the following files:

Data preparation file (data\_preparation.ipyn ):

this note book will be a starter file for the data preparation, data loading, data cleaning, data visualization, feature engineering and uploading train and test datasets to S3 to be used in the model training pipeline and to be used in the model deployment pipeline.

In this file data preparation for the first and second project was carried out

Data cleaning by removing outliers from dataset

Data augmentation for the target to be predicted variables to form time series dataset to prepare time series forecasting.

For the second project

Data visualization:

A screenshot of a computer

Description automatically generated with low confidenceThe data visualization for controlled variables:

Data visualization:

The data visualization for response variables:

A picture containing text, handwriting, font, line

Description automatically generated

Data visualization for manipulated variables:

A picture containing text, diagram, line, handwriting

Description automatically generated

To detect outliers in the data set boxplot for the dataset will be used:

A picture containing text, screenshot, parallel, number

Description automatically generated

After removing the outlies:

A picture containing text, screenshot, parallel, number

Description automatically generated

Plotting the variables after fixing outliers:

A picture containing text, handwriting, font, line

Description automatically generated