

Valresultat Riksdagen 2018

I ämnesplanerna 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 med digitala verktyg.
- Metoder för beräkning av olika lägesmått och spridningsmått inklusive standardavvikelse, med digitala verktyg.

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 2018. I princip finns alla anvisningar för eleverna i TI-Nspire dokumentet *valresultat-riksdagen_2018*.

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.

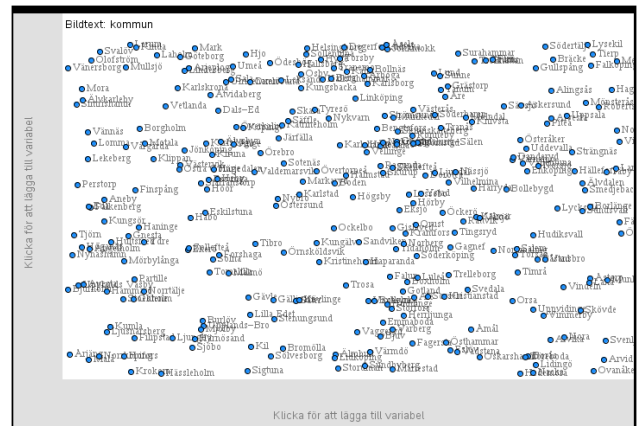
Problem 1

Beräkningar på alla röstande och länen

län	kommun	m2018	m2018_proc	c2018
1 Blekinge län	Karlshamn	3632.00	16.94	1392.0
2 Blekinge län	Karlskrona	8134.00	18.35	3177.0
3 Blekinge län	Olofström	1002.00	12.02	542.0
4 Blekinge län	Ronneby	3024.00	16.23	1313.0
5 Blekinge län	Sölvesborg	2246.00	19.05	670.0
6 Dalarnas län	Avesta	2298.00	15.77	983.0
7 Dalarnas län	Borlänge	5547.00	17.09	2411.0
8 Dalarnas län	Falun	7778.00	19.71	4669.0
9 Dalarnas län	Gagnef	991.00	14.36	946.0
10 Dalarnas län	Hedemora	1498.00	15.25	979.0
11 Dalarnas län	Leksand	2142.00	19.57	1408.0

Så här ser första sidan ut. 290 rader och 24 kolumner. Att data alltid visas med två decimaler beror på att vi i kolumner med procentvärden vill ha just två decimaler. Gå gärna in i och ändra under Arkiv och sedan Inställningar/Dokumentinställningar och ändra för att se effekten.

Om man nu infogar en *Data&Statistik*-sida ser det först ut som nedan. Vi får alla namn på kommunerna utslängda i diagramfältet.



Om man nu högerklickar under diagrammet kan man välja variabel. Se en del av variabellistan nedan. Det finns ganska många variabler i detta problem inom dokumentet. Variablerna är ordnade i bokstavsordning.

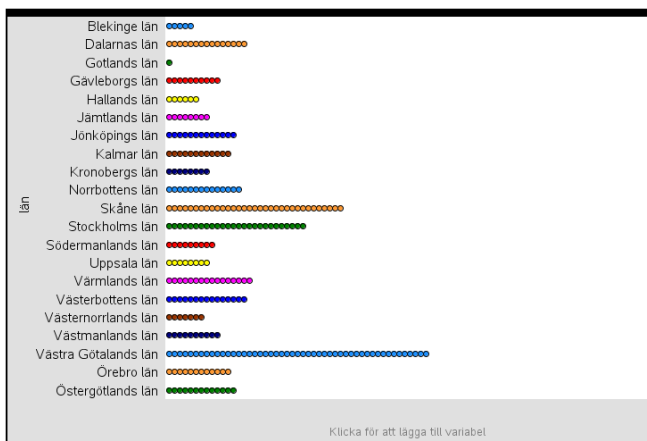
- 1: c2018
- 2: c2018_proc
- 3: fi2018
- 4: fi2018_proc
- 5: giltiga2018
- 6: kd2018
- 7: kd2018_proc
- 8: kommun
- 9: l2018
- A: l2018_proc
- B: län
- C: m2018
- D: m2018_proc
- E: mp2018
- F: mp2018_proc
- G: röstande2018
- H: röstberätt2018
- I: s2018
- J: s2018_proc
- K: sd2018
- L: sd2018_proc

Sid 4: Som första diagram vill vi nu ha en översikt över kommunerna och till vilka län de tillhör. Vi kan naturligtvis titta i själva kalkylarket men nu kan man få en snabb översikt. Klicka till vänster om den lodräta axeln och välj variabeln län.

