Time series decomposition and transformations. Practice.

2022-09-21

```
library(tibble)
library(dplyr)
library(tidyr)
library(lubridate)
library(ggplot2)

# tsibble: tidy temporal data frames and tools
library(tsibble)

# fable (forecast table)
library(fable)

# fabletools - provides tools for building modelling packages, with a focus on time series forecasting
library(fabletools)

# Feature Extraction and Statistics for Time Series in tsibble format
library(feasts)

# tsibbledata: used datasets for example global_economy
library(tsibbledata)

library(fpp3)
```

There are multiple libraries loaded and to make life easier here are some associations between used function and library name:

"Regular" tydiverce

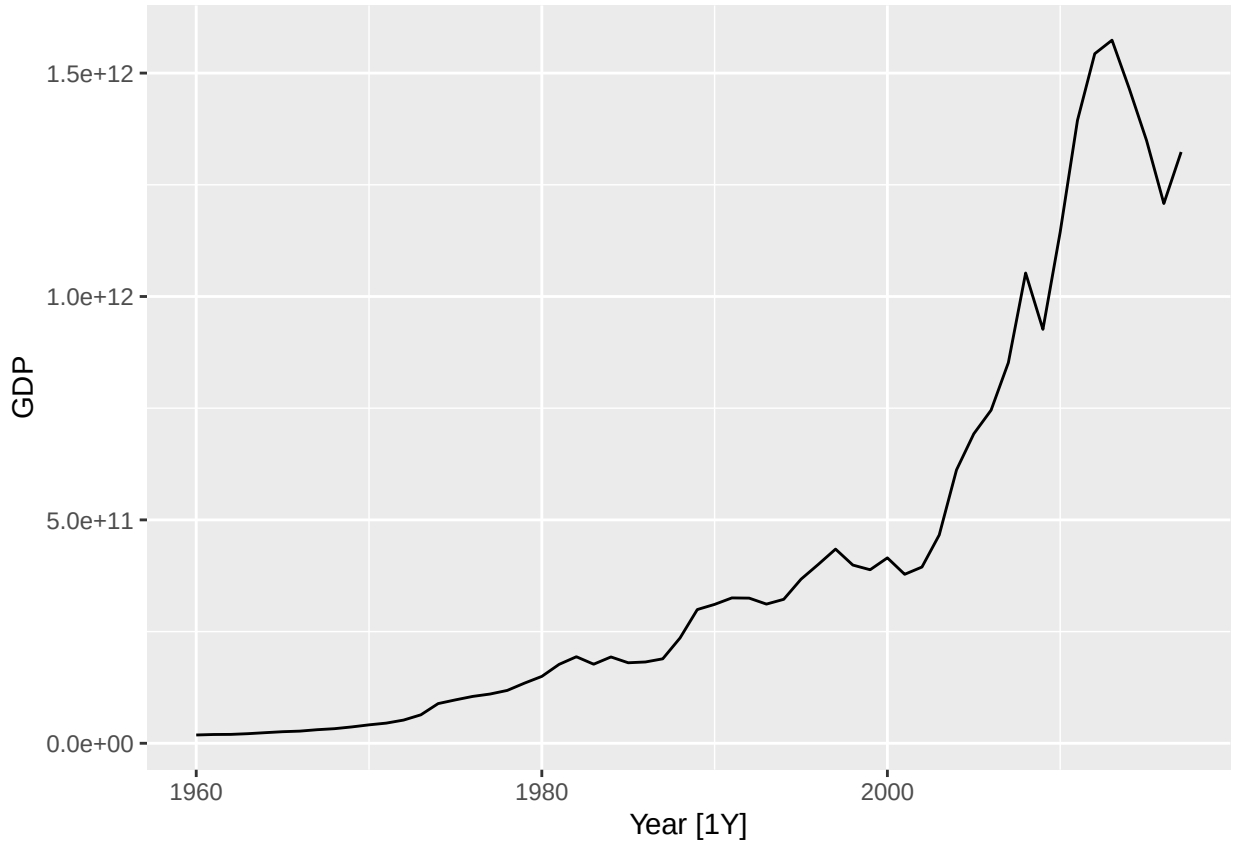
- tidyr
  - pivot_longer - Pivot data from wide to long

Time-Series tydiverce

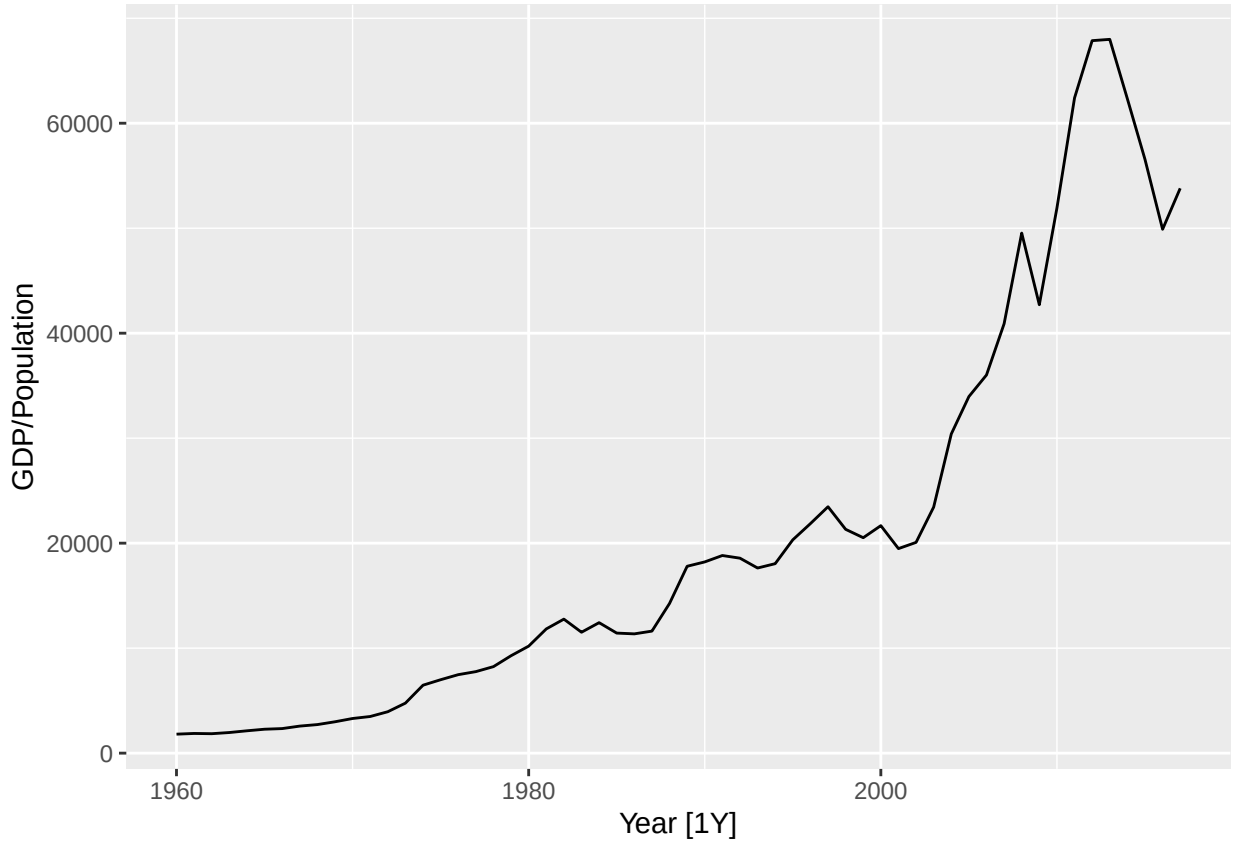
- tsibble - tidy temporal data frames and tools
  - index_by - set time index
- fable - forecast table
- fabletools - provides tools for building modelling packages, with a focus on time series forecasting
  - autoplot - also ggplot2 and feasts depending on object
  - autolayer - also ggplot2 and feasts depending on object
  - features - Extract features from a dataset
  - model - Estimate models
- feasts - Feature Extraction and Statistics for Time Series in tsibble format
  - autoplot - also ggplot2 and feasts depending on object
  - autolayer - also ggplot2 and feasts depending on object
  - guerrero - Guerrero's method for Box Cox lambda selection
  - STL - Multiple seasonal decomposition by Loess
  - gg_subseries

