

Example 1: real data report: multiple-source education register

Data set and main information needed

about: time-variables, modeled - variables, imputed-variables

Education register, after combining several administrative sources and imputing level and field of education for many records. The variable `info_source` identifies the ones which are imputed (when it has the value “tilreikna”).

For this report, a large sample from the whole register is extracted, for speed and memory reasons but this can be easily run for the whole set.

Output of this report: pdf, html, word.

Plots are now static. We will produce soon interactive ones.

Dashboards are possible

Main packages and resources:

ggplot2, DataExplorer, funModeling, tabplot, forecast, tsfeatures, anomalous

view_data

Overview of main data-set characteristics

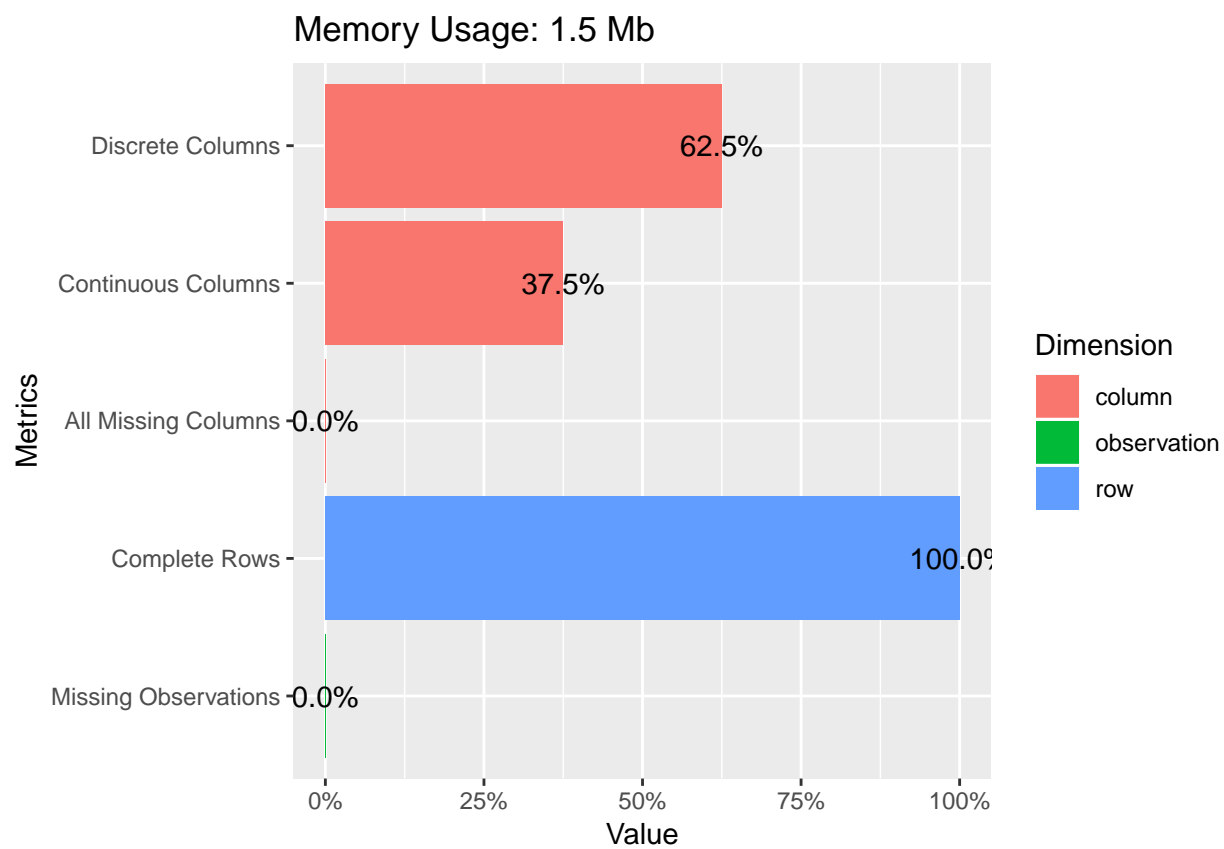
```
## df
##
## 8 Variables      30000 Observations
## -----
## gender
##      n missing distinct
## 29988      12         2
##
## Value      1      2
## Frequency 16098 13890
## Proportion 0.537 0.463
## -----
## info_source
##      n missing distinct
## 30000      0         24
##
## lowest : Atvinnuvega- og nýsköpunarráðuneyti Danska hagstofan
## highest: Sænska hagstofan      Tilreiknað
## -----
## edu
##      n missing distinct
## 30000      0         31
##
## lowest : 00 11 20 21 22, highest: 61 62 71 72 80
## -----
## edu_field
##      n missing distinct
```

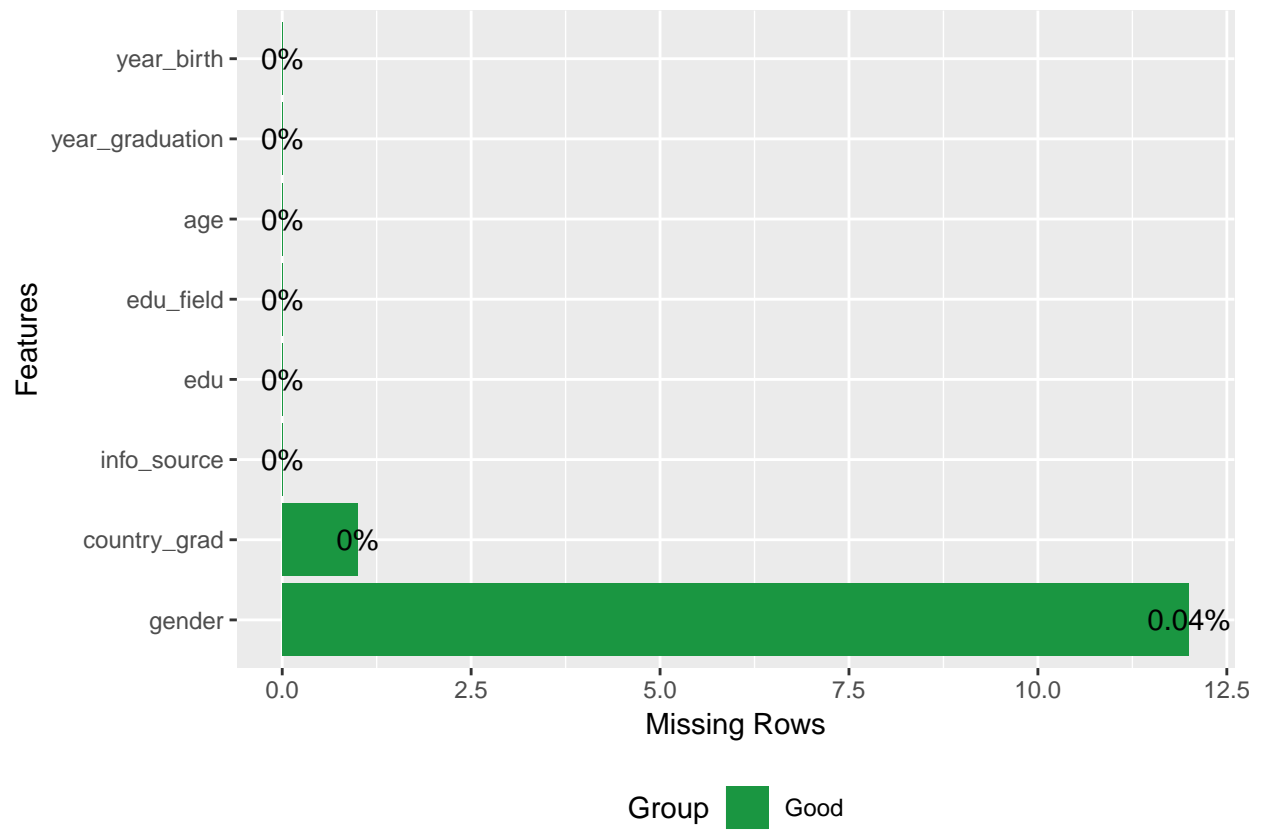
EldriPrófaskrár
Útlendingastofnun

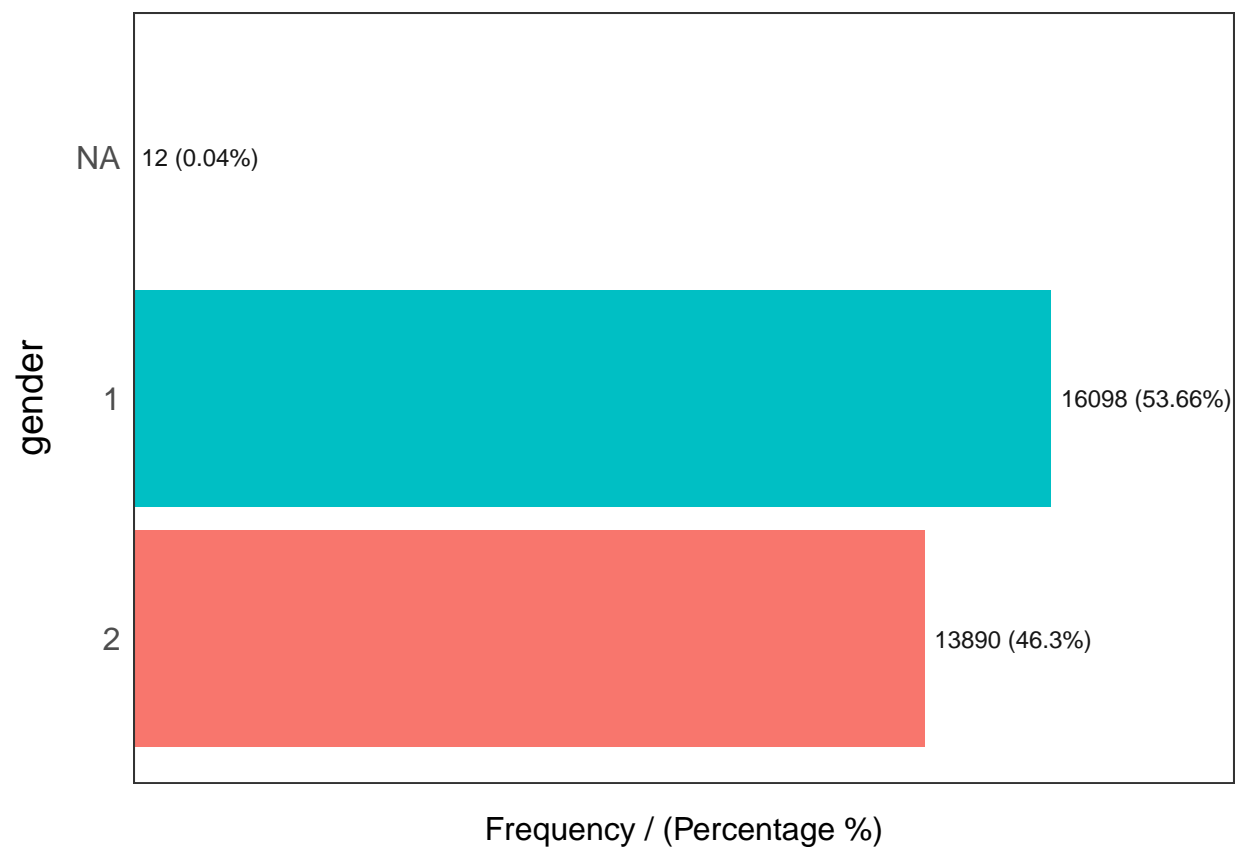
```

##      30000          0          442
##
## lowest : 0... 010. 010a 010b 010c, highest: 861. 861a 863. 999. ooo
## -----
## country_grad
##      n missing distinct
##    29999          1          95
##
## lowest : AL AR AT AU AZ, highest: VU XX YU YY ZA
## -----
## age
##      n missing distinct      Info      Mean      Gmd      .05      .10
##    30000          0          57    0.981    22.37    7.706      16      16
##      .25      .50      .75      .90      .95
##      16      20      26      31      38
##
## lowest : 13 14 15 16 17, highest: 65 66 69 71 78
## -----
## year_graduation
##      n missing distinct      Info      Mean      Gmd      .05      .10
##    30000          0          122         1    1990    27.57    1937    1951
##      .25      .50      .75      .90      .95
##    1977    1998    2011    2016    2017
##
## lowest : 1879 1880 1882 1895 1897, highest: 2014 2015 2016 2017 2018
## -----
## year_birth
##      n missing distinct      Info      Mean      Gmd      .05      .10
##    30000          0          126         1    1968    27.47    1918    1931
##      .25      .50      .75      .90      .95
##    1954    1973    1987    1996    2000
##
## lowest : 1863 1864 1866 1879 1881, highest: 2001 2002 2003 2004 2005
## -----

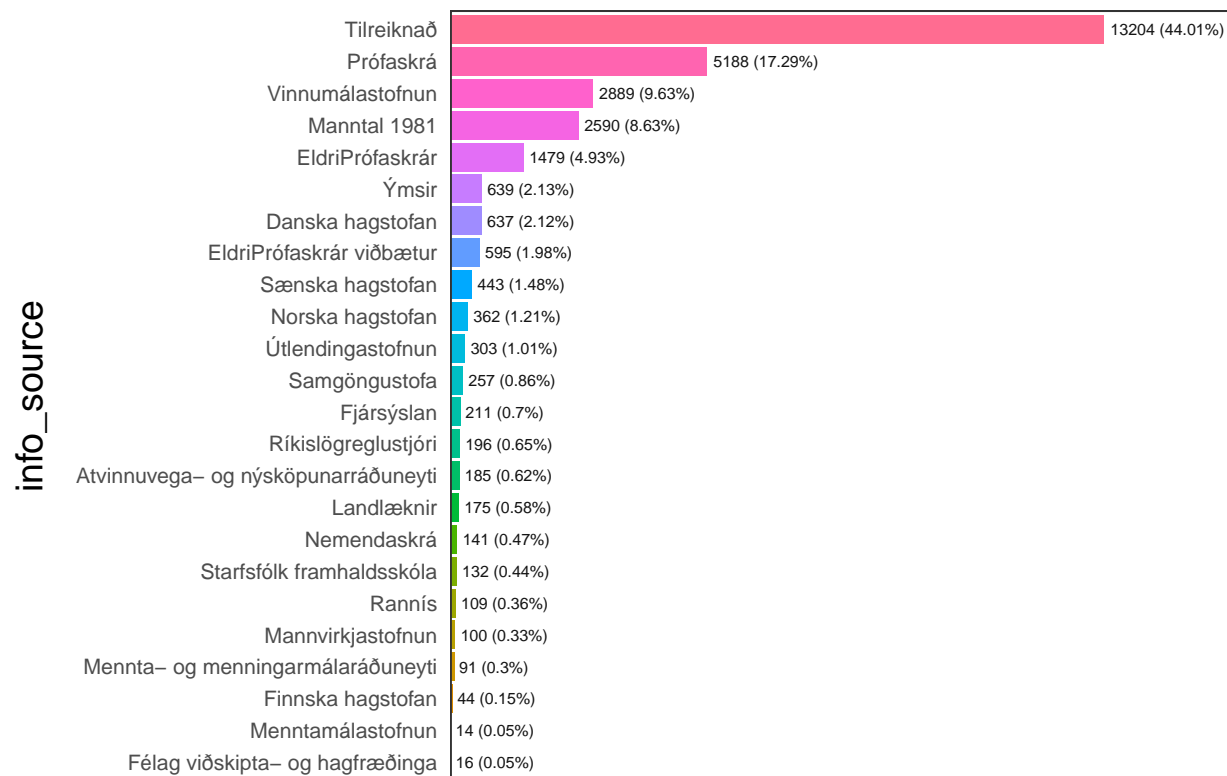
```







