

Valresultat riksdagen 2014

I lämningsplanerna i matematik betonas att eleverna ska få möjlighet att använda digitala verktyg. Ett exempel från kursen Matematik 2 är

- **Statistiska metoder för rapportering av observationer och mätdata från undersökningar inklusive regressionsanalys**
- **Metoder för beräkning av olika lägesmått och spridningsmått inklusive standardavvikelse.**

Matematisk nivå

För den inledande undersökningen behövs kunskaper motsvarande matematik kurs 2. Viss erfarenhet av TI-Nspire är nödvändig. Ett viktigt begrepp i fortsättningen är statistisk variabel.

I denna aktivitet tittar vi på en stor datamängd med mer än 7000 data i kalkylarkets celler. Det handlar om slutresultatet från riksdagsvalet 2014. I princip finns alla anvisningar i Nspire-dokumentet *valresultat_riksdagen_2014*.

Alla sidor i detta dokument visar exempel på olika statistiska undersökningar man kan göra. Om ni vill träna och göra egna undersökningar föreslår vi att ni först sparar undan originaldokumentet och sedan kopierar själva kalkylarket (sid 3 i dokumentet) från sidsorteraren till vänster på skärmen och sedan lägger in kalkylarket i ett nytt träningsdokument.

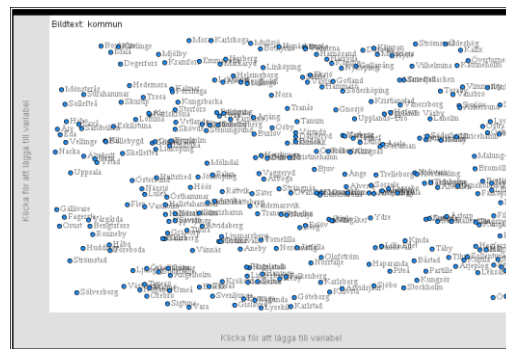
Problem 1

A kommun	B län	C m_antal	D m_proc	E c_...	F c_proc
1 Karlshamn	Blekinge län	3923	18.69	1015	4.84
2 Karlskrona	Blekinge län	8848	20.71	2493	5.83
3 Olofström	Blekinge län	1123	13.41	473	5.65
4 Ronneby	Blekinge län	3335	17.84	1289	6.89
5 Sölvesborg	Blekinge län	2595	22.69	522	4.56
6 Avesta	Dalarnas län	2264	15.45	825	5.63
7 Borlänge	Dalarnas län	5713	18	1615	5.09
8 Falun	Dalarnas län	8871	23.19	3058	7.99

Har här vi vårt stora kalkylark. Bläddra gärna genom kolumnerna så du ser vad det är för data. Vi har antal och procentandelar för alla riksdagspartier och för fi. Dessutom finns uppgifter om antal röster, röstberättigade och valjeltagande (i %).

Så här ser första sidan ut. 290 rader och 26 kolumner. Vi utnyttjar ofta programmets möjligheter att ha delade sidor. Här har vi lagt en anteckningssida under kalkylarket med förklaringar.

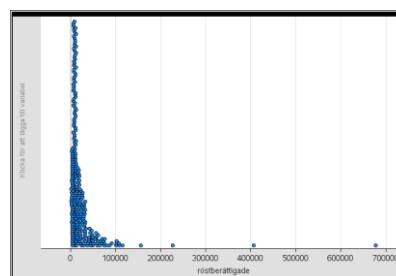
Om man nu infogar en *Data&Statistik*sida ser det först ut som ovan



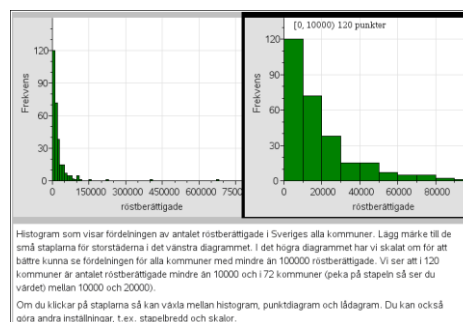
Om man nu högerklickar under diagrammet kan man välja variabel. Se en del av variabelistan nedan.

- 1: antal_röster
- 2: c_antal
- 3: c_proc
- 4: fi_antal
- 5: fi_proc
- 6: fp_antal
- 7: fp_proc
- 8: giltiga_röster
- 9: kd_antal
- A: kd_proc
- B: kommun
- C: län
- D: m_antal
- E: m_proc
- F: mp_antal

Sid 4: När man valt en variabel försöker programmet alltid rita ett punktdiagram. Eftersom vi har 290 datarader blir ett sådant diagram lite otydligt. Se diagrammen nedan. Histogram passar bättre. Diskutera med eleverna vad histogrammet visar och vad frekvens betyder här.