Adjustment

```
## GDP ------------------------------------------------------------------------
global_economy %>%
  filter(Country == "Australia") %>%
  autoplot(GDP)
```

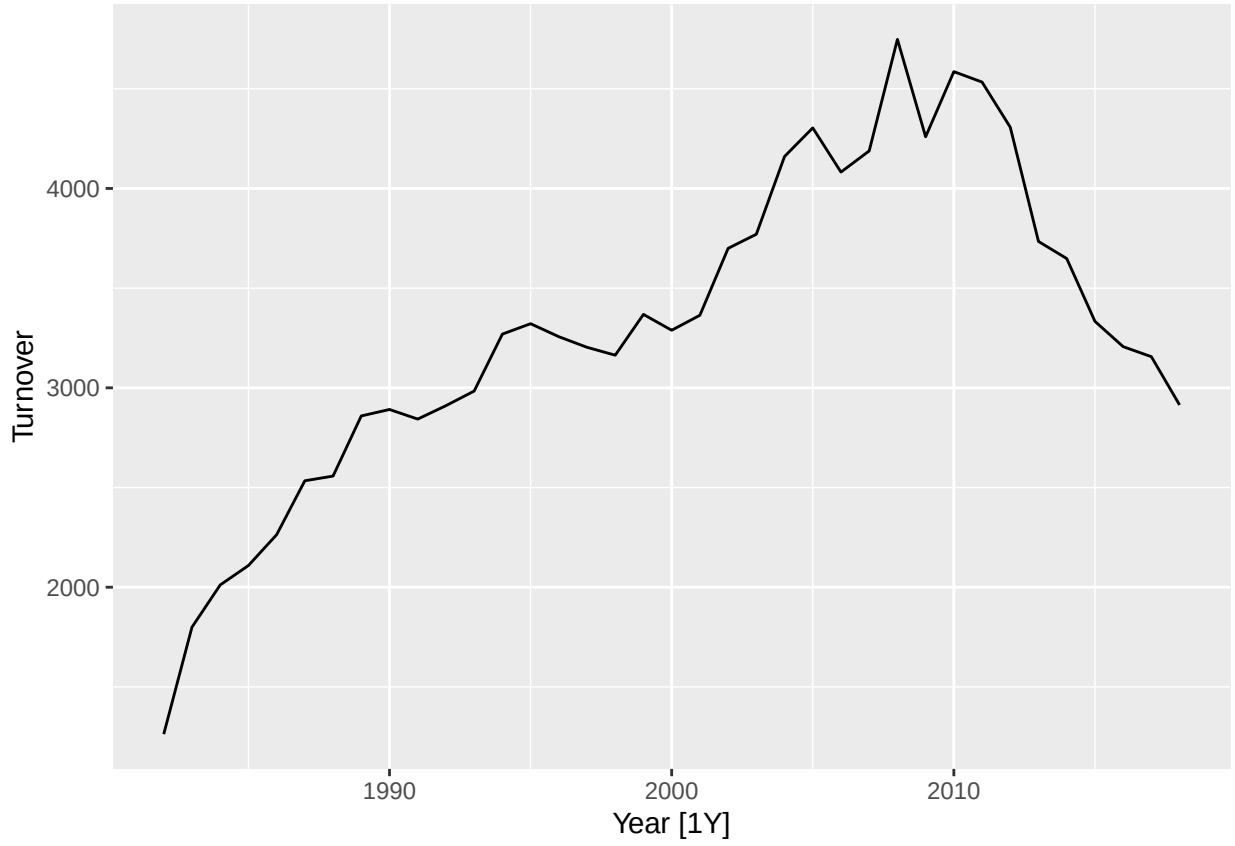



```
global_economy %>%
  filter(Country == "Australia") %>%
  autoplot(GDP / Population)
```




```
## Print retail adjusted by CPI ------------------------------------------------
print_retail <- aus_retail %>%
  filter(Industry == "Newspaper and book retailing") %>%
  group_by(Industry) %>%
  index_by(Year = year(Month)) %>%
  summarise(Turnover = sum(Turnover))

print_retail %>% autoplot(Turnover)
```

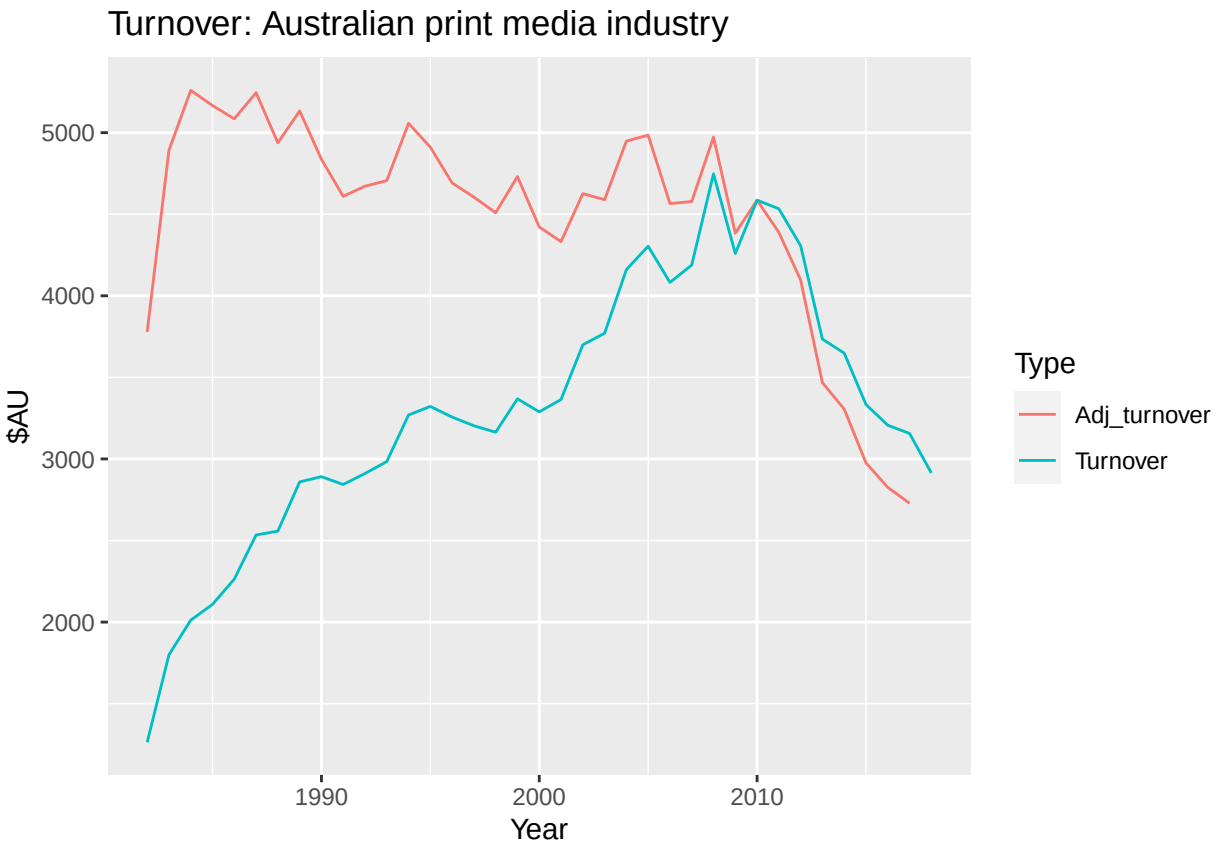



```
aus_economy <- global_economy %>%
  filter(Code == "AUS")

print_retail <- print_retail %>%
  left_join(aus_economy, by = "Year") %>%
  mutate(Adj_turnover = Turnover / CPI * 100) %>%
  pivot_longer(c(Turnover, Adj_turnover),
    names_to = "Type", values_to = "Turnover"
  )
```