Vi får ett punktdiagram där varje "prick" betyder en kommun. Vi ser att Västra Götalands län har flest kommuner.

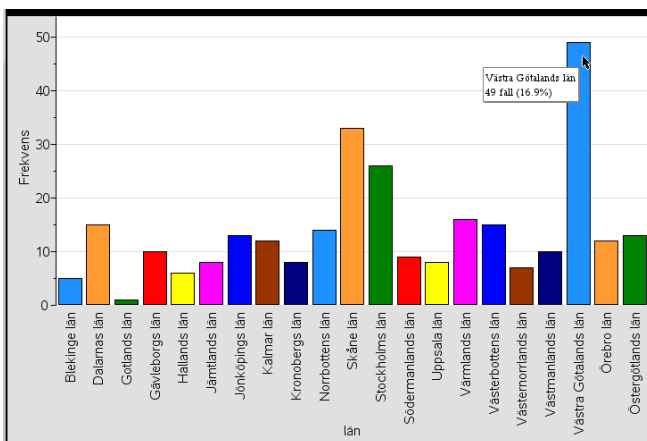
Om man högerklickar i diagrammet så får man möjligheter att byta till en annan diagramform. Vi väljer Stapeldiagram. Vi har ju här en icke numerisk variabel. Se diagram på nästa sida.

I TI-Nspire-dokumentet har vi delat sidan och lagt in en karta över de olika länen.

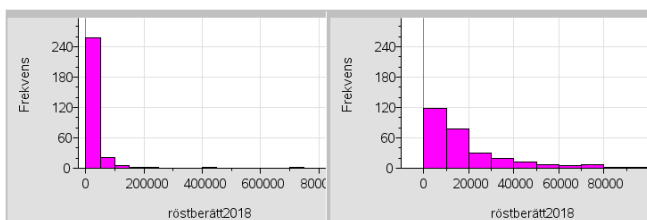


Om man rör markören över staplarna så ser man värden, både absoluta värden och relativa värden i procent.

Vi har här valt att ha staplarna lodrätt men vi får försöka läsa texten lodrätt också.



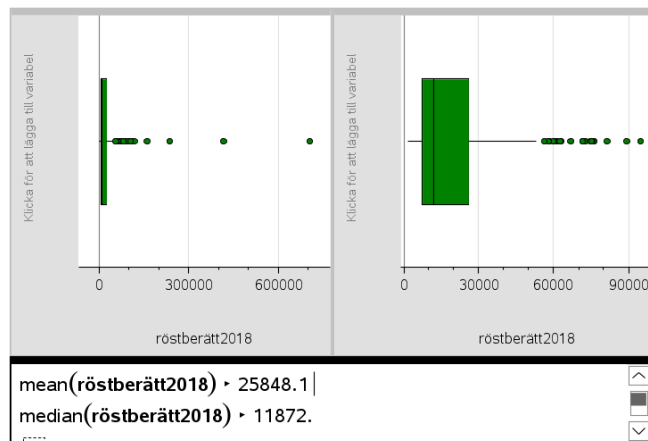
Sid 5:



Histogram som visar fördelningen av antalet röstberättigade i Sveriges alla kommuner (totalt 290 st). Lagg märke till de små knappt synliga staplarna för storstäderna i det vänstra diagram met. I det högra diagrammet har vi skalat om för att bättre kunna se fördelningen för alla kommuner med mindre än 100 000 röstberättigade. Vi ser att i 120 kommuner är antalet röstberättigade mindre än 10000 och i 77 kommuner (peka på stapeln med markörenså ser du värdet) mellan 10000 och 20000).

