MONASH University

MONASH BUSINESS SCHOOL

High dimensional time series analysis

tsibble feasts fable

5. Forecast reconciliation

robjhyndman.com/hdtsa

Outline

Outline

Australian Pharmaceutical Benefits Scheme

PBS sales

```
PBS
```

```
## # A tsibble: 65,219 x 9 [1M]
## # Key:       Concession, Type, ATC1, ATC2 [336]
##         Month Concession Type  ATC1  ATC1_desc ATC2
##         <mth> <chr>      <chr> <chr> <chr>     <chr>
##  1   1991 Jul Concessio~ Co-p~ A     Alimenta~ A01
##  2   1991 Aug Concessio~ Co-p~ A     Alimenta~ A01
##  3   1991 Sep Concessio~ Co-p~ A     Alimenta~ A01
##  4   1991 Oct Concessio~ Co-p~ A     Alimenta~ A01
##  5   1991 Nov Concessio~ Co-p~ A     Alimenta~ A01
##  6   1991 Dec Concessio~ Co-p~ A     Alimenta~ A01
##  7   1992 Jan Concessio~ Co-p~ A     Alimenta~ A01
##  8   1992 Feb Concessio~ Co-p~ A     Alimenta~ A01
##  9   1992 Mar Concessio~ Co-p~ A     Alimenta~ A01
## 10   1992 Apr Concessio~ Co-p~ A     Alimenta~ A01
## # ... with 65,209 more rows, and 3 more variables:
## #   ATC2_desc <chr>, Scripts <dbl>, Cost <dbl>
```

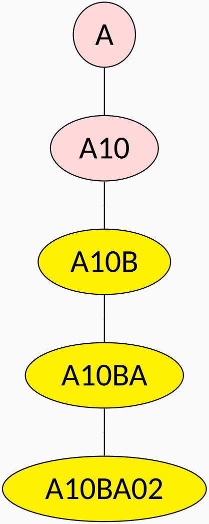
ATC drug classification

A	Alimentary tract and metabolism
B	Blood and blood forming organs
C	Cardiovascular system
D	Dermatologicals
G	Genito-urinary system and sex hormones
H	Systemic hormonal preparations, excluding sex hormones and insulins
J	Anti-infectives for systemic use
L	Antineoplastic and immunomodulating agents
M	Musculo-skeletal system
N	Nervous system
P	Antiparasitic products, insecticides and repellents
R	Respiratory system
S	Sensory organs
V	Various

ATC drug classification

ATC1: 14 classes

ATC2: 84 classes



Alimentary tract and metabolism

Drugs used in diabetes

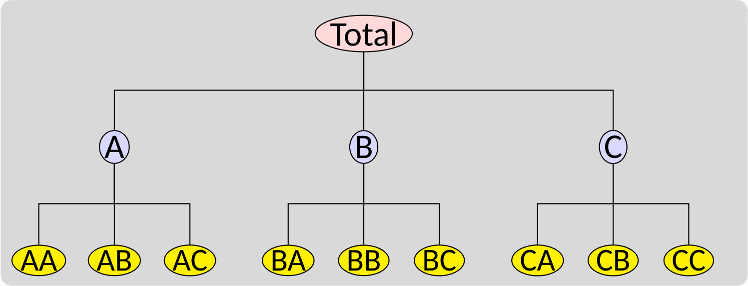
Blood glucose lowering drugs

Biguanides

Metformin

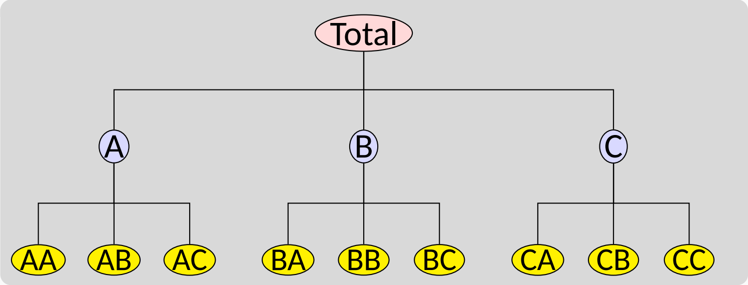
Hierarchical time series

A **hierarchical time series** is a collection of several time series that are linked together in a hierarchical structure.