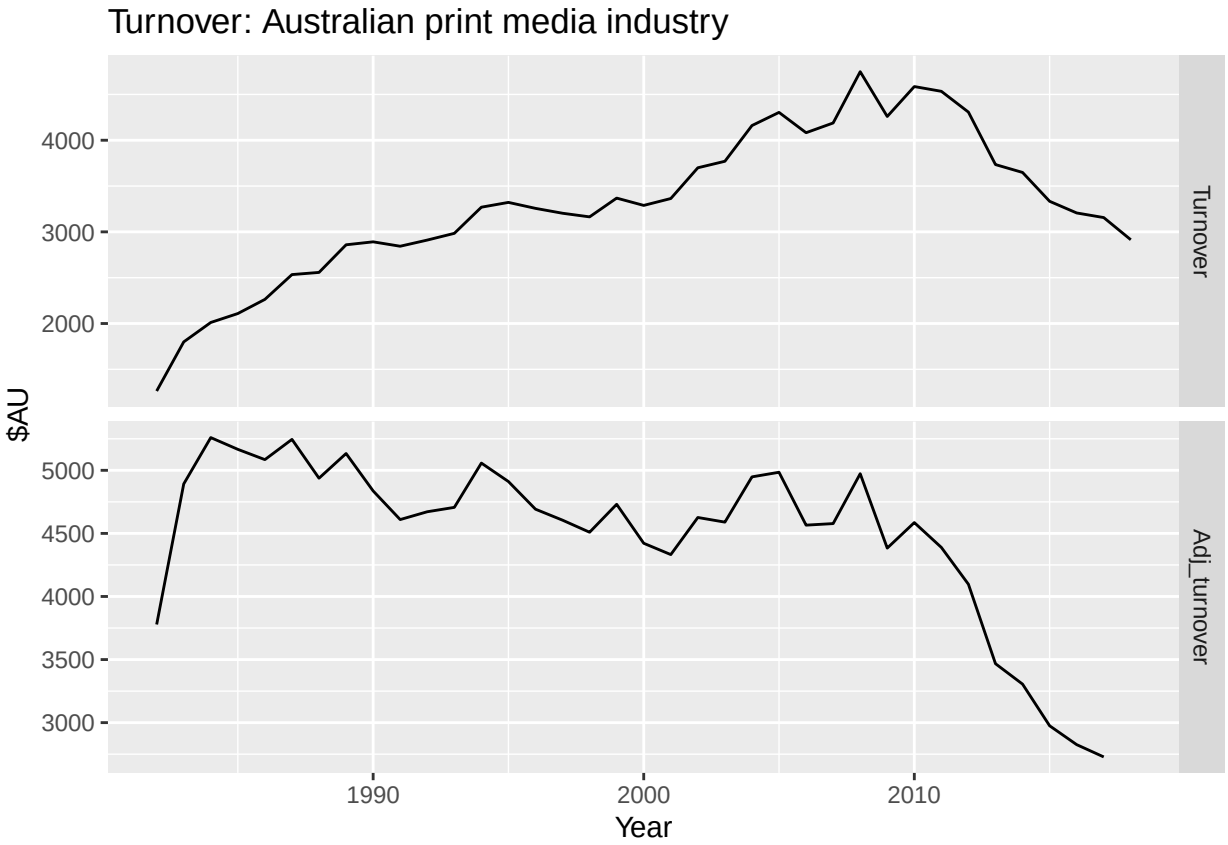
```
# Plot both on same graph
print_retail %>%
  ggplot(aes(x = Year, y = Turnover, col = Type)) +
  geom_line() +
  labs(
    title = "Turnover: Australian print media industry",
    y = "$AU"
  )
```

```
## Warning: Removed 1 row(s) containing missing values (geom_path).
```




```
# Use faceting
print_retail %>%
  mutate(Type = factor(Type,
    levels = c("Turnover", "Adj_turnover")
  )) %>%
  ggplot(aes(x = Year, y = Turnover)) +
  geom_line() +
  facet_grid(Type ~ ., scales = "free_y") +
  labs(
    title = "Turnover: Australian print media industry",
    y = "$AU"
  )
```

```
## Warning: Removed 1 row(s) containing missing values (geom_path).
```

