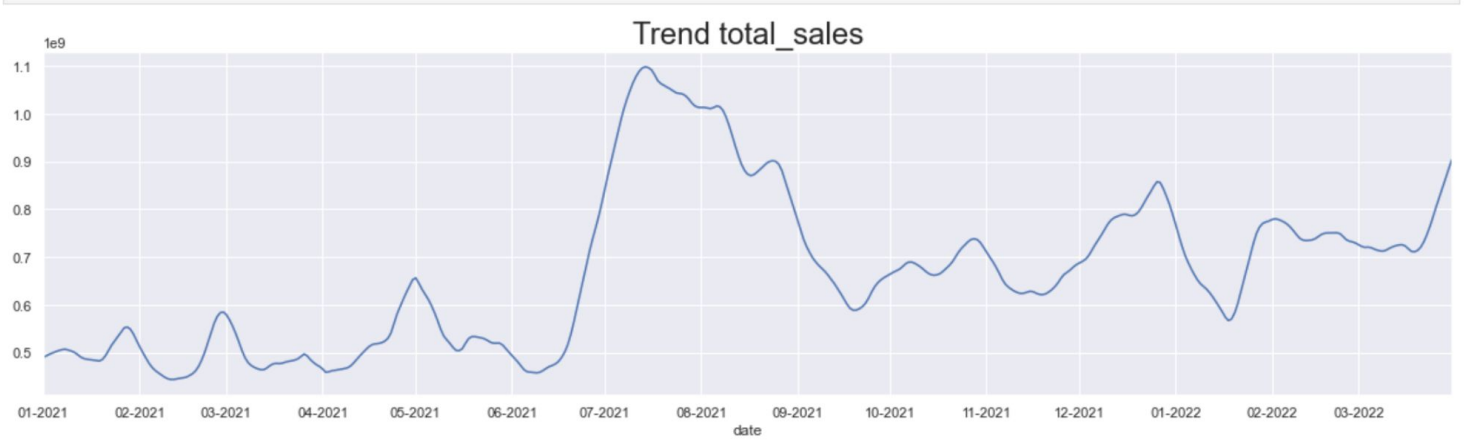
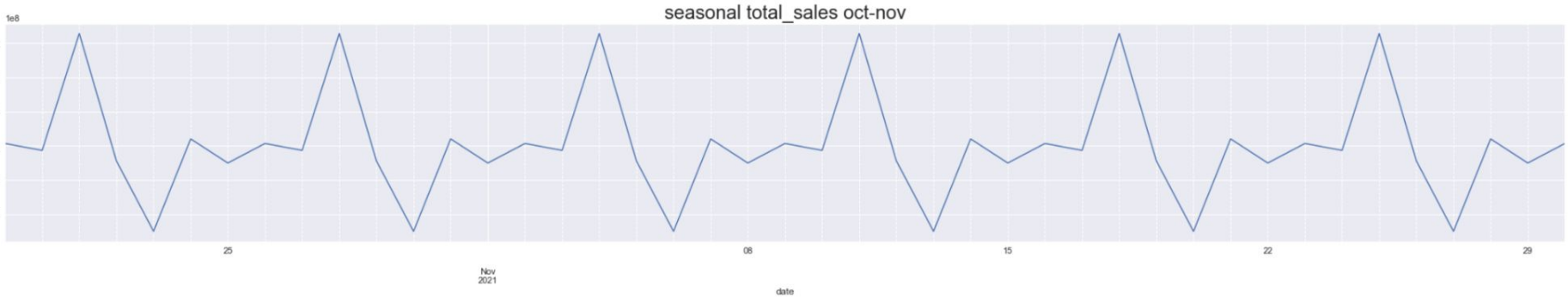
EDA (sales_total)

