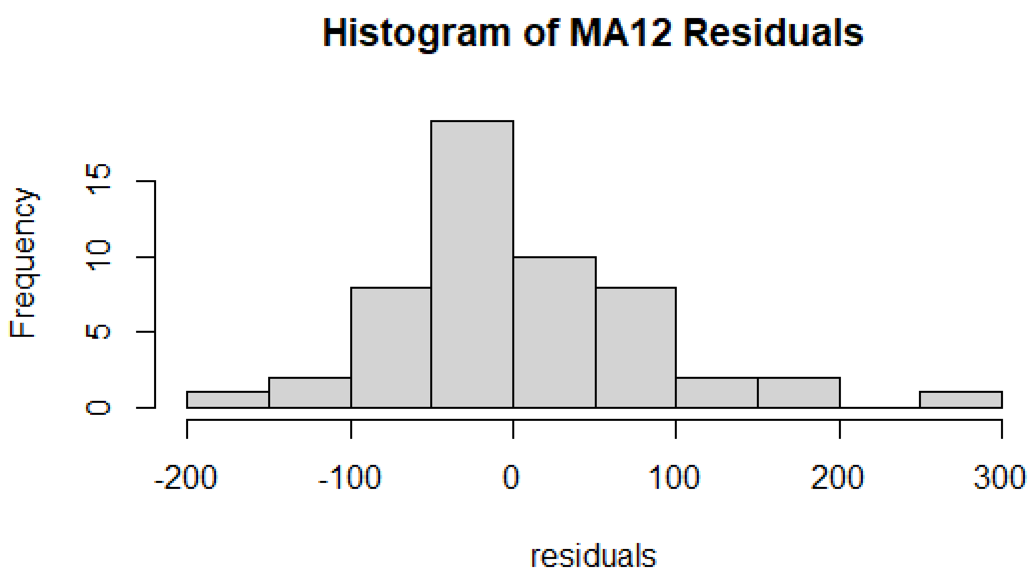
SIMPLE MOVING AVERAGES (ORDER 12)

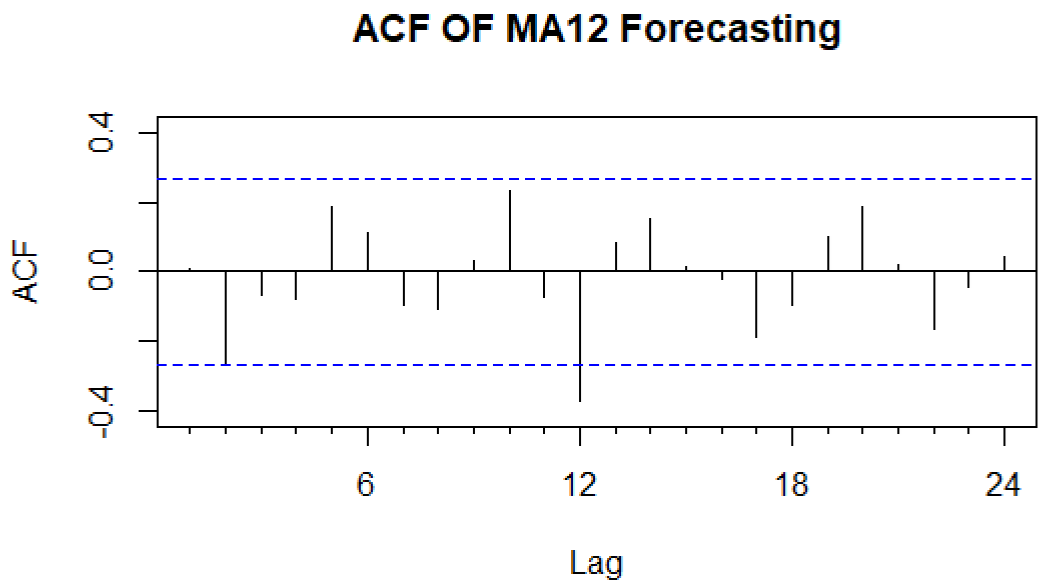


- ACCURACY VALUES

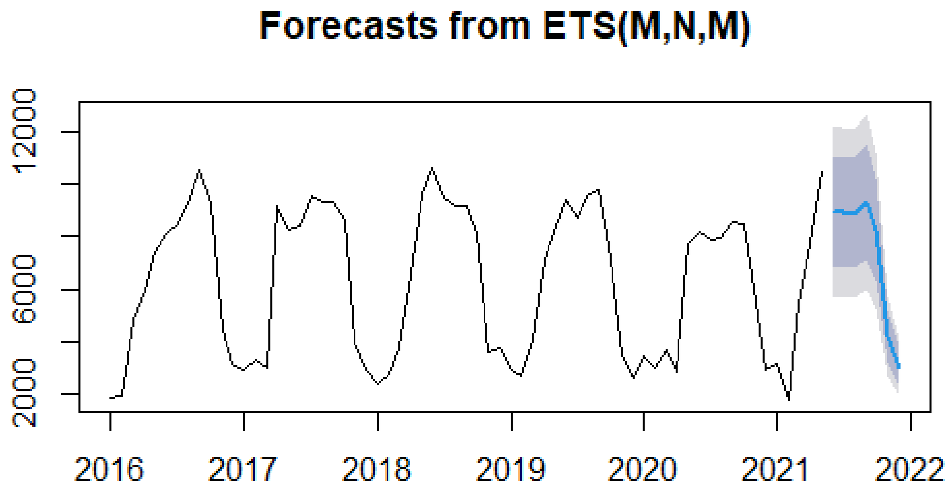
ME	RMSE	MAE	MPE	MAPE	MASE	ACF1
3.895965	82.73701	61.92499	-0.027109	0.9868236	0.2076915	0.0107897

