

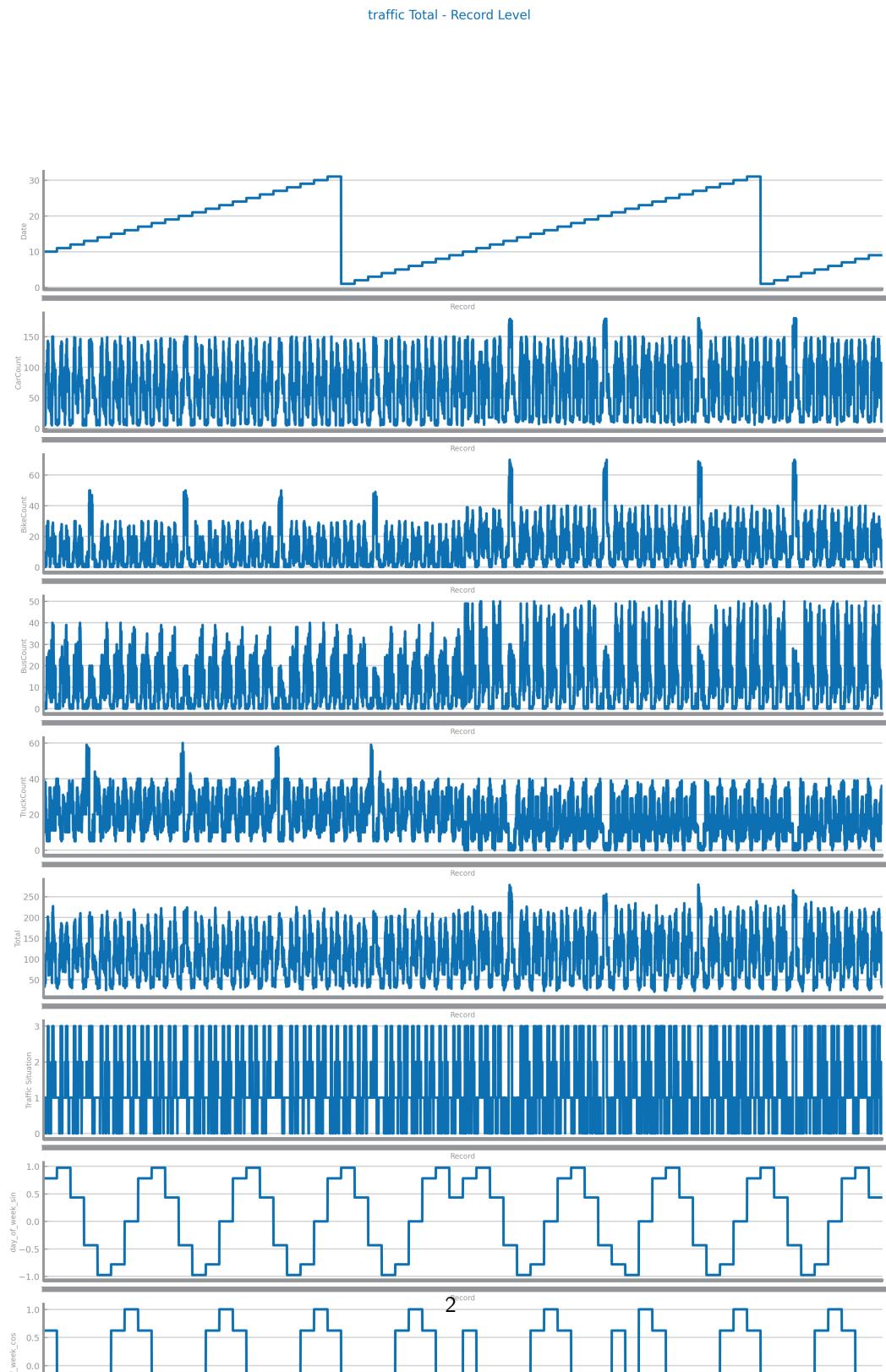
Data Science Project

Team nr: 16	Student 1: Antero Morgado IST nr: 1119213
	Student 2: David Ferreira IST nr: 1107077
	Student 3: José Fernandes IST nr: 1103727
	Student 4: Olha Buts IST nr: 1116276

