

## DATAANALYSE MED SUMPRODUKT()

### SUMPRODUCT()/SUMPRODUKT()

Læg mærke til at denne funktion ikke har punktum mellem SUM og PRODUKT i den danske version. Funktionen bruges egentlig til at beregne summen af produktet af elementerne i to eller flere matrixer. Men den kan altså også bruges til dataanalyse. Først dog den "almindelige" brug af funktionen.

Antag at du har følgende regneark:

	A	B
1	2	3
2	4	5
3	6	7

Hvis du indsætter formelen =SUMPRODUKT(A1:A3;B1:B3) vil funktionen gange tallet i A1 med tallet i B1 ( $2 * 3 = 6$ ) og tallet i A2 med tallet i B2 ( $4 * 5 = 20$ ), og endelig tallet i A3 med tallet i B3 ( $6 * 7 = 42$ ) så lægge de tre resultater sammen, så funktionens resultat bliver 68. Funktionen arbejder altså på matrixer, og den skal have mindst to matrixer, for at give mening. Fx vil =SUMPRODUKT(A1:A3) virke, men den vil bare lægge tallene sammen (12), og det kunne SUM() have gjort mere effektivt. På samme måde vil en funkti-

on, der arbejder på matrixer med kun et enkelt tal i hver matrix, fx =SUMPRODUKT(A1;B1) bare gange de tal med hinanden (6) og det kunne have været gjort bedre med PRODUKT() eller nemmere med en simpel regnefunktion =A1\*B1.

Nu var det imidlertid ikke lige det, vi ville bruge SUMPRODUKT() til i denne omgang, men derimod til dataanalyse. I denne forbindelse kan vi udnytte en anden egenskab ved et Excel regneark, nemlig at de logiske værdier SAND og FALSK, faktisk også har en talværdi. Prøv fx i to celler i et regneark at skrive =SAND(). Det skulle gerne give resultatet SAND i de to celler. Prøv dernæst at lægge dem sammen (med + ikke med SUM()). Nu skulle du gerne få resultatet 2 i formelcellen. Dette skyldes at SAND tildeles værdien 1, mens FALSK tildeles værdien 0. Prøv så i stedet at skrive ="Bil"="Bil" og ="Tog">"Bil". Begge disse formler giver også værdien SAND. Dette skyldes, at vi formlerne bruger sammenligningsoperatører. Og resultatet af en sammenligning er altid enten SAND eller FALSK. Der er ingen andre muligheder, og det er dette forhold, vi udnytter i det følgende. Når det er sandt at "Tog" > "Bil", skyldes at "T" kommer senere i alfabetet end "B", og dets binære værdi derfor er højere. Et stort B har fx den binære værdi 01000010, svarende til 66, mens T har værdien 01010100 svarende til 84. Men det var et sidespring ☺. Tilbage til SUMPRODUKT() og dens anvendelse til dataanalyse.

Ved hjælp af filtre (auto eller avanceret), kan man udvælge fx alle de poster, hvor værdien i kolonne M er 245A, og værdien i kolonne R er OK. Ved hjælp af SUBTOTAL() kan man så tælle, hvor mange der opfylder betingelserne. Har vi imidlertid kun brug for antallet, og ikke for at se de enkelte rækker er dette en lidt besværlig procedure, og i stedet kan vi bruge netop SUMPRODUKT(). Formlen kunne i givet fald have set ud som følger:

=SUMPRODUKT((M2:M26929="245-A")\*(R2:R26929="OK")).

Læg mærke til, at der ikke bruges argumentadskillere (;) i parenteser. Hele indholdet af SUMPRODUKT() er ét langt argument.

Hvad er det så, der sker her? Jo, for hvert af de to områder, M2:M26929 og R2:R26929 gennemgår Excel indholdet af cellen, og foretager den sammenligning, der er specificeret. Altså er M2 = 245-

A? Hvis Ja, "skrives" SAND og hvis Nej "skrives" FALSK. Skrivning foregår ikke på skærmen men i et hukommelsesregister. Da SAND= 1 og FALSK = 0, har vi nu en række 0'er og 1-taller i hver af de to kolonner. SUMPRODUKT gange nu hvert par med hinanden, og lægger resultatet af denne udgangning sammen. De tilfælde, hvor der står 0 og 1, 1 og 0 eller 0 og 0 giver alle 0 i udgangningen, kun de tilfælde, hvor der står 1 og 1, bliver det 1. Disse 1-taller, som repræsenterer de tilfælde, hvor begge betingelser er opfyldt, lægges sammen, og summen svarer altså til antallet af situationer, hvor der for 245-A i M-kolonnen står OK i R-kolonnen.

Vil vi udvide beregningen til kun at omhandle de af ovenstående tilfælde, hvor P-kolonnen er 1, kan formelen ændres til

=SUMPRODUKT((M2:M26929="245-A")\*(R2:R26929="OK")\*(P2:P26929=1)) eller til  
=SUMPRODUKT((M2:M26929="245-A")\*(R2:R26929="OK")\*(P2:P26929<=3)) hvis den bare skal være mindre end eller lig med 3.

Læg mærke til, at kolonnernes orden er ligegyldig. Ud over at tælle, kan funktionen faktisk også summere. Denne formel giver således summen af de tal i kolonne for de rækker, der opfylder den nederste formel ovenfor:

=SUMPRODUKT((M2:M26929="245-A")\*(R2:R26929="OK")\*(P2:P26929<=3)\*(O2:O26929)).

Der er faktisk ingen grænser for, hvor mange ting, vi kan teste for samtidigt. Vi kan dog risikere at "løbe ind i" Excels begrænsning på, at en formel højst må være 256 tegn lang (Excel 2003). Ovenstående er 76 tegn lang.

NB! Som næsten altid ved arbejde med flere matrixer, skal disse være identiske i størrelse og placering.

I Excel 2007, kan en del af de opgaver, som vi tidligere løste med SUMPRODUKT() løses med to nye funktioner SUM.HVISER og TÆL.HVISER. SUMPRODUKT kan man dog fortsat ikke undvære i mange sammenhænge. Fx kan SUMPRODUKT regne med "Eller-betingelser" og ikke kun med "OG-Betingelser". Det kan du læse mere om i artiklerne om SUMPRODUKT() under Excel.