

Tidy Forecasting in R

Rob J Hyndman

ISF 2018

forecast package

Pre 2003	Private functions used for consulting projects
July/August 2003	ets and thetaf added
August 2006	v1.0 available on CRAN
May 2007	auto.arima added
May 2010	arfima added
Feb/March 2011	tslm , stlf , naive , snaive added
August 2011	v3.0 . Box Cox transformations added
December 2011	tbats added
April 2012	Package moved to github
November 2012	v4.0 . nnetar added
June 2013	Major speed-up of ets
February 2016	v7.0 . Added ggplot2 graphics
February 2017	v8.0 . Added checkresiduals , tsCV and %>%
April 2018	v8.3 . Added mstl
June 2018	≈ 100,000 package downloads per month

fable package

A replacement for the forecast package.

Why change?

- Interacting with tidyverse packages
- Sub-daily data and multiple seasonal data handled more easily
- Consistency of interface
- Distribution forecasting rather than point+interval
- Simpler interface for hierarchical and grouped forecast reconciliation
- Designed for forecasting many related time series
- Changes will break too much existing code
- Opportunity to re-think forecasting practice

Example: Australian eating-out expenditure

```
library(fpp2)
auscafe
```

##	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug
## 1982				0.342	0.342	0.329	0.339	0.332
## 1983	0.369	0.348	0.366	0.351	0.360	0.347	0.364	0.376
## 1984	0.389	0.377	0.398	0.383	0.414	0.382	0.393	0.409
## 1985	0.426	0.392	0.416	0.420	0.446	0.407	0.449	0.466
## 1986	0.504	0.453	0.480	0.497	0.531	0.485	0.526	0.538
## 1987	0.572	0.525	0.544	0.558	0.565	0.542	0.599	0.584
## 1988	0.605	0.586	0.625	0.612	0.630	0.635	0.659	0.656
## 1989	0.733	0.661	0.713	0.694	0.710	0.722	0.741	0.746
## 1990	0.858	0.764	0.840	0.805	0.809	0.799	0.815	0.828
## 1991	0.862	0.771	0.813	0.797	0.821	0.801	0.829	0.854
## 1992	0.938	0.862	0.936	0.932	0.929	0.869	0.891	0.875
## 1993	0.918	0.838	0.870	0.862	0.852	0.828	0.882	0.867
## 1994	0.985	0.902	1.015	0.939	0.941	0.935	1.013	1.018
## 1995	1.076	0.982	1.099	1.068	1.083	1.045	1.094	1.110
## 1996	1.213	1.128	1.180	1.169	1.146	1.109	1.138	1.146
## 1997	1.180	1.060	1.148	1.141	1.170	1.113	1.165	1.173

Example: Australian eating-out expenditure

```
library(tsibble)
cafe <- as_tsibble(auscafe)
cafe
```

```
## # A tsibble: 426 x 2 [1MONTH]
##       index value
##       <mth> <dbl>
## 1 1982 Apr 0.342
## 2 1982 May 0.342
## 3 1982 Jun 0.329
## 4 1982 Jul 0.338
## 5 1982 Aug 0.332
## 6 1982 Sep 0.342
## 7 1982 Oct 0.358
## 8 1982 Nov 0.375
## 9 1982 Dec 0.433
## 10 1983 Jan 0.369
## # ... with 416 more rows
```

Example: Australian eating-out expenditure

```
library(fable)
cafe %>% ETS(value)
```

```
## # A tibble: 1 x 2
##   data                model
##   <list>              <list>
## 1 <tsibble [426 x 2]> <ETS(M,A,M)>
```

Example: Australian eating-out expenditure

```
cafe %>% ETS(value) %>% summary()
```

```
## data.Length data.Class data.Mode model.Length mode  
## 2          tbl_ts  list          19          ts_model
```

Example: Australian eating-out expenditure

```
cafe %>% ETS(value) %>% forecast()
```

```
## # A tibble: 1 x 3
```

```
##   data                                model          forecast
```

```
##   <list>                             <list>         <list>
```

```
## 1 <tsibble [426 x 2]> <ETS(M,A,M)> <tsibble [24 x 3]>
```


Example: Australian eating-out expenditure

```
cafe %>% ETS(value) %>% forecast() %>% summary()
```

```
## data.Length data.Class data.Mode model.Length model  
## 2          tbl_ts  list              19          ts_model  
## forecast.Length forecast.Class forecast.Mode  
## 3          tbl_ts  list
```

Example: Australian eating-out expenditure

```
#cafe %>% ETS(value) %>% forecast() %>% autoplot()  
## Currently not working
```

Example: Australian eating-out expenditure

```
cafe %>% ARIMA(log(value)) %>%  
  forecast() %>% summary()
```

```
## data.Length data.Class data.Mode model.Length model  
## 2      tbl_ts list              18      ts_model  
## forecast.Length forecast.Class forecast.Mode  
## 3      tbl_ts list
```

