

tidyr

**Problem:** Tables are usually not formatted for convenient analysis in R

**Solution:** “Tidy” up the table

**“Wide”**

```
table1
#> # A tibble: 6 x 4
#>   country    year  cases population
#>   <chr>    <int> <int>      <int>
#> 1 Afghanistan 1999     745  19987071
#> 2 Afghanistan 2000    2666  20595360
#> 3 Brazil      1999   37737  172006362
#> 4 Brazil      2000   80488  174504898
#> 5 China       1999  212258 1272915272
#> 6 China       2000  213766 1280428583
```

**“Gross”**

```
table3
#> # A tibble: 6 x 3
#>   country    year rate
#> * <chr>    <int> <chr>
#> 1 Afghanistan 1999 745/19987071
#> 2 Afghanistan 2000 2666/20595360
#> 3 Brazil      1999 37737/172006362
#> 4 Brazil      2000 80488/174504898
#> 5 China       1999 212258/1272915272
#> 6 China       2000 213766/1280428583
```

**“Long”**

```
table2
#> # A tibble: 12 x 4
#>   country    year type      count
#>   <chr>    <int> <chr>    <int>
#> 1 Afghanistan 1999 cases      745
#> 2 Afghanistan 1999 population 19987071
#> 3 Afghanistan 2000 cases      2666
#> 4 Afghanistan 2000 population 20595360
#> 5 Brazil      1999 cases      37737
#> 6 Brazil      1999 population 172006362
#> # ... with 6 more rows
```

**“Split”**

```
table4a # cases
#> # A tibble: 3 x 3
#>   country `1999` `2000`
#> * <chr>    <int> <int>
#> 1 Afghanistan     745     2666
#> 2 Brazil          37737    80488
#> 3 China           212258   213766
table4b # population
#> # A tibble: 3 x 3
#>   country `1999` `2000`
#> * <chr>    <int> <int>
#> 1 Afghanistan 19987071 20595360
#> 2 Brazil      172006362 174504898
#> 3 China       1272915272 1280428583
```