



EKONOMIHÖGSKOLAN  
Lunds universitet  
Statistiska institutionen

## *Slumpmässiga urval med Minitab*



# 1 OSU, obundet slumpmässigt urval

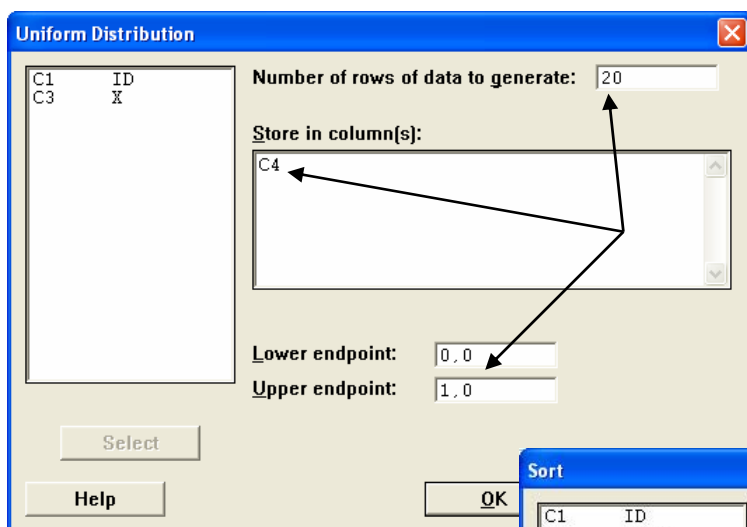
I Minitab har vi lagt upp ett register med våra tjugo bästa kompisar. Nu ska vi göra ett OSU för att välja ut fem av dem. Vi ska se på två olika sätt att göra det.

## Metod 1

Börja med att tilldela alla i registret ett slumptal, likformigt fördelat mellan noll och ett.

**Calc > Random Data > Uniform**

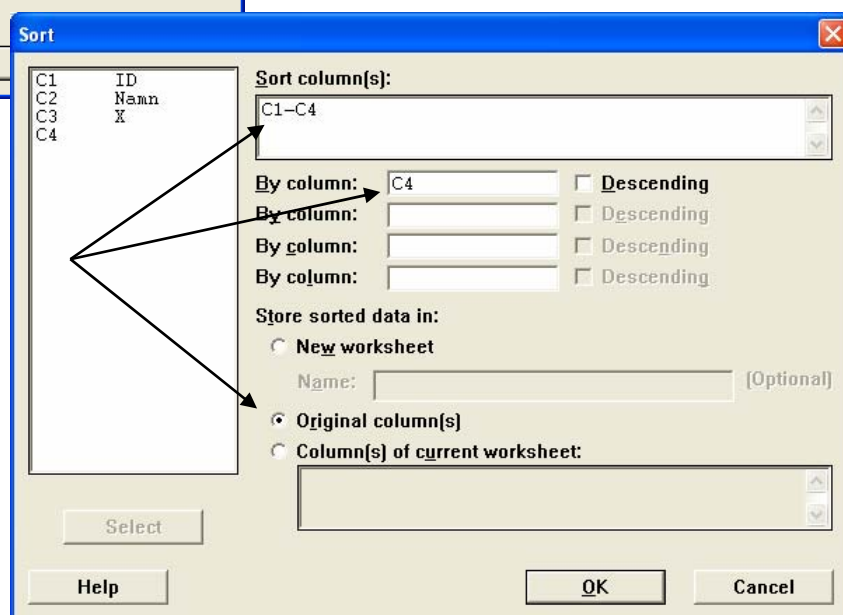
Fyll i enligt nedan.



	C1	C2-T	C3	C4	C5
	ID	Namn	X		
1	1	Azzam	50		
2	2	Beda	43		
3	3	Cilla	22		
4	4	David	30		
5	5	Esteban	47		
6	6	Fatima	37		
7	7	Gudrun	56		
8	8	Horst	44		
9	9	Ivan	32		
10	10	Josephine	26		
11	11	Katinka	20		
12	12	Lasse	58		
13	13	Melvin	54		
14	14	Nicole	41		
15	15	Olga	44		
16	16	Petros	51		
17	17	Quijote	55		
18	18	Ronja	33		
19	19	Sara	43		
20	20	Theodor	31		

När man sedan klickar på **OK** läggs slumptalen i C4. Sortera där-efter materialet efter kolumn C4.