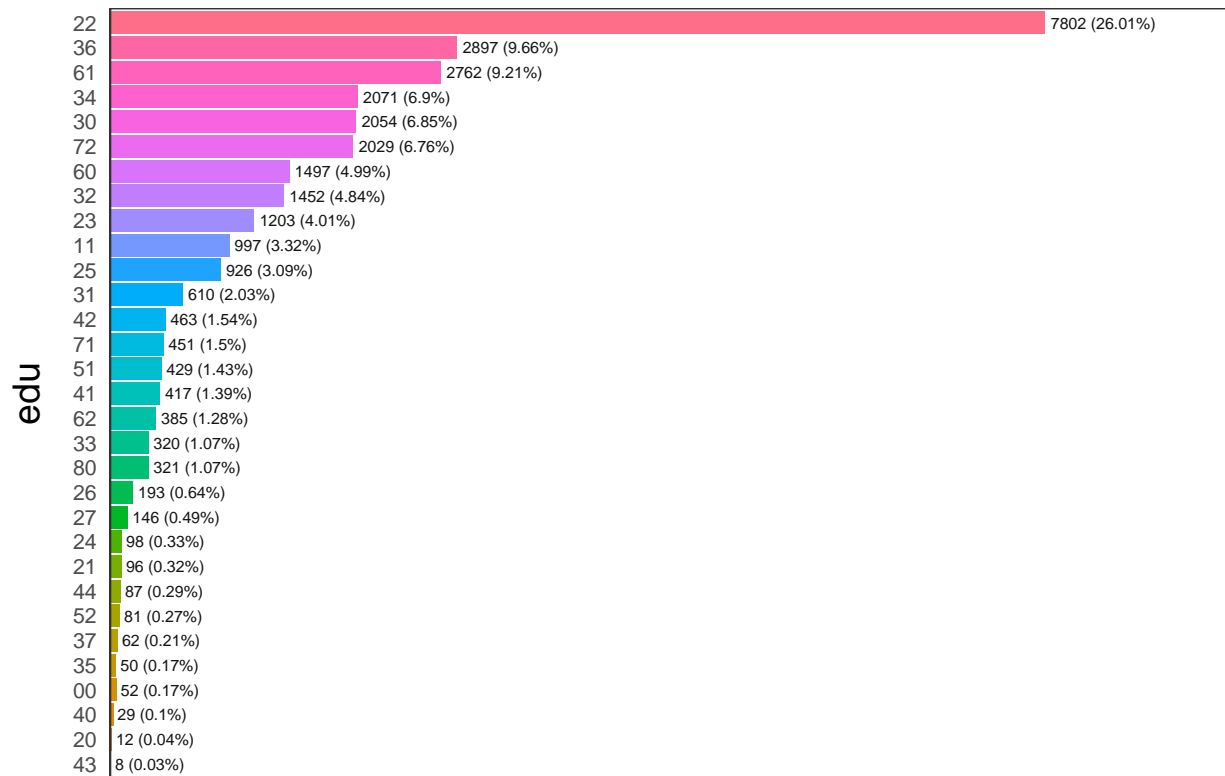
##	gender	frequency	percentage	cumulative_perc
## 1	1	16098	53.66	53.66
## 2	2	13890	46.30	99.96
## 3	<NA>	12	0.04	100.00



Frequency / (Percentage %)

##	info_source	frequency	percentage
## 1	Tilreiknað	13204	44.01
## 2	Prófaskrá	5188	17.29
## 3	Vinnumálastofnun	2889	9.63
## 4	Manntal 1981	2590	8.63
## 5	EldriPrófaskrár	1479	4.93
## 6	Ýmsir	639	2.13
## 7	Danska hagstofan	637	2.12
## 8	EldriPrófaskrár viðbætur	595	1.98
## 9	Sænska hagstofan	443	1.48
## 10	Norska hagstofan	362	1.21
## 11	Útlendingastofnun	303	1.01
## 12	Samgöngustofa	257	0.86
## 13	Fjársýslan	211	0.70
## 14	Ríkislögreglustjóri	196	0.65
## 15	Atvinnuvega- og nýsköpunarráðuneyti	185	0.62
## 16	Landlæknir	175	0.58
## 17	Nemendaskrá	141	0.47
## 18	Starfsfólk framhaldsskóla	132	0.44
## 19	Rannís	109	0.36
## 20	Mannvirkjastofnun	100	0.33
## 21	Mennta- og menningarmálaráðuneyti	91	0.30
## 22	Finnska hagstofan	44	0.15
## 23	Félag viðskipta- og hagfræðinga	16	0.05
## 24	Menntamálastofnun	14	0.05
##	cumulative_perc		

```
## 1      44.01
## 2      61.30
## 3      70.93
## 4      79.56
## 5      84.49
## 6      86.62
## 7      88.74
## 8      90.72
## 9      92.20
## 10     93.41
## 11     94.42
## 12     95.28
## 13     95.98
## 14     96.63
## 15     97.25
## 16     97.83
## 17     98.30
## 18     98.74
## 19     99.10
## 20     99.43
## 21     99.73
## 22     99.88
## 23     99.93
## 24    100.00
```



Frequency / (Percentage %)

```
##      edu frequency percentage cumulative_perc
## 1    22      7802      26.01      26.01
```

##	2	36	2897	9.66	35.67
##	3	61	2762	9.21	44.88
##	4	34	2071	6.90	51.78
##	5	30	2054	6.85	58.63
##	6	72	2029	6.76	65.39
##	7	60	1497	4.99	70.38
##	8	32	1452	4.84	75.22
##	9	23	1203	4.01	79.23
##	10	11	997	3.32	82.55
##	11	25	926	3.09	85.64
##	12	31	610	2.03	87.67
##	13	42	463	1.54	89.21
##	14	71	451	1.50	90.71
##	15	51	429	1.43	92.14
##	16	41	417	1.39	93.53
##	17	62	385	1.28	94.81
##	18	80	321	1.07	95.88
##	19	33	320	1.07	96.95
##	20	26	193	0.64	97.59
##	21	27	146	0.49	98.08
##	22	24	98	0.33	98.41
##	23	21	96	0.32	98.73
##	24	44	87	0.29	99.02
##	25	52	81	0.27	99.29
##	26	37	62	0.21	99.50
##	27	00	52	0.17	99.67
##	28	35	50	0.17	99.84
##	29	40	29	0.10	99.94
##	30	20	12	0.04	99.98
##	31	43	8	0.03	100.00
##					
##		edu_field	frequency	percentage	cumulative_perc
##	1	010.	11597	38.66	38.66
##	2	ooo	3259	10.86	49.52
##	3	010a	587	1.96	51.48
##	4	010c	479	1.60	53.08
##	5	840.	458	1.53	54.61
##	6	582b	406	1.35	55.96
##	7	814.	343	1.14	57.10
##	8	144a	301	1.00	58.10
##	9	840a	275	0.92	59.02
##	10	380a	256	0.85	59.87
##	11	340.	243	0.81	60.68
##	12	525f	241	0.80	61.48
##	13	522d	223	0.74	62.22
##	14	723.	222	0.74	62.96
##	15	723c	220	0.73	63.69
##	16	340b	216	0.72	64.41
##	17	525c	200	0.67	65.08
##	18	010b	186	0.62	65.70
##	19	481c	179	0.60	66.30
##	20	582.	166	0.55	66.85
##	21	723b	163	0.54	67.39
##	22	521i	159	0.53	67.92

## 23	521.	152	0.51	68.43
## 24	721a	144	0.48	68.91
## 25	815a	142	0.47	69.38
## 26	010i	141	0.47	69.85
## 27	010t	139	0.46	70.31
## 28	811e	134	0.45	70.76
## 29	143a	131	0.44	71.20
## 30	520.	121	0.40	71.60
## 31	343a	120	0.40	72.00
## 32	621.	117	0.39	72.39
## 33	311a	116	0.39	72.78
## 34	340a	116	0.39	73.17
## 35	145a	109	0.36	73.53
## 36	010z	105	0.35	73.88
## 37	345f	94	0.31	74.19
## 38	342a	91	0.30	74.49
## 39	523.	90	0.30	74.79
## 40	726b	88	0.29	75.08
## 41	010f	85	0.28	75.36
## 42	521j	84	0.28	75.64
## 43	34..	82	0.27	75.91
## 44	225c	80	0.27	76.18
## 45	420.	80	0.27	76.45
## 46	541.	78	0.26	76.71
## 47	721.	77	0.26	76.97
## 48	142a	76	0.25	77.22
## 49	214.	76	0.25	77.47
## 50	812a	76	0.25	77.72
## 51	811.	74	0.25	77.97
## 52	213.	73	0.24	78.21
## 53	345.	73	0.24	78.45
## 54	146a	71	0.24	78.69
## 55	010ó	69	0.23	78.92
## 56	582e	68	0.23	79.15
## 57	347a	67	0.22	79.37
## 58	090a	66	0.22	79.59
## 59	222f	66	0.22	79.81
## 60	346.	66	0.22	80.03
## 61	222b	65	0.22	80.25
## 62	762a	65	0.22	80.47
## 63	144.	64	0.21	80.68
## 64	314a	64	0.21	80.89
## 65	582d	61	0.20	81.09
## 66	522.	59	0.20	81.29
## 67	212.	58	0.19	81.48
## 68	582c	54	0.18	81.66
## 69	815b	54	0.18	81.84
## 70	144c	53	0.18	82.02
## 71	344a	52	0.17	82.19
## 72	582i	52	0.17	82.36
## 73	211.	50	0.17	82.53
## 74	222.	50	0.17	82.70
## 75	313a	50	0.17	82.87
## 76	542.	50	0.17	83.04

## 77	144b	49	0.16	83.20
## 78	541a	48	0.16	83.36
## 79	726.	48	0.16	83.52
## 80	761.	48	0.16	83.68
## 81	344b	47	0.16	83.84
## 82	52..	47	0.16	84.00
## 83	999.	47	0.16	84.16
## 84	145.	46	0.15	84.31
## 85	525.	46	0.15	84.46
## 86	421a	44	0.15	84.61
## 87	581.	44	0.15	84.76
## 88	621a	44	0.15	84.91
## 89	727b	44	0.15	85.06
## 90	010l	43	0.14	85.20
## 91	312a	43	0.14	85.34
## 92	543a	43	0.14	85.48
## 93	815.	43	0.14	85.62
## 94	840b	43	0.14	85.76
## 95	812c	42	0.14	85.90
## 96	322b	41	0.14	86.04
## 97	541b	41	0.14	86.18
## 98	811b	41	0.14	86.32
## 99	814a	40	0.13	86.45
## 100	213f	39	0.13	86.58
## 101	523a	39	0.13	86.71
## 102	211a	38	0.13	86.84
## 103	48..	38	0.13	86.97
## 104	481b	38	0.13	87.10
## 105	725a	38	0.13	87.23
## 106	581a	37	0.12	87.35
## 107	346c	36	0.12	87.47
## 108	522j	35	0.12	87.59
## 109	543.	35	0.12	87.71
## 110	723a	35	0.12	87.83
## 111	2...	34	0.11	87.94
## 112	811a	34	0.11	88.05
## 113	213d	33	0.11	88.16
## 114	223c	33	0.11	88.27
## 115	421.	33	0.11	88.38
## 116	521c	33	0.11	88.49
## 117	311.	32	0.11	88.60
## 118	349.	32	0.11	88.71
## 119	422c	32	0.11	88.82
## 120	522e	32	0.11	88.93
## 121	313d	31	0.10	89.03
## 122	314.	31	0.10	89.13
## 123	345n	31	0.10	89.23
## 124	443b	31	0.10	89.33
## 125	521e	31	0.10	89.43
## 126	582a	31	0.10	89.53
## 127	725.	31	0.10	89.63
## 128	221b	30	0.10	89.73
## 129	010h	29	0.10	89.83
## 130	22..	29	0.10	89.93

## 131	312b	29	0.10	90.03
## 132	525g	29	0.10	90.13
## 133	720a	29	0.10	90.23
## 134	010g	28	0.09	90.32
## 135	443.	28	0.09	90.41
## 136	523c	28	0.09	90.50
## 137	443a	27	0.09	90.59
## 138	582h	27	0.09	90.68
## 139	726d	27	0.09	90.77
## 140	141.	26	0.09	90.86
## 141	520b	26	0.09	90.95
## 142	010é	25	0.08	91.03
## 143	21..	25	0.08	91.11
## 144	221.	25	0.08	91.19
## 145	226a	25	0.08	91.27
## 146	346b	25	0.08	91.35
## 147	811d	25	0.08	91.43
## 148	010m	24	0.08	91.51
## 149	213e	24	0.08	91.59
## 150	223.	24	0.08	91.67
## 151	225.	24	0.08	91.75
## 152	343.	24	0.08	91.83
## 153	345c	24	0.08	91.91
## 154	422a	24	0.08	91.99
## 155	521a	24	0.08	92.07
## 156	210.	23	0.08	92.15
## 157	321b	23	0.08	92.23
## 158	442a	23	0.08	92.31
## 159	481.	23	0.08	92.39
## 160	525b	23	0.08	92.47
## 161	720.	23	0.08	92.55
## 162	212e	22	0.07	92.62
## 163	521d	22	0.07	92.69
## 164	525d	22	0.07	92.76
## 165	541e	22	0.07	92.83
## 166	621b	22	0.07	92.90
## 167	621c	22	0.07	92.97
## 168	762.	22	0.07	93.04
## 169	861.	22	0.07	93.11
## 170	341.	21	0.07	93.18
## 171	522c	21	0.07	93.25
## 172	010r	20	0.07	93.32
## 173	213c	20	0.07	93.39
## 174	622.	20	0.07	93.46
## 175	811f	20	0.07	93.53
## 176	813.	20	0.07	93.60
## 177	312.	19	0.06	93.66
## 178	313b	19	0.06	93.72
## 179	345e	19	0.06	93.78
## 180	441a	19	0.06	93.84
## 181	723d	19	0.06	93.90
## 182	726a	19	0.06	93.96
## 183	727.	19	0.06	94.02
## 184	812.	19	0.06	94.08

