

Automatic Conversion of Tables to LongForm Dataframes

Jimmy Oh

Motivation



Source: http://en.wikipedia.org/wiki/File:Kitchen_table.jpg

Hierarchical Tables

Group: Household Labour Force Survey - HLF

Table: Labour Force Status by Sex by Sing/Comb Ethnic Group (Qrtly-Mar/Jun/Sep/Dec)

Edit table

Save table

--Please select--

Table on screen

[Footnotes](#) | [Information](#) | [Print table](#) | [Help](#) | [Status flags](#)

	European Only								Maori Only					
	Persons Employed in Labour Force	Persons Unemployed in Labour Force	Not in Labour Force	Working Age Population	Labour Force Participation Rate	Unemployment Rate	Employment Rate	Total Labour Force	Persons Employed in Labour Force	Persons Unemployed in Labour Force	Not in Labour Force	Working Age Population	Labour Force Participation Rate	Unemp
2007Q4	855.8	20.0	280.0	1,155.8	75.8	2.3	74.0	875.8	71.1	6.1	28.1	105.3	73.4	
2008Q1	863.0	25.4	283.5	1,171.9	75.8	2.9	73.6	888.5	69.1	7.5	31.4	107.9	71.0	
2008Q2	850.1	26.0	280.7	1,156.8	75.7	3.0	73.5	876.1	67.2	5.7	27.4	100.2	72.7	
2008Q3	839.6	29.8	285.9	1,155.3	75.2	3.4	72.7	869.4	71.7	8.7	30.7	111.1	72.3	
2008Q4	854.8	29.5	274.7	1,158.9	76.3	3.3	73.8	884.2	76.1	8.5	28.5	113.1	74.8	
2009Q1	845.0	35.4	279.4	1,159.8	75.9	4.0	72.9	880.4	75.4	8.4	35.7	119.5	70.1	
2009Q2	831.6	34.9	279.7	1,146.2	75.6	4.0	72.6	866.5	74.2	9.9	33.1	117.3	71.8	
2009Q3	813.3	42.5	290.4	1,146.2	74.7	5.0	71.0	855.8	70.9	10.9	36.0	117.8	69.5	
2009Q4	831.1	40.1	277.0	1,148.2	75.9	4.6	72.4	871.2	71.7	13.6	33.2	118.5	71.9	
2010Q1	822.5	36.4	283.2	1,142.1	75.2	4.2	72.0	858.9	71.8	11.3	35.3	118.4	70.2	
2010Q2	825.3	39.9	290.1	1,155.3	74.9	4.6	71.4	865.2	71.9	13.7	33.7	119.2	71.8	
2010Q3	836.9	31.0	287.1	1,155.1	75.1	3.6	72.5	867.9	69.8	13.5	34.1	117.3	71.0	
2010Q4	838.1	39.6	277.1	1,154.8	76.0	4.5	72.6	877.7	70.7	14.4	36.4	121.5	70.0	
2011Q1	829.6	36.8	281.2	1,147.6	75.5	4.2	72.3	866.4	70.7	13.9	35.3	119.8	70.6	
2011Q2	838.7	41.0	279.0	1,158.7	75.9	4.7	72.4	879.6	67.1	10.5	37.1	114.7	67.6	
2011Q3	830.5	34.6	280.2	1,145.3	75.5	4.0	72.5	865.1	69.5	13.4	34.9	117.8	70.4	
2011Q4	841.8	34.8	277.5	1,154.1	76.0	4.0	72.9	876.6	69.2	12.7	36.1	118.0	69.4	
2012Q1	843.1	43.3	282.7	1,169.1	75.8	4.9	72.1	886.3	71.5	11.2	34.6	117.3	70.5	
2012Q2	837.1	38.2	296.5	1,171.8	74.7	4.4	71.4	875.3	66.2	10.6	33.0	109.7	69.9	
2012Q3	833.0	38.0	298.4	1,169.3	74.5	4.4	71.2	871.0	67.1	13.1	33.0	113.2	70.8	
2012Q4	833.4	41.0	298.2	1,172.6	74.6	4.7	71.1	874.4	63.0	12.3	35.4	110.7	68.0	
2013Q1	832.0	35.8	294.7	1,162.5	74.6	4.1	71.6	867.8	69.9	12.2	38.2	120.3	68.3	

