

Videregående
Statistik og
KeHaTools
Kapitel 1:
Mere om
beskrivende statistik

Oversigt

- Kapitel 1.1 Skævhed og kurtosis
- Kapitel 1.2 Median, kvartilsæt og IQR
- Kapitel 1.3 Boxplot
- Kapitel 1.4 Andre mål
- Kapitel 1.5 Fraktiler
- Eksempel 1.4 Ugrupperede observationer
- Eksempel 1.5 Grupperede observationer

Data

FS_Data.xlsx - Microsoft Excel

Filer Startside Indsæt Sidelayout Formler Data Gennemse Vis Udvikler Tilføjelsesprogrammer

Sæt ind Skrifttype Justering Tal Typografier Celler Redigering

G19 fx

	A	B	C	D	E	F	G	H	I	J
1	Health Care			Pengeinstitutter						
2		K/I			K/I					
3	ALK Abello	2,19		Skjern Bank	0,08					
4	Ambu	4,34		Totalbank	0,1					
5	Bavarian Nord	1,25		Saprekassen Fae	0,14					
6	BioPorto	13,24		DiBa Bank	0,16					
7	ChemoMetec	1,32		Østjysk Bank	0,18					
8	Coloplast	3,92		Vestjysk Bank	0,23					
9	Exigon	2,47		Salling Bank	0,27					
10	Genmab	1,46		Vordingborg Bar	0,3					
11	GN STore Nord	3,09		Sparekassen Hvi	0,36					
12	Lundbeck	1,76		Lollands Bank	0,38					
13	NeuroSearch	0,08		Hvidbjerg Bank	0,39					
14	Novo Nordisk	14,55		Møns Bank	0,44					
15	Novozymes	7,17		Sparekassen Hir	0,46					
16	TopoTarget	0,84		Svendborg Spar	0,57					
17	Veloxis	0,51		Nordfyns Bank	0,59					
18	William Dema	11,65		Djursland Bank	0,61					
19				BankNordik	0,66					
20				Nordjyske Bank	0,66					
21				Kreditbanken	0,76					
22				Vestfyns Bank	0,78					
23				Danske Bank	0,8					
24				Spar Nord Bank	0,89					
25				Sydbank	0,92					

Eksempel 1.2 Eksempel 1.3 Eksempel 1.4 Eksempel 1.1

Klar 100%

Skævhed og kurtosis - I

- Disse størrelser beregnes ved hjælp af indbyggede funktioner i Excel
- Skævhed: =skævhed(...)
- Kurtosis: =topstejl(...)

Skævhed og kurtosis - II

FS_Data.xlsx - Microsoft Excel

	A	B	C	D	E	F	G	H	I	J
1	Health Care			Pengeinstitutter						
2		K/I			K/I					
3	ALK Abello	2,19		Skjern Bank	0,08					
4	Ambu	4,34		Totalbank	0,1					
5	Bavarian Nord	1,25		Saprekassen Fæ	0,14					
6	BioPorto	13,24		DiBa Bank	0,16					
7	ChemoMetec	1,32		Østjyds Bank	0,18					
8	Coloplast	3,92		Vestjysk Bank	0,23					
9	Exigon	2,47		Salling Bank	0,27					
10	Genmab	1,46		Vordingborg Ba	0,3					
11	GN SStore Norc	3,09		Sparekassen Hvi	0,36					
12	Lundbeck	1,76		Lollands Bank	0,38					
13	NeuroSearch	0,08		Hvidbjerg Bank	0,39					
14	Novo Nordisk	14,55		Møns Bank	0,44					
15	Novozymes	7,17		Sparekassen Hir	0,46					
16	TopoTarget	0,84		Svendborg Spar	0,57					
17	Veloxis	0,51		Nordfyns Bank	0,59					
18	William Dema	11,65		Djursland Bank	0,61					
19				BankNordik	0,66					
20				Nordjyske Bank	0,66					
21				Kreditbanken	0,76					
22				Vestfyns Bank	0,78					
23				Danske Bank	0,8					
24				Spar Nord Bank	0,89					
25				Sydbank	0,92					

