**2019-2020 PRODUCT ANALYSIS**

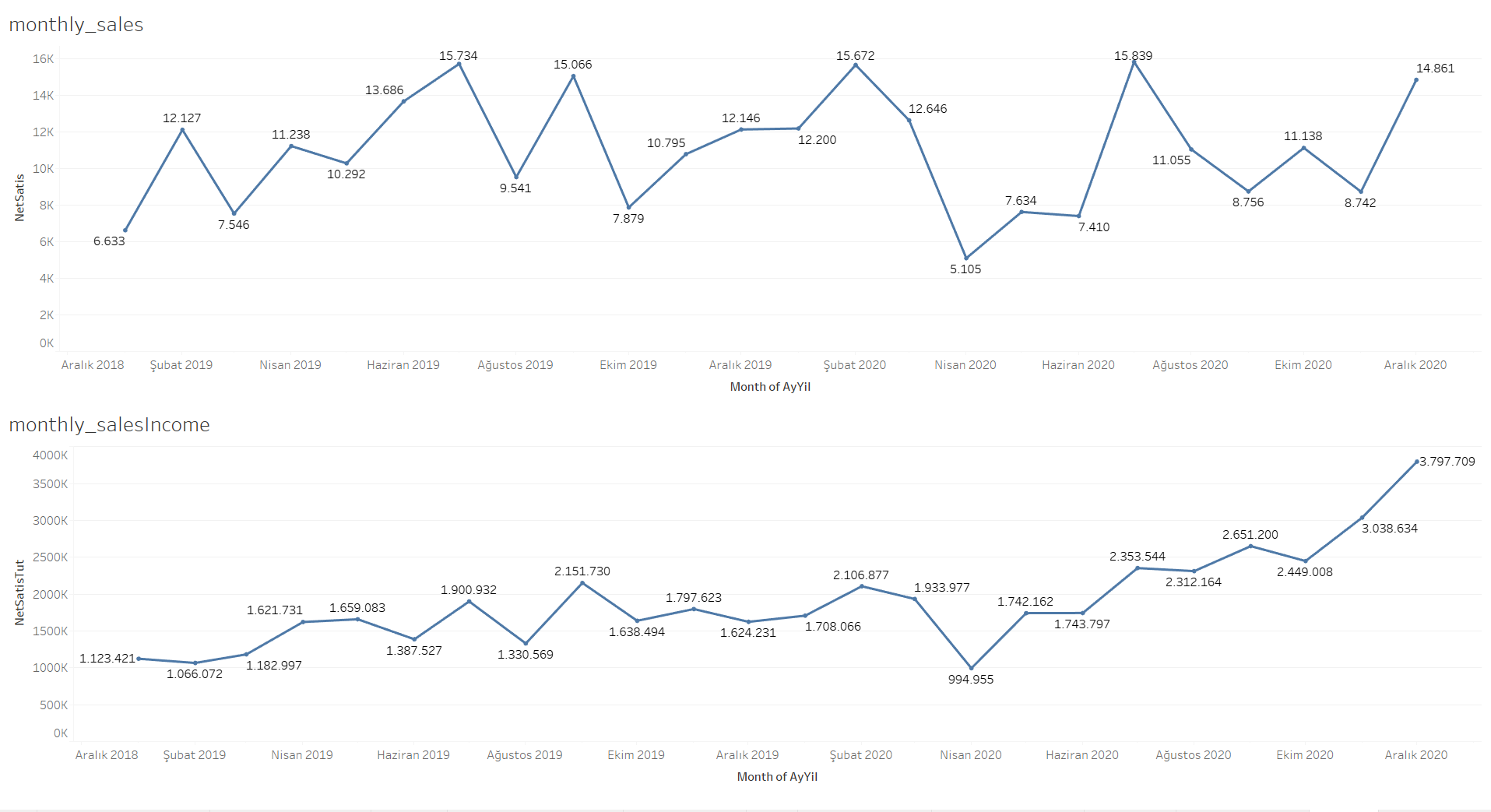
**&**

**WT Group Forecasting**

**Ugur Oguz**

**Understanding Dataset**

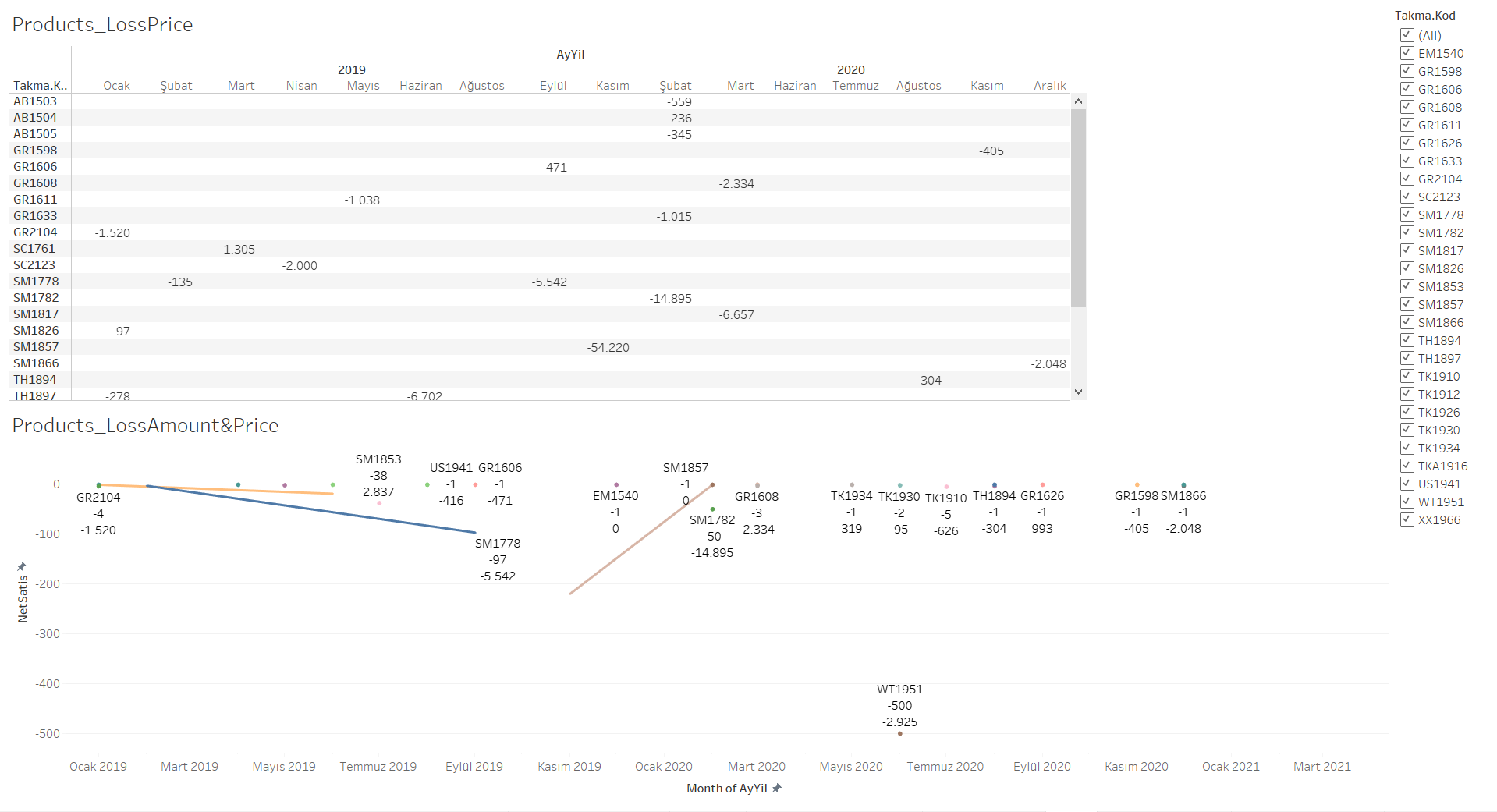
* **Malzeme Turu:** Material type.
* **Takma.Kod:** Product ID, first 2 character represents product group.
* **Kod:** Product group.
* **NetSatis:** Sales amount.
* **NetSatisTut:** Sales price, different than unit price. As in business a products can be sold different prices or in time price can vary.
* **Average Price:** average price based on total sold products. Sum (NetSatisTut)/Sum (NetSatis).