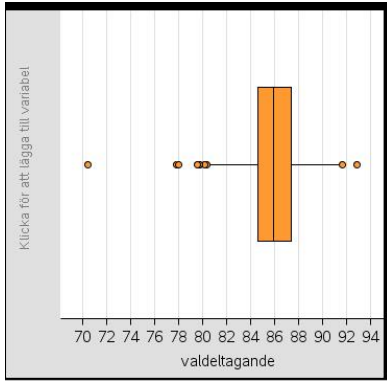


När man ska byta diagramtyp för sina data så högerklickar man i diagrammet.



Sid 5: Här har vi valt variabler på båda axlarna. Observera att vi på den vågräta axeln har en icke numerisk variabel.

Sid 6: Här kan man jämföra vilken information man kan få ut av lådagram och histogram. Diskutera vad punkterna utanför lådagramets låda betyder. Man kan klicka på olika delar och se numeriska värden.



Sid 7: Här kommer en mycket viktig sida i dokumentet.

Man ska vara försiktig och eftertänksam när man räknar med data som är procenttal.

Om vi t.ex. vill beräkna resultatet (röstandel i %) för socialdemokraterna och moderaterna för **hela landet** kan man **INTE** göra beräkningen nedan. Den ger visserligen ett medelvärde för resultaten i alla kommuner men då har man inte tagit hänsyn till antalet röstande i de olika kommunerna, d.v.s. kommunernas storlek. I Stockholms kommun röstade t.ex. ca 581 000 personer och i den minsta kommunen befolkningsmässigt, Sorsele i Västerbottens län, 1579 personer.

$\text{mean}(s_proc) \rightarrow 35.0293793103$ $\text{mean}(m_proc) \rightarrow 20.3014137931$ **FEL**

För att **korrekt** beräkna resultatet för s och m i hela landet så finns det två sätt att utföra beräkningarna:

$$\frac{\text{sum}(s_antal)}{\text{sum}(\text{giltiga_röster})} \cdot 100 = 31.014817607$$

$$\frac{\text{sum}(m_antal)}{\text{sum}(\text{giltiga_röster})} \cdot 100 = 23.3250416869$$

$\text{mean}(s_proc, \text{giltiga_röster}) \rightarrow 31.01495917$ $\text{mean}(m_proc, \text{giltiga_röster}) \rightarrow 23.3252177051$

Ovan har vi viktat procentalen mot antalet giltiga röster (frekvensen). Ger också ett korrekt värde.

Valdeltagandet för hela Sverige beräknas så här: $\frac{\text{sum}(\text{antal_röster})}{\text{sum}(\text{röstberättigade})} \cdot 100 = 85.806893782$

Resultaten ovan stämmer med den officiellt redovisade statistiken från Valmyndigheten.

Det handlar om att man ska vara försiktig när man gör beräkningar på procenttal. Om vi t.ex. vill beräkna resultatet (röstandel i %) för socialdemokraterna och moderaterna för hela landet kan man **INTE** beräkna medelvärdet av alla värden i kolumnerna med procentandelar för de två partierna. Den ger visserligen ett medelvärde för resultaten i alla kommuner men då har man inte tagit hänsyn till antalet röstande i de olika kommunerna, d.v.s. kommunernas storlek. En direkt(och felaktig) beräkning av värdet gav socialdemokraterna ca 35 %. Det korrekta värdet, t.ex. med beräkningen

$$\frac{\text{sum}(s_antal)}{\text{sum}(\text{giltiga_röster})} \cdot 100$$

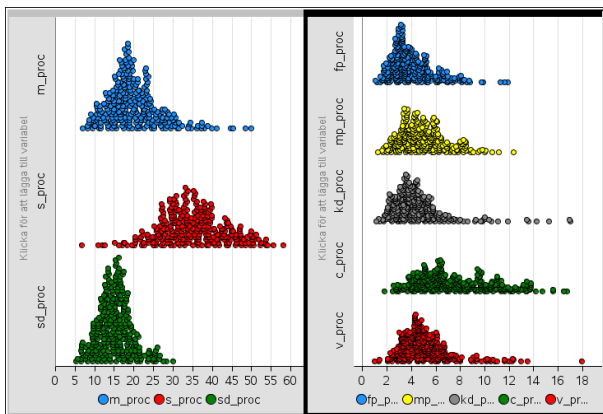
gav det riktiga värdet ca 30,1 %.

Alltså, genom att *infoga rutor för matematiska beräkningar* kan man göra beräkningar på alla statistiska variabler i kalkylarket. Om du trycker på **Ctrl m** infogas en beräkningsruta och du får tillgång till verktyg från alla applikationer. Gå till verktygslådan och välj Beräkningar och välj Statistik.