SIMPLE MOVING AVERAGES(ORDER 12)





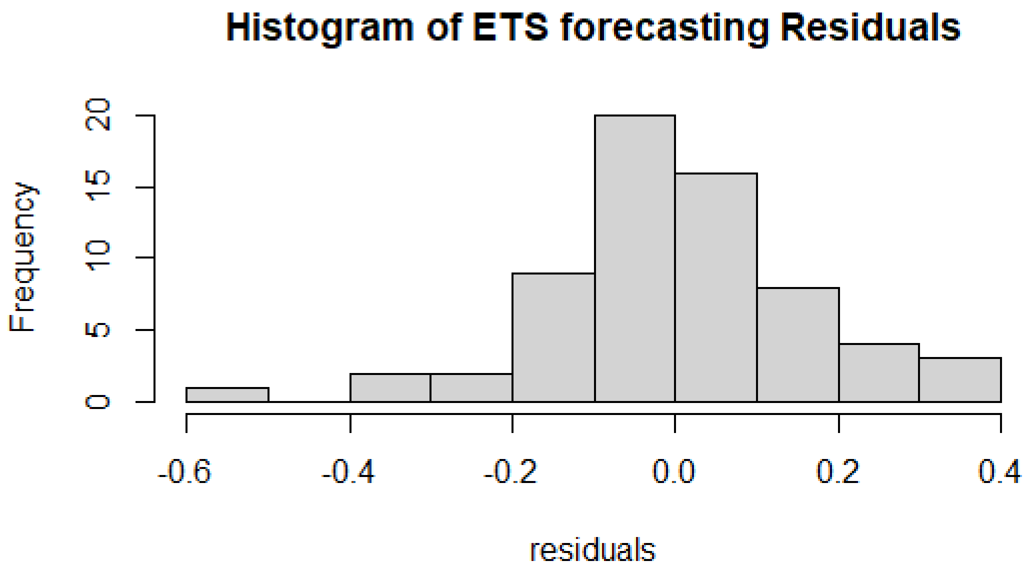
EXPONENTIAL SMOOTHING FORECASTING

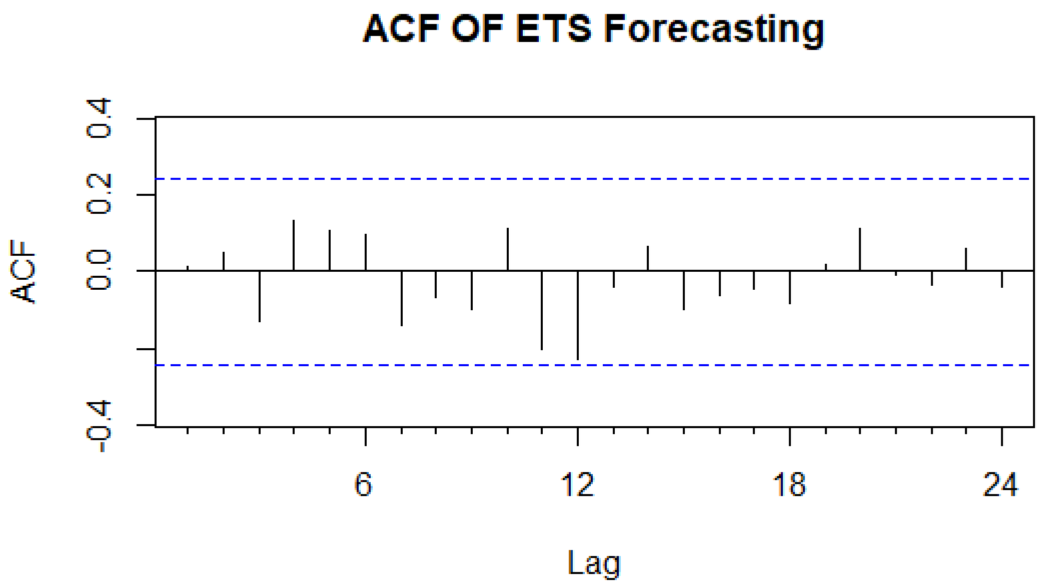


- ACCURACY VALUES

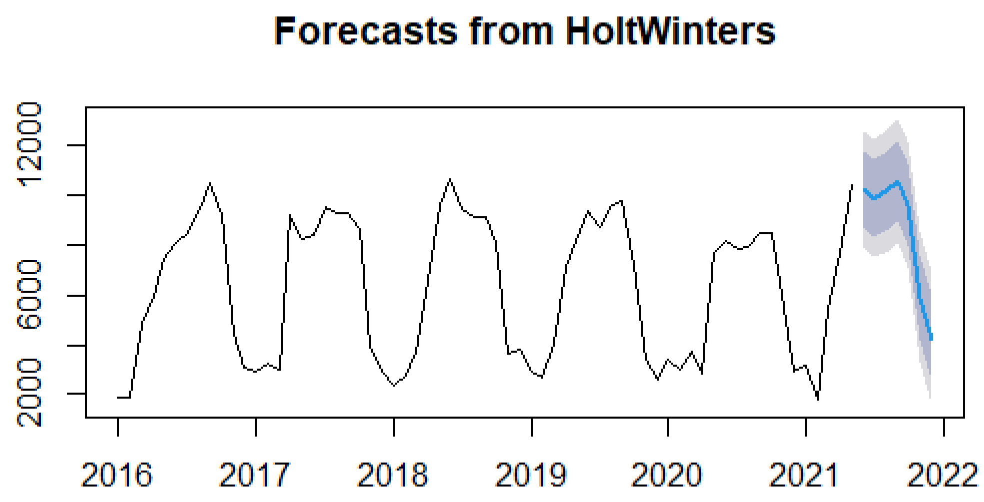
ME	RMSE	MAE	MPE	MAPE	MASE	ACF1
16.63878	921.9082	667.9987	-3.579507	13.57865	0.6312662	0.2100398