Hierarchical time series

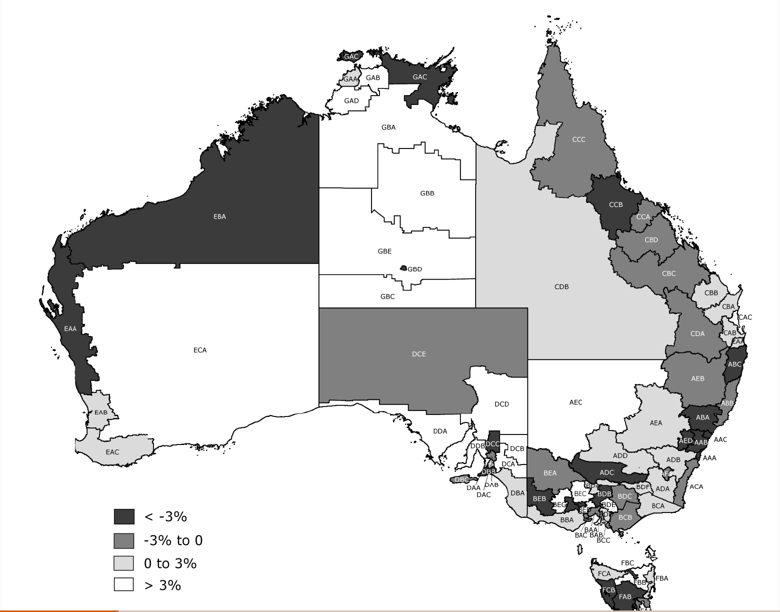
A **hierarchical time series** is a collection of several time series that are linked together in a hierarchical structure.



Examples

- PBS sales by ATC groups
- Tourism demand by states, zones, regions

Australian tourism



Australian tourism

```
tourism
```

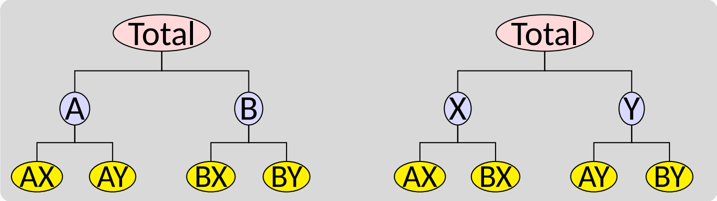
```
## # A tsibble: 24,320 x 5 [1Q]
## # Key:       Region, State, Purpose [304]
##    Quarter Region   State           Purpose  Trips
##      <qtr> <chr>    <chr>           <chr>    <dbl>
##  1 1998 Q1 Adelaide South Australia Business  135.
##  2 1998 Q2 Adelaide South Australia Business  110.
##  3 1998 Q3 Adelaide South Australia Business  166.
##  4 1998 Q4 Adelaide South Australia Business  127.
##  5 1999 Q1 Adelaide South Australia Business  137.
##  6 1999 Q2 Adelaide South Australia Business  200.
##  7 1999 Q3 Adelaide South Australia Business  169.
##  8 1999 Q4 Adelaide South Australia Business  134.
##  9 2000 Q1 Adelaide South Australia Business  154.
## 10 2000 Q2 Adelaide South Australia Business  169.
```

Australian tourism

- Quarterly data on visitor night from 1998:Q1 – 2013:Q4
- From: *National Visitor Survey*, based on annual interviews of 120,000 Australians aged 15+, collected by Tourism Research Australia.
- Split by 7 states, 27 zones and 76 regions (a geographical hierarchy)
- Also split by purpose of travel
  - Holiday
  - Visiting friends and relatives (VFR)
  - Business
  - Other
- 304 bottom-level series

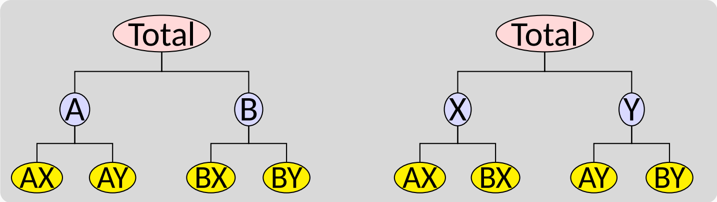
Grouped time series

A **grouped time series** is a collection of time series that can be grouped together in a number of non-hierarchical ways.



Grouped time series

A **grouped time series** is a collection of time series that can be grouped together in a number of non-hierarchical ways.



