Predicting 2020 CMS Physician Payments

Presented by: Sarah Sturgeon, Data Scientist

Agenda

- Problem Statement
- Data Source & Preparation
- EDA & Modeling
- Conclusion
- Questions

Problem Statement

Is it possible to predict CMS payments for 2020 made to physicians in Cincinnati Ohio for the top 10 most reimbursed drugs or biologics based on historical CMS payments from 2013 through 2019?

janssen
PHARMACEUTICAL COMPANIES OF
Johnson & Johnson

NOVARTIS

astellas

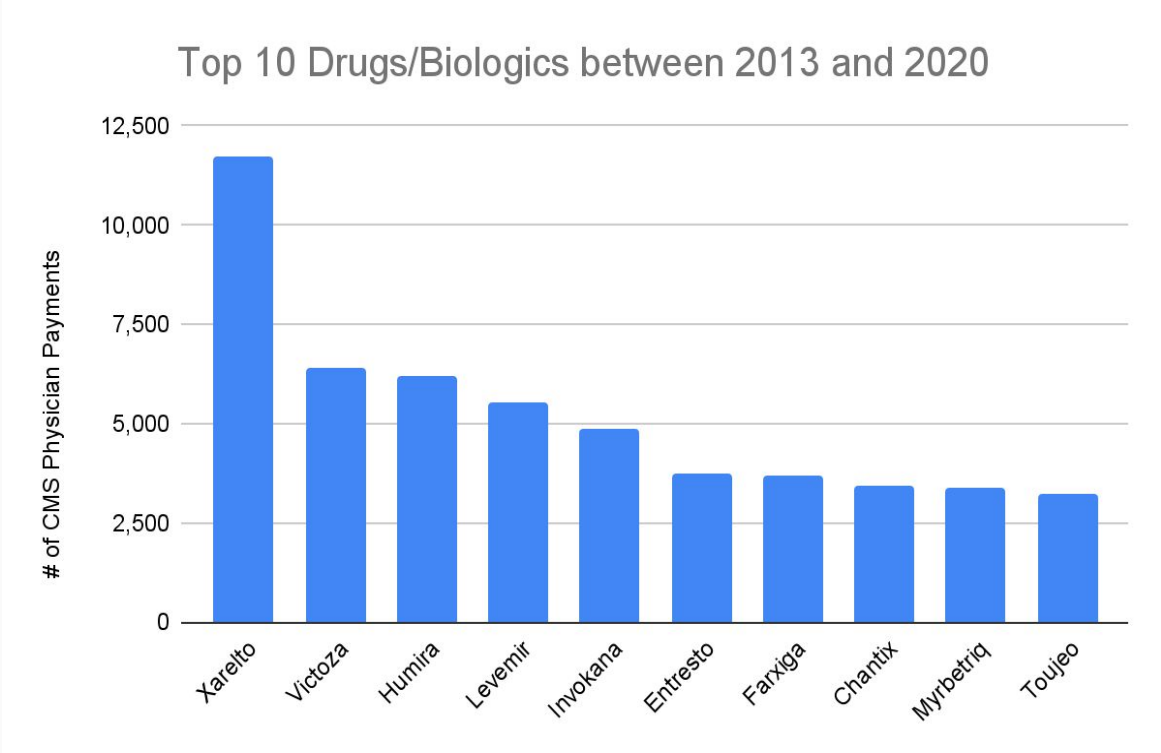
Pfizer

AstraZeneca

Top 10 drug/biologic payments

to Cincinnati physicians

between 2013 & 2020



Data Source & Preparation

The data for this project was acquired from the Center for Medicare and Medicaid Services Open Payments Dataset Downloads page (https://www.cms.gov/OpenPayments/Data/Dataset-Downloads).

- All data used is from the general payments data files from years 2013 - 2020.
    - Teaching hospital transactions were excluded
    - Payments of only one transaction were chosen
    - Payments made to physicians with a city of “Cincinnati” or “CINCINNATI” were chosen
    - Only payments related to the top 10 most reimbursed drugs/biologics were chosen
    - Payment outliers were excluded at the physician level before aggregation
    - Data from the individual files above were merged into one dataset

- 3 of the top 10 most common drug/biologic payments disappeared in 2020 in the Cincinnati area

Modeling

Cincinnati daily CMS physician payments were aggregated by date, by week and by month.