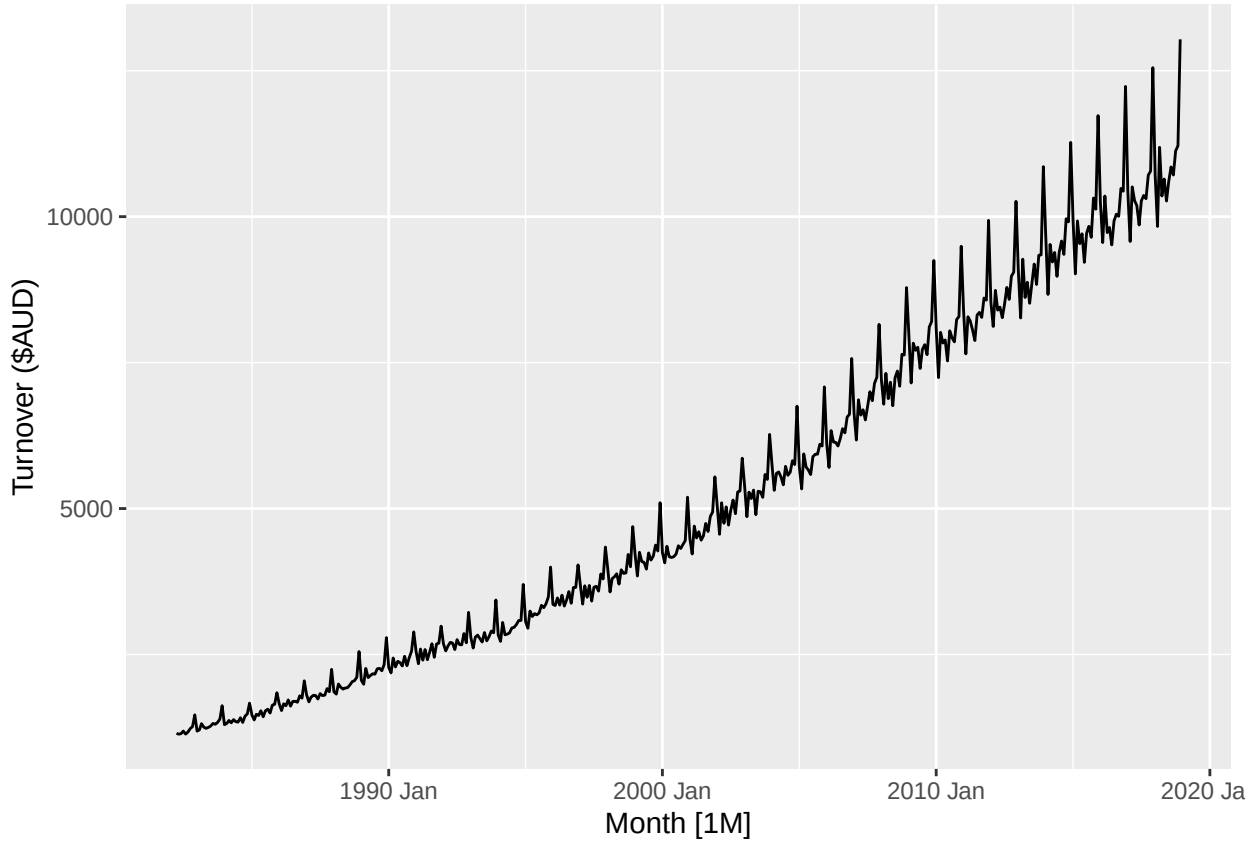


Math Transformation

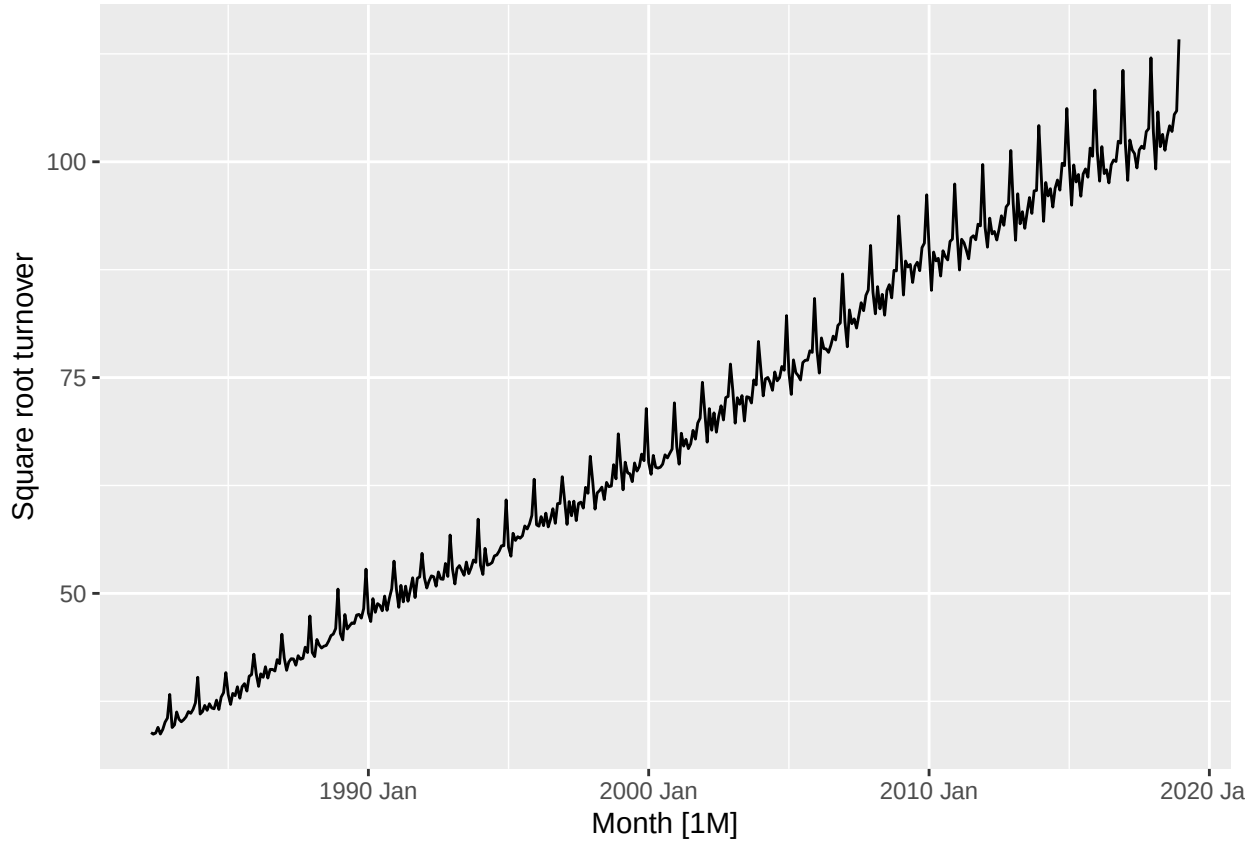
```
## Australian food retail -------------------------------------------------------

food <- aus_retail %>%
  filter(Industry == "Food retailing") %>%
  summarise(Turnover = sum(Turnover))
```