Hierarchical Tables

	1	2	3	4	5	6	7	8	9	10	11	
1	Labour Force Status by Sex by Sing/Comb Ethnic Group (Qrtly-Mar/Jun/Sep/Dec)											
2		Male										
3		European Only									Maori Only	
4		Persons Err	Persons Un	Not in Labo	Working Ag	Labour Forc	Unemploym	Employmer	Total Labou	Persons Err	Persons Un	
5	2007Q4	856	20	280	1,156	76	2	74	876	71	6	
6	2008Q1	863	25	284	1,172	76	3	74	888	69	8	
7	2008Q2	850	26	281	1,157	76	3	74	876	67	6	
8	2008Q3	840	30	286	1,155	75	3	73	869	72	9	
9	2008Q4	855	30	275	1,159	76	3	74	884	76	8	
10	2009Q1	845	35	279	1,160	76	4	73	880	75	8	
11	2009Q2	832	35	280	1,146	76	4	73	866	74	10	
12	2009Q3	813	42	290	1,146	75	5	71	856	71	11	
13	2009Q4	831	40	277	1,148	76	5	72	871	72	14	
14	2010Q1	822	36	283	1,142	75	4	72	859	72	11	
15	2010Q2	825	40	290	1,155	75	5	71	865	72	14	
16	2010Q3	837	31	287	1,155	75	4	72	868	70	14	
17	2010Q4	838	40	277	1,155	76	4	73	878	71	14	
18	2011Q1	830	37	281	1,148	76	4	72	866	71	14	
19	2011Q2	839	41	279	1,159	76	5	72	880	67	10	
20	2011Q3	830	35	280	1,145	76	4	72	865	70	13	
21	2011Q4	842	35	278	1,154	76	4	73	877	69	13	
22	2012Q1	843	43	283	1,169	76	5	72	886	72	11	
23	2012Q2	837	38	296	1,172	75	4	71	875	66	11	
24	2012Q3	833	38	298	1,169	74	4	71	871	67	13	
25	2012Q4	833	41	298	1,173	75	5	71	874	63	12	
26	2013Q1	832	36	295	1,162	75	4	72	868	70	12	

Why Tables are Bad

	1	2	3	4	5	6	7	8	9	10	11
1	Labour Force Status by Sex by Sing/Comb Ethnic Group (Qrtly-Mar/Jun/Sep/Dec)										
2		Male									
3		European Only								Maori Only	
4		Persons Err	Persons Un	Not in Labo	Working Ag	Labour Forc	Unemploym	Employer	Total Labou	Persons Err	Persons Un
5	2007Q4	856	20	280	1,156	76	2	74	876	71	6
6	2008Q1	863	25	284	1,172	76	3	74	888	69	8
7	2008Q2	850	26	281	1,157	76	3	74	876	67	6
8	2008Q3	840	30	286	1,155	75	3	73	869	72	9
9	2008Q4	855	30	275	1,159	76	3	74	884	76	8
10	2009Q1	845	35	279	1,160	76	4	73	880	75	8
11	2009Q2	832	35	280	1,146	76	4	73	866	74	10
12	2009Q3	813	42	290	1,146	75	5	71	856	71	11
13	2009Q4	831	40	277	1,148	76	5	72	871	72	14
14	2010Q1	822	36	283	1,142	75	4	72	859	72	11
15	2010Q2	825	40	290	1,155	75	5	71	865	72	14
16	2010Q3	837	31	287	1,155	75	4	72	868	70	14
17	2010Q4	838	40	277	1,155	76	4	73	878	71	14
18	2011Q1	830	37	281	1,148	76	4	72	866	71	14
19	2011Q2	839	41	279	1,159	76	5	72	880	67	10
20	2011Q3	830	35	280	1,145	76	4	72	865	70	13
21	2011Q4	842	35	278	1,154	76	4	73	877	69	13
22	2012Q1	843	43	283	1,169	76	5	72	886	72	11
23	2012Q2	837	38	296	1,172	75	4	71	875	66	11
24	2012Q3	833	38	298	1,169	74	4	71	871	67	13
25	2012Q4	833	41	298	1,173	75	5	71	874	63	12
26	2013Q1	832	36	295	1,162	75	4	72	868	70	12

Why Tables are Bad

	1	2	3	4	5	6	7	8	9	10	11
1	Labour Force Status by Sex by Sing/Comb Ethnic Group (Qrtly-Mar/Jun/Sep/Dec)										
2		Male									
3		European Only									Maori Only
4		Persons Em	Persons Un	Not in Labo	Working Ag	Labour Forc	Unemploym	Employer	Total Labou	Persons Em	Persons Un
5	2007Q4	856	20	280	1,156	76	2	74	876	71	6
6	2008Q1	863	25	284	1,172	76	3	74	888	69	8
7	2008Q2	850	26	281	1,157	76	3	74	876	67	6
8	2008Q3	840	30	286	1,155	75	3	73	869	72	9
9	2008Q4	855	30	275	1,159	76	3	74	884	76	8
10	2009Q1	845	35	279	1,160	76	4	73	880	75	8
11	2009Q2	832	35	280	1,146	76	4	73	866	74	10
12	2009Q3	813	42	290	1,146	75	5	71	856	71	11
13	2009Q4	831	40	277	1,148	76	5	72	871	72	14
14	2010Q1	822	36	283	1,142	75	4	72	859	72	11
15	2010Q2	825	40	290	1,155	75	5	71	865	72	14
16	2010Q3	837	31	287	1,155	75	4	72	868	70	14
17	2010Q4	838	40	277	1,155	76	4	73	878	71	14
18	2011Q1	830	37	281	1,148	76	4	72	866	71	14
19	2011Q2	839	41	279	1,159	76	5	72	880	67	10
20	2011Q3	830	35	280	1,145	76	4	72	865	70	13
21	2011Q4	842	35	278	1,154	76	4	73	877	69	13
22	2012Q1	843	43	283	1,169	76	5	72	886	72	11
23	2012Q2	837	38	296	1,172	75	4	71	875	66	11
24	2012Q3	833	38	298	1,169	74	4	71	871	67	13
25	2012Q4	833	41	298	1,173	75	5	71	874	63	12
26	2013Q1	832	36	295	1,162	75	4	72	868	70	12

