

A feast of time series tools

Rob J Hyndman &
Mitchell O'Hara-Wild

June 2019

Outline

- 1 Overview
- 2 Tsibbles
- 3 Graphics
- 4 Decompositions
- 5 Features

Outline

1 Overview

2 Tsibbles

3 Graphics

4 Decompositions

5 Features





Feature Extraction And Statistics for Time Series

- works with tidy temporal data provided by the tsibble package.
- produces time series features, decompositions, statistical summaries and visualisations.

Outline

1 Overview

2 Tsibbles

3 Graphics

4 Decompositions

5 Features

tsibble objects

```
global_economy
```

```
## # A tsibble: 15,150 x 6 [1Y]
## # Key:      Country [263]
##   Year Country      GDP Imports Exports Population
##   <dbl> <fct>      <dbl>   <dbl>   <dbl>      <dbl>
## 1  1960 Afghanistan 5377777811.    7.02    4.13    8996351
## 2  1961 Afghanistan 5488888896.    8.10    4.45    9166764
## 3  1962 Afghanistan 5466666678.    9.35    4.88    9345868
## 4  1963 Afghanistan 7511111191.   16.9     9.17    9533954
## 5  1964 Afghanistan 8000000044.   18.1     8.89    9731361
## 6  1965 Afghanistan 10066666638.  21.4    11.3    9938414
## 7  1966 Afghanistan 13999999967.  18.6     8.57   10152331
## 8  1967 Afghanistan 16733333418.  14.2     6.77   10372630
## 9  1968 Afghanistan 13733333367.  15.2     8.90   10604346
## 10 1969 Afghanistan 14088888922.  15.0    10.1   10854428
## # ... with 15,140 more rows
```

tsibble objects

```
global_economy
```

```
## # A tsibble: 15,150 x 6 [1Y]
```

```
## # Key:          Country [263]
```

```
##      Year Country      GDP Imports Exports Population
##      Index <fct>      <dbl>   <dbl>   <dbl>         <dbl>
##  1  1960 Afghanistan 5377777811.    7.02    4.13    8996351
##  2  1961 Afghanistan 5488888896.    8.10    4.45    9166764
##  3  1962 Afghanistan 5466666678.    9.35    4.88    9345868
##  4  1963 Afghanistan 7511111191.   16.9    9.17    9533954
##  5  1964 Afghanistan 8000000044.   18.1    8.89    9731361
##  6  1965 Afghanistan 10066666638.  21.4   11.3    9938414
##  7  1966 Afghanistan 13999999967.  18.6    8.57   10152331
##  8  1967 Afghanistan 16733333418.  14.2    6.77   10372630
##  9  1968 Afghanistan 13733333367.  15.2    8.90   10604346
## 10  1969 Afghanistan 14088888922.  15.0   10.1   10854428
## # ... with 15,140 more rows
```


tsibble objects

```
global_economy
```

```
## # A tsibble: 15,150 x 6 [1Y]
```

```
## # Key:      Country [263]
```

```
##      Year Country      GDP Imports Exports Population
##      Index  Key      <dbl>   <dbl>   <dbl>         <dbl>
##  1  1960 Afghanistan 5377777811.    7.02    4.13    8996351
##  2  1961 Afghanistan 5488888896.    8.10    4.45    9166764
##  3  1962 Afghanistan 5466666678.    9.35    4.88    9345868
##  4  1963 Afghanistan 7511111191.   16.9    9.17    9533954
##  5  1964 Afghanistan 8000000044.   18.1    8.89    9731361
##  6  1965 Afghanistan 10066666638.  21.4   11.3    9938414
##  7  1966 Afghanistan 13999999967.  18.6    8.57   10152331
##  8  1967 Afghanistan 16733333418.  14.2    6.77   10372630
##  9  1968 Afghanistan 13733333367.  15.2    8.90   10604346
## 10  1969 Afghanistan 14088888922.  15.0   10.1   10854428
## # ... with 15,140 more rows
```