**Monthly Sales & Income**

**Fig-1**

Analyzing monthly sales. The company has increasing income. Starting from November 2020 Income rate is highly increased while there isn’t any drastically change in sales amount.

Based on this graph we can assume that company has changed its products to more expensive and their customers target has change. As starting from May 2020 to September 2020 we can see that sales amount have decreased while income started to shift from stable to increase. Another option can be products are affected by stock market prices considering there isn’t any drastically change sales amount.

**PRODUCT LOSS**

**Fig-2**

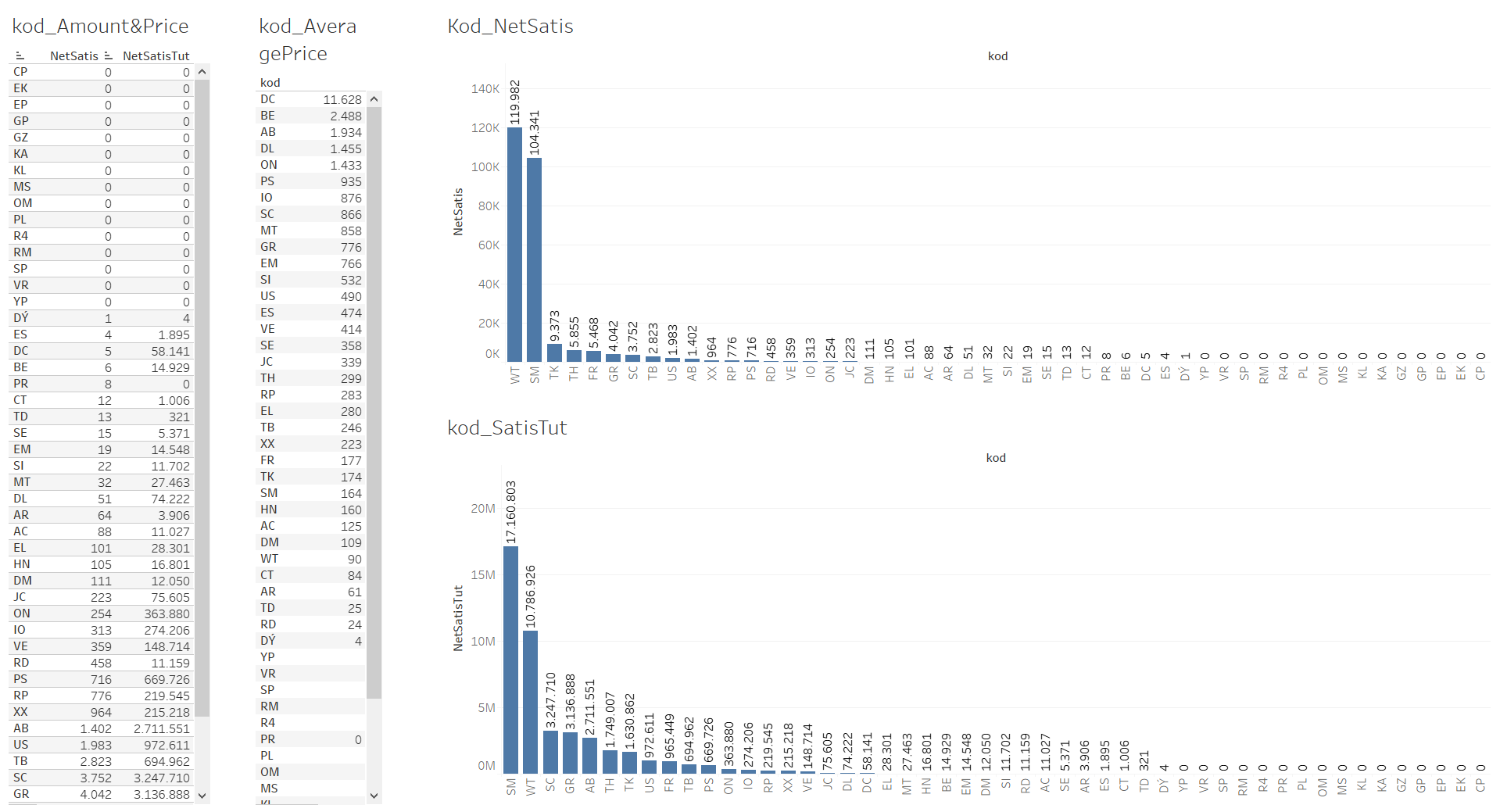
Among the products groups, 27 group’s products have been returned.

