

When **time series** meets **tibble**, it's **tsibble**!

binding univariate, multivariate, hierarchical & grouped time series into one

Earo Wang @ Monash University

@earowang

🕒 INDEX

Time index can be:

- POSIXct
- Date
- difftime/hms
- yearmonth/zoo::yearmth
- yearquarter/zoo::yearqtr
- numeric

Time interval of data automatically computed from index.

Quarter	Region	State	Purpose	Trips ('000)	...
1998 Q1	Melbourne	Victoria	Holiday	428	
1999 Q1	Geelong	Victoria	Business	15.9	
...	
2016 Q4	Sydney	New South Wales	Visiting	921	
2016 Q4	Blue Mountains	New South Wales	Other	4.73	

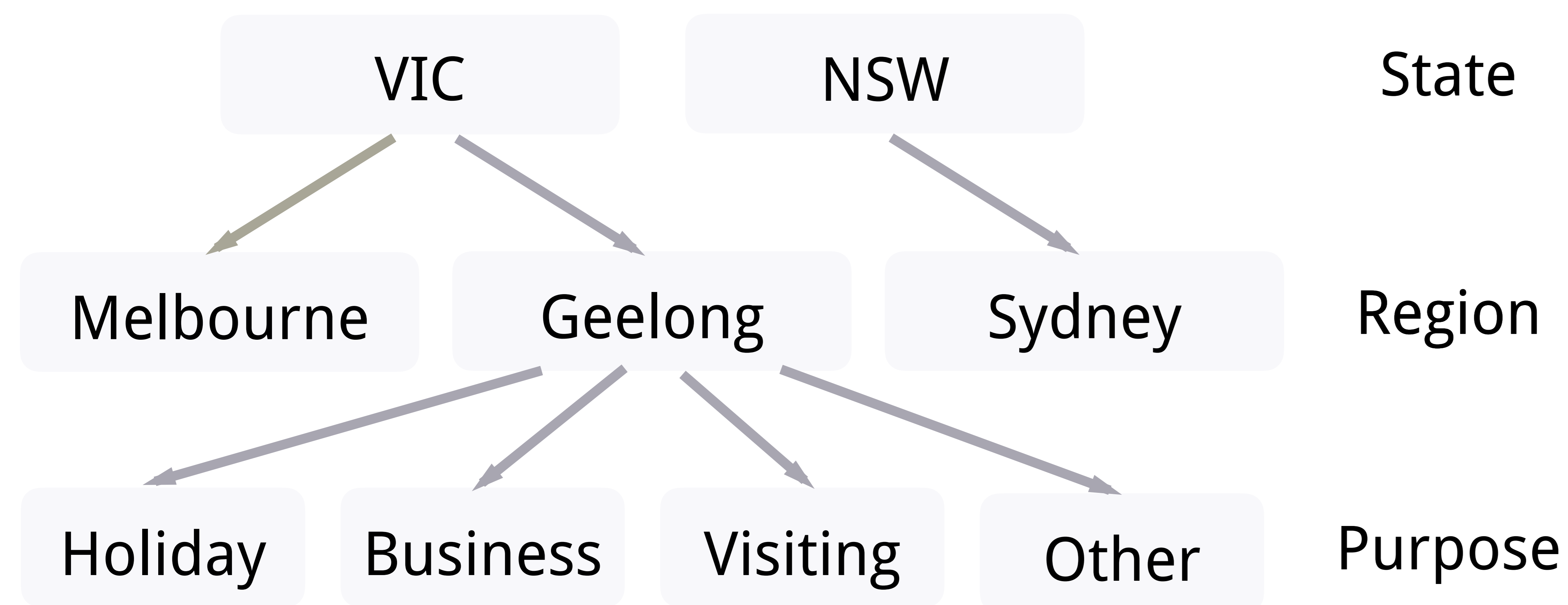
Australian domestic overnight trips*

```
as_tsibble(tourism, key = id(Region | State, Purpose),
           index = Quarter)
#> # A tsibble: 23,408 x 5 [1QUARTER]
#> # Keys:      Region | State, Purpose [308]
#>   Quarter Region   State      Purpose  Trips
#>   <qtr> <chr>    <chr>    <chr>    <dbl>
#> 1 1998 Q1 Melbourne Victoria  Holiday  428
#> 2 1998 Q1 Melbourne Victoria  Business 405
#> 3 1998 Q1 Melbourne Victoria  Visiting 666
#> 4 1998 Q1 Melbourne Victoria  Other    79.9
#> 5 1998 Q1 Sydney    New South Wales  Holiday  828
#> # ... with 2.34e+04 more rows
```

* Data source: Tourism Research Australia

Each observation is uniquely identified by INDEX & KEY in a tsibble.

key = id(Region | State, Purpose)



↑ Nested and crossed data structure



KEY

Key created via id() --- identifying variable:

- None: an implicit variable id()
- Single: an explicit variable id(Purpose)
- Nested: a nesting variable under another id(Region | State)
- Crossed: a crossing variable with another id(Region, Purpose)



MEASURES

🔧 DATA WRANGLING

- ☑ fill_na() turns implicit missing values into explicit missing values.
- ☑ tsummarise() aggregates over calendar periods.
- ☑ slide(), tile(), stretch() perform window calculations.
- ☑ support dplyr common verbs.