

**F1**

**Variabelklassifikation**

# Variabel

- Variabel: en egenskap som kan variera mellan olika enheter i populationen
- Egenskaperna hos variablerna avgör
  - Val av diagram
  - Val av läges- och spridningsmått
  - Val av analysmetod

# Klassifikation av variabler

- Vilken **information** ger en viss variabel?
- Vilka **värden** antar denna variabel?
- Jämför variablerna *kön* och *inkomst*!
- Kön antar värdena man och kvinna.
- Värdena är **egenskaper** (som inte kan uttryckas på ett meningsfullt sätt med siffror).
- Inkomst antar olika **siffror** som värden.

# Klassifikation av variabler

- Kön är en **kvalitativ** variabel (kategorivariabel; eng: categorical variable) En kvalitativ variabel antar icke-numeriska värden.
- Inkomst är en **kvantitativ** variabel (numerisk variabel; eng: numerical variable) En kvantitativ variabel antar numeriska värden.

Vi kan dela upp de numeriska värdena ytterligare.

# Olika typer av tal

- Naturliga tal

1,2,3,...

- Heltal

..., -2, -1, 0, 1, 2, ...

- Rationella tal

$p/q$ , där  $p$  och  $q$  är heltal

- Exempel på rationella tal i decimalform:

$-14/1 = -14$      $3/4 = 0,75$      $2/7 = 0, \underline{285714} 285714$

- Reella tal – icke-periodiska decimaltal (det som återstår). Tex  $e = 2.71828182845904$  och  $\pi = 3.14159265358979$

# Uppräkneliga mängder och överuppräkneliga mängder

- Mängden av alla naturliga tal

$$\{1, 2, 3, \dots\}$$

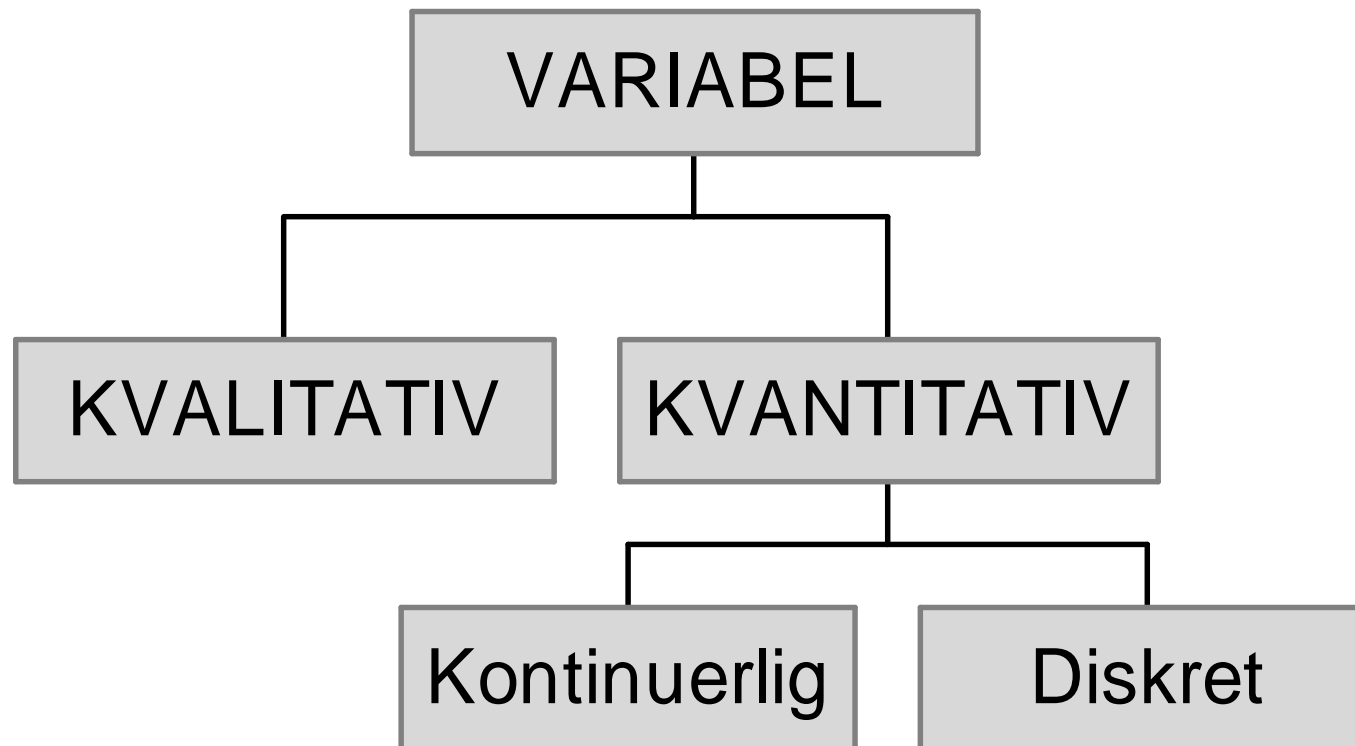
innehåller oändligt många tal. Den är **oändlig**.

- Mängden av alla reella tal är också oändlig.
- De naturliga talen kan **räknas upp**, men det kan inte de reella talen.

# Diskreta och kontinuerliga variabler

- Beroende på vilken sorts tal en variabel antar får de två olika namn.
- Vi sammanför de naturliga, hela och rationella talen, eftersom de alla är uppräkneliga. En variabel som antar ett av dessa tal kallas **diskret**.
- De reella talen är överuppräkneliga och en variabel som antar reella tal kallas **kontinuerlig**.

# Variabelbegreppet



# Datanivåer

- Vilka beräkningar kan göras?
- **Nominalskala** – om observationerna endast kan klassificeras i kategorier.
- Om mätvärdena resulterar i en rangordning föreligger en **ordinalskala**. Kan jämföra variabelvärden.
- De mätvärden en termometer anger befinner sig på en **intervallskala**. Godtycklig nollpunkt.
- **Kvotskalan** har en absolut nollpunkt. Samtliga fyra räknesätt ger meningsfulla resultat.

# Exempel

- Är följande variabler kategoriska eller numeriska, och om numeriska, är de diskreta eller kontinuerliga?
  - Vikt
  - Färgpreferens
  - Tid för att springa 100 meter
  - Antal hemmavarande barn
- Vilken datanivå mäts följande variabler på?
  - Kön
  - Längd (i cm)
  - Årtal
  - Partisympati (S, V, Mp, M, osv.)
  - Attityd till statistik (Mycket positiv, ganska positiv, v/e, osv)
  - Poäng på prov