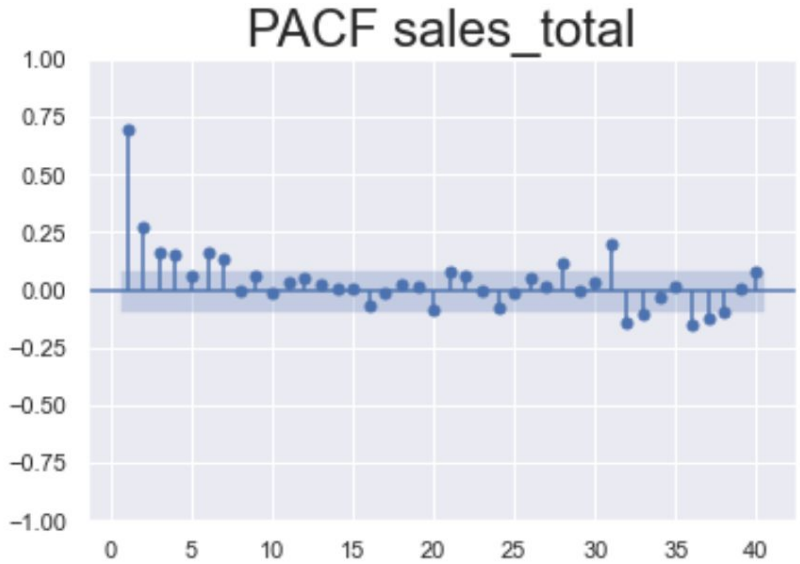
- Stationarity: stationary hingga 10% (kurang ideal)





- Weekly Seasonality
- Senin-Selasa-Rabu ga terlalu beda, sharp rise di kamis(peak), drop smp sabtu, rise lumayan di minggu

ACF signifikan up to 20an lag



PACF signifikan up to 7 lag

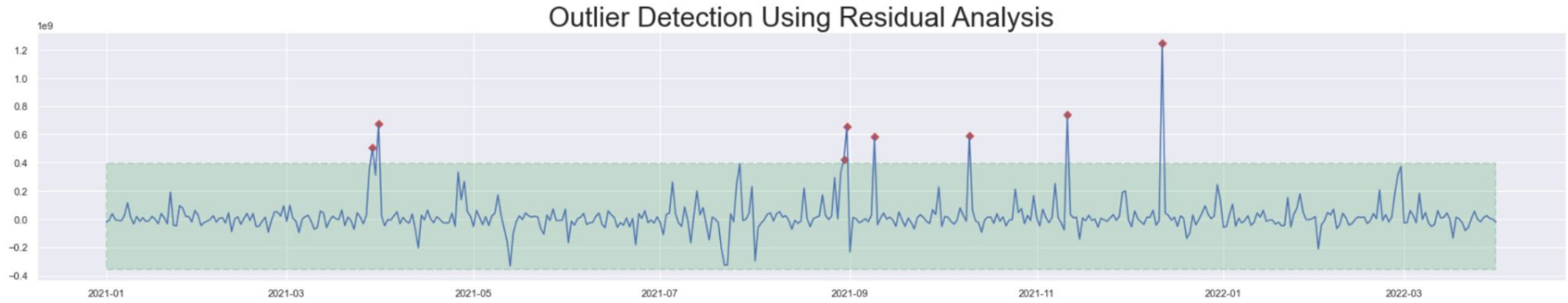
Anomaly Detection

```
resid_mu = resid_STL.mean()
resid_dev = resid_STL.std()

lower = resid_mu - 3*resid_dev
upper = resid_mu + 3*resid_dev
```

Dikategorikan Anomali apabila lebih besar/kurang dari 3 kali simpangan baku

```
2021-03-29
2021-03-31
2021-08-30
2021-08-31
2021-09-09
2021-10-10
2021-11-11
2021-12-12
```