EXPONENTIAL SMOOTHING FORECASTING





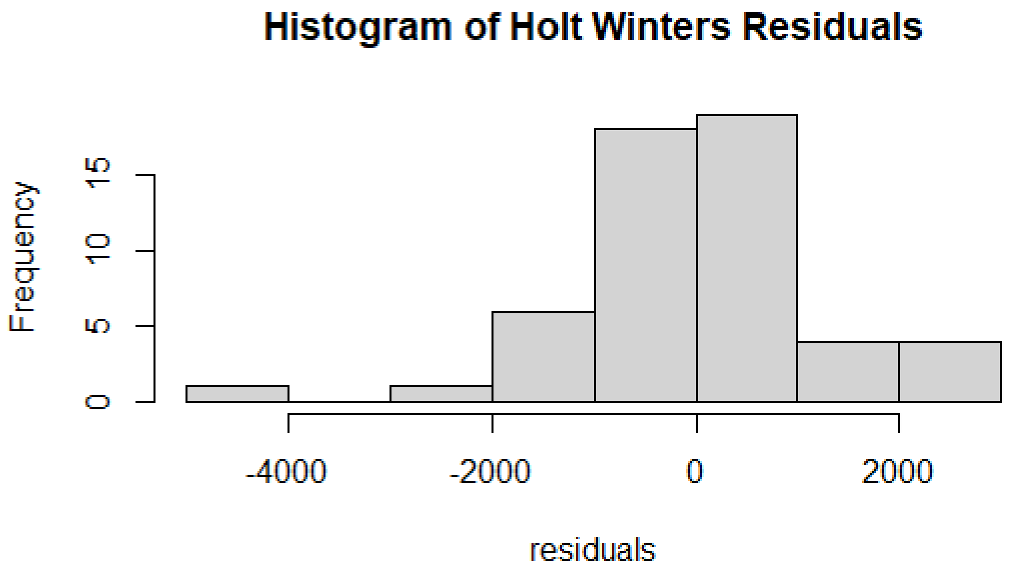
HOLT WINTERS FORECASTING

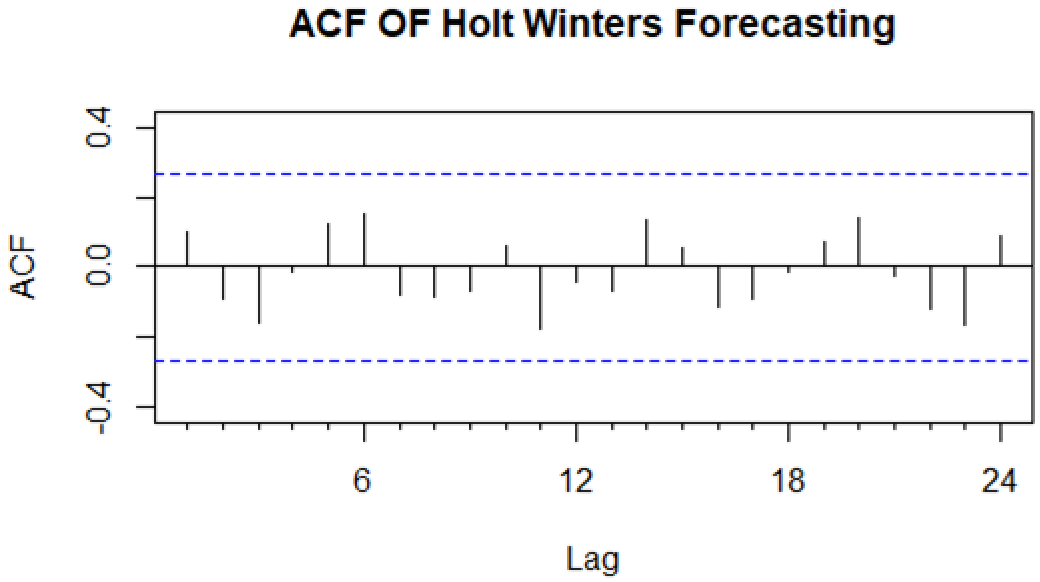


- ACCURACY VALUES

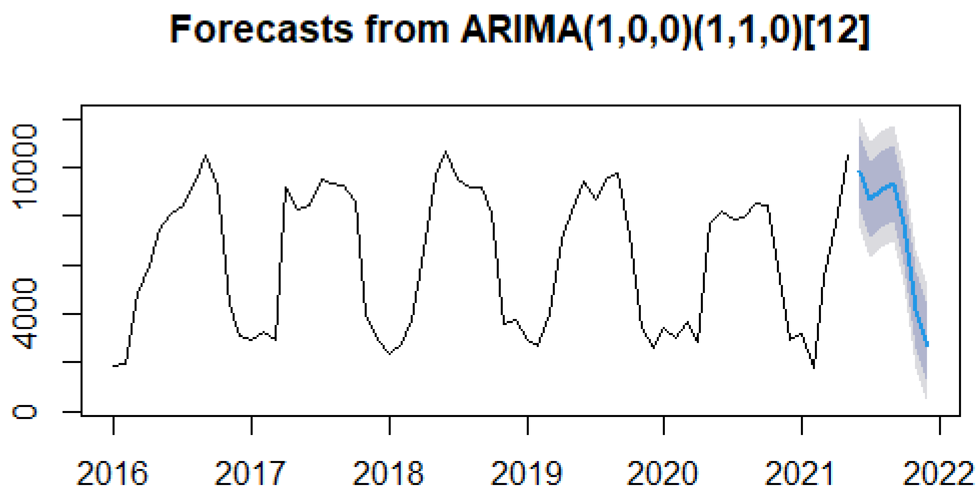
ME	RMSE	MAE	MPE	MAPE	MASE	ACF1
-20.55108	1179.395	795.9391	-2.795843	15.84126	0.7521712	0.1013042

HOLT WINTERS FORECASTING





