FUTURE-DATA
Dublin Bikes

Ariana Alves Antunes, Imelda Finn, Marcus Ò Faolain

Due: April 21, 2023

1 Technical/Data issues

Missing data ...

Picked stations by clustering on ....

Excluded stations with less than 300,000 data points. None of the stations were open the whole time. 2 stations from each cluster. The cluster 2 stations had similar pre-pandemic characteristics and both serve hospitals.

Station Name	cluster	bike stands
FITZWILLIAM SQUARE EAST	0	40
HANOVER QUAY	0	40
MATER HOSPITAL	2	40
NEW CENTRAL BANK	1	40
PARNELL SQUARE NORTH	2	20
YORK STREET EAST	1	32

Variable modelled is usage, which is the absolute difference in available bikes from one time point to the next. For modelling, this is normalised by dividing by the number of bike stands.

2 The impact of the pandemic on bike usage

3 Estimated non-pandemic usage for 2020

The bike usage was way down compared to the prediction.[1] The prediction isn't too bad for January, February, but usage drops at the end of February and stays down for the rest of the year.

---

[1]RidgeCV(alphas=(0.1, 1.0, 10.0), cv=None, fit_intercept=True, gcv_mode=None, normalize=False, scoring=None, store_cv_values=False)
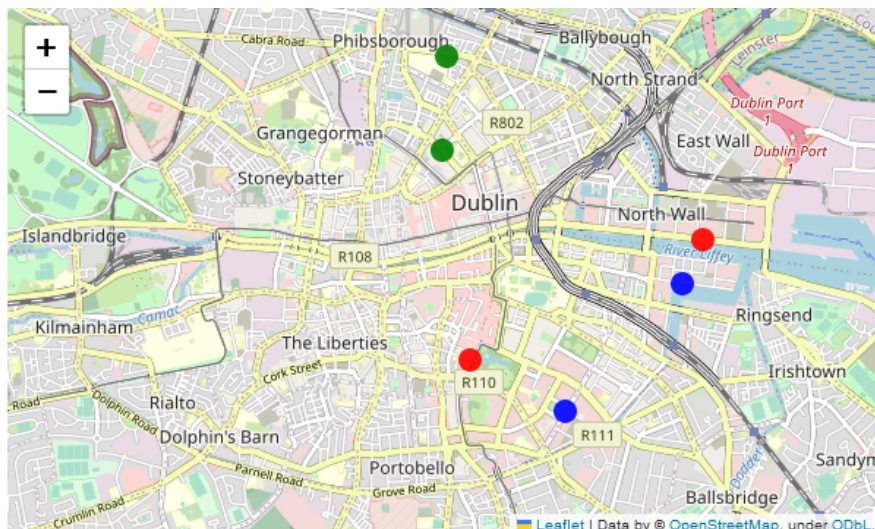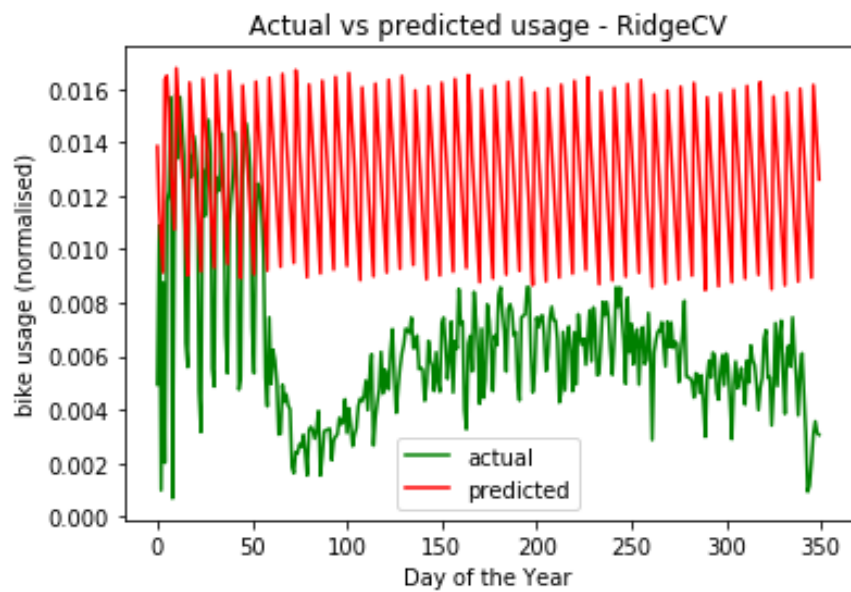
Figure 1: Selected Bike Stations

Figure 2: Predicted bike usage in 2020, compared to actual usage

notes

https://github.com/Arianaxsz/ML-Project
    Code in `fileName.ipynb`, `py/fileName.py`