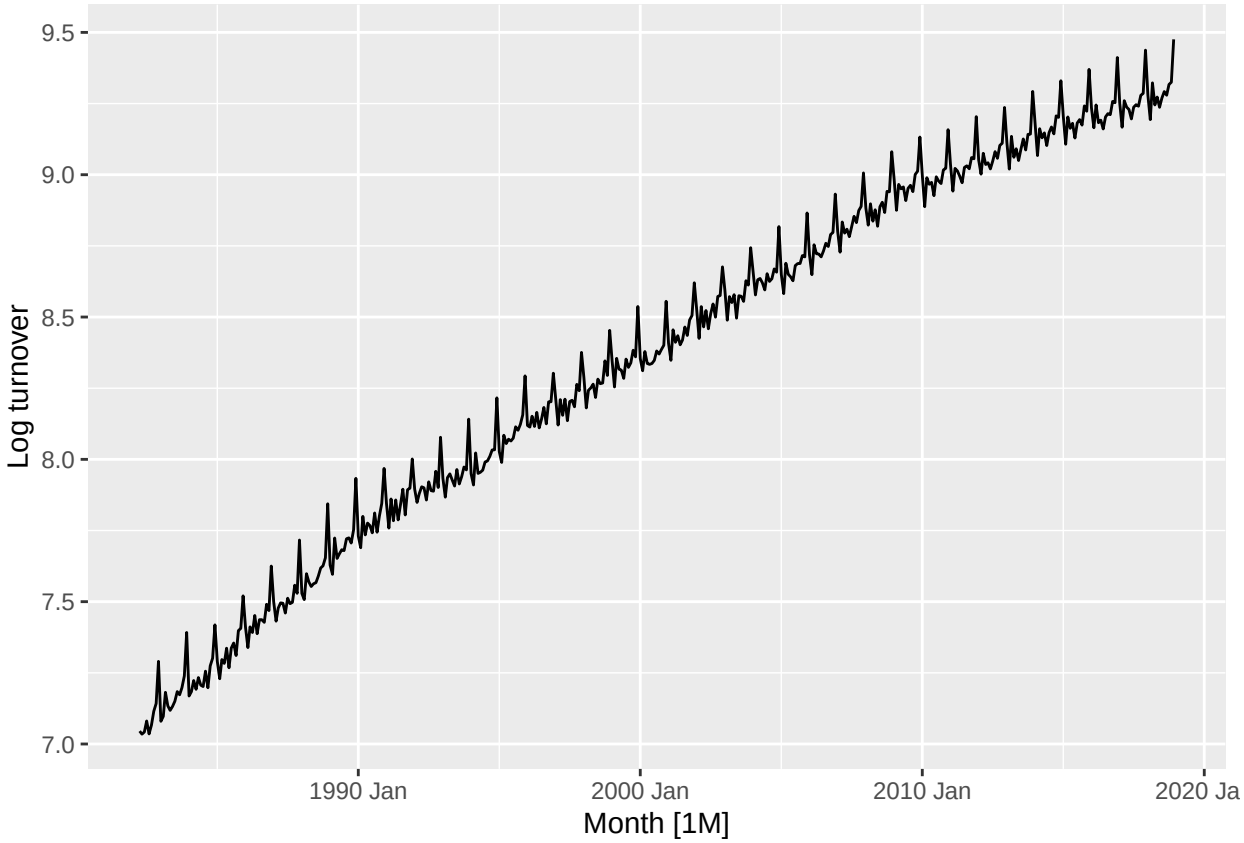
```
food %>% autoplot(Turnover) +
  labs(y = "Turnover ($AUD)")
```




```
food %>% autoplot(sqrt(Turnover)) +
  labs(y = "Square root turnover")
```



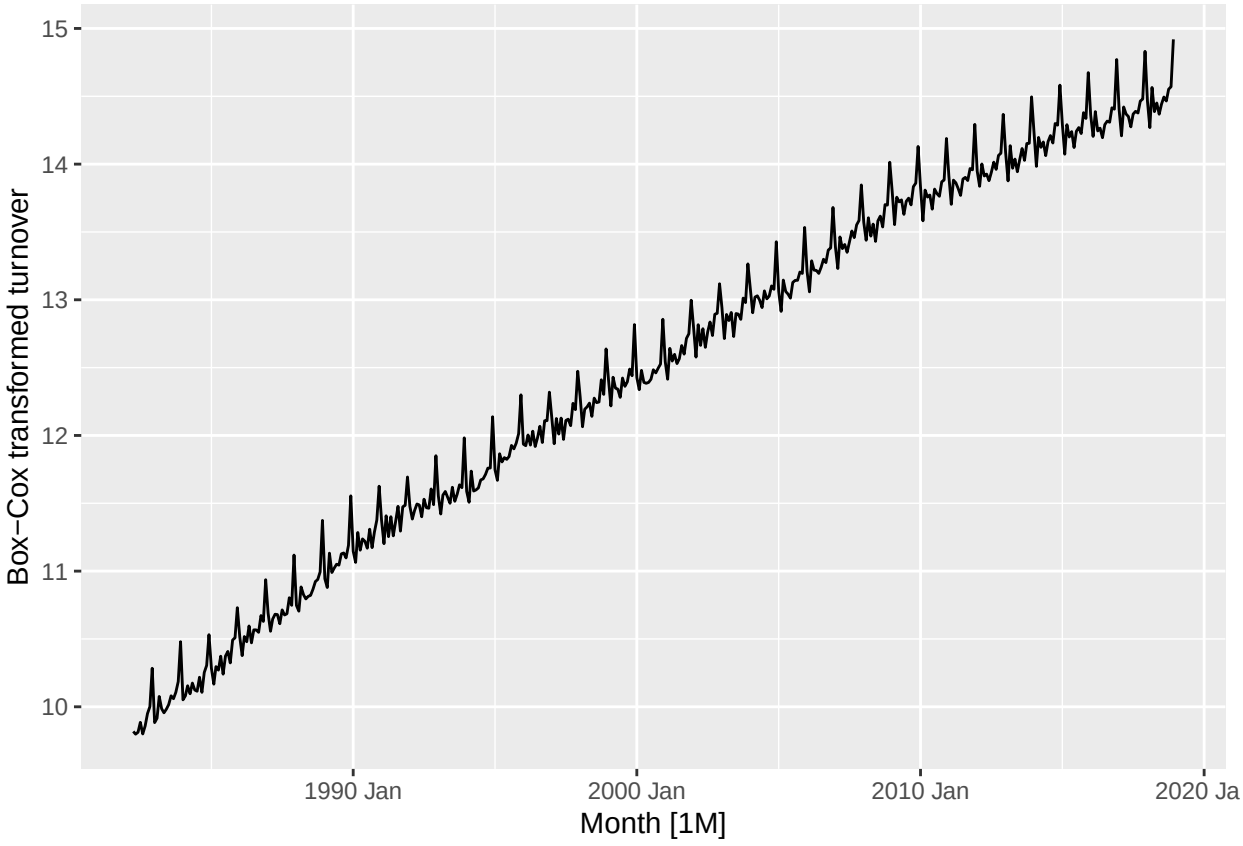

```
food %>% autoplot(log(Turnover)) +
  labs(y = "Log turnover")
```




```
food %>%
  features(Turnover, features = guerrero)
```

```
## # A tibble: 1 x 1
##   lambda_guerrero
##             <dbl>
## 1          0.0895
```

```
food %>% autoplot(box_cox(Turnover, 0.08952696)) +
  labs(y = "Box-Cox transformed turnover")
```