**Data > Sort**



Och eftersom ordningen nu är slumpmässig kan man betrakta de översta som ett slumpmässigt urval bland kompisarna.

Grattis Nicole, David, Cilla, Quijote och Azzam !

	C1	C2-T	C3	C4
	ID	Namn	X	
1	14	Nicole	41	0,007293
2	4	David	30	0,042828
3	3	Cilla	22	0,240981
4	17	Quijote	55	0,264217
5	1	Azzam	50	0,324608
6	20	Theodor	31	0,344221
7	8	Horst	44	0,350359
8	15	Olga	44	0,365627
9	19	Sara	43	0,378058

## Metod 2

Egentligen behöver man inte vara så kreativ som vi var på föregående sida. Minitab har faktiskt en inbyggd funktion för att göra OSU (simple random sampling).

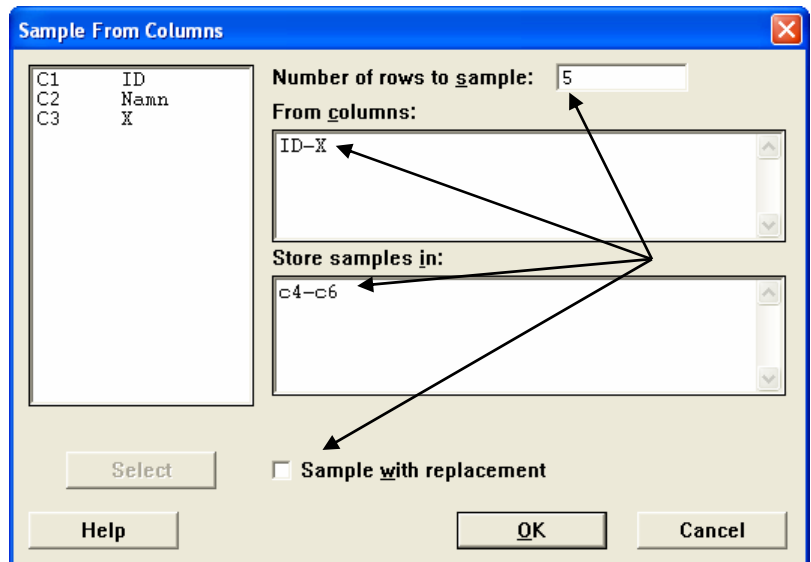
### Calc > Random Data > Sample from Columns

Till höger har vi angett att 5 observationer från kolumnerna ID-X (dvs C1–C3) ska väljas och läggas i C4–C6.

Urvalet görs *utan* återläggning om man inte kryssar för rutan ”Sample with replacement”.

Resultat blir att fem av observationerna väljs slumpmässigt av Minitab och värdena ser du längre ner.

Den här gången blev Azzam, Sara, Theodor, Cilla och Lasse de utvalda.



	C1	C2-T	C3	C4	C5-T	C6	C7	C8
	ID	Namn	X					
1	1	Azzam	50	1	Azzam	50		
2	2	Beda	43	19	Sara	43		
3	3	Cilla	22	20	Theodor	31		
4	4	David	30	3	Cilla	22		
5	5	Esteban	47	12	Lasse	58		
6	6	Fatima	37					
7	7	Gudrun	56					
8	8	Horst	44					
9	9	Ivan	32					
10	10	Josephine	26					
11	11	Katinka	20					
12	12	Lasse	58					
13	13	Melvin	54					
14	14	Nicole	41					
15	15	Olga	44					
16	16	Petros	51					
17	17	Quijote	55					
18	18	Ronja	33					
19	19	Sara	43					
20	20	Theodor	31					