Median og kvartil - I

- Disse størrelser beregnes ligeledes vha. indbyggede funktioner
- Median: `=median(...)`
- 1. kvartil: `=kvartil(...; 1)`
- 3. kvartil: `=kvartil(...; 3)`
- IQR beregnes som differensen mellem 1. og 3. kvartil

Median og kvartil - II

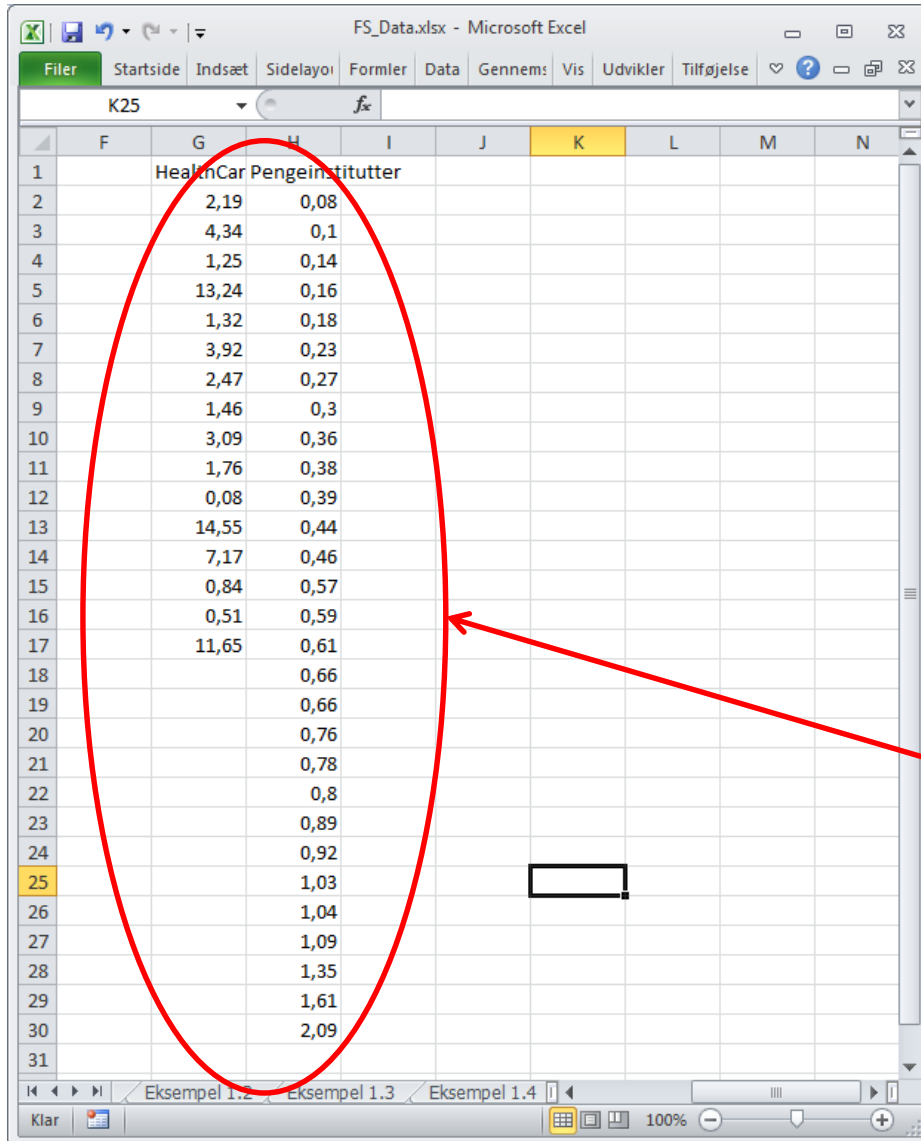
The screenshot shows an Excel spreadsheet titled 'FS_Data.xlsx'. The data is organized in columns B through K. Column B contains 'K/I' values, column C contains bank names, column D contains 'K/I' values, column E contains 'Health Ca' values, column F contains 'Pengeinstitutter' values, column G contains 'skævhed' values, column H contains 'kurtosis' values, column I contains 'median' values, column J contains '1. kvartil' values, and column K contains '3. kvartil' values. The IQR is calculated as the difference between the 3. kvartil and the 1. kvartil. The value 0,59 in cell I10 is highlighted with a black box.