Why Tables are Bad

	1	2	3	4	5	6	7	8	9	10	11
1	Labour Force Status by Sex by Sing/Comb Ethnic Group (Qrtly-Mar/Jun/Sep/Dec)										
2		Male									
3		European Only								Maori Only	
4		Persons Em	Persons Un	Not in Labo	Working Ag	Labour Forc	Unemploym	Employer	Total Labou	Persons Em	Persons Un
5	2007Q4	856	20	280	1,156	76	2	74	876	71	6
6	2008Q1	863	25	284	1,172	76	3	74	888	69	8
7	2008Q2	850	26	281	1,157	76	3	74	876	67	6
8	2008Q3	840	30	286	1,155	75	3	73	869	72	9
9	2008Q4	855	30	275	1,159	76	3	74	884	76	8
10	2009Q1	845	35	279	1,160	76	4	73	880	75	8
11	2009Q2	832	35	280	1,146	76	4	73	866	74	10
12	2009Q3	813	42	290	1,146	75	5	71	856	71	11
13	2009Q4	831	40	277	1,148	76	5	72	871	72	14
14	2010Q1	822	36	283	1,142	75	4	72	859	72	11
15	2010Q2	825	40	290	1,155	75	5	71	865	72	14
16	2010Q3	837	31	287	1,155	75	4	72	868	70	14
17	2010Q4	838	40	277	1,155	76	4	73	878	71	14
18	2011Q1	830	37	281	1,148	76	4	72	866	71	14
19	2011Q2	839	41	279	1,159	76	5	72	880	67	10
20	2011Q3	830	35	280	1,145	76	4	72	865	70	13
21	2011Q4	842	35	278	1,154	76	4	73	877	69	13
22	2012Q1	843	43	283	1,169	76	5	72	886	72	11
23	2012Q2	837	38	296	1,172	75	4	71	875	66	11
24	2012Q3	833	38	298	1,169	74	4	71	871	67	13
25	2012Q4	833	41	298	1,173	75	5	71	874	63	12
26	2013Q1	832	36	295	1,162	75	4	72	868	70	12

Why Tables are Bad

	1	2	3	4	5	6	7	8	9	10	11
1	Labour Force Status by Sex by Sing/Comb Ethnic Group (Qrtly-Mar/Jun/Sep/Dec)										
2		Male									
3		European C								Maori Only	
4		Persons En	Persons Un	Not in Labo	Working Ag	Labour Forc	Unemploym	Employer	Total Labou	Persons En	Persons Un
5	2007Q4	856	20	280	1,156	76	2	74	876	71	6
6	2008Q1	863	25	284	1,172	76	3	74	888	69	8
7	2008Q2	850	26	281	1,157	76	3	74	876	67	6
8	2008Q3	840	30	286	1,155	75	3	73	869	72	9
9	2008Q4	855	30	275	1,159	76	3	74	884	76	8
10	2009Q1	845	35	279	1,160	76	4	73	880	75	8
11	2009Q2	832	35	280	1,146	76	4	73	866	74	10
12	2009Q3	813	42	290	1,146	75	5	71	856	71	11
13	2009Q4	831	40	277	1,148	76	5	72	871	72	14
14	2010Q1	822	36	283	1,142	75	4	72	859	72	11
15	2010Q2	825	40	290	1,155	75	5	71	865	72	14
16	2010Q3	837	31	287	1,155	75	4	72	868	70	14
17	2010Q4	838	40	277	1,155	76	4	73	878	71	14
18	2011Q1	830	37	281	1,148	76	4	72	866	71	14
19	2011Q2	839	41	279	1,159	76	5	72	880	67	10
20	2011Q3	830	35	280	1,145	76	4	72	865	70	13
21	2011Q4	842	35	278	1,154	76	4	73	877	69	13
22	2012Q1	843	43	283	1,169	76	5	72	886	72	11
23	2012Q2	837	38	296	1,172	75	4	71	875	66	11
24	2012Q3	833	38	298	1,169	74	4	71	871	67	13
25	2012Q4	833	41	298	1,173	75	5	71	874	63	12
26	2013Q1	832	36	295	1,162	75	4	72	868	70	12

What are LongForm Dataframes?