## 2 Systematiskt urval

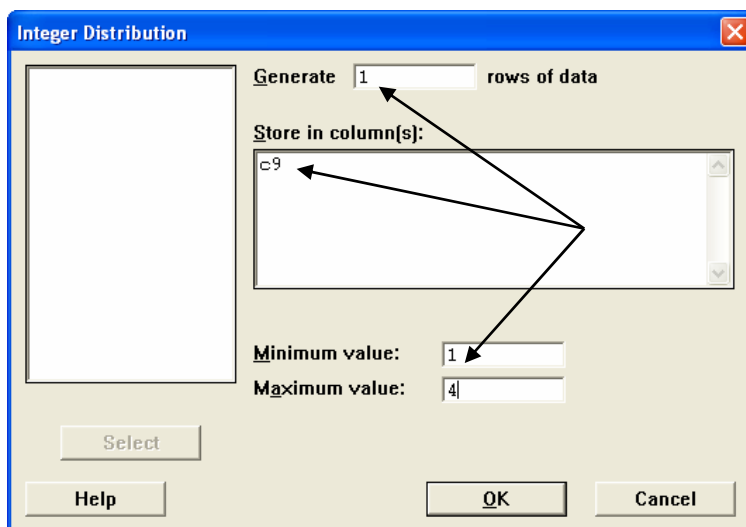
Det finns flera olika sätt att lösa det på i Minitab. MEN man måste själv göra tankearbetet!

Att välja fem av 20 är detsamma som att välja var fjärde från listan ( $5/20 = 1/4$ ) med en slumpmässig startpunkt bland de fyra första.

Den slumpmässiga startpunkten får du genom att kasta med en fyrsidig tärning. Använder du Minitab kan det göras som till höger.

### Calc > Random Data > Integer.

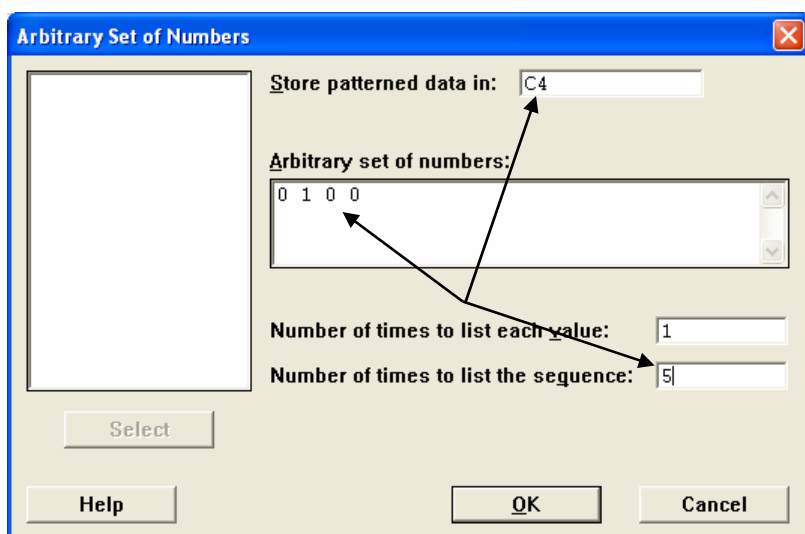
Ett slumptal (översta pilen) mellan 1 och 4 (nedersta pilen) väljs och läggs i kolumn C9.



Vi säger att det blev en 2:a och då ska vi alltså välja rad 2, 6, 10, 14 och 18 bland våra ursprungliga datavärden. Nu ska vi se hur man går till väga när starten väl är bestämd.

### Metod 3 (Copy med hjälpkolumn)

### Calc > Make Patterned Data > Arbitrary Set of Numbers



De fyra värdena 0 1 0 0 läggs i C4 och hela sekvensen upprepas fem gånger vilket blir sammanlagt 20 värden – precis så många som observationerna i datafilen.

