

Anomaly Detection

```
# Load tidyverse and anomalize
# ---
#
library(tidyverse)
```

```
## Warning: package 'tidyverse' was built under R version 3.6.3
```

```
## -- Attaching packages ----- tidyverse 1.3.0 --
```

```
## v ggplot2 3.3.2      v purrr  0.3.4
## v tibble  3.0.3      v dplyr  1.0.2
## v tidyr   1.1.2      v stringr 1.4.0
## v readr   1.3.1      v forcats 0.5.0
```

```
## Warning: package 'ggplot2' was built under R version 3.6.3
```

```
## Warning: package 'tibble' was built under R version 3.6.3
```

```
## Warning: package 'tidyr' was built under R version 3.6.3
```

```
## Warning: package 'readr' was built under R version 3.6.3
```

```
## Warning: package 'purrr' was built under R version 3.6.3
```

```
## Warning: package 'dplyr' was built under R version 3.6.3
```

```
## Warning: package 'stringr' was built under R version 3.6.3
```

```
## Warning: package 'forcats' was built under R version 3.6.3
```

```
## -- Conflicts ----- tidyverse_conflicts() --
```

```
## x dplyr::filter() masks stats::filter()
```

```
## x dplyr::lag()     masks stats::lag()
```

```
library(anomalize)
```

```
## Warning: package 'anomalize' was built under R version 3.6.3
```

```
## == 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(tibbletime)
```

```
## Warning: package 'tibbletime' was built under R version 3.6.3
```

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

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

```
library(lubridate)
```

```
## Warning: package 'lubridate' was built under R version 3.6.3
```

```
##  
## Attaching package: 'lubridate'
```

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

```
#loading or dataset
```

```
sale = read.csv("http://bit.ly/CarreFourSalesDataset")
```

```
head(sale)
```

```
##      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
```

```
sale$Date <- as.Date(sale$Date )
```

```
sale_tb <- as_tibble(sale)
```

```
head(sale_tb)
```

```
## # A tibble: 6 x 2  
##   Date      Sales  
##   <date>    <dbl>  
## 1 0001-05-20 549.  
## 2 0003-08-20  80.2  
## 3 0003-03-20 341.  
## 4 NA        489.  
## 5 0002-08-20 634.  
## 6 NA        628.
```

```
sapply(sale_tb, class)
```

```
##      Date      Sales  
##      "Date" "numeric"
```

```
sale_tb <- sale_tb %>%  
  tibbltime::as_tbl_time(index = Date)
```

```
head(sale_tb)
```

```
## # A time tibble: 6 x 2  
## # Index: Date  
##   Date      Sales  
##   <date>    <dbl>  
## 1 0001-05-20 549.  
## 2 0003-08-20  80.2  
## 3 0003-03-20 341.  
## 4 NA        489.  
## 5 0002-08-20 634.  
## 6 NA        628.
```

```
sale_tbl <- sale %>%  
  tibbltime::as_tbl_time(index = Date)
```

```
sale_tb %>%  
  time_decompose(Date, method = "stl") %>%  
  anomalize(remainder) %>%  
  time_recompose() %>%  
  plot_anomalies(time_recomposed = TRUE, ncol = 3, alpha_dots = 0.5)
```

```
## Error: 'nm' must be 'NULL' or a character vector the same length as 'x'
```