	1	2	3	4	5	6	7	8	9	10	11
1				Persons En	Persons Un	Not in Labo	Working Ag	Labour Forc	Unemploym	Employmer	Total Labou
2	Male	European C	2007Q4	856	20	280	1,156	76	2	74	876
3	Male	European C	2008Q1	863	25	284	1,172	76	3	74	888
4	Male	European C	2008Q2	850	26	281	1,157	76	3	74	876
5	Male	European C	2008Q3	840	30	286	1,155	75	3	73	869
6	Male	European C	2008Q4	855	30	275	1,159	76	3	74	884
7	Male	European C	2009Q1	845	35	279	1,160	76	4	73	880
8	Male	European C	2009Q2	832	35	280	1,146	76	4	73	866
9	Male	European C	2009Q3	813	42	290	1,146	75	5	71	856
10	Male	European C	2009Q4	831	40	277	1,148	76	5	72	871
11	Male	European C	2010Q1	822	36	283	1,142	75	4	72	859
12	Male	European C	2010Q2	825	40	290	1,155	75	5	71	865
13	Male	European C	2010Q3	837	31	287	1,155	75	4	72	868
14	Male	European C	2010Q4	838	40	277	1,155	76	4	73	878
15	Male	European C	2011Q1	830	37	281	1,148	76	4	72	866
16	Male	European C	2011Q2	839	41	279	1,159	76	5	72	880
17	Male	European C	2011Q3	830	35	280	1,145	76	4	72	865
18	Male	European C	2011Q4	842	35	278	1,154	76	4	73	877
19	Male	European C	2012Q1	843	43	283	1,169	76	5	72	886
20	Male	European C	2012Q2	837	38	296	1,172	75	4	71	875
21	Male	European C	2012Q3	833	38	298	1,169	74	4	71	871
22	Male	European C	2012Q4	833	41	298	1,173	75	5	71	874
23	Male	European C	2013Q1	832	36	295	1,162	75	4	72	868
24	Male	Maori Only	2007Q4	71	6	28	105	73	8	68	77
25	Male	Maori Only	2008Q1	69	8	31	108	71	10	64	77
26	Male	Maori Only	2008Q2	67	6	27	100	73	8	67	73

Why LongForm is better?

	1	2	3	4	5	6	7	8	9	10	11
1				Persons En	Persons Un	Not in Labo	Working Ag	Labour Forc	Unemployr	Employer	Total Labou
2	Male	European C	2007Q4	856	20	280	1,156	76	2	74	876
3	Male	European C	2008Q1	863	25	284	1,172	76	3	74	888
4	Male	European C	2008Q2	850	26	281	1,157	76	3	74	876
5	Male	European C	2008Q3	840	30	286	1,155	75	3	73	869
6	Male	European C	2008Q4	855	30	275	1,159	76	3	74	884
7	Male	European C	2009Q1	845	35	279	1,160	76	4	73	880
8	Male	European C	2009Q2	832	35	280	1,146	76	4	73	866
9	Male	European C	2009Q3	813	42	290	1,146	75	5	71	856
10	Male	European C	2009Q4	831	40	277	1,148	76	5	72	871
11	Male	European C	2010Q1	822	36	283	1,142	75	4	72	859
12	Male	European C	2010Q2	825	40	290	1,155	75	5	71	865
13	Male	European C	2010Q3	837	31	287	1,155	75	4	72	868
14	Male	European C	2010Q4	838	40	277	1,155	76	4	73	878
15	Male	European C	2011Q1	830	37	281	1,148	76	4	72	866
16	Male	European C	2011Q2	839	41	279	1,159	76	5	72	880
17	Male	European C	2011Q3	830	35	280	1,145	76	4	72	865
18	Male	European C	2011Q4	842	35	278	1,154	76	4	73	877
19	Male	European C	2012Q1	843	43	283	1,169	76	5	72	886
20	Male	European C	2012Q2	837	38	296	1,172	75	4	71	875
21	Male	European C	2012Q3	833	38	298	1,169	74	4	71	871
22	Male	European C	2012Q4	833	41	298	1,173	75	5	71	874
23	Male	European C	2013Q1	832	36	295	1,162	75	4	72	868
24	Male	Maori Only	2007Q4	71	6	28	105	73	8	68	77
25	Male	Maori Only	2008Q1	69	8	31	108	71	10	64	77
26	Male	Maori Only	2008Q2	67	6	27	100	73	8	67	73

TableToLongForm

- ▶ TableToLongForm is an R package that automatically converts hierarchical *Tables* into *LongForm* R `data.frame` objects.
- ▶ Once so converted, the data is machine readable and can be fed into your favourite tool, whatever that may be.