39 models were created for each of 7 different drugs/biologics

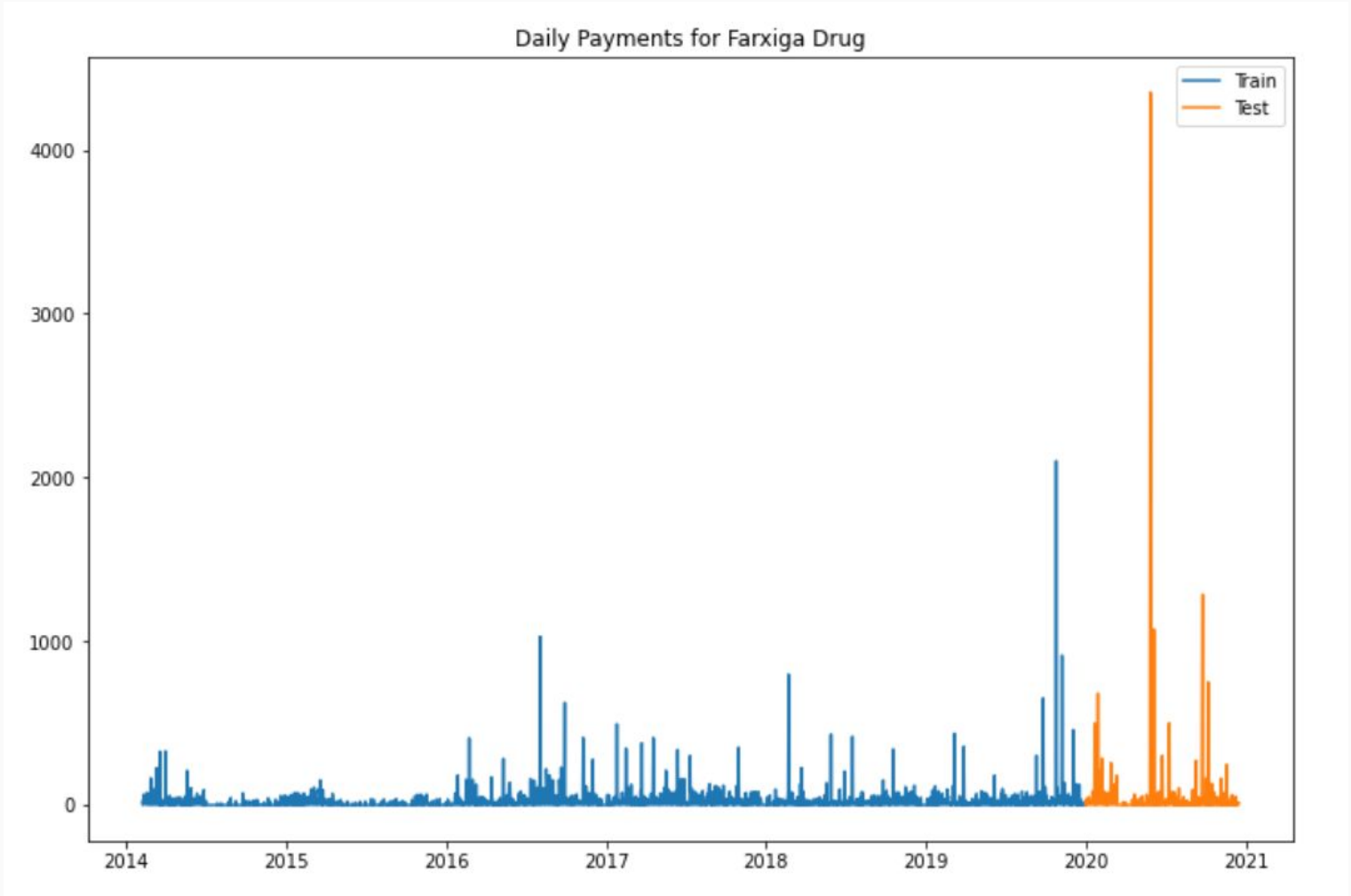
11 for aggregated daily data

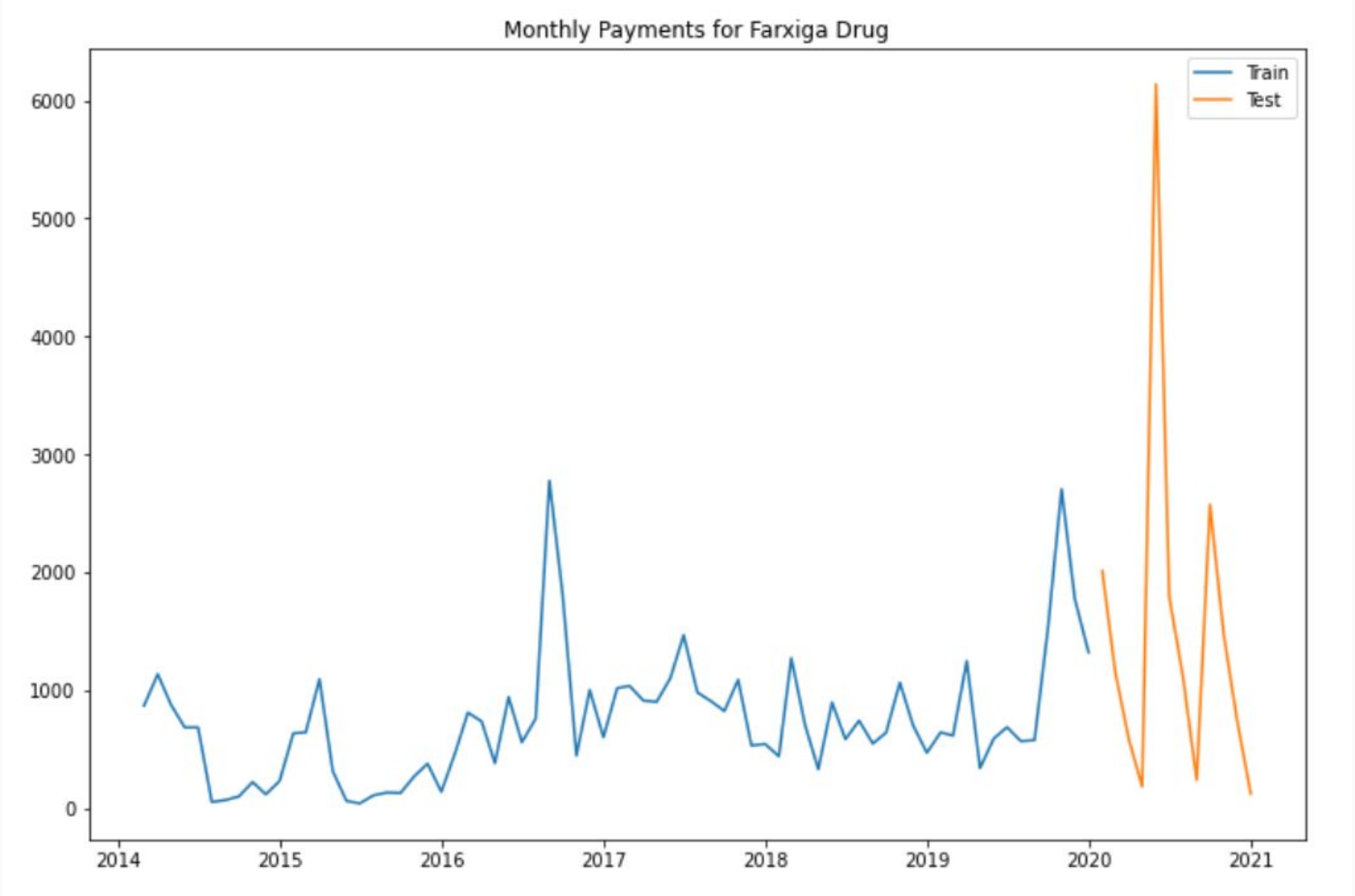
10 for aggregated weekly data

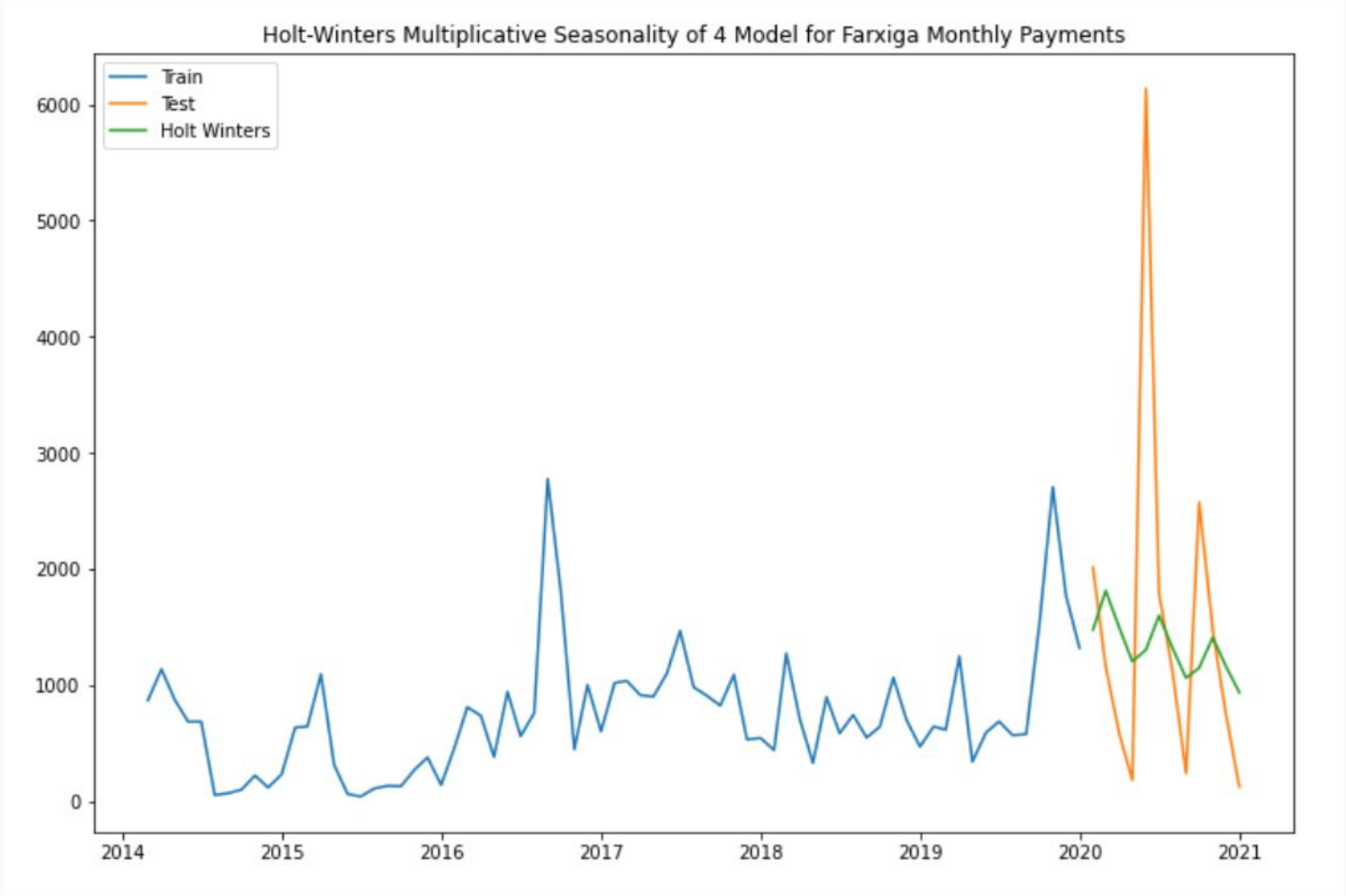
18 for aggregated monthly data

The drug/biologic payments modeled were:

- Farxiga
- Humira
- Xarelto
- Invokana
- Entresto
- Chantix
- Myrbetriq

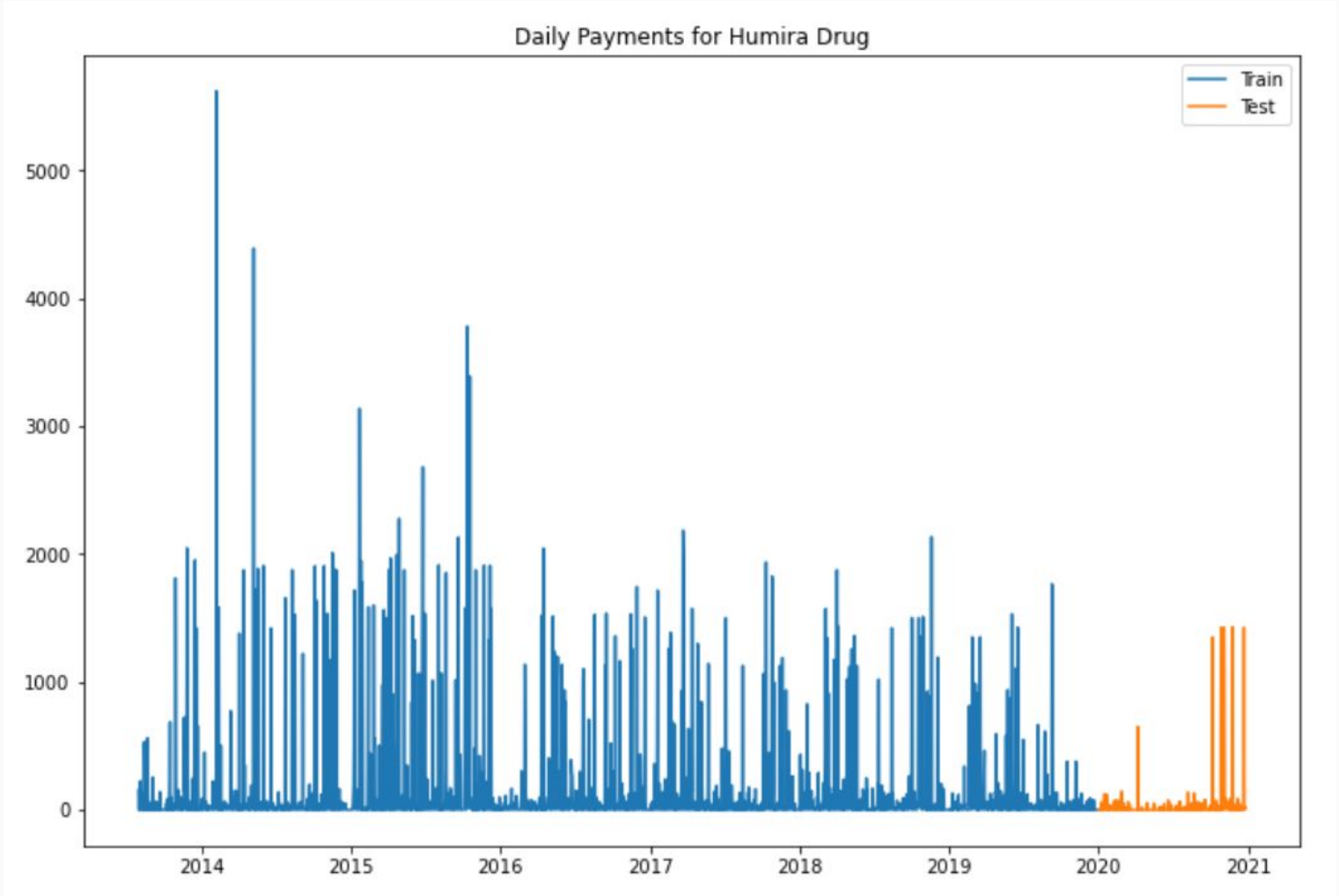

Daily Payments for Farxiga Drug
Train
Test
4000
3000
2000
1000
0
2014
2015
2016
2017
2018
2019
2020
2021

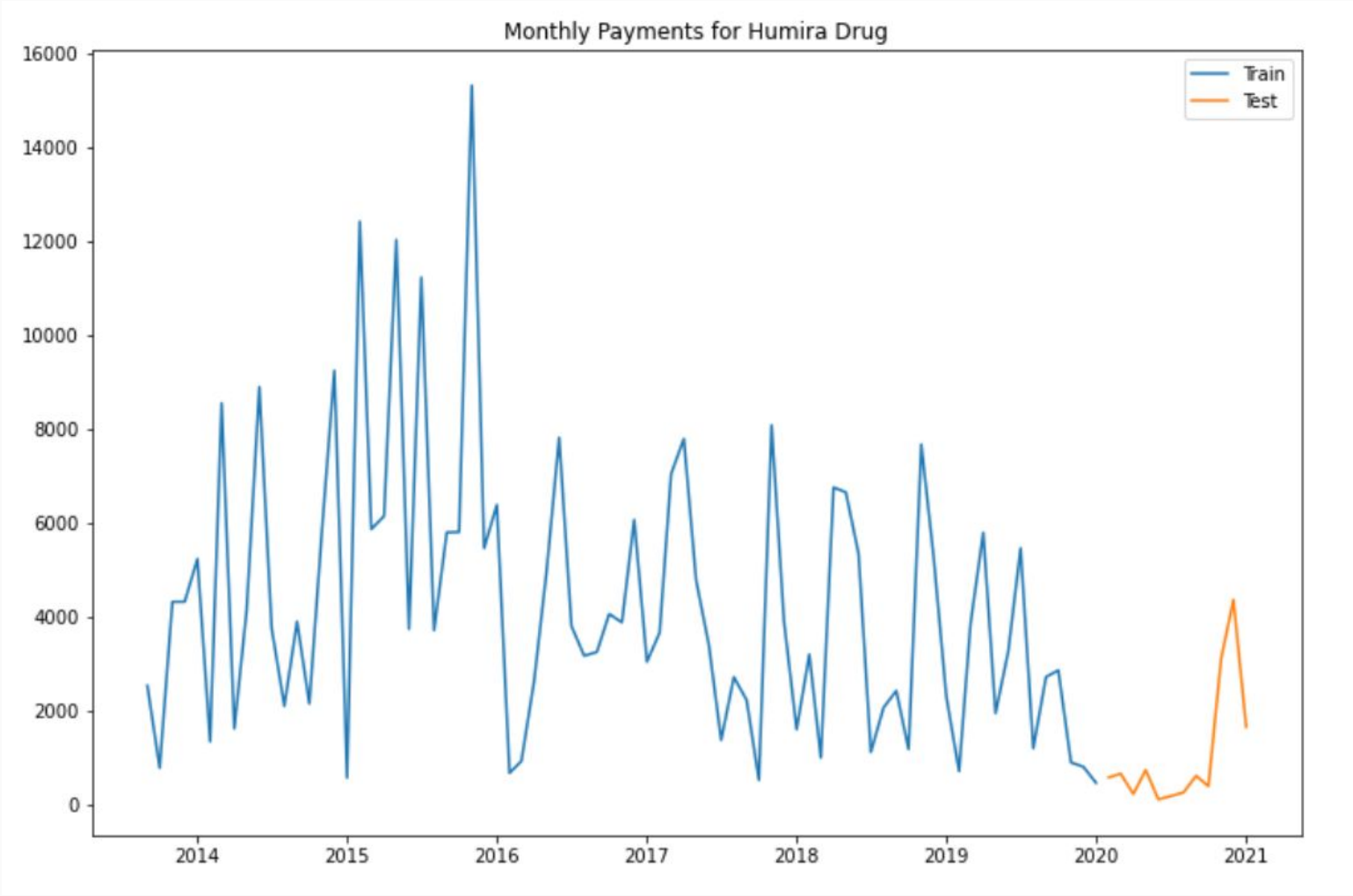

Monthly Payments for Farxiga Drug
Train
Test
6000
5000
4000
3000
2000
1000
0
2014
2015
2016
2017
2018
2019
2020
2021