Examples

- Tourism by state and purpose of travel
- Retail sales by product groups/sub groups, and by countries/regions

Creating aggregates

```
PBS %>%
  aggregate_key(ATC1/ATC2, Scripts = sum(Scripts)) %>%
  filter(Month == yearmonth("1991 Jul")) %>% print(n=18)
```

```
## # A tsibble: 98 x 4 [?]
## # Key:       ATC1, ATC2 [98]
##    ATC1         ATC2            Month Scripts
##    <chr>        <chr>           <mth>   <dbl>
##  1 <aggregated> <aggregated> 1991 Jul 8090395
##  2 A            <aggregated> 1991 Jul  799025
##  3 B            <aggregated> 1991 Jul  109227
##  4 C            <aggregated> 1991 Jul 1794995
##  5 D            <aggregated> 1991 Jul  299779
##  6 G            <aggregated> 1991 Jul  300931
##  7 H            <aggregated> 1991 Jul  112114
##  8 J            <aggregated> 1991 Jul 1151681
##  9 L            <aggregated> 1991 Jul   24580
## 10 M            <aggregated> 1991 Jul  562956
## 11 N            <aggregated> 1991 Jul 1546023
## 12 P            <aggregated> 1991 Jul   47661
## 13 R            <aggregated> 1991 Jul  859273
## 14 S            <aggregated> 1991 Jul  391639
## 15 V            <aggregated> 1991 Jul   38705
## 16 Z            <aggregated> 1991 Jul   51806
## 17 A            A01          1991 Jul   22615
## 18 A            A02          1991 Jul  299251
## # ... with 80 more rows
```

Creating aggregates

```
tourism %>%
  aggregate_key(Purpose * (State / Region), Trips = sum(Trips)) %>%
  filter(Quarter == yearquarter("1998 Q1")) %>% print(n=15)
```

```
## # A tsibble: 425 x 5 [?]
## # Key:       Purpose, State, Region [425]
##    Purpose      State           Region        Quarter  Trips
##    <chr>        <chr>           <chr>           <qtr>  <dbl>
##  1 <aggregated> <aggregated>    <aggregated>  1998 Q1 23182.
##  2 Business     <aggregated>    <aggregated>  1998 Q1  3599.
##  3 Holiday      <aggregated>    <aggregated>  1998 Q1 11806.
##  4 Other        <aggregated>    <aggregated>  1998 Q1   680.
##  5 Visiting     <aggregated>    <aggregated>  1998 Q1  7098.
##  6 <aggregated> ACT           ~ <aggregated>  1998 Q1   551.
##  7 <aggregated> New South Wale~ <aggregated>  1998 Q1  8040.
##  8 <aggregated> Northern Terri~ <aggregated>  1998 Q1   181.
##  9 <aggregated> Queensland    ~ <aggregated>  1998 Q1  4041.
## 10 <aggregated> South Australi~ <aggregated>  1998 Q1  1735.
## 11 <aggregated> Tasmania      ~ <aggregated>  1998 Q1   982.
## 12 <aggregated> Victoria      ~ <aggregated>  1998 Q1  6010.
## 13 <aggregated> Western Austra~ <aggregated>  1998 Q1  1641.
## 14 <aggregated> ACT           ~ Canberra    ~ 1998 Q1   551.
## 15 <aggregated> New South Wale~ Blue Mounta~  1998 Q1   196.
## # ... with 410 more rows
```

Creating aggregates

- Similar to `summarise()` but using the key structure
- A grouped structure is specified using `grp1 * grp2`
- A nested structure is specified via `parent / child`.
- Groups and nesting can be mixed:

```
(country/region/city) * (brand/product)
```

- All possible aggregates are produced.
- These are useful when forecasting at different levels of aggregation.

Outline

The problem

1. How to forecast time series at all nodes such that the forecasts add up in the same way as the original data?
2. Can we exploit relationships between the series to improve the forecasts?

The problem

1. How to forecast time series at all nodes such that the forecasts add up in the same way as the original data?
2. Can we exploit relationships between the series to improve the forecasts?

The solution

1. Forecast all series at all levels of aggregation using an automatic forecasting algorithm.
(e.g., `ETS`, `ARIMA`, ...)
2. Reconcile the resulting forecasts so they add up correctly using least squares optimization (i.e., find closest reconciled forecasts to the original forecasts).
3. This is available using `reconcile()`.