As we look into this graph we can see that in SM1853 product 38 items have returned and company has profited. Assuming product is affected by stock exchange. Another possibility is that refund policy.

In November 2019, 220 SM1857 products have been returned causing -54.220 losses which is the biggest loss among the products.

WT1951 product has the highest return amount (500).

**Kod**

****

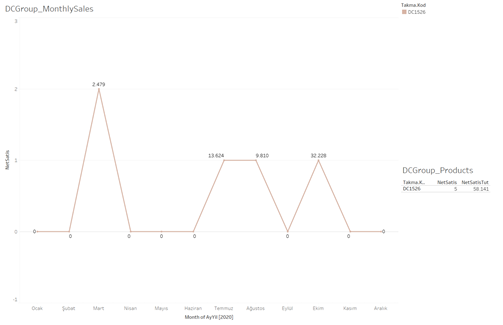
**Fig-3**

WT group has the most sold items while SM group comes second. We can see that the clients are more focused in these product groups. When we evaluate the income SM has the highest income while it is second in highest sales.

DC group has the highest average price. 8 PR group products have been sold but income is 0 we will look into these groups in more detail.

When we look into kod\_amount&price “CP, EK, EP, GP, GZ, KA, KL, MS, OM, PL, R4, RM, SP, VR, YP” these product group sales amount and income is 0. There is chance that the sales amount might me equal to return amount so we will see them more detailed in monthly time base.

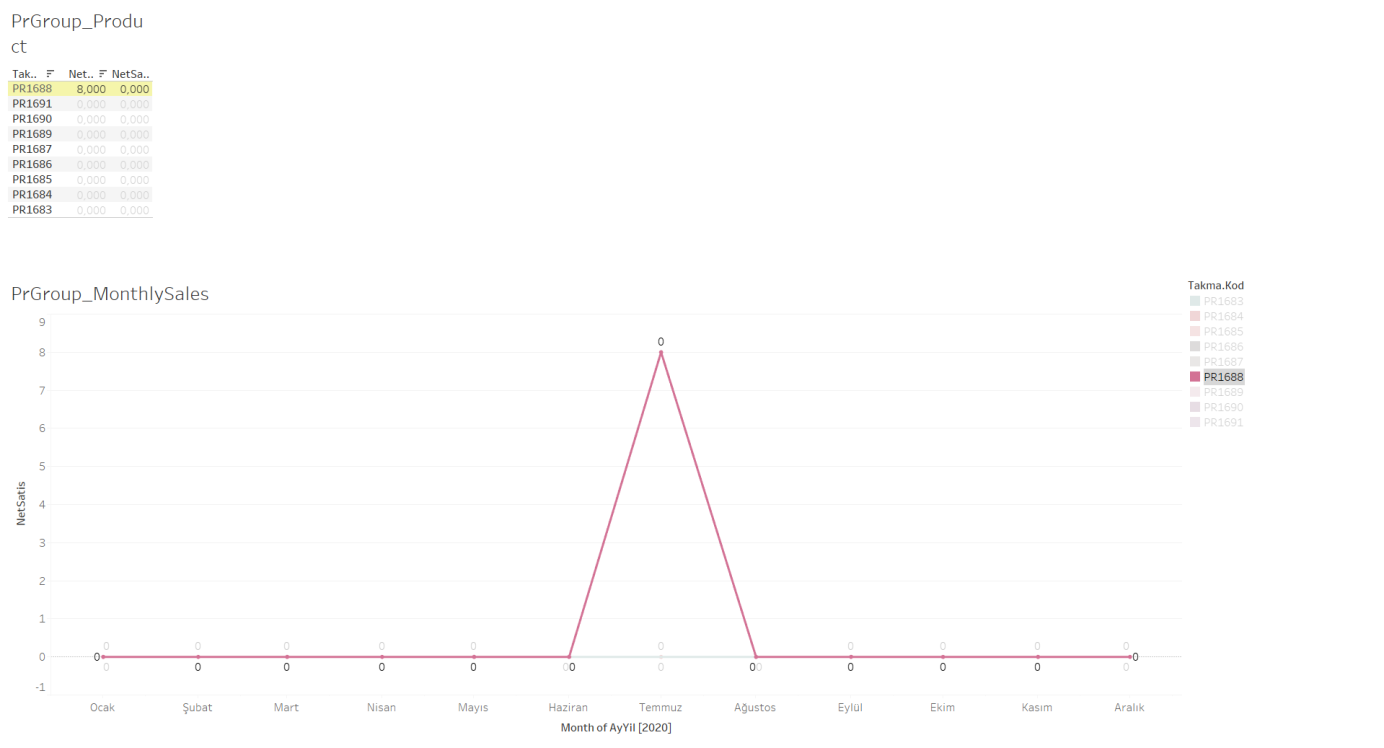
**DC Group**

****

**Fig-4**

In DC group we can see that the product group has only one product that came out in 2020 and 5 items have been sold.