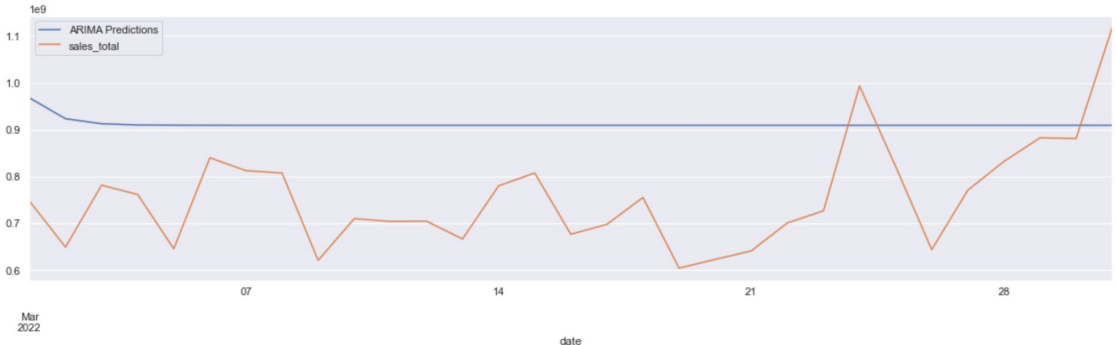


Models

- Model yang digunakan:
  - **ARIMA**: Assumption ANOVA tidak terpenuhi
    - Non-homogeneity of variance (levene test)
    - Non-normal data (shapiro-wilk)
    - Ada dependency antar variabel (Pearson Rank)

ARIMA tidak dapat fit dengan baik karena

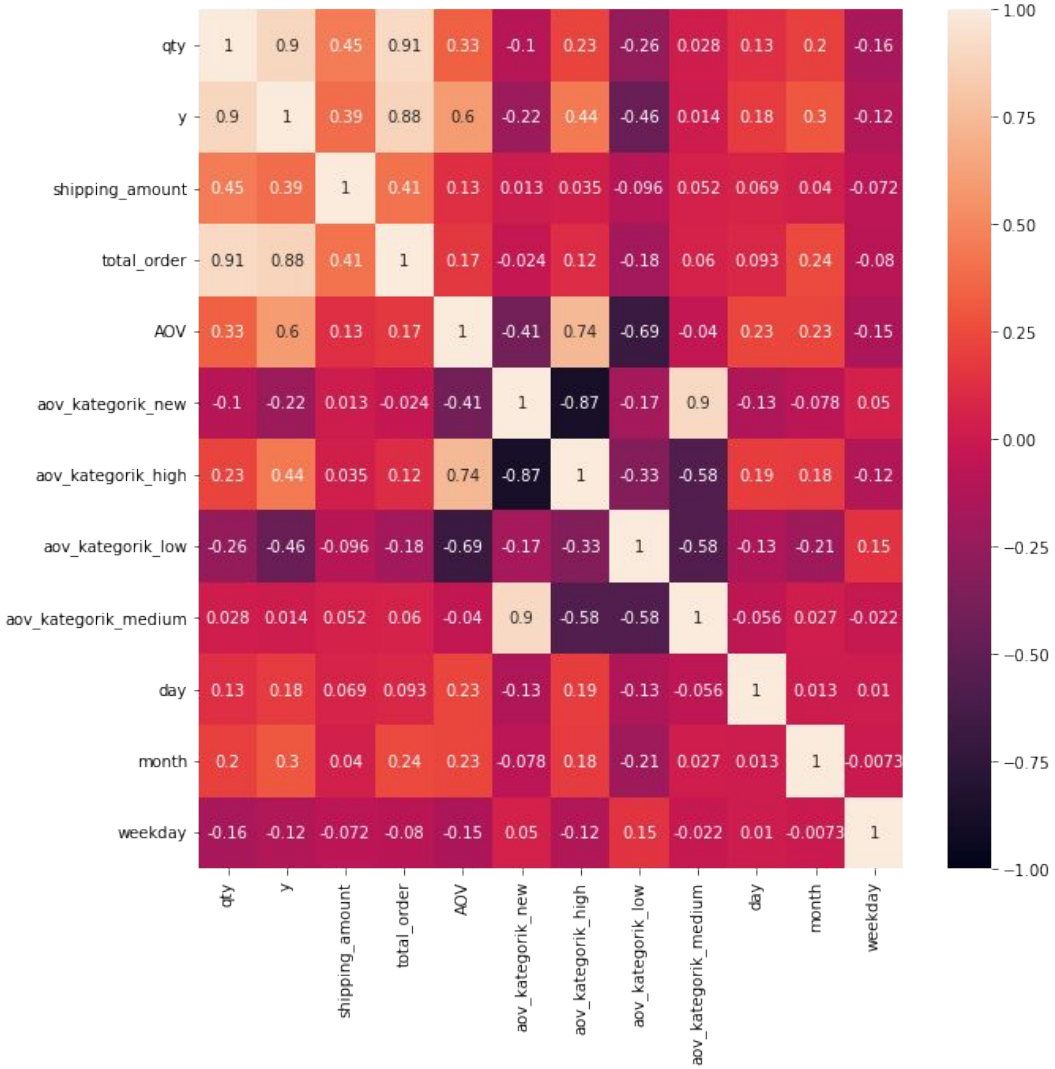
1. asumsi tidak terpenuhi
2. Data kurang stasioner
3. Doesn't have strong seasonality (Sehingga ARIMA hanya mengambil average of previous values (stackoverflow)