Forecast reconciliation

```
tourism %>%
  aggregate_key(Purpose*(State/Region), Trips=sum(Trips)) %>%
  model(ets = ETS(Trips)) %>%
  reconcile(ets_adjusted = min_trace(ets)) %>%
  forecast(h = 2)
```

```
## # A fable: 1,700 x 7 [1Q]
## # Key:     Purpose, State, Region, .model [850]
##    Purpose    State      Region     .model     Quarter Trips
##    <chr>      <chr>      <chr>      <chr>        <qtr> <dbl>
##  1 Business   ACT       ~ Canberra ~ ets        2018 Q1 144.
##  2 Business   ACT       ~ Canberra ~ ets        2018 Q2 203.
##  3 Business   ACT       ~ <aggregat~ ets        2018 Q1 144.
##  4 Business   ACT       ~ <aggregat~ ets        2018 Q2 203.
##  5 Business   New South~ Blue Moun~ ets        2018 Q1  19.7
##  6 Business   New South~ Blue Moun~ ets        2018 Q2  19.7
##  7 Business   New South~ Capital C~ ets        2018 Q1  36.1
##  8 Business   New South~ Capital C~ ets        2018 Q2  36.1
```

Hierarchical and grouped time series

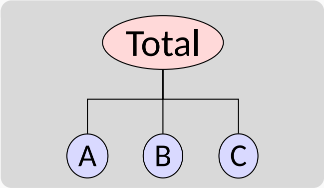
Every collection of time series with aggregation constraints can be written as

$$\boldsymbol{y}_t = \boldsymbol{S}\boldsymbol{b}_t$$

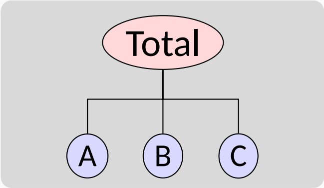
where

- $\boldsymbol{y}_t$ is a vector of all series at time t
- $\boldsymbol{b}_t$ is a vector of the most disaggregated series at time t
- $\boldsymbol{S}$ is a “summing matrix” containing the aggregation constraints.

Hierarchical time series



Hierarchical time series

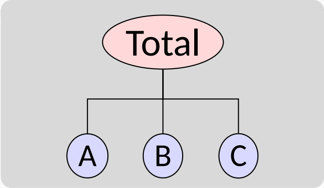


y_t : observed aggregate of all series at time t.

$y_{X,t}$: observation on series X at time t.

$\boldsymbol{b}_t$: vector of all series at bottom level in time t.

Hierarchical time series



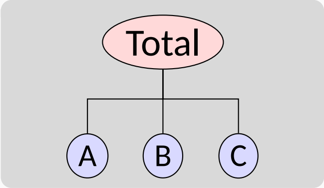
y_t : observed aggregate of all series at time t.

$y_{X,t}$: observation on series X at time t.

$\boldsymbol{b}_t$: vector of all series at bottom level in time t.

$$\boldsymbol{y}_t = \begin{pmatrix} y_t \\ y_{A,t} \\ y_{B,t} \\ y_{C,t} \end{pmatrix} = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} Y_{A,t} \\ y_{B,t} \\ y_{C,t} \end{pmatrix}$$

Hierarchical time series



y_t : observed aggregate of all series at time t.

$y_{X,t}$: observation on series X at time t.

$\boldsymbol{b}_t$: vector of all series at bottom level in time t.

$$\boldsymbol{y}_t = \begin{pmatrix} y_t \\ y_{A,t} \\ y_{B,t} \\ y_{C,t} \end{pmatrix} = \underbrace{\begin{pmatrix} 1 & 1 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}}_{\boldsymbol{S}} \underbrace{\begin{pmatrix} y_{A,t} \\ y_{B,t} \\ y_{C,t} \end{pmatrix}}_{\boldsymbol{b}_t}$$

$$\boldsymbol{y}_t = \boldsymbol{S}\boldsymbol{b}_t$$

Forecasting notation

Let $\hat{\boldsymbol{y}}_n(h)$ be vector of initial h-step forecasts, made at time n, stacked in same order as $\boldsymbol{y}_t$.

Forecasting notation

Let $\hat{\boldsymbol{y}}_n(h)$ be vector of initial h-step forecasts, made at time n, stacked in same order as $\boldsymbol{y}_t$.
(In general, they will not "add up".)

Forecasting notation

Let $\hat{\boldsymbol{y}}_n(h)$ be vector of initial h-step forecasts, made at time n, stacked in same order as $\boldsymbol{y}_t$.
(In general, they will not "add up".)

Reconciled forecasts must be of the form:

$$\tilde{\boldsymbol{y}}_n(h) = \boldsymbol{S}\boldsymbol{G}\hat{\boldsymbol{y}}_n(h)$$

for some matrix $\boldsymbol{G}$.

Forecasting notation

Let $\hat{\boldsymbol{y}}_n(h)$ be vector of initial h-step forecasts, made at time n, stacked in same order as $\boldsymbol{y}_t$.
(In general, they will not "add up".)

Reconciled forecasts must be of the form:

$$\tilde{\boldsymbol{y}}_n(h) = \boldsymbol{S}\boldsymbol{G}\hat{\boldsymbol{y}}_n(h)$$

for some matrix $\boldsymbol{G}$.