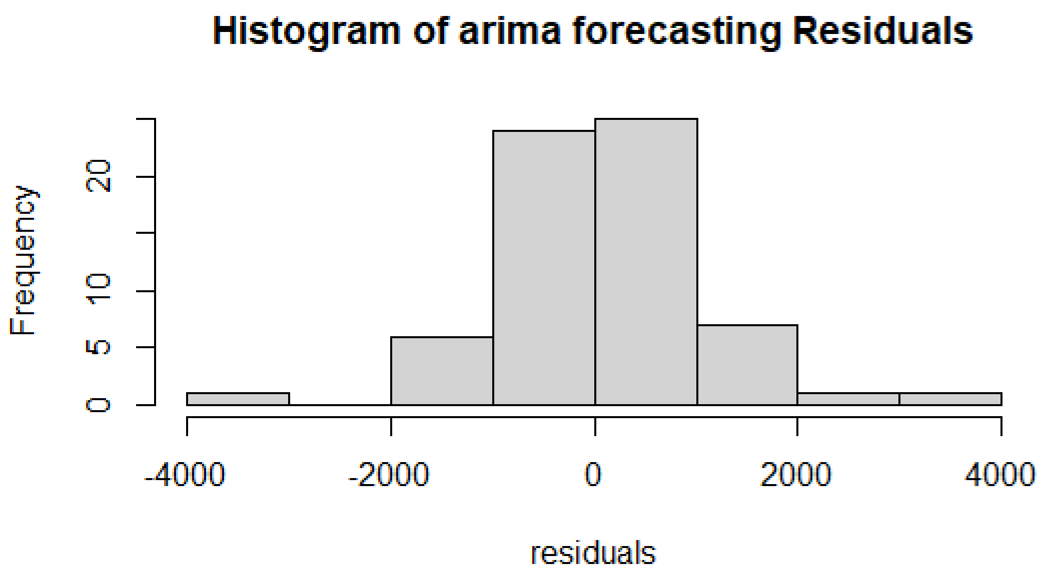
ARIMA



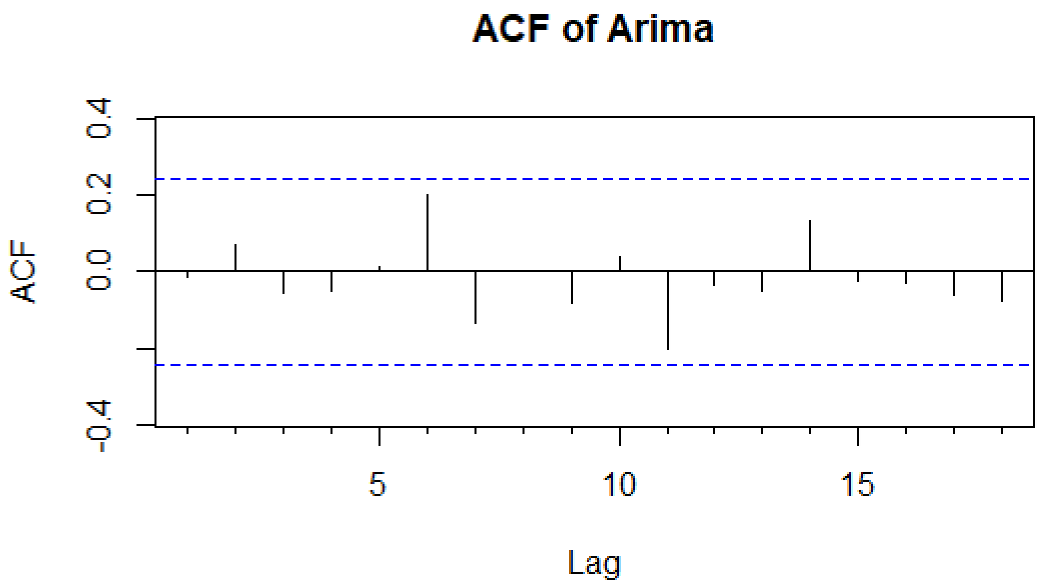
- ACCURACY VALUES

ME	RMSE	MAE	MPE	MAPE	MASE	ACF1
9.87972	1017.617	676.7122	-2.182681	13.46607	0.6395005	-0.014160

ARIMA



Histogram of arima forecasting Residuals
Frequency
0
5
10
20
-4000
-2000
0
2000
4000
residuals