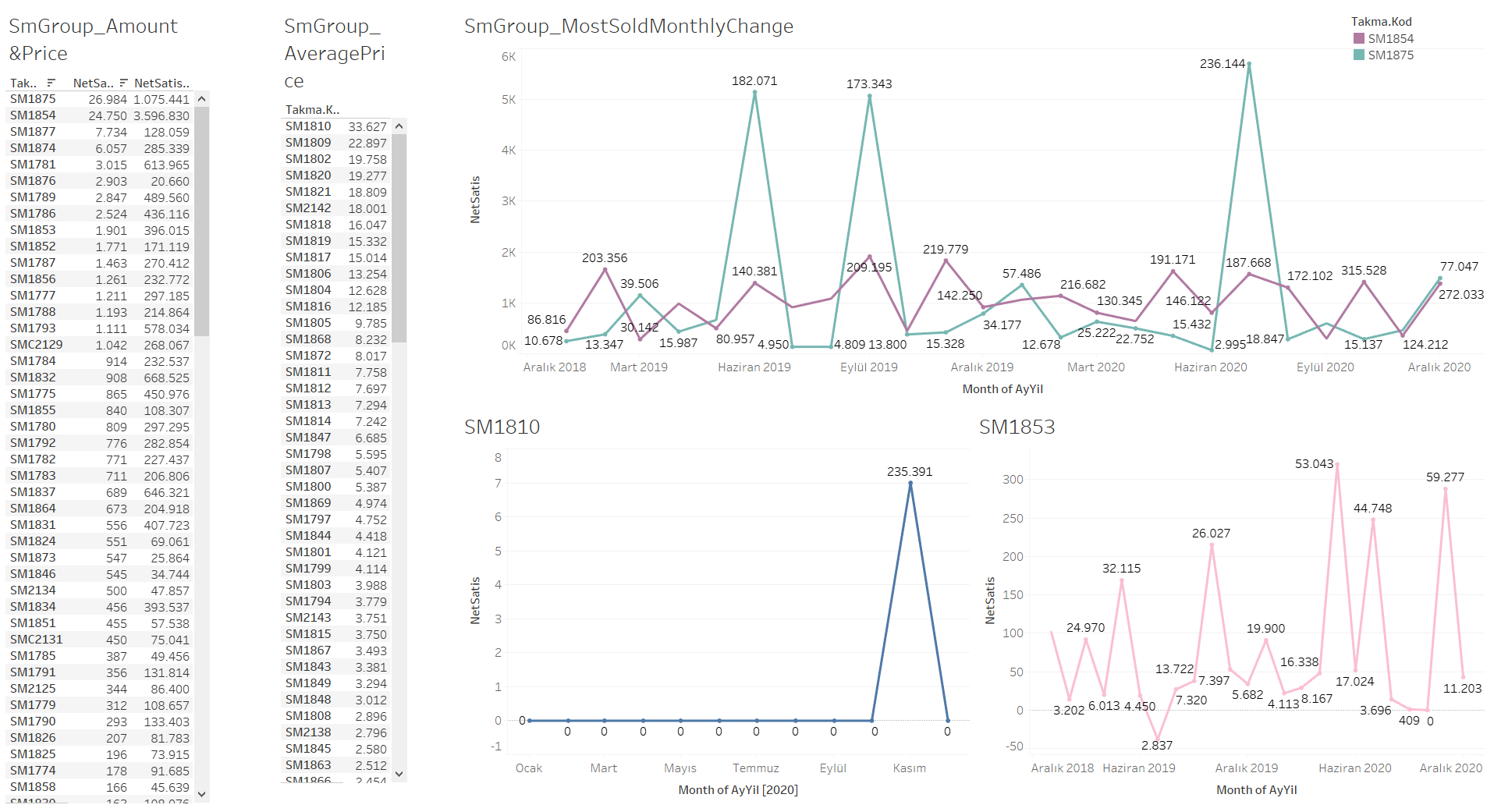
**PR Group**

****

**Fig-5**

When we looked into PR Group only PR1688 product has been sold and there is no income. Based on the graph we are assuming that new group has been introduced in 2020 and 8 products have been given as sample/gift but there wasn’t any demand.

**SM Group**

****

**Fig-6**

SM group is second most sold item group and highest group that bring income to the company.

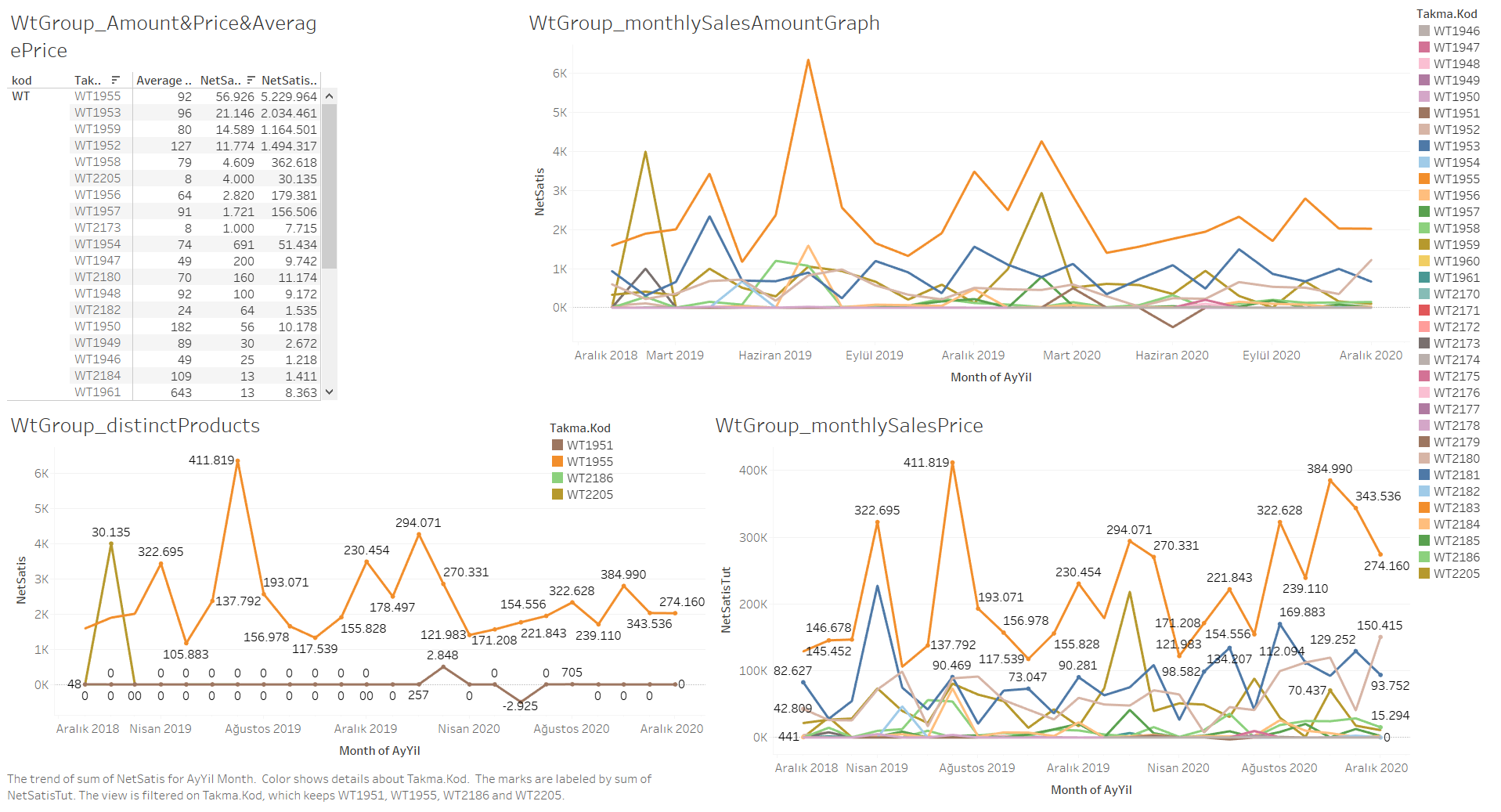
SM1875 have the highest sell rate. From Product Loss section SM1875 had highest loss rate and that occurred in November 2020 and 220 products have been returned. When we check the graph we can see that there are high pikes, product has been batch bought or during June there is high demand for this product.