- $\boldsymbol{G}$ extracts and combines base forecasts $\hat{\boldsymbol{y}}_n(h)$ to get bottom-level forecasts.
- $\boldsymbol{S}$ adds them up

Optimal combination forecasts

Main result

The best (minimum sum of variances) unbiased forecasts are obtained when $\boldsymbol{G} = (\boldsymbol{S}'\Sigma_h^{-1}\boldsymbol{S})^{-1}\boldsymbol{S}'\Sigma_h^{-1}$, where Σ_h is the h-step base forecast error covariance matrix.

Optimal combination forecasts

Main result

The best (minimum sum of variances) unbiased forecasts are obtained when $\boldsymbol{G} = (\boldsymbol{S}'\Sigma_h^{-1}\boldsymbol{S})^{-1}\boldsymbol{S}'\Sigma_h^{-1}$, where Σ_h is the h-step base forecast error covariance matrix.

$$\tilde{\boldsymbol{y}}_n(h) = \boldsymbol{S}(\boldsymbol{S}'\Sigma_h^{-1}\boldsymbol{S})^{-1}\boldsymbol{S}'\Sigma_h^{-1}\hat{\boldsymbol{y}}_n(h)$$

Problem: Σ_h hard to estimate, especially for $h > 1$.

Solutions:

- Ignore Σ_h (OLS) [`min_trace(method='ols')`]
- Assume $\Sigma_h = k_h\Sigma_1$ is diagonal (WLS) [`min_trace(method='wls')`]
- Assume $\Sigma_h = k_h\Sigma_1$ and estimate it (GLS) [`min_trace(method='shrink')` (the default)]

Features

- Covariates can be included in initial forecasts.
- Adjustments can be made to initial forecasts at any level.
- Very simple and flexible method. Can work with *any* hierarchical or grouped time series.
- Conceptually easy to implement: regression of base forecasts on structure matrix.

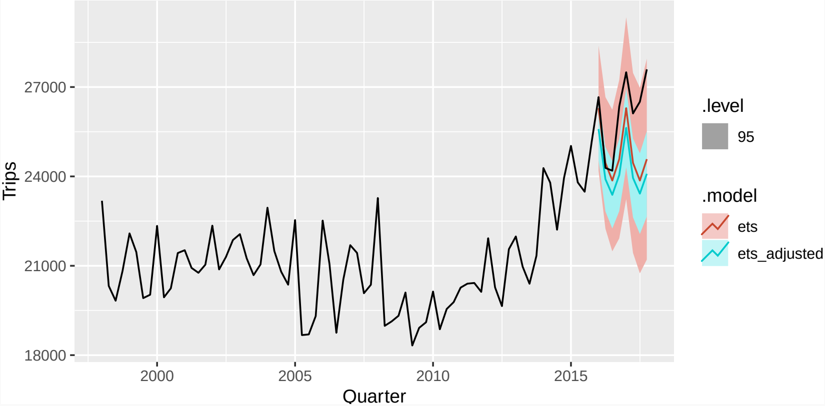
Outline

Example: Australian tourism

```r
tourism_agg <- tourism %>%
  aggregate_key(Purpose * (State / Region),
                Trips = sum(Trips))
fc <- tourism_agg %>%
  filter_index(. ~ yearquarter("2015 Q4")) %>%
  model(ets = ETS(Trips)) %>%
  reconcile(ets_adjusted = min_trace(ets)) %>%
  forecast(h = "2 years")
```

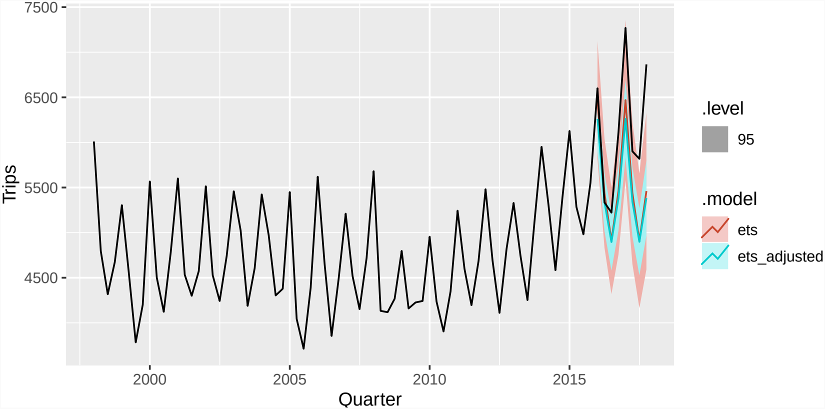
Example: Australian tourism

```
fc %>%
  filter(is_aggregated(Purpose) & is_aggregated(State)) %>%
  autoplot(tourism_agg, level=95)
```