User Manual

Step 1

- ▶ Read in the original Table.

```
LabourForce = read.csv("StatsNZLabourForce.csv",  
                      header = FALSE, na.strings = c("", ".."))
```

Step 2

- ▶ Call `TableToLongForm`

```
TableToLongForm(LabourForce)
```

Comparison with Manual Conversion

	1	2	3	4	5	6	7	8	9	10	11
1	Labour Force Status by Sex by Sing/Comb Ethnic Group (Qrtly-Mar/Jun/Sep/Dec)										
2		Male									
3		European Only								Maori Only	
4		Persons Err	Persons Un	Not in Labo	Working Ag	Labour Forc	Unemploym	Employer	Total Labou	Persons Err	Persons Un
5	2007Q4	856	20	280	1,156	76	2	74	876	71	6
6	2008Q1	863	25	284	1,172	76	3	74	888	69	8
7	2008Q2	850	26	281	1,157	76	3	74	876	67	6
8	2008Q3	840	30	286	1,155	75	3	73	869	72	9
9	2008Q4	855	30	275	1,159	76	3	74	884	76	8
10	2009Q1	845	35	279	1,160	76	4	73	880	75	8
11	2009Q2	832	35	280	1,146	76	4	73	866	74	10
12	2009Q3	813	42	290	1,146	75	5	71	856	71	11
13	2009Q4	831	40	277	1,148	76	5	72	871	72	14
14	2010Q1	822	36	283	1,142	75	4	72	859	72	11
15	2010Q2	825	40	290	1,155	75	5	71	865	72	14
16	2010Q3	837	31	287	1,155	75	4	72	868	70	14
17	2010Q4	838	40	277	1,155	76	4	73	878	71	14
18	2011Q1	830	37	281	1,148	76	4	72	866	71	14
19	2011Q2	839	41	279	1,159	76	5	72	880	67	10
20	2011Q3	830	35	280	1,145	76	4	72	865	70	13
21	2011Q4	842	35	278	1,154	76	4	73	877	69	13
22	2012Q1	843	43	283	1,169	76	5	72	886	72	11
23	2012Q2	837	38	296	1,172	75	4	71	875	66	11
24	2012Q3	833	38	298	1,169	74	4	71	871	67	13
25	2012Q4	833	41	298	1,173	75	5	71	874	63	12
26	2013Q1	832	36	295	1,162	75	4	72	868	70	12

Comparison with Manual Conversion

We note after careful observation, that there are:

- ▶ 3 gender categories: 'Male', 'Female' and 'Total Both Sexes', each a consistent 80 columns in width.
- ▶ 10 ethnic categories, each a consistent 8 columns in width.
- ▶ The data are found in rows 5 to 26.

Comparison with Manual Conversion

```
LFout = NULL
chYear = LabourForce[5:26, 1]
for(Gender in 0:2)
  for(Ethni in 0:9){
    chGender = LabourForce[2, 2 + Gender * 80]
    chEthni = LabourForce[3, 2 + Ethni * 8]
    LFout = rbind(LFout,
      cbind(chGender, chEthni, chYear,
        LabourForce[5:26, 2 + Gender * 80 +
          (Ethni * 8):((Ethni + 1) * 8 - 1)])
    )
  }
colnames(LFout) = c("Gender", "Ethnicity",
  "Time.Period", LabourForce[4, 2:9])
```

Comparison with Manual Conversion

Even for a relatively tidy Table,
Manual Conversion is:

- ▶ Time-consuming.
- ▶ Only works for this **specific table** with the **exact specifications** mentioned in a previous slide.
- ▶ Prone to error.

Recognised Patterns

- ▶ TableToLongForm consists of a number of algorithms that can collectively process a variety of so-called *recognised patterns*.

Understanding Parentage

	1	2
1	Accounting	
2	NZ Maori	Male
3		Female
4	NZ European	Male
5		Female
6	Biology	
7	NZ Maori	Male
8		Female
9	NZ European	Male
10		Female

Identify Tell-tale Features

	1	2
1	Accounting	
2	NZ Maori	Male
3		Female
4	NZ European	Male
5		Female
6	Biology	
7	NZ Maori	Male
8		Female
9	NZ European	Male
10		Female

Divide and Conquer

	1	2
1	Accounting	
2	NZ Maori	Male
3		Female
4	NZ European	Male
5		Female
6	Biology	
7	NZ Maori	Male
8		Female
9	NZ European	Male
10		Female

Divide and Conquer

	1	2
1	Accounting	
2	NZ Maori	Male
3		Female
4	NZ European	Male
5		Female
6	Biology	
7	NZ Maori	Male
8		Female
9	NZ European	Male
10		Female

