

# Tilbrigði

## í íslenskri setningagerð

### II. Helstu niðurstöður

#### Tölfræðilegt yfirlit

*Ritstjórar*

Höskuldur Þráinsson,  
Ásgrímur Angantýsson, Einar Freyr Sigurðsson

*Aðrir höfundar*

Ásta Svavarsdóttir, Eiríkur Rögnvaldsson,  
Guðrún Þórðardóttir, Heimir Freyr Viðarsson, Hlíf Árnadóttir,  
Jóhannes Gísli Jónsson, Matthew J. Whelpton, †Salbjörg Óskarsdóttir,  
Sigríður Sigurjónsdóttir, Sigrún Steingrímsdóttir, Tania E. Strahan,  
Theódóra A. Torfadóttir, Þórhallur Eyþórsson, Þórunn Blöndal

Málvísindastofnun Háskóla Íslands  
Reykjavík 2015

*Tilbrigði í íslenskri setningagerð II*

Útgefandi: Málvísindastofnun Háskóla Íslands  
[www.malvis.hi.is](http://www.malvis.hi.is)

©2015 Málvísindastofnun Háskóla Íslands

Dreifing: Háskólaútgáfan (Distribution: University of Iceland Press)  
Dunhaga 18, IS-107 Reykjavík, ICELAND  
<http://haskolautgafan.hi.is/>, [hu@hi.is](mailto:hu@hi.is)

Öll réttindi áskilin.

Bók þessa má ekki afrita með neinum hætti,  
svo sem ljósmyndun, prentun, hljóðritun  
eða á annan sambærilegan hátt, að hluta  
eða í heild, án skriflegs leyfis  
Málvísindastofnunar Háskóla Íslands.

Tölvuvinnsla: Höskuldur Þráinsson  
Kápa: Margrét Lára Höskuldsdóttir  
Prentun: Litlaprent  
[www.prent.is](http://www.prent.is)

ISBN: 978-9979-853-52-7

Listi yfir höfunda efnis eins og hann birtist í I. bindi:

### Höfundar efnis

(hér er í aðalatriðum miðað við þau störf eða stöðu sem höfundar höfðu meðan á rannsókninni stóð, en upplýsingar og netföng hafa sums staðar verið uppfærð)

**Ásgrímur Angantýsson**, doktorsnemi, Íslensku- og menningardeild, Háskóla Íslands, nú lektor við Háskólann á Akureyri: *asgrimur@unak.is*

**Ásta Svavarsdóttir**, rannsóknardósent, Stofnun Árna Magnússonar í íslenskum fræðum: *asta@hi.is*

**Einar Freyr Sigurðsson**, meistaranemi, Íslensku- og menningardeild, Háskóla Íslands, nú doktorsnemi við Univ. of Pennsylvania: *einarsig@babel.ling.upenn.edu*

**Eiríkur Rögnvaldsson**, prófessor, Íslensku- og menningardeild, Háskóla Íslands: *eirikur@hi.is*

**Guðrún Þórðardóttir**, meistaranemi, Íslensku- og menningardeild, Háskóla Íslands: *gudrut@ismennt.is*

**Heimir Freyr Viðarsson**, meistaranemi, Íslensku- og menningardeild, Háskóla Íslands, nú doktorsnemi í sömu deild: *hfv3@hi.is*

**Hlíf Árnadóttir**, meistaranemi, Íslensku- og menningardeild, Háskóla Íslands: *hlifa@hi.is*

**Höskuldur Þráinsson**, prófessor, Íslensku- og menningardeild, Háskóla Íslands: *hoski@hi.is*

**Jóhannes Gisli Jónsson**, aðjunkt, Íslensku- og menningardeild, Háskóla Íslands: *jj@hi.is*

**Matthew J. Whelpton**, dósent, Deild erlendra tungumála, bókmennta og málvísinda, Háskóla Íslands: *whelpton@hi.is*

†**Salbjörg Óskarsdóttir**, BA-nemi, Íslensku- og menningardeild, Háskóla Íslands (Salbjörg lést áður en rannsóknarverkefningu lauk)

**Sigríður Sigurjónsdóttir**, prófessor, Íslensku- og menningardeild, Háskóla Íslands: *siggasig@hi.is*

**Sigrún Steingrímsdóttir**, meistaranemi, Íslensku- og menningardeild, Háskóla Íslands, nú kennari við Kvenmaskólann í Reykjavík: *sigrumst@kvenmo.is*

**Tania E. Strahan**, nýdoktor, Málvísindastofnun Háskóla Íslands, nú starfsmaður Appen Butler Hill, Sydney: *tania.strahan@gmail.com*

**Theódóra A. Torfadóttir**, doktorsnemi, Íslensku- og menningardeild, Háskóla Íslands: *tat@hi.is*

**Þórhallur Eypórsson**, sérfræðingur, Málvísindastofnun Háskóla Íslands, nú dósent, Deild erlendra tungumála, bókmennta og málvísinda, Háskóla Íslands: *tolli@hi.is*

**Þórunn Blöndal**, dósent, Kennaradeild, Menntavísindasviði Háskóla Íslands: *thorumb@hi.is*

# Efnisyfirlit

## II. Helstu niðurstöður – Tölfræðilegt yfirlit

<b>Inngangur að II. bindi</b>	
Hvað er í þessu bindi?	13
Hvaða gagn má hafa af tölfræðilega yfirlitinu?	13
Hvernig er vísað í einstakar kannanir og dæmi?	14
Summary	14
<b>6. Um tölfræðiyfirlitið (Höskuldur Þráinsson, Einar Freyr Sigurðsson)</b>	
<b>6.0 Inngangur</b>	15
<b>6.1 Gagnavinnslan</b>	15
<b>6.2 Nokkur tölfræðihugtök og notkun þeirra í bókinni</b>	20
<b>6.3 Vixlverkun breyta</b>	23
<b>6.4 Frekari dæmi um framsetningu á niðurstöðum</b>	24
6.4.1 <i>Að skilgreina nýjar breytur og reikna meðaltal</i>	24
6.4.2 <i>Að búa til súlurit og línurit</i>	29
<b>6.5 Lokaorð</b>	32
Rit sem vísað er til í þessum kafla	32
Summary	32
<b>7. Fallmörkun (Höskuldur Þráinsson, Þórhallur Eyþórsson, Ásta Svavarsdóttir, Þórunn Blöndal)</b>	
<b>7.1 Athugunarefnið</b>	33
7.1.1 <i>Þágufallshneigð</i>	33
7.1.2 <i>Nefnifallshneigð</i>	34
7.1.3 <i>Andlagsfall</i>	35
7.1.4 <i>Fallglötun og fallflutningur</i>	37
<b>7.2 Þágufallshneigð</b>	38
7.2.1 <i>Almennt yfirlit</i>	38
7.2.2 <i>Tengsl við aldur, menntun og kyn</i>	42
7.2.2.1 Aldur	42
7.2.2.2 Menntun	45
7.2.2.3 Kyn	48
7.2.3 <i>Tengsl við búsetu</i>	49
7.2.4 <i>Samantekt</i>	51
<b>7.3 Nefnifallshneigð</b>	52
7.3.1 <i>Almennt yfirlit</i>	52
7.3.2 <i>Tengsl við aldur, menntun og kyn</i>	53
7.3.2.1 Aldur	53
7.3.2.2 Menntun	54
7.3.2.3 Kyn	55
7.3.3 <i>Tengsl við búsetu</i>	55
7.3.4 <i>Samantekt</i>	56
<b>7.4 Andlagsfall</b>	57
7.4.1 <i>Almennt yfirlit</i>	57
7.4.2 <i>Tengsl við aldur, menntun og kyn</i>	59
7.4.2.1 Aldur	59
7.4.2.2 Menntun	62
7.4.2.3 Kyn	63
7.4.3 <i>Tengsl við búsetu</i>	65
7.4.4 <i>Samantekt</i>	67

<b>7.5</b>	<b>Fallglötun og fallflutningur</b>	68
7.5.1	<i>Almennt yfirlit</i>	68
7.5.2	<i>Tengsl við aldur, menntun og kyn</i>	69
7.5.2.1	Aldur	69
7.5.2.2	Menntun	70
7.5.2.3	Kyn	71
7.5.3	<i>Tengsl við búsetu</i>	71
7.5.4	<i>Samantekt</i>	72
<b>7.6</b>	<b>Lokaorð</b>	72
	<b>Rit sem vísað er til í þessum kafla</b>	73
	<b>Summary</b>	75
<b>8.</b>	<b>Um þolmynd, germynd og það</b> (Höskuldur Þráinsson, Sigríður Sigurjónsdóttir, Hlíf Árnadóttir, Þórhallur Eypórsson)	
<b>8.1</b>	<b>Athugunarefnið</b>	77
8.1.0	<i>Imgangur</i>	77
8.1.1	<i>Leppsetningar og skyld dæmi</i>	78
8.1.2	<i>Hefðbundin þolmynd og þolmörk hennar</i>	80
8.1.3	<i>Nýja þolmyndin</i>	82
<b>8.2</b>	<b>Leppsetningar og skyld dæmi</b>	84
8.2.1	<i>Almennt yfirlit</i>	84
8.2.2	<i>Tengsl við aldur, menntun og kyn</i>	86
8.2.2.1	Aldur	86
8.2.2.2	Menntun	87
8.2.2.3	Kyn	89
8.2.3	<i>Tengsl við búsetu</i>	90
8.2.4	<i>Samantekt</i>	90
<b>8.3</b>	<b>Hefðbundin þolmynd og þolmörk hennar</b>	91
8.3.1	<i>Almennt yfirlit</i>	91
8.3.2	<i>Tengsl við aldur, menntun og kyn</i>	93
8.3.2.1	Aldur	93
8.3.2.2	Menntun	94
8.3.2.3	Kyn	96
8.3.3	<i>Tengsl við búsetu</i>	97
8.3.4	<i>Samantekt</i>	99
<b>8.4</b>	<b>Nýja þolmyndin</b>	99
8.4.1	<i>Almennt yfirlit</i>	99
8.4.2	<i>Tengsl við aldur, menntun og kyn</i>	102
8.4.2.1	Aldur	102
8.4.2.2	Menntun	106
8.4.2.3	Kyn	108
8.4.3	<i>Tengsl við búsetu</i>	110
8.4.4	<i>Samantekt</i>	115
<b>8.5</b>	<b>Lokaorð</b>	116
	<b>Rit sem vísað er til í þessum kafla</b>	117
	<b>Summary</b>	118
<b>9.</b>	<b>Um vera að og vera búinn að</b> (Höskuldur Þráinsson, Theódóra A. Torfadóttir)	
<b>9.1</b>	<b>Athugunarefnið</b>	121
9.1.0	<i>Imgangur</i>	121
9.1.1	<i>Vera að í íþróttamáli</i>	123
9.1.2	<i>Útvíkkun framvinduhorfs</i>	124

9.1.3	Vera búinn að og hafa	126
<b>9.2</b>	<b>Vera að í íþróttamáli</b>	126
9.2.1	Almennt yfirlit	126
9.2.2	Tengsl við aldur, menntun og kyn	128
9.2.2.1	Aldur	128
9.2.2.2	Menntun	130
9.2.2.3	Kyn	131
9.2.3	Tengsl við búsetu	132
9.2.4	Samantekt	133
<b>9.3</b>	<b>Útvíkkun framvinduhorfs</b>	133
9.3.1	Almennt yfirlit	133
9.3.2	Tengsl við aldur, menntun og kyn	137
9.3.2.1	Aldur	137
9.3.2.2	Menntun	141
9.3.2.3	Kyn	142
9.3.3	Tengsl við búsetu	144
9.3.4	Samantekt	146
<b>9.4</b>	<b>Um vera búinn að og hafa</b>	147
9.4.1	Almennt yfirlit	147
9.4.2	Tengsl við aldur, menntun og kyn	147
9.4.2.1	Aldur	147
9.4.2.2	Menntun	148
9.4.2.3	Kyn	149
9.4.3	Tengsl við búsetu	149
9.4.4	Samantekt	150
<b>9.5</b>	<b>Lokaorð</b>	151
	<b>Rit sem vísað er til í þessum kafla</b>	151
	<b>Summary</b>	152
<b>10.</b>	<b>Fornöfn (Höskuldur Þráinsson, Tania E. Strahan)</b>	
<b>10.1</b>	<b>Athugunarefni</b>	155
10.1.0	Imngangur	155
10.1.1	Persónufornöfn og afturbeygð fornöfn í ósamsettum málsgreinum	156
10.1.2	Langdræg afturbeyging	158
<b>10.2</b>	<b>Almennt yfirlit</b>	160
10.2.1	Persónufornöfn og afturbeygð fornöfn í ósamsettum málsgreinum	160
10.2.2	Langdræg afturbeyging	163
10.2.3	Samantekt	164
<b>10.3</b>	<b>Tengsl við aldur, menntun og kyn</b>	165
10.3.1	Aldur	165
10.3.2	Menntun	169
10.3.3	Kyn	170
10.3.4	Samantekt	171
<b>10.4</b>	<b>Tengsl við búsetu</b>	172
<b>10.5</b>	<b>Lokaorð</b>	173
	<b>Rit sem vísað er til í þessum kafla</b>	174
	<b>Summary</b>	175
<b>11.</b>	<b>Hættir og tíðir (Höskuldur Þráinsson, Guðrún Þórðardóttir)</b>	
<b>11.1</b>	<b>Athugunarefni</b>	177
11.1.0	Imngangur	177