Man kan, som vi tidigare påpekat, använda anteckningssidor för många saker. Man kan skriva kommentarer till andra sidor i dokumentet och man kan göra beräkningar allmänt och på variabler i dokumentet. Dessutom kan man få korrekt formaterade formler i dokumentet och till hjälp har man olika mallar.

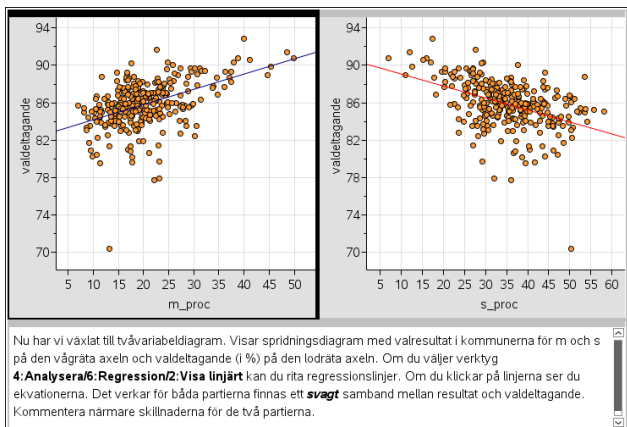
Här är ett exempel till på beräkningar man kan göra. Först har vi gjort en "manuell" beräkning av standardavvikelsen för Centerpartiets resultat i alla kommuner i kalkylarket. Sedan har vi på en anteckningssida infogat *envariabelstatistik* för variabeln *c_proc*. Vi får samma resultat.

A	B	C	D	E
c_proc	avvikelse	avv_kvadrat		
=	=c_proc-mean('c_proc)	=avvikelse^2		
1	4.84	-2.76672	7.65476	
2	5.83	-1.77672	3.15675	3.07814
3	5.65	-1.95672	3.82877	
4	6.89	-0.716724	0.513693	
5	4.56	-3.04672	9.28253	
6	5.63	-1.97672	3.90744	
7	5.09	-2.51672	6.3339	
8	7.99	0.383276	0.1469	
9	11.61	4.00328	16.0262	
10	9.27	1.66278	2.76649	
11				
12				
13				
14				
15				
16				
17				
18				
19				
20				
21				
22				
23				
24				
25				
26				
27				
28				
29				
30				
31				
32				
33				
34				
35				
36				
37				
38				
39				
40				
41				
42				
43				
44				
45				
46				
47				
48				
49				
50				
51				
52				
53				
54				
55				
56				
57				
58				
59				
60				
61				
62				
63				
64				
65				
66				
67				
68				
69				
70				
71				
72				
73				
74				
75				
76				
77				
78				
79				
80				
81				
82				
83				
84				
85				
86				
87				
88				
89				
90				
91				
92				
93				
94				
95				
96				
97				
98				
99				
100				
101				
102				
103				
104				
105				
106				
107				
108				
109				
110				
111				
112				
113				
114				
115				
116				
117				
118				
119				
120				
121				
122				
123				
124				
125				
126				
127				
128				
129				
130				
131				
132				
133				
134				
135				
136				
137				
138				
139				
140				
141				
142				
143				
144				
145				
146				
147				
148				
149				
150				
151				
152				
153				
154				
155				
156				
157				
158				
159				
160				
161				
162				
163				
164				
165				
166				
167				
168				
169				
170				
171				
172				
173				
174				
175				
176				
177				
178				
179				
180				
181				
182				
183				
184				
185				
186				
187				
188				
189				
190				
191				
192				
193				
194				
195				
196				
197				
198				
199				
200				
201				
202				
203				
204				
205				
206				
207				
208				
209				
210				
211				
212				
213				
214				
215				
216				
217				
218				
219				
220				
221				
222				
223				
224				
225				
226				
227				
228				
229				
230				
231				
232				
233				
234				
235				
236				
237				
238				
239				
240				
241				
242				
243				
244				
245				
246				
247				
248				
249				
250				
251				
252				
253				
254				
255				
256				
257				
258				
259				
260				
261				
262				
263				
264				
265				
266				
267				
268				
269				
270				
271				
272				
273				
274				
275				
276				
277				
278				
279				
280				
281				
282				
283				
284				
285				
286				
287				
288				
289				
290				
291				
292				
293				
294				
295				
296				
297				
298				
299				
300				
301				
302				
303				
304				
305				
306				
307				
308				
309				
310				
311				
312				
313				
314				
315				
316				
317				
318				
319				
320				
321				
322				
323				
324				
325				
326				
327				
328				
329				
330				
331				
332				
333				
334				
335				
336				
337				
338				
339				
340				
341				
342				
343				
344				
345				
346				
347				
348				
349				
350				
351				
352				
353				
354				
355				
356				
357				
358				