ACF of Arima
ACF
-0.4
0.0
0.2
0.4
5
10
15
Lag

ACCURACIES OF FORECASTING USING MEAN PERCENTAGE ERROR

FORECASTING MODELS	MPE (%)
Simple Moving Averages (Order 12)	-0.0271099
Simple Moving Averages (Order 6)	-0.067717
Simple Moving Averages (Order 3)	-0.214658
ARIMA	-2.182681
Seasonal Naive	-2.659102
Holt Winters	-2.795843
Exponential Smoothing	-3.579507
Naïve	-4.691763
Random Walk	-4.691763
Mean Forecast	-31.73938

DRILLING DOWN TO THE BEST MODEL

2021	MA3	MA12
Jan	3469.664	5912.93
Feb	3527.736	5982.704
Mar	4850.309	6038.524
Apr	6702.962	6083.18
May	8561.114	6118.905
Jun	9513.069	6147.484
Jul	9737.048	6170.348
Aug	9797.566	6188.639
Sep	9508.67	6203.272
Oct	7785.793	6214.978
Nov	5630.931	6224.343
Dec	3778.935	6231.835

Why **Moving Averages of order 3** is our best forecasting model

- It is the best for forecasting commodities with constant demand, where there is a seasonality or slight trend.
- Useful for separating out random variations.
- More accurate monthly insights lead to better business decisions.
- Simplicity of application and interpretation makes it possible to plot several different moving average lines at the same time.
- Gives constant forecasts.
- Great at smoothing data, form trend lines, and create an easily interpreted visual aid.

CONCLUSION

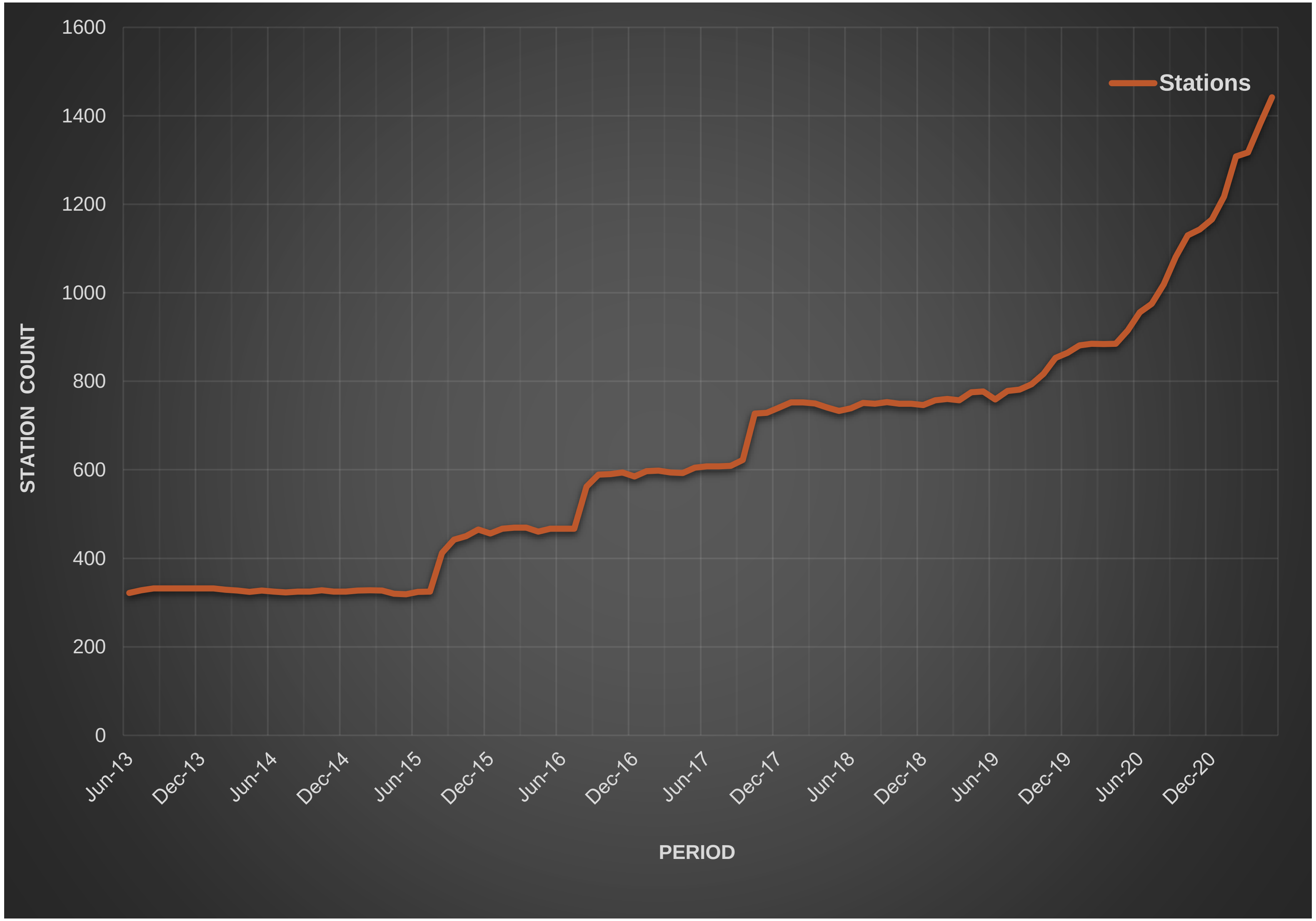
- For a yearly overview of Citi bike trips usage, simple moving averages forecast of order 12 provides the highest accuracy with a margin of 0.02% error being under-forecasted.