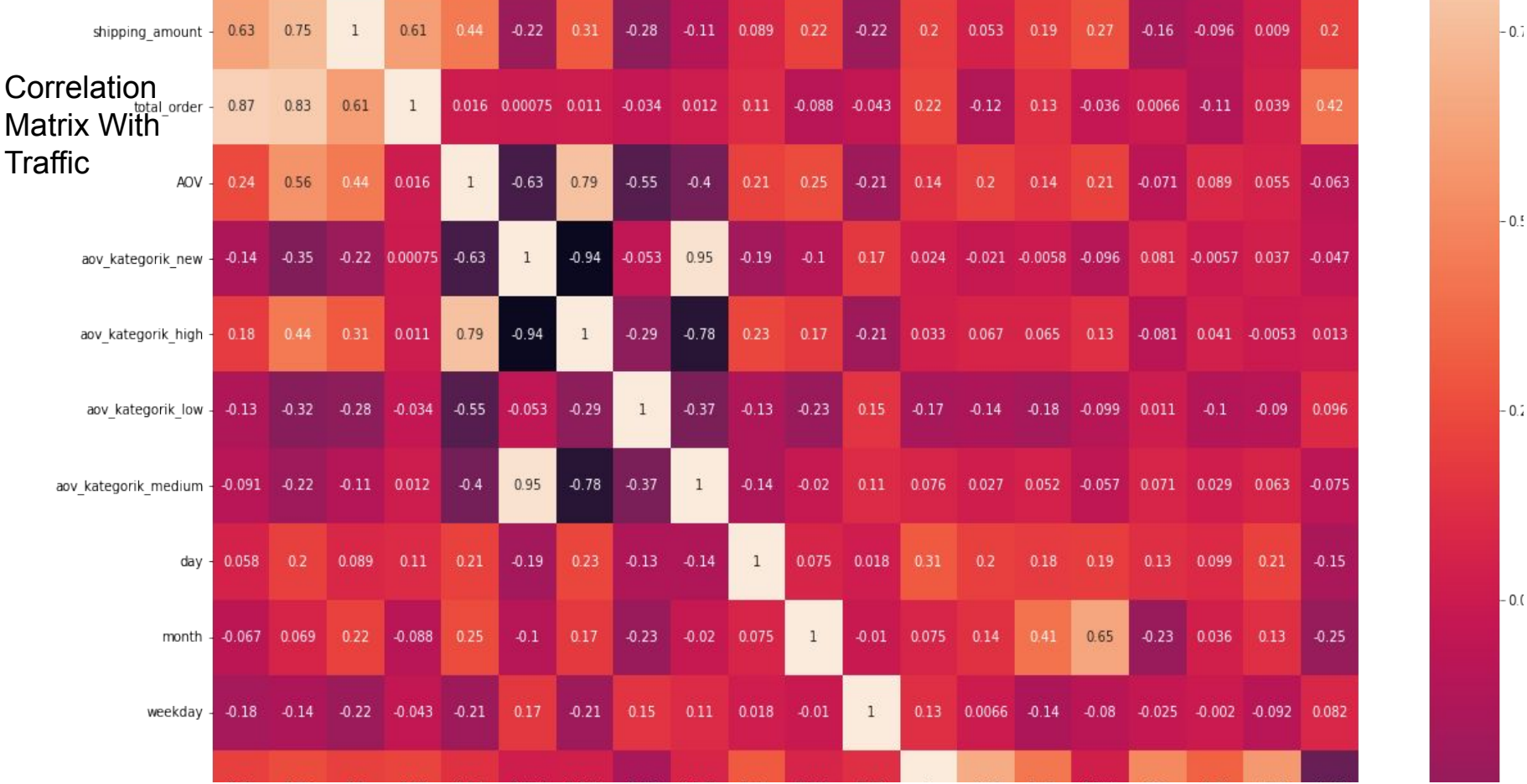
Models

- Model yang digunakan:
  - **Prophet**: interpretability, mudah dimanipulasi, use case sesuai
  - Notes:
    - Training menggunakan data Januari 2021 - April 2022 (kecuali data yang melibatkan traffic, kl itu dari 19 oktober 2021)
    - Testing menggunakan data Mei 2022
    - Metrics: RMSE,MAE,MAPE
    - Penambahan regressor di model prophet memerlukan controlled variabel yaitu value regressor tersebut untuk bulan depannya (diinput secara manual)

Correlation Matrix Without Traffic



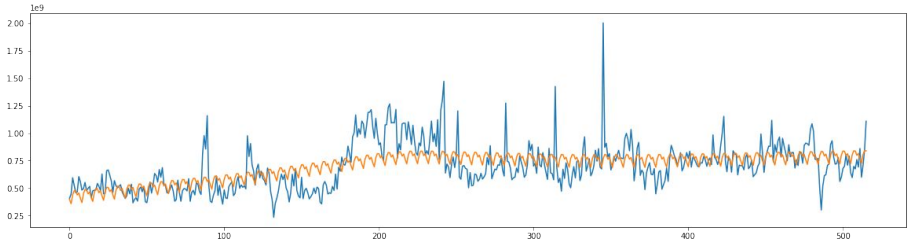
Correlation Matrix With Traffic



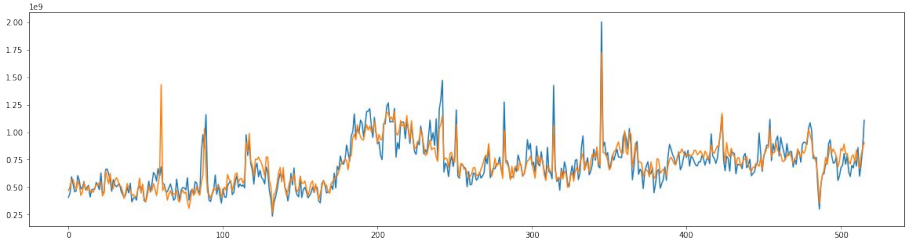
Beberapa Performance Prophet

No	Model	Testing			Training		
		RMSE	MAE	MAPE	RMSE	MAE	MAPE
1	Prophet(Baseline)	172375665	133798410.8	23.51778319	172207818.6	125522798	18.61226997
2	Prophet(Add Seasonality)	172375665	133798410.8	23.51778319	172207818.6	125522798	18.61226997
3	Regressor (total_order)	95482535.73	80157980.62	11.74862379	92012285.63	63435779.29	9.204028294
4	Regressor (AOV)	150558276.7	101011580.1	19.01717967	133002357.1	95949752.62	14.04029014
5	Regressor (AOV, label encoder)	150558276.7	101011580.1	19.01717967	133002357.1	95949752.62	14.04029014
6	Regressor (AOV, one hot)	141096133.9	101554913.2	18.25388762	146204027.2	106992071	15.84731476
7	Regressor (date decomp, aov label)	180252752.3	145738280.3	25.44985342	158882059.7	114461852.5	17.18407947
8	Baseline(dr 19 okt 2021)	181842033.2	145327305.1	25.11130419	159997166.5	106282728.3	14.08542843
9	Regressor(Conversion)	161429713.9	114725108.8	17.68159828	155078495.2	107567821.5	14.32358972

