Saudi Central Bank Treasury Bill Prediction

What is treasury bill?

- Offered by central banks of all countries
- Unlike bonds, has a maturity term 1 year or less
- The return on them is called yields

Data

	date	repo_rate	inflation_rate	money_rate	1st Week(%)	4th Week(%)	13th Week(%)	26th Week(%)	52nd Week(%)
0	2021-10-01	1.0	0.80	7.9	0.51218	0.54208	0.57601	0.60201	0.6180
1	2021-09-30	1.0	0.80	7.9	0.51218	0.54208	0.57601	0.60201	0.6180
2	2021-09-29	1.0	0.80	7.9	0.51218	0.54208	0.57601	0.60201	0.6180
3	2021-09-28	1.0	0.80	7.9	0.51218	0.54208	0.57601	0.60201	0.6180
4	2021-09-27	1.0	0.80	7.9	0.51218	0.54208	0.57601	0.60201	0.6180
...	...	...	...	...	...	...	...	...	...
4287	2010-01-05	2.0	3.77	5.6	0.13400	0.18080	0.41260	0.53580	0.7674
4288	2010-01-04	2.0	3.77	5.6	0.13400	0.18080	0.41260	0.53580	0.7674
4289	2010-01-03	2.0	3.77	5.6	0.13400	0.18080	0.41260	0.53580	0.7674
4290	2010-01-02	2.0	3.77	5.6	0.13400	0.18080	0.41260	0.53580	0.7674
4291	2010-01-01	2.0	3.27	8.3	0.13750	0.19050	0.43250	0.55275	0.7395

- Obtained by SAMA and GASTAT
- 4000+ records, 8 features
  - Inflation rate
  - Repo rate (Repurchasing agreement): Interest rate collected by central bank to encourage commercial bank to borrow money
  - Money supply rate: Circulating cash and other assets
  - Maturity terms:
    - 1st week
    - 4th week
    - 13th week
    - 26th week
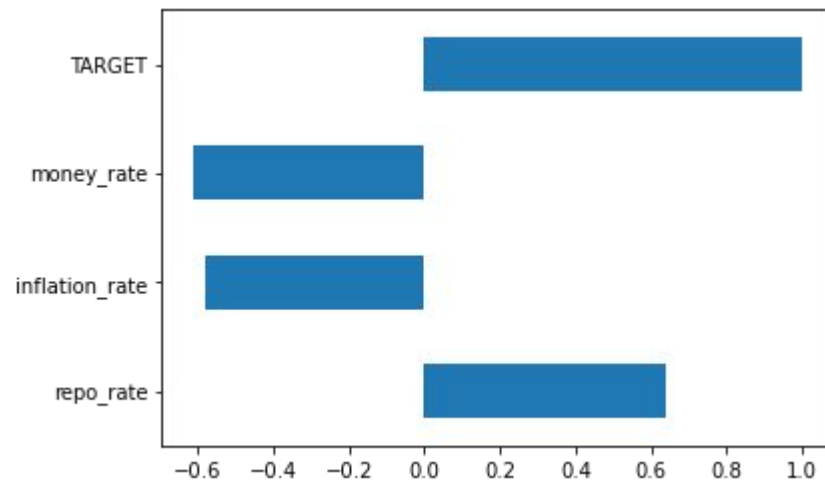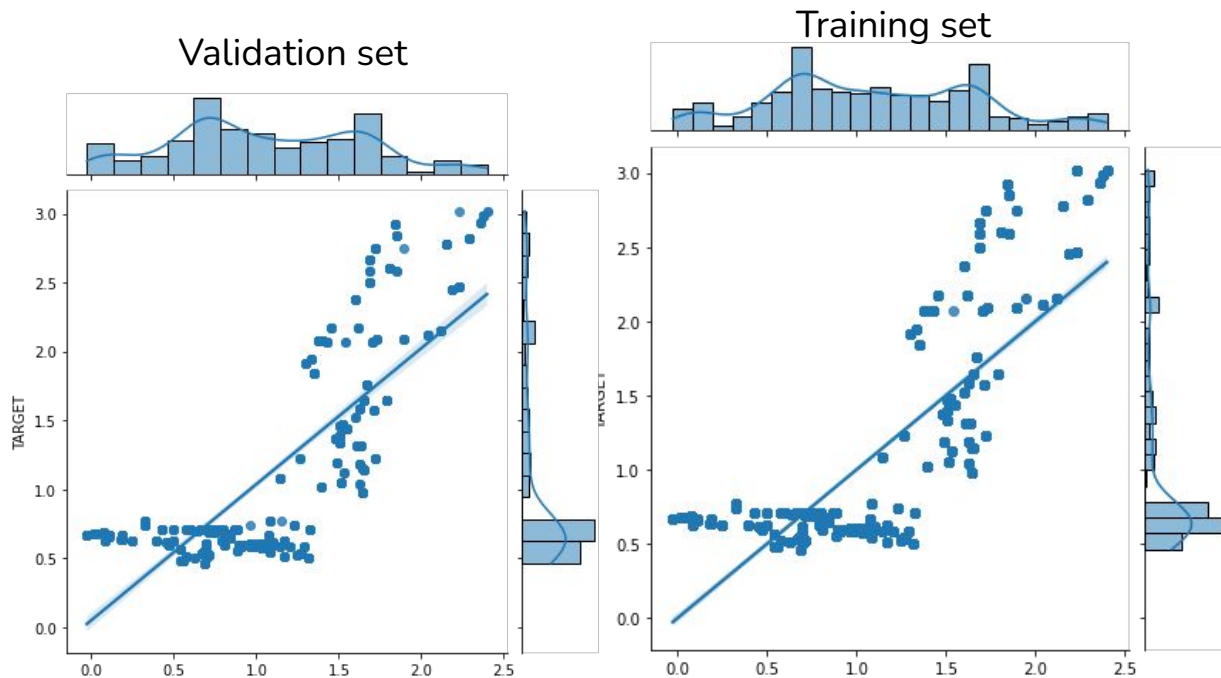   - 52nd week (our target)