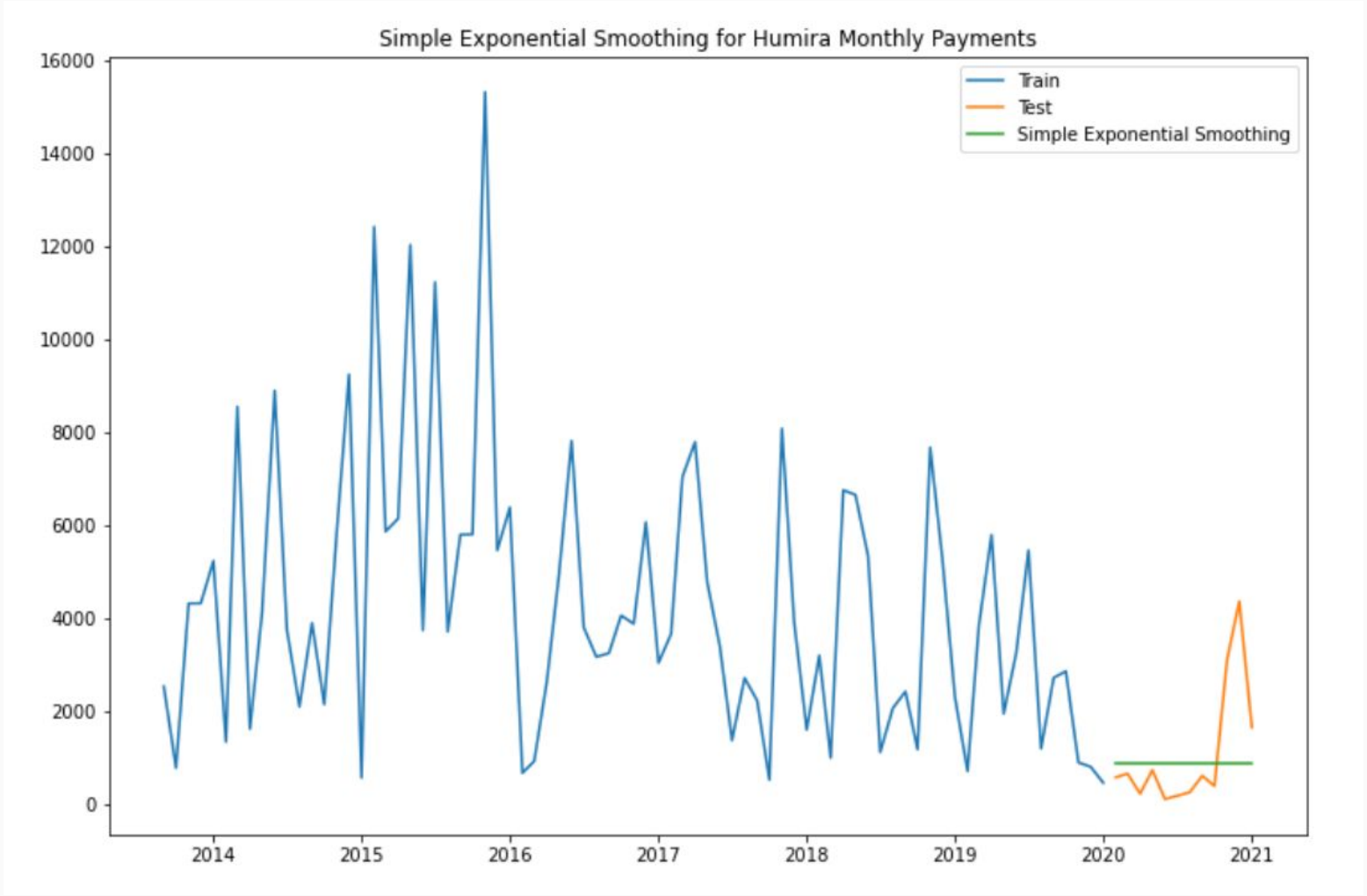

Holt-Winters Multiplicative Seasonality of 4 Model for Farxiga Monthly Payments
Train
Test
Holt Winters
6000
5000
4000
3000
2000
1000
0
2014
2015
2016
2017
2018
2019
2020
2021

	RMSE
MODEL	FARXIGA
Baseline (Mean)	1,754.97
Baseline (Shift) 4	2,178.44
Baseline (Shift) 6	1,816.32
Baseline (Shift) 12	1,827.36
Baseline (Shift) 18	1,773.59
Simple Exp Smoothing	1,577.67
HW M 4	1,568.13
HW M 6	1,672.99
HW M 12	1,695.23
HW M 18	1,693.55
HW A 4	1,607.81
HW A 6	1,621.86
HW A 12	1,625.30
HW A 18	1,590.60
SARIMA 4	1,598.30
SARIMA 6	1,598.30
SARIMA 12	1,598.30
SARIMA 18	1,598.30

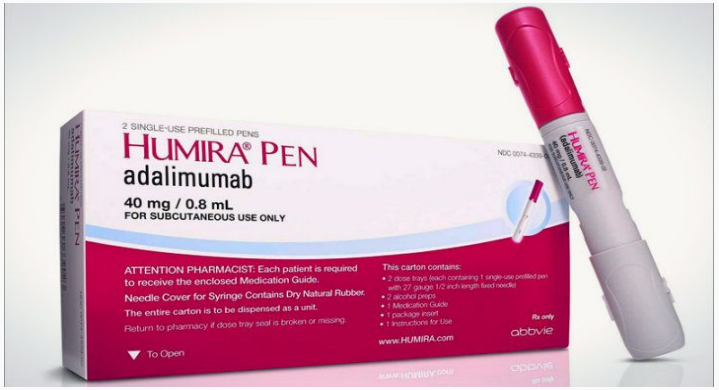
The Holt-Winters Multiplicative Model with a period of 4 was the strongest model.


Daily Payments for Humira Drug
Train
Test
5000
4000
3000
2000
1000
0
2014
2015
2016
2017
2018
2019
2020
2021

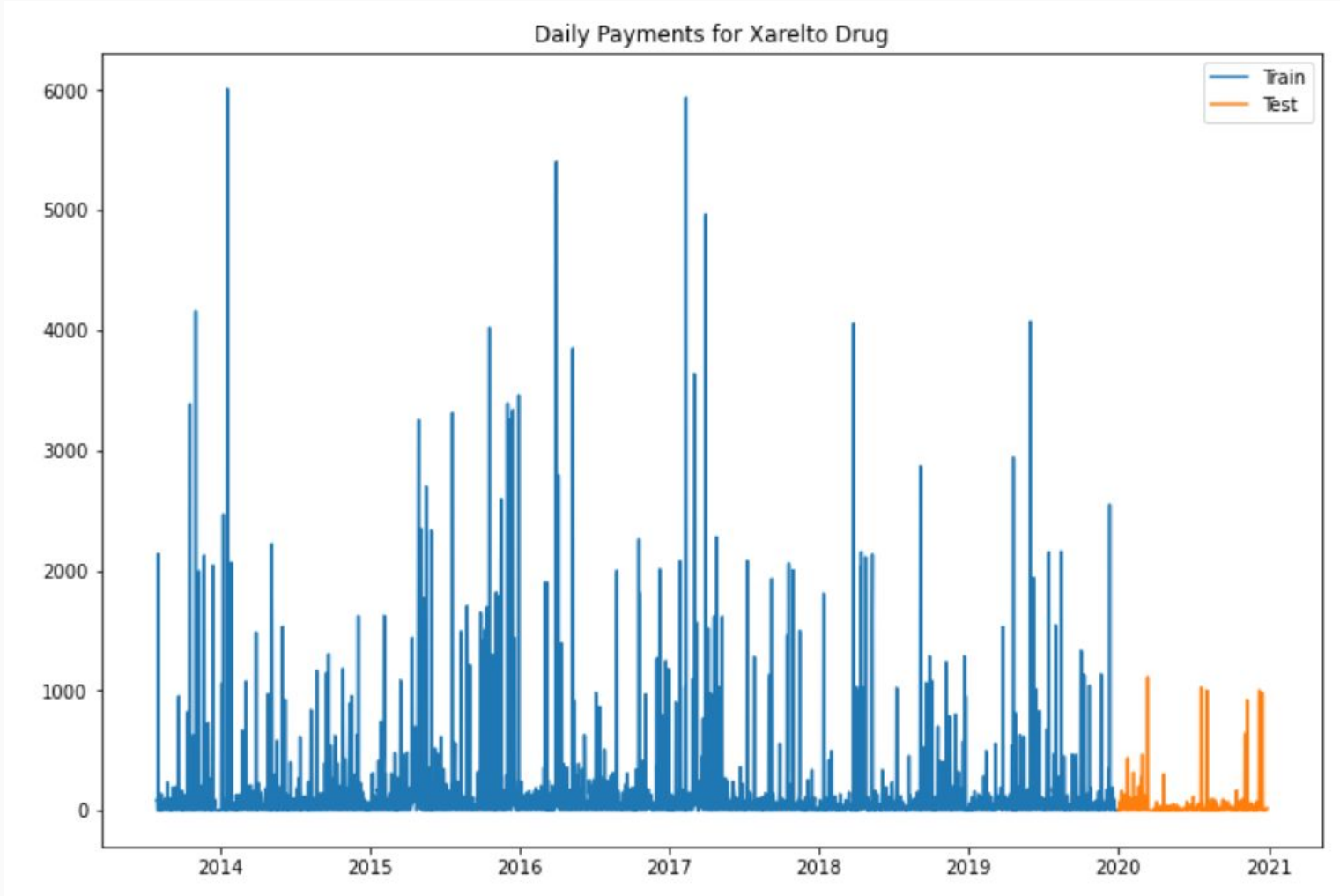

Monthly Payments for Humira Drug
Train
Test
16000
14000
12000
10000
8000
6000
4000
2000
0
2014
2015
2016
2017
2018
2019
2020
2021

