**Innføringskurs i statistikk og SPSS**

**Helse Nord (Teams)**

**2.-3. desember 2025**

 ***Litteratur s 2***

 ***Timeplan s 3***

 ***Database s 4***

 ***Øvelser s 5***

### Tom Wilsgaard og Marina Espinasse

***Litteratur***

Julie Palant: SPSS Survival Manual. 5th Edition. 2013 (McGraw-Hill)

Arild Bjørndal, Dag Hofoss: Statistikk for helse- og sosialfagene (Gyldendal Akademisk)

Betty R. Kirkwood, Jonathan AC Sterne: Essential Medical Statistics. 2003. (Blackwell)

Odd O. Aalen, og andre: Statistiske metoder i medisin og helsefag. 2.utgave 2018. (Gyldendal Akademisk)

Asbjørn Johannessen: Introduksjon til SPSS. Versjon 17 (Abstrakt forlag).

Randi Hammervold: En kort innføring i SPSS. Utgave 2 2012 (Fagbokforlaget)

Knut-Andreas Christophersen: Databehandling og statistisk analyse med SPSS. 4. utgave. 2009 (Unipub).

Petter Laake, Anette Hjartaker, Dag S. Thelle og Marit B. Veierød (red.): Epidemiologiske og kliniske forskningsmetoder. Gyldendal 2007.

## Timeplan

# Dag 1

09.15 **Beskrivende statistikk.**

* Forelesning 1
* PC demonstrasjon, SPSS 1.

11.00 **PC øvelse 1.**

12.00 Lunsjpause.

12.45 **Statistiske tester. Sammenligning mellom to grupper**

* Forelesning 2
* PC demonstrasjon, SPSS 2.

14.30 **PC øvelse 2.**

# Dag 2

09.15 **Parvis informasjon og krysstabell.**

* Forelesning 3
* PC demonstrasjon, SPSS 3.

11.00 **PC øvelse 3.**

12.00 Lunsjpause.

12.45 **Korrelasjon og regresjon**

* Forelesning 4
* PC demonstrasjon, SPSS 4

14.30 **PC øvelse 4.**

***Database***

Følgende data benyttes i kurset



**Variabelbeskrivelse:**

kjønn 1 mann, 2 kvinne

alder i hele år ved undersøkelse

røyk røyker daglig, 0 nei, 1 ja

fysakt fysisk aktivitetsnivå i fritida, 1 rolig, 2 moderat, 3 aktiv, 4 hard trening

ggt gammaglutamyltransferase (U/l)

høyde kroppshøyde i hele cm

vekt1 kroppsvekt i halve kg, undersøkelse 1

vekt2 kroppsvekt i halve kg, undersøkelse 2

kol1 total kolesterol (mmol/l), undersøkelse 1

kol2 total kolesterol (mmol/l), undersøkelse 2

kaffe kaffekonsum, 1 ikke, 2 1-4 kopper, 3 5-8 kopper, 4 9+ kopper

død død under oppfølging, 0 nei, 1 ja

## Øvelser

***PC øvelse 1***

1. Åpne datasettet ”kursfilUNN.sav”.
2. Presenter deskriptive karakteristikker (gj.snitt, std. avvik, maks, min, variasjonsbredde, median og 5, 10, 90 og 95 prosentilen) for variabelen *vekt1*.
3. Bruk ”compute” til å beregne BMI, *BMI1* = *vekt1* / (*høyde*/100)2.
4. Lag histogram for *BMI1,* samt inkluder normalkurven
5. Gjenta oppgave 2 for variablene *BMI1* og *ggt*.
6. Lag histogram for *ggt.*
7. Lag frekvensfordeling for *fysakt* og *røyk*.
8. Lag søylediagram for *fysakt* og *røyk*.

***PC øvelse 2***

1. Test om det er kjønnsforskjell for variabelen *BMI1*. Bruk to-uvalgs t-test.
2. Bruk Wilcoxon test.
3. Test om det er BMI forskjell mellom de som rapporterer aktiv/ hard trening og rolig/moderat fysisk aktivitet (dvs *fysakt* = 1,2 mot *fysakt* = 3,4).
4. Samme som pkt 3, men utfør separat analyse for hvert kjønn. Hint: Bruk ”split file” for kjønn.
5. Test om det er kjønnsforskjell for variabelen *ggt.* (NB! Husk å fjerne ”split file”).
6. Anta at vi kun er interessert i personer som har rapportert hard fysisk aktivitetsnivå i fritiden. Utfør pkt 5, hvor du først selekterer personer med *fysakt*=4.

***PC øvelse 3***

1. Test om det er en signifikant endring av *BMI* mellom undersøkelsene.
2. Lag endringsvariabel for BMI. Bruk en-utvalgs t-test til å svare på pkt 1.
3. Lag søylediagram over andel røykere i hver kategori av *fysakt*.
4. Lag krysstabell mellom *fysakt* og *røyk*. Test om det er noen signifikant sammenheng mellom variablene.
5. Test om det er noen signifikant sammenheng mellom variablene *kjonn* og *røyk*.
6. Bruk ”recode” til å lage en ny variabel, *ald10*, dvs alder i 10 års grupper.
7. Test om det er noen signifikant sammenheng mellom variablene *ald10* og *røyk*.

***PC øvelse 4***

1. Konstruer spredningsplott mellom *kol1* og *BMI1*, samt beregn korrelasjonskoeffisient (Pearson og Spearman). Er det signifikant lineær sammenheng mellom variablene?
2. Utfør lineær regresjonsanalyse hvor *kol1* er avhengig variable. Hvor mye øker kolesterolet hvis BMI økes med 1 enhet?
3. Tegn regresjonslinjen inn i spredningsplottet fra pkt. 1.
4. Lagre residualene fra analysen i pkt 2. Tegn histogram over residualene.
5. Plott residualene på y-aksen mot BMI1.
6. Kommenter plottene i pkt 4 og 5 i forhold til modellantakelsene.
7. Utfør en multivariabel lineær regresjonsanalyse hvor du inkluderer alder og BMI1 som uavhengige variabler. Tolk resultatene!