En sida delad i tre. Förklaringar finns i texten under diagrammen.

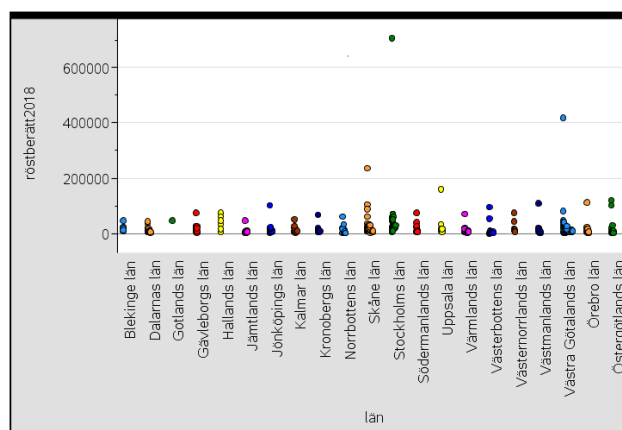
Ett alternativ till histogrammen ovan kan vara lådagram. Se nästa spalt.



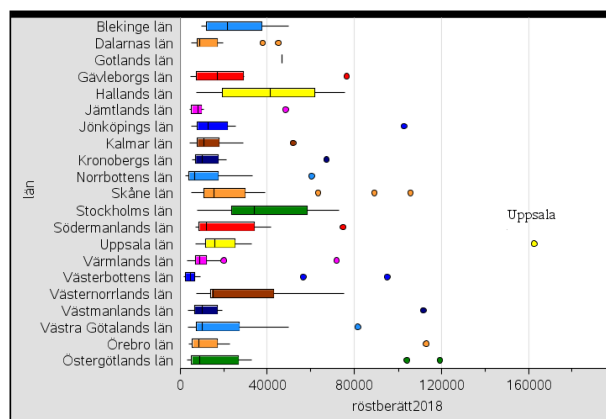
Genom att peka med markören i lådagrammen kan du avläsa under och övre kvartiler och median. Diskutera med eleverna skillnaden mellan median och medelvärde i detta fall. Hur påverkas t.ex. medianen om man skulle plocka bort Stockholm från kommunlistan.

Sid 6:

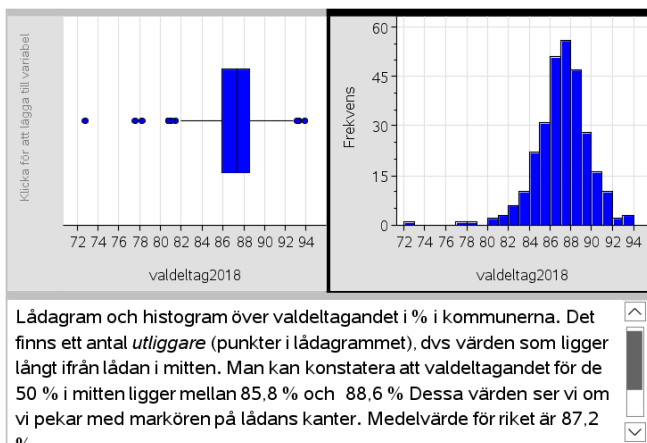
Här har vi delat upp variabeln röstberättigade på län. Även här Sticker storstäderna ut. De är lätta att identifiera. För län med många kommuner så trängs prickarna på ett litet utrymme.



Här har vi istället plottat lådagram och valt ett mindre intervall för antal röstberättigade. Lådagram ger större överskådlighet men man förlorar i detaljer.



Sid 7: Här har vi nu ännu ett dubbel-diagram där man kan jämföra information från två diagramtyper.



Sid 8-9: Detta är tvåviktiga sidor. Man kan inte göra beräkningar på statistiska mått hur som helst.