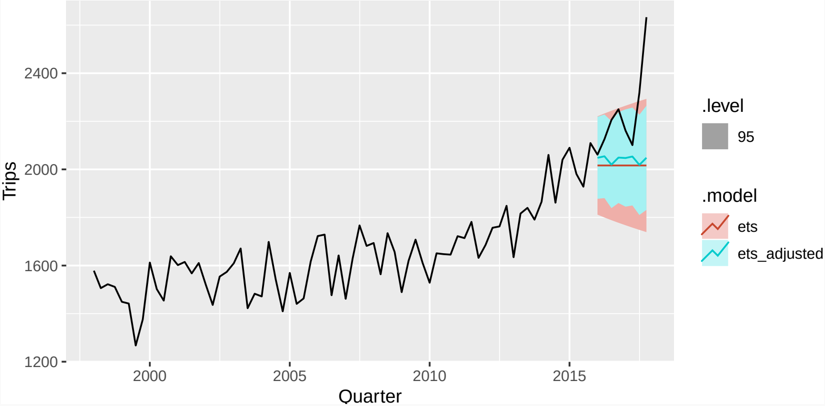
Example: Australian tourism

```r
fc %>%
  filter(is_aggregated(Purpose) & State=="Victoria" &
         is_aggregated(Region)) %>%
  autoplot(tourism_agg, level=95)
```



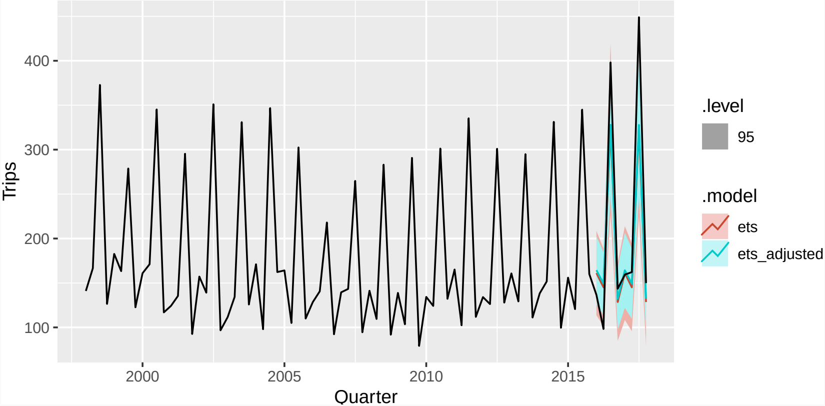
Example: Australian tourism

```
fc %>%
  filter(is_aggregated(Purpose) & Region=="Melbourne") %>%
  autoplot(tourism_agg, level=95)
```



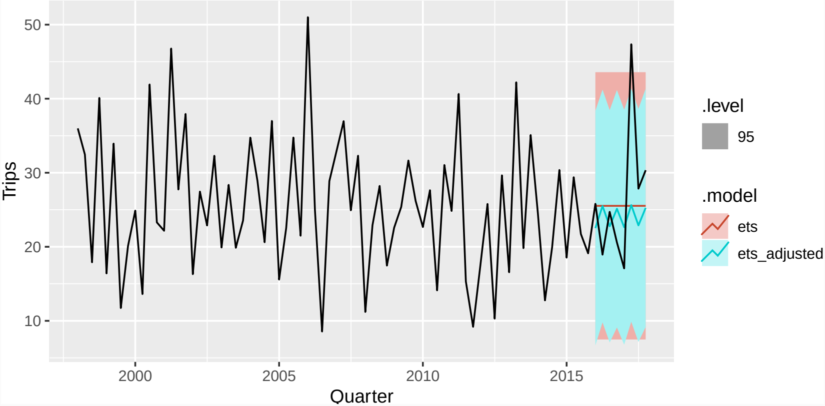
Example: Australian tourism

```
fc %>%
  filter(is_aggregated(Purpose) & Region=="Snowy Mountains") %>%
  autoplot(tourism_agg, level=95)
```