- For any monthly, weekly or daily forecasts involving business decisions, moving averages of order 3 stands the best model to predict Citi bike trips data with an accuracy of 99.78%.

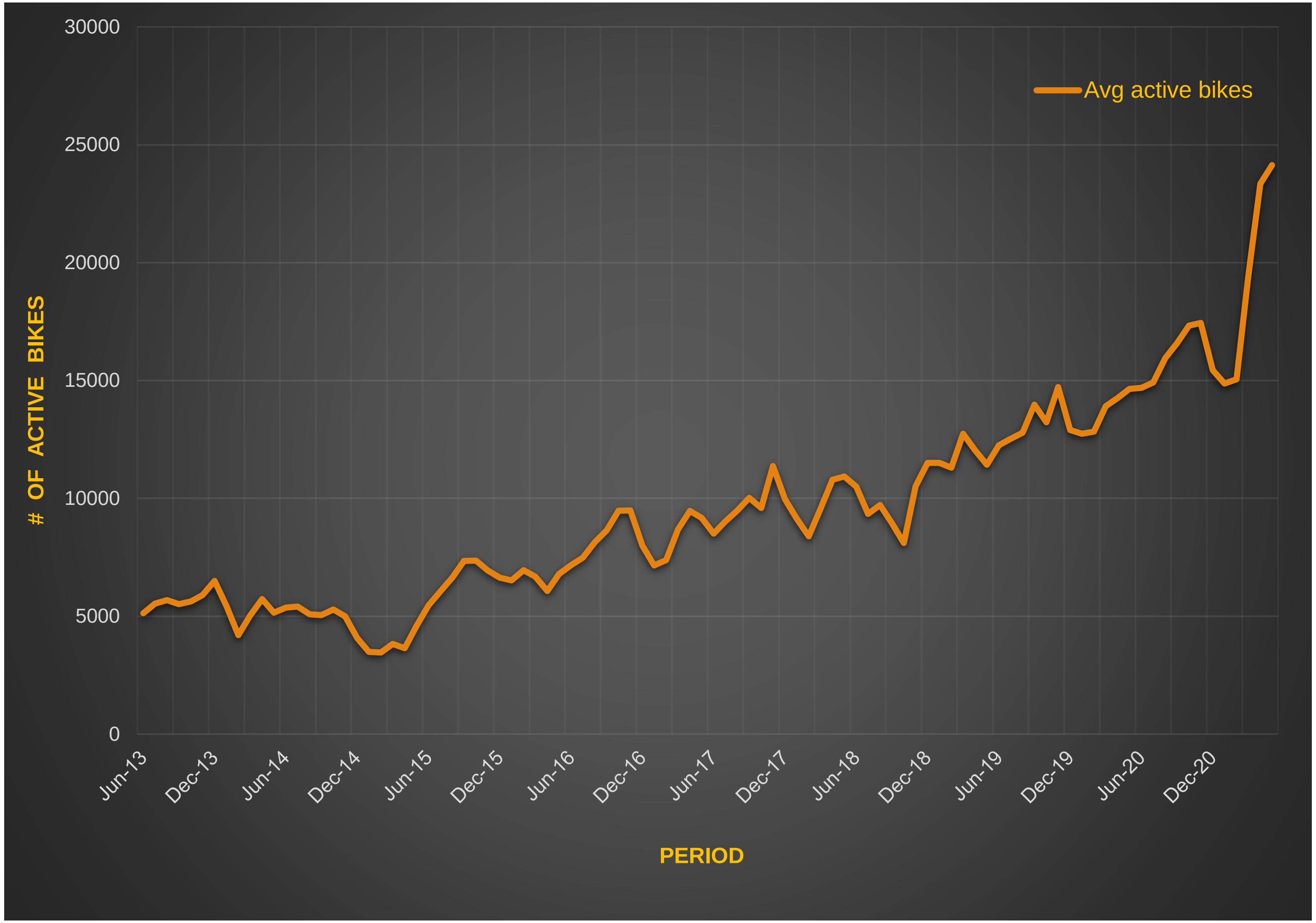
THANK YOU

APPENDIX

DOWNLOAD THE APP
citibikenyc.com/app
12 Avenue & W 40 Street
Pier 79 Ferry Terminal
citi bike
NY WATERWAY