Divide and Conquer

	1	2
1	Accounting	
2	NZ Maori	Male
3		Female
4	NZ European	Male
5		Female
6	Biology	
7	NZ Maori	Male
8		Female
9	NZ European	Male
10		Female

NZQA Table

	1	2	3	4	5	6	7	8	9	10	11	12	13
1	Scholarship Entries and Results by Gender and Ethnicity (Broken down by Decile)												
2													
3				Decile 1-3									Decile 4-7
4			# of	#	#		#	# Not	#	#		# of	#
5	Results		Entries	Absent	SNA		Assessed	Achieved	Scholarship	Outstanding		Entries	Absent
6													
7	All Subjects		714	148	13		553	462	81	10		6,482	1,772
8													
9	Accounting		22	4	0		18	16	2	0		156	41
10	NZ Maori	Male	2	1	0		1	1	0	0		2	1
11		Female	0	0	0		0	0	0	0		7	2
12	NZ European	Male	2	0	0		2	1	1	0		51	13
13		Female	3	0	0		3	2	1	0		44	12
14		Unknown	0	0	0		0	0	0	0		0	0
15	Pasifika Pacific	Male	2	0	0		2	2	0	0		3	0
16		Female	6	2	0		4	4	0	0		4	2
17	Asian	Male	5	0	0		5	5	0	0		29	4
18		Female	2	1	0		1	1	0	0		15	7
19	Other/Unspecified	Male	0	0	0		0	0	0	0		0	0
20		Female	0	0	0		0	0	0	0		1	0
21													
22	Agricultural & Horticulture		0	0	0		0	0	0	0		15	3
23	NZ Maori	Male	0	0	0		0	0	0	0		0	0
24		Female	0	0	0		0	0	0	0		0	0
25	NZ European	Male	0	0	0		0	0	0	0		10	1
26		Female	0	0	0		0	0	0	0		5	2
27		Unknown	0	0	0		0	0	0	0		0	0
28	Pasifika Pacific	Male	0	0	0		0	0	0	0		0	0
29		Female	0	0	0		0	0	0	0		0	0
30	Asian	Male	0	0	0		0	0	0	0		0	0

NZQA Table

	1	2	3	4	5	6	7	8	9	10	11	12	13
1	Scholarship Entries and Results by Gender and Ethnicity (Broken down by Decile)												
2													
3				Decile 1-3									Decile 4-7
4			# of	#	#		#	# Not	#	#		# of	#
5	Results		Entries	Absent	SNA		Assessed	Achieved	Scholarship	Outstanding		Entries	Absent
6													
7	All Subjects		714	148	13		553	462	81	10		6,482	1,772
8													
9	Accounting		22	4	0		18	16	2	0		156	41
10	NZ Maori	Male	2	1	0		1	1	0	0		2	1
11		Female	0	0	0		0	0	0	0		7	2
12	NZ European	Male	2	0	0		2	1	1	0		51	13
13		Female	3	0	0		3	2	1	0		44	12
14		Unknown	0	0	0		0	0	0	0		0	0
15	Pasifika Pacific	Male	2	0	0		2	2	0	0		3	0
16		Female	6	2	0		4	4	0	0		4	2
17	Asian	Male	5	0	0		5	5	0	0		29	4
18		Female	2	1	0		1	1	0	0		15	7
19	Other/Unspecified	Male	0	0	0		0	0	0	0		0	0
20		Female	0	0	0		0	0	0	0		1	0
21													
22	Agricultural & Horticulture		0	0	0		0	0	0	0		15	3
23	NZ Maori	Male	0	0	0		0	0	0	0		0	0
24		Female	0	0	0		0	0	0	0		0	0
25	NZ European	Male	0	0	0		0	0	0	0		10	1
26		Female	0	0	0		0	0	0	0		5	2
27		Unknown	0	0	0		0	0	0	0		0	0
28	Pasifika Pacific	Male	0	0	0		0	0	0	0		0	0
29		Female	0	0	0		0	0	0	0		0	0
30	Asian	Male	0	0	0		0	0	0	0		0	0