TIME SERIES ANALYSIS

1 DATA PROFILING

Data Dimensionality and Granularity



traffic Total - Daily

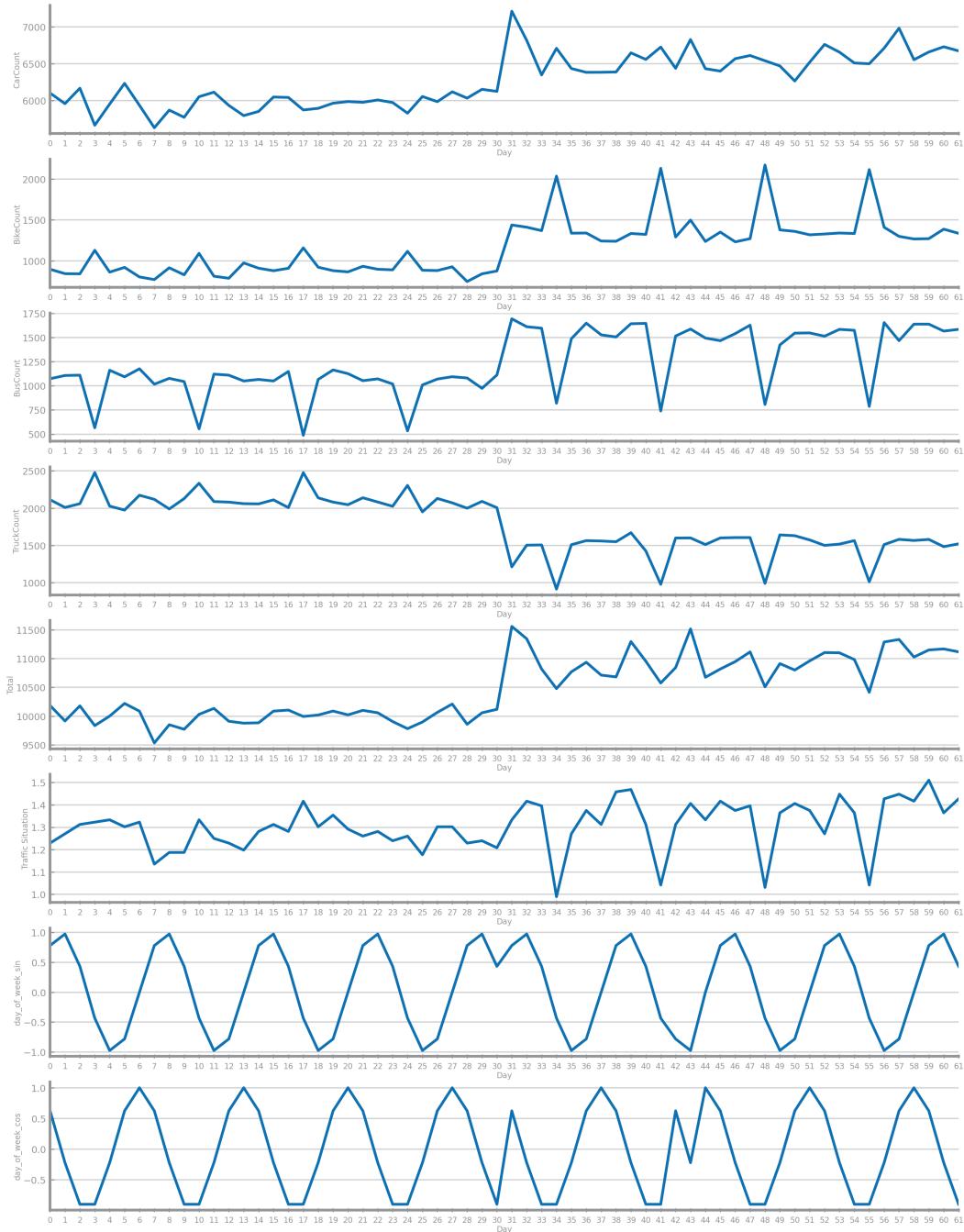


Figure 2: Time series 1 at the second chosen granularity

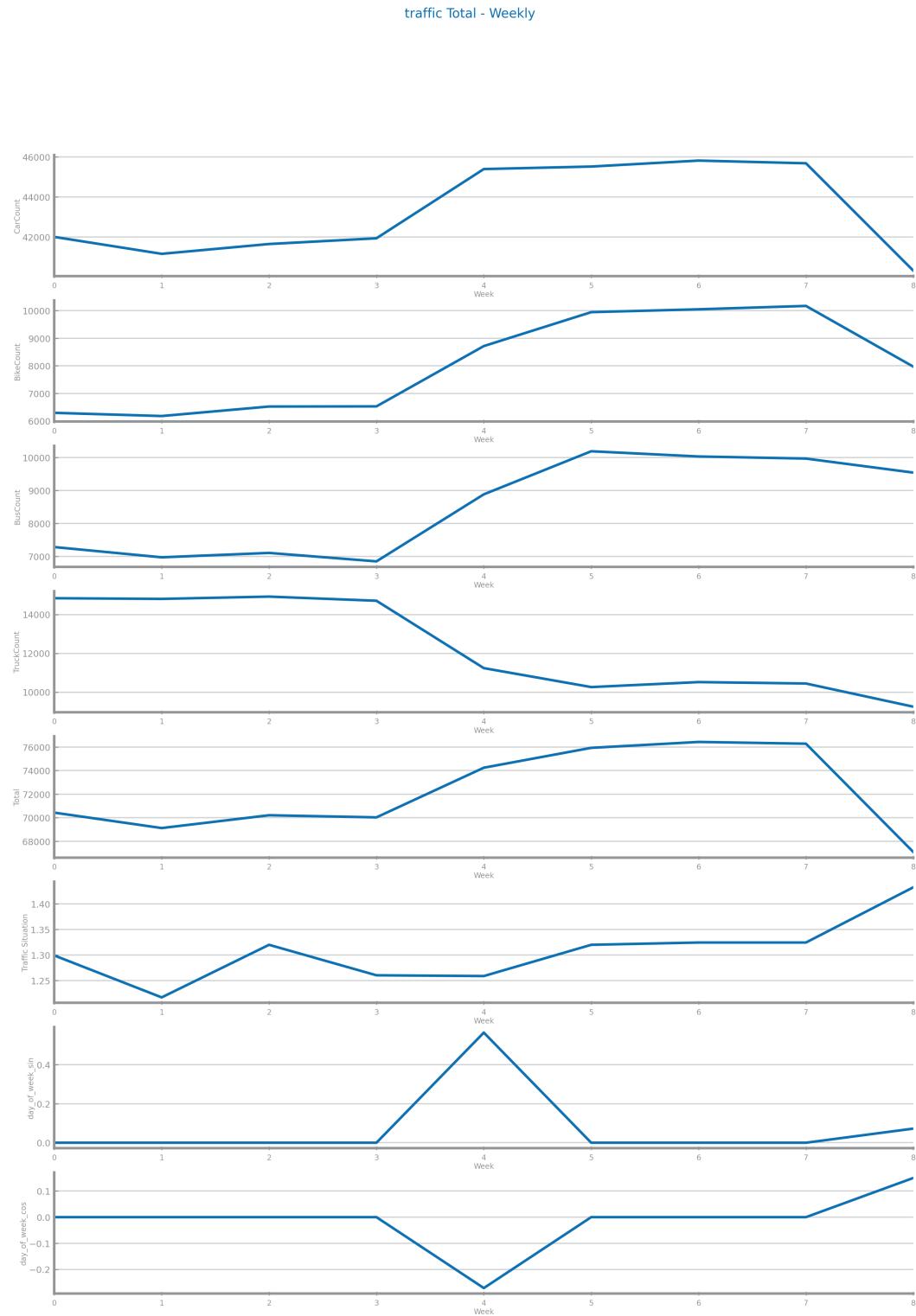


Figure 3: Time series 1 at the third chosen granularity

Data Distribution