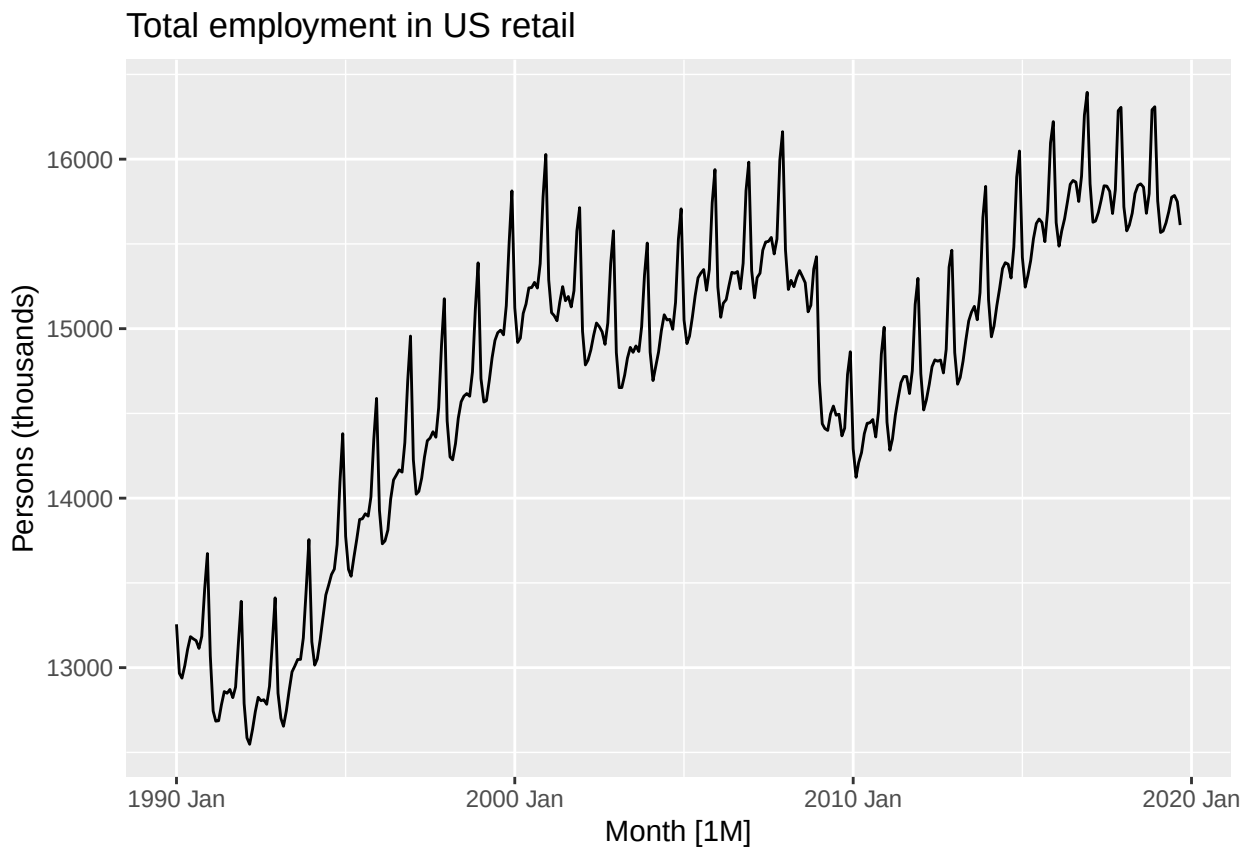


Decomposition

```
## US retail employment -----------------------------------------------------------

us_retail_employment <- us_employment %>%
  filter(year(Month) >= 1990, Title == "Retail Trade") %>%
  select(-Series_ID)

us_retail_employment %>%
  autoplot(Employed) +
  labs(
    y = "Persons (thousands)",
    title = "Total employment in US retail"
  )
```



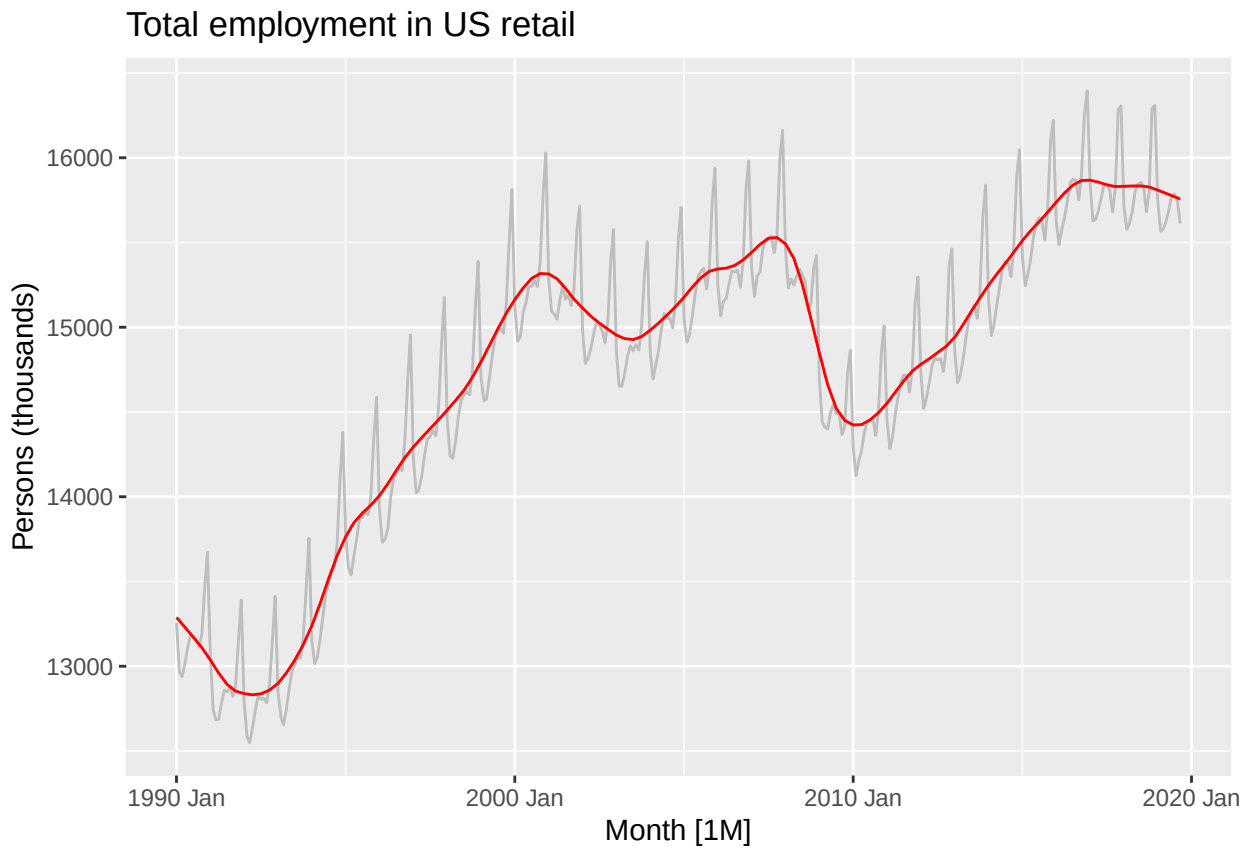

```
dcmp <- us_retail_employment %>%
  model(stl = STL(Employed))
components(dcmp)
```

```
## # A dable: 357 x 7 [1M]
## # Key:     .model [1]
## # :        Employed = trend + season_year + remainder
##    .model    Month Employed  trend season_year remainder season_adjust
##    <chr>     <mth>    <dbl>  <dbl>       <dbl>     <dbl>         <dbl>
##  1 stl    1990 Jan   13256. 13288.      -33.0      0.836        13289.
##  2 stl    1990 Feb   12966. 13269.     -258.     -44.6          13224.
##  3 stl    1990 Mar   12938. 13250.     -290.     -22.1          13228.
##  4 stl    1990 Apr   13012. 13231.     -220.       1.05         13232.
##  5 stl    1990 May   13108. 13211.     -114.      11.3          13223.
##  6 stl    1990 Jun   13183. 13192.      -24.3     15.5          13207.
##  7 stl    1990 Jul   13170. 13172.      -23.2     21.6          13193.
##  8 stl    1990 Aug   13160. 13151.       -9.52    17.8          13169.
##  9 stl    1990 Sep   13113. 13131.      -39.5     22.0          13153.
## 10 stl    1990 Oct   13185. 13110.       61.6     13.2          13124.
## # ... with 347 more rows
```