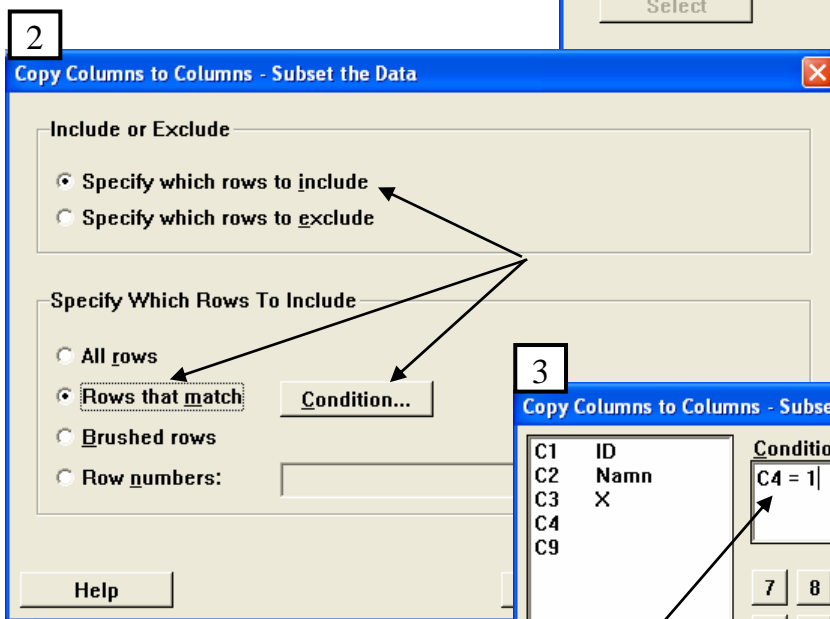
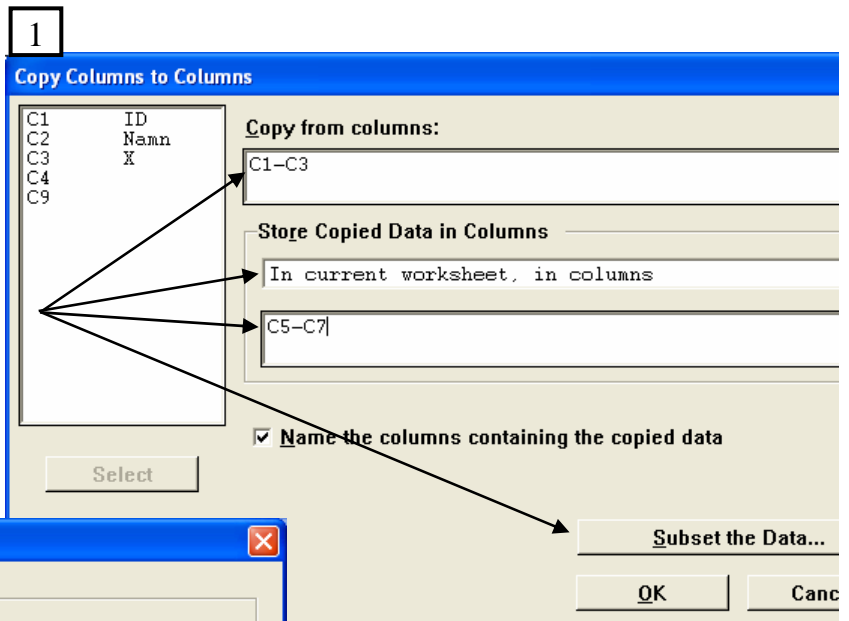
Nästa steg blir att kopiera enbart de rader som har värdet 1 (ett) i kolumn C4.

Lösningen kommer på nästa sida.