	B	C	D	E	F	G	H	I	J	K
1			Pengeinstitutter							
2	K/I			K/I						
3	2,19		Skjern Bank	0,08			Health Ca	Pengeinstitutter		
4	4,34		Totalbank	0,1		skævhed	1,355531	1,246324		
5	1,25		Saprekassen Fae	0,14		kurtosis	0,503616	1,861173		
6	13,24		DiBa Bank	0,16						
7	1,32		Østjydsk Bank	0,18		median	2,33	0,59		
8	3,92		Vestjysk Bank	0,23		1. kvartil	1,3025	0,3		
9	2,47		Salling Bank	0,27		3. kvartil	5,0475	0,89		
10	1,46		Vordingborg Bar	0,3		IQR	3,745	0,59		
11	3,09		Sparekassen Hvi	0,36						
12	1,76		Lollands Bank	0,38						
13	0,08		Hvidbjerg Bank	0,39						
14	14,55		Møns Bank	0,44						
15	7,17		Sparekassen Hir	0,46						
16	0,84		Svendborg Spar	0,57						
17	0,51		Nordfyns Bank	0,59						
18	11,65		Djursland Bank	0,61						
19			BankNordik	0,66						
20			Nordjyske Bank	0,66						
21			Kreditbanken	0,76						
22			Vestfyns Bank	0,78						
23			Danske Bank	0,8						
24			Spar Nord Bank	0,89						
25			Sydbank	0,92						

Boxplot - I

- Boxplots kan ikke umiddelbart laves i Excel, men sagtens i KeHaTools
- Observationerne skal placeres i kolonner, med én kolonne pr. datasæt, og overskrifter i øverste række

Boxplot - II



KeHaTools 2.3

- 1: Deskriptiv statistik
- 2: Sandsynlighed
- 3: Konfidensinterval
- 4: Hypotesetest
- 5: Goodness-of-fit test
- 6: Linear regression
- 7: Normalfraktildiagram
- 8: Boxplot