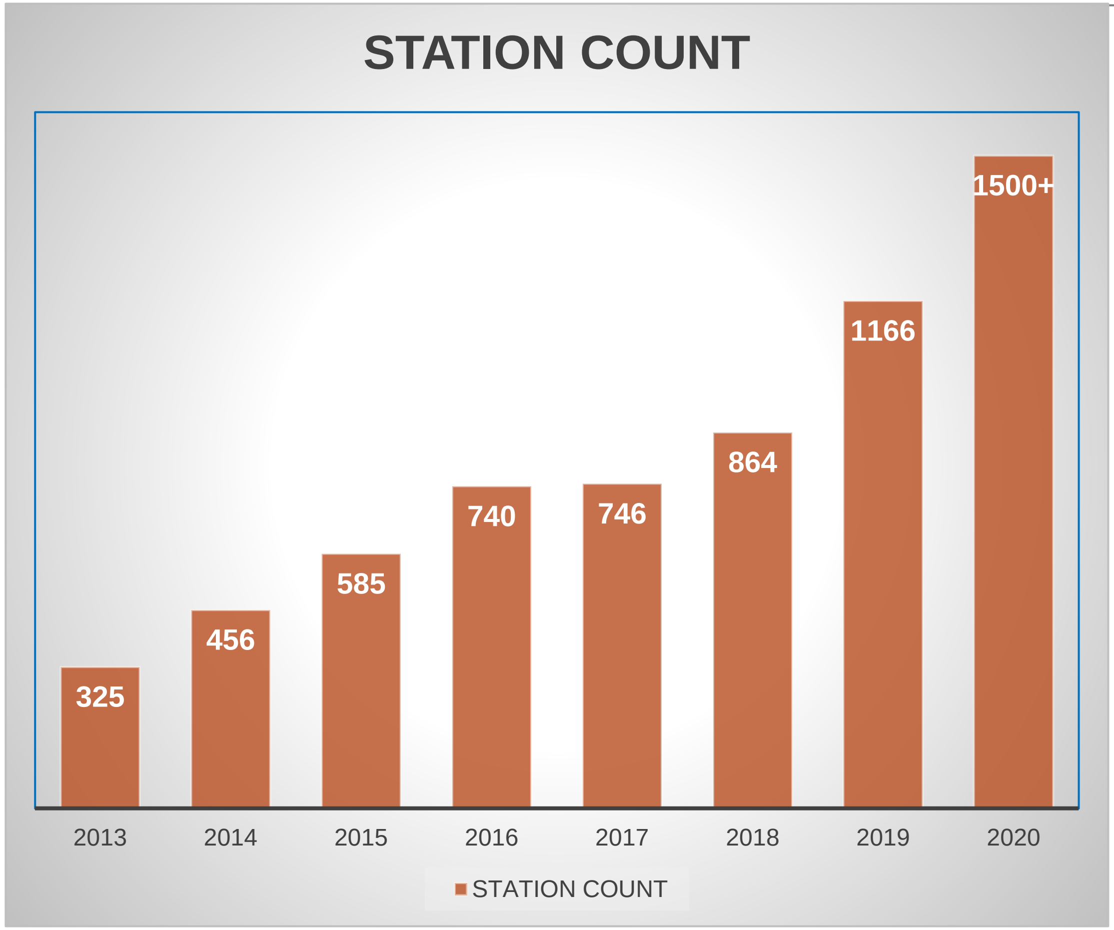


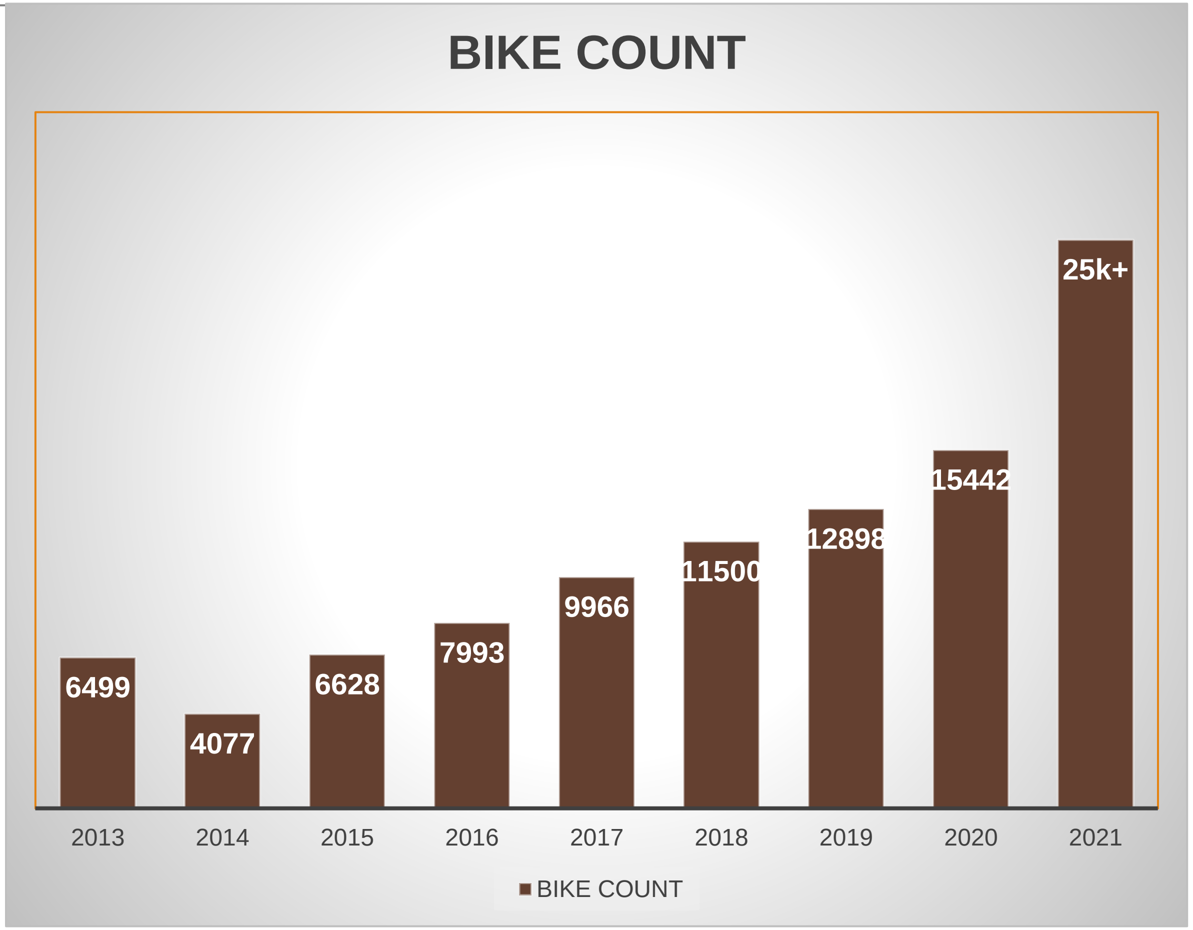
NEW STATION DISTRIBUTION
(MONTH OVER MONTH)