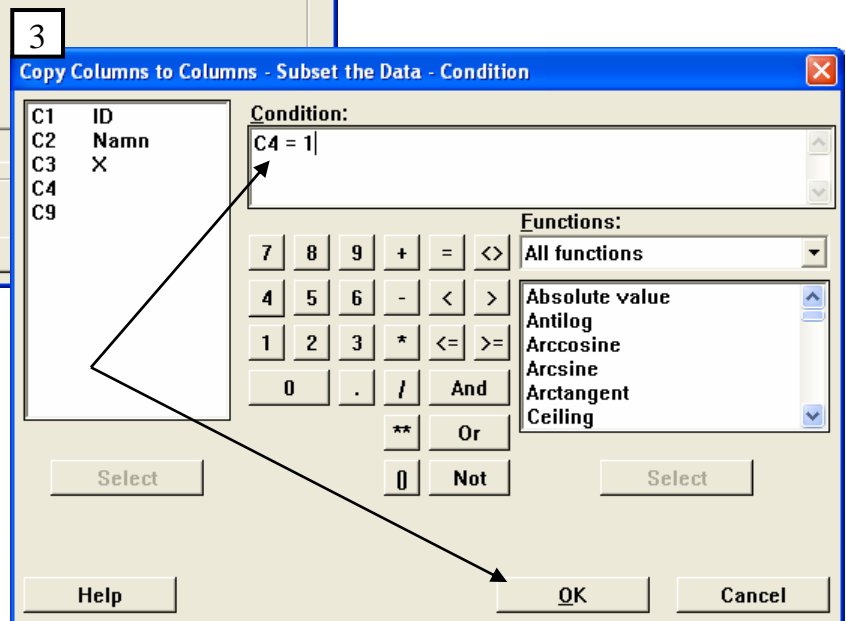
## Data > Copy > Columns to Columns

1. Fyll i att C1–C3 ska kopieras till C5–C7 “In current worksheet” och klicka sedan på **Subset the Data**.

2. I nästa meny kryssar man för ”Rows that match” och klickar på **Condition...**



3. ... och där berättar man till slut att villkoret som ska vara uppfyllt är att C4 = 1.

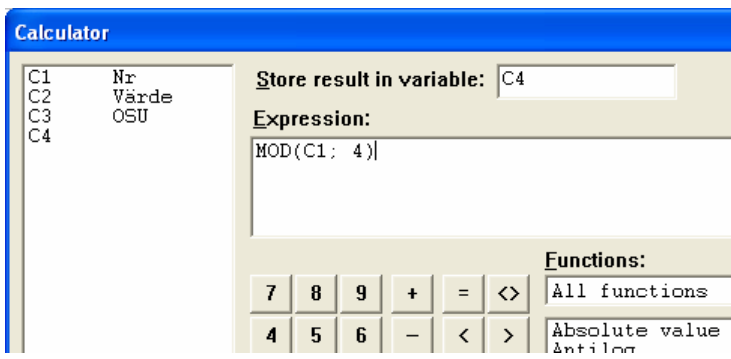


Några klick på **OK** och de utvalda visar sig vara Beda, Fatima, Josephine, Nicole och Ronja.

	C1	C2-T	C3	C4	C5	C6-T	C7	C8	C9	C10	C11
	ID	Namn	X		ID_1	Namn_1	X_1				
1	1	Azzam	50	0	2	Beda	43		2		
2	2	Beda	43	1	6	Fatima	37				
3	3	Cilla	22	0	10	Josephine	26				
4	4	David	30	0	14	Nicole	41				
5	5	Esteban	47	0	18	Ronja	33				
6	6	Fatima	37	1							
7	7	Guðrun	56	0							
8	8	David	44	0							

Det lilla tricket att lägga nollor och ettor i kolumn C4 är ju inte så dumt. Det finns ett annat sätt som man också kan använda och det är den matematiska modulo-funktionen som ger resten vid heltalsdivision. Om man dividerar radnumret (i C1) med periodiciteten (4 i det här fallet, vi skulle ju ha vart fjärde värde) blir *resten vid heltalsdivision* med talet 4 i tur och ordning värdena 1, 2, 3, 0, 1, 2, 3, 0, 1, 2, 3, 0, etc. Det fungerar så här i Minitab.