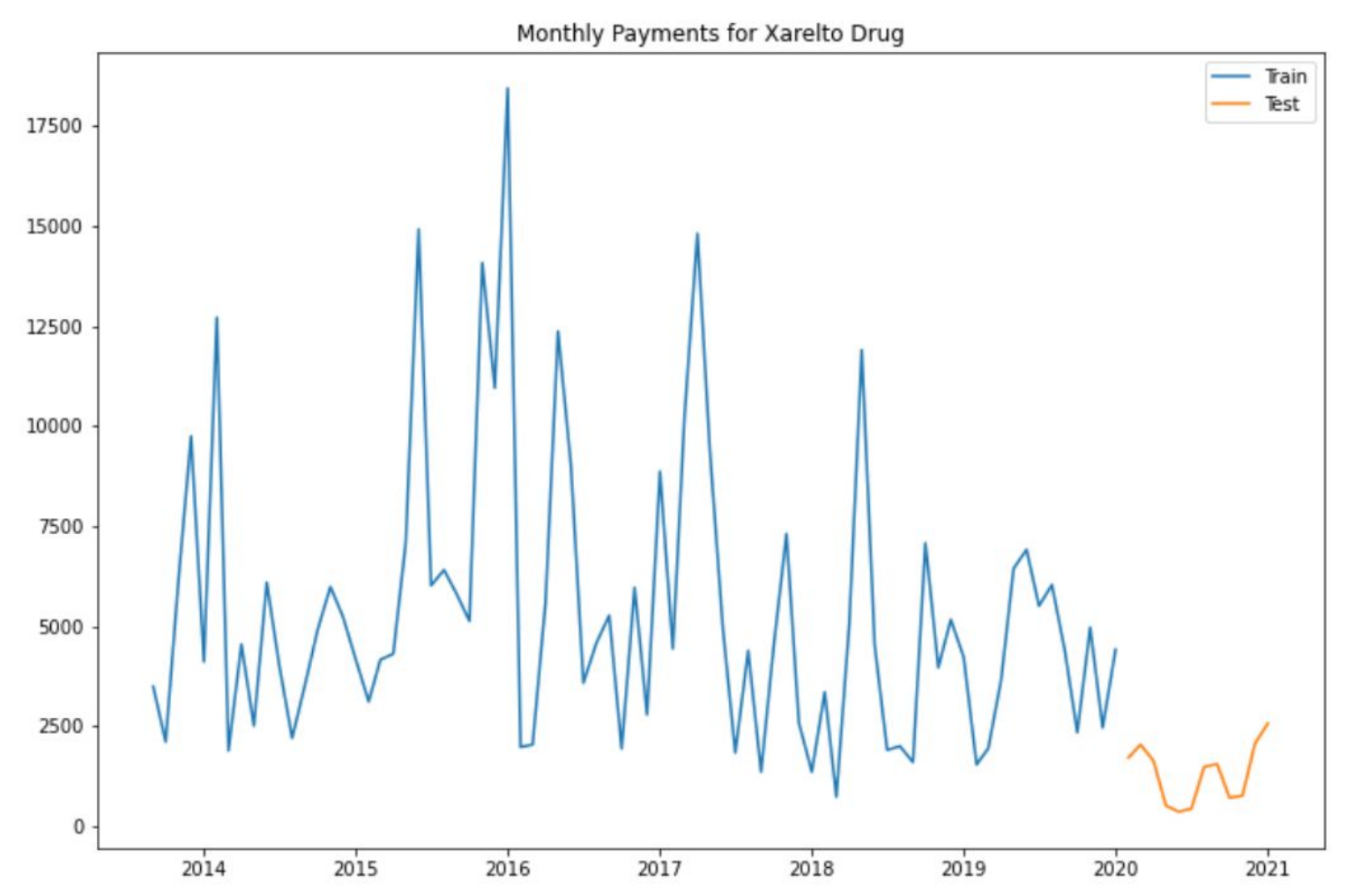

Simple Exponential Smoothing for Humira Monthly Payments
Train
Test
Simple Exponential Smoothing
16000
14000
12000
10000
8000
6000
4000
2000
0
2014
2015
2016
2017
2018
2019
2020
2021

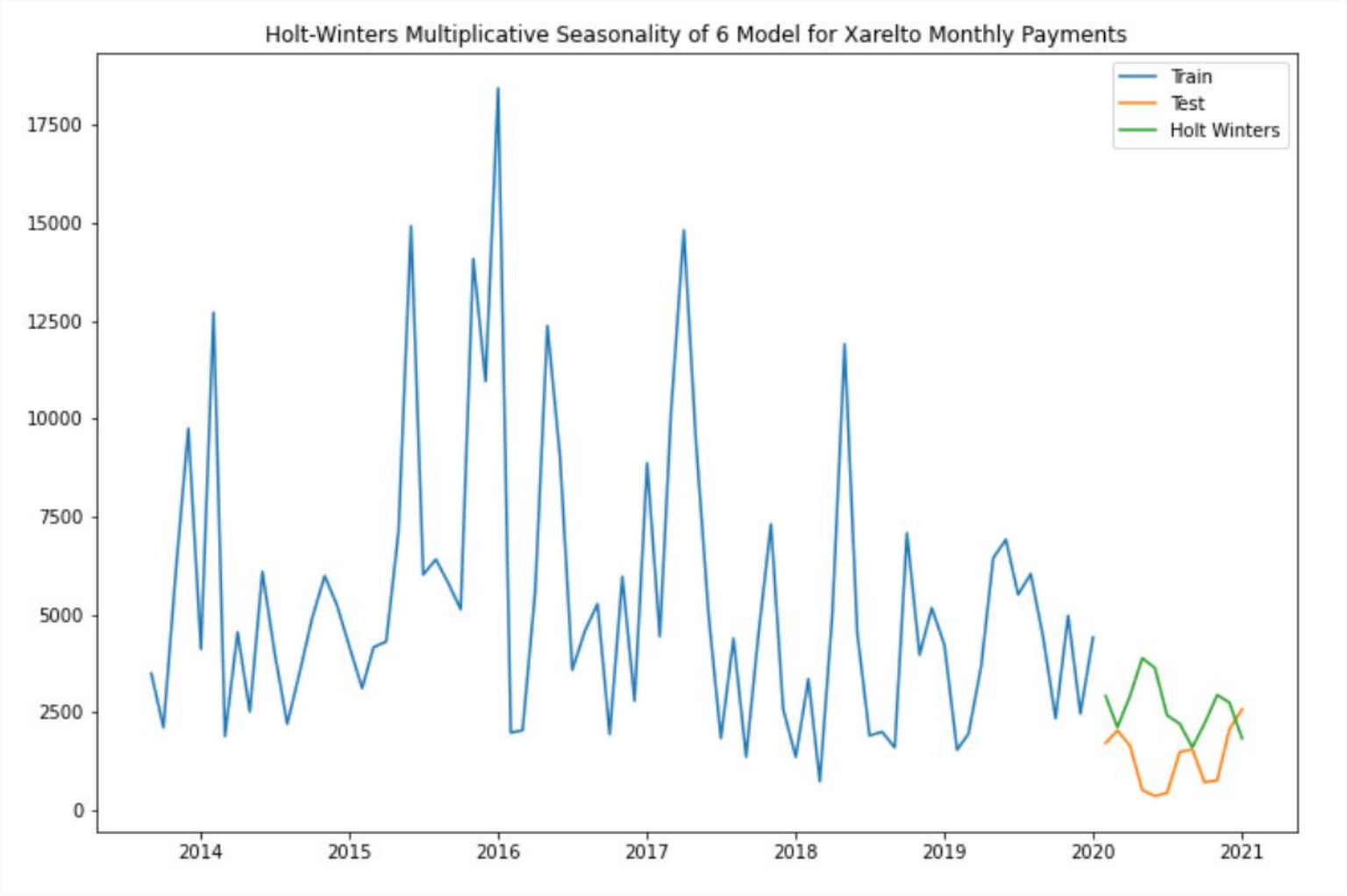
	RMSE
MODEL	HUMIRA
Baseline (Mean)	3,450.91
Baseline (Shift) 4	1,774.58
Baseline (Shift) 6	1,919.65
Baseline (Shift) 12	3,038.17
Baseline (Shift) 18	3,443.70
Simple Exp Smoothing	1,290.19
HW M 4	1,557.79
HW M 6	1,886.97
HW M 12	1,551.92
HW M 18	2,885.98
HW A 4	1,626.42
HW A 6	1,452.41
HW A 12	1,720.23
HW A 18	2,523.96
SARIMA 4	1,834.11
SARIMA 6	1,620.61
SARIMA 12	1,801.46
SARIMA 18	1,801.46



The Simple Exponential Smoothing Model was the strongest model for Humira.

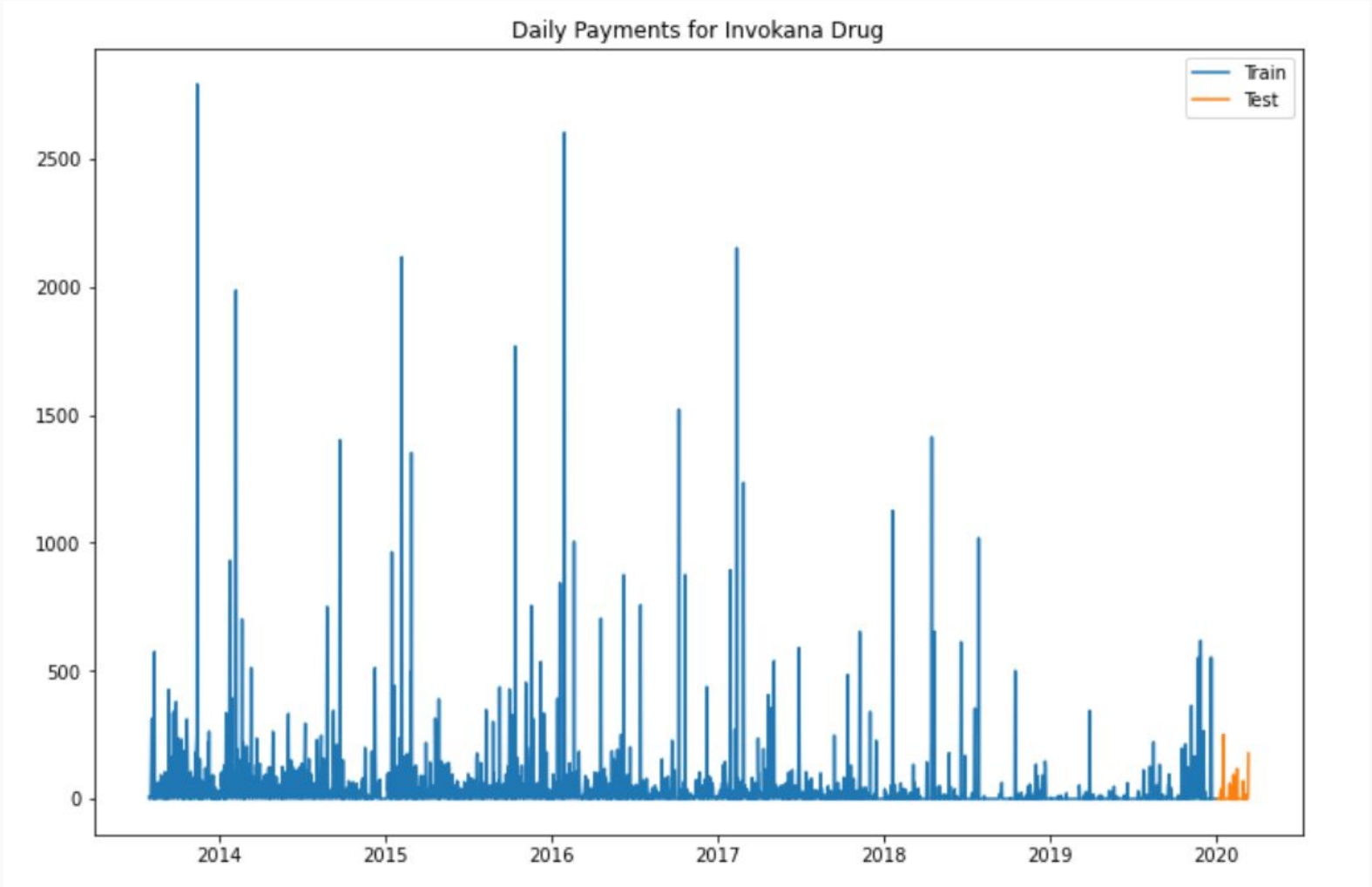

Daily Payments for Xarelto Drug
Train
Test
6000
5000
4000
3000
2000
1000
0
2014
2015
2016
2017
2018
2019
2020
2021