Example: prison data

```
prisonLF
```

```
## # A tibble: 1,536 x 5
```

```
##   state gender legal      t      count
```

```
##   <fct> <fct>  <fct>   <date>   <dbl>
```

```
## 1 ACT    Female Remanded 2005-03-01      2
```

```
## 2 ACT    Female Remanded 2005-06-01      4
```

```
## 3 ACT    Female Remanded 2005-09-01      1
```

```
## 4 ACT    Female Remanded 2005-12-01      4
```

```
## 5 ACT    Female Remanded 2006-03-01      4
```

```
## 6 ACT    Female Remanded 2006-06-01      6
```

```
## 7 ACT    Female Remanded 2006-09-01      9
```

```
## 8 ACT    Female Remanded 2006-12-01      6
```

```
## 9 ACT    Female Remanded 2007-03-01      4
```

```
## 10 ACT   Female Remanded 2007-06-01      4
```

```
## # ... with 1,526 more rows
```

Example: prison data

```
prison <- prisonLF %>%  
  mutate(qtr=yearquarter(t)) %>%  
  select(-t) %>%  
  as_tsibble(index=qtr, key=id(state,gender,legal))  
prison
```

```
## # A tsibble: 1,536 x 5 [1QUARTER]  
## # Keys:      state, gender, legal [32]  
##   state gender legal    count    qtr  
##   <fct> <fct>  <fct>    <dbl>  <qtr>  
## 1 ACT   Female Remanded      2 2005 Q1  
## 2 ACT   Female Remanded      4 2005 Q2  
## 3 ACT   Female Remanded      1 2005 Q3  
## 4 ACT   Female Remanded      4 2005 Q4  
## 5 ACT   Female Remanded      4 2006 Q1  
## 6 ACT   Female Remanded      6 2006 Q2  
## 7 ACT   Female Remanded      9 2006 Q3
```

Example: prison data

```
prison %>% ETS(count)
```

```
## # A tibble: 32 x 5
```

##	state	gender	legal	data	model
##	<fct>	<fct>	<fct>	<list>	<list>
## 1	ACT	Female	Remanded	<tsibble [48 x 2]>	<ETS(M,A,N)>
## 2	ACT	Female	Sentenced	<tsibble [48 x 2]>	<ETS(A,A,N)>
## 3	ACT	Male	Remanded	<tsibble [48 x 2]>	<ETS(M,N,N)>
## 4	ACT	Male	Sentenced	<tsibble [48 x 2]>	<ETS(A,N,N)>
## 5	NSW	Female	Remanded	<tsibble [48 x 2]>	<ETS(M,N,M)>
## 6	NSW	Female	Sentenced	<tsibble [48 x 2]>	<ETS(M,N,M)>
## 7	NSW	Male	Remanded	<tsibble [48 x 2]>	<ETS(M,A,A)>
## 8	NSW	Male	Sentenced	<tsibble [48 x 2]>	<ETS(M,A,A)>
## 9	NT	Female	Remanded	<tsibble [48 x 2]>	<ETS(M,N,N)>
## 10	NT	Female	Sentenced	<tsibble [48 x 2]>	<ETS(M,A,A)>
## #	... with 22 more rows				

Example: prison data

```
prison %>% ETS(count) %>% forecast()
```

```
## # A tibble: 32 x 6
##   state gender legal      data      model      forecast
##   <fct> <fct>  <fct>    <list>    <list>    <list>
## 1 ACT    Female Remanded <tsibble [48 x 2]> <ETS(M,A,N)> <tsibble [8 x 3~
## 2 ACT    Female Sentenced <tsibble [48 x 2]> <ETS(A,A,N)> <tsibble [8 x 3~
## 3 ACT    Male    Remanded <tsibble [48 x 2]> <ETS(M,N,N)> <tsibble [8 x 3~
## 4 ACT    Male    Sentenced <tsibble [48 x 2]> <ETS(A,N,N)> <tsibble [8 x 3~
## 5 NSW    Female Remanded <tsibble [48 x 2]> <ETS(M,N,M)> <tsibble [8 x 3~
## 6 NSW    Female Sentenced <tsibble [48 x 2]> <ETS(M,N,M)> <tsibble [8 x 3~
## 7 NSW    Male    Remanded <tsibble [48 x 2]> <ETS(M,A,A)> <tsibble [8 x 3~
## 8 NSW    Male    Sentenced <tsibble [48 x 2]> <ETS(M,A,A)> <tsibble [8 x 3~
## 9 NT     Female Remanded <tsibble [48 x 2]> <ETS(M,N,N)> <tsibble [8 x 3~
## 10 NT    Female Sentenced <tsibble [48 x 2]> <ETS(M,A,A)> <tsibble [8 x 3~
## # ... with 22 more rows
```

Aggregation and reconciliation not yet implemented.

Equivalent methods

- `auto.arima` → ARIMA
- `ets` → ETS
- `tbats` → TBATS
- `stlm` → STL???

All modelling functions produce mable class objects.

Equivalent methods

- `stlf` → STL %>% forecast
- `thetaf` → ??
- `dshw`, `hw`, `holt`, `ses` ??
- `splinef` → ??
- `rwf`, `naive` → RW %>% forecast
- `croston` →

forecast produces fable class objects.

Download

```
devtools::install_github("tidyverts/tsibble")  
devtools::install_github("tidyverts/fable")
```

NUMBATS



Di Cook



Earo Wang



Mitchell O'Hara-Wild

More information

robjhyndman.com

OTexts.org/fpp2

tidyverts.org