## Calc > Calculator



	C1	C2-T	C3	C4	C5	C6
	ID	Namn	X			
1	1	Azzam	50	1		
2	2	Beda	43	2		
3	3	Cilla	22	3		
4	4	David	30	0		
5	5	Esteban	47	1		
6	6	Fatima	37	2		
7	7	Gudrun	56	3		
8	8	Horst	44	0		
9	9	Ivan	32	1		
10	10	Josephine	26	2		
11	11	Katinka	20	3		
12	12	Lasse	58	0		
13	13	Melvin	54	1		
14	14	Nicole	41	2		
15	15	Olga	44	3		
16	16	Petros	51	0		
17	17	Quijote	55	1		
18	18	Ronja	33	2		
19	19	Sara	43	3		
20	20	Theodor	31	0		

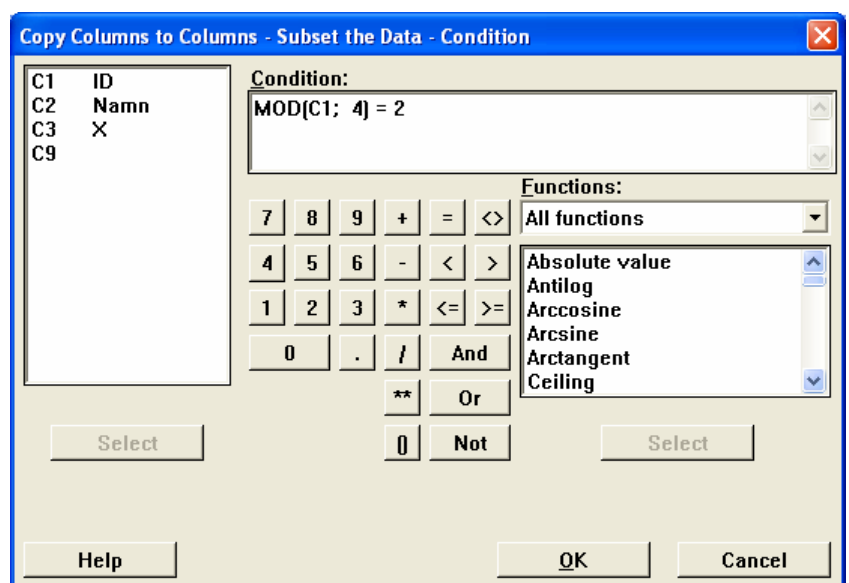
Resultatet har du till höger och sen är det bara att använda **Copy** på samma sätt som tidigare.

## Metod 4 (Copy med villkor direkt)

Hoppsan, nu har du naturligtvis redan kommit på hur man kunde underlättat arbetet genom att använda **Copy** med ett villkor direkt i stället för att gå omvägen över en hjälpkolumn.

Gå direkt till **Data > Copy > Columns to Columns** och fyll i som till höger.

Smart va!



### 3 Stratifierat urval

Vid stratifierat urval har man en variabel som visar stratumtillhörighet, kolumnen C4 i exemplet till höger. Ettorna bor på landet och tvåorna bor i stan.

Innan man gör sina urval inom varje stratum måste man först dela upp datamaterialet så att varje stratum finns antingen i egna kolumner eller i egna worksheets. Den uppdelningen kan man göra så här.

#### i) Strata i olika kolumner

Antingen använder man

**Data > Copy > Column to Column**

på samma sätt som på sidan 5. Ett **Copy** för varje stratum måste man i så fall göra.

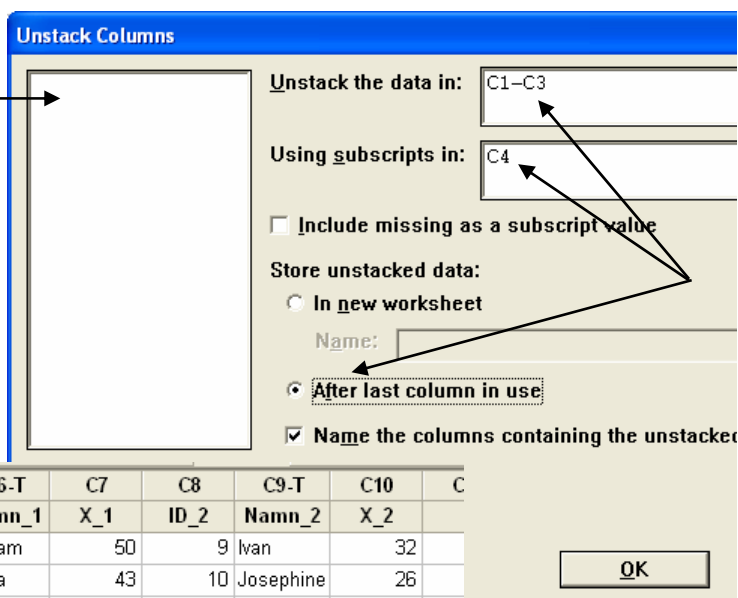
Alternativt kan man använda

**Data > Unstack Columns**

Kolumnerna C1–C3 delas upp i grupper efter vilka värden som finns i kolumn C4 som är stratumvariabeln.