tsibble objects

```
global_economy
```

```
## # A tsibble: 15,150 x 6 [1Y]
```

```
## # Key:      Country [263]
```

```
##      Year Country      GDP Imports Exports Population
```

```
##      Index  Key      Measured variables
```

```
## 1  1960 Afghanistan 5377777811.    7.02    4.13    8996351
```

```
## 2  1961 Afghanistan 5488888896.    8.10    4.45    9166764
```

```
## 3  1962 Afghanistan 546666678.    9.35    4.88    9345868
```

```
## 4  1963 Afghanistan 7511111191.   16.9    9.17    9533954
```

```
## 5  1964 Afghanistan 8000000044.   18.1    8.89    9731361
```

```
## 6  1965 Afghanistan 1006666638.   21.4   11.3    9938414
```

```
## 7  1966 Afghanistan 1399999967.   18.6    8.57   10152331
```

```
## 8  1967 Afghanistan 1673333418.   14.2    6.77   10372630
```

```
## 9  1968 Afghanistan 1373333367.   15.2    8.90   10604346
```

```
## 10 1969 Afghanistan 1408888922.   15.0   10.1   10854428
```

```
## # ... with 15,140 more rows
```

tsibble objects

```
aus_retail
```

```
## # A tsibble: 64,088 x 4 [1M]
## # Key:      State, Industry [152]
##           Month State Industry      Turnover
##           <mth> <chr> <chr>      <dbl>
## 1  1982 Apr ACT    Cafes, restaurants and catering~  4.4
## 2  1982 May ACT    Cafes, restaurants and catering~  3.4
## 3  1982 Jun ACT    Cafes, restaurants and catering~  3.6
## 4  1982 Jul ACT    Cafes, restaurants and catering~   4
## 5  1982 Aug ACT    Cafes, restaurants and catering~  3.6
## 6  1982 Sep ACT    Cafes, restaurants and catering~  4.2
## 7  1982 Oct ACT    Cafes, restaurants and catering~  4.8
## 8  1982 Nov ACT    Cafes, restaurants and catering~  5.4
## 9  1982 Dec ACT    Cafes, restaurants and catering~  6.9
## 10 1983 Jan ACT    Cafes, restaurants and catering~  3.8
## # ... with 64,078 more rows
```

tsibble objects

```
aus_retail
```

```
## # A tsibble: 64,088 x 4 [1M]
## # Key:      State, Industry [152]
##           Month State Industry      Turnover
##           Index  <chr> <chr>      <dbl>
## 1  1982 Apr ACT    Cafes, restaurants and catering~  4.4
## 2  1982 May ACT    Cafes, restaurants and catering~  3.4
## 3  1982 Jun ACT    Cafes, restaurants and catering~  3.6
## 4  1982 Jul ACT    Cafes, restaurants and catering~   4
## 5  1982 Aug ACT    Cafes, restaurants and catering~  3.6
## 6  1982 Sep ACT    Cafes, restaurants and catering~  4.2
## 7  1982 Oct ACT    Cafes, restaurants and catering~  4.8
## 8  1982 Nov ACT    Cafes, restaurants and catering~  5.4
## 9  1982 Dec ACT    Cafes, restaurants and catering~  6.9
## 10 1983 Jan ACT    Cafes, restaurants and catering~  3.8
## # ... with 64,078 more rows
```

tsibble objects

```
aus_retail
```

```
## # A tsibble: 64,088 x 4 [1M]
## # Key:      State, Industry [152]
##           Month State Industry      Turnover
##           Index  Keys      <dbl>
## 1  1982 Apr ACT    Cafes, restaurants and catering~  4.4
## 2  1982 May ACT    Cafes, restaurants and catering~  3.4
## 3  1982 Jun ACT    Cafes, restaurants and catering~  3.6
## 4  1982 Jul ACT    Cafes, restaurants and catering~   4
## 5  1982 Aug ACT    Cafes, restaurants and catering~  3.6
## 6  1982 Sep ACT    Cafes, restaurants and catering~  4.2
## 7  1982 Oct ACT    Cafes, restaurants and catering~  4.8
## 8  1982 Nov ACT    Cafes, restaurants and catering~  5.4
## 9  1982 Dec ACT    Cafes, restaurants and catering~  6.9
## 10 1983 Jan ACT    Cafes, restaurants and catering~  3.8
## # ... with 64,078 more rows
```