NZQA Table

	1	2	3	4	5	6	7	8	9	10	11	12	13
1	Scholarship Entries and Results by Gender and Ethnicity (Broken down by Decile)												
2													
3				Decile 1-3									Decile 4-7
4			# of	#	#		#	# Not	#	#		# of	#
5	Results		Entries	Absent	SNA		Assessed	Achieved	Scholarship	Outstanding		Entries	Absent
6													
7	All Subjects		714	148	13		553	462	81	10		6,482	1,772
8													
9	Accounting		22	4	0		18	16	2	0		156	41
10	NZ Maori	Male	2	1	0		1	1	0	0		2	1
11		Female	0	0	0		0	0	0	0		7	2
12	NZ European	Male	2	0	0		2	1	1	0		51	13
13		Female	3	0	0		3	2	1	0		44	12
14		Unknown	0	0	0		0	0	0	0		0	0
15	Pasifika Pacific	Male	2	0	0		2	2	0	0		3	0
16		Female	6	2	0		4	4	0	0		4	2
17	Asian	Male	5	0	0		5	5	0	0		29	4
18		Female	2	1	0		1	1	0	0		15	7
19	Other/Unspecified	Male	0	0	0		0	0	0	0		0	0
20		Female	0	0	0		0	0	0	0		1	0
21													
22	Agricultural & Horticulture		0	0	0		0	0	0	0		15	3
23	NZ Maori	Male	0	0	0		0	0	0	0		0	0
24		Female	0	0	0		0	0	0	0		0	0
25	NZ European	Male	0	0	0		0	0	0	0		10	1
26		Female	0	0	0		0	0	0	0		5	2
27		Unknown	0	0	0		0	0	0	0		0	0
28	Pasifika Pacific	Male	0	0	0		0	0	0	0		0	0
29		Female	0	0	0		0	0	0	0		0	0
30	Asian	Male	0	0	0		0	0	0	0		0	0

NZQA Table

	1	2	3	4	5	6	7	8	9	10	11	12	13
1	Scholarship Entries and Results by Gender and Ethnicity (Broken down by Decile)												
2													
3				Decile 1-3									Decile 4-7
4			# of	#	#		#	# Not	#	#		# of	#
5	Results		Entries	Absent	SNA		Assessed	Achieved	Scholarship	Outstanding		Entries	Absent
6													
7	All Subjects		714	148	13		553	462	81	10		6,482	1,772
8													
9	Accounting		22	4	0		18	16	2	0		156	41
10	NZ Maori	Male	2	1	0		1	1	0	0		2	1
11		Female	0	0	0		0	0	0	0		7	2
12	NZ European	Male	2	0	0		2	1	1	0		51	13
13		Female	3	0	0		3	2	1	0		44	12
14		Unknown	0	0	0		0	0	0	0		0	0
15	Pasifika Pacific	Male	2	0	0		2	2	0	0		3	0
16		Female	6	2	0		4	4	0	0		4	2
17	Asian	Male	5	0	0		5	5	0	0		29	4
18		Female	2	1	0		1	1	0	0		15	7
19	Other/Unspecified	Male	0	0	0		0	0	0	0		0	0
20		Female	0	0	0		0	0	0	0		1	0
21													
22	Agricultural & Horticulture		0	0	0		0	0	0	0		15	3
23	NZ Maori	Male	0	0	0		0	0	0	0		0	0
24		Female	0	0	0		0	0	0	0		0	0
25	NZ European	Male	0	0	0		0	0	0	0		10	1
26		Female	0	0	0		0	0	0	0		5	2
27		Unknown	0	0	0		0	0	0	0		0	0
28	Pasifika Pacific	Male	0	0	0		0	0	0	0		0	0
29		Female	0	0	0		0	0	0	0		0	0
30	Asian	Male	0	0	0		0	0	0	0		0	0

NZQA Table

	1	2	3	4	5	6	7	8	9	10	11	12	13
1	Scholarship Entries and Results by Gender and Ethnicity (Broken down by Decile)												
2													
3				Decile 1-3									Decile 4-7
4			# of	#	#		#	# Not	#	#		# of	#
5	Results		Entries	Absent	SNA		Assessed	Achieved	Scholarship	Outstanding		Entries	Absent
6													
7	All Subjects		714	148	13		553	462	81	10		6,482	1,772
8													
9	Accounting		22	4	0		18	16	2	0		156	41
10	NZ Maori	Male	2	1	0		1	1	0	0		2	1
11		Female	0	0	0		0	0	0	0		7	2
12	NZ European	Male	2	0	0		2	1	1	0		51	13
13		Female	3	0	0		3	2	1	0		44	12
14		Unknown	0	0	0		0	0	0	0		0	0
15	Pasifika Pacific	Male	2	0	0		2	2	0	0		3	0
16		Female	6	2	0		4	4	0	0		4	2
17	Asian	Male	5	0	0		5	5	0	0		29	4
18		Female	2	1	0		1	1	0	0		15	7
19	Other/Unspecified	Male	0	0	0		0	0	0	0		0	0
20		Female	0	0	0		0	0	0	0		1	0
21													
22	Agricultural & Horticulture		0	0	0		0	0	0	0		15	3
23	NZ Maori	Male	0	0	0		0	0	0	0		0	0
24		Female	0	0	0		0	0	0	0		0	0
25	NZ European	Male	0	0	0		0	0	0	0		10	1
26		Female	0	0	0		0	0	0	0		5	2
27		Unknown	0	0	0		0	0	0	0		0	0
28	Pasifika Pacific	Male	0	0	0		0	0	0	0		0	0
29		Female	0	0	0		0	0	0	0		0	0
30	Asian	Male	0	0	0		0	0	0	0		0	0