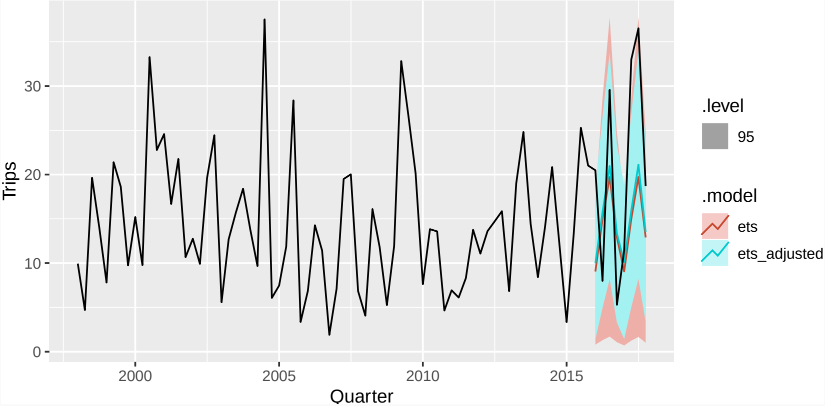
Example: Australian tourism

```
fc %>%
  filter(Purpose=="Holiday" & Region=="Barossa") %>%
  autoplot(tourism_agg, level=95)
```



Example: Australian tourism

```
fc %>%
  filter(is_aggregated(Purpose) & Region=="MacDonnell") %>%
  autoplot(tourism_agg, level=95)
```



Forecast evaluation

```
fc %>% accuracy(tourism_agg)
```

```
## # A tibble: 850 x 12
##    .model Purpose  State      Region     .type    ME  RMSE
##    <chr>  <chr>    <chr>      <chr>      <chr> <dbl> <dbl>
##  1 ets    Business ACT      ~ Canberra ~ Test  51.7   56.4
##  2 ets    Business ACT      ~ <aggregat~ Test  51.7   56.4
##  3 ets    Business New South~ Blue Moun~ Test   1.93  10.6
##  4 ets    Business New South~ Capital C~ Test   8.83  16.1
##  5 ets    Business New South~ Central C~ Test   9.62  14.2
##  6 ets    Business New South~ Central N~ Test  19.3   31.9
##  7 ets    Business New South~ Hunter   ~ Test  21.4   33.8
##  8 ets    Business New South~ New Engla~ Test  22.7   33.4
##  9 ets    Business New South~ North Coa~ Test  11.8   27.9
## 10 ets    Business New South~ Outback N~ Test  17.3   19.2
## # ... with 840 more rows, and 5 more variables: MAE <dbl>,
## #   MPE <dbl>, MAPE <dbl>, MASE <dbl>, ACF1 <dbl>
```

Forecast evaluation

```
fc %>% accuracy(tourism_agg) %>%
  group_by(.model) %>%
  summarise(MASE = mean(MASE))
```

```
## # A tibble: 2 x 2
##   .model         MASE
##   <chr>         <dbl>
## 1 ets           1.04
## 2 ets_adjusted 0.984
```

Forecast evaluation

```
tourism_state <- tourism %>%
  group_by(State, Purpose) %>%
  summarise(Trips = sum(Trips)) %>%
  aggregate_key(State * Purpose,
                Trips = sum(Trips))
tourism_state
```

```
## # A tsibble: 3,600 x 4 [1Q]
## # Key:       State, Purpose [45]
##    State        Purpose      Quarter  Trips
##    <chr>        <chr>          <qtr>  <dbl>
##  1 <aggregated> <aggregated> 1998 Q1 23182.
##  2 <aggregated> <aggregated> 1998 Q2 20323.
##  3 <aggregated> <aggregated> 1998 Q3 19827.
##  4 <aggregated> <aggregated> 1998 Q4 20830.
##  5 <aggregated> <aggregated> 1999 Q1 22087.
##  6 <aggregated> <aggregated> 1999 Q2 21458.
##  7 <aggregated> <aggregated> 1999 Q3 19914.
```

Forecast evaluation

```r
tourism_stretch <- tourism_state %>%
  stretch_tsibble(.init = 16)
tourism_stretch
```

```
## # A tsibble: 140,400 x 5 [1Q]
## # Key:       .id, State, Purpose [2,925]
##    State Purpose  Quarter Trips   .id
##    <chr> <chr>      <qtr> <dbl> <int>
##  1 ACT   Business 1998 Q1 150.      1
##  2 ACT   Business 1998 Q2  99.9     1
##  3 ACT   Business 1998 Q3 130.      1
##  4 ACT   Business 1998 Q4 102.      1
##  5 ACT   Business 1999 Q1  95.5     1
##  6 ACT   Business 1999 Q2 229.      1
##  7 ACT   Business 1999 Q3 109.      1
##  8 ACT   Business 1999 Q4 159.      1
##  9 ACT   Business 2000 Q1 105.      1
## 10 ACT   Business 2000 Q2 202.      1
```

