France regional electricity consumption clustering using Generalised Cross Correlation.

Pierre Mercatoris

<2018-05-19 Sat>

1 Introduction

1.1 Cluster electricity consumption using GCC

1.2 Clustering time series

2 Methodology

2.1 Describe GCC

2.2 Describe data

The electricity consumption was available at a 30 minutes frequency for each of the 12 regions of France from 2013 to 2017. Each year of each region can be downloaded from the French transmission operator (Rte) download portal[1].

Périmètre	Nature	Date	Heures	Consommation
Grand-Est	Données définitives	2016-01-01	00:00	5130
Grand-Est	Données définitives	2016-01-01	00:15	
Grand-Est	Données définitives	2016-01-01	00:30	5130
Grand-Est	Données définitives	2016-01-01	00:45	
Grand-Est	Données définitives	2016-01-01	01:00	5014

3 Results

3.1 Data preparation

3.1.1 Cleaning

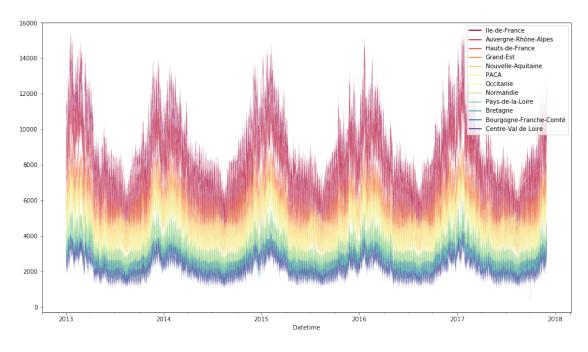1. Append all regions and years together

2. Clean the region names

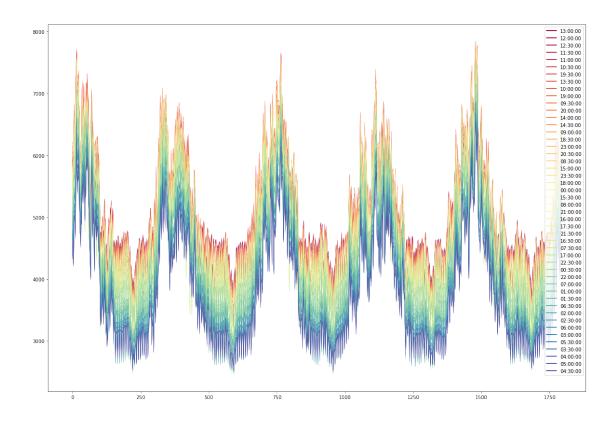[1] `http://www.rte-france.com/en/eco2mix/eco2mix-telechargement-en]]`

3. Format each column to appropriate data type

4. Set UTC time to correct summer/winter time changes

5. Pivot table so that the columns are the regions and the rows are consumption values

6. Resample the date as 30 minutes intervals

7. Pivot the table again so that we get daily value for each row

8.

9.

10.

This resulted in a table of 576 columns (48 x 12 regions) and 1794 rows/days.

Périmètre	Auvergne-Rhône-Alpes					
time	00:00:00	00:30:00	01:00:00	01:30:00	02:00:00	02:30:00
2013-01-02	7847.0	7674.0	7427.0	7441.0	7467.0	7550.0
2013-01-03	9028.0	8839.0	8544.0	8560.0	8569.0	8667.0
2013-01-04	8982.0	8754.0	8476.0	8480.0	8453.0	8554.0
2013-01-05	8625.0	8465.0	8165.0	8134.0	8087.0	8149.0
2013-01-06	8314.0	8097.0	7814.0	7791.0	7785.0	7842.0

3.1.2 Transformation

3.2 Distance calculation

3.2.1 Selecting k

3.2.2 Distance matrix

3.3 Clustering

3.4 Cluster analysis

4 Conclusion