11.1.1	<i>Viðtengingarháttur og framsöguháttur í fallsetningum</i>	179
11.1.2	<i>Viðtengingarháttur og framsöguháttur í atvikssetningum</i>	180
11.1.3	<i>Tíðasamræmi</i>	181
<b>11.2</b>	<b>Almennt yfirlit</b>	182
11.2.1	<i>Viðtengingarháttur og framsöguháttur í fallsetningum</i>	182
11.2.2	<i>Viðtengingarháttur og framsöguháttur í atvikssetningum</i>	184
11.2.3	<i>Tíðasamræmi</i>	188
11.2.4	<i>Samantekt</i>	189
<b>11.3</b>	<b>Tengsl við aldur, menntun og kyn</b>	190
11.3.1	<i>Aldur</i>	190
11.3.2	<i>Menntun</i>	193
11.3.3	<i>Kyn</i>	195
11.3.4	<i>Samantekt</i>	196
<b>11.4</b>	<b>Tengsl við búsetu</b>	197
<b>11.5</b>	<b>Lokaorð</b>	199
	<b>Rit sem vísað er til í þessum kafla</b>	199
	<b>Summary</b>	200
<b>12.</b>	<b>Samræmi (Höskuldur Þráinsson, Einar Freyr Sigurðsson, Jóhannes Gísli Jónsson)</b>	
<b>12.1</b>	<b>Athugunarefni</b>	203
12.1.0	<i>Inngangur</i>	203
12.1.1	<i>Samræmi við nefnifallsandlag</i>	203
12.1.2	<i>Samræmi við nefnifall í nafnháttarsambandi eða smáklausu</i>	205
12.1.3	<i>Samræmi við hlutaeignarfall</i>	207
12.1.4	<i>Samræmismörk: Hljómsveitir, lið og samlagning</i>	208
<b>12.2</b>	<b>Samræmi við nefnifallsandlag</b>	209
12.2.1	<i>Almennt yfirlit</i>	209
12.2.2	<i>Tengsl við aldur, menntun og kyn</i>	211
	12.2.2.1 Aldur	211
	12.2.2.2 Menntun	213
	12.2.2.3 Kyn	214
12.2.3	<i>Tengsl við búsetu</i>	215
12.2.4	<i>Samantekt</i>	215
<b>12.3</b>	<b>Samræmi við nefnifall í nafnháttarsambandi eða smáklausu</b>	216
12.3.1	<i>Almennt yfirlit</i>	216
12.3.2	<i>Tengsl við aldur, menntun og kyn</i>	218
	12.3.2.1 Aldur	218
	12.3.2.2 Menntun	218
	12.3.2.3 Kyn	219
12.3.3	<i>Tengsl við búsetu</i>	220
12.3.4	<i>Samantekt</i>	220
<b>12.4</b>	<b>Samræmi við hlutaeignarfall</b>	221
12.4.1	<i>Almennt yfirlit</i>	221
12.4.2	<i>Tengsl við aldur, menntun og kyn</i>	222
	12.4.2.1 Aldur	222
	12.4.2.2 Menntun	222
	12.4.2.3 Kyn	223
12.4.3	<i>Tengsl við búsetu</i>	223
12.4.4	<i>Samantekt</i>	224



<b>12.5 Samræmismörk</b>	225
12.5.1 <i>Almennt yfirlit</i>	225
12.5.2 <i>Tengsl við aldur, menntun og kyn</i>	225
12.5.2.1 Aldur	225
12.5.2.2 Menntun	226
12.5.2.3 Kyn	227
12.5.3 <i>Tengsl við búsetu</i>	227
12.5.4 <i>Samantekt</i>	228
<b>12.6 Lokaorð</b>	228
<b>Rit sem vísað er til í þessum kafla</b>	229
<b>Summary</b>	230
<b>13. Eignarsambönd</b> (Höskuldur Þráinsson, Einar Freyr Sigurðsson, Eiríkur Rögnvaldsson)	
<b>13.1 Athugunarefni</b>	233
13.1.1 <i>Staða eignarfor nafns og eignarfalls</i>	233
13.1.2 <i>Samspil greinis, eignarfor nafns og eignarfalls</i>	234
13.1.3 <i>Ólafsfjarðareignarfallið</i>	236
13.1.4 <i>Hjá-liðir og aðrir forsetningarliðir í eignarmerkingu</i>	237
13.1.5 <i>Eignarfall persónufornafns og eignarfor nafn í tengdum liðum</i>	239
<b>13.2 Staða eignarfor nafns og eignarfalls</b>	239
13.2.1 <i>Almennt yfirlit</i>	239
13.2.2 <i>Tengsl við aldur, menntun og kyn</i>	240
13.2.2.1 Aldur	240
13.2.2.2 Menntun	241
13.2.2.3 Kyn	241
13.2.3 <i>Tengsl við búsetu</i>	241
13.2.4 <i>Samantekt</i>	242
<b>13.3 Samspil greinis, eignarfor nafns og eignarfalls</b>	242
13.3.1 <i>Almennt yfirlit</i>	242
13.3.2 <i>Tengsl við aldur, menntun og kyn</i>	244
13.3.2.1 Aldur	244
13.3.2.2 Menntun	247
13.3.2.3 Kyn	249
13.3.3 <i>Tengsl við búsetu</i>	250
13.3.4 <i>Samantekt</i>	253
<b>13.4 Ólafsfjarðareignarfallið</b>	253
13.4.1 <i>Almennt yfirlit</i>	253
13.4.2 <i>Tengsl við aldur, menntun og kyn</i>	253
13.4.2.1 Aldur	253
13.4.2.2 Menntun	254
13.4.2.3 Kyn	255
13.4.3 <i>Tengsl við búsetu</i>	256
13.4.4 <i>Samantekt</i>	258
<b>13.5 Hjá-liðir og aðrir forsetningarliðir í eignarmerkingu</b>	259
13.5.1 <i>Almennt yfirlit</i>	259
13.5.2 <i>Tengsl við aldur, menntun og kyn</i>	260
13.5.2.1 Aldur	260
13.5.2.2 Menntun	263
13.5.2.3 Kyn	264
13.5.3 <i>Tengsl við búsetu</i>	265
13.5.4 <i>Samantekt</i>	266



<b>13.6 Eignarfall persónufornafns og eignarforanafn í tengdum liðum</b>	267
13.6.1 <i>Almennt yfirlit</i>	267
13.6.2 <i>Tengsl við aldur, menntun og kyn</i>	267
13.6.2.1 Aldur	267
13.6.2.2 Menntun	268
13.6.2.3 Kyn	269
13.6.3 <i>Tengsl við búsetu</i>	270
13.6.4 <i>Samantekt</i>	270
<b>13.7 Lokaorð</b>	271
<b>Rit sem vísað er til í þessum kafla</b>	271
<b>Summary</b>	272
<b>14. Kjarnafærsla, stílfærsla, leppsetningar og frumlagseyða (Höskuldur Þráinsson, Ásgrímur Angantýsson, Heimir Freyr Viðarsson)</b>	
<b>14.1 Athugunarefni</b>	275
14.1.0 <i>Inngangur</i>	275
14.1.1 <i>Kjarnafærsla í aukasetningum</i>	276
14.1.2 <i>Stílfærsla</i>	278
14.1.3 <i>Leppinnskot</i>	278
14.1.4 <i>Frumlagseyða</i>	279
<b>14.2 Almenn yfirlit</b>	280
14.2.1 <i>Kjarnafærsla í aukasetningum</i>	280
14.2.2 <i>Stílfærsla</i>	282
14.2.3 <i>Leppinnskot</i>	282
14.2.4 <i>Frumlagseyða</i>	283
14.2.5 <i>Samantekt</i>	283
<b>14.3 Tengsl við aldur, menntun og kyn</b>	284
14.3.1 <i>Aldur</i>	284
14.3.2 <i>Menntun</i>	288
14.3.3 <i>Kyn</i>	290
14.3.4 <i>Samantekt</i>	291
<b>14.4 Tengsl við búsetu</b>	292
<b>14.5 Lokaorð</b>	294
<b>Rit sem vísað er til í þessum kafla</b>	295
<b>Summary</b>	296
<b>15. Orðaröð í aukasetningum (Höskuldur Þráinsson, Ásgrímur Angantýsson)</b>	
<b>15.1 Athugunarefni</b>	299
15.1.1 <i>Gerðir aukasetninga og aukatengingar</i>	299
15.1.2 <i>S2 og S3 í ýmsum gerðum aukasetninga</i>	300
15.1.3 <i>Innskots-að og staða sagnar</i>	302
15.1.4 <i>S2 og S3 í nafnháttarsamböndum</i>	303
<b>15.2 Brottfall og innskot tengingar</b>	303
15.2.1 <i>Almennt yfirlit</i>	303
15.2.2 <i>Tengsl við aldur, menntun og kyn</i>	305
15.2.2.1 Aldur	305
15.2.2.2 Menntun	305
15.2.2.3 Kyn	306
15.2.3 <i>Tengsl við búsetu</i>	307
15.2.4 <i>Samantekt</i>	308
<b>15.3 S2 og S3 í ýmsum gerðum aukasetninga</b>	308
15.3.1 <i>Almennt yfirlit</i>	308

15.3.2	<i>Tengsl við aldur, menntun og kyn</i>	313
15.3.2.1	Aldur	313
15.3.2.2	Menntun	315
15.3.2.3	Kyn	316
15.3.3	<i>Tengsl við búsetu</i>	317
15.3.4	<i>Samantekt</i>	319
<b>15.4</b>	<b>Innskots-að og staða sagnar</b>	319
15.4.1	<i>Almennt yfirlit</i>	319
15.4.2	<i>Tengsl við aldur, menntun og kyn</i>	321
15.4.2.1	Aldur	321
15.4.2.2	Menntun	322
15.4.2.3	Kyn	323
15.4.3	<i>Tengsl við búsetu</i>	323
15.4.4	<i>Samantekt</i>	325
<b>15.5</b>	<b>S2, S3 og þolfall með nafnhætti</b>	325
15.5.1	<i>Almennt yfirlit</i>	325
15.5.2	<i>Tengsl við aldur, menntun og kyn</i>	326
15.5.2.1	Aldur	326
15.5.2.2	Menntun	326
15.5.2.3	Kyn	326
15.5.3	<i>Tengsl við búsetu</i>	327
15.5.4	<i>Samantekt</i>	327
<b>15.6</b>	<b>Lokaorð</b>	327
	<b>Rit sem vísað er til í þessum kafla</b>	328
	<b>Summary</b>	329
<b>16. Ýmislegt</b>	(Höskuldur Bráinsson, Matthew J. Whelpton, Jóhannes Gísli Jónsson)	
<b>16.1</b>	<b>Athugunarefni</b>	331
16.1.1	<i>Útkomuumsagnir</i>	331
16.1.2	<i>Lyfting út úr nafnhætti með lýsingarorði</i>	333
16.1.3	<i>Ósögð aukafallsfrumlög</i>	334
16.1.4	<i>Staða agna</i>	335
16.1.5	<i>Tvöföldun forsetinga, strand og lokkun</i>	336
16.1.6	<i>Ýmis atriði tengd formöfnum og töluorðum</i>	337
<b>16.2</b>	<b>Útkomuumsagnir</b>	338
16.2.1	<i>Almennt yfirlit</i>	338
16.2.2	<i>Tengsl við aldur, menntun og kyn</i>	340
16.2.2.1	Aldur	340
16.2.2.2	Menntun	341
16.2.2.3	Kyn	342
16.2.3	<i>Tengsl við búsetu</i>	343
16.2.4	<i>Samantekt</i>	343
<b>16.3</b>	<b>Lyfting út úr nafnhætti með lýsingarorði</b>	344
16.3.1	<i>Almennt yfirlit</i>	344
16.3.2	<i>Tengsl við aldur, menntun og kyn</i>	344
16.3.2.1	Aldur	344
16.3.2.2	Menntun	345
16.3.2.3	Kyn	346
16.3.3	<i>Tengsl við búsetu</i>	346
16.3.4	<i>Samantekt</i>	346
<b>16.4</b>	<b>Ósögð aukafallsfrumlög</b>	347
16.4.1	<i>Almennt yfirlit</i>	347