Options Retur Annullér

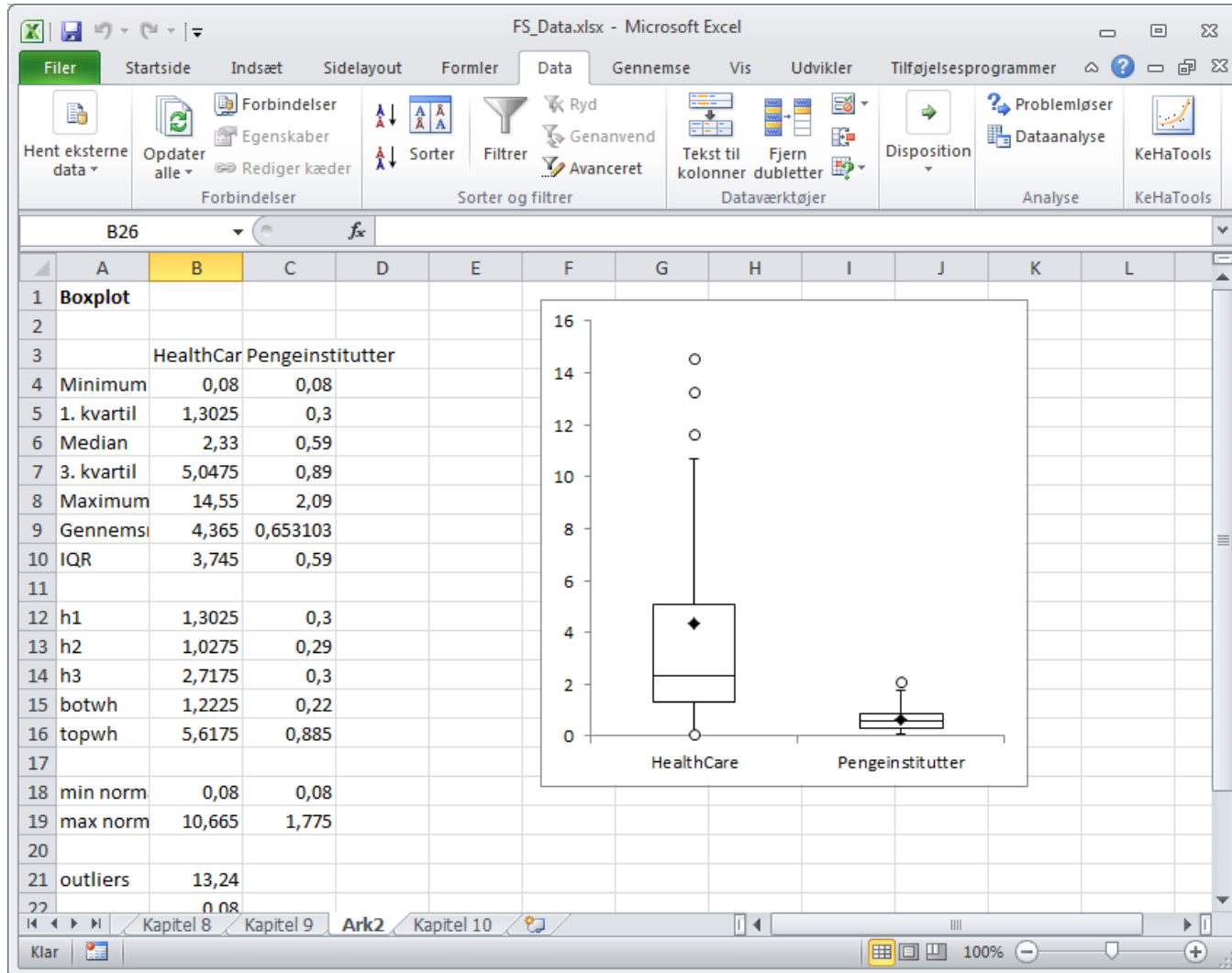
Boxplot

Observationer: 10!:\$G\$1:\$H\$30

Etiketter i øverste række

OK Annullér

Boxplot - III



Andre centralmål

- Typetallet = hyppigst.enkelt(...)
- Trimmet middelværdi
= trimmiddelværdi(...; procent)

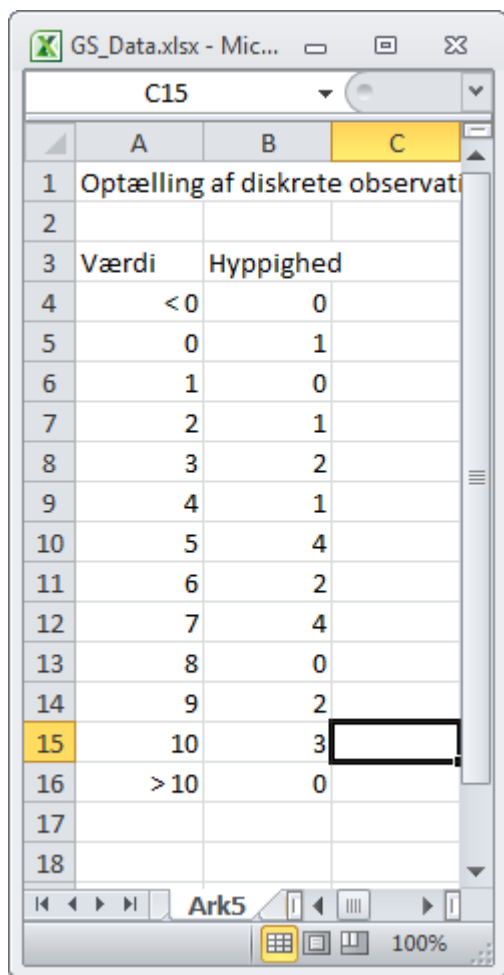
Andre spredningsmål

- Variationsbredde kan ikke umiddelbart beregnes, men findes vha. funktionerne $=\min(\dots)$ og $=\max(\dots)$
- MAD $=\text{MAD}(\dots)$