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, dvs kommunernas storlek. I Stockholms kommun röstade t.ex. ca 615 000 personer och i den minsta kommunen befolkningsmässigt, Sorsele i Västerbottens län, 1553 personer.

mean(s2018_proc) > 30.26 FEL
mean(m2018_proc) > 17.46 FEL

För att *korrekt* beräkna resultatet för s och m i hela landet så finns det två sätt att utföra beräkningarna. Se nästa sida!

1) $\frac{\text{sum}(s2018)}{\text{sum}(\text{giltiga2018})} \cdot 100 > 28.26$ $\frac{\text{sum}(m2018)}{\text{sum}(\text{giltiga2018})} \cdot 100 > 19.84$

2) **mean(s2018_proc, giltiga2018) > 28.26**
mean(m2018_proc, giltiga2018) > 19.84

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

Valdeltagandet för hela Sverige beräknas så här:
 $\frac{\text{sum}(\text{röstande2018})}{\text{sum}(\text{röstberätt2018})} \cdot 100 > 87.18$

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

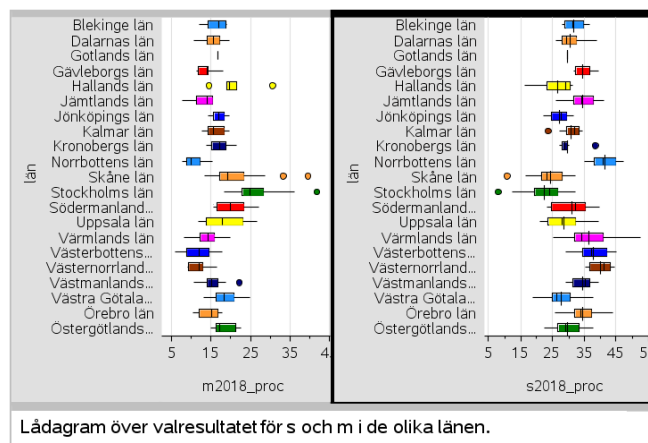
direkt (och felaktig) beräkning av värdet gav socialdemokraterna ca 30,26 %.

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.

Beräkningar på partiernas resultat

Sid 10: Lådagram över valresultatet för s och m i de olika länen. Punkterna i diagrammen är *utliggare*, dvs värden som är väldigt mycket större eller mindre än resten av värdena. I vilka län är spridningen stor och vilka län liknar resultatet för hela landet?



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

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 *c2018_proc*. Vi får samma resultat.

Se nästa sida!

sultat_c	avvikelse	avv_kvad		
=	=c2018_proc-mean(c2018_proc)	=avvikelse^2		
1	6.49	-2.82	7.95	2.96
2	7.17	-2.14	4.60	
3	6.50	-2.81	7.91	
4	7.05	-2.27	5.14	
5	5.68	-3.63	13.17	
6	6.75	-2.57	6.59	
7	7.43	-1.89	3.56	
8	11.83	2.52	6.34	
9	13.71	4.39	19.31	
10	9.97	0.65	0.43	
sum('avv_kvad')				290

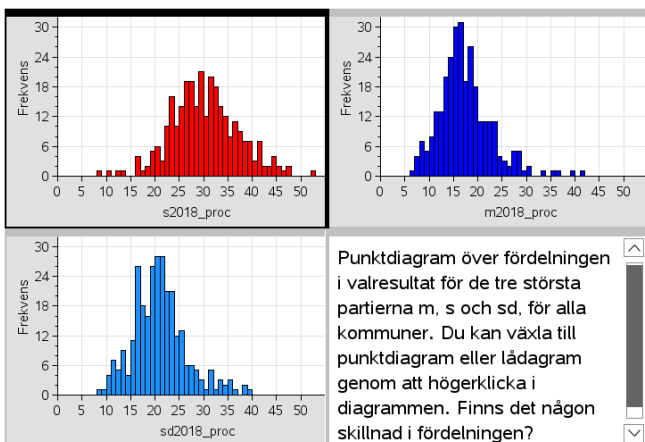
Observera att Centerpartiets medelvärde här är ett direkt medelvärde beräknat på resultaten i samtliga kommuner. Riksmedelvärdet var ju 8,61 %. Diskutera med eleverna varför man kan få denna differens.