16.4.2	<i>Tengsl við aldur, menntun og kyn</i>	347
16.4.2.1	Aldur	347
16.4.2.2	Menntun	348
16.4.2.3	Kyn	349
16.4.3	<i>Tengsl við búsetu</i>	349
16.4.4	<i>Samantekt</i>	349
<b>16.5</b>	<b>Staða agna</b>	350
16.5.1	<i>Almennt yfirlit</i>	350
16.5.2	<i>Tengsl við aldur, menntun og kyn</i>	350
16.5.2.1	Aldur	350
16.5.2.2	Menntun	351
16.5.2.3	Kyn	351
16.5.3	<i>Tengsl við búsetu</i>	352
16.5.4	<i>Samantekt</i>	352
<b>16.6</b>	<b>Tvöföldun forsetninga, strand og lokkun</b>	352
16.6.1	<i>Almennt yfirlit</i>	352
16.6.2	<i>Tengsl við aldur, menntun og kyn</i>	353
16.6.2.1	Aldur	353
16.6.2.2	Menntun	354
16.6.2.3	Kyn	355
16.6.3	<i>Tengsl við búsetu</i>	355
16.6.4	<i>Samantekt</i>	356
<b>16.7</b>	<b>Ýmis atriði tengd fornöfnum og töluorðum</b>	356
16.7.1	<i>Almennt yfirlit</i>	356
16.7.2	<i>Tengsl við aldur, menntun og kyn</i>	358
16.7.2.1	Aldur	358
16.7.2.2	Menntun	359
16.7.2.3	Kyn	360
16.7.3	<i>Tengsl við búsetu</i>	360
16.7.4	<i>Samantekt</i>	361
<b>16.8</b>	<b>Lokaorð</b>	362
	<b>Rit sem vísað er til í þessum kafla</b>	363
	<b>Summary</b>	364

## 6. Um tölfræðifyfirlitið

Höskuldur Þráinsson, Einar Freyr Sigurðsson

### 6.0 Inngangur

Eins og rakið er í inngangi hér framfar er markmið þessa kafla fyrst og fremst að búa lesendur undir að lesa og skilja eftirfarandi tölfræðifyfirlit og þau hugtök sem þar eru notuð.<sup>1</sup> Um leið ætti kaflinn að auðvelda lesendum að nýta sér niðurstöður Tilbrigðarannsóknarinnar og hliðstæðra rannsókna í eigin skrifum ef þeir hafa hug á því. Í þessu skyni er byrjað á því að kynna hvernig gögnin úr Tilbrigðarannsókninni voru meðhöndluð í tölfræðiforritinu SPSS. Í framhaldi af því eru svo kynnt helstu tölfræðihugtök sem verða notuð í eftirfarandi köflum. Þá er rætt um víxlverkun breyta, þ.e. þegar samband á milli breyta fer eftir gildi þriðju breytu, og bent á leiðir til að bregðast við því. Í öllum þessum útskýringum er stuðst við dæmi úr rannsókninni, einkum með því að skoða ýmiss konar töflur yfir niðurstöður, og í síðasta undirkaflanum eru svo kynntar nokkrar aðferðir til viðbótar til þess að setja niðurstöður fram.

### 6.1 Gagnavinnslan

Unnið var úr úrlausnum þátttakenda í skriflegu könnunum í tölfræðiforritinu SPSS (*Statistical Package for the Social Sciences*). Þetta er mjög öflugt og þekkt tölfræðiforrit sem gefur mikla möguleika á ýmiss konar útreikningum og vinnslu með gögnin (sjá nánar í kafla 6.2 og ýmis dæmi í eftirfarandi köflum). Gallinn er hins vegar sá að forritið er nokkuð dýrt og tiltölulega fáir hafa þess vegna aðgang að því. Í því sambandi má þó benda á að auðvelt er að færa gögn úr SPSS yfir í ýmis forrit sem fleiri þekkja eða hafa aðgang að, t.d. Excel, R eða GNU PSPP.

Áður en hægt er að byrja að slá inn niðurstöður í SPSS þarf að skilgreina hverja breytu fyrir sig.<sup>2</sup> Það er gert í **breytuglugga** (*variable view*) forritsins. Þannig þarf að búa til breytur sem lúta að upplýsingum um þátttakendur, en um þær var spurt á sérstökum blöðum í upphafi kampananna. Mest af þeim upplýsingum eru **félagsbreytur** (aldur, búseta, kyn, menntun). Einnig þarf að skilgreina **málfræðilegu breytur** sjálfar, þ.e. þau setningafræðilegu atriði sem voru til skoðunar, en þannig getum við skráð svör þátttakenda í þeim könnunum sem hér eru á dagskrá, t.d. hvort þeir merktu við *já*, *?* eða *nei* þegar þeir voru beðnir að meta dæmi. Þessar breytur verða oftast kallaðar **setningabreytur** hér á eftir, en til einföldunar verða aðeins tekin skýringardæmi með setningabreytum þar sem hægt var að svara með *já*, *?* eða *nei*.<sup>3</sup>

Þessum breytum má (gróflega) skipta í tvennt: **talnabreytur** og **textabreytur** (þær síðarnefndu eru t.d. fyrir athugasemdir þátttakenda). Talnabreytur eru skilgreindar

<sup>1</sup> Sérstakar þakkir til Guðbjargar Andreu Jónsdóttur, forstöðumanns Félagsvísindastofnunar Háskóla Íslands, en hún las þennan kafla í handriti og benti á margt sem betur mátti fara. Það er ekki hennar sök ef ekki hefur tekist að fara að öllu leyti að þeim ráðum.

<sup>2</sup> Sú lýsing á gagnavinnslu í SPSS sem hér fer á eftir er einkum miðuð við Tilbrigðakönnunina. Fyrir þá sem vilja svólitið almennari lýsingu, en samt stutta og aðgengilega, má benda á grein Kjartans Ólafssonar 2003. Ítarlegri umfjöllun er í handbókum um notkun SPSS, t.d. bók Einars Guðmundssonar og Árna Kristjánssonar 2005.

<sup>3</sup> Ekki er nauðsynlegt að fjalla hér um annars konar breytur, t.d. breytur þar sem velja átti milli tveggja kosta, fylla í eyður o.s.frv., en þær þarf auðvitað líka að skilgreina í SPSS.

Þannig að þeim eru gefin tiltekin töluleg gildi. Þannig koma aðeins tvö gildi til greina þegar spurt er um kyn og þar létum við **karl** hafa gildið 1 og **konu** 2. Við skráningu á búsetu fékk hver könnunarstaður sitt númer og við skiptum landinu þar að auki í átta hluta eins og áður er lýst (sjá (2.2) og *töflur 2.1–2.3* í 2. kafla). Gildi af þessu tagi eru stundum kölluð **flokkabreytur** (flokkakvarðar) eða **nafnbreytur** (e. *nominal*, sjá t.d. Field 2005:48–49, Einar Guðmundsson og Árna Kristjánsson 2005:53–54). Hvaða númer er valið á gildum fyrir þær skiptir engu máli — þær eru bara nöfn. Þannig hefðum við alveg eins getað látið flokkinn **karlar** fá gildið 2 og **konur** gildið 1 og Austurland hefði getað haft gildið 1, 8 eða jafnvel 28.<sup>4</sup> Í þessum gildum felst engin mæling. Röðin á númerum landshlutanna skiptir ekki heldur máli og er í sjálfu sér geðþóttaákvörðun (þótt við höfum reyndar byrjað á Reykjavík og farið hringinn suður- og austurum og landfræðileg nálægð geti skipt einhverju máli um málfar íbúanna). Öðru máli gegnir með gildi fyrir **aldursflokkana**. Þar fékk yngsti aldursflokkurinn gildið 1 (nemendur í 9. bekk, þ.e. fólk sem var um 15 ára gamalt), sá næstýngsti gildið 2, sá næstelsti gildið 3 og sá elsti (65–70 ára) gildið 4. Í þessu felst ákveðin röð og þess vegna eru breytur af þessu tagi kallaðar **raðbreytur** (raðkvarðar, e. *ordinal*, sbr. tilvitnaða staði hér ofar). Í Tilbrigðakönnuninni var síðan ekki nákvæmlega sama bil á milli flokkanna (15 ára, 20–25 ára, 40–45 ára, 65–70 ára) og það er einmitt eitt einkenni á raðbreytum. Þetta er þó kannski ennþá augljósara þegar breytan **menntun** á í hlut. Þar létum við gildið 1 tákna grunnskólamenntun, 2 tákna menntun í framhaldsskóla og 3 háskólamenntun. Í því felst greinilega ákveðin röð, en hins vegar er ekki hægt að segja að framhaldsskólamenntun (gildið 2) sé tvisvar sinnum meiri en grunnskólamenntun (gildið 1). Þegar **þrepabreytur** (þrepakvarðar, hlutfallskvarðar, jafnbilakvarðar, e. *interval/ratio scale*) eiga í hlut á hins vegar að vera jafnt bil á milli allra skilgreindra þrepa í kvarðanum. Ef slíkur kvarði nær t.d. frá 1,00 og upp í 3,00 (eða 100 upp í 300) á að vera sami munur á 1,20 og 1,40 annars vegar og svo 2,50 og 2,70 hins vegar (sjá Field 2005:49, Einar Guðmundsson og Árna Kristjánsson 2005:54). Í könnunum á sviði hug- og félagsvísinda er oft erfitt að segja til um hvort tölulegir kvarðar eru í raun þrepakvarðar fremur en raðkvarðar.

Þegar gildi voru ákveðin fyrir setningabreytur fékk *já* gildið 1, *? gildið 2* og *nei* 3. Á mynd 6.1 má sjá dæmi um það hvernig við skilgreindum nokkrar breytur:<sup>5</sup>

<sup>4</sup> Í raun voru landshlutunum gefin gildi frá einum og upp í 8. Reykjavík fékk gildið 1, Reykjanes 2, Suðurland 3 og síðan áfram rangsælis austur og norður um land og endað á Vesturlandi, sem fékk gildið 8.

<sup>5</sup> Eins og áður er rakið stendur stafurinn *T* í heiti breytu (t.d. *T3* á mynd 6.1) fyrir *Tilbrigðakönnun*. Skriflegu kannanirnar voru alls þrjár, eins og fram kemur í 2. kafla. Þess vegna voru síðar notuð fimm stafa heiti fyrir breytur. Þar merkir þá *T2001* ‘Tilbrigðakönnun 2, dæmi númer 1’, *T3104* ‘Tilbrigðakönnun 3, dæmi númer 104’ o.s.frv. Þegar vísað er í dæmi úr Tilbrigðakönnun 1 í textanum hér á eftir eru þau auðkennð á þennan hátt. Þess vegna væri þar vísað í þriðja dæmið úr Tilbrigðakönnun 1 sem *T1003* en ekki bara *T3* þótt það hafi verið númerað þannig í innsláttarforritinu.



	Name	Type	Width	Decimals	Label	Values	Missing
5	Dagsetning	Date	10	0	Könnunardagur	None	None
6	Stadur	Numeric	8	0	Könnunarstaður	{1, Akureyri}...	None
7	Landshlutar	Numeric	8	0	Landshlutar	{1, Reykjavík}...	None
8	Kyn	Numeric	8	0	Kyn þátttakanda	{1, Karl}...	None
9	Fæðingarar	Numeric	4	0	Fæðingarár þátttakanda	None	None
10	Aldursflokk	Numeric	8	0	Aldursflokkun	{1, 9. bekkur}...	9
11	Buseta	String	60	0	Hvar býrðu núna?	None	None
12	Buseta2	String	150	0	Hvar ólst þú upp?	None	None
13	Buseta3	Numeric	8	0	Hefur þú búið eða dvalist lengur en	{1, Nei}...	9
14	Buseta4	String	180	0	Ef já hvar, hvenær og hversu lengi?	None	None
15	Modurmal	Numeric	8	0	Móðurmál þátttakanda	{1, Íslenska}...	None
16	Modurmal2	String	20	0	Hvert ef ekki íslenska?	None	None
17	Modurmal3	Numeric	8	0	Er talað annað tungumál en íslensk	{1, Já}...	9
18	Modurmal4	String	200	0	Ef já hvaða tungumál?	None	None
19	Starf	String	101	0	Starf þátttakanda	None	None
20	Menntun	Numeric	8	0	Menntun þátttakanda	{1, Grunnskóla}...	9
21	Menntun2	Numeric	8	0	Menntun föður	{1, Grunnskóla}...	9
22	Menntun3	Numeric	8	0	Menntun móður	{1, Grunnskóla}...	9
23	T1	Numeric	8	0	Hann notar aldrei hana.	{1, Já = Eðlileg}...	9
24	T1aths	String	78	0		None	None
25	T2	Numeric	8	0	Hann notar hana aldrei.	{1, Já = Eðlileg}...	9
26	T2aths	String	61	0		None	None
27	T3	Numeric	8	0	Þeir eru bara ekki að átta sig á aðal	{1, Já = Eðlileg}...	9
28	T3aths	String	200	0		None	None
29	T4	Numeric	8	0	Bíllinn hjá henni er í viðgerð.	{1, Já = Eðlileg}...	9
30	T4aths	String	95	0		None	None
31	T5	Numeric	8	0	Það var rekið manninn út af staðnu	{1, Já = Eðlileg}...	9
32	T5aths	String	40	0		None	None
33	T6	Numeric	8	0	Það var strax dæmt vítaspjrymu.	{1, Já = Eðlileg}...	9
34	T6aths	String	64	0		None	None

Mynd 6.1: Breytugluggi í SPSS (dæmi úr Tilbrigðakönnun 1).