SM1854 have the highest income rate, looking from the graph while it has more stable demand during season it is the second most sold product.

SM1810 has highest average income rate. On November 2020, 7 products have been sold.

From Product Loss SM1853 is the product that company made profit from returned item, we have graphed to see more closely how product sales and income changed in time.

**WT Group**

****

**Fig-7**

Most sold product group. There are 4 distinct products that we would to look closely.

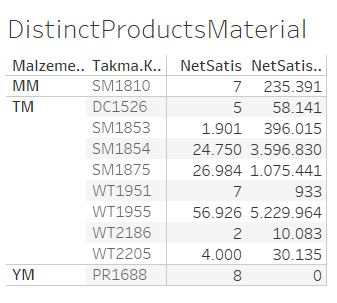
WT1955 have the highest sale rate and income rate as we can see in the graph this product stand out from rest of the products. Second highest product “WT1953” has approximately half of the sale rate and income rate of WT1955.

WT2205 in February 2019 there was a high demand for this product but the product has been removed from 2020. Assuming product material became expensive or product wasn’t profitable than the similar products that are sold.

WT1951 is the only product in WT group that have been returned, as we checked in Product loss the return amount is 500.

WT2186 have the highest average income rate (5.042), the company have sold 2 products in June 2019.

**Distinct Products Material**

****

**Fig-8**

As conclusion to our research, the distinct products that have impact on company sales. We can see the material groups.

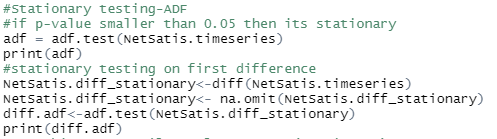
**FORECAST**

In this section we will talk about WT group 2019 & 2020 forecasting. Finding Optimal model for forecasting 2021.

* Testing Models
  + Stationary Testing by using ADF (Augmented Dickey-Fuller) and UR.DF (Dickey-Fuller unit root test)
  + Finding Lags & Fitting ARIMA models
  + Forecasting
  + Model Graphs
* Final Model
  + Stationary Testing by using ADF (Augmented Dickey-Fuller)
  + Finding Lags & Fitting ARIMA models
  + Residual diagnostic
  + Forecasting & Model Graphs
* All Model Results

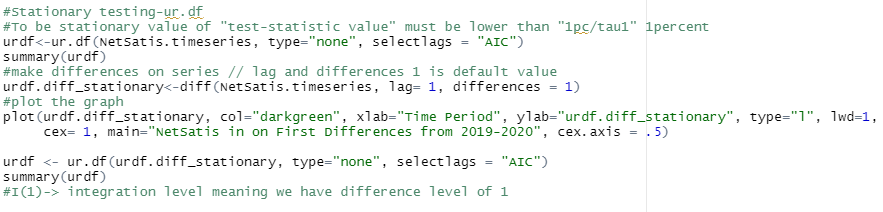
**TESTING MODELS**

**ADF-Stationary testing**



In adf-test we have stationary data at I(4).

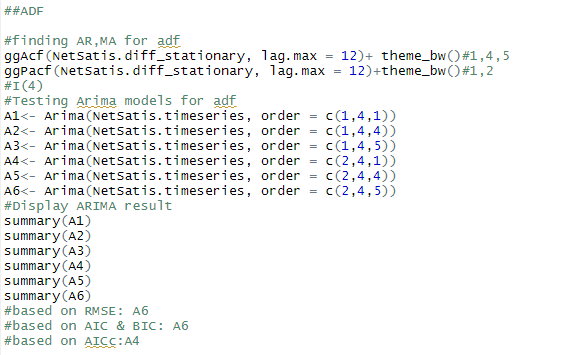
**Ur.df-Stationary testing**



In ur.df testing we have stationary data I(1).

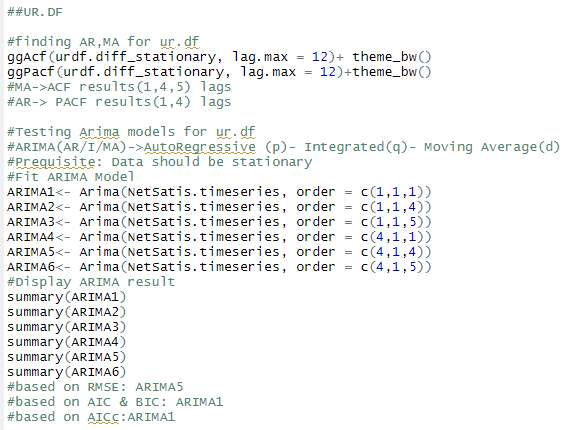
**Fitting In ARIMA models**

**ADF- Finding Lags & Model Fitting**



Based on result we will be using Arima(2,4,5) for RMSE and AIC/BIC, Arima(2,4,1) based on AICc.