## 185	215.	18	0.06	94.14
## 186	225a	18	0.06	94.20
## 187	345j	18	0.06	94.26
## 188	380c	18	0.06	94.32
## 189	461a	18	0.06	94.38
## 190	529x	18	0.06	94.44
## 191	541f	18	0.06	94.50
## 192	542b	18	0.06	94.56
## 193	542c	18	0.06	94.62
## 194	146b	17	0.06	94.68
## 195	223d	17	0.06	94.74
## 196	380.	17	0.06	94.80
## 197	521b	17	0.06	94.86
## 198	582f	17	0.06	94.92
## 199	724.	17	0.06	94.98
## 200	312c	16	0.05	95.03
## 201	313.	16	0.05	95.08
## 202	520a	16	0.05	95.13
## 203	524.	16	0.05	95.18
## 204	582l	16	0.05	95.23
## 205	724b	16	0.05	95.28
## 206	724c	16	0.05	95.33
## 207	810.	16	0.05	95.38
## 208	812f	16	0.05	95.43
## 209	010p	15	0.05	95.48
## 210	142.	15	0.05	95.53
## 211	523b	15	0.05	95.58
## 212	6...	15	0.05	95.63
## 213	622b	15	0.05	95.68
## 214	622d	15	0.05	95.73
## 215	143.	14	0.05	95.78
## 216	212b	14	0.05	95.83
## 217	342.	14	0.05	95.88
## 218	345h	14	0.05	95.93
## 219	521f	14	0.05	95.98
## 220	522a	14	0.05	96.03
## 221	720b	14	0.05	96.08
## 222	725b	14	0.05	96.13
## 223	010k	13	0.04	96.17
## 224	14..	13	0.04	96.21
## 225	140a	13	0.04	96.25
## 226	140b	13	0.04	96.29
## 227	213g	13	0.04	96.33
## 228	31..	13	0.04	96.37
## 229	321.	13	0.04	96.41
## 230	345a	13	0.04	96.45
## 231	482a	13	0.04	96.49
## 232	624a	13	0.04	96.53
## 233	624b	13	0.04	96.57
## 234	640.	13	0.04	96.61
## 235	861a	13	0.04	96.65
## 236	213b	12	0.04	96.69
## 237	222d	12	0.04	96.73
## 238	581b	12	0.04	96.77

## 239	72..	12	0.04	96.81
## 240	762b	12	0.04	96.85
## 241	010y	11	0.04	96.89
## 242	146d	11	0.04	96.93
## 243	221c	11	0.04	96.97
## 244	225b	11	0.04	97.01
## 245	441.	11	0.04	97.05
## 246	443c	11	0.04	97.09
## 247	461.	11	0.04	97.13
## 248	521m	11	0.04	97.17
## 249	522i	11	0.04	97.21
## 250	525h	11	0.04	97.25
## 251	726e	11	0.04	97.29
## 252	851.	11	0.04	97.33
## 253	010j	10	0.03	97.36
## 254	145e	10	0.03	97.39
## 255	146.	10	0.03	97.42
## 256	345b	10	0.03	97.45
## 257	4...	10	0.03	97.48
## 258	727a	10	0.03	97.51
## 259	811c	10	0.03	97.54
## 260	813a	10	0.03	97.57
## 261	146c	9	0.03	97.60
## 262	146e	9	0.03	97.63
## 263	212a	9	0.03	97.66
## 264	214a	9	0.03	97.69
## 265	222m	9	0.03	97.72
## 266	322.	9	0.03	97.75
## 267	421b	9	0.03	97.78
## 268	521k	9	0.03	97.81
## 269	525a	9	0.03	97.84
## 270	582g	9	0.03	97.87
## 271	640a	9	0.03	97.90
## 272	76..	9	0.03	97.93
## 273	210a	8	0.03	97.96
## 274	211e	8	0.03	97.99
## 275	212d	8	0.03	98.02
## 276	226.	8	0.03	98.05
## 277	310.	8	0.03	98.08
## 278	521l	8	0.03	98.11
## 279	541c	8	0.03	98.14
## 280	599.	8	0.03	98.17
## 281	726c	8	0.03	98.20
## 282	010p	7	0.02	98.22
## 283	010ö	7	0.02	98.24
## 284	142d	7	0.02	98.26
## 285	146f	7	0.02	98.28
## 286	214c	7	0.02	98.30
## 287	214d	7	0.02	98.32
## 288	225e	7	0.02	98.34
## 289	312h	7	0.02	98.36
## 290	341a	7	0.02	98.38
## 291	522b	7	0.02	98.40
## 292	620.	7	0.02	98.42

## 293	724a	7	0.02	98.44
## 294	010o	6	0.02	98.46
## 295	010æ	6	0.02	98.48
## 296	140.	6	0.02	98.50
## 297	213a	6	0.02	98.52
## 298	213h	6	0.02	98.54
## 299	215c	6	0.02	98.56
## 300	222k	6	0.02	98.58
## 301	223a	6	0.02	98.60
## 302	223e	6	0.02	98.62
## 303	310b	6	0.02	98.64
## 304	319x	6	0.02	98.66
## 305	344.	6	0.02	98.68
## 306	345d	6	0.02	98.70
## 307	422b	6	0.02	98.72
## 308	522g	6	0.02	98.74
## 309	542a	6	0.02	98.76
## 310	582j	6	0.02	98.78
## 311	623.	6	0.02	98.80
## 312	145b	5	0.02	98.82
## 313	212f	5	0.02	98.84
## 314	215a	5	0.02	98.86
## 315	215f	5	0.02	98.88
## 316	220.	5	0.02	98.90
## 317	222ö	5	0.02	98.92
## 318	223h	5	0.02	98.94
## 319	225d	5	0.02	98.96
## 320	312f	5	0.02	98.98
## 321	421g	5	0.02	99.00
## 322	442.	5	0.02	99.02
## 323	520c	5	0.02	99.04
## 324	540.	5	0.02	99.06
## 325	542d	5	0.02	99.08
## 326	58..	5	0.02	99.10
## 327	624.	5	0.02	99.12
## 328	81..	5	0.02	99.14
## 329	840c	5	0.02	99.16
## 330	0...	4	0.01	99.17
## 331	010e	4	0.01	99.18
## 332	010n	4	0.01	99.19
## 333	222a	4	0.01	99.20
## 334	222j	4	0.01	99.21
## 335	312e	4	0.01	99.22
## 336	312g	4	0.01	99.23
## 337	312i	4	0.01	99.24
## 338	319.	4	0.01	99.25
## 339	341d	4	0.01	99.26
## 340	343b	4	0.01	99.27
## 341	345g	4	0.01	99.28
## 342	346a	4	0.01	99.29
## 343	421f	4	0.01	99.30
## 344	44..	4	0.01	99.31
## 345	462.	4	0.01	99.32
## 346	525e	4	0.01	99.33

## 347	529.	4	0.01	99.34
## 348	543b	4	0.01	99.35
## 349	727c	4	0.01	99.36
## 350	812b	4	0.01	99.37
## 351	010d	3	0.01	99.38
## 352	010v	3	0.01	99.39
## 353	142b	3	0.01	99.40
## 354	145d	3	0.01	99.41
## 355	145f	3	0.01	99.42
## 356	146g	3	0.01	99.43
## 357	211c	3	0.01	99.44
## 358	212h	3	0.01	99.45
## 359	213i	3	0.01	99.46
## 360	222l	3	0.01	99.47
## 361	222o	3	0.01	99.48
## 362	226b	3	0.01	99.49
## 363	3...	3	0.01	99.50
## 364	310c	3	0.01	99.51
## 365	312d	3	0.01	99.52
## 366	345k	3	0.01	99.53
## 367	345m	3	0.01	99.54
## 368	422.	3	0.01	99.55
## 369	521h	3	0.01	99.56
## 370	524a	3	0.01	99.57
## 371	544.	3	0.01	99.58
## 372	726f	3	0.01	99.59
## 373	762d	3	0.01	99.60
## 374	762f	3	0.01	99.61
## 375	860.	3	0.01	99.62
## 376	863.	3	0.01	99.63
## 377	010w	2	0.01	99.64
## 378	145c	2	0.01	99.65
## 379	211b	2	0.01	99.66
## 380	212c	2	0.01	99.67
## 381	212g	2	0.01	99.68
## 382	215e	2	0.01	99.69
## 383	221a	2	0.01	99.70
## 384	32..	2	0.01	99.71
## 385	322a	2	0.01	99.72
## 386	345l	2	0.01	99.73
## 387	349x	2	0.01	99.74
## 388	421d	2	0.01	99.75
## 389	421e	2	0.01	99.76
## 390	422d	2	0.01	99.77
## 391	443e	2	0.01	99.78
## 392	481g	2	0.01	99.79
## 393	5...	2	0.01	99.80
## 394	522f	2	0.01	99.81
## 395	522h	2	0.01	99.82
## 396	522l	2	0.01	99.83
## 397	599x	2	0.01	99.84
## 398	622c	2	0.01	99.85
## 399	7...	2	0.01	99.86
## 400	840d	2	0.01	99.87

## 401	850a	2	0.01	99.88
## 402	010i	1	0.00	99.88
## 403	010q	1	0.00	99.88
## 404	010u	1	0.00	99.88
## 405	149x	1	0.00	99.88
## 406	210b	1	0.00	99.88
## 407	213j	1	0.00	99.88
## 408	214b	1	0.00	99.88
## 409	214e	1	0.00	99.88
## 410	215g	1	0.00	99.88
## 411	221d	1	0.00	99.88
## 412	222á	1	0.00	99.88
## 413	222p	1	0.00	99.88
## 414	223b	1	0.00	99.88
## 415	223f	1	0.00	99.88
## 416	223g	1	0.00	99.88
## 417	310a	1	0.00	99.88
## 418	312j	1	0.00	99.88
## 419	314b	1	0.00	99.88
## 420	340c	1	0.00	99.88
## 421	341c	1	0.00	99.88
## 422	347.	1	0.00	99.88
## 423	421i	1	0.00	99.88
## 424	440.	1	0.00	99.88
## 425	481e	1	0.00	99.88
## 426	481h	1	0.00	99.88
## 427	482.	1	0.00	99.88
## 428	499x	1	0.00	99.88
## 429	521o	1	0.00	99.88
## 430	541d	1	0.00	99.88
## 431	542e	1	0.00	99.88
## 432	581d	1	0.00	99.88
## 433	581e	1	0.00	99.88
## 434	623a	1	0.00	99.88
## 435	720c	1	0.00	99.88
## 436	8...	1	0.00	99.88
## 437	812g	1	0.00	99.88
## 438	814c	1	0.00	99.88
## 439	815c	1	0.00	99.88
## 440	840e	1	0.00	99.88
## 441	850.	1	0.00	99.88
## 442	853.	1	0.00	100.00



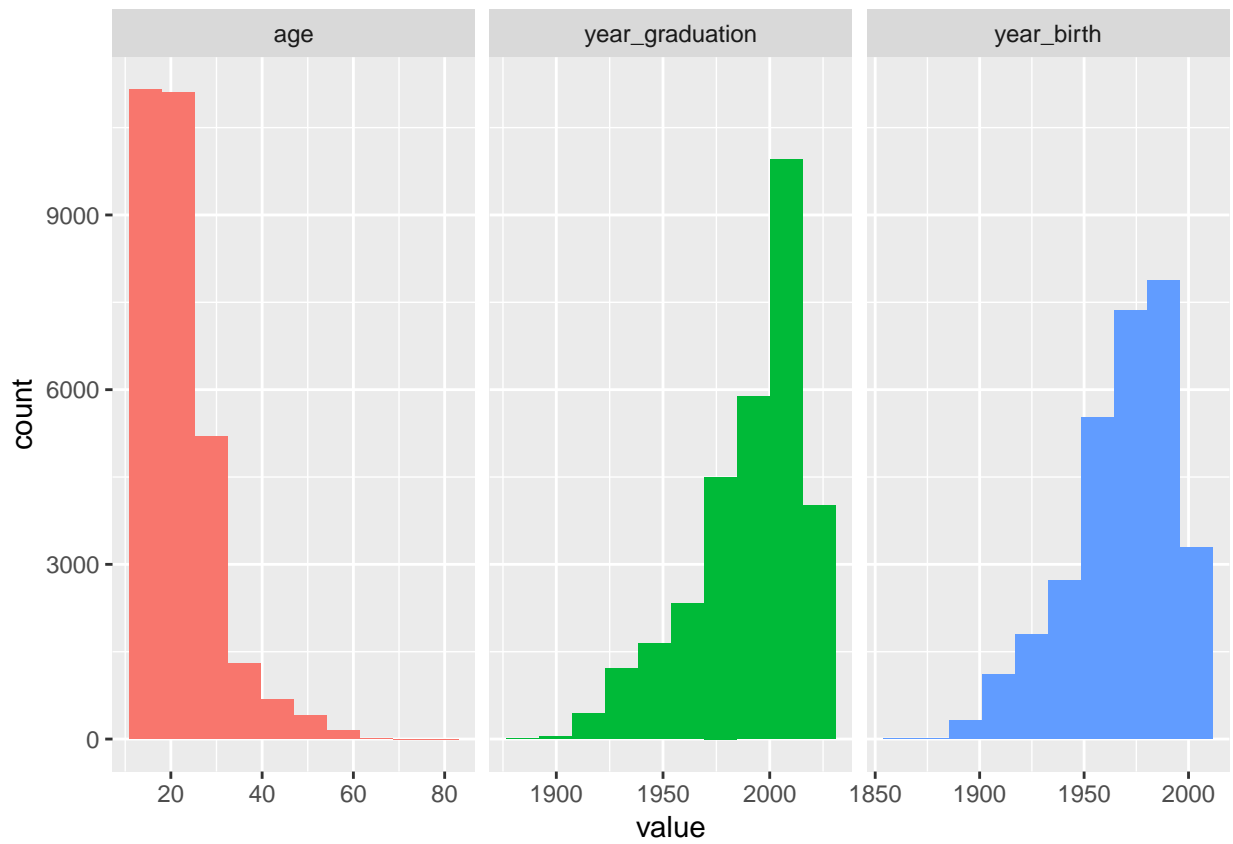
Frequency / (Percentage %)