Í hverri röð er ein breyta skilgreind. Henni þarf að gefa  **nafn**  (*name*) en það birtist svo í sérstökum gagnaglugga. Hægt er að skýra breytuna frekar, en merkingin (*label*) kemur þó aðeins fyrir í breytuglugga (líkt og gildir um aðra flokka en nafn). Einnig verður að ákveða **gerð** (*type*) breytunnar. Aðeins talnabreytur fá **gildi** (*values*). Síðan er hægt að ákveða að slá inn tiltekna tölu í hvert skipti sem upplýsingar vantar um tiltekið atriði hjá einhverjum þátttakanda (*missing value*).

Þegar breytur hafa verið skilgreindar getur innsláttur hafist í **gagnaglugga** (*data view*). Dæmi um slíkan glugga má sjá á mynd 6.2:

Tilbrigði í íslenskrí setningagerð II

File Edit View Data Transform Analyze Graphs Utilities Window Help

32 : Stadur 2

	Stadur	Landshlutar	Kyn	Faedingarar	Aldursflokkur	Buseta	Buseta2
32	2	6	1	1991	1	Blönduós	Blönduósi
33	2	6	1	1991	1	Blönduósi	Blönduósi
34	2	6	1	1991	1	Blönduósi	Blönduósi
35	2	6	1	1991	1	Blönduós	Blönduósi, fy
36	2	6	1	1991	1	Blönduósi	Blönduósi og
37	2	6	1	1991	1	Blönduósi	Blönduósi
38	2	6	2	1991	1	Blönduós	Blönduósi
39	2	6	2	1991	1	Blönduósi	Blönduósi (o
40	2	6	1	1982	2	Blönduósi	Blönduósi og
41	2	6	1	1983	2	Blönduósi	Blönduósi

Mynd 6.2: Gagnagluggi í SPSS (dæmi úr Tilbrigðakönnun 1).

Hver dálkur í gagnaglugga tákna eina breytu sem hefur áður verið skilgreind í breytuglugga. Í hverri röð má svo slá inn upplýsingar og svör fyrir einn þáttakanda. Á myndinni getur því að líta upplýsingar um 10 þátttakendur á Blönduósi í Tilbrigðakönnun 1. Við sjáum það á því að í dálkinum *Buseta* (label Hvar býrðu núna?) stendur Blönduós(i) en líka á því að *Stadur* (label Könnunarstaður) hefur gildið 2 en það tákna einmitt Blönduós. Þátttakendurnir eru úr landshluta 6 sem er Norðvesturland og eru ýmist í aldursflokki 1 (fæddir 1991) eða 2 (fæddir 1982–1983).

En lítum nú á það hvernig þeir 10 einstaklingar sem koma við sögu á mynd 6.2 dæmdu sex fyrstu setningarnar í Tilbrigðakönnun 1:

File Edit View Data Transform Analyze Graphs Utilities Window Help

33 : T5aths hent honum út

	T1	T1aths	T2	T2aths	T3	T3aths	T4	T4aths	T5	T5aths	T6	T6aths
32	3		1		1		3		1		2	
33	3		1		1		3	hennar	2	hent h	1	
34	3		1		1		3		2		1	
35	3		1		9		2		2		2	
36	3		1		3		3		1		1	
37	3		1		2		3		3		2	
38	2		1		1		1		3		1	
39	3		1		2		3		3		2	
40	3		1		1		3		3		2	
41	3		1		1		3		3		3	

Mynd 6.3: Gagnagluggi í SPSS (dæmi úr Tilbrigðakönnun 1).

Á mynd 6.3 má sjá hvernig setningabreytur T1–T6 eru skilgreindar. Þar sést þó ekki að gildi þessara breyta eru 1 fyrir *já*, 2 fyrir *?* og 3 fyrir *nei*. Þegar þær upplýsingar liggja fyrir er hægt að átta sig á dómum þátttakenda. Á mynd 6.3 má þannig sjá að enginn



Þessara 10 þátttakenda samþykkti setningu T1 (*Hann notar aldrei hana*), einum þykir hún vafasöm (2 = ?) og níu segja hana ótæka (3 = *nei*). Aftur á móti samþykkja allir T2 (*Hann notar hana aldrei*) (1 = *já*). Setningar T1 og T2 voru þrjúsetningar (sjá umræðu í kafla 2.1), en gert var ráð fyrir að nánast enginn gæti sagt fyrri setninguna en að nær allir gætu sagt þá síðari. Hugmyndin var sú að slíkar setningar gætu gefið vísbendingar um það hvort þátttakendur tækju könnunina alvarlega og væru með hugann við verkið. Svo reyndist jafnan vera.

Eins og lesa má úr mynd 6.1 eru T1aths–T6aths textabreytur, en þá er hægt að færa inn athugasemdir sem þátttakendum gafst kostur á að gera við dæmasetningar. Einn þessara 10 þátttakenda (í röð 33) gerir athugasemdir við tvær setningar. Í athugasemdadálk við hliðina á setningu T4 (*Billim hjá henni er í viðgerð*) á könnunarblaði skrifar hann „hennar“ og við setningu T5 (*Það var rekið mannum út af staðnum*) gerir hann athugasemdina „hent honum út“ (athugasemdin sést ekki öll á myndinni).

Setning T4 var notuð til að kanna *hjá*-eignarliði (sjá nánar um þá í köflum 13 og 19). Út frá athugasemd umrædds þátttakanda má álykta að hann hafi hafnað setningunni vegna þess að hann kjósi frekar *Billim hennar er í viðgerð*. Í setningu T5 var mat þátttakanda á nýju þolmyndinni kannað (sjá nánar um hana í köflum 8 og 26). Sama þátttakanda þykir setningin vafasöm. Ef hann hefði ekki gert athugasemd hefði mátt ætla að hann vildi ekkert með nýju þolmyndina hafa. Af athugasemd hans má þó ráða annað. Þátttakandinn virðist vilja breyta setningunni í *Það var hent honum út (af staðnum)* en það er einmitt dæmi um nýja þolmynd!

Þegar tölfræði er reiknuð út í SPSS vinnur forritið aðeins með talnabreytur — það „skilur“ ekki textann sem skráður er í textabreytureiti. Ef ætlunin er að reikna út hve margir Vestfirðingar samþykktu tiltekna setningu vinnur forritið með gildi tveggja breyta. Þá reiknar það út hve margir þátttakendur sem hafa gildið 7 (= ‘Vestfirðir’) fyrir landshlutabreytuna (sbr. það sem segir hér að framan) hafa jafnframt gildið 1 (= *já*) fyrir viðkomandi setningabreytu.

Nánar er fjallað um útreikninga í SPSS í næstu undirköflum, en í (6.2) eru gefin dæmi um atriði sem auðveldlega er hægt að láta forritið reikna út eða gera:

- (6.1) a. **Hlutfall** (prósenta) þátttakenda sem samþykkja tiltekna setningu, telja hana vafasama eða ótæka (og þá má annaðhvort miða við allan hópinn eða tiltekinn hluta hans).
- b. Samanburður á dómum þátttakenda miðað við ýmsar **félagslegar breytur**, svo sem aldur (aldurshópa), kyn, staði, landshluta, menntun, menntun foreldra o.s.frv. Þetta má ýmist gera með töflum eða með súluritum, línuritum eða á annan myndrænan hátt eins og nánar er lýst hér á eftir og fram kemur í ýmsum yfirlitsköflum.
- c. Athugun á **fylgni** milli tiltekinna breyta, t.d. hversu mikil fylgni er á milli aldurs og tiltekinna dóma og mat á því hvort sú fylgni er **tölfræðilega marktæk** eða ekki (þ.e. hversu miklar líkur eru til þess að fylgni í úrtaki sé tilviljun og ekki til staðar í þýði). Forritið gefur líka kost á því að könnuð sé fylgni á milli tiltekinna mállegra breyta.
- d. Samanburður á niðurstöðum eftir því hvaða **aðferð** er notuð (mat á setningum hverri fyrir sig, val á milli kosta o.s.frv., sbr. umræðu um aðferðir í 2. kafla, einkum 2.2 og 2.3).
- e. Hægt er að búa til **undirflokka þátttakenda** og skoða niðurstöður fyrir þá sérstaklega. Þannig má skoða tiltekinn aldurshóp eða -hópa, sleppa t.d.

ýngsta hópnunum ef verið er að skoða áhrif menntunar. Eins má skoða sérstaklega þá sem búa á ákveðnum stað eða svæði, jafnvel einstökum hverfum í Reykjavík (póstnúmer voru skráð í Reykjavík).

Sú fyrirhöfn sem er fólgin í því að slá gögnin inn í þetta öflugra forrit er því fljót að borga sig vegna þeirra miklu möguleika sem forritið gefur á því að greina gögnin.

Í lýsingunni á möguleikum forritsins hér fyrir ofan eru nefnd nokkur tölfræðihugtök, svo sem **fylgni** og **marktækni**. Til að geta haft fullt gagn af því sem segir í næstu köflum um niðurstöður úr Tilbrigðakönnuninni er nauðsynlegt að hafa hugmynd um hvað þessi hugtök merkja. Þess vegna verða nokkur algeng tölfræðihugtök skýrð stuttlega í næsta undirkafla.

## 6.2 Nokkur tölfræðihugtök og notkun þeirra í bókinni

Auðveldast er að skýra tölfræðihugtök í tengslum við ákveðin dæmi. Lítum fyrst á eftirfarandi töflu (þessi dæmi eru til umræðu í 8. kafla):<sup>6</sup>

Númer	Setning	9. b	20–25	40–45	65–70	r	p	N
a. T1093	Það var strítt <b>honum</b> á hverjum degi.	<b>52,9</b>	18,4	1,6	1,1	<b>.549</b>	<b>.000</b>	<b>768</b>
b. T1019	Það var <b>borðað</b> svo margar kjötbollar.	<b>36,8</b>	12,7	3,7	5,6	<b>.401</b>	<b>.000</b>	<b>769</b>
c. L3.6	Það er kominn <b>leigubíllinn</b> .	7,5	7,3	9,0	<b>16,4</b>	-.074	.047	<b>712</b>

Tafla 6.1: *Jákvætt mat á nokkrum dæmum, flokkað eftir aldri þátttakenda.*

Eins og lýst er í 2. kafla og í kafla 6.1 voru oftast þrenns konar svör í boði í skriflegu könnunum: *já* (= eðlileg setning), *?* (= vafasöm setning) og *nei* (= ótæk setning). Í þessari töflu er mat þátttakenda flokkað eftir aldri þeirra og tölurnar í dálkunum sem eru merktir 9.b. (= þátttakendur í 9. bekk, þ.e. unglingar um 15 ára aldur), 20–25 ára, 40–45 ára og 65–70 ára í *töflu 6.1* sýna **hlutfall** (prósentu) þátttakenda í hverjum aldurshóp sem taldi viðkomandi setningu eðlilegt mál (merkti við *já* á eyðublaðinu). Þess vegna segir í skýringum við töfluna að hún sýni „jákvætt mat“ þátttakenda. Síðan er hæsta hlutfallið feitletrað fyrir hvert dæmi og við sjáum að það er hæst fyrir yngsta aldurshópinn í dæmum *a* og *b* en fyrir þann elsta í dæmi *c*. Talan í aftasta dálkinum, sem er merktur með N, táknar fjölda þeirra sem svöruðu viðkomandi spurningu (mátu setninguna), eins og venja er í töflum af þessu tagi.

