Goal: Predict the bike renting studied in previous homework using Time Series techniques.

Data: The same data as for homework on Advanced regression

Target: `cnt` - Number of bikes rented per hour.

Endogenous variables: All variables engineered from target and time

Exogenous variables: Other than endogenous (e.g., something about weather)

Test sample: last month of data

Metrics: MAE

To-do:

- Perform EDA using Time Series Analysis techniques (any results of EDA from the previous homework can be used as well – please provide a link). EDA for this homework should include decomposition, correlation analysis, and tests for stationarity.
- Construct the baseline model using Linear Regression keeping in data only endogenous features.
- Improve your baseline model using results of target decomposition and exogenous features.
- Train ARIMA model.
- Perform analysis of results of modelling for each model. Do not forget to perform residual analysis (in ideal world residuals should look like white noise).
- Provide analytical comments and conclusions to all decisions and results.
- Compare obtained results with ones obtained using Advanced Regression techniques in the previous homework.

Additional

- Try use Prophet and/or SARIMAX.
- Try use advanced regression model with the time-based features engineered for linear model. You can also add new ones.

Some tips

In general, testing of residuals is a very broad topic, don't spend a lot of time here, the main point is: your model must be adequate and the residuals are a good source of information to check it.

Criteria

Total (max): 24 points

Criteria	Points	
EDA and Feature engineering:	**3**	
- Trend and seasonality components are found and discussed		2
- Correlation analysis is performed		1
Quality of the prediction:	**9**	
- MAE of model with extra features lower 180		1
- MAE of model with extra features lower 150		1
- Residuals of the model are close to white noise (considered only in case when the “whiteness” is proven)		1
- Model with exogenous features (X-model) is created		1
- MAE of X-model lower 100		1
- Arima model is created		1
- The choice of ARIMA coefficients is explained.		1
- MAE of ARIMA model lower 120		1
- Analysis of residuals is performed		1
Validation:	**1**	
- Data splitting for validation period considers time-components		1
Quality of delivered work:	**4**	
- Analytical comments provided		2
- The experiment is structured (file is readable, pictures have titles)		1
- Code is clear (reusable code in functions, comments, code is easy readable)		1
Additional:	**4**	
- Prophet		1
- SARIMAX		1
- Advanced Regression model is created		1
Extra points for improvements not considered in the criteria	**4**	
TOTAL:	**24**	