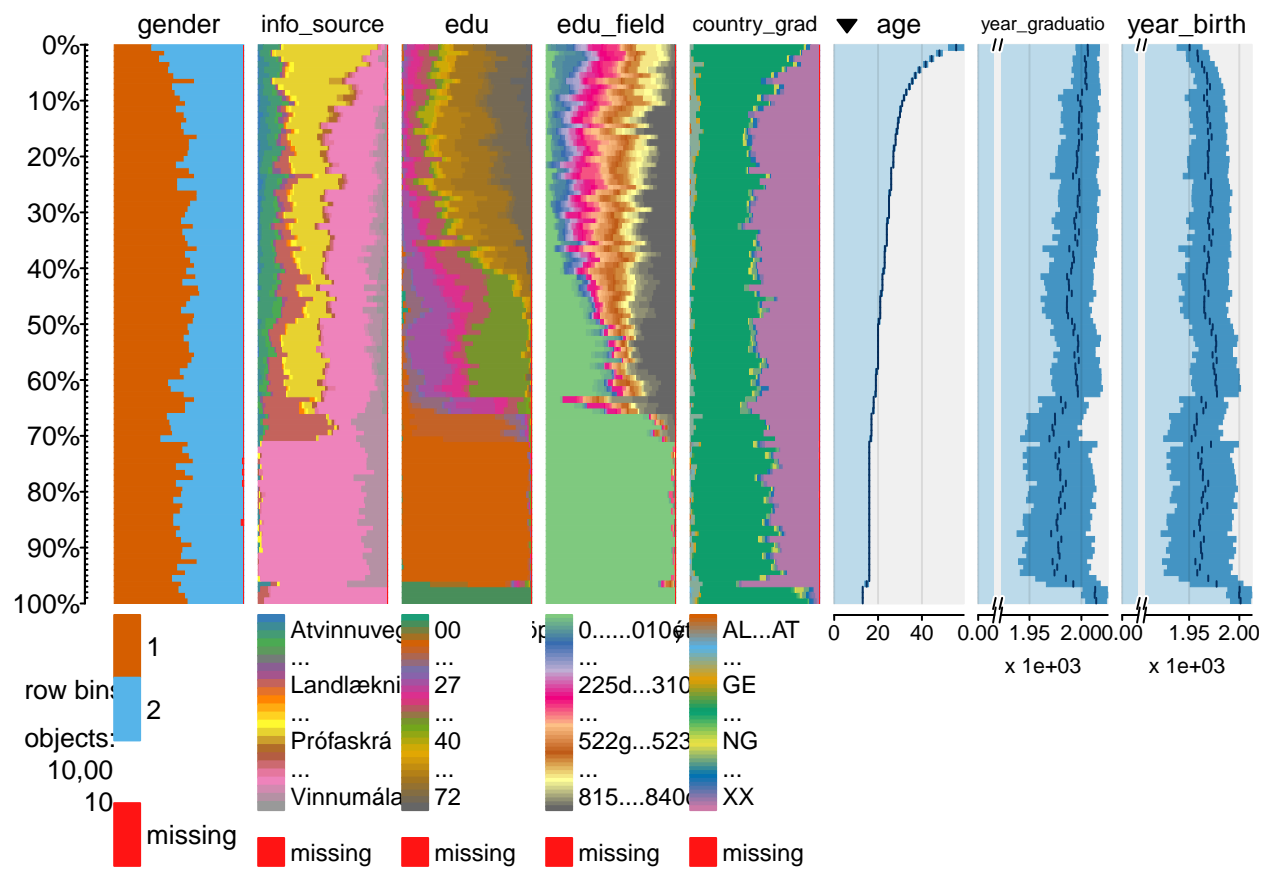
##	country_grad	frequency	percentage	cumulative_perc
## 1	IS	15069	50.23	50.23
## 2	XX	11873	39.58	89.81
## 3	DK	784	2.61	92.42
## 4	PL	471	1.57	93.99
## 5	SE	374	1.25	95.24
## 6	NO	279	0.93	96.17
## 7	US	239	0.80	96.97
## 8	DE	117	0.39	97.36
## 9	GB	93	0.31	97.67
## 10	FI	48	0.16	97.83
## 11	LT	45	0.15	97.98
## 12	PH	39	0.13	98.11
## 13	TH	37	0.12	98.23
## 14	CA	31	0.10	98.33
## 15	PT	29	0.10	98.43
## 16	SK	29	0.10	98.53
## 17	VN	28	0.09	98.62
## 18	CN	26	0.09	98.71
## 19	IT	26	0.09	98.80
## 20	FR	25	0.08	98.88
## 21	LV	20	0.07	98.95
## 22	ES	17	0.06	99.01
## 23	AT	14	0.05	99.06
## 24	CZ	13	0.04	99.10
## 25	NL	13	0.04	99.14

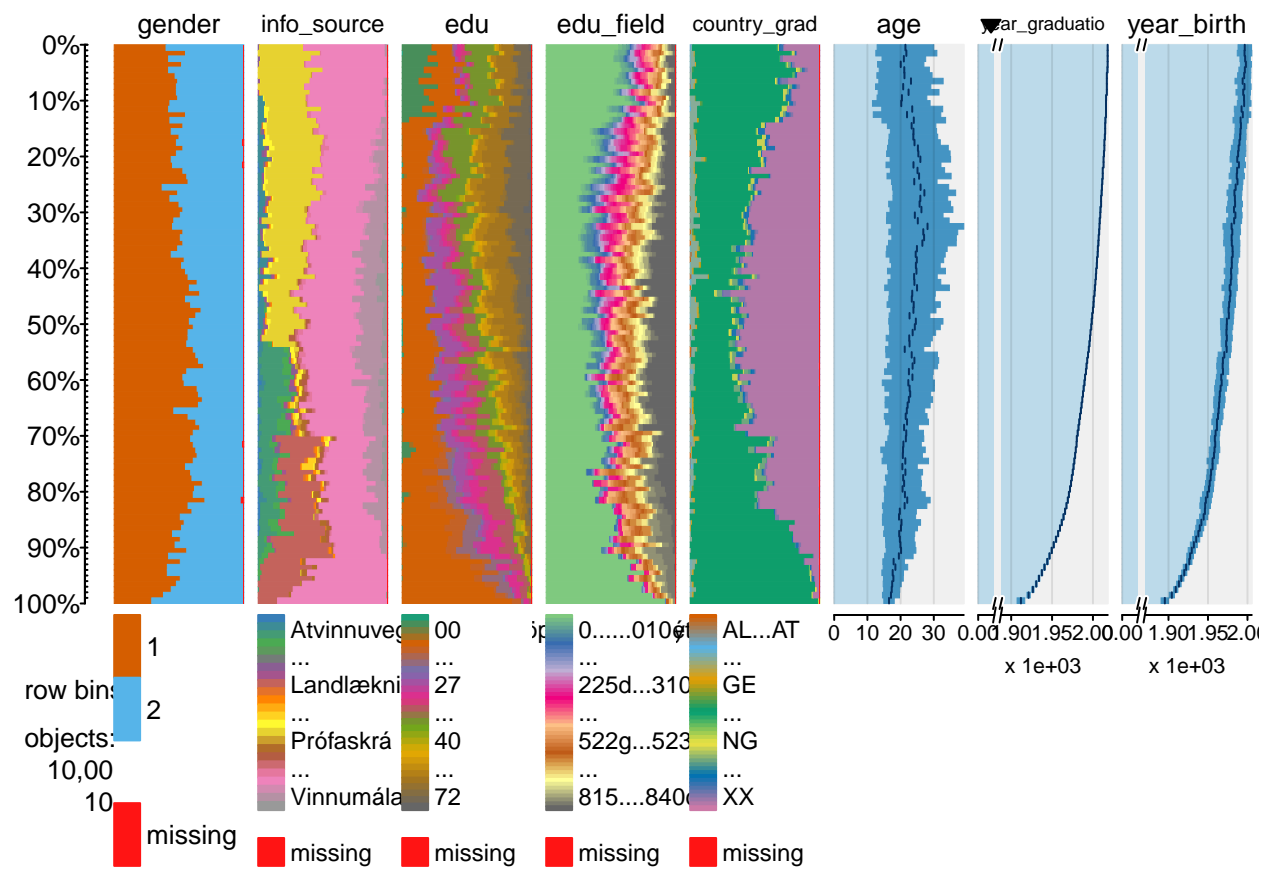
## 26	RS	13	0.04	99.18
## 27	HU	12	0.04	99.22
## 28	NZ	12	0.04	99.26
## 29	IN	11	0.04	99.30
## 30	RO	11	0.04	99.34
## 31	CH	10	0.03	99.37
## 32	HR	10	0.03	99.40
## 33	BA	9	0.03	99.43
## 34	CO	9	0.03	99.46
## 35	LU	8	0.03	99.49
## 36	RU	8	0.03	99.52
## 37	AU	7	0.02	99.54
## 38	FO	7	0.02	99.56
## 39	NP	6	0.02	99.58
## 40	BE	5	0.02	99.60
## 41	BG	5	0.02	99.62
## 42	IE	5	0.02	99.64
## 43	SI	5	0.02	99.66
## 44	SY	5	0.02	99.68
## 45	AL	4	0.01	99.69
## 46	EE	4	0.01	99.70
## 47	ID	4	0.01	99.71
## 48	PK	4	0.01	99.72
## 49	UA	4	0.01	99.73
## 50	BR	3	0.01	99.74
## 51	BY	3	0.01	99.75
## 52	CL	3	0.01	99.76
## 53	KE	3	0.01	99.77
## 54	MK	3	0.01	99.78
## 55	MX	3	0.01	99.79
## 56	NG	3	0.01	99.80
## 57	QR	3	0.01	99.81
## 58	YU	3	0.01	99.82
## 59	YY	3	0.01	99.83
## 60	AR	2	0.01	99.84
## 61	CR	2	0.01	99.85
## 62	DZ	2	0.01	99.86
## 63	ET	2	0.01	99.87
## 64	GL	2	0.01	99.88
## 65	GR	2	0.01	99.89
## 66	HN	2	0.01	99.90
## 67	IL	2	0.01	99.91
## 68	IR	2	0.01	99.92
## 69	JP	2	0.01	99.93
## 70	LK	2	0.01	99.94
## 71	MA	2	0.01	99.95
## 72	MY	2	0.01	99.96
## 73	SL	2	0.01	99.97
## 74	SU	2	0.01	99.98
## 75	ZA	2	0.01	99.99
## 76	AZ	1	0.00	99.99
## 77	BO	1	0.00	99.99
## 78	CY	1	0.00	99.99
## 79	DD	1	0.00	99.99

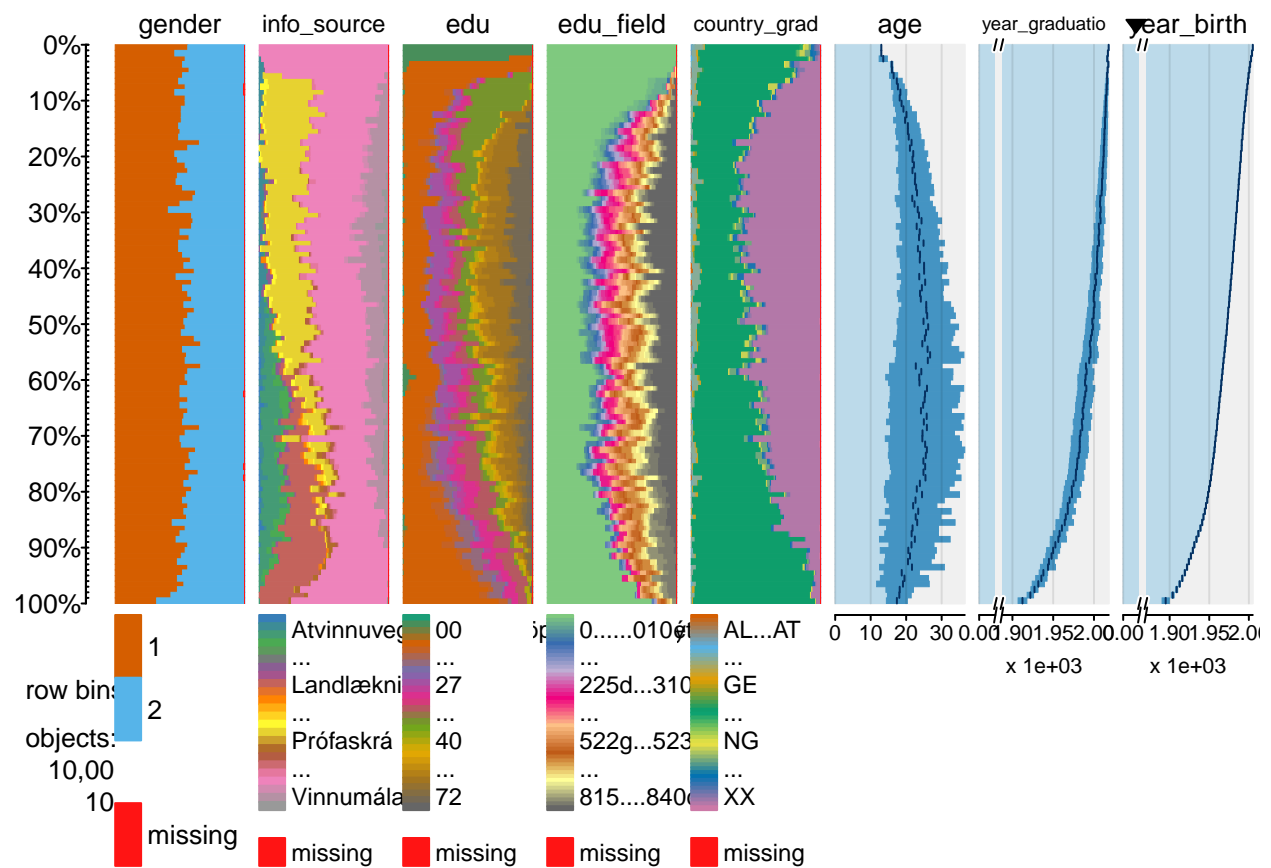
```
## 80      EC      1      0.00      99.99
## 81      GE      1      0.00      99.99
## 82      IQ      1      0.00      99.99
## 83      JM      1      0.00      99.99
## 84      JO      1      0.00      99.99
## 85      KZ      1      0.00      99.99
## 86      LB      1      0.00      99.99
## 87      MN      1      0.00      99.99
## 88      PE      1      0.00      99.99
## 89      SG      1      0.00      99.99
## 90      SN      1      0.00      99.99
## 91      SO      1      0.00      99.99
## 92      TR      1      0.00      99.99
## 93      TZ      1      0.00      99.99
## 94      UG      1      0.00      99.99
## 95      VU      1      0.00      99.99
## 96      <NA>    1      0.00     100.00
```

```
## [1] "Variables processed: gender, info_source, edu, edu_field, country_grad"
```







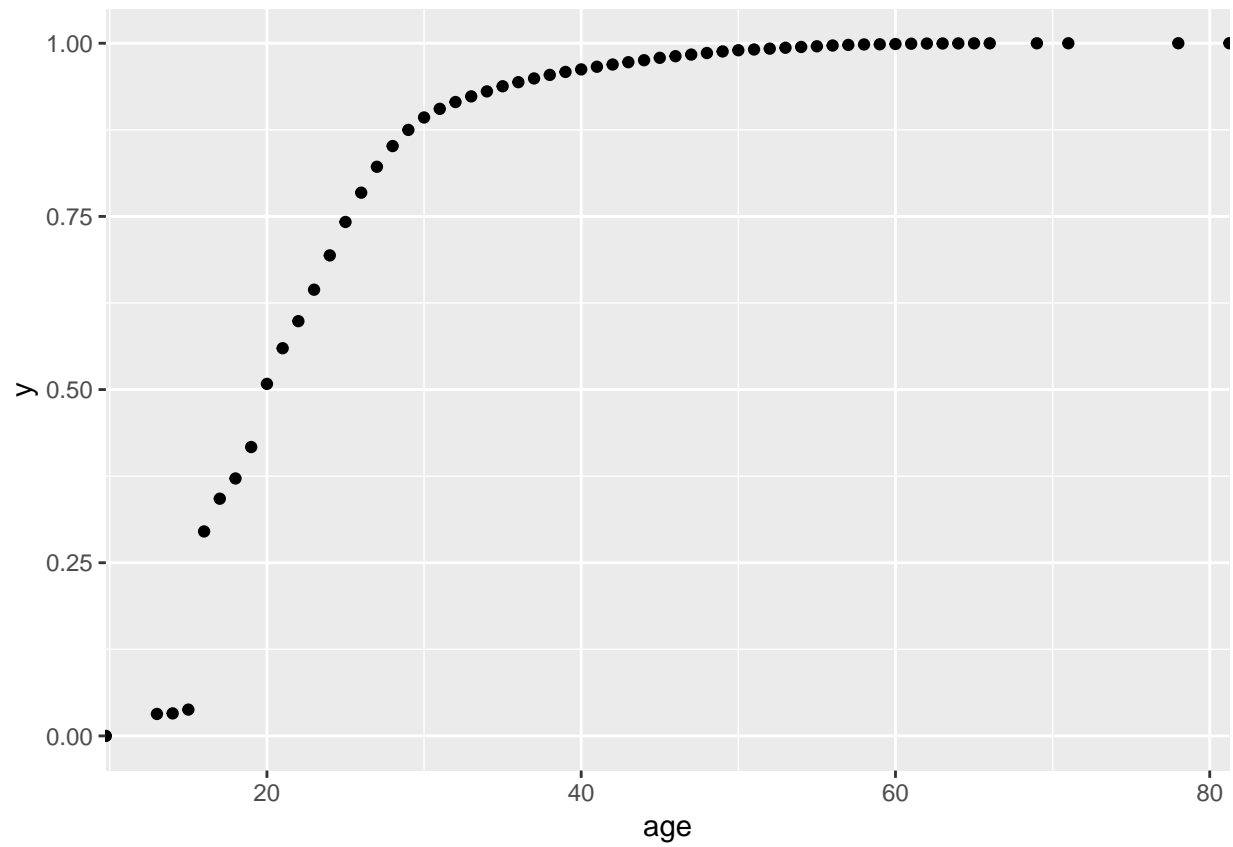


```
## [[1]]
## NULL
##
## [[2]]
## NULL
##
## [[3]]
## NULL
```