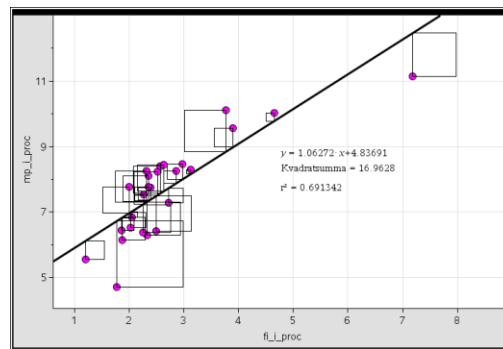
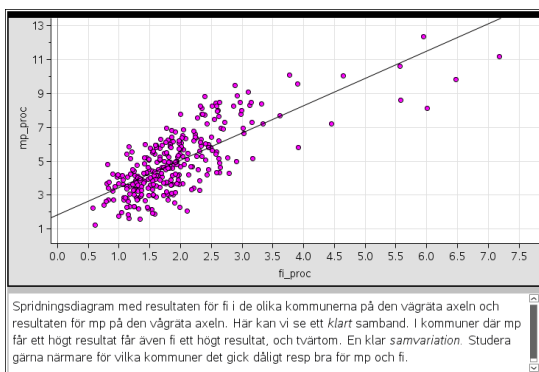


Ovan visar vi fördelningen för alla riksdagspartierna. Viket parti har en fördelning som avviker från de andra?

Sid 12-13: Här har vi de första tvåvariabeldiagrammen. Vi visar nedan ett spridningsdiagram med valresultat för s och m i alla kommuner. Om man klickar på de regressionslinjerna ser man linjernas ekvationer och r^2 (kvadraten på korrelationskoefficienten).

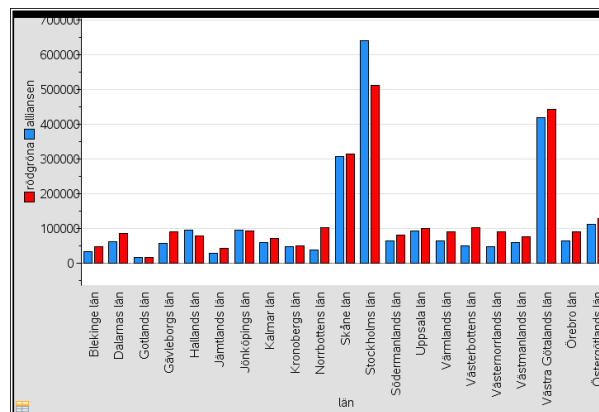


Här kan man diskutera samvariationen som verkar finnas mellan resultaten för mp och fi. Korrelationskoefficienten är ung 0,75. I diagrammet i nästa spalt har vi ritat s.k. *residualkvadrater*, som med kvadrater visar hur nära datapunkterna ligger från regressionslinjen. Regressionslinjen beräknas så att summan av kvadraternas area blir så liten som möjligt. Vi har här valt att titta på resultaten för kommunerna i Stockholms län därför att diagrammet annars skulle bli för plottrigt med alla kvadrater.



Sid 14-15: Alla kommentarer ligger i anteckningsfält på sidorna.

Sid 15-16: Man kan även rita diagram på *sammanfattningsdata*. Först har vi på sid 15 skapat två nya variabler för de bägge blocken. Se formlerna i formelfälten. På sid 16 har vi först valt variabeln län på vågräta axeln och sedan på lodräta axeln valt alternativet "Lägg till Y-sammanfattningslista" och sedan valt de bägge variablerna. Diagrammet visar då totala antalet röster för de bägge blocken i alla län. Om man klickar på staplarna så ser man även andelarna i procent.



Problem 2

Sid 1-4: Kan man undersöka den svenska politiska opinionen genom att titta närmare på en enda kommun? Ja, kanske kan man det. Vi har nedan gjort en undersökning där vi beräknat differensen mellan partiernas valresultat i de olika kommunerna och riksresultatet. Vi har sedan summerat dessa differenser för alla partier. Även fi, som inte kom in i riksdagen, är med i denna undersökning.

Vi har då skapat en ny variabel "Närmaste":

$$\text{Närmaste} := \text{abs}(m_proc - 23.32) + \text{abs}(c_proc - 5.42) + \dots + \text{abs}(sd_proc - 12.86) + \text{abs}(fi_proc - 3.12)$$

där vi har beräknat avvikelsen från riksgenomsnittet för varje parti och sedan har vi summerat och fått ett nytt mått. Vilka kommuner liknar mest hela Sverige?

Problem 3

I Problem 3 har lagt till en variabel, *Kommunindelning*, för att kunna jämföra valresultat efter typ av kommun. Tidigare i dokumentet tittade vi bara på regionala skillnader. Nu tittar vi istället hur man röstade man i glesbygdskommuner jämfört med storstäder osv. Indelningen är från 2013.