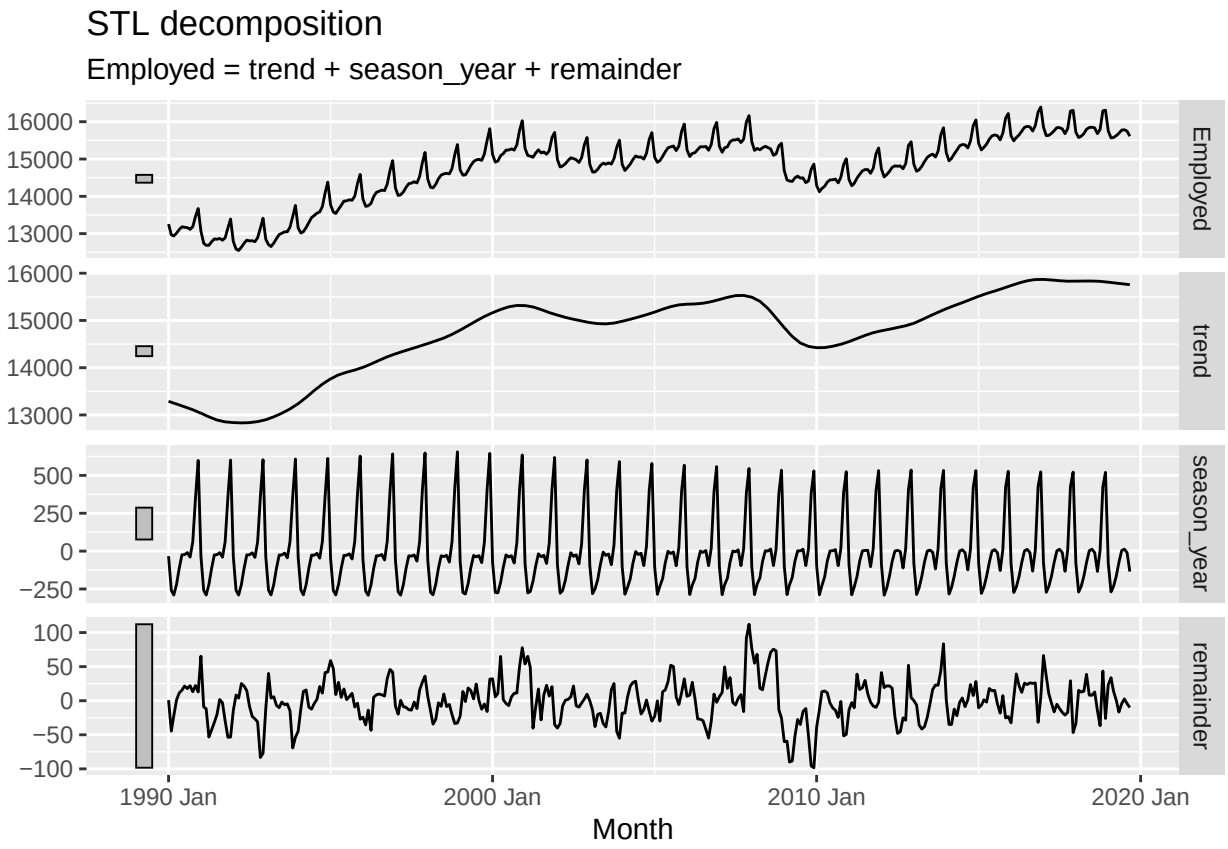
```
# %>% head()
```

```
us_retail_employment %>%
  autoplot(Employed, color = "gray") +
  autolayer(components(dcmp), trend, color = "red") +
  labs(
```

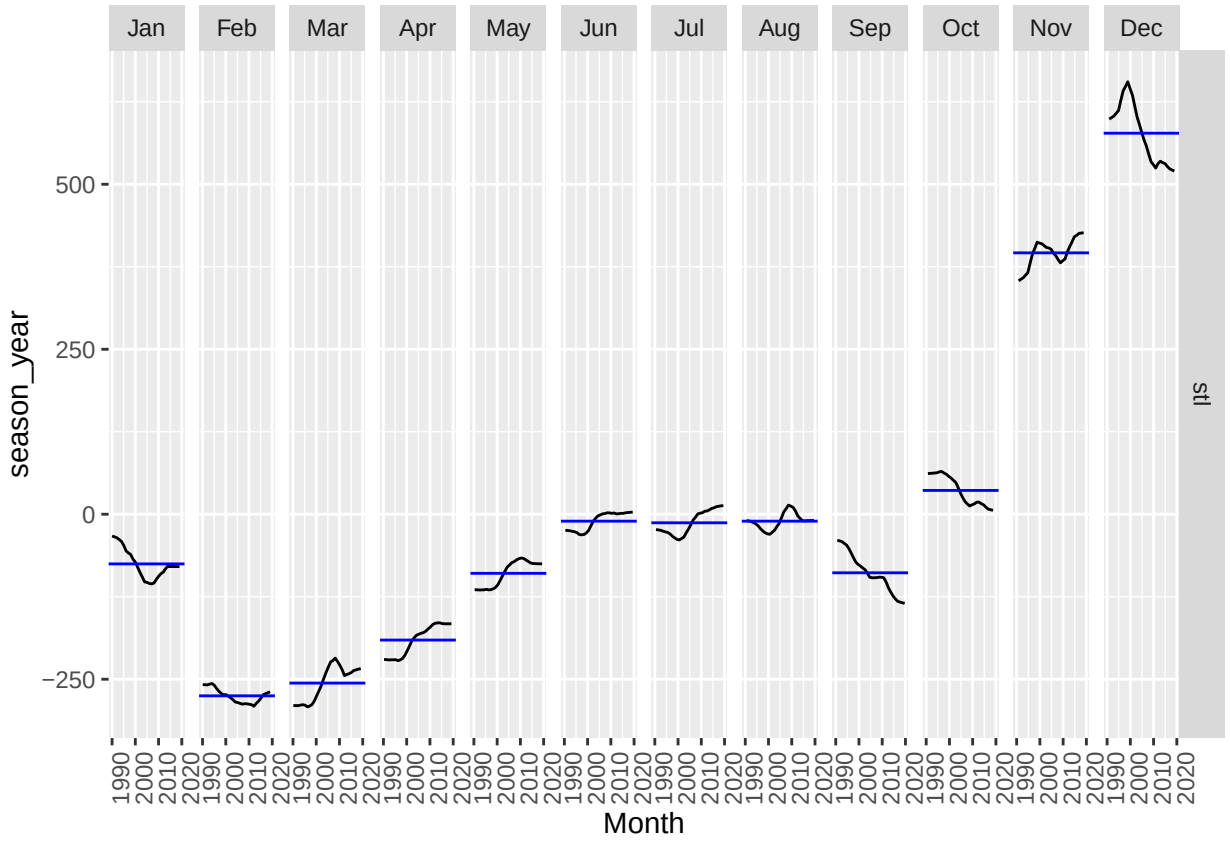
```
    y = "Persons (thousands)",
    title = "Total employment in US retail"
  )
```