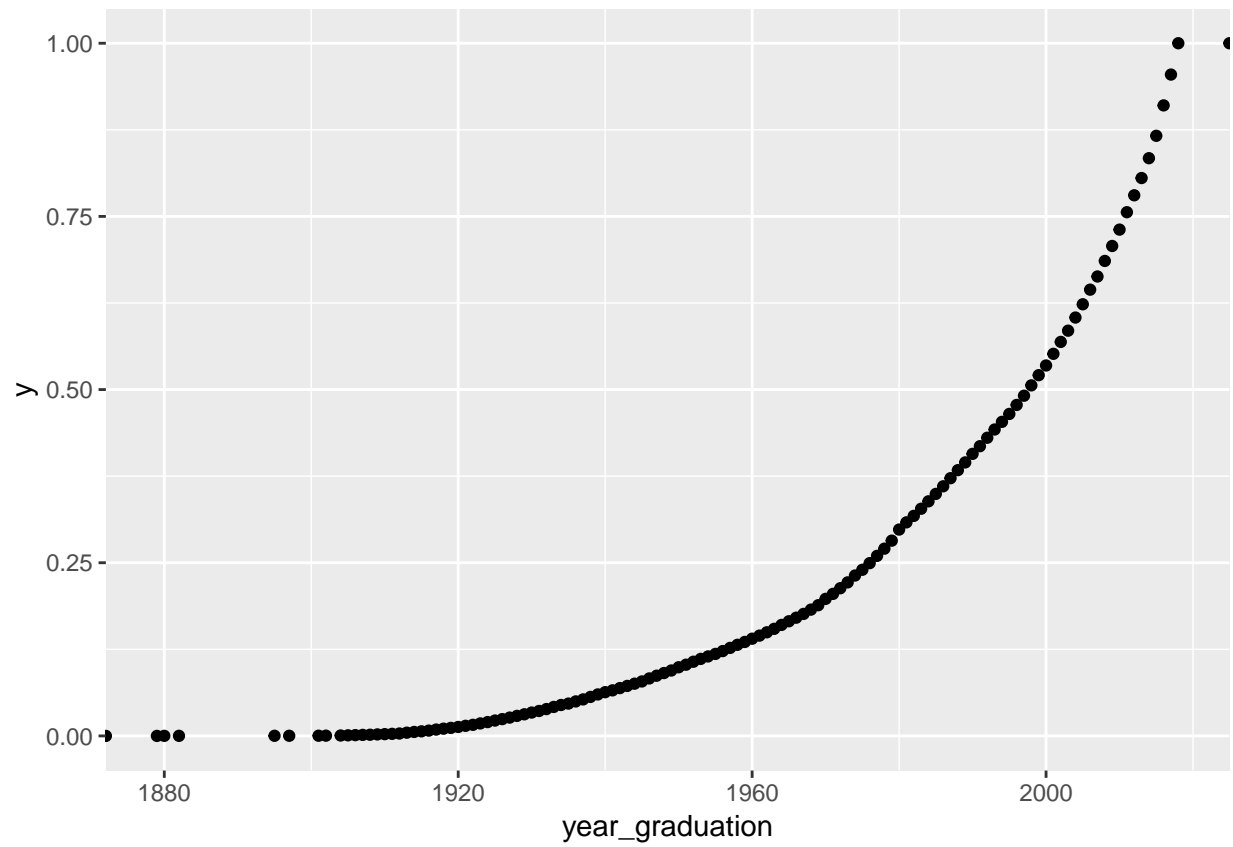
view__univar

Marginal cumulative distributions

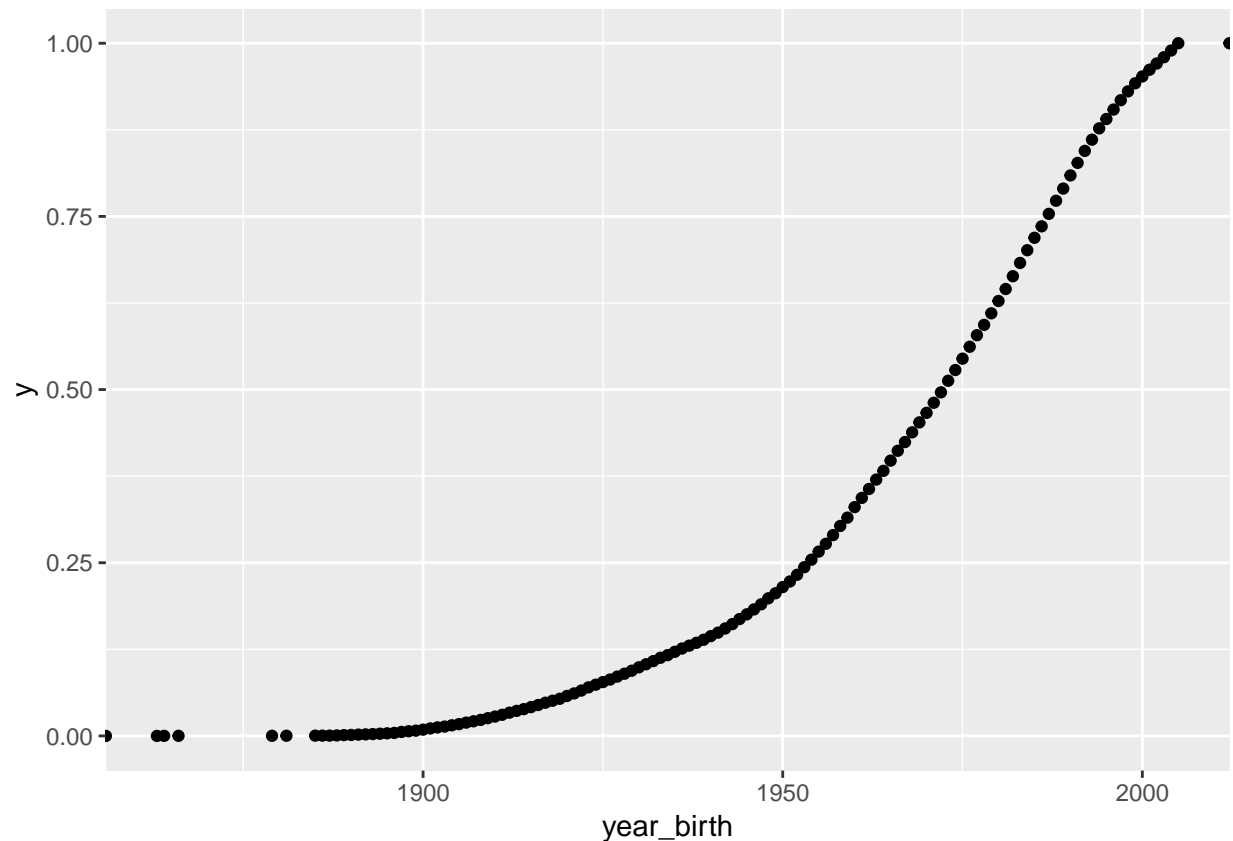
```
## [[1]]
```



```
##  
## [[2]]
```



```
##  
## [[3]]
```

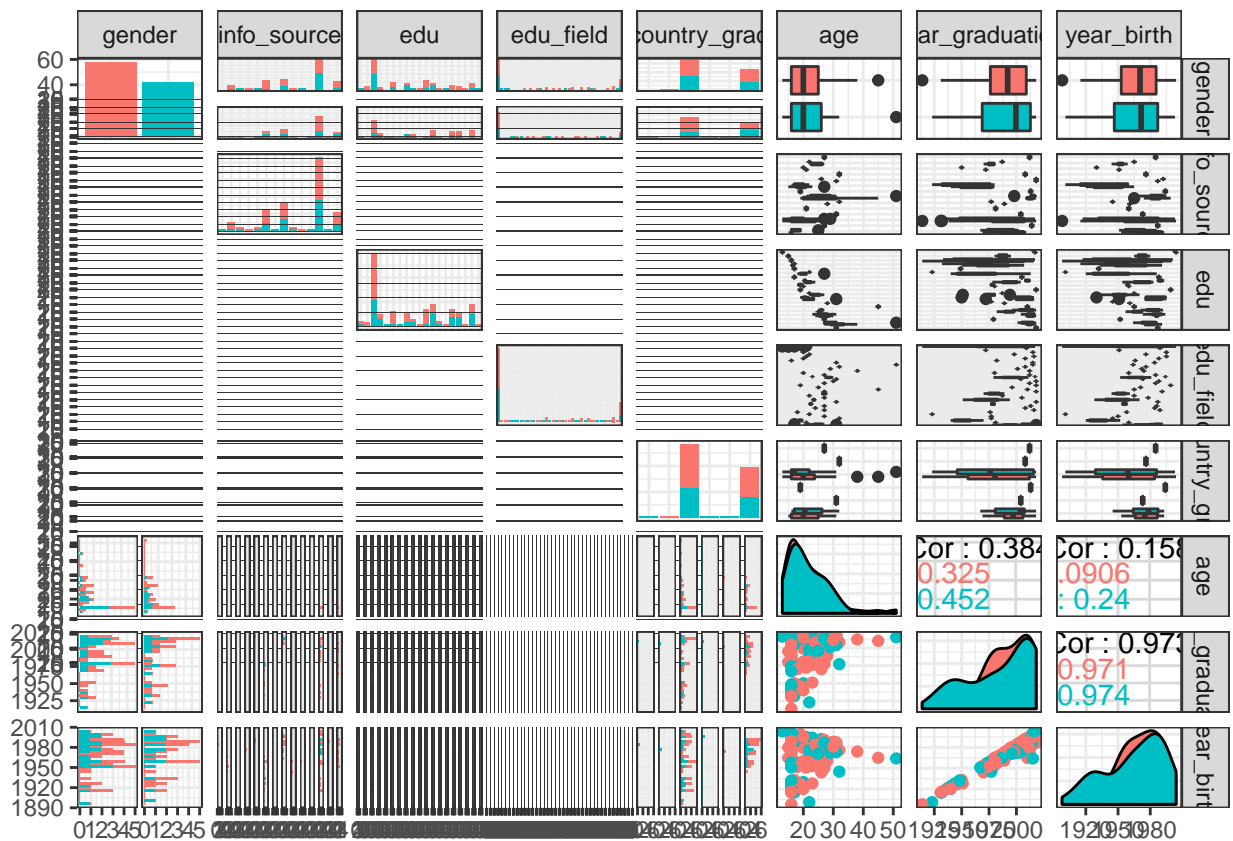



view_multivar

Pairwise bivariate distributions and correlation plots

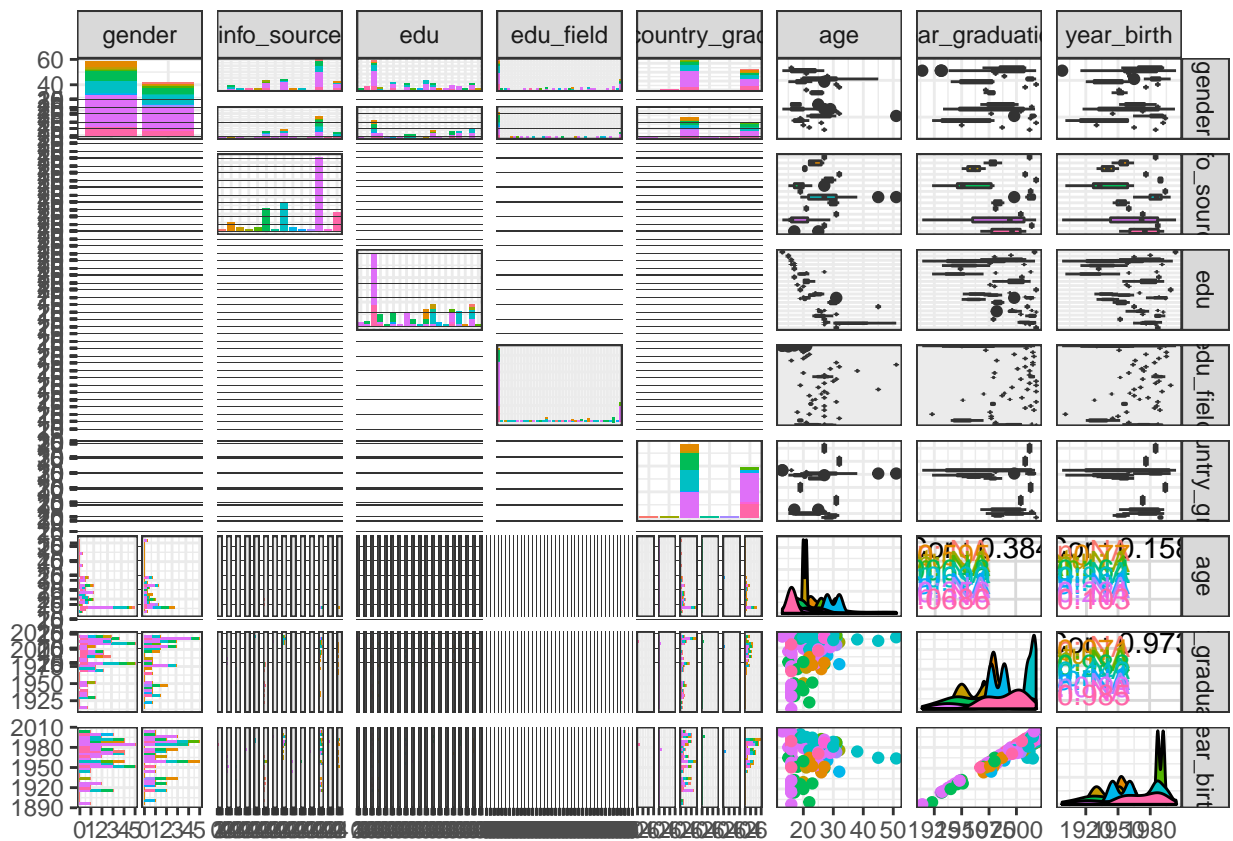
Note that: printing is not yet “addapted” to size of data and paper

```
## [[1]]
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
```