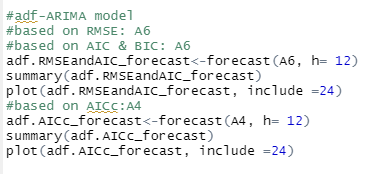
**URDF- Finding Lags & Model Fitting**



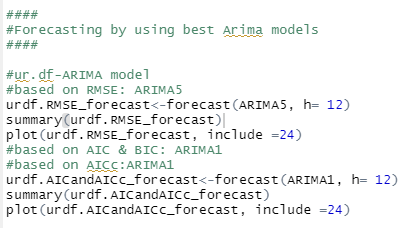
Based on result we will be using Arima(4,1,4) for RMSE, Arima(1,1,1) for AIC/BIC and AICc.

**Forecasting**

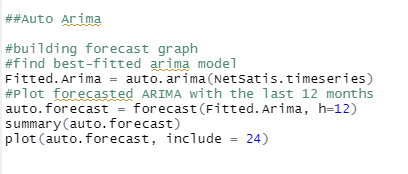
**ADF**

****

**URDF**

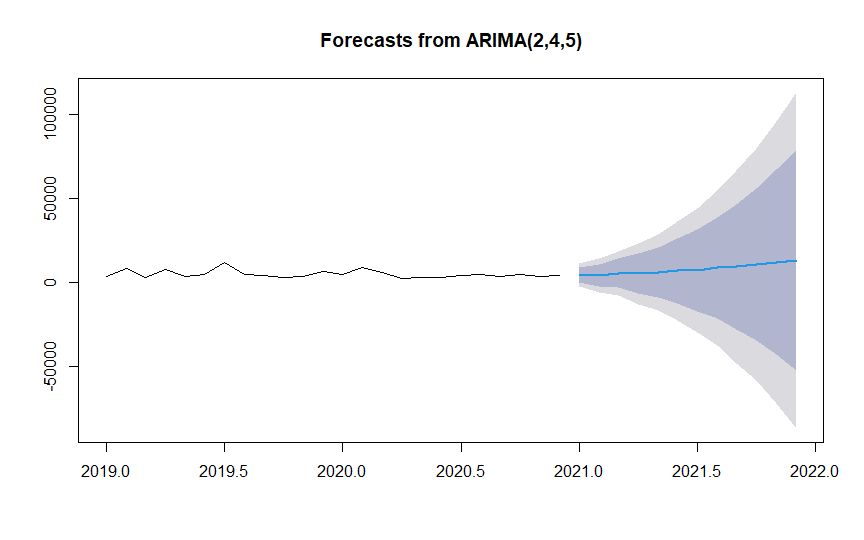
****

**Auto Arima**

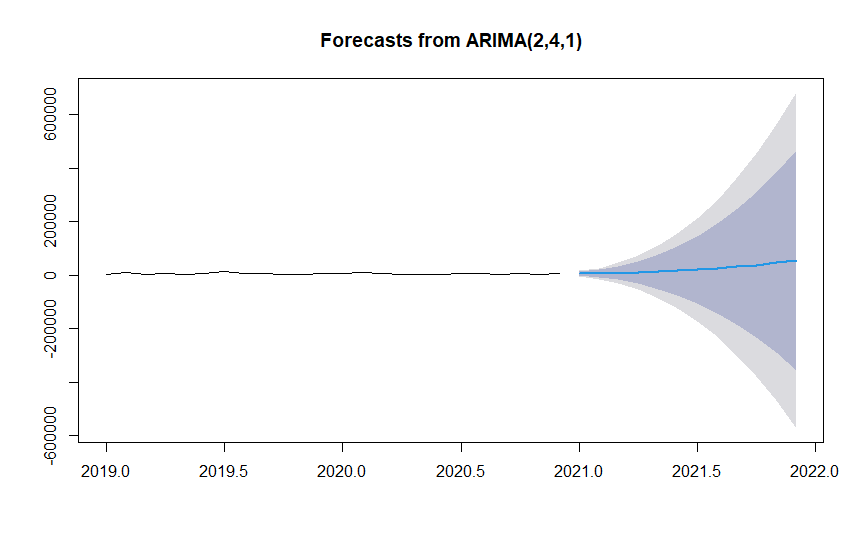


**Model Graph**

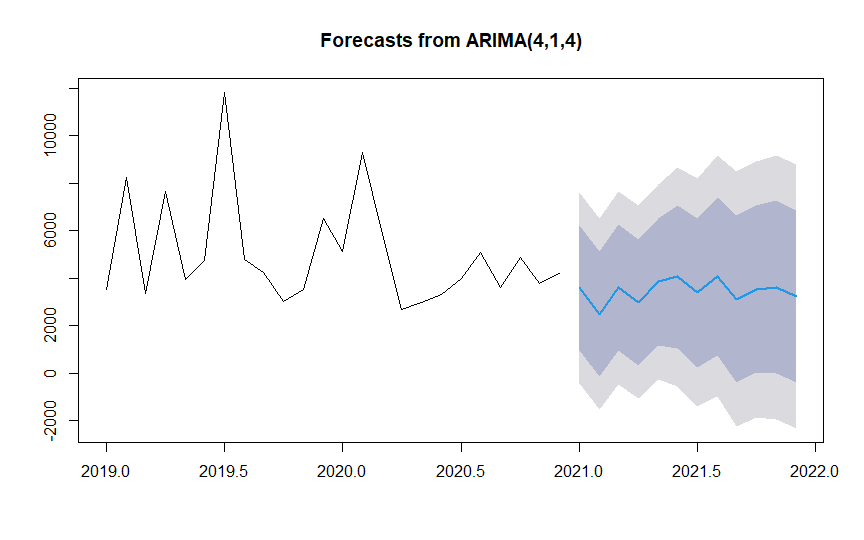
**ADF- RMSE and AIC/BIC forecast**



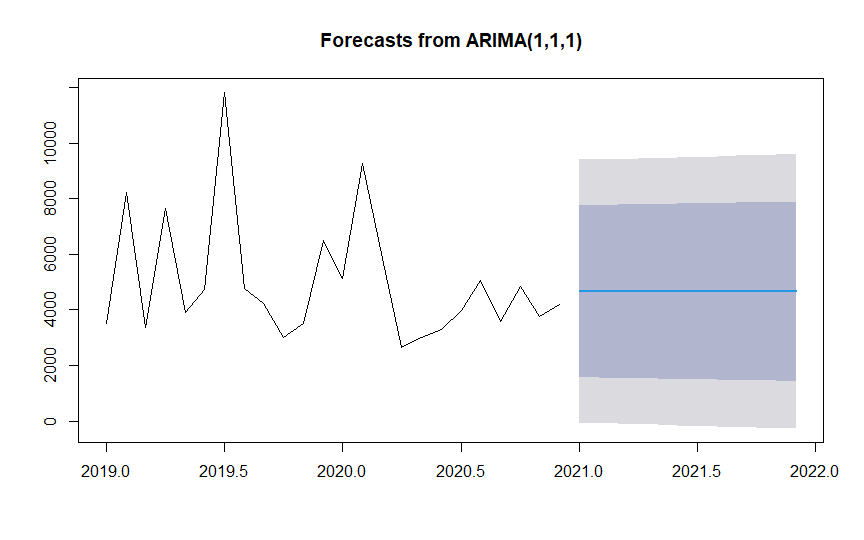
