Prediction of ozone level in Boston

Lucas Emanuel Resck Domingues   Lucas Machado Moscheb*

Predicting O3 in Boston

```
##
## Attaching package: 'zoo'

## The following objects are masked from 'package:base':
##
##     as.Date, as.Date.numeric

##
## Attaching package: 'dplyr'

## The following objects are masked from 'package:stats':
##
##     filter, lag

## The following objects are masked from 'package:base':
##
##     intersect, setdiff, setequal, union
```

---
*Escola de Matemática Aplicada

Load and visualize

Daily average level of O3 in Boston
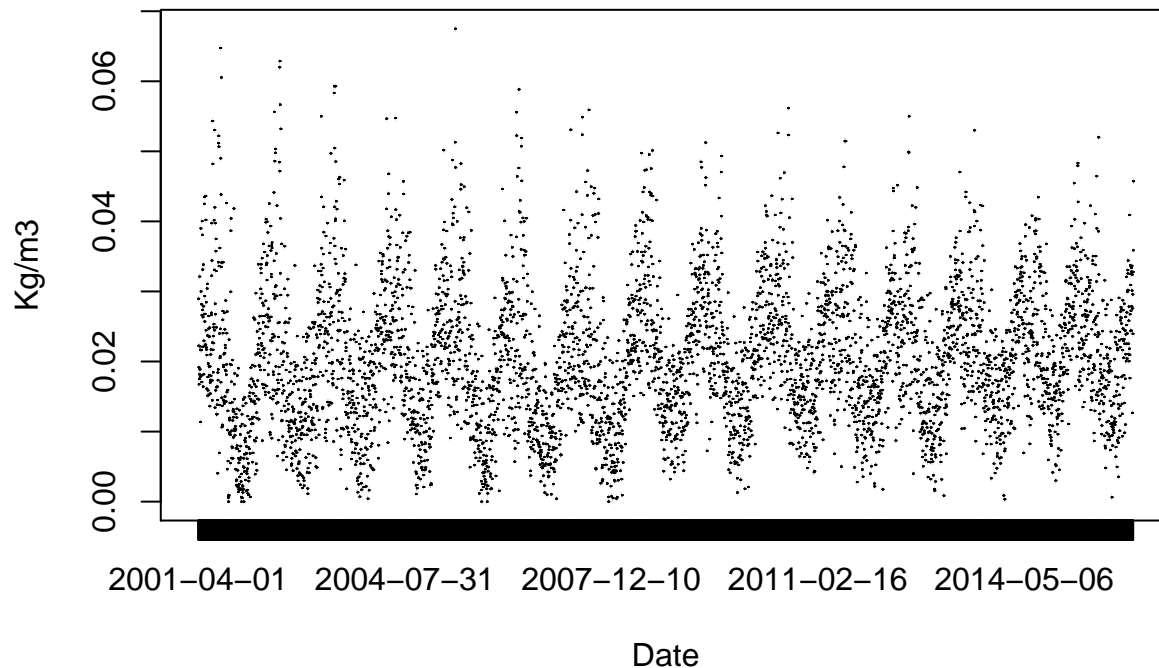
Data treatment

We found that no NaN value exist in the dataset. However there was duplicated days. Comparing with the neighbors, we saw that the mean of the two values would be a good option.

```
sum(is.na(bos$O3.Mean))
```

```
## [1] 0
```

```
bos$Date.Local[duplicated(bos$Date.Local)]
```

```
## [1] 2002-06-09
## 5235 Levels: 2001-04-01 2001-04-02 2001-04-03 2001-04-04 ... 2016-03-31
```

```
bos[bos$Date.Local %in% c("2002-06-08", "2002-06-09", "2002-06-10"),]
```

```
##       X   City          State Site.Num Date.Local  O3.Mean
## 412 412 Boston Massachusetts        42 2002-06-08 0.022917
## 413 413 Boston Massachusetts        42 2002-06-09 0.036190
## 414 414 Boston Massachusetts        42 2002-06-09 0.037000
## 415 415 Boston Massachusetts        42 2002-06-10 0.023389
```

```
##       X   City          State Site.Num Date.Local  O3.Mean
## 412 412 Boston Massachusetts        42 2002-06-08 0.022917
## 413 413 Boston Massachusetts        42 2002-06-09 0.036595
## 414 415 Boston Massachusetts        42 2002-06-10 0.023389
```

We use the Zoo library handle the data from now on. We also separate the train and test (last 365 days).

Daily average level of O3 in Boston
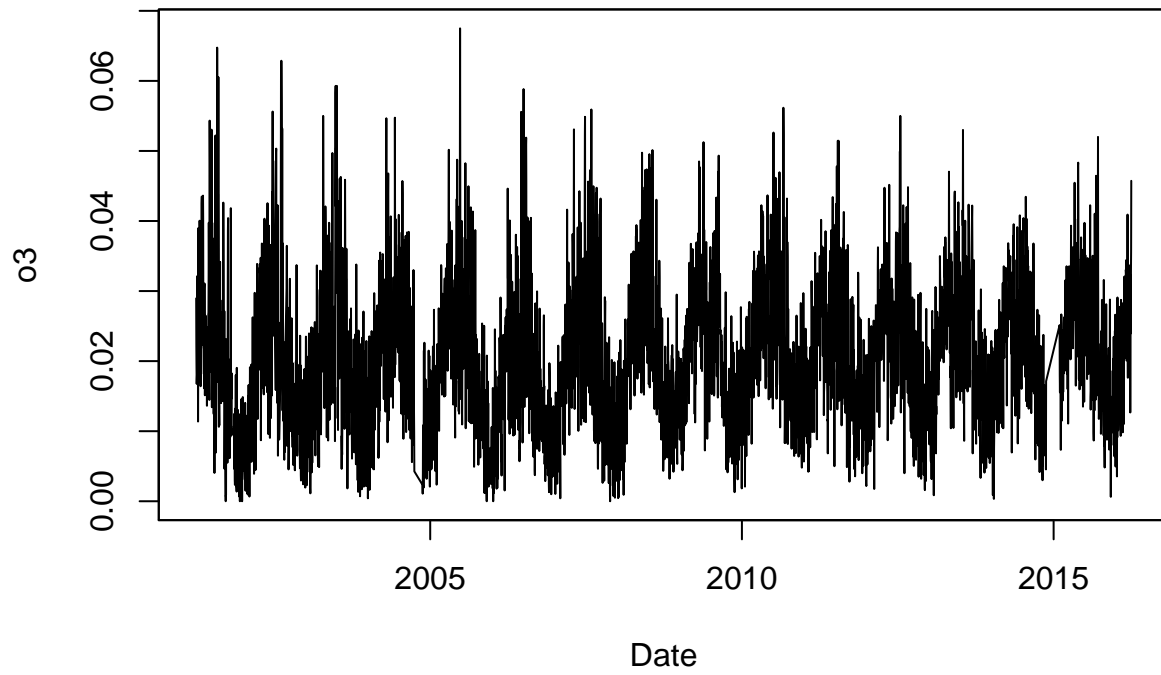
Models

Regression

Decompose

Holt-Winters

ARMA