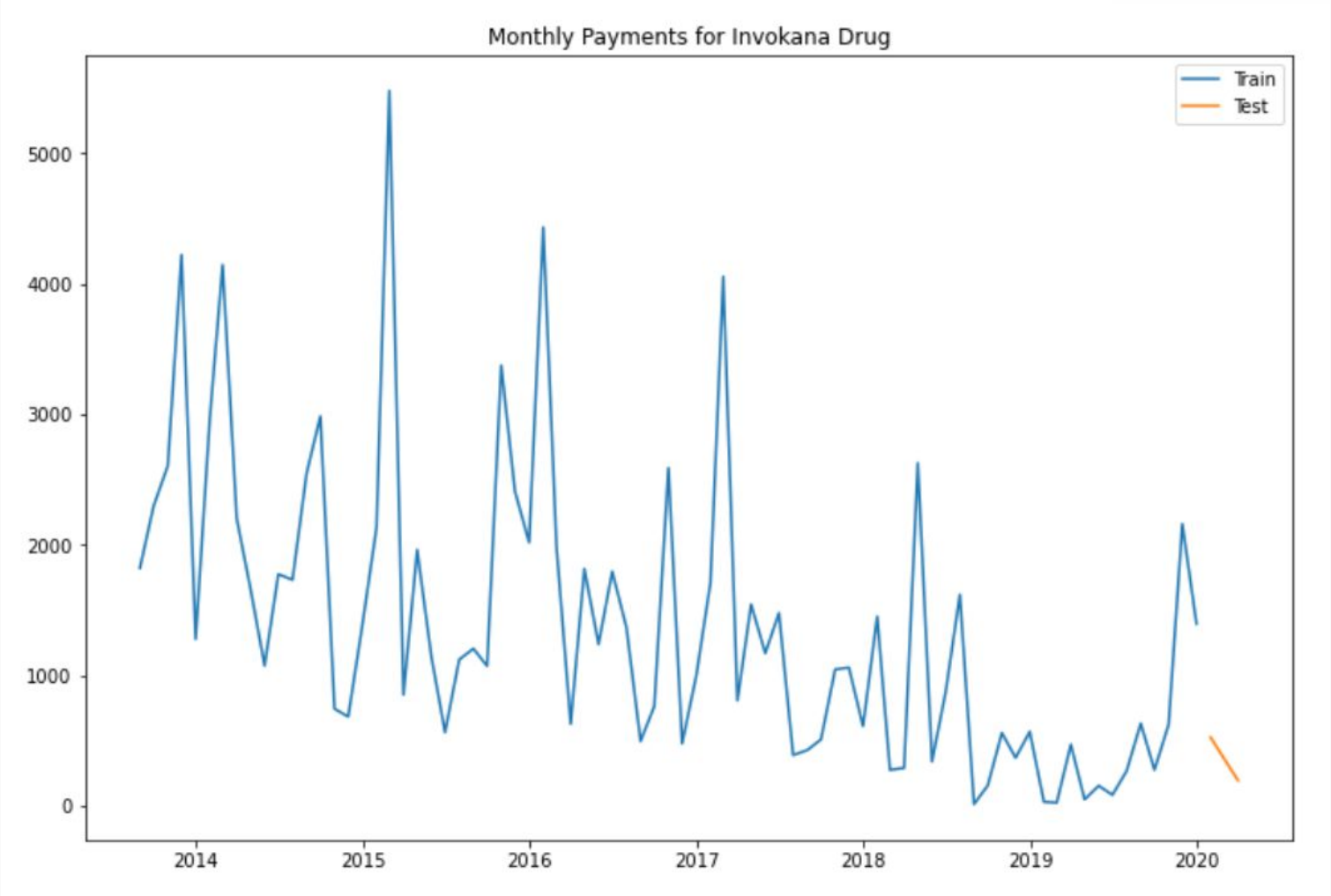

Monthly Payments for Xarelto Drug
Train
Test
17500
15000
12500
10000
7500
5000
2500
0
2014
2015
2016
2017
2018
2019
2020
2021

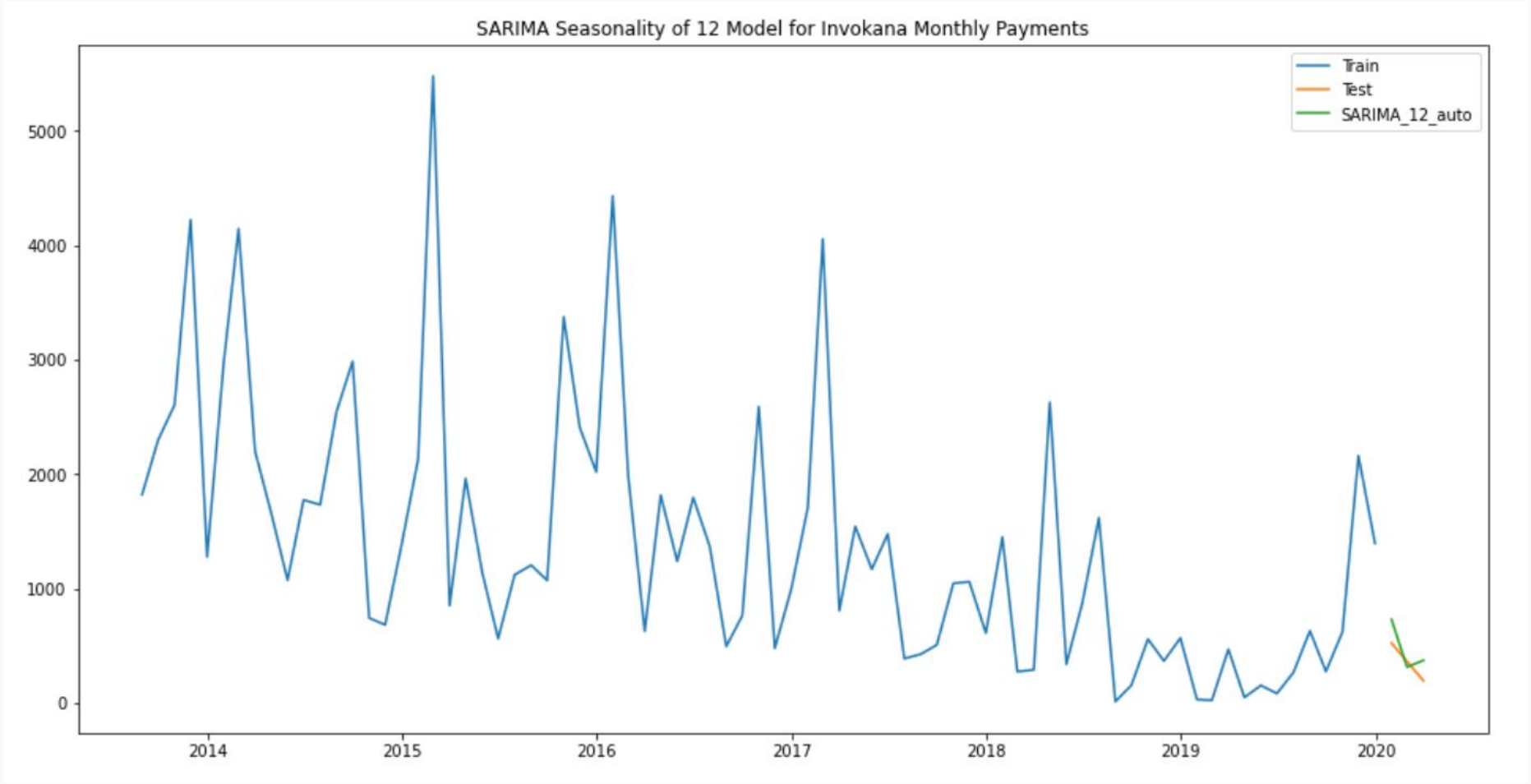

Holt-Winters Multiplicative Seasonality of 6 Model for Xarelto Monthly Payments
Train
Test
Holt Winters
17500
15000
12500
10000
7500
5000
2500
0
2014
2015
2016
2017
2018
2019
2020
2021

	RMSE
MODEL	XARELTO
Baseline (Mean)	4,156.02
Baseline (Shift) 4	1,864.04
Baseline (Shift) 6	2,602.50
Baseline (Shift) 12	3,656.89
Baseline (Shift) 18	3,562.85
Simple Exp Smoothing	2,684.34
HW M 4	2,392.69
HW M 6	1,769.89
HW M 12	2,619.00
HW M 18	4,381.07
HW A 4	2,382.07
HW A 6	2,709.44
HW A 12	2,663.74
HW A 18	4,173.43
SARIMA 4	4,126.28
SARIMA 6	4,126.28
SARIMA 12	3,736.63
SARIMA 18	4,126.28

The Holt-Winters Multiplicative Model with period 6 was the strongest model for Xarelto.

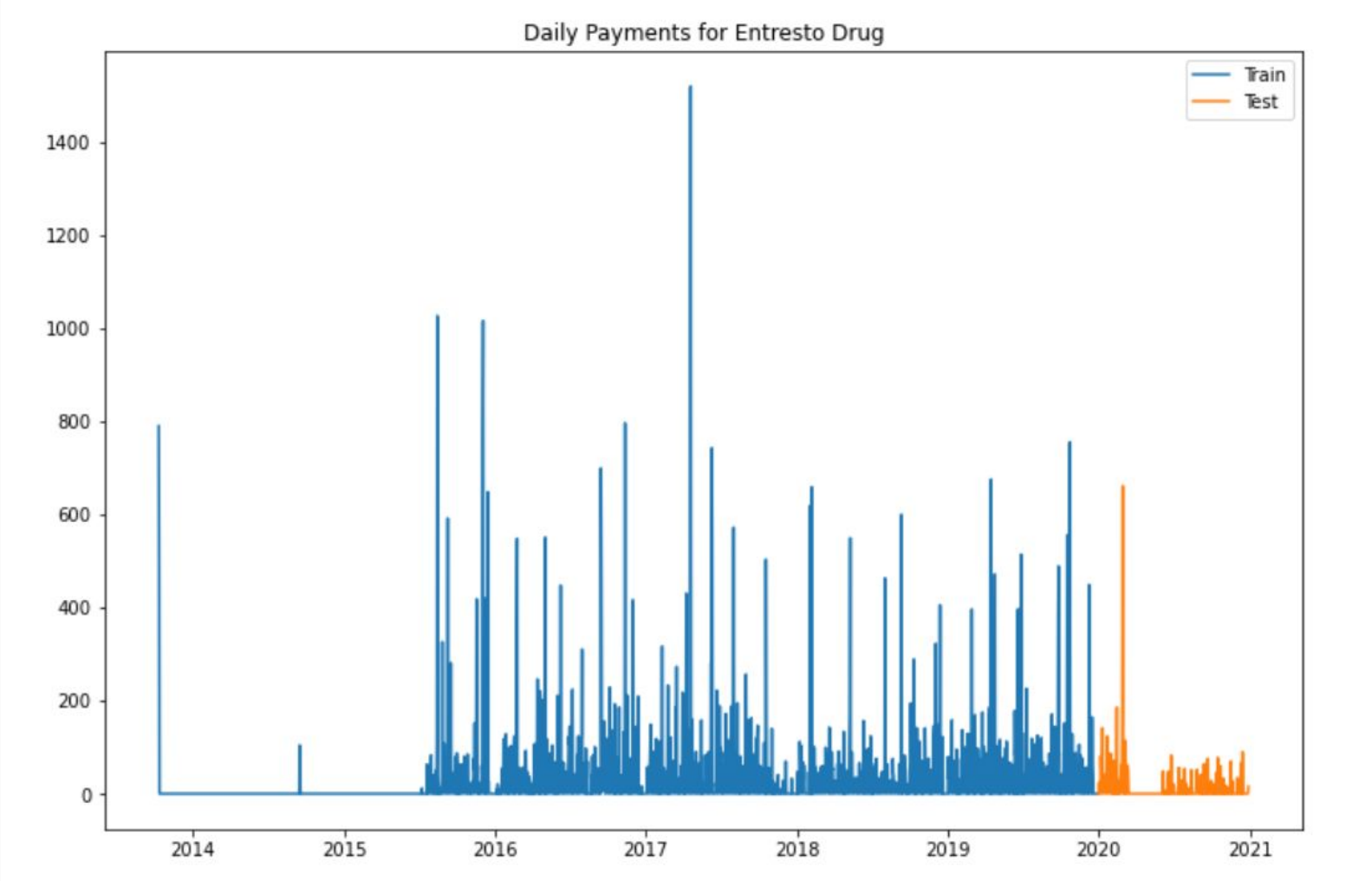

Daily Payments for Invokana Drug
Train
Test
2500
2000
1500
1000
500
0
2014
2015
2016
2017
2018
2019
2020

