

# INNHOOLD

FORORD.....	5
DEL I	
<b>KOM I GANG</b> .....	17
KAPITTEL 1	
<b>HVA ER R?</b> .....	19
<b>1.1 Et statistikkprogram</b> .....	19
1.1.1 Installere programmet.....	20
1.1.1.1 Installere R.....	20
1.1.1.2 Installere og oppdatere RStudio.....	20
<b>1.2 Et programmeringsspråk</b> .....	22
KAPITTEL 2	
<b>R SOM KALKULATOR</b> .....	25
<b>2.1 Åpne RStudio for første gang</b> .....	26
<b>2.2 Datatyper i R</b> .....	30
2.2.1 Vektorer.....	30
2.2.1.1 Indeksering .....	31
2.2.1.2 Regneoperasjoner.....	32
2.2.1.3 Funksjoner og argumenter.....	33
2.2.1.4 Sortere vektorer .....	35
2.2.2 Matriser .....	36
2.2.2.1 Opprette en matrise.....	36
2.2.2.2 Opprette et datasett .....	37
2.2.2.3 Indeksering .....	39
2.2.2.4 Målenivå .....	40
2.2.3 Lister .....	43
2.2.3.1 Opprette en liste.....	44
2.2.3.2 Indeksering .....	44
<b>Opgaver</b> .....	46

KAPITTEL 3	
<b>R SOM TEGNEPROGRAM</b> .....	47
<b>3.1 Plassere punkter i et koordinatsystem</b> .....	47
<b>3.2 Opprette tegneark for så å legge inn informasjon</b> .....	50
3.2.1 Opprette og tilrettelegge tegnearket .....	50
3.2.2 Legge til grafikk på eksisterende ark .....	51
3.2.2.1 Legge til punkter, linjer og tekst .....	51
3.2.2.2 Spesifisere farge og størrelse .....	53
3.2.2.3 Gi tegnforklaringer .....	54
<b>Oppgaver</b> .....	56
KAPITTEL 4	
<b>INN OG UT AV R</b> .....	57
<b>4.1 Arbeidsflyt</b> .....	58
4.1.1 Lagre eget arbeid .....	58
4.1.1.1 Organisere kodefiler .....	58
4.1.1.2 Lagre kodefiler .....	59
4.1.1.3 Organisere arbeid i mapper .....	59
4.1.2 Bruke andres arbeid: R-pakker .....	61
<b>4.2 Import/eksport av datafiler</b> .....	64
4.2.1 R på hjemmebane (.rda) .....	65
4.2.2 Skrive data ut av R (.txt og .csv) .....	66
4.2.3 Åpne data i Excel/OpenOffice .....	67
4.2.4 Åpne data i R .....	68
4.2.4.1 Import fra tekstfiler (.txt og .csv) .....	69
4.2.4.2 Import fra andre filformater .....	70
<b>Oppgaver</b> .....	71
DEL II	
<b>UTFORSKE OG TILRETTELEGGE DATA</b> .....	73
KAPITTEL 5	
<b>BESKRIVENDE STATISTIKK</b> .....	75
<b>5.1 Datastrukturen</b> .....	76
<b>5.2 Univariate fordelinger</b> .....	77
5.2.1 Kategoriske variabler .....	78
5.2.1.1 Frekvenstabell .....	78
5.2.1.2 Kake- og søylediagram .....	79
5.2.2 Kontinuerlige variabler .....	80
5.2.2.1 Numeriske beskrivelser .....	80
5.2.2.2 Grafiske alternativer .....	82
<b>5.3 Bivariate sammenhenger</b> .....	86
5.3.1 To kategoriske variabler .....	86
5.3.1.1 Bivariat sammenheng med tall .....	86
5.3.1.2 Bivariat sammenheng i tegning .....	89

5.3.2	Én kategorisk og én kontinuerlig variabel .....	90
5.3.3	To kontinuerlige variabler .....	93
5.3.3.1	Numeriske beskrivelser .....	94
5.3.3.2	Grafiske alternativer .....	96
<b>Oppgaver</b>	.....	100

KAPITTEL 6

**TILRETTELEGGE DATA** .....

<b>6.1 Omkode eksisterende variabler</b>	.....	101
6.1.1 Kontinuerlige variabler .....	102	
6.1.1.1 Additive indekser .....	102	
6.1.1.2 Snu skalaretning .....	105	
6.1.1.3 Endre skalaenheter .....	106	
6.1.2 Kategoriske variabler (betinget omkoding) .....	106	
6.1.2.1 Omkode ved hjelp av indeksering .....	107	
6.1.2.2 Omkode med regelmessige uttrykk («regex») .....	109	
6.1.2.3 Automatiske varianter .....	112	
<b>6.2 Slå sammen data og endre verdier</b>	.....	113
6.2.1 Forprosessering – tilrettelegge for sammenslåing .....	114	
6.2.1.1 Aggregere data når observasjonsnivåene er ulike .....	114	
6.2.1.2 Nøkkelvariabler for observasjoner som hører sammen .....	115	
6.2.2 Data fra andre kilder .....	116	
6.2.2.1 Legge til enkeltvariabler («match») .....	117	
6.2.2.2 Slå sammen hele datasett («merge») .....	117	
<b>Oppgaver</b>	.....	119