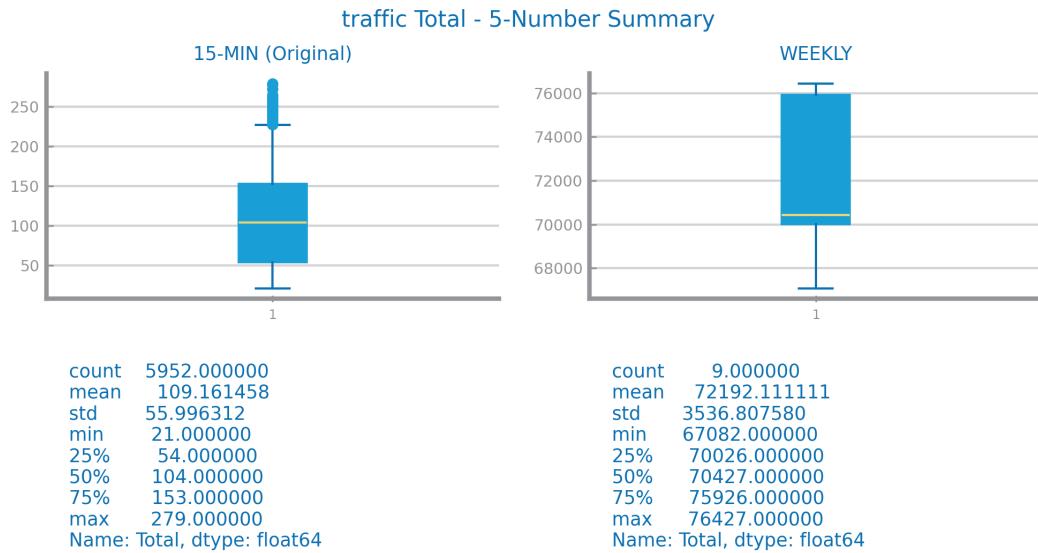


Figure 4: Boxplot(s) for time series 1

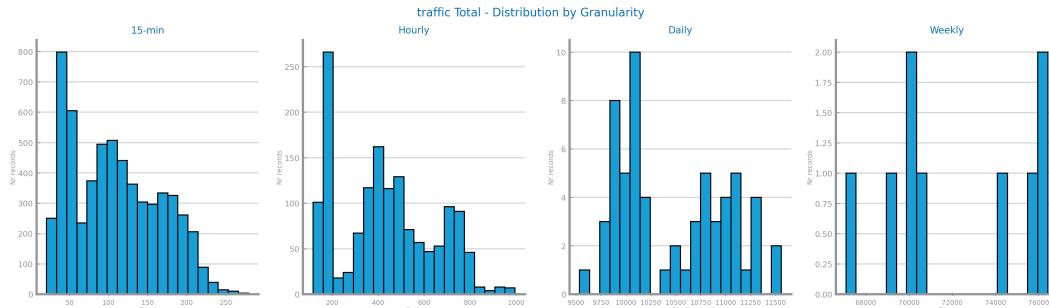


Figure 5: Histogram(s) for time series 1

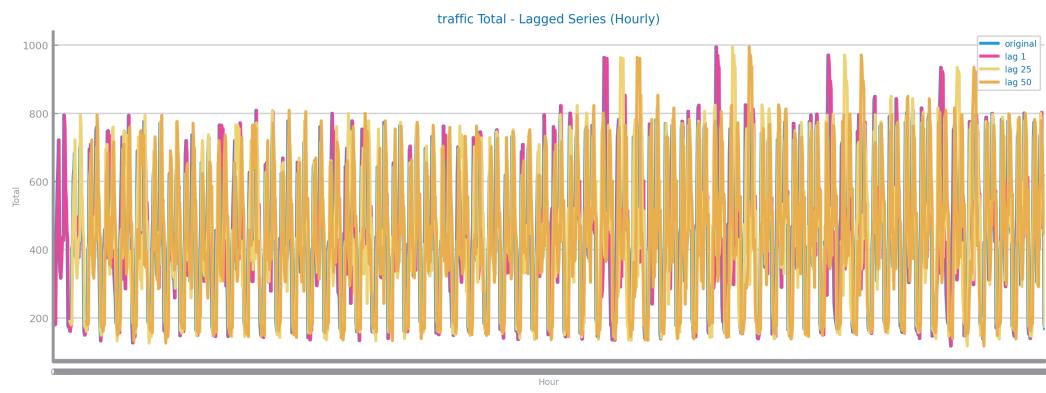


Figure 6: Autocorrelation lag-plots for original time series 1

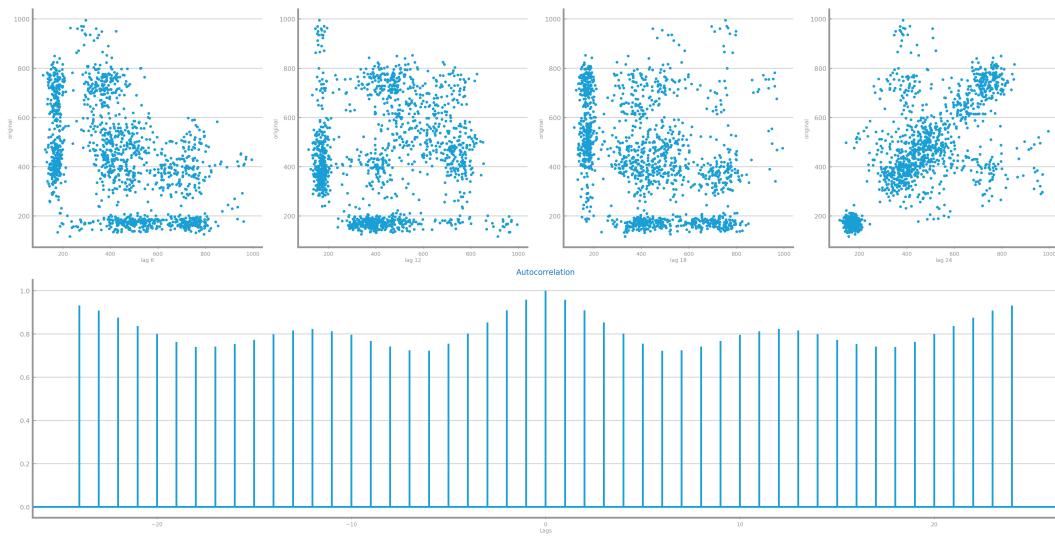