Forecast evaluation

```r
fits <- tourism_stretch %>%
  model(ets = ETS(Trips)) %>%
  reconcile(ets_adjusted = min_trace(ets, method='wls'))
fits
```

```
## # A mable: 2,925 x 5
## # Key:     .id, State, Purpose [2,925]
##      .id State           Purpose      ets          ets_adjusted
##    <int> <chr>           <chr>        <model>      <model>
##  1     1 ACT             Business     <ETS(A,N,N)> <ETS(A,N,N)>
##  2     1 ACT             Holiday      <ETS(A,N,N)> <ETS(A,N,N)>
##  3     1 ACT             Other        <ETS(A,N,N)> <ETS(A,N,N)>
##  4     1 ACT             Visiting     <ETS(A,N,N)> <ETS(A,N,N)>
##  5     1 ACT             <aggregated> <ETS(A,N,N)> <ETS(A,N,N)>
##  6     1 New South Wales Business     <ETS(M,N,N)> <ETS(M,N,N)>
##  7     1 New South Wales Holiday      <ETS(M,N,M)> <ETS(M,N,M)>
##  8     1 New South Wales Other        <ETS(A,N,N)> <ETS(A,N,N)>
##  9     1 New South Wales Visiting     <ETS(M,N,N)> <ETS(M,N,N)>
## 10     1 New South Wales <aggregated> <ETS(A,N,N)> <ETS(A,N,N)>
```

Forecast evaluation

```
fc <- fits %>% forecast(h = 1)
fc
```

```
## # A fable: 5,850 x 7 [1Q]
## # Key:     .id, State, Purpose, .model [5,850]
##        .id State       Purpose     .model     Quarter  Trips
##      <int> <chr>       <chr>       <chr>        <qtr>  <dbl>
##  1      1 ACT        ~ Business    ets        2002 Q1  138.
##  2      1 ACT        ~ Holiday     ets        2002 Q1  158.
##  3      1 ACT        ~ Other       ets        2002 Q1   25.3
##  4      1 ACT        ~ Visiting    ets        2002 Q1  180.
##  5      1 ACT        ~ <aggregat~ ets        2002 Q1  501.
##  6      1 New South~ Business    ets        2002 Q1 1350.
##  7      1 New South~ Holiday     ets        2002 Q1 3699.
##  8      1 New South~ Other       ets        2002 Q1  275.
##  9      1 New South~ Visiting    ets        2002 Q1 2356.
## 10      1 New South~ <aggregat~ ets        2002 Q1 7114.
## # ... with 5,840 more rows, and 1 more variable:
```

Forecast evaluation

```r
fc %>% accuracy(tourism_state)
```

```
## # A tibble: 90 x 11
##    .model State        Purpose     .type    ME  RMSE    MAE
##    <chr>  <chr>        <chr>       <chr> <dbl> <dbl>  <dbl>
##  1 ets    ACT        ~ Business    Test   8.89  34.1  27.9
##  2 ets    ACT        ~ Holiday     Test   3.06  44.7  32.1
##  3 ets    ACT        ~ Other       Test   2.51  12.4   9.96
##  4 ets    ACT        ~ Visiting    Test   4.89  31.1  24.4
##  5 ets    ACT        ~ <aggregat~ Test   8.54  61.2  48.1
##  6 ets    New South~ Business    Test   4.50 141.  111.
##  7 ets    New South~ Holiday     Test   7.32 176.  144.
##  8 ets    New South~ Other       Test   6.23  44.9  38.7
##  9 ets    New South~ Visiting    Test  32.1  165.  134.
## 10 ets    New South~ <aggregat~ Test  50.1  345.  282.
## # ... with 80 more rows, and 4 more variables: MPE <dbl>,
## #   MAPE <dbl>, MASE <dbl>, ACF1 <dbl>
```

Forecast evaluation

```
fc %>% accuracy(tourism_state) %>%
  group_by(.model) %>%
  summarise(MASE = mean(MASE))
```

```
## # A tibble: 2 x 2
##   .model        MASE
##   <chr>        <dbl>
## 1 ets          0.833
## 2 ets_adjusted 0.952
```

Outline

Lab Session 10

- Create a new tsibble for PBS data using only ATC1 groups (combining over ATC2 groups).
- Use forecast reconciliation with this data, using both ETS and SNAIVE models.
- Which type of model works best?
- Does the reconciliation improve the forecast accuracy?
- Why doesn't reconciliation make any difference to SNAIVE forecasts?