Här under ser du resultatet och nu är det bara att göra urvalen från varje stratum separat.

	C1	C2-T	C3	C4
	ID	Namn	X	Bor
1	1	Azzam	50	1
2	2	Beda	43	1
3	3	Cilla	22	1
4	4	David	30	1
5	5	Esteban	47	1
6	6	Fatima	37	1
7	7	Gudrun	56	1
8	8	Horst	44	1
9	9	Ivan	32	2
10	10	Josephine	26	2
11	11	Katinka	20	2
12	12	Lasse	58	2
13	13	Melvin	54	2
14	14	Nicole	41	2
15	15	Olga	44	2
16	16	Petros	51	2
17	17	Quijote	55	2
18	18	Ronja	33	2
19	19	Sara	43	2
20	20	Theodor	31	2



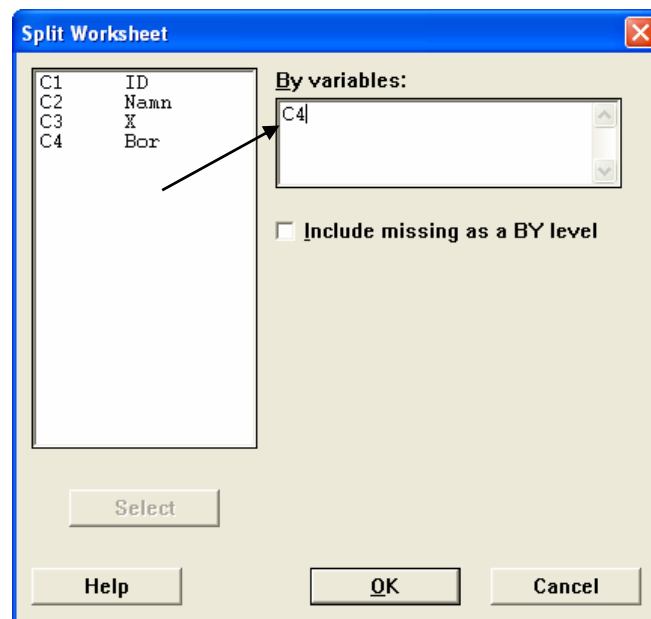
	C1	C2-T	C3	C4	C5	C6-T	C7	C8	C9-T	C10	C
	ID	Namn	X	Bor	ID_1	Namn_1	X_1	ID_2	Namn_2	X_2	
1	1	Azzam	50	1	1	Azzam	50	9	Ivan	32	
2	2	Beda	43	1	2	Beda	43	10	Josephine	26	
3	3	Cilla	22	1	3	Cilla	22	11	Katinka	20	
4	4	David	30	1	4	David	30	12	Lasse	58	
5	5	Esteban	47	1	5	Esteban	47	13	Melvin	54	
6	6	Fatima	37	1	6	Fatima	37	14	Nicole	41	
7	7	Gudrun	56	1	7	Gudrun	56	15	Olga	44	
8	8	Horst	44	1	8	Horst	44	16	Petros	51	
9	9	Ivan	32	2				17	Quijote	55	
10	10	Josephine	26	2				18	Ronja	33	
11	11	Katinka	20	2				19	Sara	43	
12	12	Lasse	58	2				20	Theodor	31	

## ii) Strata i olika worksheets

Genom att välja

### Data > Split Worksheet

kan man få sina strata i separata worksheets i samma Minitab Project i stället för att arbeta med olika kolumner.



När man klickar på **OK** ovan skapar Minitab ett nytt worksheet för varje stratum.

---