ACTIVE BIKE DISTRIBUTION
(MONTH OVER MONTH)

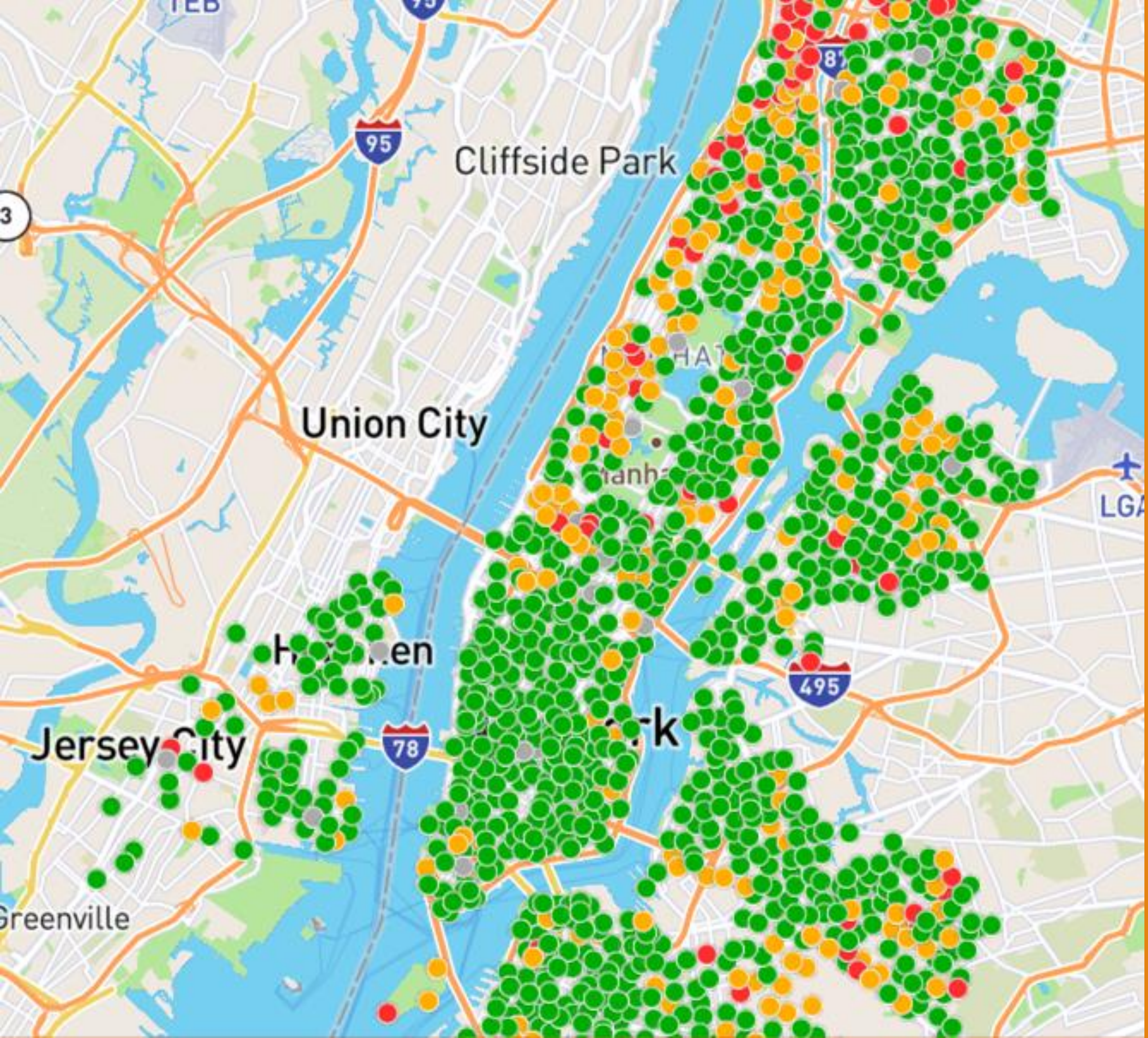
Growth and Expansion Since onset





Year	Year Ending Memberships (Annual)	Total Annual Revenue
2013	96,125	$5.8 million
2014	88,405	$8.0 million
2015	92,781	$10.0 million
2016	119,681	$14.1 million
2017	136,702	$16.7 million
2018	147,090	$17.9 million
2019	149,740	$20.5 million
2020	167,556	$19.5 million
2021	150,000 (as of Oct)	> $25 million

ANNUAL MEMBERSHIP AND REVENUE



CITI BIKE STATION MAP

1500+ STATIONS IN NYC

80+ STATIONS IN NJ

STATION USED FOR OUR ANALYSIS