OneVar c2018_proc,1: stat.results	"Envariabelstatistik"
" \bar{x} "	9.31
" Σx "	2700.85
" Σx^2 "	27700.15
" $s_x := s_{n-1}x$ "	2.97
" $\sigma_x := \sigma_n x$ "	2.96
"n"	290.00
"MinX"	3.28
" $Q_1 X$ "	7.06
"MedianX"	8.91
" $Q_3 X$ "	11.34
"MaxX"	21.61
" $SSX := \Sigma(x-\bar{x})^2$ "	2546.37

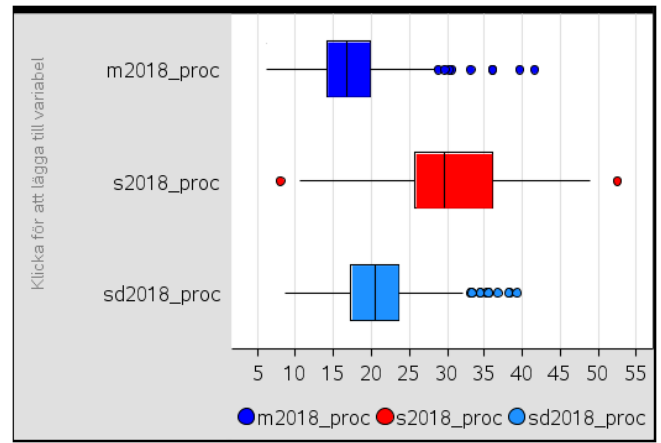
Sid 11:

Här kommer nu histogram som visar fördelningen av andel (i %) för de tre största partierna. Vi har de absoluta frekvensen på den lodräta axeln.

Klassbredden är 1. Det är samma skala i diagrammen för att det ska bli lättare att jämföra.



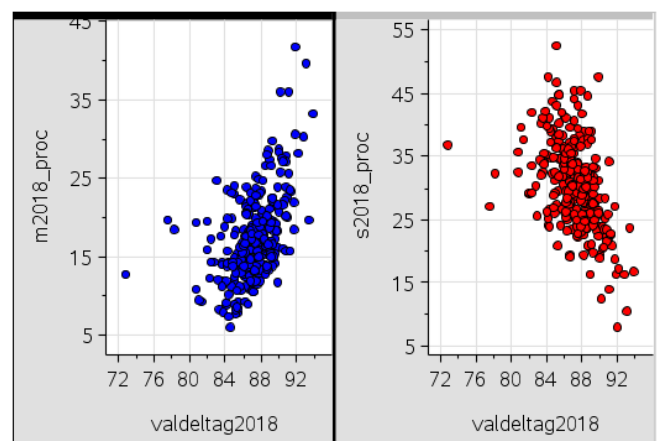
Man kan placera diagrammen i ett diagram om man väljer att plotta lådagram. Lådagram passar mycket bra om man vill göra jämförelser. Se nästa spalt.



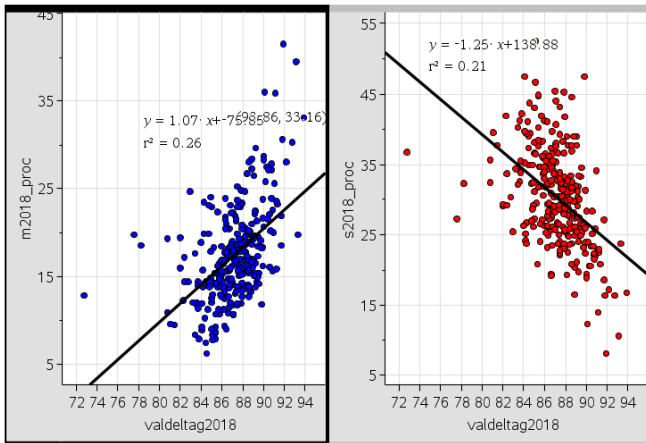
Hur ligger medianerna i förhållande till medelvärdet? För vilket parti är spridningen störst? Vi tittar på detta!