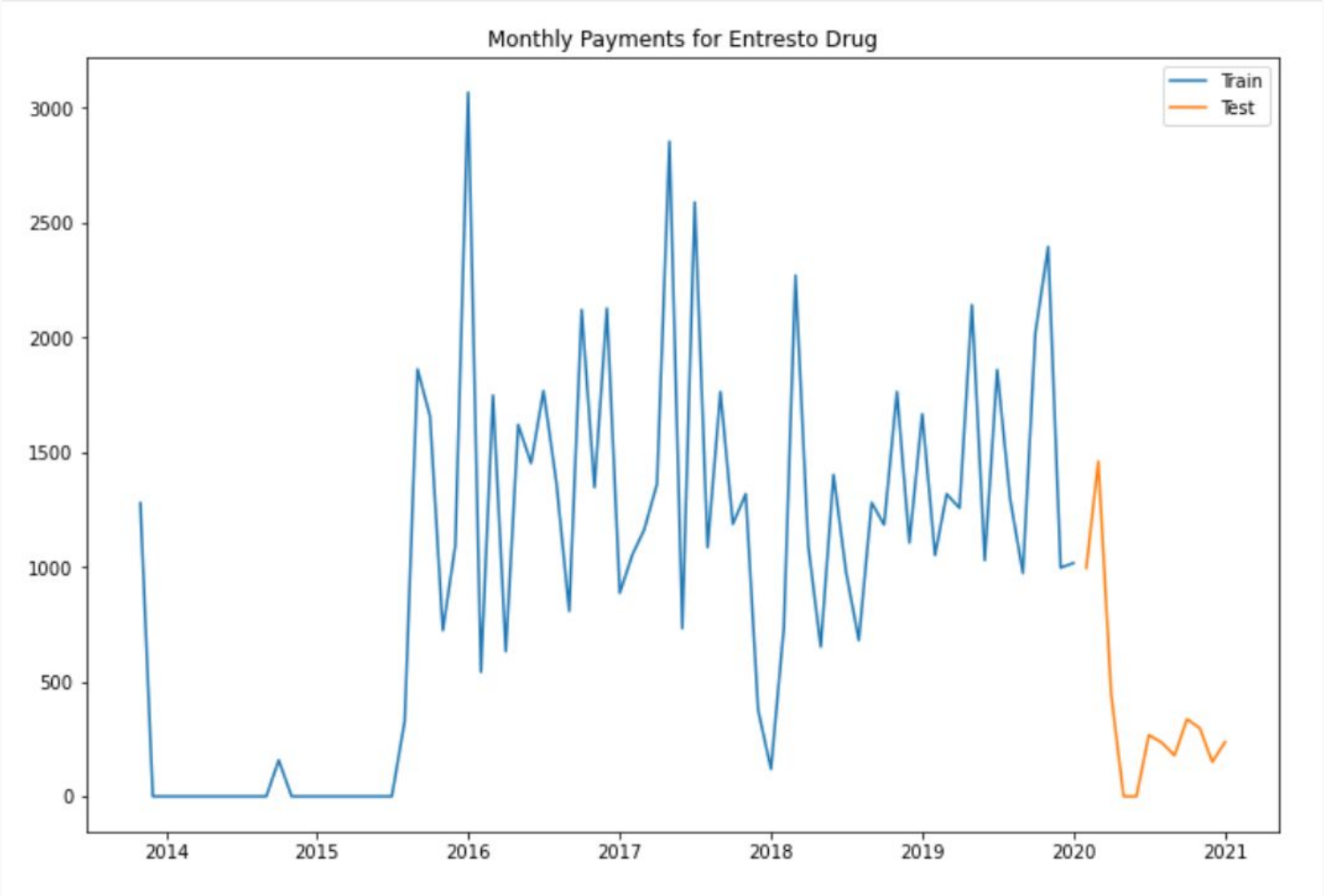

Monthly Payments for Invokana Drug
Train
Test
5000
4000
3000
2000
1000
0
2014
2015
2016
2017
2018
2019
2020

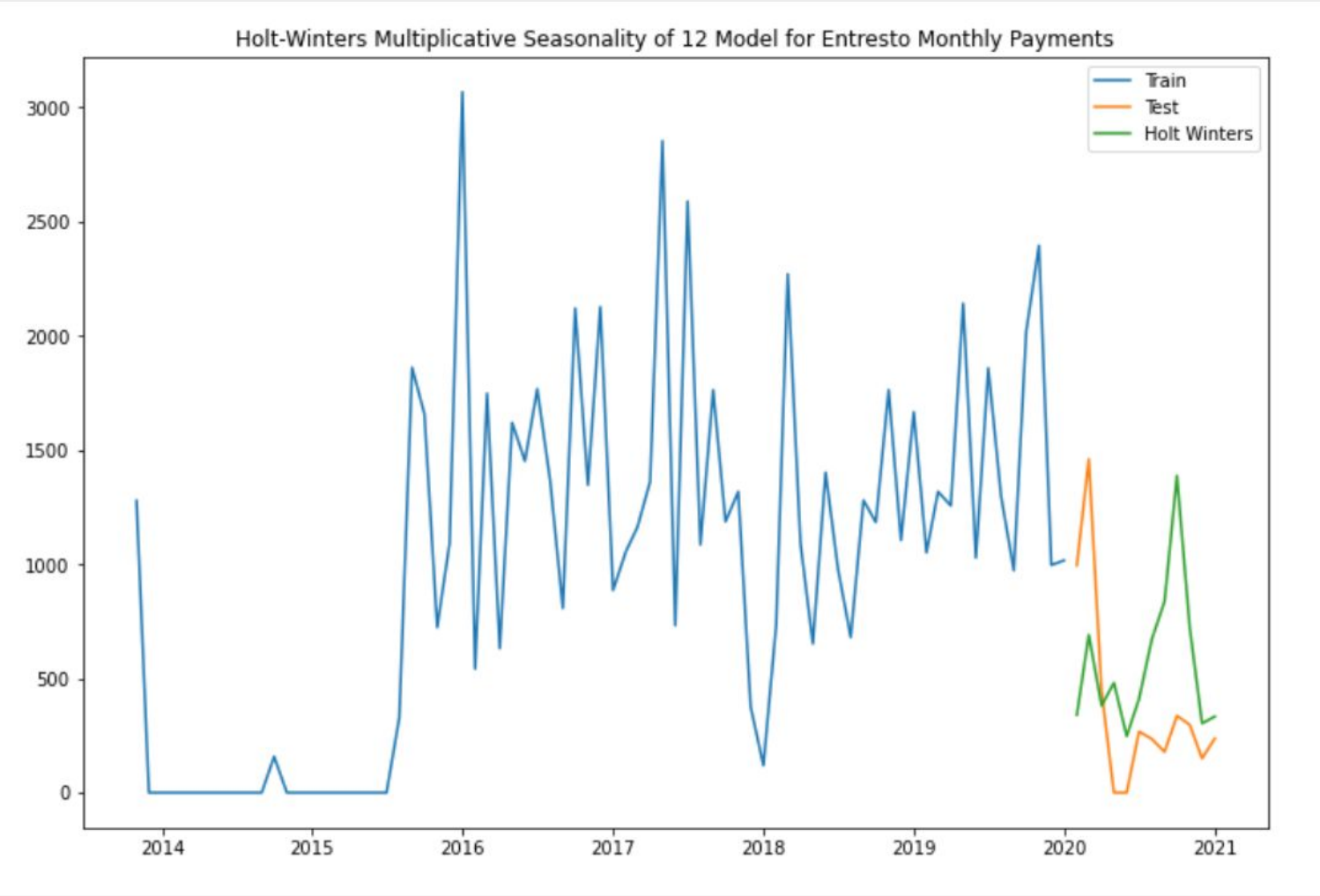

SARIMA Seasonality of 12 Model for Invokana Monthly Payments
Train
Test
SARIMA_12_auto
5000
4000
3000
2000
1000
0
2014
2015
2016
2017
2018
2019
2020

	RMSE
MODEL	INVOKANA
Baseline (Mean)	1,052.47
Baseline (Shift) 4	1,153.62
Baseline (Shift) 6	219.33
Baseline (Shift) 12	380.68
Baseline (Shift) 18	663.74
Simple Exp Smoothing	1,007.17
HW M 4	296.01
HW M 6	321.13
HW M 12	1,532.41
HW M 18	383.47
HW A 4	429.92
HW A 6	268.03
HW A 12	1,331.02
HW A 18	375.33
SARIMA 4	404.16
SARIMA 6	404.16
SARIMA 12	160.27
SARIMA 18	404.16

The SARIMA model with period 12 was the strongest model for Invokana.

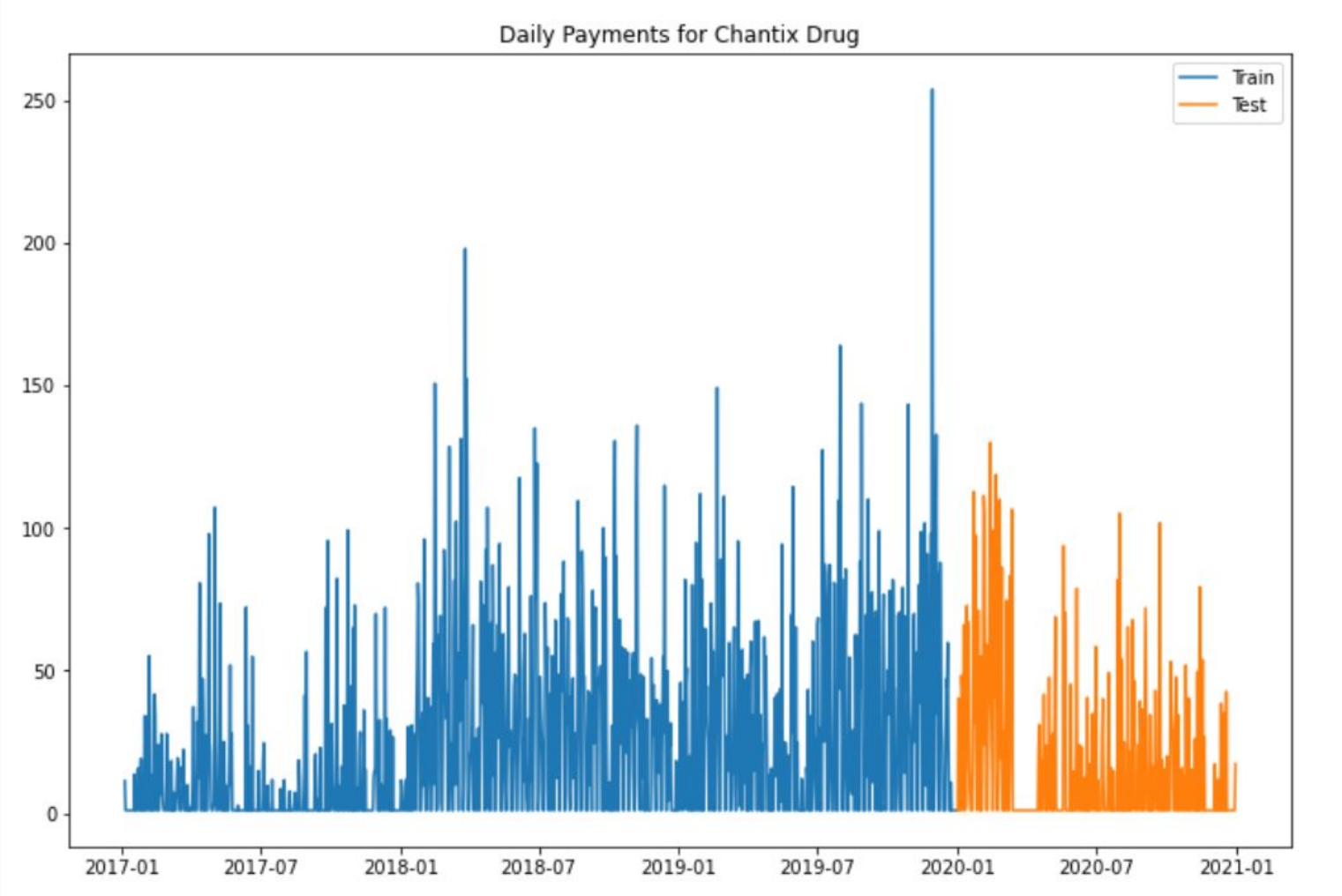

Daily Payments for Entresto Drug
Train
Test
1400
1200
1000
800
600
400
200
0
2014
2015
2016
2017
2018
2019
2020
2021