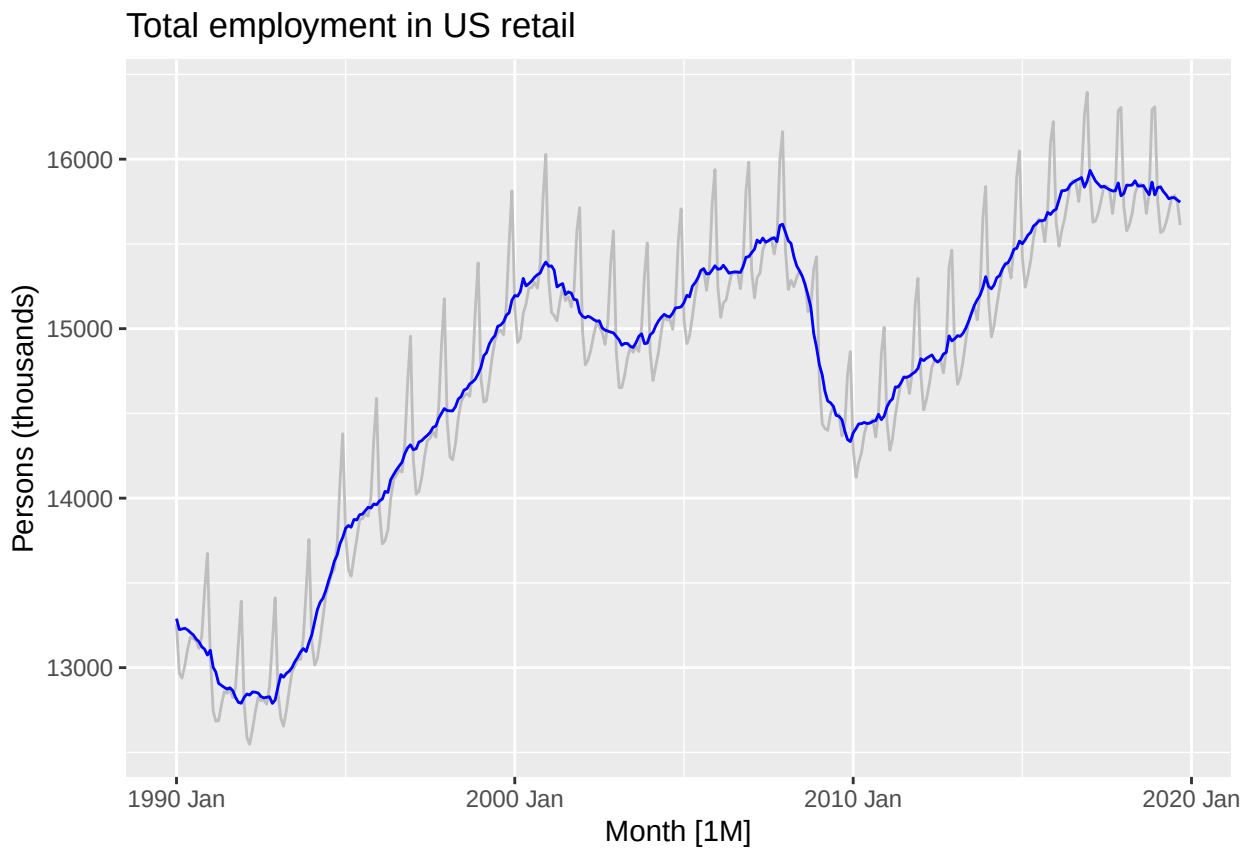

```
components(dcmp) %>% autoplot()
```



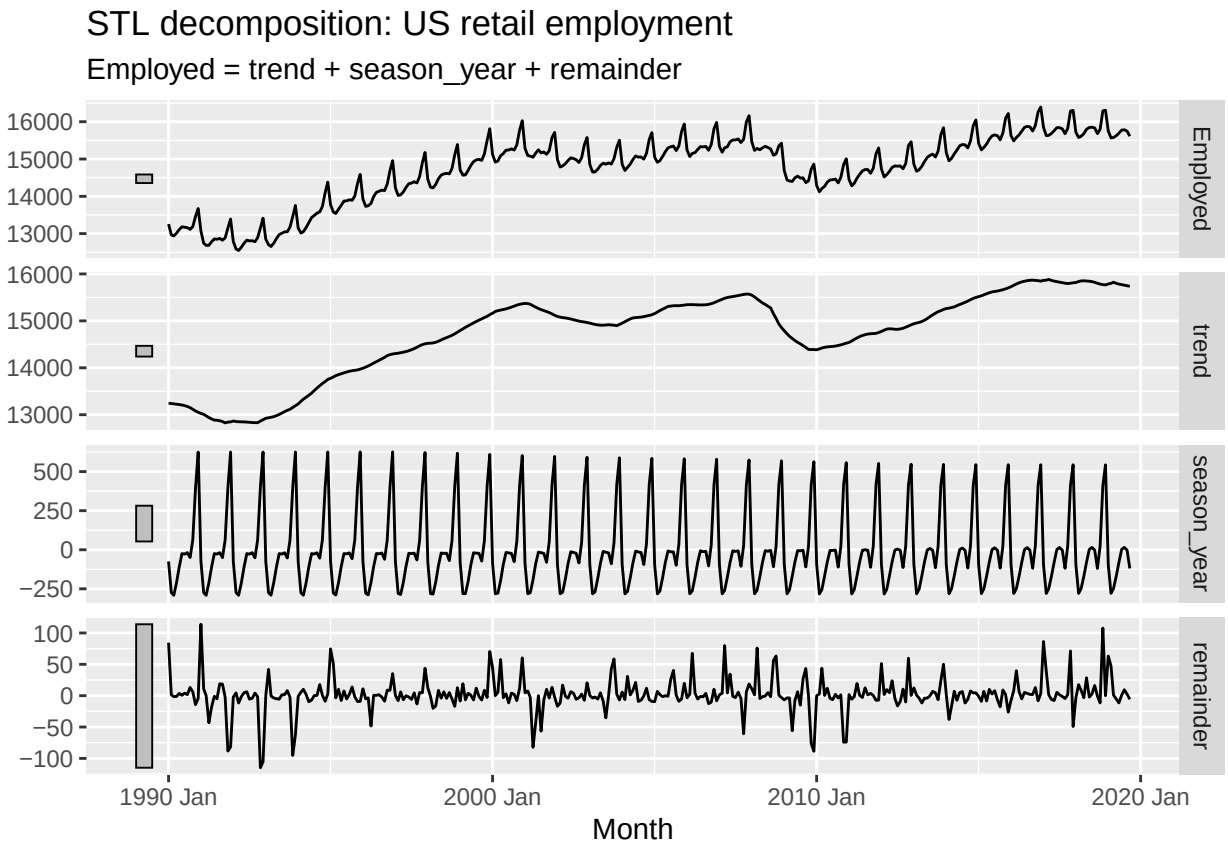

```
components(dcmp) %>% gg_subseries(season_year)
```




```
us_retail_employment %>%
  autoplot(Employed, color = "gray") +
  autolayer(components(dcmp), season_adjust, color = "blue") +
  labs(
    y = "Persons (thousands)",
    title = "Total employment in US retail"
  )
```




```
us_retail_employment %>%
  model(STL(Employed ~ season(window = 13) + trend(window = 7), robust = TRUE)) %>%
  components() %>%
  autoplot() +
  labs(title = "STL decomposition: US retail employment")
```


STL decomposition: US retail employment
Employed = trend + season_year + remainder
Employed
trend
season_year
remainder
16000
15000
14000
13000
500
250
0
−250
100
50
−50
−100
1990 Jan
2000 Jan
2010 Jan
2020 Jan
Month