# Anomaly Detection

Alex

15/07/2021

```
library(tidyverse)
```

```
## -- Attaching packages --------------------------------------- tidyverse 1.3.1 --

## v ggplot2 3.3.5     v purrr   0.3.4
## v tibble  3.1.2     v dplyr   1.0.7
## v tidyr   1.1.3     v stringr 1.4.0
## v readr   1.4.0     v forcats 0.5.1

## -- Conflicts ------------------------------------------ tidyverse_conflicts() --
## x dplyr::filter() masks stats::filter()
## x dplyr::lag()    masks stats::lag()
```

```
library(anomalize)
```

```
## == Use anomalize to improve your Forecasts by 50%! ============================
## Business Science offers a 1-hour course - Lab #18: Time Series Anomaly Detection!
## </> Learn more at: https://university.business-science.io/p/learning-labs-pro </>
```

```
library(dplyr)
```

```
path <-"http://bit.ly/CarreFourSalesDataset"

anomaly <-read.csv(path)

head(anomaly)
```

```
##        Date    Sales
## 1  1/5/2019 548.9715
## 2  3/8/2019  80.2200
## 3  3/3/2019 340.5255
## 4 1/27/2019 489.0480
## 5  2/8/2019 634.3785
## 6 3/25/2019 627.6165
```

## Converting date

```
anomaly$Date<- as.Date(anomaly$Date,format = "%m/%d/%Y")
anomaly[["Date"]] <- as.POSIXct(anomaly$Date)
```

## converting into a tibble

```
library(tibbletime)
```

```
##
## Attaching package: 'tibbletime'

## The following object is masked from 'package:stats':
##
##     filter
```

## Summary

```
summary(anomaly)
```

```
##       Date                          Sales
##  Min.   :2019-01-01 03:00:00   Min.   :  10.68
##  1st Qu.:2019-01-24 03:00:00   1st Qu.: 124.42
##  Median :2019-02-13 03:00:00   Median : 253.85
##  Mean   :2019-02-14 03:05:45   Mean   : 322.97
##  3rd Qu.:2019-03-08 03:00:00   3rd Qu.: 471.35
##  Max.   :2019-03-30 03:00:00   Max.   :1042.65
```

```
library(tidyverse)
library(anomalize)
library(dplyr)
```

```
library(Rcpp)
```

```
anomaly <- anomaly %>%
tibbletime::as_tbl_time(index = Date)
```