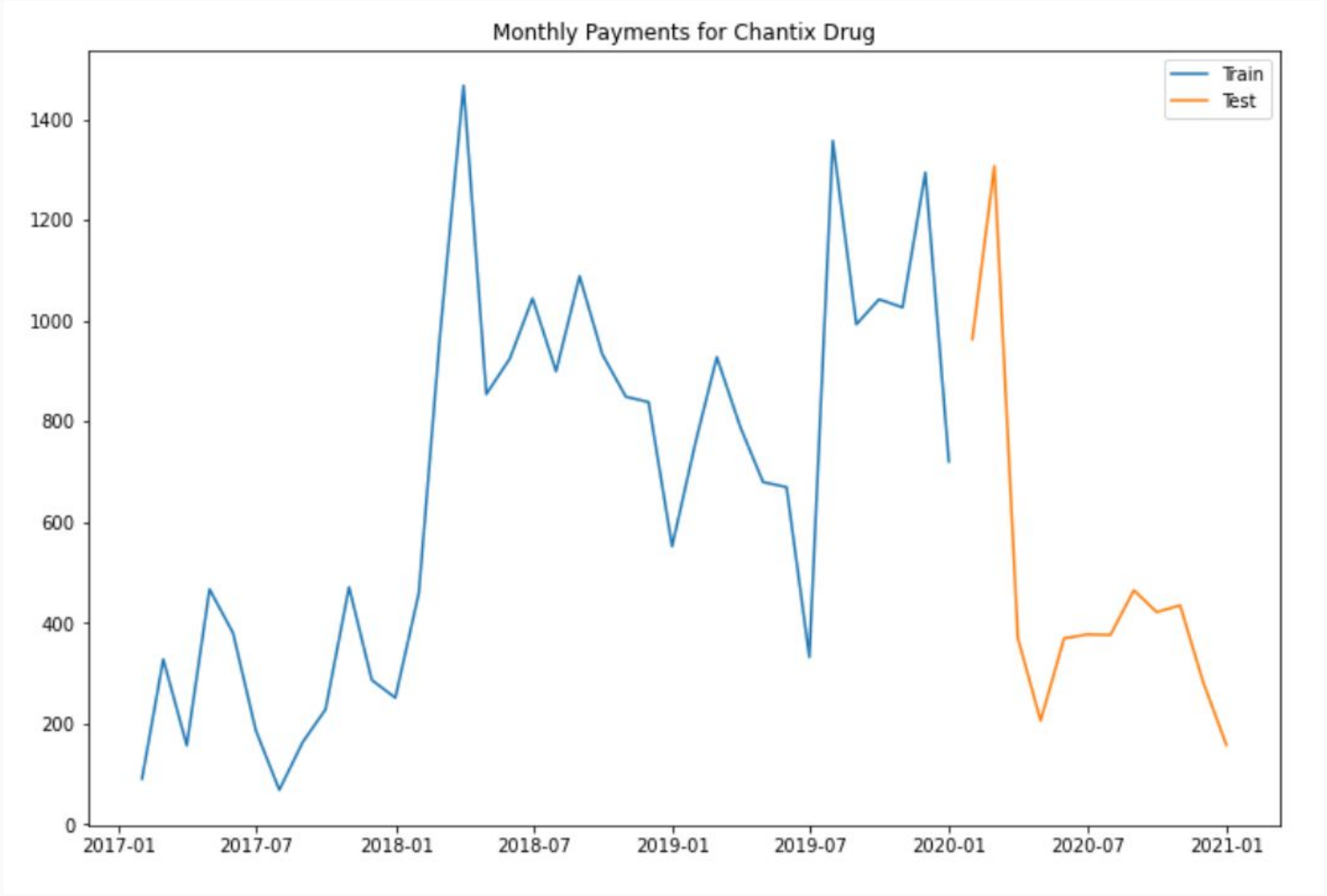

Monthly Payments for Entresto Drug
Train
Test
3000
2500
2000
1500
1000
500
0
2014
2015
2016
2017
2018
2019
2020
2021

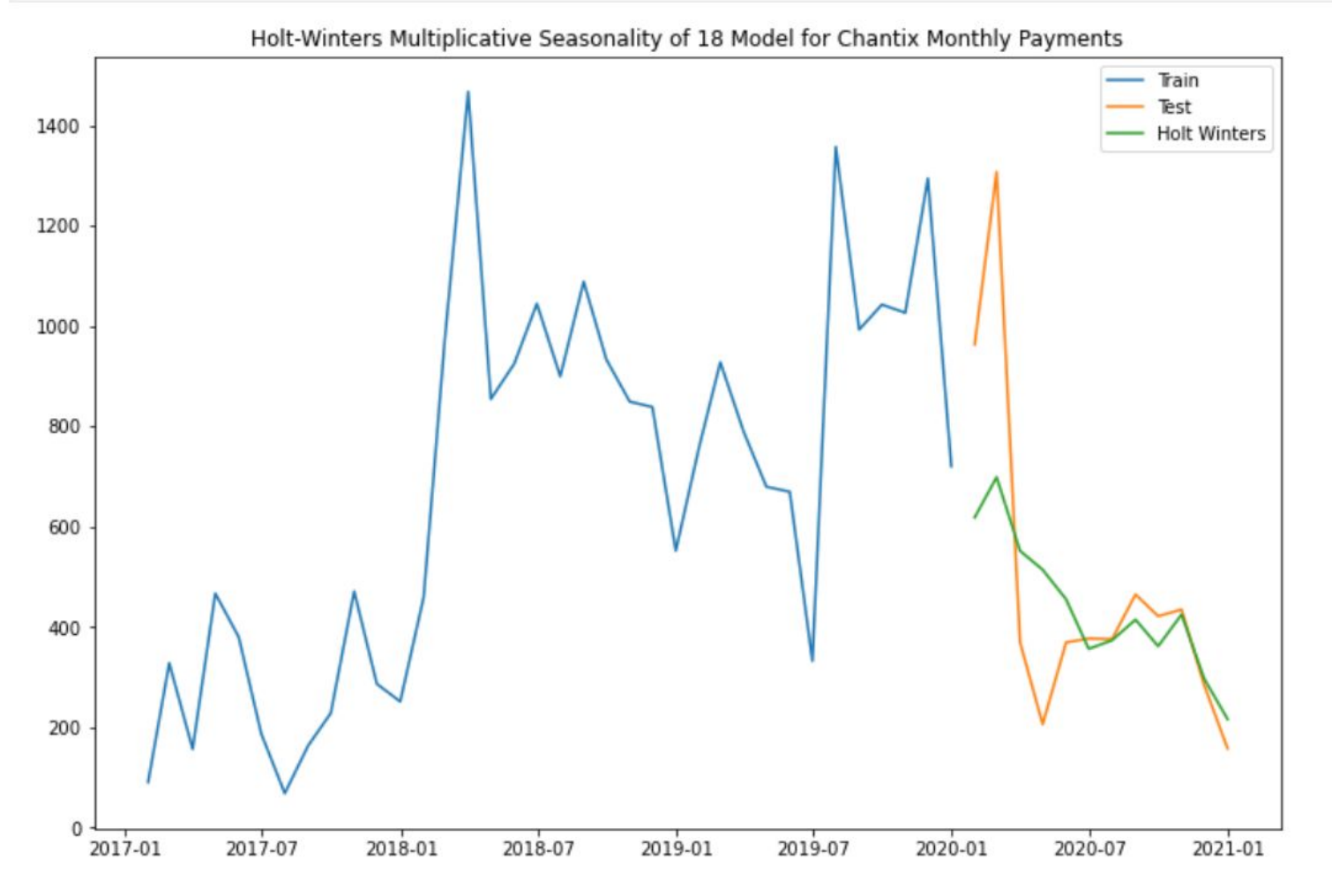

Holt-Winters Multiplicative Seasonality of 12 Model for Entresto Monthly Payments
Train
Test
Holt Winters
3000
2500
2000
1500
1000
500
0
2014
2015
2016
2017
2018
2019
2020
2021

	RMSE
MODEL	ENTRESTO
Baseline (Mean)	731.58
Baseline (Shift) 4	554.98
Baseline (Shift) 6	933.34
Baseline (Shift) 12	1,263.49
Baseline (Shift) 18	1,174.89
Simple Exp Smoothing	965.23
HW M 4	956.72
HW M 6	1,137.92
HW M 12	522.82
HW M 18	
HW A 4	1,280.87
HW A 6	1,275.59
HW A 12	1,269.44
HW A 18	1,550.18
SARIMA 4	1,136.19
SARIMA 6	1,107.07
SARIMA 12	1,107.07
SARIMA 18	1,107.07

The Holt-Winters Multiplicative model with a period of 12 was the strongest model for Entresto.

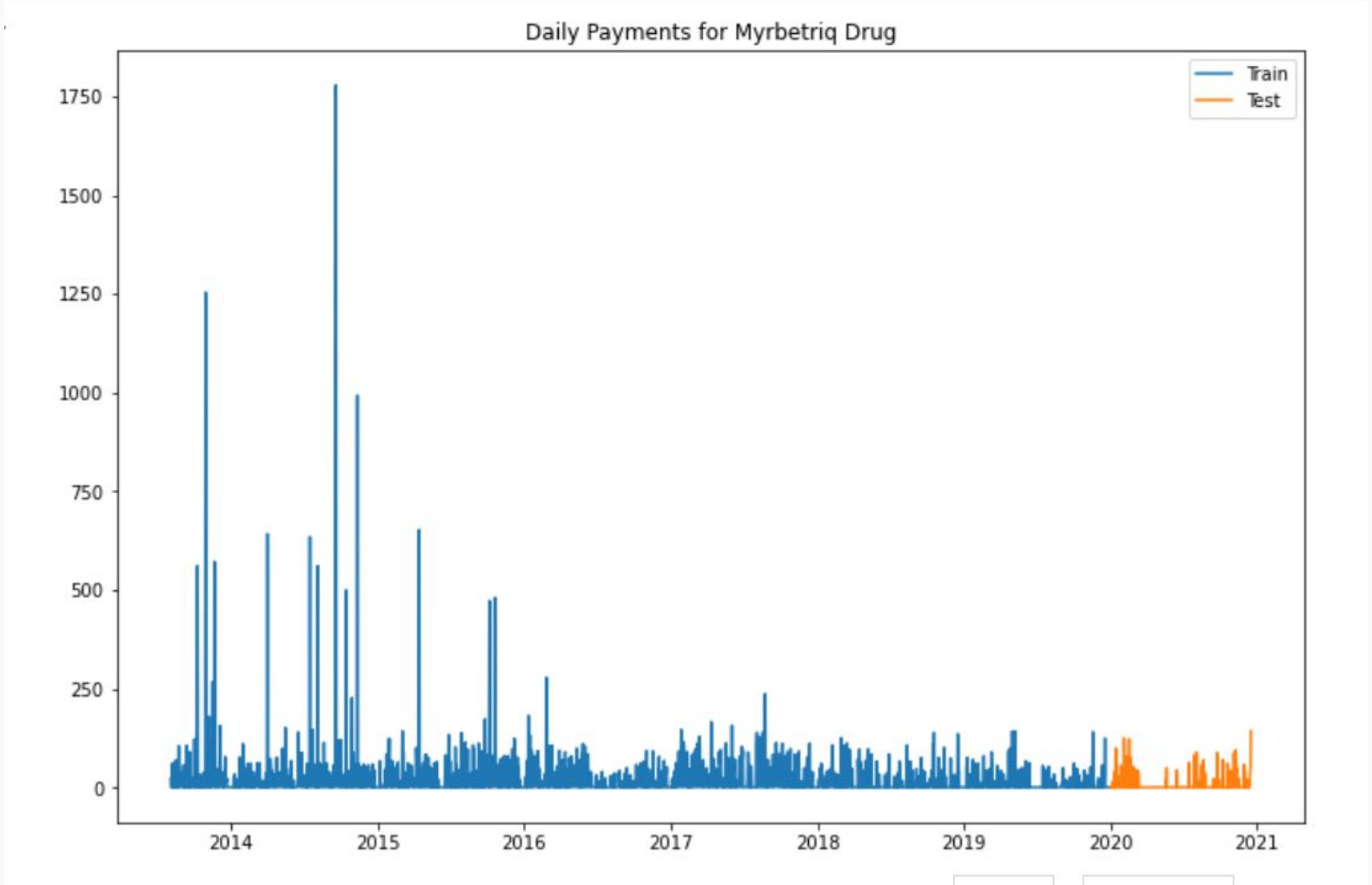

Daily Payments for Chantix Drug
Train
Test
250
200
150
100
50
0
2017-01
2017-07
2018-01
2018-07
2019-01
2019-07
2020-01
2020-07
2021-01