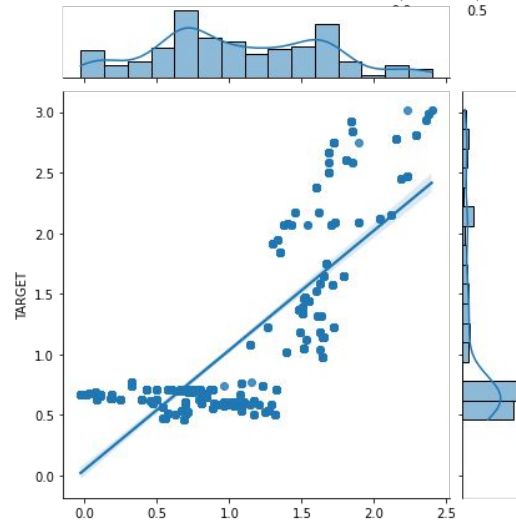# Data correlation
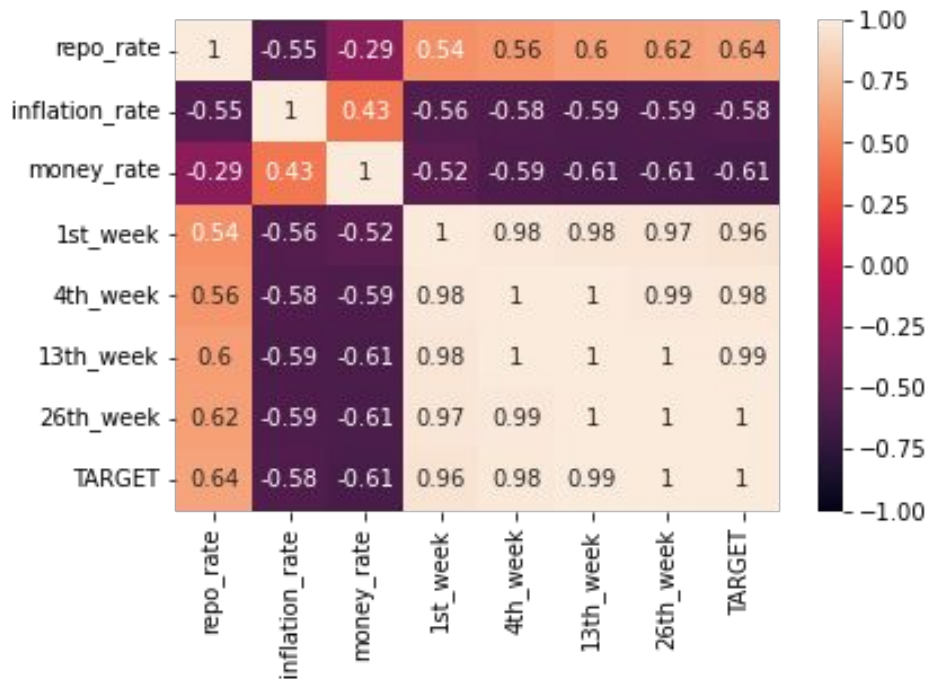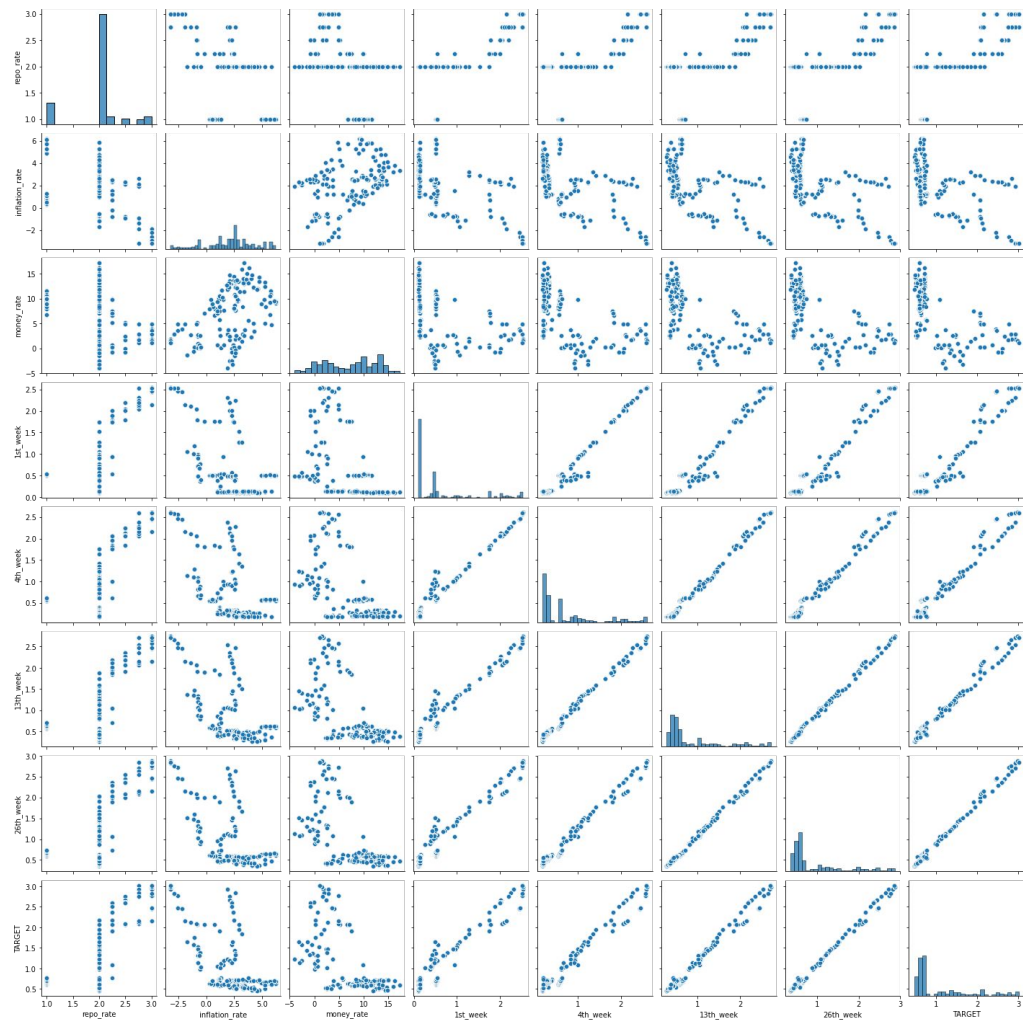
Inference:
- Money rate +, inflation +

