Project Data Mining

Contents

Inflation Time Series Analysis

```
library(arrow)
library(dplyr)
inflation = read_parquet("https://storage.dosm.gov.my/cpi/cpi_2d.parquet")

str(inflation)
```

```
## tibble [2,506 x 3] (S3: tbl_df/tbl/data.frame)
##  $ date    : POSIXct[1:2506], format: "2010-01-01 08:00:00" "2010-02-01 08:00:00" ...
##  $ division: chr [1:2506] "overall" "overall" "overall" "overall" ...
##  $ index   : num [1:2506] 99.4 99.4 99.4 99.4 99.6 ...
```

```
unique(inflation$division)
```

```
##  [1] "overall" "01"      "02"      "03"      "04"      "05"      "06"
##  [8] "07"      "08"      "09"      "10"      "11"      "12"      "13"
```

Take only the CPI value for the overall market.

```
inflation = inflation %>%
  subset(division == 'overall')
```

```
unique(inflation$division)
```

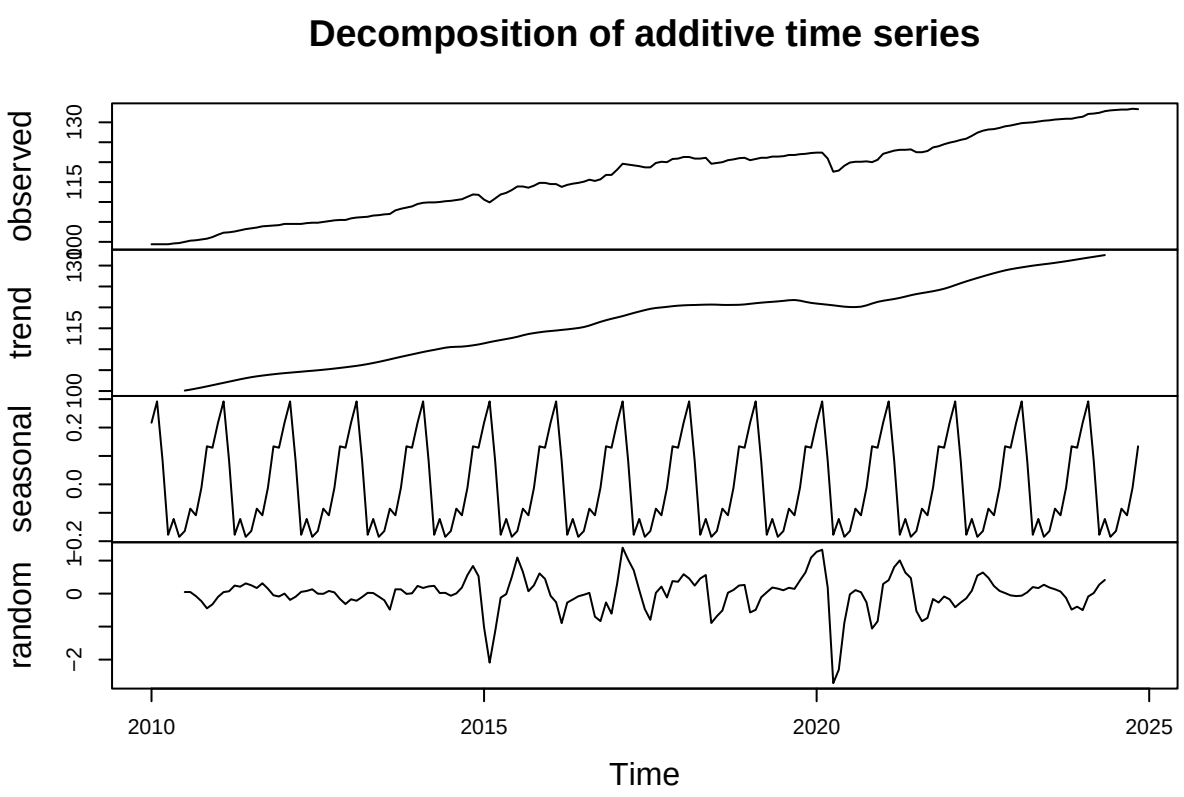
```
## [1] "overall"
```

Convert data into time series

```
inflation = inflation[,3]
inflation_ts = ts(inflation, start = c(2010,1), frequency = 12)
```

See components of the data

```
plot(decompose(inflation_ts))
```

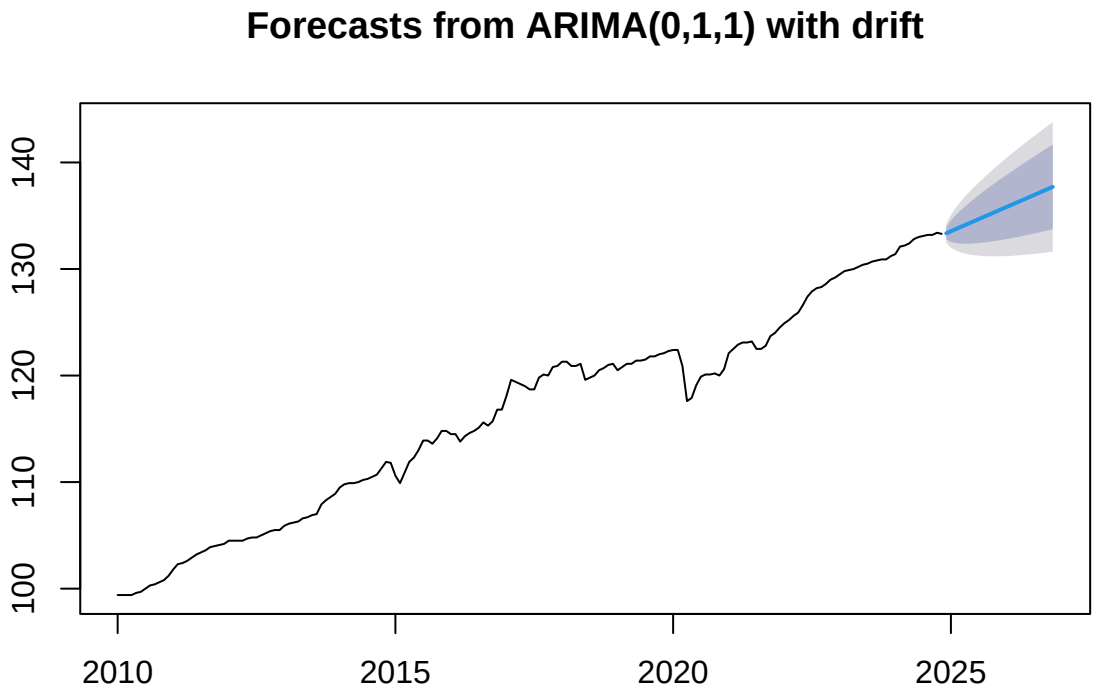


Forecast using ARIMA

```
library(forecast)
```

```
## Registered S3 method overwritten by 'quantmod':
##   method            from
##   as.zoo.data.frame zoo
```

```
plot(forecast(auto.arima(inflation_ts)))
```



Forecast using TBATS

```
plot(forecast(tbats(inflation_ts)))
```