tsibble objects

```
aus_retail
```

```
## # A tsibble: 64,088 x 4 [1M]
```

```
## # Key:      State, Industry [152]
```

```
##           Month State Industry Turnover
```

```
##           Index  Keys Measure
```

```
## 1 1982 Apr ACT Cafes, restaurants and catering~ 4.4
```

```
## 2 1982 May ACT Cafes, restaurants and catering~ 3.4
```

```
## 3 1982 Jun ACT Cafes, restaurants and catering~ 3.6
```

```
## 4 1982 Jul ACT Cafes, restaurants and catering~ 4
```

```
## 5 1982 Aug ACT Cafes, restaurants and catering~ 3.6
```

```
## 6 1982 Sep ACT Cafes, restaurants and catering~ 4.2
```

```
## 7 1982 Oct ACT Cafes, restaurants and catering~ 4.8
```

```
## 8 1982 Nov ACT Cafes, restaurants and catering~ 5.4
```

```
## 9 1982 Dec ACT Cafes, restaurants and catering~ 6.9
```

```
## 10 1983 Jan ACT Cafes, restaurants and catering~ 3.8
```

```
## # ... with 64,078 more rows
```

Pharmaceutical expenditure

```
pharma <- aus_retail %>%  
  filter(  
    # State %in% c("VIC", "NSW"),  
    str_detect(Industry, "Pharmaceutical")  
  ) %>%  
  select(-Industry)
```

```
## # A tsibble: 3,432 x 3 [1M]
```

```
## # Key:      State [8]
```

```
##      Month State Turnover  
##      <mtm> <chr>    <dbl>
```

```
## 1 1982 Apr ACT      2.5
```

```
## 2 1982 May ACT      2.5
```

```
## 3 1982 Jun ACT      2.5
```

```
## 4 1982 Jul ACT      2.6
```

```
## 5 1982 Aug ACT      2.8
```

```
## 6 1982 Sep ACT      2.8
```

```
## 7 1982 Oct ACT      2.6
```

```
## 8 1982 Nov ACT      3.1
```

Outline

1 Overview

2 Tsibbles

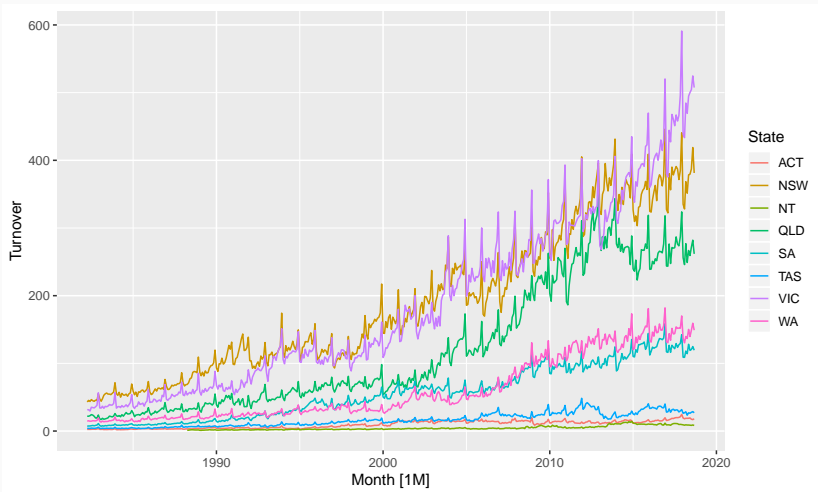
3 Graphics

4 Decompositions

5 Features

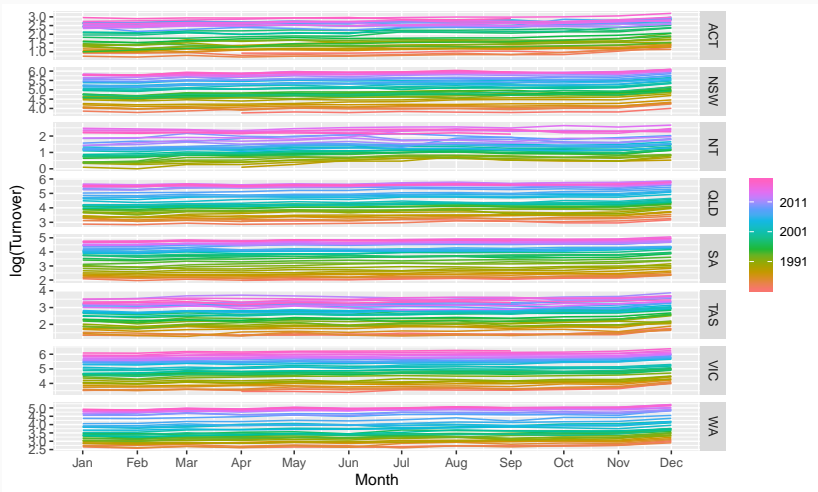
Time plots

```
pharma %>% autoplot(Turnover)
```