NZQA Table After Conversion

1	2	3	4	5	6	7	8	9	10	11	12	13
					# of Entries	# Absent	# SNA	# Assessed	# Not Achieved	# Scholars	# Outstanding	
2	Decile 1-3	All Subjects	Accounting	NZ Maori	Male	2	1	0	1	1	0	0
3	Decile 1-3	All Subjects	Accounting	NZ Maori	Female	0	0	0	0	0	0	0
4	Decile 1-3	All Subjects	Accounting	NZ Europe	Male	2	0	0	2	1	1	0
5	Decile 1-3	All Subjects	Accounting	NZ Europe	Female	3	0	0	3	2	1	0
6	Decile 1-3	All Subjects	Accounting	NZ Europe	Unknown	0	0	0	0	0	0	0
7	Decile 1-3	All Subjects	Accounting	Pasifika Pe	Male	2	0	0	2	2	0	0
8	Decile 1-3	All Subjects	Accounting	Pasifika Pe	Female	6	2	0	4	4	0	0
9	Decile 1-3	All Subjects	Accounting	Asian	Male	5	0	0	5	5	0	0
10	Decile 1-3	All Subjects	Accounting	Asian	Female	2	1	0	1	1	0	0
11	Decile 1-3	All Subjects	Accounting	Other/Unsp	Male	0	0	0	0	0	0	0
12	Decile 1-3	All Subjects	Accounting	Other/Unsp	Female	0	0	0	0	0	0	0
13	Decile 1-3	All Subjects	Agricultural	NZ Maori	Male	0	0	0	0	0	0	0
14	Decile 1-3	All Subjects	Agricultural	NZ Maori	Female	0	0	0	0	0	0	0
15	Decile 1-3	All Subjects	Agricultural	NZ Europe	Male	0	0	0	0	0	0	0
16	Decile 1-3	All Subjects	Agricultural	NZ Europe	Female	0	0	0	0	0	0	0
17	Decile 1-3	All Subjects	Agricultural	NZ Europe	Unknown	0	0	0	0	0	0	0
18	Decile 1-3	All Subjects	Agricultural	Pasifika Pe	Male	0	0	0	0	0	0	0
19	Decile 1-3	All Subjects	Agricultural	Pasifika Pe	Female	0	0	0	0	0	0	0
20	Decile 1-3	All Subjects	Agricultural	Asian	Male	0	0	0	0	0	0	0
21	Decile 1-3	All Subjects	Agricultural	Asian	Female	0	0	0	0	0	0	0
22	Decile 1-3	All Subjects	Agricultural	Other/Unsp	Male	0	0	0	0	0	0	0
23	Decile 1-3	All Subjects	Agricultural	Other/Unsp	Female	0	0	0	0	0	0	0
24	Decile 1-3	All Subjects	Art History	NZ Maori	Male	0	0	0	0	0	0	0
25	Decile 1-3	All Subjects	Art History	NZ Maori	Female	2	0	0	2	2	0	0
26	Decile 1-3	All Subjects	Art History	NZ Europe	Male	1	1	0	0	0	0	0
27	Decile 1-3	All Subjects	Art History	NZ Europe	Female	3	0	0	3	3	0	0
28	Decile 1-3	All Subjects	Art History	NZ Europe	Unknown	0	0	0	0	0	0	0
29	Decile 1-3	All Subjects	Art History	Pasifika Pe	Male	0	0	0	0	0	0	0
30	Decile 1-3	All Subjects	Art History	Pasifika Pe	Female	0	0	0	0	0	0	0

Concluding Remarks

- ▶ TableToLongForm
- ▶ An R package
- ▶ Automatically convert Tables to LongForm Dataframe

Concluding Remarks

- ▶ Can handle a variety of Tables automatically
- ▶ But not without limitations

Concluding Remarks

- ▶ Ultimately, still uses algorithms
- ▶ Detect a **known set** of *recognised patterns*
- ▶ If the Table deviates from these patterns, TableToLongForm fails.

Concluding Remarks

- ▶ Diagnostics needs more work

```
Error in 1:ncol(datbit) : argument of length 0
In addition: Warning message:
In rbind(matColLabel[!fullrows, , drop = FALSE],
  collapsedlabels) :
number of columns of result is not a multiple
of vector length (arg 2)
```

Concluding Remarks

- ▶ `TableToLongForm` is available from CRAN.

References

- ▶ NZQA (2012) URL <http://www.nzqa.govt.nz/>
- ▶ R Development Core Team (2012) URL <http://www.R-project.org/>
- ▶ Statistics New Zealand (2013) URL <http://www.stats.govt.nz/infoshare/>