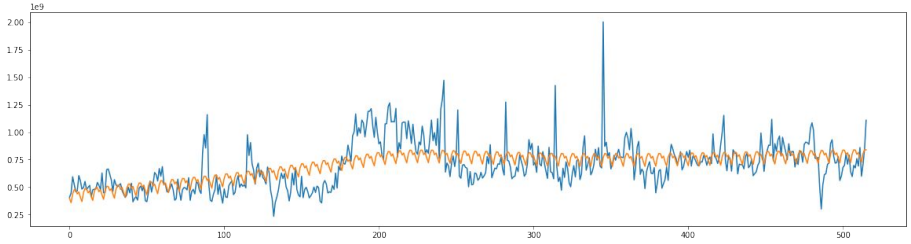
1



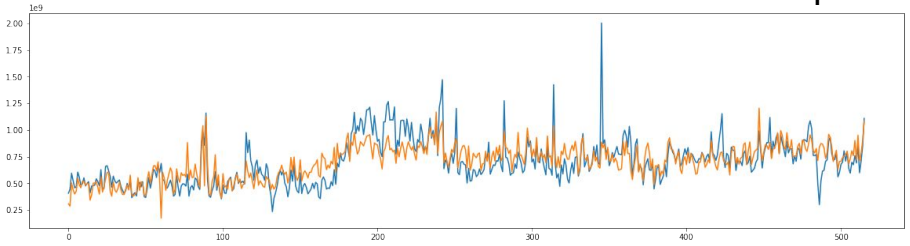
3



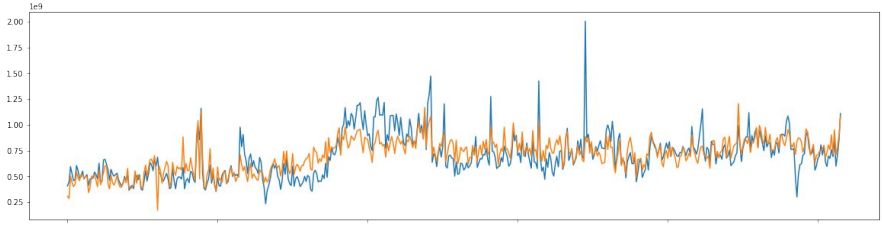
2



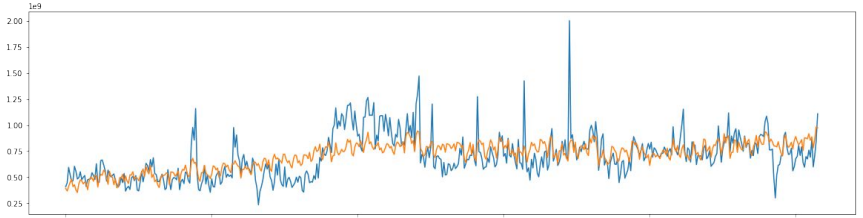
4



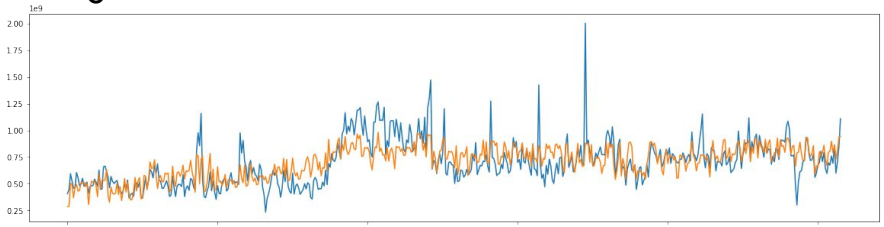
5


1e9
2.00
1.75
1.50
1.25
1.00
0.75
0.50
0.25
0
100
200
300
400
500

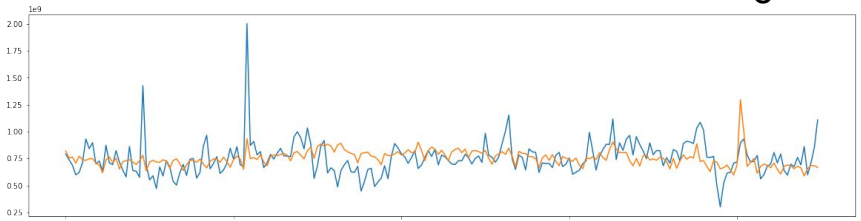
7


1e9
2.00
1.75
1.50
1.25
1.00
0.75
0.50
0.25
0
100
200
300
400
500

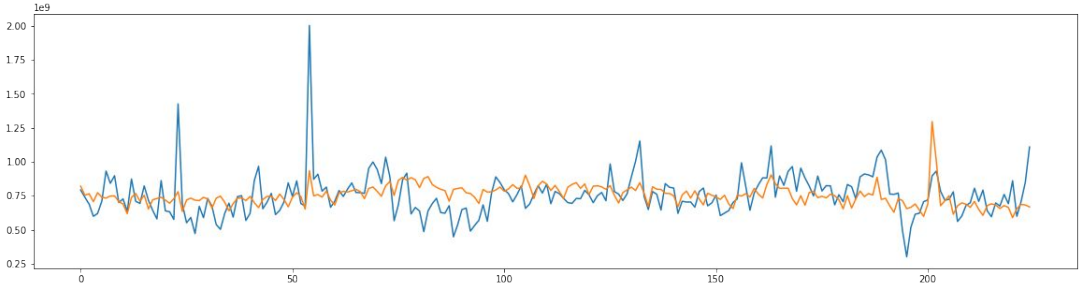
6


1e9
2.00
1.75
1.50
1.25
1.00
0.75
0.50
0.25
0
100
200
300
400
500

8

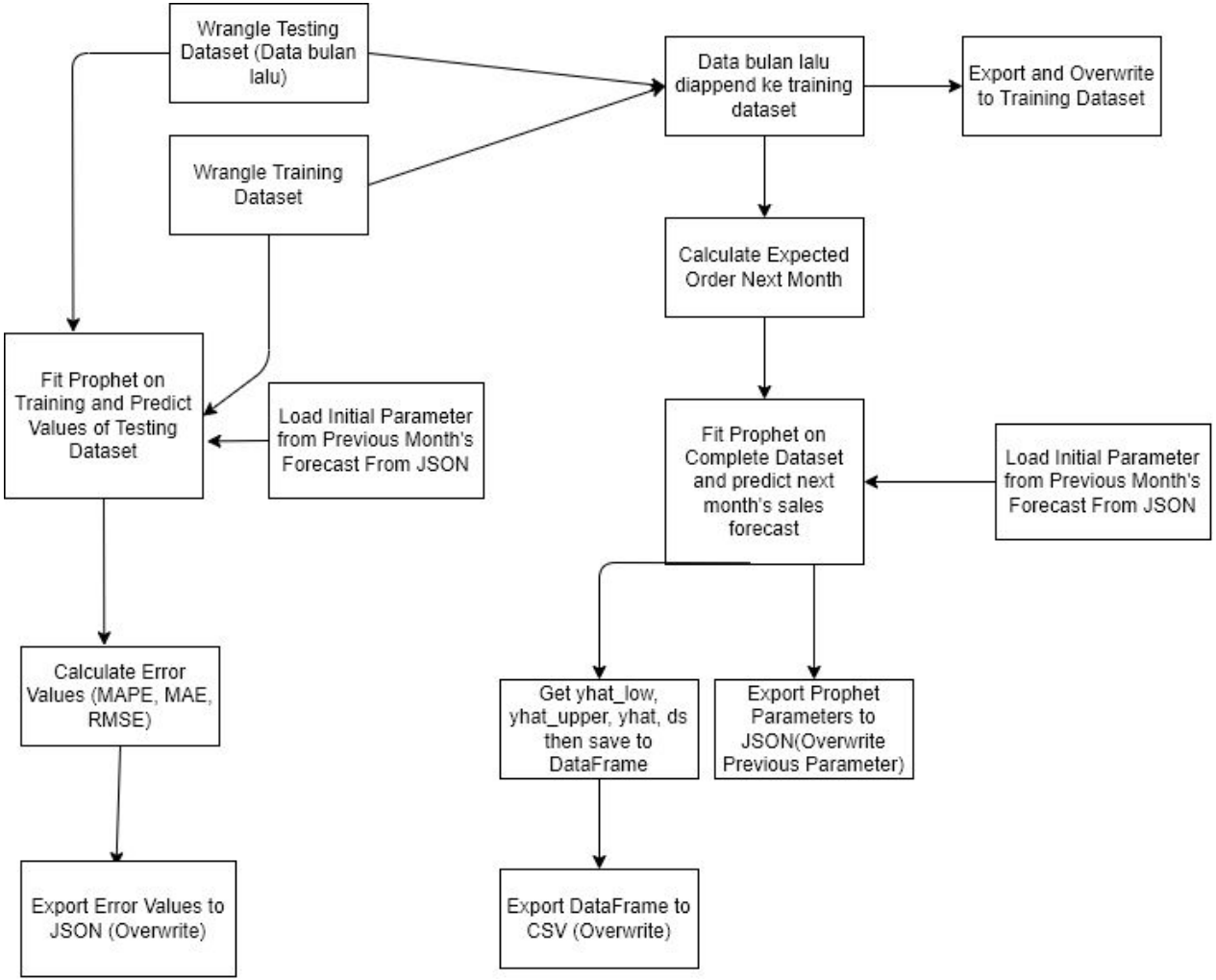

1e9
2.00
1.75
1.50
1.25
1.00
0.75
0.50
0.25
0
50
100
150
200

9


1e9
2.00
1.75
1.50
1.25
1.00
0.75
0.50