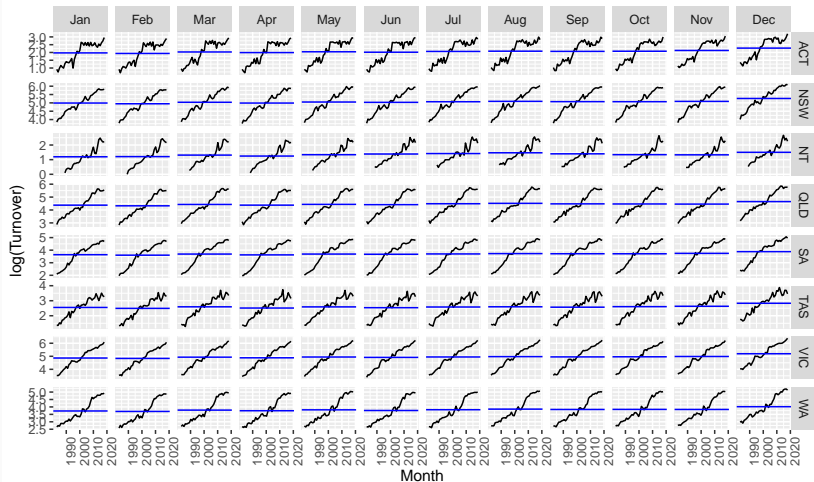
Season plots

```
pharma %>% gg_season(log(Turnover))
```



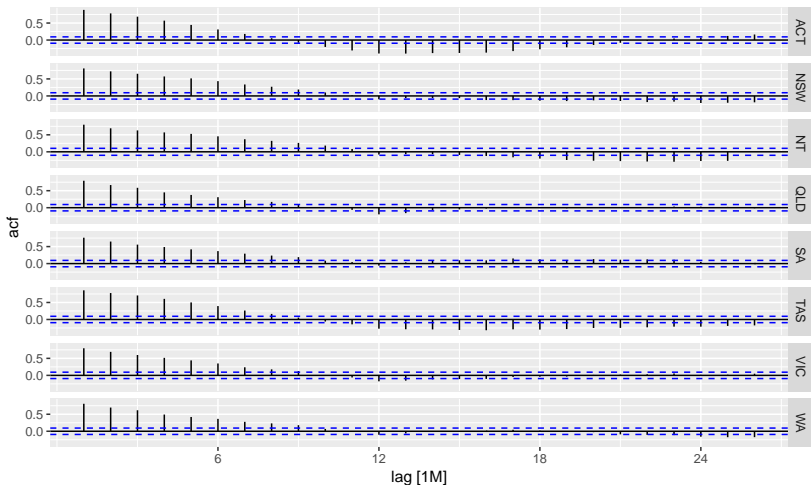
Graphics

```
pharma %>% gg_subseries(log(Turnover))
```



Graphics

```
pharma %>% ACF(difference(log(Turnover), 12)) %>% autoplot()
```



Outline

- 1 Overview
- 2 Tsibbles
- 3 Graphics
- 4 Decompositions**
- 5 Features

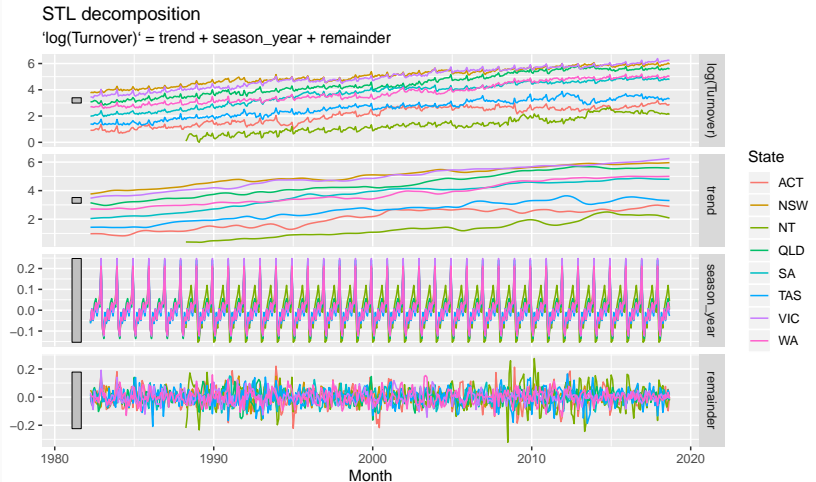
Decompositions

The feasts package supports four common time series decomposition methods:

- Classical decomposition
- STL decomposition
- X11 decomposition
- X-13ARIMA-SEATS decomposition

Decompositions

```
pharma %>% STL(log(Turnover) ~ season(window = "periodic")) %>%  
  autoplot()
```



Outline

- 1 Overview
- 2 Tsibbles
- 3 Graphics
- 4 Decompositions
- 5 Features**

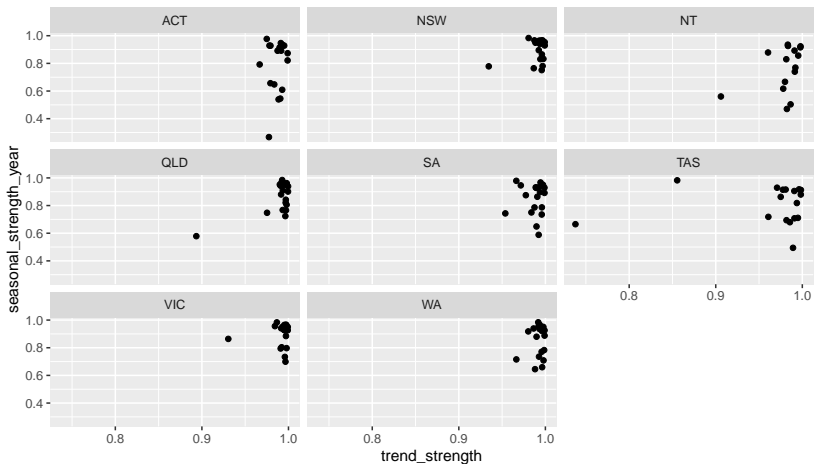
Feature extraction and statistics

```
aus_retail %>% features(Turnover, features_stl)
```

```
## # A tibble: 152 x 9
##   State Industry trend_strength seasonal_streng~ spike
##   <chr> <chr>          <dbl>          <dbl>    <dbl>
## 1 ACT   Cafes, ~           0.989           0.540 6.17e-5
## 2 ACT   Cafes, ~           0.993           0.609 1.16e-4
## 3 ACT   Clothin~          0.990           0.914 4.93e-6
## 4 ACT   Clothin~          0.991           0.947 2.60e-5
## 5 ACT   Departm~          0.975           0.977 2.85e-5
## 6 ACT   Electri~          0.991           0.925 3.19e-5
## 7 ACT   Food re~          0.999           0.874 2.50e-4
## 8 ACT   Footwea~          0.978           0.929 7.75e-6
## 9 ACT   Furnitu~          0.979           0.657 4.87e-5
## 10 ACT  Hardwar~          0.992           0.892 1.56e-5
## # ... with 142 more rows, and 4 more variables:
## #   linearity <dbl>, curvature <dbl>,
## #   seasonal_peak_year <dbl>, seasonal_trough_year <dbl>
```

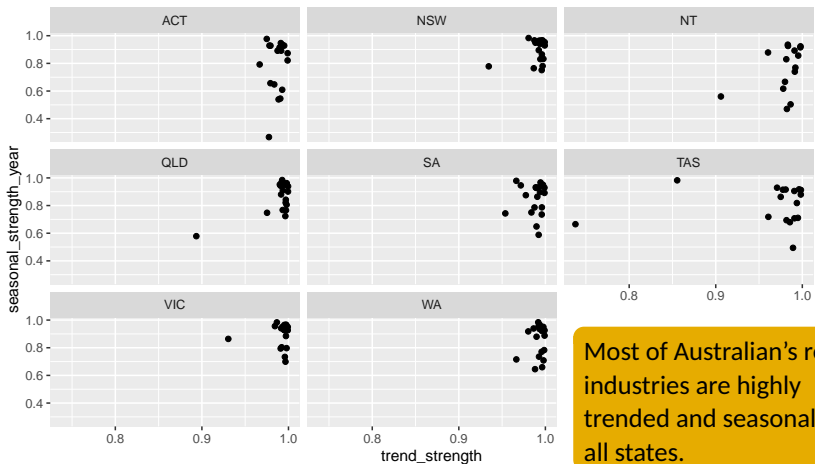
Feature extraction and statistics

```
aus_retail %>% features(Turnover, features_stl) %>%  
  ggplot(aes(x = trend_strength, y = seasonal_strength_year)) +  
  geom_point() + facet_wrap(vars(State))
```



Feature extraction and statistics

```
aus_retail %>% features(Turnover, features_stl) %>%  
  ggplot(aes(x = trend_strength, y = seasonal_strength_year)) +  
  geom_point() + facet_wrap(vars(State))
```



Most of Australian's retail industries are highly trended and seasonal for all states.