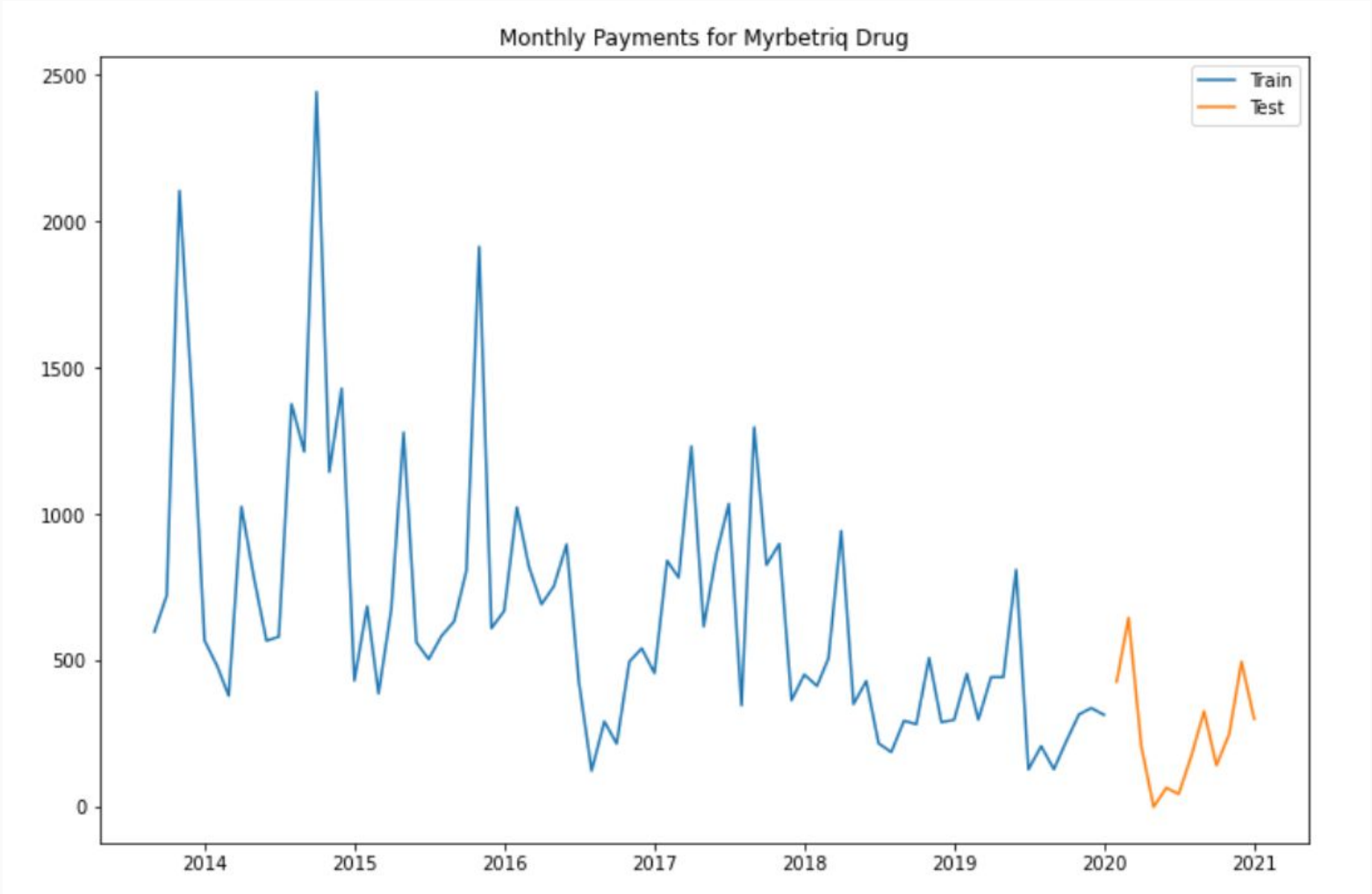

Monthly Payments for Chantix Drug
Train
Test
1400
1200
1000
800
600
400
200
0
2017-01
2017-07
2018-01
2018-07
2019-01
2019-07
2020-01
2020-07
2021-01

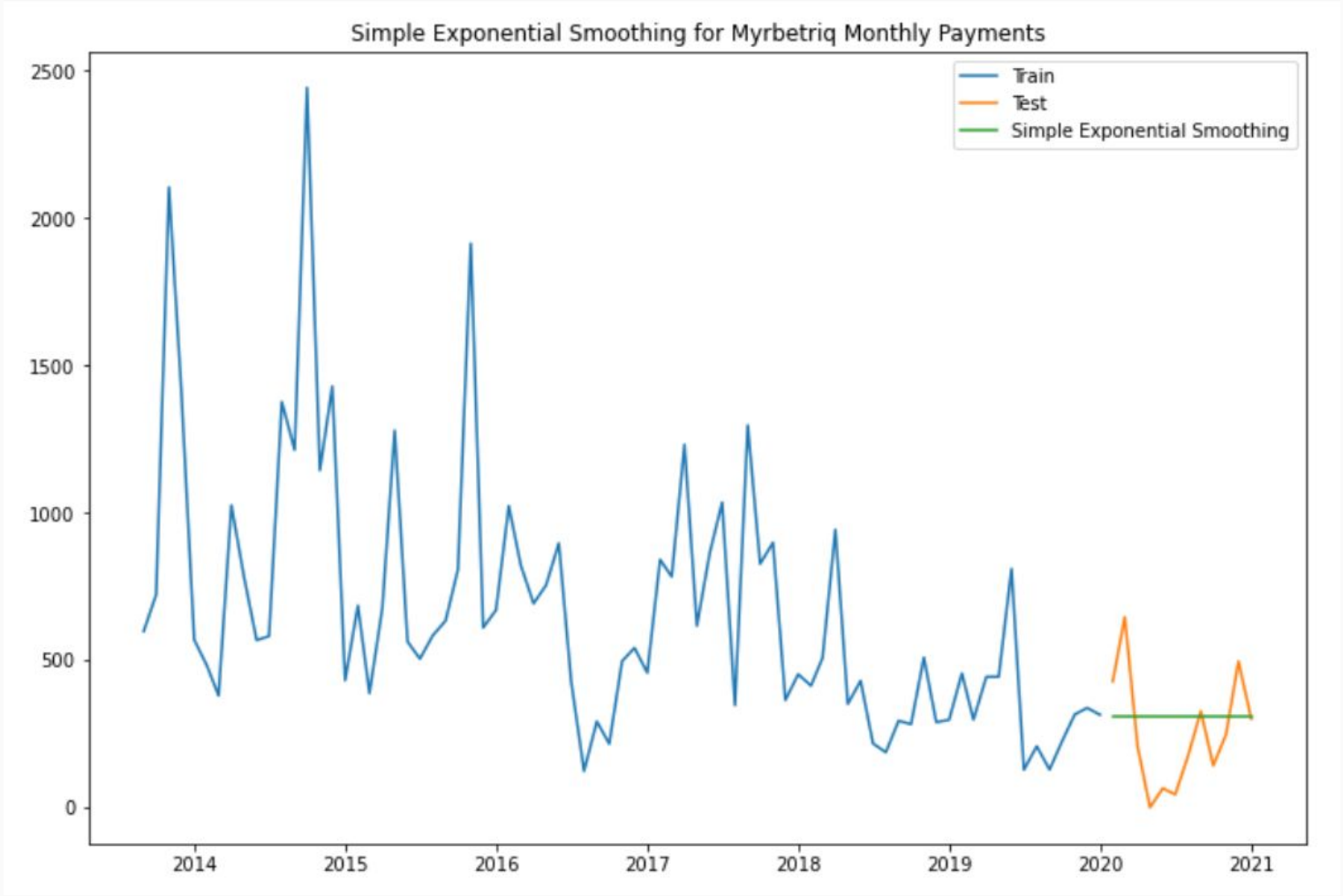

Holt-Winters Multiplicative Seasonality of 18 Model for Chantix Monthly Payments
Train
Test
Holt Winters
1400
1200
1000
800
600
400
200
0
2017-01
2017-07
2018-01
2018-07
2019-01
2019-07
2020-01
2020-07
2021-01

	RMSE
MODEL	CHANTIX
Baseline (Mean)	374.95
Baseline (Shift) 4	438.91
Baseline (Shift) 6	514.91
Baseline (Shift) 12	576.67
Baseline (Shift) 18	383.84
Simple Exp Smoothing	557.02
HW M 4	1,005.91
HW M 6	1,124.23
HW M 12	1,773.16
HW M 18	229.99
HW A 4	722.13
HW A 6	845.00
HW A 12	884.01
HW A 18	309.19
SARIMA 4	555.96
SARIMA 6	555.96
SARIMA 12	555.96
SARIMA 18	555.96

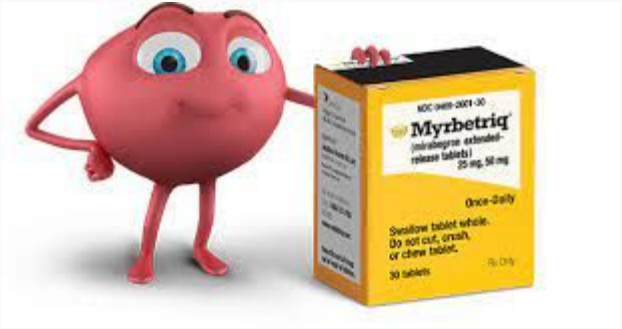
A Holt-Winters Multiplicative model with period 18 was the strongest for Chantix.


Daily Payments for Myrbetriq Drug
Train
Test
1750
1500
1250
1000
750
500
250
0
2014
2015
2016
2017
2018
2019
2020
2021


Monthly Payments for Myrbetriq Drug
Train
Test
2500
2000
1500
1000
500
0
2014
2015
2016
2017
2018
2019
2020
2021


Simple Exponential Smoothing for Myrbetriq Monthly Payments
Train
Test
Simple Exponential Smoothing
2500
2000
1500
1000
500
0
2014
2015
2016
2017
2018
2019
2020
2021

	RMSE
MODEL	MYRBETRIQ
Baseline (Mean)	454.36
Baseline (Shift) 4	263.22
Baseline (Shift) 6	305.73
Baseline (Shift) 12	290.35
Baseline (Shift) 18	273.35
Simple Exp Smoothing	192.98
HW M 4	193.12
HW M 6	234.54
HW M 12	232.89
HW M 18	286.01
HW A 4	194.65
HW A 6	253.67
HW A 12	267.53
HW A 18	323.03
SARIMA 4	231.72
SARIMA 6	219.88
SARIMA 12	219.88
SARIMA 18	219.88



The Simple Exponential Smoothing model was the strongest model for Myrbetriq.

Conclusion

- It was tough to create good predictions using time series models for 2020 CMS drug/biologic physician payments from historical 2013-2019 payment data.

For Further Consideration…

- Try a supervised learning model instead
- Also try some ensemble models
- How will 2021 data change these models?
- Can we apply Benford's Law to physician payments?
- How do Cincinnati CMS physician payments compare with those in other Ohio cities? With Florida cities?

Health = Wealth

Resources

CMS Data: https://www.cms.gov/OpenPayments/Data/Dataset-Downloads

Farxiga study:
https://www.astrazeneca.com/media-centre/press-releases/2021/update-on-farxiga-covid-19-dare-19-phase-iii-trial.html

Humira article:
https://www.statnews.com/2022/01/31/rheumatoid-arthritis-humira-immune-supression-covid19/

Thanks!

Questions?