Nú er munurinn á dómum aldurshópanna mjög greinilegur hér og þess vegna má ætla að hann sé **tölfræðilega marktækur** eins og það er kallað. Þegar slíkt er rannsakað er metið hversu líklegt er að tiltekinn munur geti stafað af tilviljun. Í þessari töflu er þetta sýnt í tveim næstsíðustu dálkunum. Sá fyrri er merktur með *r* og talan í honum sýnir svokallaðan **fylgnistuðul** (e. *correlation coefficient*). Í þessu tilviki er nánar tiltekið um að ræða raðfylgnistuðul Spearmans (*Spearman's rho*), en rétt þykir að nota hann hér af því að breytur sem um er að ræða í *töflu 6.1* eru ekki mældar á jafnbilakvarða (sjá t.d. umræðu um fylgni hjá Field 2005:32 og ýmis dæmi um

<sup>6</sup> Íslenska ritvenjan er að skrifa 0,549 en ekki .549 o.s.frv. eins og hér er gert til að spara pláss. Það væri líka nóg að hafa tvo aukastafi til að sýna fylgnina, en hér og í flestum samsvarandi töflum eru fylgnitölurnar teknar upp eins og þær koma úr forritinu.

fylgniútreikninga í bók Einarar Guðmundssonar og Árna Kristjánssonar 2005).<sup>7</sup> Þessi stuðull er mælikvarði á **fylgni** milli breyta og gildi hans liggur einhvers staðar á milli gildanna  $-1$  og  $+1$ . Málið er einfaldast ef verið er að skoða fylgni milli tveggja breyta eins og hér (aldursflokkur annars vegar og hins vegar „einkunnagjöf“ (þ.e. gildið 1, 2 eða 3) fyrir dæmi um nýju þolmyndina). Ef fylgnistuðullinn er 0 er ekki um neina fylgni að ræða, ef hann er  $-1$  er um að ræða fullkomna neikvæða fylgni (ef önnur breytan hækkar um eitthvað tiltekið lækkar hin á sama hátt) og ef fylgnistuðullinn er  $+1$  er um fullkomna jákvæða fylgni að ræða (breyturnar fylgjast nákvæmlega að, þ.e. hækka eða lækka á sama hátt). Jákvæð fylgni gefur hér til kynna að eftir því sem þátttakendur eru eldri, því líklegri séu þeir til að hafna setningunum. Neikvæð fylgni táknar að með hækkingu aldri aukast líkur á jákvæðu mati á setningu. Í rannsóknunum á borð við þá sem hér er á dagskrá koma gildin  $-1$ ,  $+1$  og  $0$  yfirleitt aldrei fyrir heldur liggur gildi fylgnistuðulsins einhvers staðar á bilinu milli  $-1$  og  $+1$ . Þá vaknar auðvitað sú spurning hvað teljist mikil fylgni eða lítil í þessu samhengi. Samkvæmt Field (2005:32) hefur verið stungið upp á eftirfarandi mælikvarða fyrir kannanir af þessu tagi (ekki skiptir máli hvort gildið á  $r$  er jákvætt eða neikvætt):

(6.2) Mælikvarði á fylgni:

- $r = 0,1$ : **lítil fylgni** (útskýrir 1% af öllum breytileikanum:  $0,1 \times 0,1 = 0,01$ )
- $r = 0,3$ : **meðalfylgni** (útskýrir 9% af öllum breytileikanum:  $0,3 \times 0,3 = 0,09$ )
- $r = 0,5$ : **mikil fylgni** (útskýrir 25% breytileikans:  $0,5 \times 0,5 = 0,25$ )

Við getum haft þetta til hliðsjónar við mat á þeirri fylgni milli aldurs þátttakenda og mats þeirra á dæmum um nýju þolmyndina svokölluðu sem fram kemur í *töflu 6.1*. Samkvæmt þessu er fylgnin mikil fyrir  $a$ -dæmið, milli þess að vera í meðallagi og mikil fyrir  $b$ -dæmið en mjög lítil ( $r < 0,1$ ) fyrir  $c$ -dæmið (þar er hún  $-0,074$ ).<sup>8</sup>

Nú hafa einhverjir lesendur sjálfsagt rekið augun í það að fylgnistuðullinn fyrir  $c$ -dæmið er neikvæður (hann er  $-0,074$ ). Ástæðan er sú að því hærri sem aldur þátttakenda er því lægri var „einkunnin“ sem þeir gáfu. Innsláttargildin fyrir aldurshópana (þ.e. þær tölur sem voru notaðar í kóðuninni á aldursflokkunum) voru nefnilega 1 = yngsti flokkur, 2 = næstýngsti flokkur, 3 = næstelsti flokkur, 4 = elsti flokkur. Þess vegna sýnir neikvæð fylgni á milli einkunna og aldurs að mat eldri þátttakenda var jákvæðara en þeirra yngri (þ.e. hærri tala fyrir aldursflokk eykur líkur á lágru (= jákvæðari) tölu fyrir matið).

<sup>7</sup> Annar vel þekktur fylgnistuðull er fylgnistuðull Pearsons (*Pearson's correlation coefficient*). Hann er einkum notaður þegar verið er að skoða fylgni á milli þrepabreyta (breyta á jafnbilakvarða, sbr. umræðu hjá Amalíu Björnsdóttur 2003:120 til dæmis). Í ýmsum tilvikum getur verið álitamál hvor stuðullinn á betur við og oft fæst mjög svipuð útkoma hvor aðferðin sem er notuð, enda eru þær skyldar í eðli sínu. SPSS reiknar yfirleitt gildi fyrir báða þessa stuðla þegar verið er að meta fylgni og fyrir setningarnar í *töflu 6.1* eru þeir 0,545/0,549 fyrir fyrsta dæmið, 0,392/0,401 fyrir annað dæmið og  $-0,091/-0,074$  fyrir það þriðja (fylgnistuðull Pearsons talinn á undan, enda skilar SPSS stuðlunum jafnan í þeirri röð). Marktækniþölurnar eru sama eðlis í báðum tilvikum. Um fleiri tegundir fylgnistuðla og notkun þeirra í SPSS má t.d. lesa hjá Kjartani Ólafssyni 2003.

<sup>8</sup> Hér er vert að athuga að þessi fylgni er reiknuð út frá öllum dómum allra þátttakenda í þessum könnunum, þ.e. dómumum *já* (= eðlileg setning, tölugildi 1 í úrvinnslu), *?* (= vafasöm setning, tölugildi 2) og *nei* (= ótæk setning, tölugildi 3) en ekki bara þeim hlutfallstölum um jákvæða dóma sem eru sýndar í *töflu 6.1*.



En mikil fylgni getur í sjálfu sér verið tilviljun. Það gæti gerst vegna þess að úrtakið væri mjög lítið (fáir þátttakendur) eða vegna þess að ákveðnar tölfræðilegar líkur eru á því að úrtök séu skökk eða ólík þýði þrátt fyrir að vera valin með tilviljunaraðferð. Í dálknum sem er merktur með  $p$  kemur fram mat á þessu. Þetta svokallaða  $p$ -gildi sýnir hversu miklar líkur eru til að fylgnin sé tilviljun. Því lægra sem það er, því minni líkur eru til þess að það sé tilviljun. Þetta gildi getur hæst verið 1, en það myndi þýða 100% líkur á því að fylgnin (eða munurinn á breytunum eða hvað það er sem verið er að skoða) væri tilviljun. Í *töflu 6.1* er þetta gildi 0,000 fyrir dæmi  $a$  og  $b$ , en það merkir í raun að það er minna en 0,001. Í því felst að líkurnar fyrir því að aldursfylgnin sé tilviljun eru minni en 0,1% (minni en 1‰). Þegar tölfræðingar eru spurðir hvort tölfræðilegar niðurstöður séu marktækar miða þeir yfirleitt annaðhvort við  $p \leq 0,05$  eða  $p \leq 0,01$ . Ef  $p$ -gildið er minna en 0,05 merkir það að líkurnar á að niðurstöðurnar séu tilviljun (fylgnin sé tilviljun eða munurinn á breytunum tilviljun) eru minni en 5%. Ef  $p$ -gildið er minna en 0,01 eru líkurnar á tilviljun minni en 1%. Ef við lítum á  $c$ -dæmið í *töflu 6.1* sjáum við að þar er  $p$ -gildið 0,047. Það merkir að sú litla fylgni sem þar er að finna (-0,074) er marktæk miðað við  $p \leq 0,05$  markið en ekki miðað við  $p \leq 0,01$  markið.

Eins og áður sagði hefur fjöldi þátttakenda í könnun mikil áhrif á  $p$ -gildið eins og augljóst er: Því fleiri sem þátttakendur eru, því minni líkur eru á að einhver tiltekin fylgni (eða tölfræðilegur munur) sé tilviljun. Eins má segja að því sterkari sem fylgnin er (því meiri sem munurinn er), því færri þátttakendur þurfi til þess að hún teljist marktæk. Þetta metur tölfræðiforritið fyrir okkur.<sup>9</sup> Þetta sést vel ef mat þátttakenda á Blönduósi í *töflu 6.2* er borið saman við *töflu 6.1*, en á Blönduósi voru þátttakendur aðeins 30 í þeim könnunum sem hér eru á dagskrá (Könnun 1 og 3):

Númer	Setning	9. b	20–25	40–45	65–70	r	p	N
a. T1093	Það var strítt honum á hverjum degi.	75,0	12,5	0,0	0,0	.672	.000	30
b. T1019	Það var borðað svo margar kjötbollar.	50,0	12,5	0,0	16,7	.514	.004	30
c. L3.6	Það er kominn leigubíllinn.	0,0	0,0	14,3	28,6	-.250	.183	30

Tafla 6.2: *Jákvætt mat á nokkrum dæmum, flokkað eftir aldri þátttakenda á Blönduósi.*

Hér eru þátttakendur aðeins 30 en ekki yfir 700 eins og í *töflu 6.1*. Þrátt fyrir það er fylgni dóma um  $a$ -dæmið ennþá mjög vel marktæk ( $p = 0,000$ ) og sömuleiðis þegar  $b$ -dæmið á í hlut ( $p = 0,004$ ). Fylgnistuðullinn fyrir  $b$ -dæmið er reyndar aðeins hærra en í *töflu 6.1* en  $p$ -gildið er samt aðeins hærra hér þó enn sé það vel undir viðmiðunarmarkinu  $p \leq 0,01$ . En þótt fylgni milli aldurs og dómanna um  $c$ -dæmið sé nú talsvert meiri en í *töflu 6.1* (fylgnistuðullinn er -0,250 hér en var -0,074 þar) dugir það ekki til þess að fylgnin teljist núna tölfræðilega marktæk við  $p \leq 0,05$  markið ( $p$ -gildi er hér 0,183).

Af því dæmi sem nú var sýnt sést líka að vegna hins mikla fjölda þátttakenda sem var í þessum könnunum getur jafnvel mjög lítil fylgni talist tölfræðilega marktæk. Þess

<sup>9</sup> Vegna þessa er ekki hægt að gefa einhlítt svar við spurningum eins og: „Hvað þarf ég að hafa marga þátttakendur í minni könnun til þess að niðurstöðurnar verði tölfræðilega marktækar?“ Ef línur eru skýrar og einfaldar (mikil fylgni, mikill munur) er hægt að fá tölfræðilega marktækar niðurstöður þótt þátttakendur séu fáir. Annars þurfa þeir að vera fleiri, eins og skýrt er með dæmum í meginmálinu.

vegna er yfirleitt aðeins vakin athygli á því með feitlitrún í töflum í eftirfarandi köflum ef niðurstöður eru marktækar miðað við  $p \leq 0,01$  markið. Hins vegar er vakin athygli á fylgnistuðlum sem ná 0,3 markinu (meðalfylgni skv. viðmiðum í (6.2)).

### 6.3 Víxlverkun breyta

Loks er vert að vekja athygli á því að í ákveðnum tilvikum getur verið erfitt að átta sig á raunverulegri fylgni milli breyta vegna þess að stundum getur verið um víxlverkun að ræða, þ.e. samband á milli tveggja breyta getur verið háð þriðju breytu. Þetta er t.d. nokkuð augljóst þegar félagslegu breytunar **aldur** og **menntun** eiga í hlut. Menntun þátttakenda er grófluga flokkuð í þrennt, eins og áður var nefnt: grunnskólamenntun, framhaldsskóli, háskólamenntun. Þar sem þátttakendur í yngsta aldurshópnum voru aðeins um 15 ára gamlir er augljóst að þeir geta aðeins verið með grunnskólamenntun (eða hluta hennar). Ef við ætlum að skoða fylgni menntunar við tiltekna málbreytu, t.d. nýju þolmyndina sem áður var nefnd, þarf að hafa þetta í huga. Þetta má skýra með dæmum. Í *töflu 6.3* er sýnd niðurstaðan af því að biðja forritið að reikna út fylgni menntunar við dóma um setningar *a* og *b* í undanfarandi töflum:

Númer	Setning	grsk	frhsk	hásk	r	p	N
a. T1093	Það var strítt honum á hverjum degi.	32,9	7,9	0,0	.369	.000	750
b. T1019	Það var borðað svo margar kjötbollur.	26,2	5,5	0,0	.339	.000	751

Tafla 6.3: *Jákvætt mat á dæmum um nýju þolmyndina, flokkað eftir menntun þátttakenda.*

Samkvæmt þessu mætti ætla að veruleg tengsl væru á milli menntunar og dóma um þessar setningar. Hins vegar höfum við þegar séð skýr tengsl milli mats á dæmum af þessu tagi og aldurs þátttakenda og af þeim ástæðum sem þegar voru raktar gætu þau áhrif spilað hér inn í: Yngsti aldurshópurinn hefur ekki haft tækifæri á meiri menntun en grunnskóla þannig að hann fellur allur í þann menntunarflokk. Til að koma í veg fyrir þessa víxlverkun má nota tvær leiðir. Önnur er sú að sleppa yngsta aldurshópnum alveg úr og skoða aðeins áhrif menntunar í hinum aldurshópnum þrem. Þá verður niðurstaðan eins og sýnt er í *töflu 6.4*:

Númer	Setning	grsk	frhsk	hásk	r	p	N
a. T1093	Það var strítt honum á hverjum degi.	11,0	7,3	0,0	.131	.002	546
b. T1019	Það var borðað svo margar kjötbollur.	14,7	5,2	0,0	.184	.000	547

Tafla 6.4: *Jákvætt mat á dæmum um nýju þolmyndina, flokkað eftir menntun þátttakenda. Yngsta aldurshóp sleppt.*

Hér sjáum við að fylgni við menntun hefur minnkað verulega en hún er samt mjög vel marktæk.<sup>10</sup>

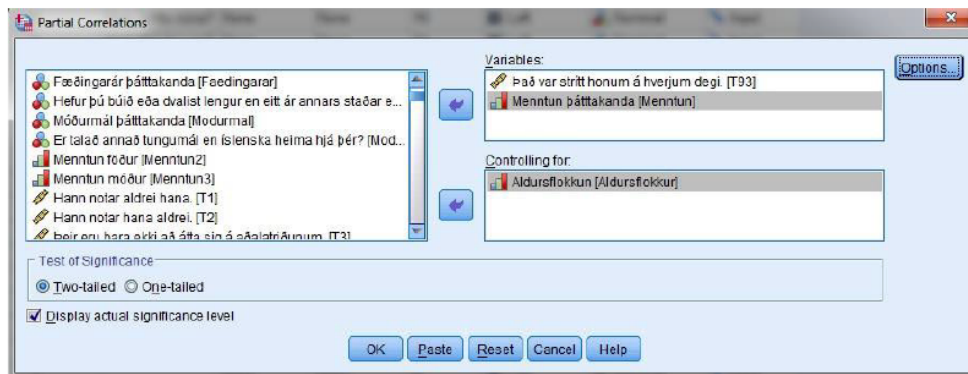
Önnur leið til að losna við víxlverkun milli breyta er að láta forritið reikna út svokallaða **hlutfylgni** (e. *partial correlation*, sbr. t.d. Field 2005:134 o.áfr.; Einar Guðmundsson og Árna Kristjánsson 2005:311 o.áfr.). Þetta er gert á eftirfarandi hátt:

<sup>10</sup> Hér hefði jafnvel verið rétt að sleppa tveimur yngri hópnum því einungis hluti af næstýngsta hópnum er nógu gamall til að hafa lokið háskólanámi. Þó er auðvitað slæmt að sleppa svo stórum hóp úr þýðinu. Auk þess hefði það tæplega gengið í þessu tilviki þar sem eldri hóparnir tveir hafna nýju þolmyndinni að mestu leyti (sjá *töflu 6.1*).

(6.3) Reiknuð hlutfylgni í SPSS:

- velja *Analyze* í SPSS-glugganum
- velja *Correlate* undir *Analyze*
- velja *Partial* undir *Correlate* (ef valið er *Bivariate* kemur sama niðurstaða og þegar beðið er um *Correlation* undir *Crosstabs* til dæmis, sbr. nánari umræðu í næsta undirkafli)

Ef við viljum t.d. fara þessa leið til að láta forritið reikna fylgni menntunar við dóma um setningu *T1093* (a-dæmið í undanfarandi töflum) gerum við það í valglugga sem lítur út eins og sýnt er á mynd 6.4:



Mynd 6.4: Sýnishorn af því hvernig má láta SPSS reikna út hlutfylgni.

Hér er þá ætlunin að skoða fylgnina á milli mats á setningunni *Það var strítt honum á hverjum degi* (dæmi *T1093*, eða *T93* eins og það heitir þarna) og menntunar þátttakenda **að teknu tilliti til** (e. *controlling for*) aldurs. Þegar þessari aðferð er beitt segir forritið að fylgnistuðullinn sé 0,229 og *p*-gildið 0,000, en það er heldur meiri fylgni en fékkst með því einu að útiloka yngsta hópinn eins og gert var í *töflu 6.3*. Ástæðan er væntanlega sú að aldursbreytan hefur haft einhver áhrif á fylgnistuðulinn í *töflu 6.3* þótt búið væri að útiloka yngsta hópinn. Þrátt fyrir það munum við oft nota þá einföldu aðferð að útiloka yngsta hópinn og skoða fylgni við menntun hjá hinum hópunum þrem án þess að hafa áhyggjur af samspili menntunar og aldurs í þeim hópum.

Annars konar víxlverkun breyta er síðan eitt af því sem vert væri að skoða í fleiri tilvikum en gert er í eftirfarandi köflum. Þar má t.d. nefna tengslin á milli búsetu og menntunarstigs þátttakenda. Hlutfall háskólamenntaðra þátttakenda var nokkuð mismunandi í einstökum landshlutum og ekki hið sama frá einni könnun til annarrar.

## 6.4 Frekari dæmi um framsetningu á niðurstöðum

Í síðasta kafla voru allar helstu niðurstöður settar fram í töfluformi. SPSS gefur auðvitað kost á margs konar framsetningu af öðru tagi, en í þessari bók eru aðeins fáir af þeim möguleikum notaðir. Hér verða þó nefnd nokkur dæmi til skýringar á því sem fram kemur í síðari köflum.

### 6.4.1 Að skilgreina nýjar breytur og reikna meðaltal

Í þeim könnunum sem hér eru á dagskrá voru oftast notuð nokkur dæmi um sama fyrirbærið. Þannig voru allmörg dæmi nýtt til þess að kanna útbreiðslu nýju þol-

myndarinnar t.d. Þótt fróðlegt sé að skoða mat þátttakenda á þessum dæmum hverju fyrir sig getur líka verið gagnlegt að reyna að fá heildarmynd af fyrirbærinu. Þetta má gera með því að reikna nokkurs konar „meðaleinkunn“ fyrir þau dæmi sem eiga saman. Á þann hátt er t.d. hægt að fá gleggri sýn á mun einstakra hópa þátttakenda eða tengsl félagslegra breyta við mállegar. Lítum fyrst á nokkur dæmi um nýju þolmyndina og flokkum niðurstöðurnar eftir landshlutum að þessu sinni (sjá nánari umræðu í 8. kafla):

Númer	Setning	Rvk	Rnes	Suð	Aust	NE	NV	Vfir	Vlan
a. T1031	Það var hrint <b>stelpunni</b> af hjólinu svo hún meiddist.	14,4	<b>24,1</b>	19,8	15,2	11,9	20,5	10,2	8,8
b. T1093	Það var strítt <b>honum</b> á hverjum degi.	21,6	<b>35,2</b>	21,6	14,2	16,5	26,8	10,2	15,2
c. T1005	Það var <b>rekið manninn</b> út af staðnum.	18,9	<b>25,9</b>	19,1	17,1	15,8	19,5	20,7	13,2
d. T1006	Það var strax <b>dæmt vítaspyrnu</b> .	25,2	<b>59,3</b>	38,7	32,7	36,6	35,4	22,0	28,1
e. T1019	Það var <b>borðað</b> svo margar <b>kjötbollar</b> .	15,3	<b>35,2</b>	18,0	10,6	12,7	18,3	10,2	10,5
f. T1038	Það var strax <b>rekið hana</b> út af.	18,9	<b>33,3</b>	19,8	14,3	12,0	26,5	10,2	11,5
g. T1043	Það var <b>beðið mig</b> að vaska upp.	13,6	24,1	22,5	18,1	17,2	<b>25,3</b>	15,3	12,4
h. T1055	Það var <b>skammað mig</b> fyrir letina.	12,6	16,7	18,2	14,2	9,7	<b>21,7</b>	11,9	9,8
i. T1051	Jón var sendur heim en <b>það</b> var bara <b>skammað Óla</b> .	22,7	<b>44,4</b>	24,5	21,0	18,8	34,9	32,2	21,4
j. T1114	Það voru nokkrir reknir en bara <b>áminnt aðra</b> .	14,8	22,6	15,6	19,0	20,1	<b>36,6</b>	22,4	15,9

Tafla 6.5: *Jákvætt mat á dæmigerðum setningum með nýju þolmyndinni, flokkað eftir landshlutum.*

Hér gefa feitletuðu tölurnar vísbendingar um að þessar setningar hafi yfirleitt fengið einna besta dóma á Reykjanesi en einnig stundum á Norðvesturlandi. Það er samt ekki gott að átta sig á því hvort þarna er um marktækan mun að ræða, hvorki fyrir einstök dæmi né fyrir dæmasafnið í heild. Hér er lærdómsríkt að bera saman tvær aðferðir til að skoða slíkt, þ.e. fylgnistuðulinn (*Pearson's correlation coefficient*), sem við höfum skoðað fyrir í þessum kafla, og kí-kvaðratprófið (*(Pearson's) chi-square*), til að átta sig á eðlismun þeirra.

Ef forritið er bæði beðið að reikna fylgni (e. *correlation*) og beita kí-kvaðratprófinu á dæmin sem tekin eru í *töflu 6.5* fæst eftirfarandi niðurstaða:<sup>11</sup>

<sup>11</sup> Eins og fram kemur hér á eftir er ekkert vit í því að skoða fylgnistuðul í þessu tilviki — það er aðeins gert hér til þess að útskýra muninn á nafnbreytum annars vegar og raðbreytum og þrepabreytum hins vegar.



Númer	Setning	Kí-kvaðrat		Pearsons r	
		kí-kvaðr.	P	r	P
a. T1031	<b>Það</b> var hrint <b>stelpunni</b> af hjólinu svo hún meiddist.	24,89	.036	.044	.227
b. T1093	<b>Það</b> var stritt <b>honum</b> á hverjum degi.	36,73	<b>.001</b>	.038	.297
c. T1005	<b>Það</b> var <b>rekið manninn</b> út af staðnum.	16,91	.261	.038	.298
d. T1006	<b>Það</b> var strax <b>dæmt vítaspyrnu</b> .	36,56	<b>.001</b>	.029	.419
e. T1019	<b>Það var borðað</b> svo margar <b>kjötbollur</b> .	30,00	<b>.008</b>	.065	.071
f. T1038	<b>Það</b> var strax <b>rekið hana</b> út af.	30,09	<b>.007</b>	.073	.043
g. T1043	<b>Það</b> var <b>beðið mig</b> að vaska upp.	17,62	.225	-.001	.974
h. T1055	<b>Það</b> var <b>skammað mig</b> fyrir letina.	19,89	.134	.011	.756
i. T1051	Jón var sendur heim en <b>það</b> var bara <b>skammað Óla</b> .	33,08	<b>.003</b>	-.014	.707
j. T1114	<b>Það</b> voru nokkrir reknir en bara <b>áminnt aðra</b> .	35,47	<b>.001</b>	-.081	.025

Tafla 6.6: Kí-kvaðratprófi og fylgniprófi (Pearsons r) beitt á gögnin í töflu 6.5.

Samkvæmt kí-kvaðratprófinu er landshlutamunurinn marktækur við  $p \leq 0,01$  markið í sex tilvikum (og reyndar líka stundum við  $p \leq 0,001$  markið) en hér kemur aldrei fram fylgni sem er tölfræðilega marktæk á sama hátt samkvæmt fylgnistuðli Pearsons. Ef við myndum hins vegar beita sömu prófum á mat þátttakenda á sömu setningum flokkað eftir aldurshópum myndu þau skila svipuðum niðurstöðum (munurinn á mati aldurshópanna/fylgni dómanna við aldurshópa mælist marktækur við  $p \leq 0,01$  markið í báðum tilvikum og reyndar líka við  $p \leq 0,001$  markið). Hvernig getur staðið á þessu?