Figure 7: Autocorrelation correlogram for original time series 1

Data Stationarity

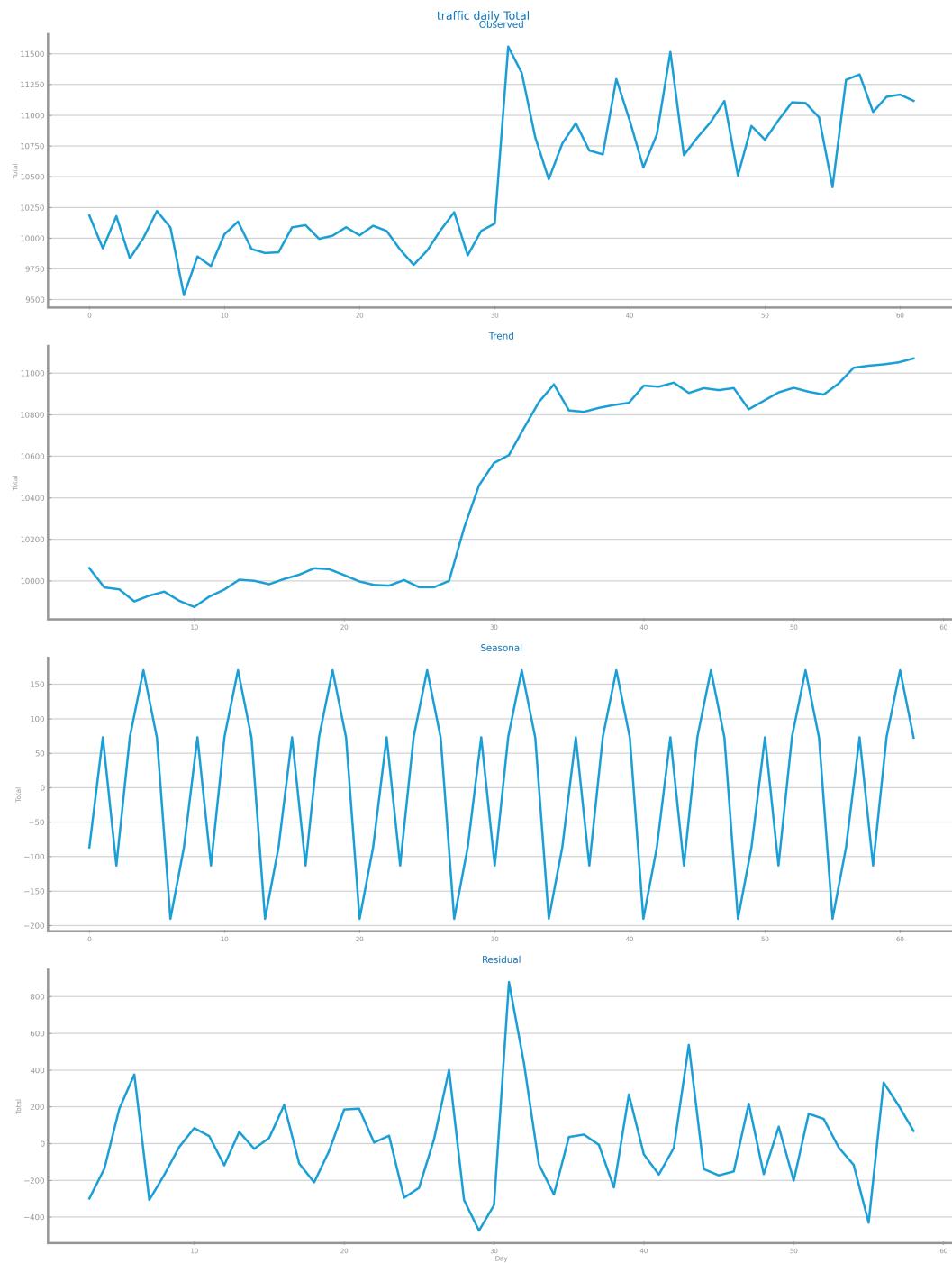


Figure 8: Components study for time series 1

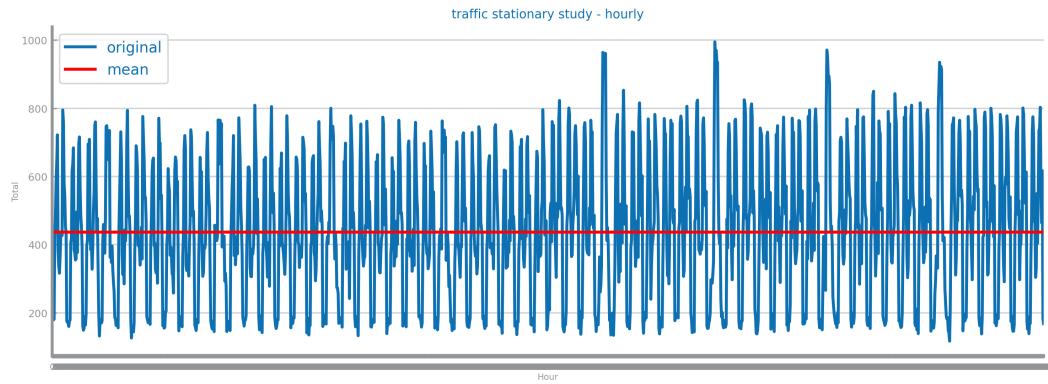


Figure 9: Stationarity study for time series 1

Augmented Dickey-Fuller Test Results:

Original (15-min):

- ADF Statistic: -14.441
- p-value: 0.000
- Critical Values: 1%: -3.431, 5%: -2.862, 10%: -2.567
- **The series IS stationary**

Hourly:

- ADF Statistic: -8.903
- p-value: 0.000
- Critical Values: 1%: -3.435, 5%: -2.864, 10%: -2.568
- **The series IS stationary**

Daily:

- ADF Statistic: -0.826
- p-value: 0.811
- Critical Values: 1%: -3.548, 5%: -2.913, 10%: -2.594
- **The series IS NOT stationary**

Deployment

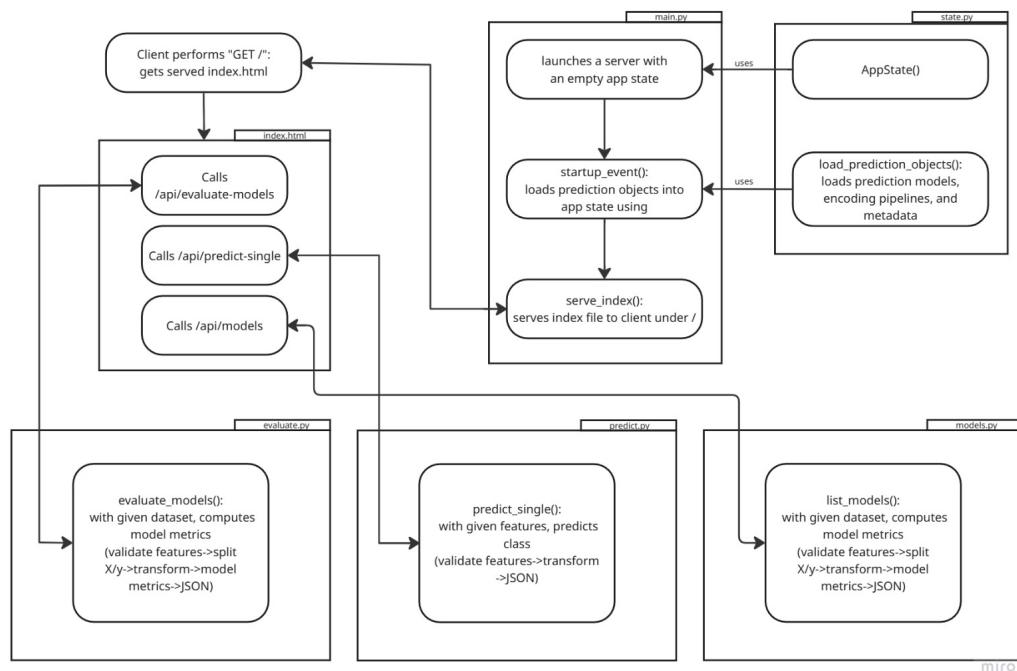


Figure 10: Architecture diagram

```
ds-deployment/
└── backend/
    ├── app/
    │   ├── main.py
    │   ├── state.py
    │   └── routes/
    │       ├── models.py
    │       ├── predict.py
    │       └── evaluate.py
    ├── models/
    │   └── traffic_lr_l1_500.joblib
    ├── pipeline/
    │   └── pipeline.joblib
    └── prediction_objects.json
└── frontend/
    └── index.html
└── requirements.txt
└── README.md
```

miro

Figure 11: File formats