Fraktiler

- Fraktiler findes i Excel vha. indbyggede funktioner:
- =fraktil (... ; andel)
- =percentil(... ; procenttal)

Eksempel 1.4 - I

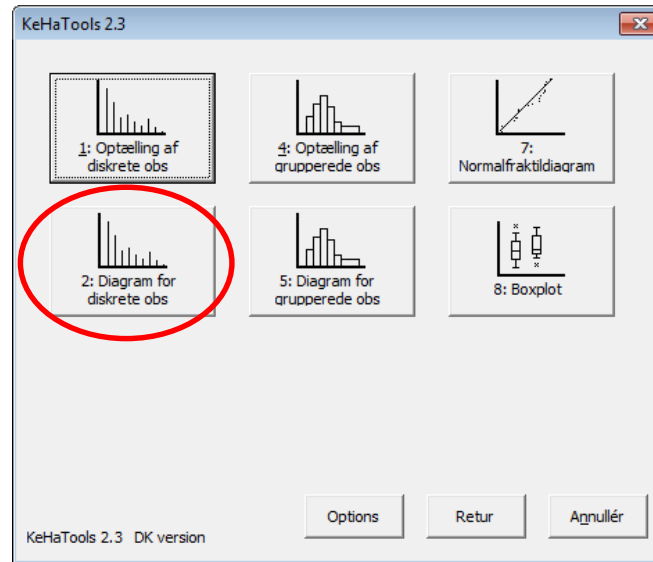
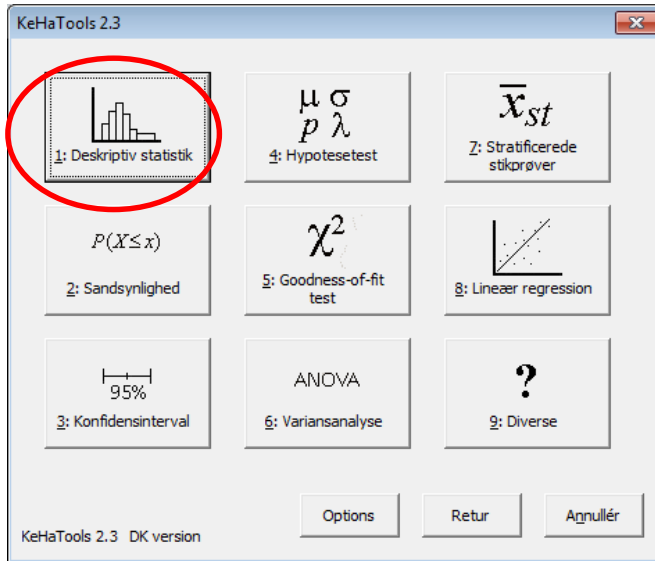


The image shows a screenshot of an Excel spreadsheet window titled "GS_Data.xlsx - Mic...". The active cell is C15. The spreadsheet contains a table with two columns: "Værdi" (Value) in column A and "Hyppighed" (Frequency) in column B. The data is as follows:

	A	B	C
1	Optælling af diskrete observat		
2			
3	Værdi	Hyppighed	
4	< 0	0	
5	0	1	
6	1	0	
7	2	1	
8	3	2	
9	4	1	
10	5	4	
11	6	2	
12	7	4	
13	8	0	
14	9	2	
15	10	3	
16	> 10	0	
17			
18			

The spreadsheet interface includes a formula bar at the top showing "C15", a sheet tab labeled "Ark5", and a status bar at the bottom showing "100%".

Eksempel 1.4 - II



Eksempel 1.4 - III

GS_Data.xlsx - Mic...