0.25
0
50
100
150
200

Models

- Model yang pernah dicoba digunakan:
  - **Prophet + LightGBM**: Menggunakan feature prophet sebagai regressor
  - Notes:
    - Flow lebih rumit
    - Performance comparable to Prophet Sendiri
  - **FlaML (Microsoft AutoML)**: Menggunakan XGBoost Limited Depth
  - Notes:
    - Tanpa regressor yang highly correlated performance lebih buruk dari prophet dengan regressor


Wrangle Testing Dataset (Data bulan lalu)
Wrangle Training Dataset
Data bulan lalu diappend ke training dataset
Export and Overwrite to Training Dataset
Calculate Expected Order Next Month
Fit Prophet on Training and Predict Values of Testing Dataset
Load Initial Parameter from Previous Month's Forecast From JSON
Fit Prophet on Complete Dataset and predict next month's sales forecast
Load Initial Parameter from Previous Month's Forecast From JSON
Calculate Error Values (MAPE, MAE, RMSE)
Get yhat_low, yhat_upper, yhat, ds then save to DataFrame
Export Prophet Parameters to JSON(Overwrite Previous Parameter)
Export Error Values to JSON (Overwrite)
Export DataFrame to CSV (Overwrite)

To Do

- Select kombinasi feature yang dapat dimanipulasi sebagai regressor
- Mencoba hyperparameter tuning prophet agar hasil lebih akurat lagi
  - Menambah kalendar High leverage observations (promosi 10/10 dkk)
  - Mencari seasonality sebenarnya