stDevPop(m2018_proc)	▶ 5.37
stDevPop(s2018_proc)	▶ 7.34
stDevPop(sd2018_proc)	▶ 5.42

Sid 12-13: På sidan finns två plottningar som visar spridningsdiagram. Vi har valdeltagande på den vågräta axeln och andelen röster för m och s på den lodräta. Kan man se någon tendens i diagrammen? Hur är resultaten samlade? Kan det finnas bakomliggande faktorer som förklarar sambanden?

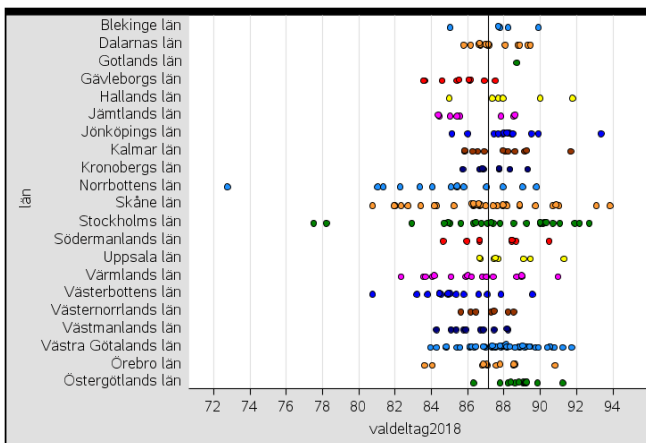


Här kommer man in på begreppen regression och korrelation? Visa eleverna hur man kan analysera sambandet mellan variablerna.

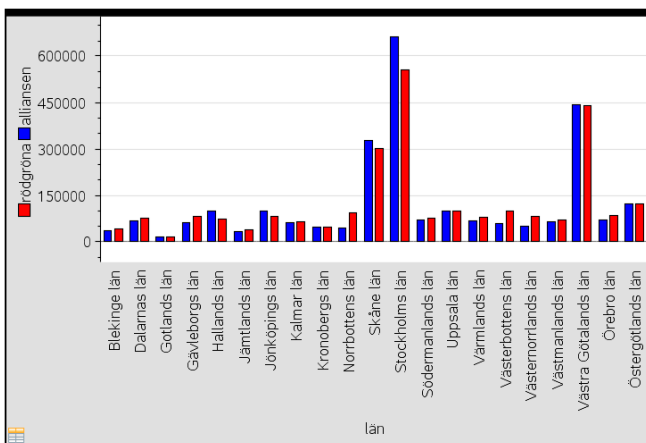


Här har vi lagt in regressionslinjer och beräknat r^2 . Det verkar finnas ett svagt samband men vi kan inte med säkerhet uttala oss om ev. bakomliggande faktorer som förklarar sambandet.

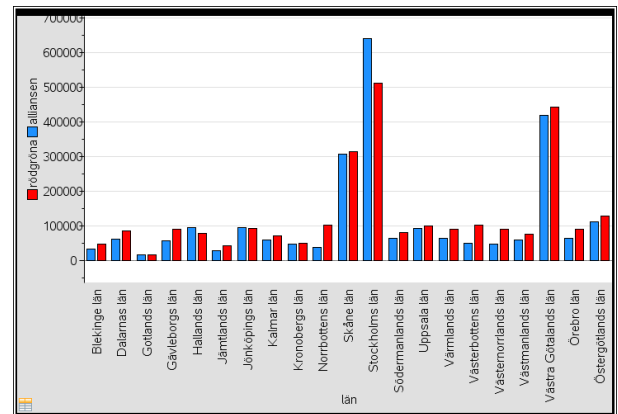
Sid 14: Vi har även undersökt valdeltagandet i de olika länen. Vi kan tydligt se att vi vissa län är spridningen mycket stor medan andra län har en betydligt mindre spridning mellan kommunerna. Riksgenomsnittet är 87,2 %. Vi har lagt in en linje som visar detta.