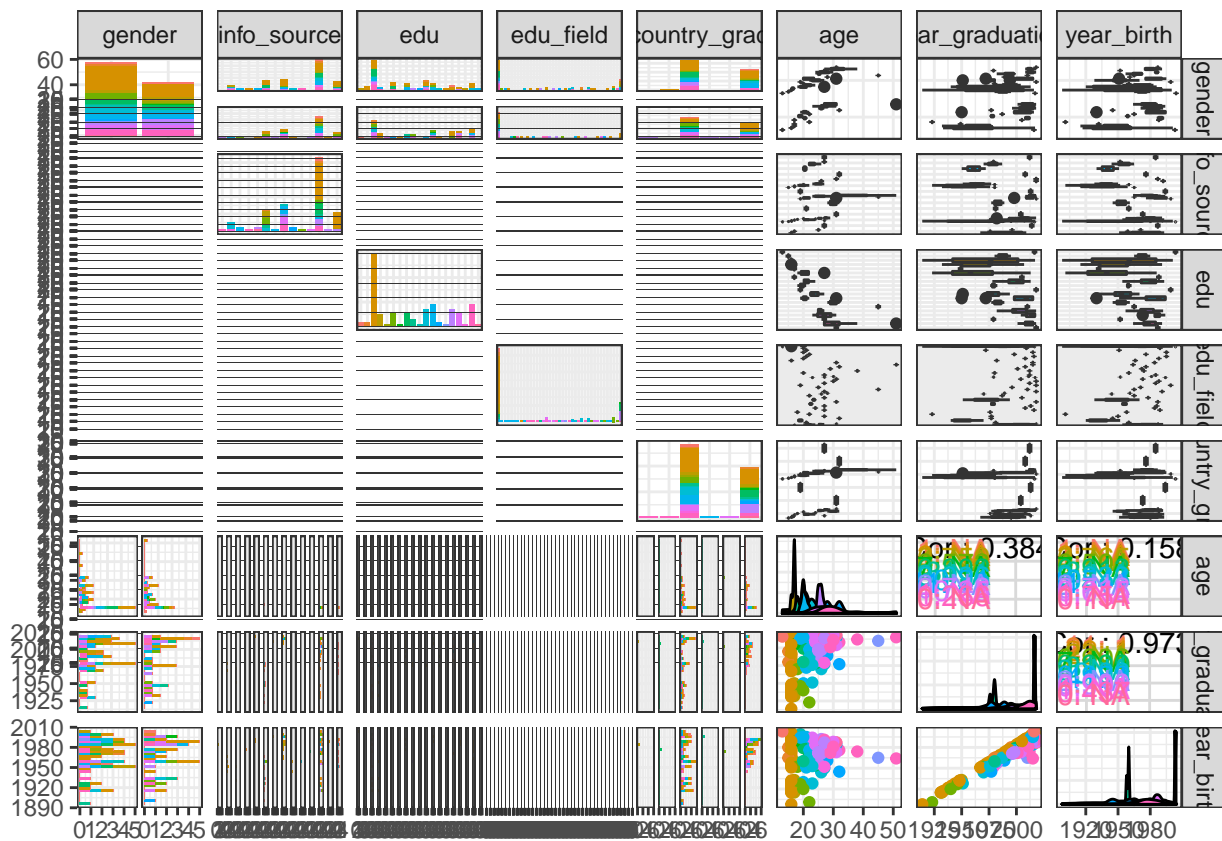
```
##
## [[2]]

## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
```



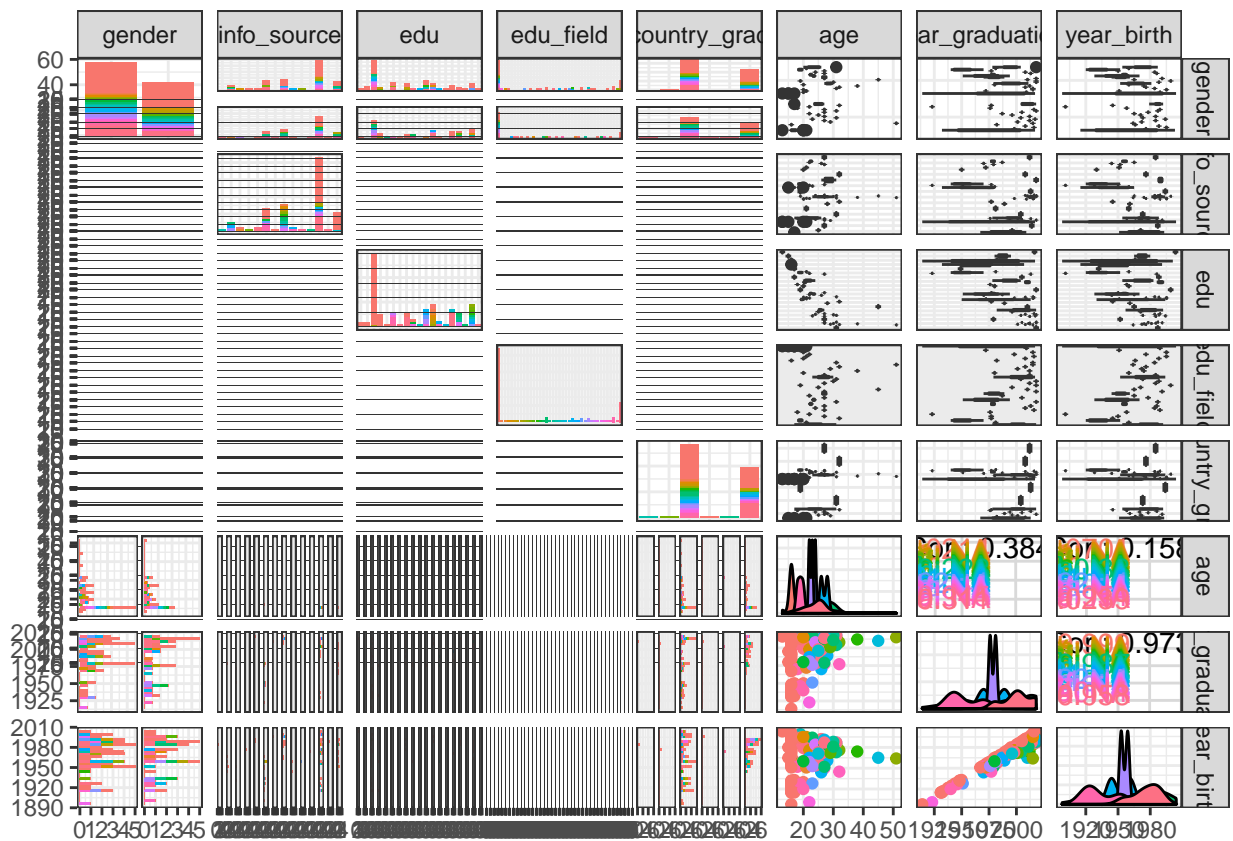
```
##
## [[3]]

## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
```



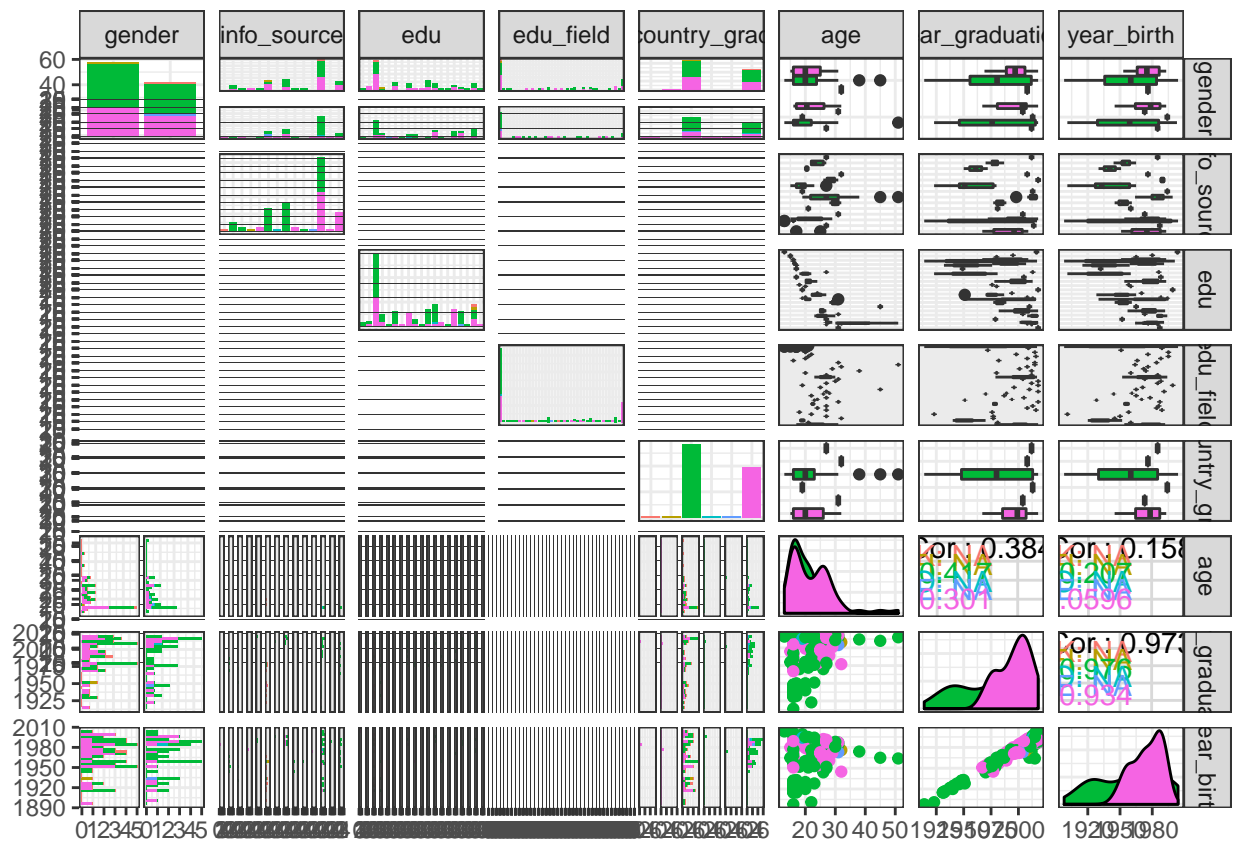
```
##
## [[4]]

## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
```



```
##
## [[5]]

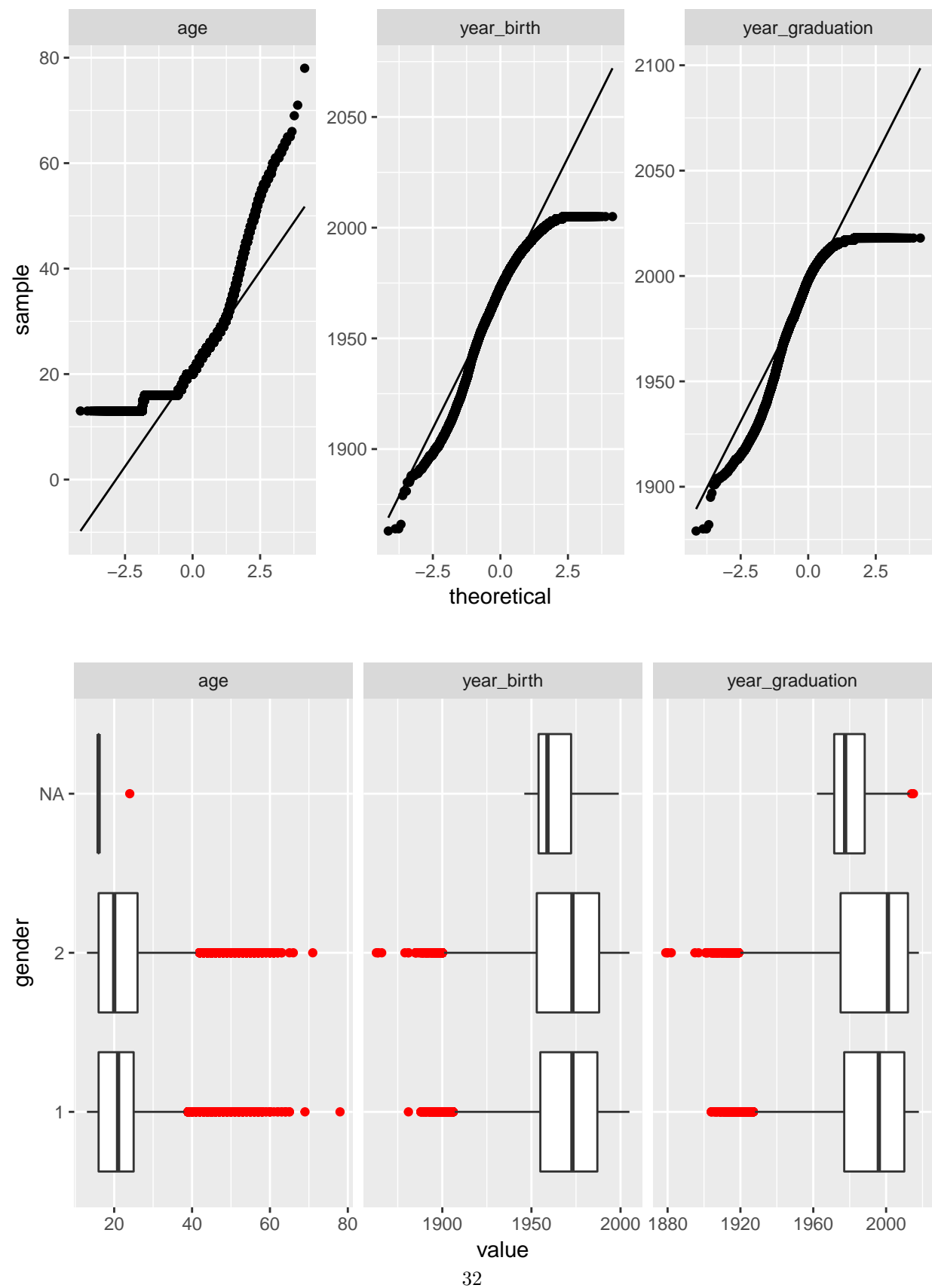
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
```

