

	NAVN	Beskrivelse
1	idnummer	Kunde-ID
2	aar	Skadesåret
3	kat	Skadeskategori. Om det er tyveri eller indbrud
4	expo	Observationens eksponering i forsikringsår
5	skadesantal	Antal skader i eksponeringsperioden
6	alder	Kundens alder
7	forsum	Forsikringssum i kr. (indeks = 1990)
8	zone	Risikoområde defineret ud fra kundens postnummer. Antager værdier fra 1 til 9
9	anc	Kundens anciennitet hos selskabet i år
10	afst_brand	Afstand til nærmeste brandstation
11	afst_politi	Afstand til nærmeste politistation
12	anthund	Antallet af hunde i boligen
13	boligareal	Boligareal
14	selvrisk	Selvrisiko i kr. (indeks = 1990)
15	segment	Gruppering af kunder i tre niveauer
16	ant_vaadrum	Antallet af vådrum i boligen
17	betfrekvens	Betalingsfrekvens for kunden. Antager værdierne 1 til 3 (årlig, kvartalvis og månedlig)
18	byg_anvend_kode	Byggeanvendelseskode
19	kon	Kundens køn
20	lejlighed	Er boligen en lejlighed? (1 = ja, 0 = nej).
21	opfaar	Opførelsesår
22	parcel	Er boligen et parcelhus? (1 = ja, 0 = nej)
23	tagtype	Tagtype
24	varmeinst	Gruppering af hvordan huset bliver varmet op
25	geo	Områdetype (1 = Land, 2 = Landsby, 3 = By, 4 = Storby, 5 = Metropol)

Tabel 1: Beskrivelse af variablene i datasættene **freq.csv**.

	NAVN	Beskrivelse
1	<code>idnummer</code>	ID-nummer
2	<code>kat</code>	Skadeskategori. Om det er tyveri eller indbrud
3	<code>aar</code>	Skadesåret
4	<code>skadesudgift</code>	Skadens størrelse i kroner efter selvriskoen er trukket fra

Tabel 2: Beskrivelse af variablene i datasættet `claims.csv`.