Ástæðan er sú að landshlutabreytan er hvorki þrepabreyta né raðbreyta í eðli sínu heldur **nafnbreyta** (sjá umræðu um breytugerðir í kafla 6.1). Þegar um er að ræða raðbreytu eins og aldursflokkaðbreytuna eru ákveðin rökleg tengsl á milli þeirra gilda sem er úthlutað í gagnavinnslnu, eins og bent var á hér framar. Við innslátt á upplýsingum um aldur þátttakenda fær yngsti aldursflokkurinn gildið 1, sá næstýngsti 2, sá þriðji 3 og sá elsti 4. Elsti hópurinn fær hæsta gildið, sá yngsti það lægsta og það hefði ekki gefið rétta mynd að úthluta þessum gildum af handahófi (láta t.d. næstýngsta flokkinn fá gildið 4 eða þann elsta gildið 2). Sama á við um menntunarbreytuna. Þannig er þetta ekki þegar nafnbreytur eiga í hlut — gildum fyrir þær er í raun úthlutað að geðþótta. Þess vegna gefur það ekki rétta mynd að reikna fylgni við þær á sama hátt og við raðbreytur eða þrepabreytur.<sup>12</sup> Í útreikningum með nafnbreytum má hins vegar nota kí-kvaðratprófið. Þegar niðurstöður úr því prófi eru skoðaðar er tiltölulega hátt kí-kvaðratgildi jafnan samferða lágu  $p$ -gildi ef um marktækan mun er að ræða, eins og sjá má í *töflu 6.6*. Reyndar þarf kí-kvaðratgildið að vera því hærra sem **frelsisgráðurnar** svokölluðu eru fleiri (líka kallaðar **frigráður** — þær ráðast af fjölda raða og dálka í töflunni og eru kallaðar á ensku *degrees of freedom*, skammstafað *df*) og í þessu tilviki eru þær 14.<sup>13</sup> Ef flett er upp í töflu sem sýnir svokallað **vendigildi** (e. *critical value*) fyrir kí-kvaðrat miðað við fjölda frelisisgráða (t.d. hjá Field

<sup>12</sup> Það er hægt að reikna fylgni við nafnbreytur eins og kyn af því að þær taka bara tvö gildi. Þá segir tölugildi fylgnistuðulsins til um fylgnina en formerkið á honum (plús eða mínus) ræðst bara af því hvaða tölugildi við höfum valið, t.d. hvort við ákveðum að gefa kynbreytunni gildið 1 fyrir karla og 2 fyrir konur eða öfugt. Það skiptir engu máli í sjálfu sér (sjá kafla 6.1).

<sup>13</sup> Um nánari umræðu um hugtakið **frelsisgráða** (eða **frigráða**) má vísa í kennslubækur eða handbækur um tölfræði (sjá t.d. Field 2005:319–320). SPSS-forritið lætur þess jafnan getið í tölfræðilegum útreikningum með kí-kvaðratprófinu hversu margar frelisisgráðurnar eru í hverju tilviki svo við þurfum ekki að hafa miklar áhyggjur af því.

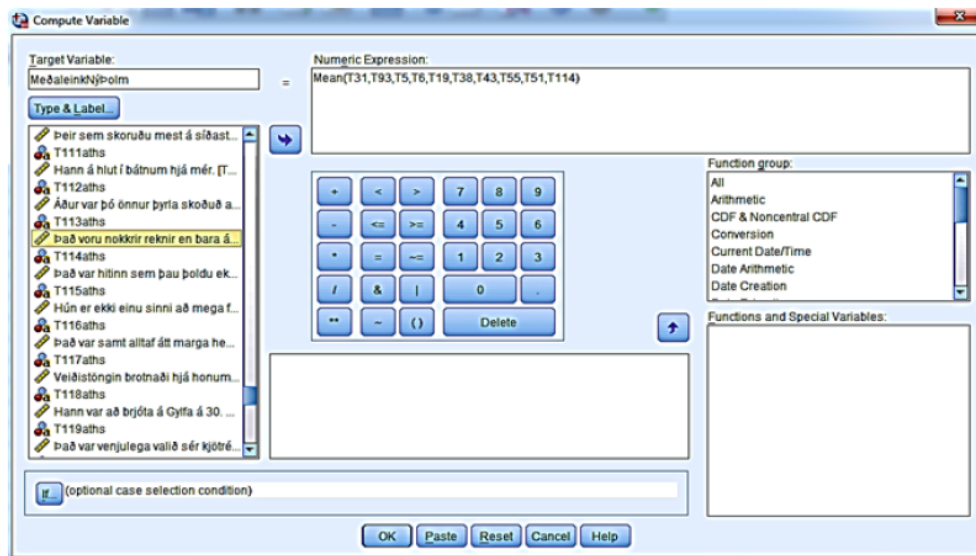
2005:760) má sjá að fyrir 14 frelsisgráður er vendigildið fyrir kí-kvaðrat 23,68 miðað við  $p \leq 0,05$  markið. Það passar við útkomuna í *töflu 6.6*. Aftur á móti þyrfti kí-kvaðratgildið að ná 29,14 til að munurinn yrði marktækur við  $p \leq 0,01$  markið.<sup>14</sup>

Þótt við höfum fengið að vita að munurinn á dómum þátttakenda í ólíkum landshlutum um sumar af setningunum í *töflu 6.5* sé tölfraeðilega marktækur samkvæmt kí-kvaðratprófinu verður ekki sagt að sú tafla gefi skýra mynd af þessum mun. Til þess að gera hann sýnilegri má byrja á því að reikna nokkurs konar „meðaleinkunn“ fyrir öll dæmin sem eru talin í *töflu 6.5*. Það er gert með því að búa til nýja breytu eins og hér verður rakið:

(6.4) Að skilgreina nýja breytu og búa til meðaleinkunn:

- velja *Transform* í SPSS-glugganum
- velja *Compute Variable*

Þá kemur upp valgluggi þar sem hægt er að gefa nýju breytunni heiti og skilgreina hana. Á *mynd 6.5* er sýnt hvernig hægt er að búa til nýja breytu með því að nota skipunina *Mean*. Þá verður til ný breyta sem er meðaltal fyrir þau gildi sem hver þátttakandi gefur þessum 10 dæmum sem þarna eru á dagskrá (setningarnar í *töflu 6.5*). Þá býr forritið til nýja breytu með því heiti sem þarna er búið að skilgreina og bætir henni aftan við breytulistann í gagnaglugganum.<sup>15</sup>



*Mynd 6.5: Sýnishorn af því hvernig búa má til nýja breytu (í þessu tilviki sameinaða breytu (meðaltal) fyrir dæmi um nýju þolmyndina).*

<sup>14</sup> Með **vendigildi** er átt við það mark sem kí-kvaðratgildið þarf að ná til þess að munurinn teljist marktækur miðað við  $p = 0,05$  eða  $p = 0,01$ . Töflur fyrir slík gildi má finna í ýmsum tölfraeðibókum og einnig á Netinu. Þegar þetta er ritað (í júlí 2012) er slík tafla t.d. á slóðinni <http://home.comcast.net/~sharov/PopEcol/tables/chisq.html>.

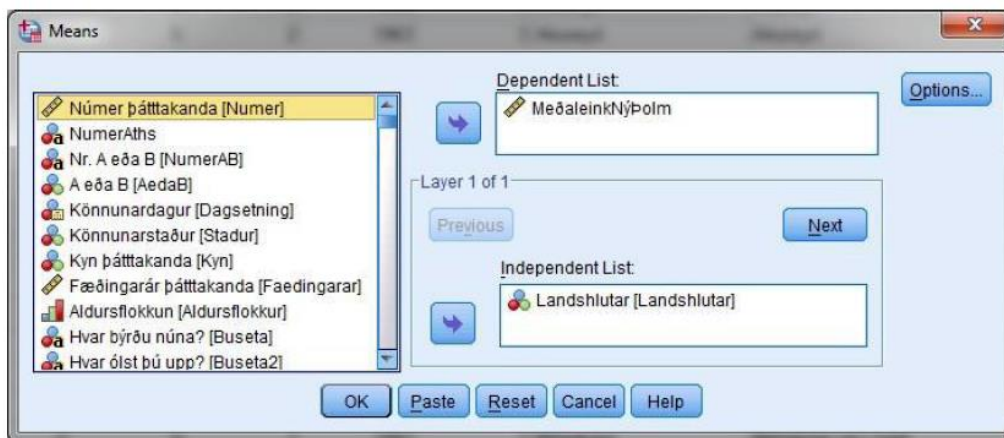
<sup>15</sup> Þetta má líka gera með því að láta forritið leggja saman gildi allra breytanna sem eru til skoðunar og deila síðan í með fjölda þeirra. Það liti þá svona út í þessu tilviki:  $(T31+T93+T5+T6+T19+T38+T43+T55+T51+T114)/10$ . Gallinn við þá aðferð er sá að þá sleppir forritið að reikna meðaltöl fyrir þá einstaklinga sem ekki hafa metið öll dæmin.

Við getum síðan notað þessa nýju breytu (meðalbreytu eða „meðaleinkunn“ fyrir þessi dæmi um nýju þolmyndina) til þess að bera landshlutana saman. Það má t.d. gera með eftirfarandi aðgerð:

(6.5) Meðaltöl borin saman:

- velja *Analyze* í SPSS-glugganum
- velja *Compare means*
- velja *Means*

Þá kemur upp valgluggi sem hægt er að nota til að láta forritið reikna út meðaltöl fyrir hvern landshluta til dæmis. Þetta er sýnt á mynd 6.6:



Mynd 6.6: Dæmi um samanburð meðaltala (hér eru þetta meðaltöl fyrir sameinaða breytu fyrir nýju þolmyndina).

Þá reiknar forritið meðaltöl fyrir hvern landshluta eins og sýnt er í töflu 6.7:

Landshluti	Meðaltal	N	Staðalfrávik
1. Reykjavík	2,5781	111	,63757
<b>2. Reykjanes</b>	<b>2,2280</b>	54	,67272
3. Suðurland	2,4510	111	,61816
4. Austurland	2,5604	106	,60783
5. Norðurland eystra	2,5402	134	,58478
6. Norðurland vestra	<b>2,3418</b>	83	,65741
7. Vestfirðir	2,5305	59	,54433
8. Vesturland	2,5640	114	,54126

Tafla 6.7: Meðaltalsgildi („meðaleinkunnir“) fyrir dæmi um nýju þolmyndina.

Hér er vert að hafa í huga að meðaltölin í töflunni eru „meðaleinkunnir“, þ.e. meðaltal þeirra gilda sem þátttakendur í viðkomandi landshluta gáfu þessum dæmum, þar sem 1 merkir ‘já, eðlileg setning’, 2 merkir ‘vafasöm setning’ og 3 ‘nei, ótæk setning’. Lág meðaleinkunn merkir því að tiltölulega margir þátttakendur hafa talið dæmin eðlilegt mál en há meðaleinkunn að tiltölulega margir hafi hafnað þeim. Taflan staðfestir þannig það hugboð sem við fengum eftir að hafa skoðað *töflu 6.5*, nefnilega að þessar setningar fengu hvað besta dóma á Reykjanesi og á Norðurlandi vestra.



Í framhaldi af þessu má svo finna út hvort munurinn á þessum meðaltölum er tölfræðilega marktækur. Það má gera með því að fara alveg sömu skref og sýnd voru í (6.5) en bæta síðan við því sem fram kemur í (6.6):

(6.6) Meðaltöl borin saman:

- velja *Analyze* í SPSS-glugganum og *Compare Means* og *Means* (sbr. (6.5))
- velja síðan *Options* (sjá mynd 6.6) og haka við *Anova table and eta* og síðan *Continue* og loks *OK* þegar valglugginn á mynd 6.6 birtist aftur

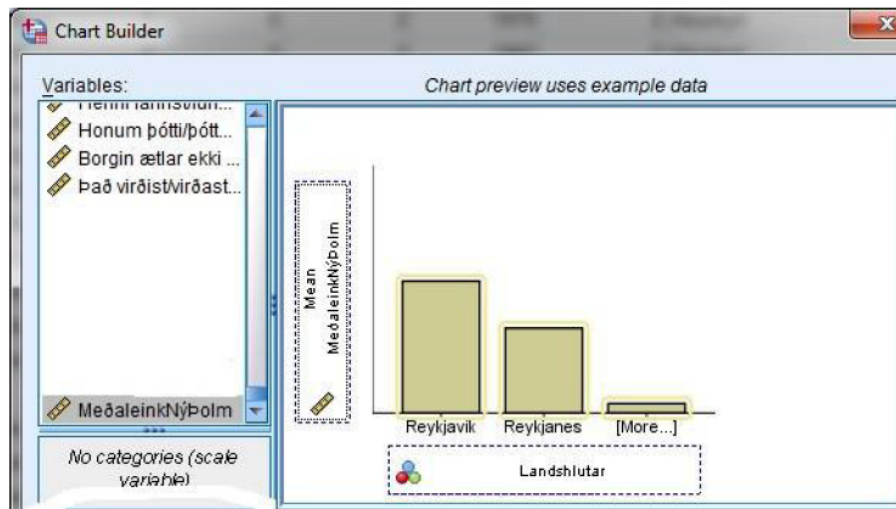
Ef þetta er gert verður til ANOVA-tafla sem sýnir að munurinn á meðaltalseinknum landshlutanna er mjög vel marktækur ( $F_{(7, 764)} = 3,159, p = 0,003$ ).<sup>16</sup>

#### 6.4.2 Að búa til súlurit og línurit

Þennan mun á landshlutunum má þó gera ennþá sýnilegri en hér er gert með því að láta forritið búa til súlurit eða línurit. Það er gert á þessa leið:

(6.7) Súlurit búið til:

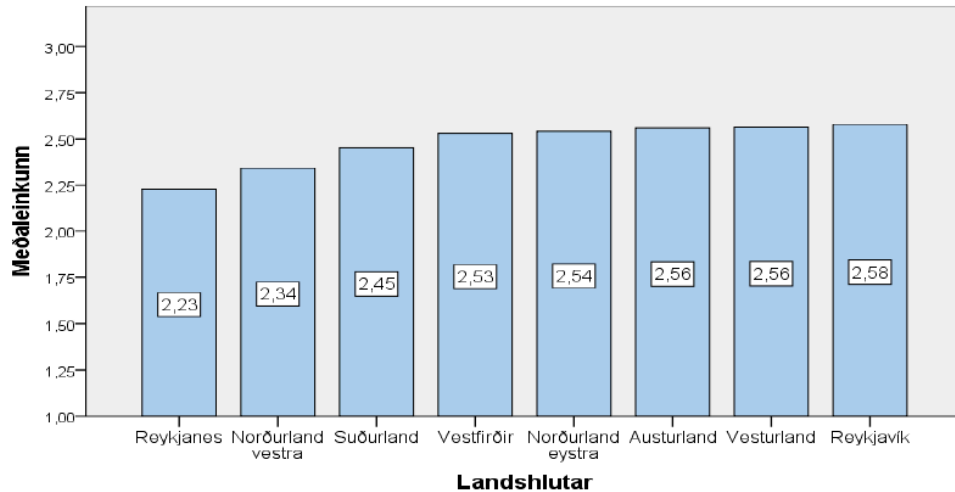
- velja *Graphs* í SPSS-glugganum
- velja *Chart Builder*
- velja *OK*
- velja súlurit af þeirri gerð sem óskað er eftir
- setja viðeigandi breytur inn á X-ásinn og Y-ásinn, sbr. mynd 6.7:



Mynd 6.7: Búið til súlurit fyrir meðalsgildi nýju þolmyndarinnar í einstökum landshlutum.

