Metro.digital Skills Test

Project Structure

---

data

eda_output.csv - For some reason the STL from statsmodel did not work with the datatype of datetime64. So had to save and load it again for decomposing the time series.

take_home_final_data.csv The dataset provided for the skills test containing of weekly sales data for 25 articles in the category of "Processed Foods" for a certain country for approximately two years.

test_predictions.csv - Output csv for forecasting of 3 weeks using the trained model.

model

Folder containing the trained model.

notebooks

exploration.ipynb - Exploratory data analysis to identify the significance of the features and drawing some useful insights from the data.

modelling.ipynb - Feature transformation and model training, prediction and evaluation.

src

train.py - Replica of modelling.ipynb but used for dockerization and automated scripts.

docker-compose.yaml

Running two docker containers one for training (train-model) and other for notebooks.

Dockerfile

Commands to build the docker image

requirements.txt

All necessary packages for the project to run flawlessly.

Execution Details

Build docker image

From the base directory run the following command in order to build docker image

```
docker build -t metrodigital-task .
```

Train the final model

```
docker-compose run train-model
```

Run jupyter notebook

```
docker-compose up notebook
```

Stop and Remove containers

```
docker compose down
```

```
docker image rm metrodigital-task
```

Training and Validation results