På sidorna 15-16 finns det också med en beräkning och presentation av andelar för de rödgröna partierna respektive allianspartierna i de olika länen.



Motsvarande siffror för 2014 års riksdagsval. Lite svårt att jämföra eftersom antalet röstande totalt var större i 2018 års val. Ska man jämföra bör man arbeta med andelar.



Problem 2: Vilken kommun liknar mest hela landet?

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 rikslösultatet. 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":

$$\text{Närmaste} = \text{abs}(m-19.84) + \text{abs}(c-8.61) + \dots + \text{abs}(sd-17.53) + \text{abs}(fi-0.46)$$

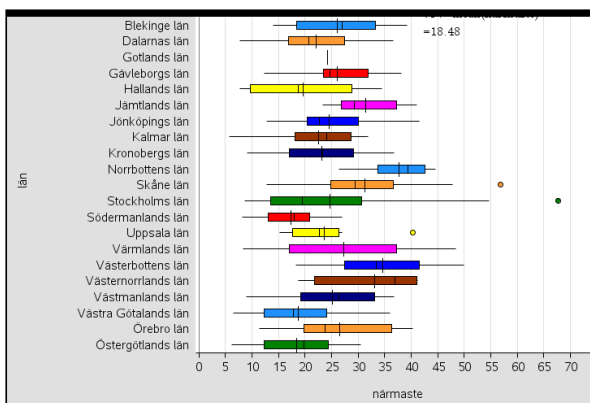
där vi har beräknat absolutbeloppet av avvikelserna från riksgenomsnittet för varje parti och sedan har vi summerat värdet för alla partier och fått ett nytt mått. Vilka kommuner liknar mest hela Sverige?

I bilden har vi markerat värdena för toptre-kommunerna.

	G s	H v	I mp	J sd	K fi	L närmaste	M
1	30.35	6.71	4.17	17.09	0.40	5.66	
2	28.34	8.13	4.04	19.61	0.47	6.20	
3	29.53	7.02	3.66	18.37	0.34	6.51	
4	28.81	5.63	3.40	16.94	0.50		7.68
5	30.43	5.52	3.69	17.71	0.41		7.69
6	28.02	8.18	4.88	15.30	0.51		7.74
7	31.96	6.44	4.48	16.11	0.41		8.21
8	31.73	7.88	4.84	14.05	0.50		8.27
9	26.97	8.01	3.92	18.04	0.37		8.57
10	29.12	7.15	3.40	17.51	0.37		8.96
11	30.01	7.19	4.16	15.38	0.32		9.09

Be eleverna göra ett diagram som visar fördelningen för variabeln *närmaste* på alla kommuner. Så här ser det ut

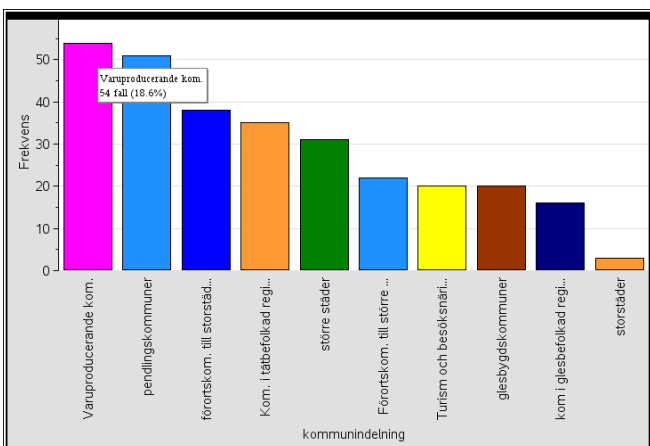
Så här ser det ut på länsnivå.