C15

	A	B	C
1	Optælling af diskrete observat		
2			
3	Værdi	Hyppighed	
4	< 0	0	
5	0	1	
6	1	0	
7	2	1	
8	3	2	
9	4	1	
10	5	4	
11	6	2	
12	7	4	
13	8	0	
14	9	2	
15	10	3	
16	> 10	0	
17			
18			

Ark5 100%

Deskriptiv statistik for diskrete observationer

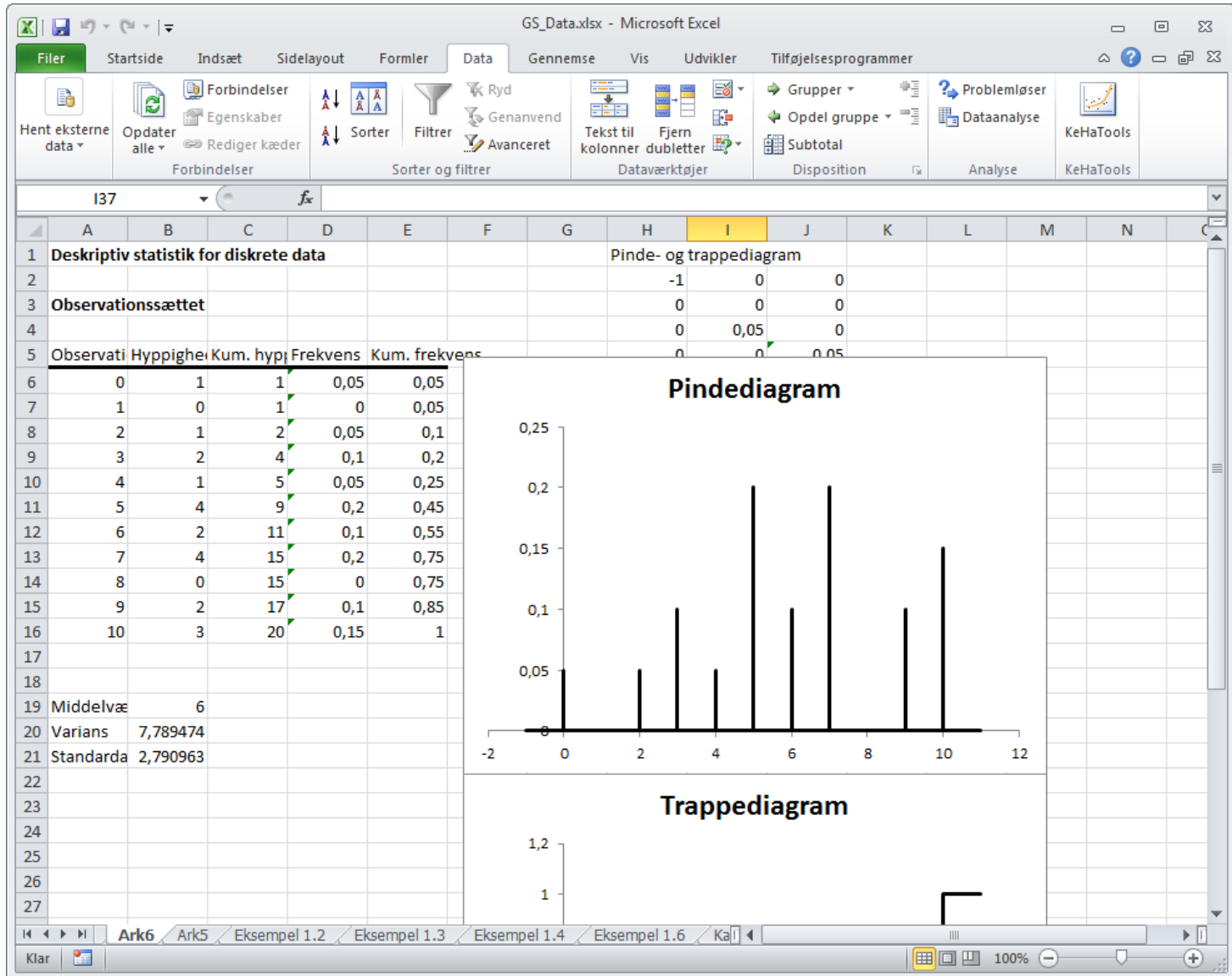
Observerede værdier → Ark5!\$A\$5:\$A\$15