Feature extraction and statistics

Find the most seasonal time series:

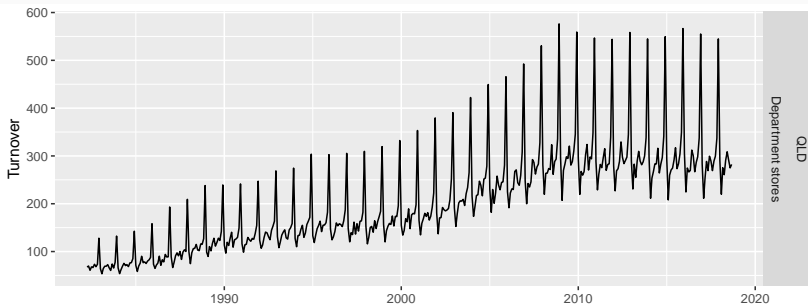
```
most_seasonal <- aus_retail %>%  
  features(Turnover, features_stl) %>%  
  filter(seasonal_strength_year == max(seasonal_strength_year))
```

Feature extraction and statistics

Find the most seasonal time series:

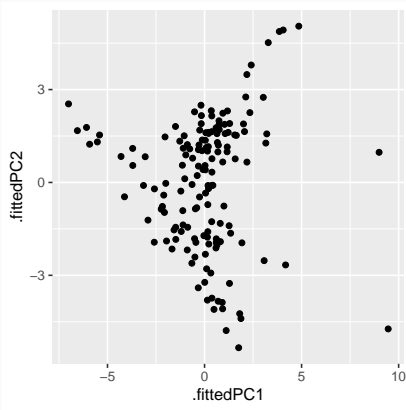
```
most_seasonal <- aus_retail %>%  
  features(Turnover, features_stl) %>%  
  filter(seasonal_strength_year == max(seasonal_strength_year))
```

```
aus_retail %>%  
  right_join(most_seasonal, by = c("State", "Industry")) %>%  
  ggplot(aes(x = Month, y = Turnover)) + geom_line() +  
  facet_grid(vars(State, Industry))
```



Feature extraction and statistics

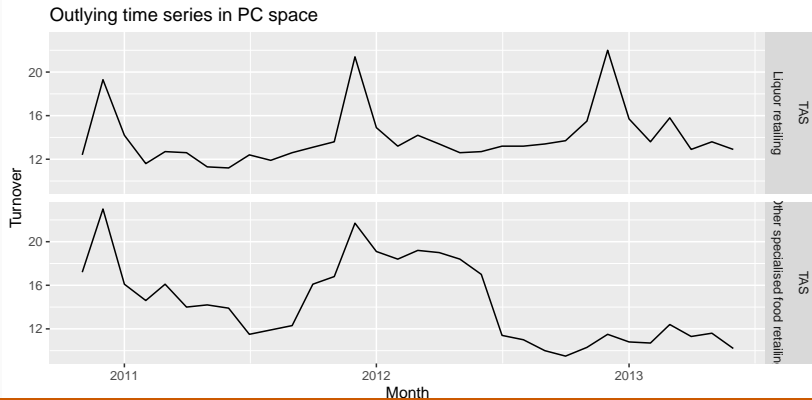
```
retail_features <- aus_retail %>%  
  features(Turnover, list(features_stl, features_acf, features_pacf))  
pcs <- retail_features %>% select(-State, -Industry) %>%  
  prcomp(scale=TRUE) %>% augment(retail_features)  
pcs %>% ggplot(aes(x=.fittedPC1, y=.fittedPC2)) + geom_point()
```



Principal components
based on STL, ACF and
PACF features

Feature extraction and statistics

```
outliers <- pcs %>% filter(.fittedPC1 > 5)
aus_retail %>%
  right_join(outliers, by = c("State", "Industry")) %>%
  ggplot(aes(x = Month, y = Turnover)) + geom_line() +
  facet_grid(vars(State, Industry)) +
  ggtitle("Outlying time series in PC space")
```



Acknowledgements



Mitchell O'Hara-Wild



Earo Wang

feasts.tidyverts.org
robjhyndman.com