**ADF- AICc forecast**



**URDF- RMSE forecast**



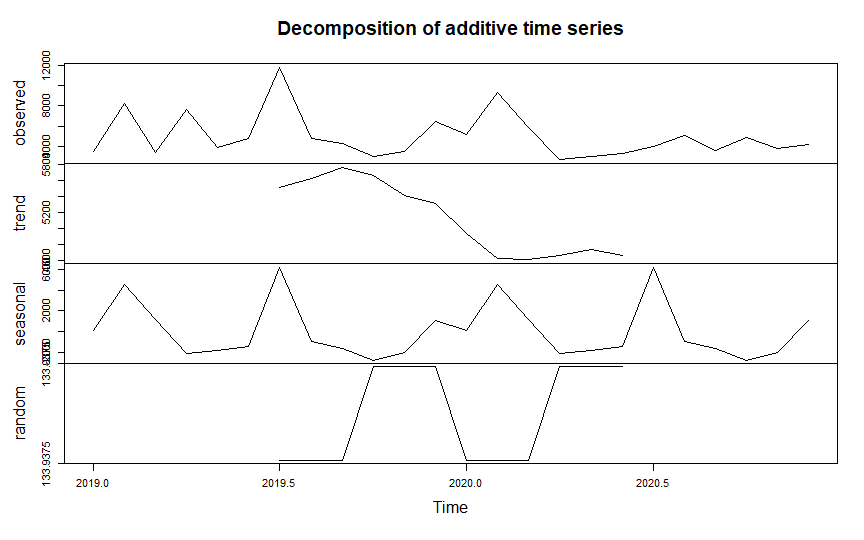
**URDF- AIC/BIC and AICc forecast**



**Final Model**

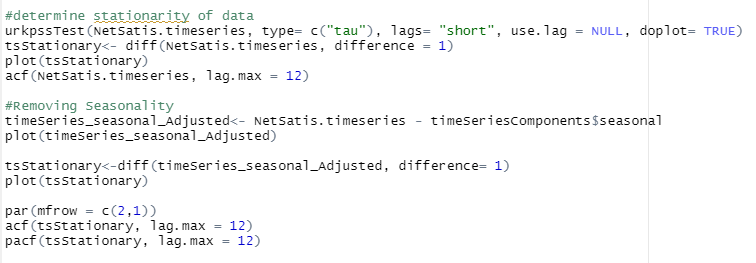
Looking into time series data

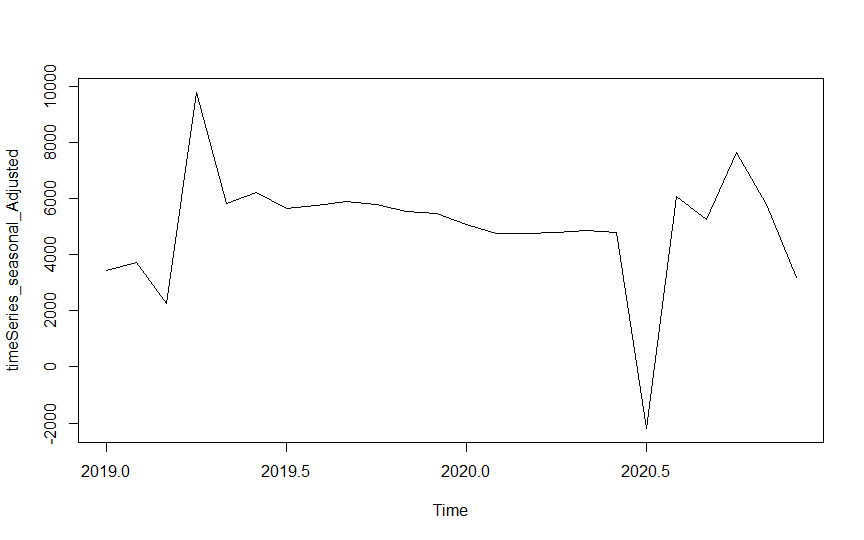
C:\Users\ugur\Desktop\GokceAbla\Analytic\PredictionImg\TimeseriesComponentsCode.png



As we look into components of time series we can see there is seasonality. Starting from 6th month of 2019 till 2nd month of 2020 there is decreasing trend.