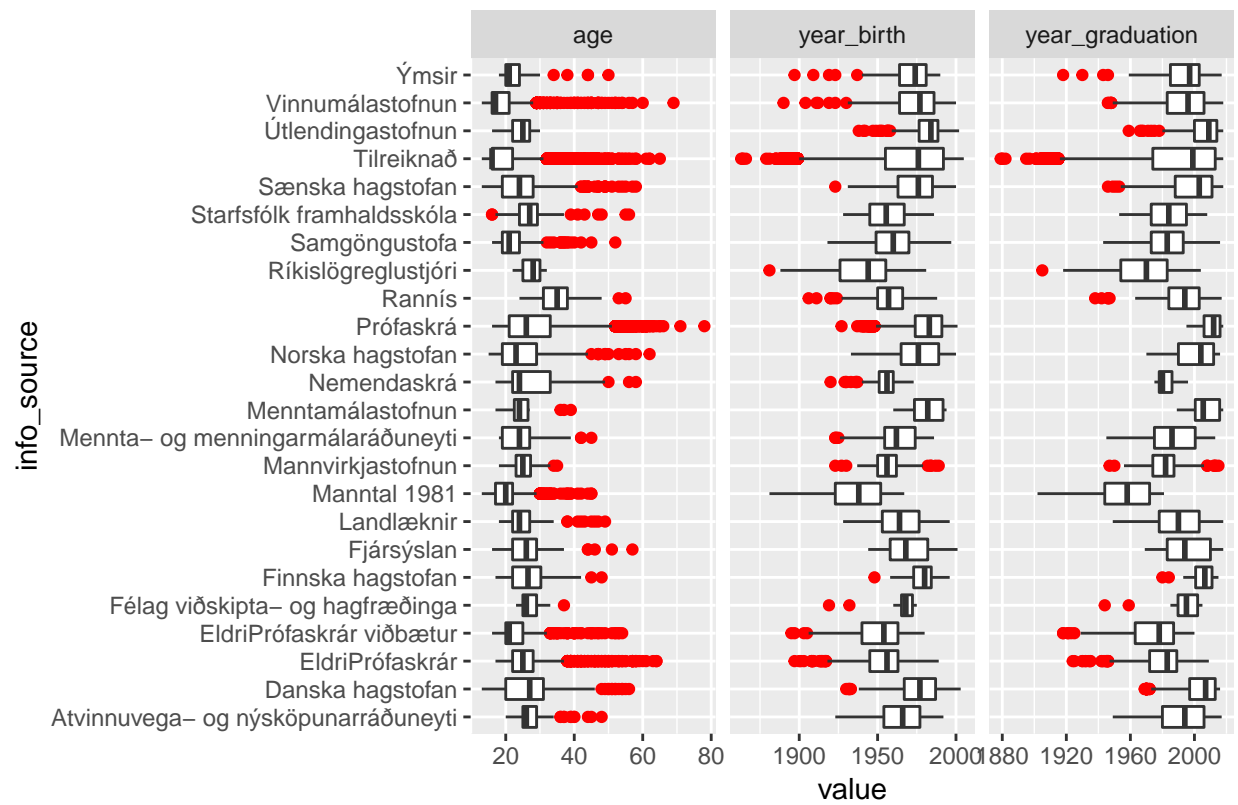


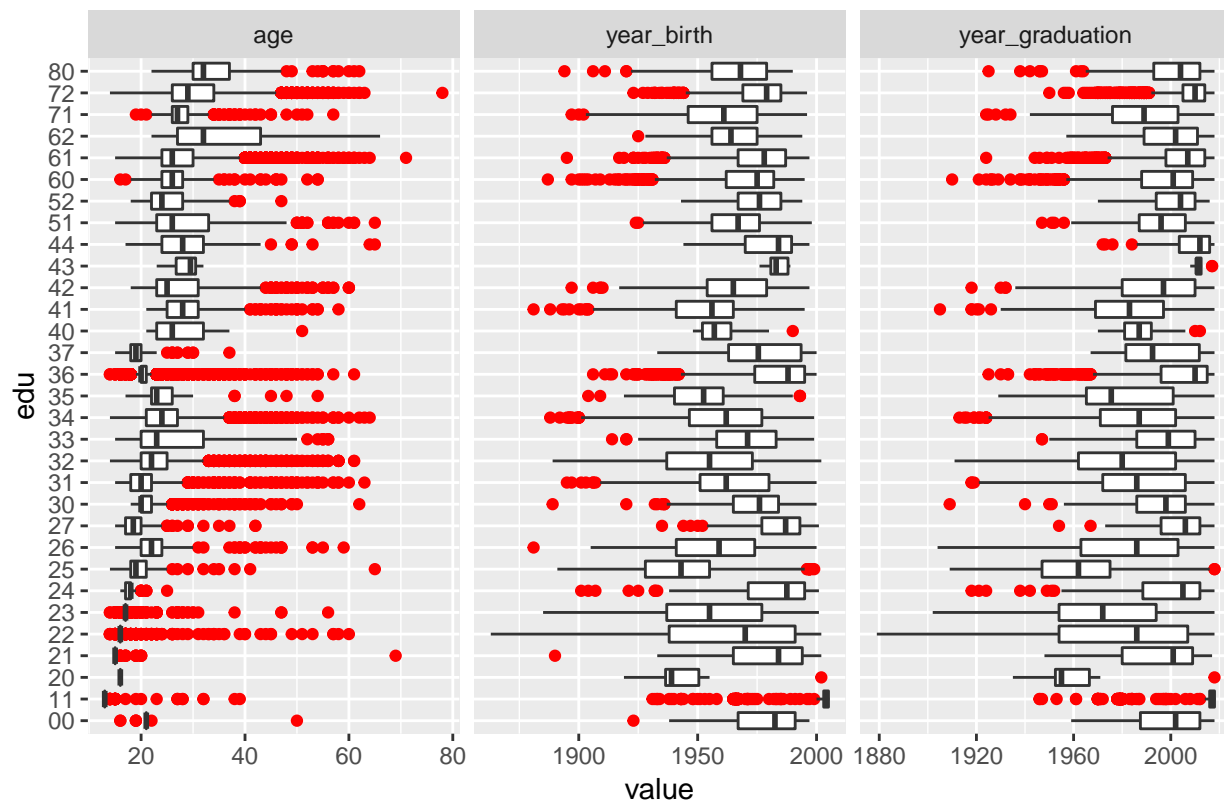
```
## 2 features with more than 50 categories ignored!
## edu_field: 442 categories
## country_grad: 96 categories
```

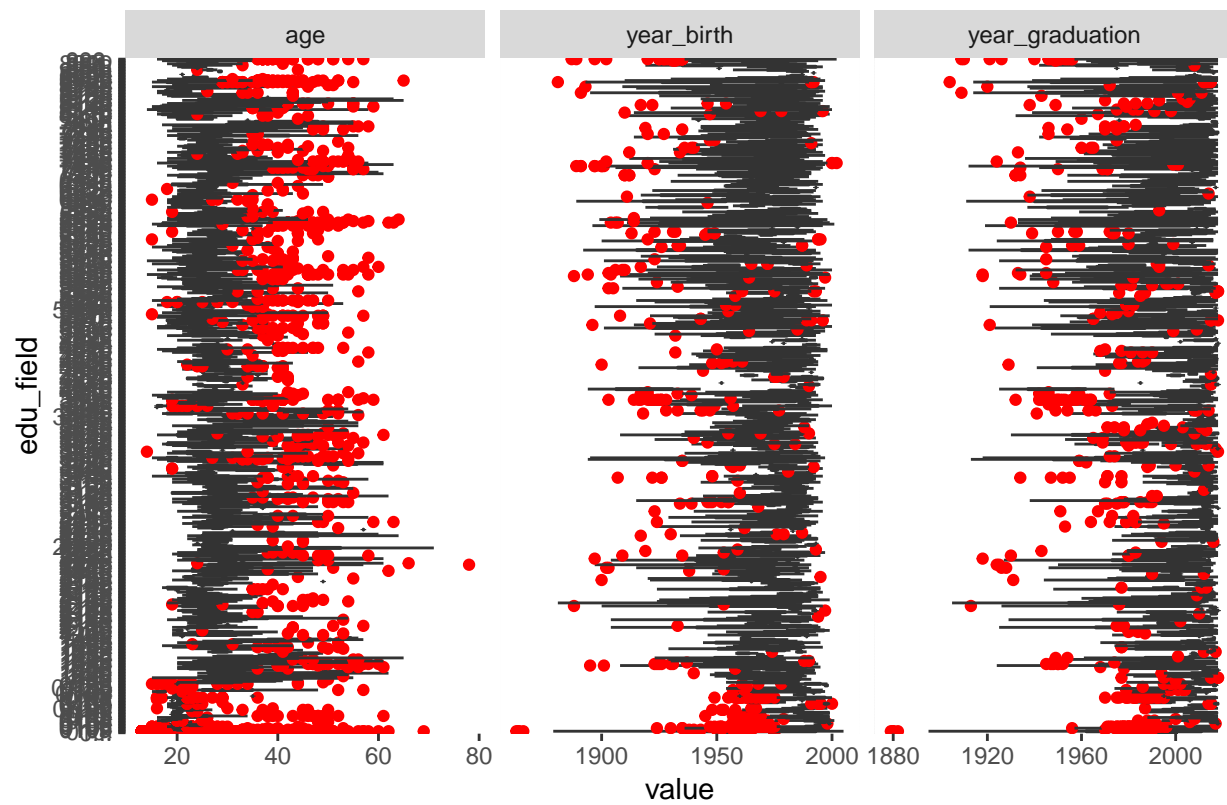

view_outliers

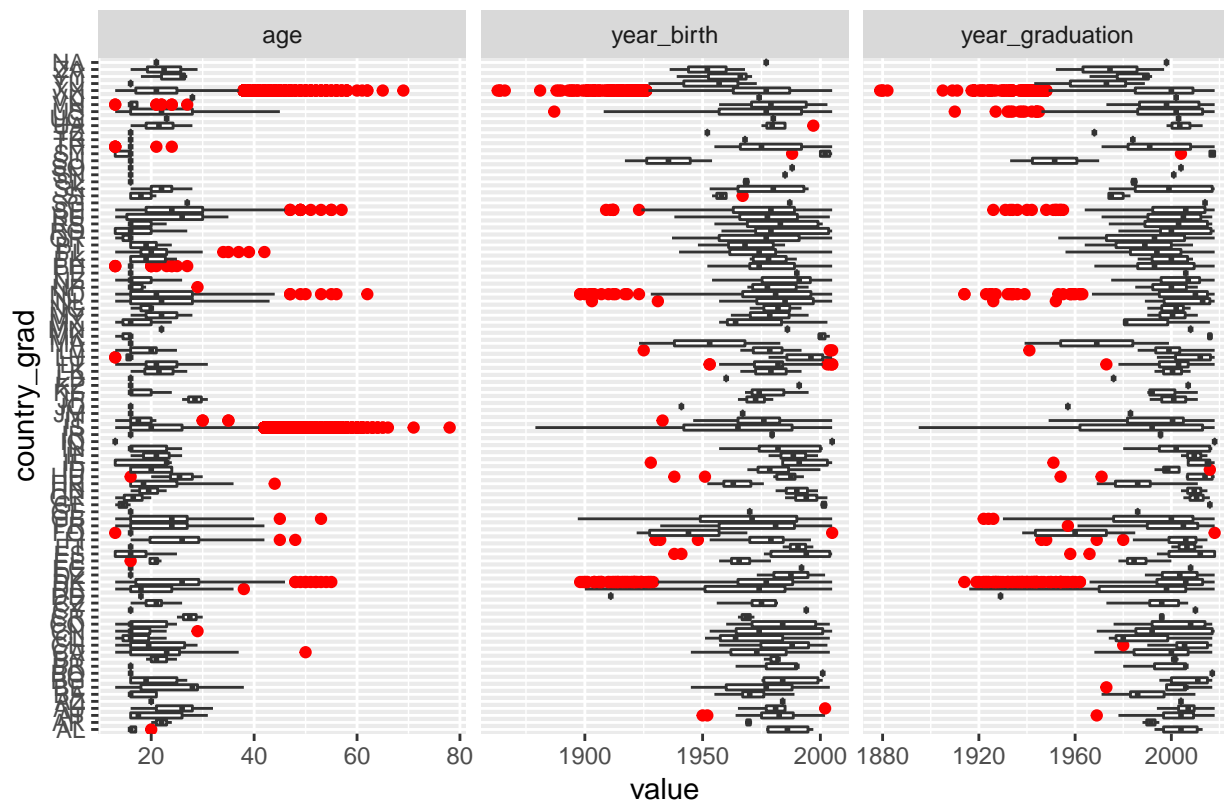
Plots and boxplots,limits based on Tukey and Hampel methods