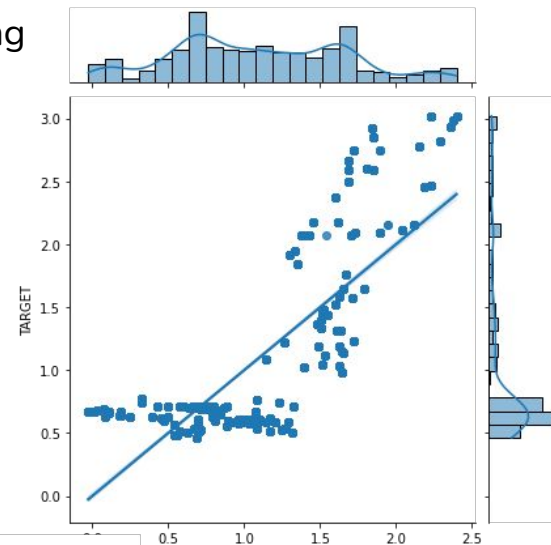# **Modeling**

- Linear Regression:
- Training set:
  - R2: 0.6247428
  - RMSE: 0.4389
- Validation set:
  - R2: 0.6169012
  - RMSE: 0.4422

Validation set

Training set

Ridge Model

Training

- Scale data
- Training set:
  - R2 on training set: 0.624742
  - RMSE on training set: 0.4389
- Validation set:
  - R2 on validation set: 0.6168989
  - RMSE on validation set: 0.442286

Validation

Polynomial regression model

- Training set:
  - R2 on training set: 0.829457
  - RMSE on training set: 0.295880
- Validation set:
  - R2 on validation set: 0.815630
  - RMSE on validation set: 0.3068257

K folds on all models

Running splits of 10, with shuffling

- Linear Regression: Mean: 0.6217198, Standard deviation: 0.01584672
- Ridge: Mean: 0.621721, Standard deviation: 0.0158376
- Polynomial Regression:
  - Mean of degree 1: 0.621719, Standard deviation: 0.0158467
  - Mean of degree 2: 0.825274, Standard deviation: 0.0153684
  - Mean of degree 3: 0.896315, Standard deviation: 0.0096554

Conclusion

- When money supply increase, inflation increases in a linear relationship
- Repo rate are not explainable by our data.
- Our models collectively tends to miss our target by approximately 0.39 points
- Polynomial regression with a degree of 2 is our choice of model but could be misleading since we have low feature model being simple, we might have bias-variance tradeoff.