<sup>16</sup> Hér táknar F svonefint F-gildi því tölfræðiprófið sem SPSS notar þarna er kallað F-próf. Tölurnar í niðurskrifaða sviganum við F eru frelsisgráðurnar (e. *degrees of freedom, df*), en þær eru oft kallaðar *df1* og *df2* í þessu samhengi (sjá umræðu hjá Einari Guðmundssyni og Árna Kristjánssyni 2005:370 til dæmis). Sú fyrri er hér fjöldi landshluta mínus einn en sú síðari fjöldi þátttakenda (N). Niðurstöður úr F-prófum eru yfirleitt sýndar á þennan hátt en fyrir óinnvígða er nóg að huga að *p*-gildinu. ANOVA stendur fyrir *analysis of variance*.

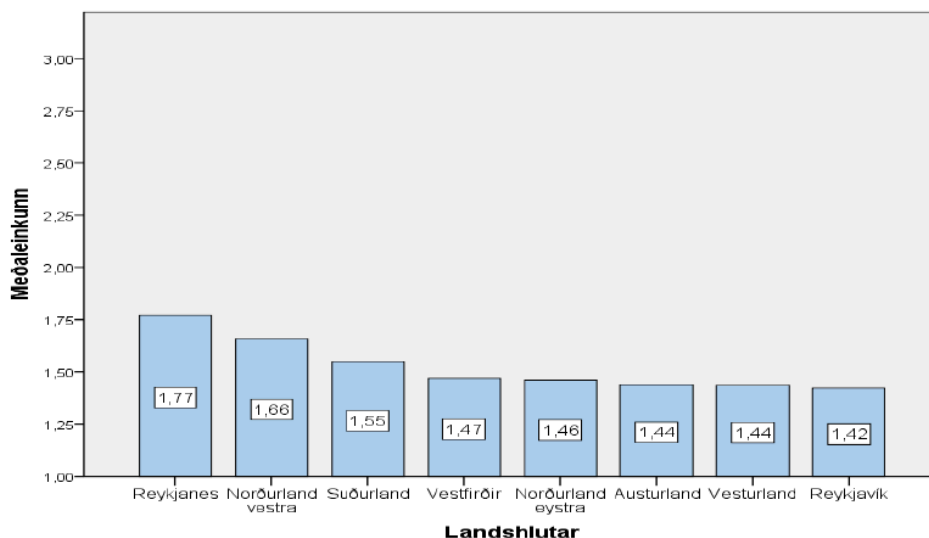
Út úr þessu gæti þá komið súluritið sem er sýnt á mynd 6.8. Þá er búið að snyrta það dálítið til, t.d. laga skalann fyrir Y-ásinn (hann á að byrja á 1, ekki 0 eins og er sjálfgefið, og ná upp í 3), láta forritið raða landshlutunum eftir „meðaleinkunnum“, velja lit á súlurnar, setja meðaleinkunnina inn á súlurnar (*data labels*) o.s.frv.



Mynd 6.8: Meðaltalsgildi fyrir nýju þolmyndina í öllum landshlutum.

Hér sést munur landshlutanna betur en í töflunum hér á undan.

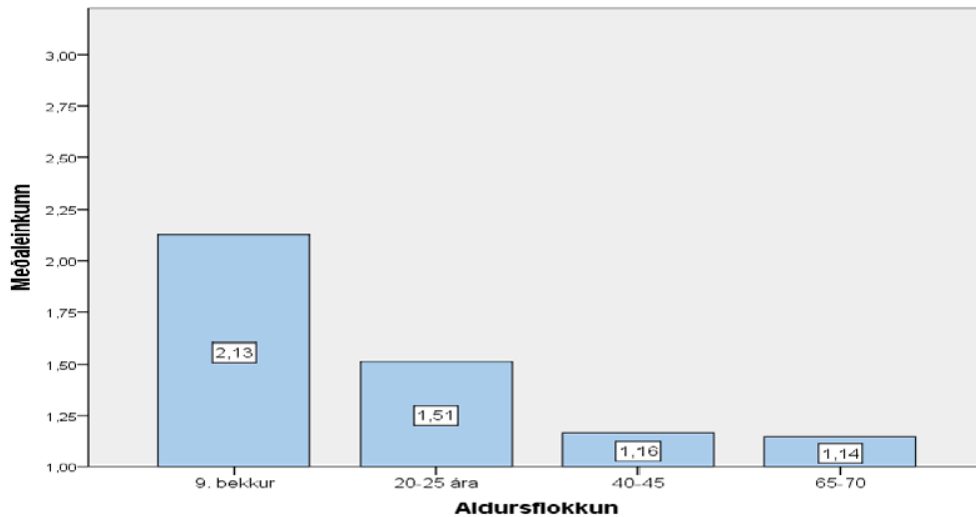
Nú má hins vegar segja að það væri „eðlilegra“ að súlan væri hæst (en ekki lægst) fyrir þann landshluta þar sem nýja þolmyndin fær besta dóma og það er auðvelt að láta forritið snúa því þannig með því að breyta gildinu 1 í 3 og gildinu 3 í 1 fyrir viðkomandi breytur áður en farið er að reikna. Ef við gerum það í þessu tilviki lítur súluritið svona út (sjá mynd 8.2 í 8. kafla og umræður í kringum hana):



Mynd 6.9: Samanburður á meðaleinkunn landshluta fyrir dæmigerðar setningar með nýrri þolmynd (breytunni snúið við).

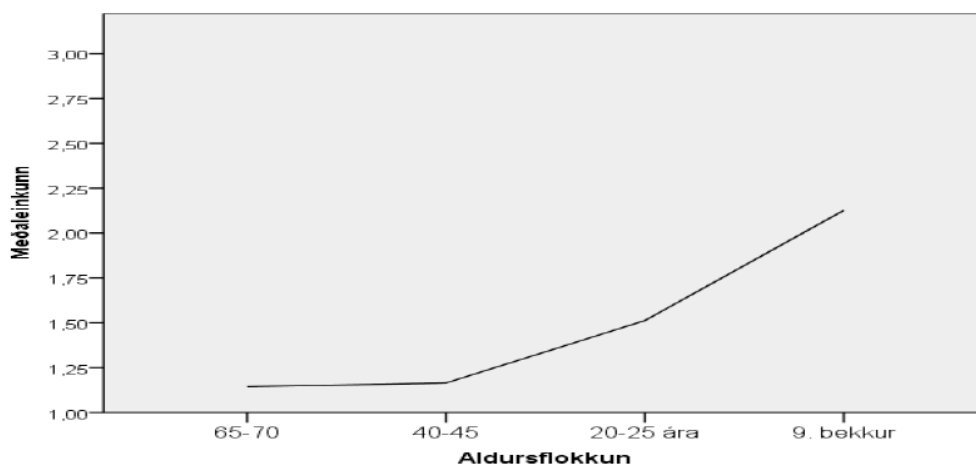
Þessi aðferð er notuð í eftirfarandi köflum þegar gagnlegt þykir að sýna á myndrænan hátt eftirtektarverðan mun á þátttakendum eftir aldri, menntun, kyni eða búsetu.

SPSS gefur kost á því að búa til ýmsar aðrar gerðir mynda en þessar, t.d. með línunum í staðinn fyrir súlur. Það væri lítið vit í því að búa til línurit sem samsvaraði súluritinu á mynd 6.9, af því að það eru engin „rökleg tengsl“ á milli landshlutanúmeranna. Öðru máli gegnir um aldurshópa. Súlurit sem sýnir meðaleinkunn fyrir nýju þolmyndina hjá einstökum aldurshópum sést á mynd 6.10:



Mynd 6.10: Súlurit sem sýnir meðaleinkunn aldurshópanna fyrir nýju þolmyndina.

Hér sést skýr fylgni við aldur. Hún kæmi líka skýrt út á línuriti og þá mætti eins raða aldurshópunum á hinn veginn, þ.e. frá þeim elsta til þess yngsta. Þá myndi línan auðvitað fara hækkandi og þannig endurspeglar tímaþróunina betur ef lítið er svo á að myndin sýni **málbreytingu í sýndartíma** (e. *language change in apparent time*, sbr. umræðu hjá Höskuldi Þráinssyni, Ástu Svavarsdóttur o.fl. 2013):



Mynd 6.11: Línurit sem sýnir meðaleinkunn aldurshópanna fyrir nýju þolmyndina (breytummi snúid við eins og á mynd 6.9, sbr. skýringu þar).

Það er þá í raun smekksatriði hvort valið er að sýna niðurstöður af þessu tagi á línuriti eða súluriti. Í þessari bók eru oftast notuð súluriti.

## 6.5 Lokaorð

Hér hafa aðeins verið gefin dæmi um örfáar af þeim aðferðum sem hægt er að nota í SPSS til þess að vinna með rammsóknagögn af því tagi sem hér eru á dagskrá og setja fram tölfraðilegar niðurstöður um þau á lýsandi hátt. Þeim sem vilja fræðast nánar um þessa hluti er bent á þau rit sem talin eru hér í ritaskrá eða aðrar greinar, handbækur og kennslubækur um tölfraði. Frá sjónarmiði málfræðinga er gallinn við flest af slíkum ritum sá að þar er ekki verið að fjalla um málfræðileg gögn og því ekki alltaf auðvelt að sjá hvað á best við þegar unnið er úr málfræðilegum könnunum. Hins vegar má þar fræðast frekar um almenn hugtök og aðferðir í tölfraði.

### Rit sem vísað er til í þessum kafla

- Amalía Björnsdóttir. 2003. Útskýringar á helstu tölfraðihugtökum. Sigríður Halldórsdóttir og Kristján Kristjánsson (ritstj.): *Handbók í aðferðafræði og rammsóknum í heilbrigðisvísindum*, bls. 115–129. Háskólinn á Akureyri, Akureyri.
- Einar Guðmundsson og Árni Kristjánsson. 2005. *Gagnavinnsla í SPSS*. Háskólaútgáfan, Reykjavík.
- Field, Andy. 2005. *Discovering Statistics Using SPSS (and Sex, Drugs and Rock 'n' Roll)*. 2. útg. Sage Publications, London.
- Höskuldur Þráinsson, Ásta Svavarsdóttir, Eiríkur Rögnvaldsson, Jóhannes Gísli Jónsson, Sigríður Sigurjónsdóttir og Þórunn Blöndal. 2013. Hvert stefnir í íslenskri setningagerð? Um samtímalegar kannanir og málbreytingar. *Íslenskt mál* 35:57–127.
- Kjartan Ólafsson. 2003. Grunnatriði í SPSS. Sigríður Halldórsdóttir og Kristján Kristjánsson (ritstj.): *Handbók í aðferðafræði og rammsóknum í heilbrigðisvísindum*, bls. 131–159. Háskólinn á Akureyri, Akureyri.

## SUMMARY

This chapter describes how the data from the Variation Study reported on in this work were entered into the computer program used for the statistical analysis (SPSS, *Statistical Package for the Social Sciences* as it is usually called, although it may officially have a different name now). The difference between types of variables is briefly explained (*nominal, ordinal, scale*), as well as the meaning of some key concepts in statistics, such as statistical significance and correlation. This is followed by a brief demonstration and explanation of the main methods used in this study to calculate the results and present the data.