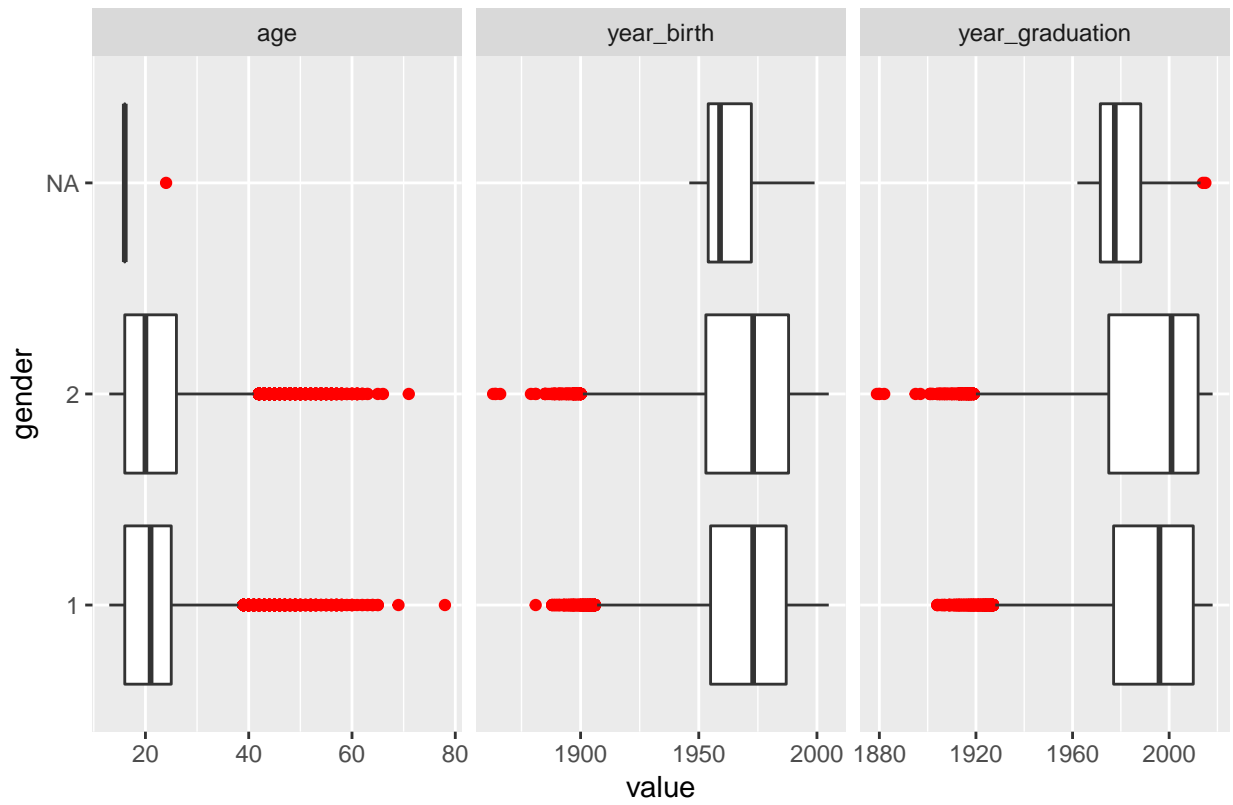




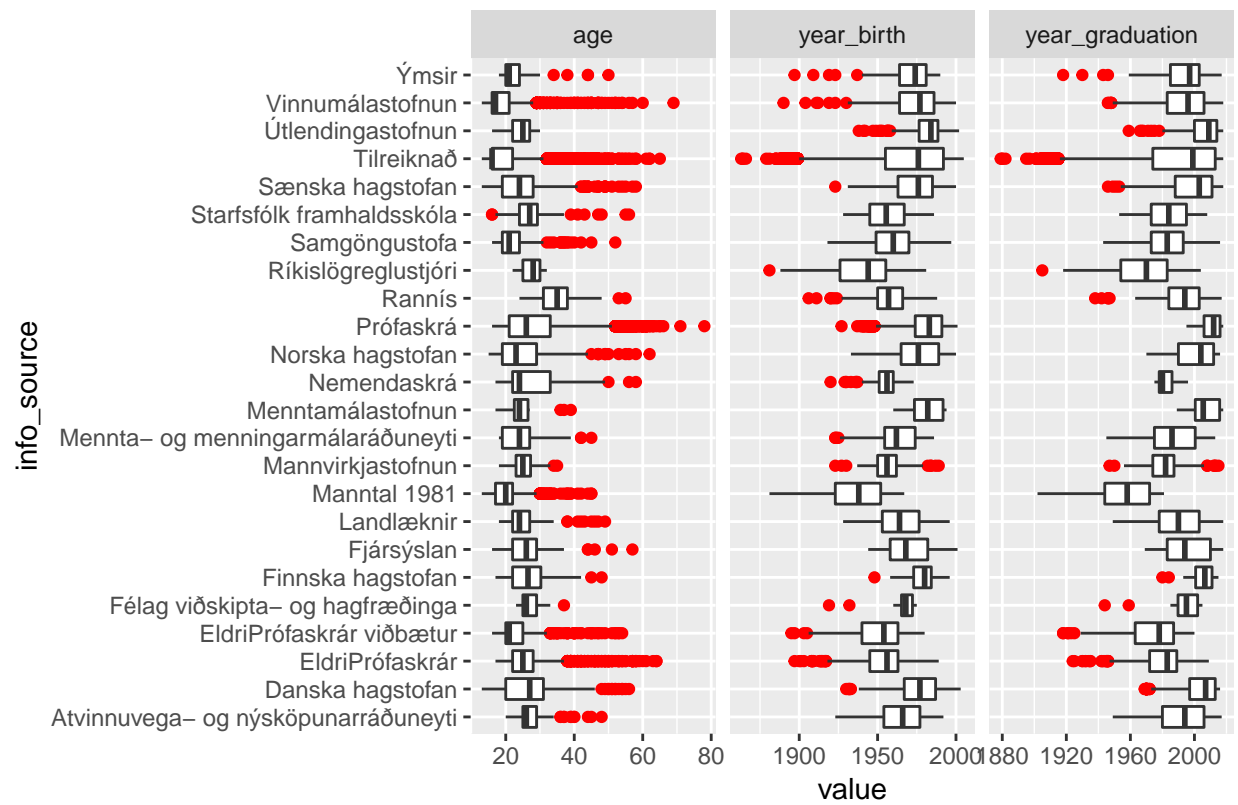




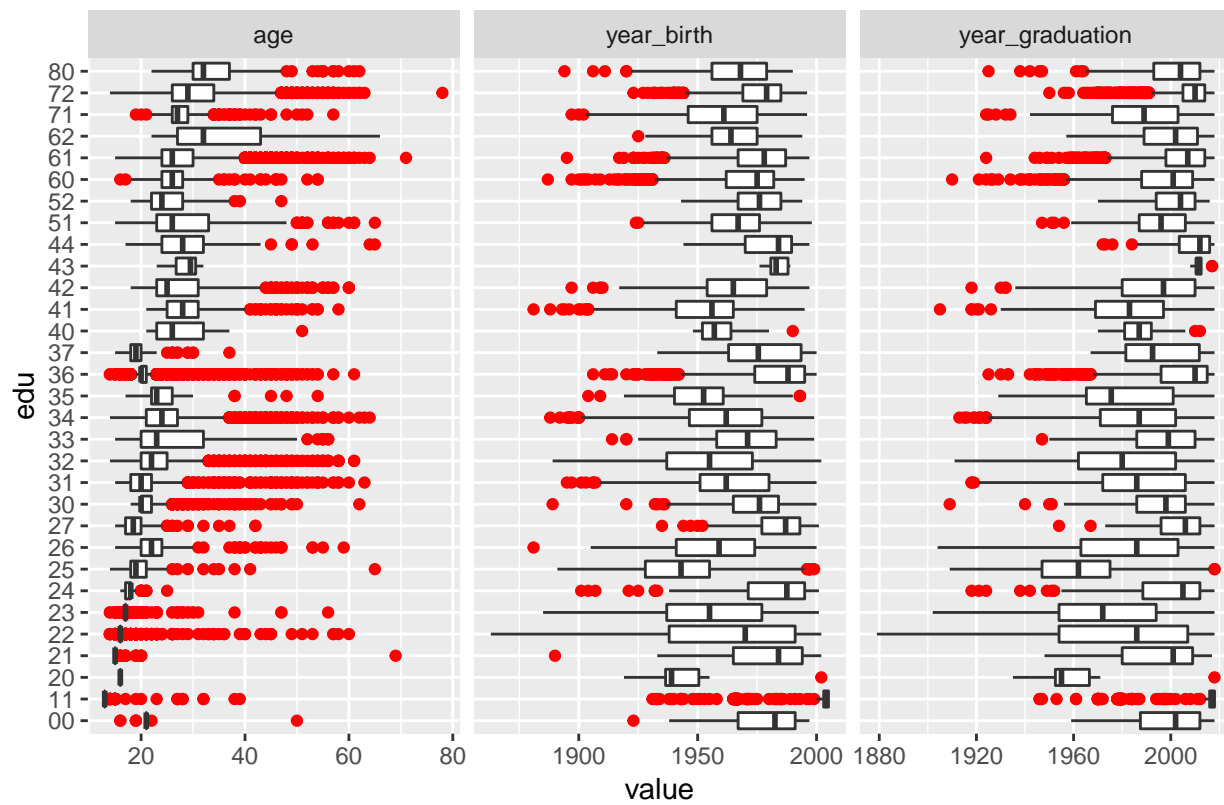
```
## [[1]]
## [[1]]$page_1
```



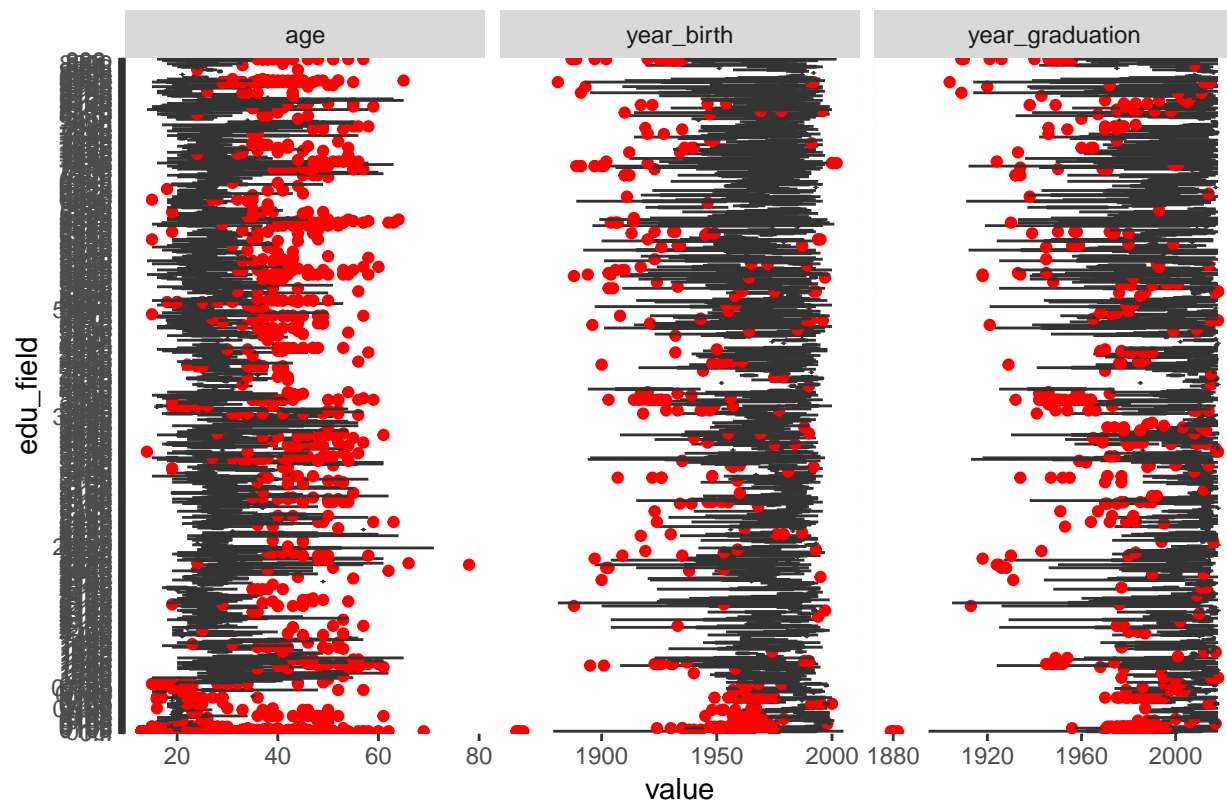
```
##
##
## [[2]]
## [[2]]$page_1
```



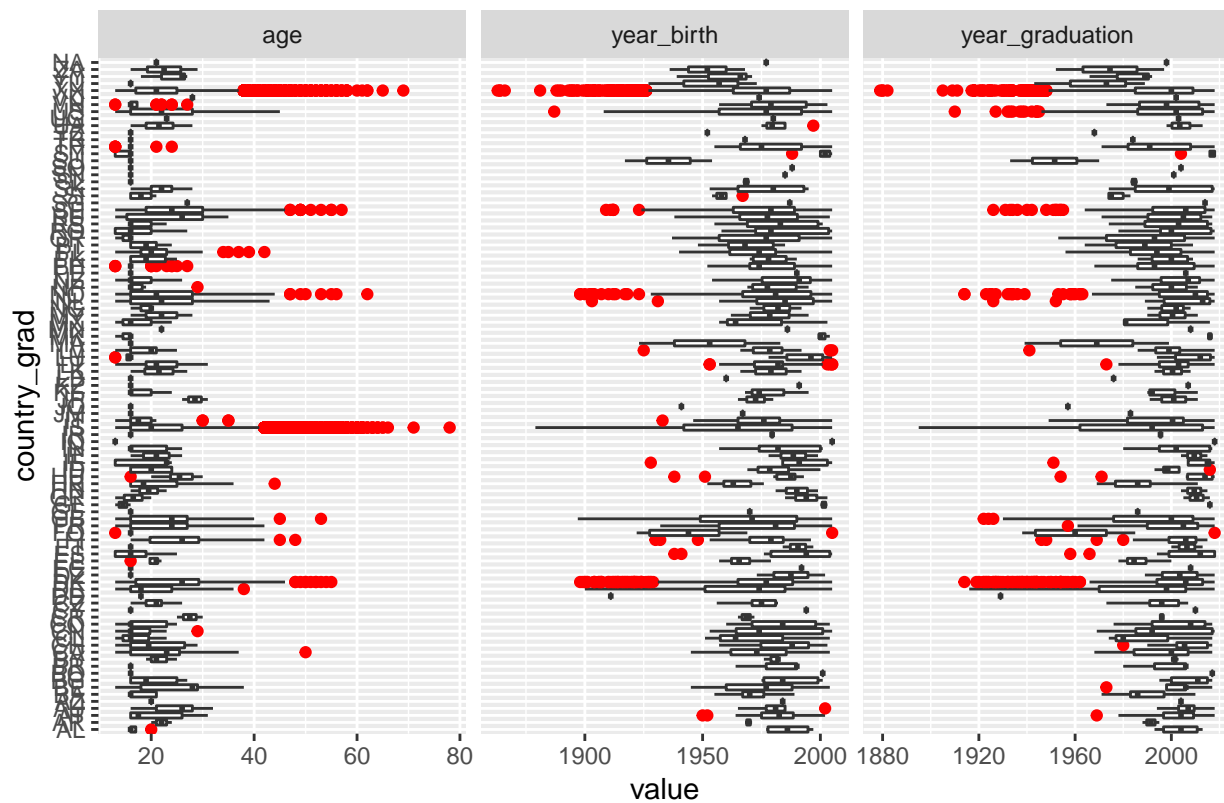
```
##
##
## [[3]]
## [[3]]$page_1
```



```
##
##
## [[4]]
## [[4]]$page_1
```



```
##
##
## [[5]]
## [[5]]$page_1
```

	age
bottom_threshold	-14
top_threshold	56

	year_graduation
bottom_threshold	1875
top_threshold	2113

	year_birth
bottom_threshold	1855
top_threshold	2086

	age
bottom_threshold	2.2088
top_threshold	37.7912

	year_graduation
bottom_threshold	1912.2157
top_threshold	2076.7843

	year_birth
bottom_threshold	1895.3874
top_threshold	2046.6126

view__assoc

With validation rules mining potential.Under development.

view__clusters

Potentially identifying unwanted structures or confirming known ones. Under development.

rev__variability

information theory based measures, for categorical variables

```
##      d_res
##      "var1" "gender"           "gender"           "gender"
##      "var2" "info_source"      "edu"           "edu_field"
## en " "      "3.775"           "4.758"         "5.955"
## mi " "      "0.029"           "0.056"         "0.195"
## ig " "      "0.0289222787196088" "0.055532621834947" "0.193285070225961"
## gr " "      "0.0290359305692985" "0.0557508406430445" "0.194044595641006"
##
##      "gender"           "info_source"      "info_source"
##      "country_grad"     "gender"           "edu"
## en "2.75"               "3.775"           "5.591"
## mi "0.014"              "0.029"           "1.034"
## ig "0.0147054975009455" "0.0293616545707784" "1.03438577710177"
## gr "0.0147632836459375" "0.0104553981667661" "0.368334663551355"
##
##      "info_source"      "info_source"      "edu"
##      "edu_field"        "country_grad"     "gender"
## en "6.684"              "3.72"             "4.758"
## mi "1.276"              "0.857"            "0.056"
## ig "1.27628896650394"   "0.856809862763192" "0.0562102501816651"
## gr "0.454474024564323" "0.30510161635499" "0.0147241479311459"
##
##      "edu"              "edu"              "edu"
##      "info_source"      "edu_field"        "country_grad"
## en "5.591"              "6.643"            "5.302"
## mi "1.034"              "2.326"            "0.284"
## ig "1.03438577710177"   "2.32603173840855" "0.284380106977406"
## gr "0.270955015334333" "0.609298754198382" "0.0744927259757373"
##
##      "edu_field"        "edu_field"        "edu_field"
##      "gender"           "info_source"      "edu"
## en "5.955"              "6.684"            "6.643"
## mi "0.195"              "1.276"            "2.326"
## ig "0.194459242812064"   "1.27628896650394" "2.32603173840855"
## gr "0.0377446573022946" "0.247728464652864" "0.451484175145851"
##
##      "edu_field"        "country_grad"     "country_grad"
##      "country_grad"     "gender"           "info_source"
## en "6.491"              "2.75"             "3.72"
```

```
## mi "0.43" "0.014" "0.857"
## ig "0.429554988421615" "0.0143514685183413" "0.856984861777206"
## gr "0.0833768845131961" "0.008113523021323" "0.484491632063189"
##
## "country_grad" "country_grad"
## "edu" "edu_field"
## en "5.302" "6.491"
## mi "0.284" "0.43"
## ig "0.284379250978661" "0.429490700221313"
## gr "0.160772230148655" "0.24281018205463"
```

rev_ts

univariate and multivariate time series: detection of anomalous features, tests of stationarity and (auto/cross)-correlation

rev_model()

Model testing

check_assumptions()

Checking test or model assumptions about data

reviewed

Reporting function