Hyppigheder → Ark5!\$B\$5:\$B\$15

OK

Annuler

Eksempel 1.4 - IV

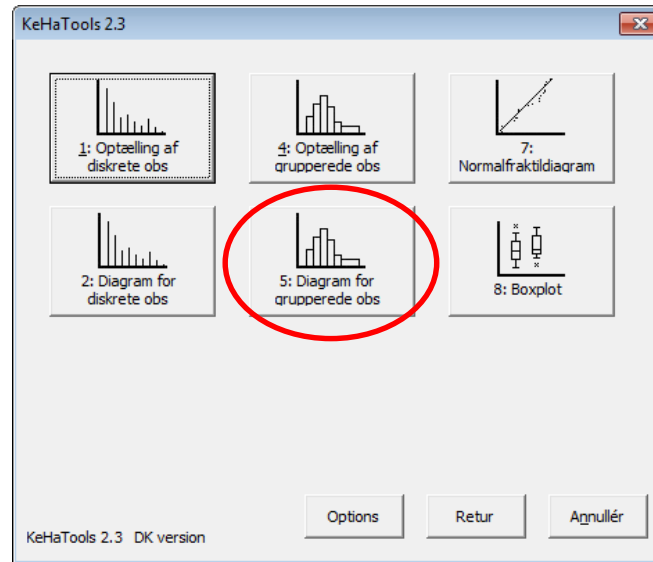
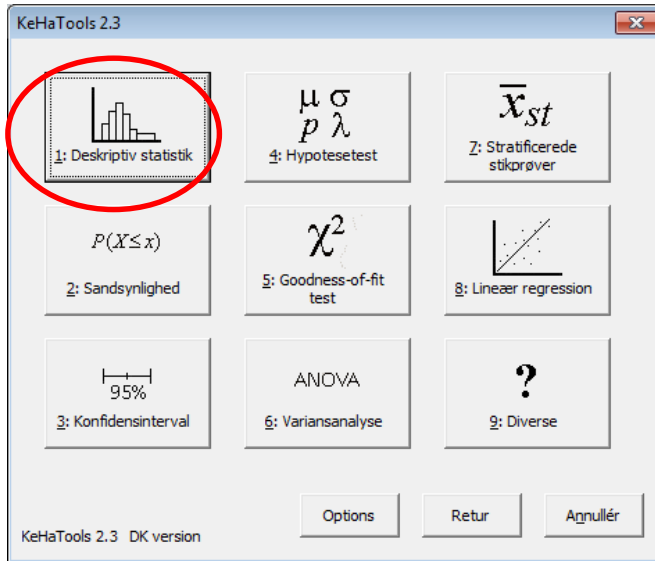


Eksempel 1.5 - I

The screenshot shows an Excel spreadsheet with the following data:

	A	B	C	D
1	Optælling af grupperede observationer			
2				
3	Min	Max	Hyppighed	
4	-	10	0	
5	10	10,1	0	
6	10,1	10,2	6	
7	10,2	10,3	24	
8	10,3	10,4	68	
9	10,4	10,5	60	
10	10,5	10,6	35	
11	10,6	10,7	7	
12	10,7	10,8	0	
13	10,8	10,9	0	
14	10,9	11	0	
15	11	-	0	
16				
17				
18				
19				

Eksempel 1.5 - II



Eksempel 1.5 - III

GS_Dat...

Filer Start Ind Side For

Hent eksterne data Opdater alle Forbindelser Sorter og filtre Dataværktøjer

D13 fx

	A	B	C	D
1	Optælling af grupperede observationer			
2				
3	Min	Max	Hyppighed	
4	-	10	0	
5	10	10,1	0	
6	10,1	10,2	6	
7	10,2	10,3	24	
8	10,3	10,4	68	
9	10,4	10,5	60	
10	10,5	10,6	35	
11	10,6	10,7	7	
12	10,7	10,8	0	
13	10,8	10,9	0	
14	10,9	11	0	
15	11	-	0	
16				
17				
18				
19				

Eksempel 1.2

Klar 100%

Deskriptiv statistik for grupperede observationer

Intervallgrænser

Min. grænse

Max. grænse

Hyppigheder

OK

annullér

Eksempel 1.5 - IV