DEL III

**REGRESJONSANALYSER** .....

KAPITTEL 7

**ESTIMERE OG TOLKE EN LINEÆR MODELL (OLS)** .....

<b>7.1 Introduksjon til OLS</b>	.....	124
7.1.1 Estimere en lineær modell .....	125	
7.1.2 Lese modellsammendraget .....	127	
<b>7.2 Tolke resultatene</b>	.....	128
7.2.1 Tolke marginaleffektene .....	129	
7.2.2 Tolke med konfidensintervall .....	130	
7.2.3 Tolke med predikerte verdier .....	131	
7.2.3.1 Beregne predikerte verdier .....	131	
7.2.3.2 Tegne predikerte verdier .....	132	
7.2.3.3 Beregne nøyaktighet .....	133	
7.2.3.4 Tegne usikkerhet .....	134	
7.2.3.5 Legge til observerte verdier .....	135	
<b>Oppgaver</b>	.....	138

## KAPITTEL 8

**MUTTEL REGRESJON: FLERE VARIABLER**

<b>OG IKKE-LINEÆRE EFFEKTER</b> .....	139
<b>8.1 Effekten av å kontrollere for andre variabler</b> .....	140
8.1.1 Estimere en muttel lineær modell.....	141
8.1.2 Tolke kontrollerte effekter i lineær modell.....	143
8.1.3 Sammenlikne effekter .....	144
8.1.3.1 Sammenlikning mellom modeller krever et likt utvalg .....	145
8.1.3.2 Sammenlikning i samme modell krever standardiserte variabler .....	145
8.1.4 Sammenlikne konfidensintervaller («coefplot») .....	147
<b>8.2 Ikke-lineære sammenhenger</b> .....	148
8.2.1 Samspillsledd .....	148
8.2.1.1 Estimere samspill.....	149
8.2.1.2 Tolke et samspill.....	150
8.2.2 Kurvilineære sammenhenger .....	155
8.2.2.1 Annengradsledd .....	155
8.2.2.2 Eksponentielle sammenhenger .....	159
<b>Oppgaver</b> .....	164

## KAPITTEL 9

**MODELLVURDERING**

.....	165
<b>9.1 Hvor godt predikerer modellen?</b> .....	166
9.1.1 Fordelingen til predikerte verdier .....	167
9.1.2 Korrelasjon mellom predikerte og observerte verdier.....	168
9.1.3 Prediksjon for undergrupper av data .....	169
<b>9.2 Residualenes fordeling</b> .....	170
9.2.1 Er residualene normalfordelte? .....	170
9.2.1.1 Beregne residualene .....	171
9.2.1.2 Kvantitative mål på skjevhet og kurtose .....	172
9.2.1.3 Histogram med hjelpelinje.....	173
9.2.1.4 Kvantil – kvantilgrafikk.....	174
9.2.1.5 Hva gjør vi hvis residualene ikke er normalfordelte? .....	174
9.2.2 Er residualene likt fordelt over avhengig variabel?.....	175
<b>9.3 Uteliggere og innflytelsesrike observasjoner</b> .....	177
9.3.1 Innflytelse – hatverdier.....	177
9.3.1.1 Beregne hatverdier .....	178
9.3.1.2 Identifisere kritiske hatverdier .....	179
9.3.2 Uteliggere – t-standardiserte residualer .....	182
9.3.2.1 Estimere t-standardiserte residualer .....	183
9.3.2.2 Identifisere kritiske verdier .....	184
9.3.3 Uteliggere med innflytelse – Cook's D .....	185
9.3.3.1 Beregne Cook's D .....	186
9.3.3.2 Identifisere kritiske verdier .....	186
9.3.4 Hva gjør vi med innflytelsesrike uteliggere?.....	190

<b>9.4 Kolinearitet</b> .....	192
9.4.1 Tegn på kolinearitet .....	192
9.4.2 Løsninger på kolinearitet .....	193
<b>Oppgaver</b> .....	194
KAPITTEL 10	
<b>BINOMISK LOGISTISK REGRESJON</b> .....	195
<b>10.1 Estimere en binomisk logistisk modell</b> .....	196
<b>10.2 Tolke en binomisk logistisk modell</b> .....	198
10.2.1 Tolke marginaleffektene .....	199
10.2.1.1 Retning og signifikans (endring i logodds) .....	199
10.2.1.2 Relativ endring (endring i odds, oddsratio) .....	199
10.2.2 Predikerte sannsynligheter .....	201
10.2.2.1 Regne ut predikerte sannsynligheter .....	202
10.2.2.2 Regne ut konfidensintervallet til predikerte verdier .....	203
10.2.2.3 Tegne predikerte verdier .....	204
<b>10.3 Vurdere en binomisk logistisk modell</b> .....	210
10.3.1 Hvor ofte predikerer modellen riktig? .....	210
10.3.1.1 Definere kuttpunkt og predikere utfall .....	210
10.3.1.2 Sammenlikne predikerte og observerte utfall .....	211
10.3.2 Avansert modellstatistikk .....	214
10.3.2.1 ROC-kurve .....	214
10.3.2.2 Hosmer-Lemeshows «Goodness-of-fit» .....	215
<b>Oppgaver</b> .....	216
<b>BIBLIOGRAFI</b> .....	217
<b>REGISTER FOR R-KOMMANDOER</b> .....	219