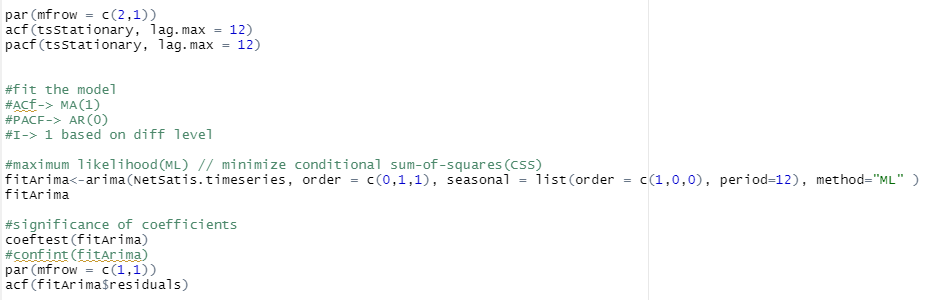
**Removing Seasonality**

****

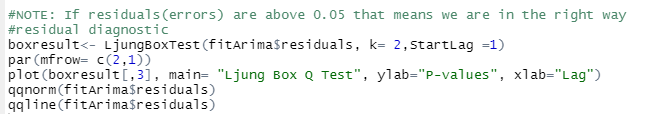
****

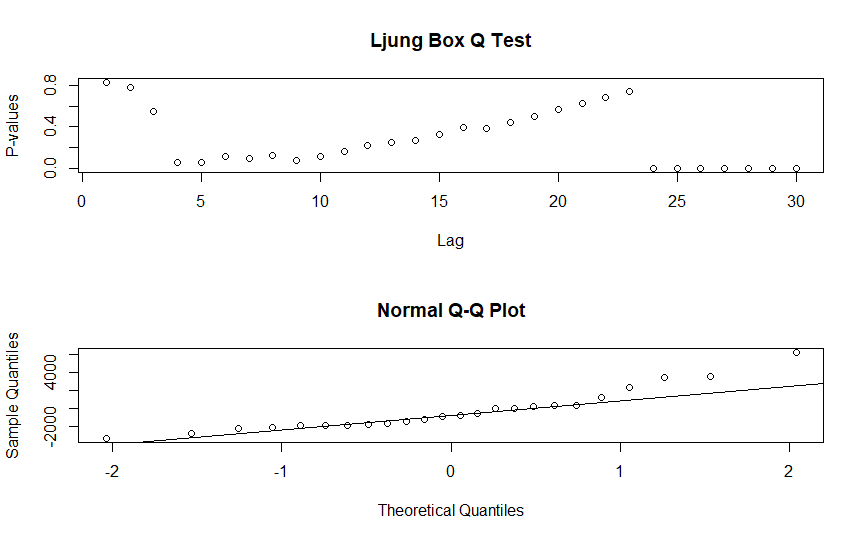
After removing seasonality we have more stable data that we can work on.

**Model Fitting**



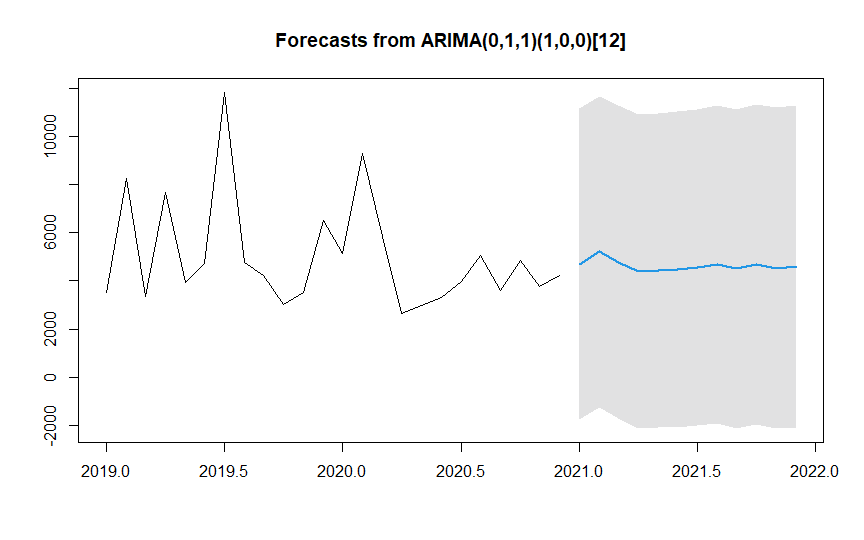
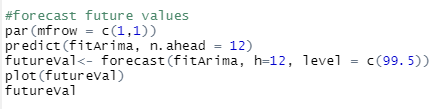
**Residual Diagnostic**

****

****

Based on residual diagnostic we have determined our model accuracy rate is 86%.

**Forecast**



**Final Results**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Month** | **Final Model** | **AutoForecast** | **urdf.RMSE\_forecas** | **urdf.AICandAICc\_forecast** |
| Jan 2021 | 4696 | 4999 | 3585,427 | 4681,987 |
| Feb 2021 | 5200 | 4999 | 2484,926 | 4662,419 |
| Mar 2021 | 4781 | 4999 | 3594,407 | 4663,225 |
| Apr 2021 | 4395 | 4999 | 2985,071 | 4663,192 |
| May 2021 | 4435 | 4999 | 3843,019 | 4663,193 |
| Jun 2021 | 4474 | 4999 | 4047,524 | 4663,193 |
| Jul 2021 | 4556 | 4999 | 3382,626 | 4663,193 |
| Aug 2021 | 4688 | 4999 | 4073,155 | 4663,193 |
| Sep 2021 | 4508 | 4999 | 3116,027 | 4663,193 |
| Oct 2021 | 4662 | 4999 | 3528,326 | 4663,193 |
| Nov 2021 | 4528 | 4999 | 3616,543 | 4663,193 |
| Dec 2021 | 4583 | 4999 | 3220,388 | 4663,193 |

|  |  |  |
| --- | --- | --- |
| **Month** | **adf.RMSEandAIC\_forecast** | **adf.AICc\_forecast** |
| Jan 2021 | 4808,912 | 5224,242 |
| Feb 2021 | 4371,773 | 5407,572 |
| Mar 2021 | 5561,307 | 7486,987 |
| Apr 2021 | 5436,883 | 9259,656 |
| May 2021 | 6169,218 | 11742,721 |
| Jun 2021 | 7102,904 | 15494,496 |
| Jul 2021 | 7301,095 | 19289,218 |
| Aug 2021 | 8856,353 | 24500,671 |
| Sep 2021 | 9398,128 | 30516,989 |
| Oct 2021 | 10734,146 | 37334,284 |
| Nov 2021 | 12252,083 | 45636,669 |
| Dec 2021 | 13280,714 | 54789,07 |

Base on models forecast urdf.RMSE\_forecast model has similar close forecast to Final Model. Adf models are exponentially increasing which is unlikely to happen.