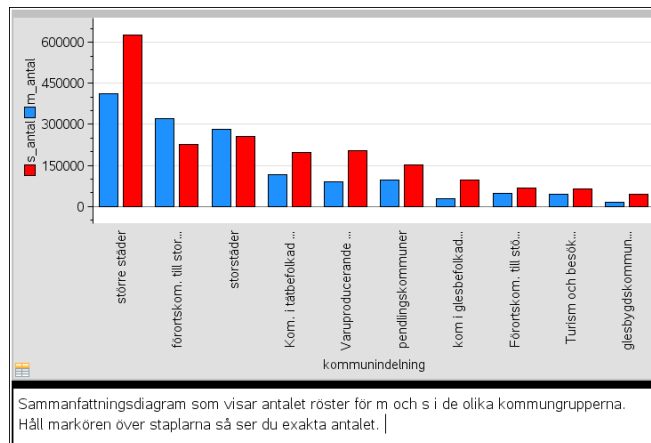
Problem 3: Hur ser opinionen ut i landsbygdskommunerna?

I Problem 3 har lagt till en variabel, *kommuntyp*, 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 i landsbygdskommuner jämfört med storstäder osv. Indelningen är från 2017 (Sveriges Kommuner och Landsting).

län	kommun	kommuntyp
Blekinge län	Karlshamn	Mindre stad/tätort
Blekinge län	Karlskrona	Mindre stad/tätort
Blekinge län	Olofström	Pendlingskommun nära mindre ...
Blekinge län	Ronneby	Pendlingskommun nära mindre ...
Blekinge län	Sölvesborg	Pendlingskommun nära mindre ...
Dalarnas län	Avesta	Mindre stad/tätort
Dalarnas län	Borlänge	Större stad
Dalarnas län	Falun	Mindre stad/tätort
Dalarnas län	Gagnef	Pendlingskommun nära större s...
Dalarnas län	Hedemora	Pendlingskommun nära mindre ...
Dalarnas län	Leksand	Landsbygdskommun med besök...



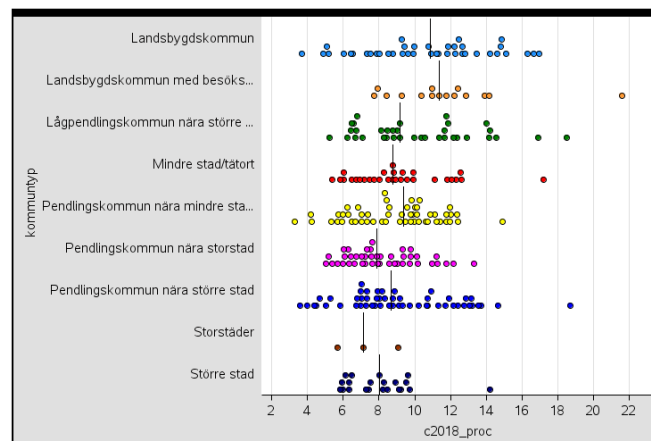
Kommunerna uppdelade efter kommuntyp



Antal röstberättigade i de olika kommuntyperna

I TI-Nspire-dokumentet har vi sedan tre diagram som visar resultat för de tre största partierna uppdelat på kommuntyp.

Här har vi gjort samma sak för Centerpartiet. Vi har också lagt in lodräta linjer som visar medianen inom varje kommuntyp. Sådana lodräta linjer kan man göra från Analysera i menyn och sedan verktyget Rita värde. Man skriver `variabelnamn:=median(c2018_proc)`.



Titta också på de beräkningarna nedan. Finns mycket att diskutera här. Varför avviker t.ex. resultatet

$$\text{mean}(c2018_proc) \approx 9.31$$

från det korrekta värdet 8,61 % för hela landet.

$$\begin{aligned} \text{median}(c2018_proc) &\triangleright 8.91 \\ \text{mean}(c2018_proc) &\triangleright 9.31 \\ \text{mean}(c2018_proc, \text{giltiga2018}) &\triangleright 8.61 \\ \frac{\text{sum}(c2018)}{\text{sum}(\text{giltiga2018})} \cdot 100 &\triangleright 8.61 \end{aligned}$$