Informasjon
TMA4240 Statistikk

H 2025

Jo Eidsvik
Institutt for Matematiske Fag, NTNU

man-

tors-fredag

Emneuka

Her finner du en oversikt over hvordan undervisningen i TMA4240 Statistikk er organisert. Hver undervisningsuke består av følgende elementer (og typisk i den rekkefølgen som her er angitt):

- **Introvideoer**: Før hver av forelesningene skal du se en del introvideoer.
  - Oversikt over hvilke introvideoer som skal sees de ulike ukene finner du i emnematrisa.
  - Introvideoene gir deg en introduksjon til nye begreper og/eller regneregler som skal være i fokus den uka.

- **Forelesninger**: I dobbelttimer med (tradisjonell) forelesninger (parallell 1) eller (parallell 2) diskuterer vi nye begreper, regning og illustrerer disse via relevante problemstillinger /eksempler.
  - Forelesningene tar utgangspunkt i at du allerede har sett ukas introvideoer.
  - Se når og hvor dine forelesninger er på parallell 1 eller parallell 2.

- **Elektronisk øving**: Hver uke blir det en elektronisk øving.
  - Hver elektronisk øving består av seks relativt kortfattede oppgaver hvor svaret typisk er et (eller flere) tall. Du får umiddelbart tilbakemelding på om svaret du angir er korrekt, og om svaret ditt er feil kan du prøve på nytt (så mange ganger du vil).
  - Link til ukas elektroniske øving med tilhørende innleveringsfrist finner du i emnematrisa.
  - For å få en elektronisk øving godkjent kreves det at minst fire av de seks oppgavene er korrekt besvart.
  - Hvis du trenger hjelp med en elektronisk øving kan du få dette på statistikklab.
  - De elektroniske øvingene er obligatoriske, se beskrivelse på siden obligatoriske aktiviteter for hvor mange som kreves godkjent.

- **Skriftlig innlevering**: Hver annen uke blir det en skriftlig innlevering.
  - En skriftlig innlevering består typisk av litt større oppgaver hvor du skal regne på flere problemstillinger relatert til samme situasjon, og/eller hvor du skal benytte ulike metoder til å finne svar på den samme problemstillingen.
  - De skriftlige innleveringene løses delvis ved hjelp av regning med papir og blyant og delvis ved å skrive og kjøre python-kode.
  - De skriftlige innleveringene skal leveres inn og vil bli rettet (og godkjent).
  - Link til de skriftlige innleveringene med tilhørende innleveringsfrist finner du i emnematrisa.
  - Merk at en del av oppgavene i en skriftlig innlevering kan løses kun basert på begrepene og regnereglene introdusert i den først undervisningsuka tilhørende den skriftlige innleveringen. Disse oppgave er markert med en stjerne etter oppgavenummeret. Det anbefales at du utnytter dette til å få en jevn arbeidsbelastning og dermed unngår at du får for mye å gjøre like før innleveringsfrist for skriftlig innlevering.
  - For å få en skriftlig innlevering godkjent kreves det at du har gjort et ærlig forsøk på (nesten) samtlige oppgaver og at minst omtrent halvparten er korrekt. Og korrekt betyr ikke bare at svaret er korrekt, men at også argumentasjonen og fremgangsmåten er korrekt, og at besvarelsen er fornuftig ført.
  - Hvis du trenger hjelp med en skriftlig innlevering kan du få dette på statistikklab.
  - De skriftlige innleveringene er obligatoriske, se beskrivelse på siden obligatoriske aktiviteter for hvor mange som kreves godkjent.

man-

Emneuka

Her finner du en oversikt over hvordan undervisningen i TMA4240 Statistikk er organisert. Hver undervisningsuke består av følgende elementer (og typisk i den rekkefølgen som her er angitt):

- **Introvideoer**: Før hver av forelesningene skal du se en del introvideoer.
  - Oversikt over hvilke introvideoer som skal sees de ulike ukene finner du i emnematrisa.
  - Introvideoene gir deg en introduksjon til nye begreper og/eller regneregler som skal være i fokus den uka.

- **Forelesninger**: I dobbelttimer med (tradisjonell) forelesninger (parallell 1) eller (parallell 2) diskuterer vi nye begreper, regning og illustrerer disse via relevante problemstillinger /eksempler.
  - Forelesningene tar utgangspunkt i at du allerede har sett ukas introvideoer.
  - Se når og hvor dine forelesninger er på parallell 1 eller parallell 2.

Essensen på videoer.
Mye eksempler og regning på forelesning.

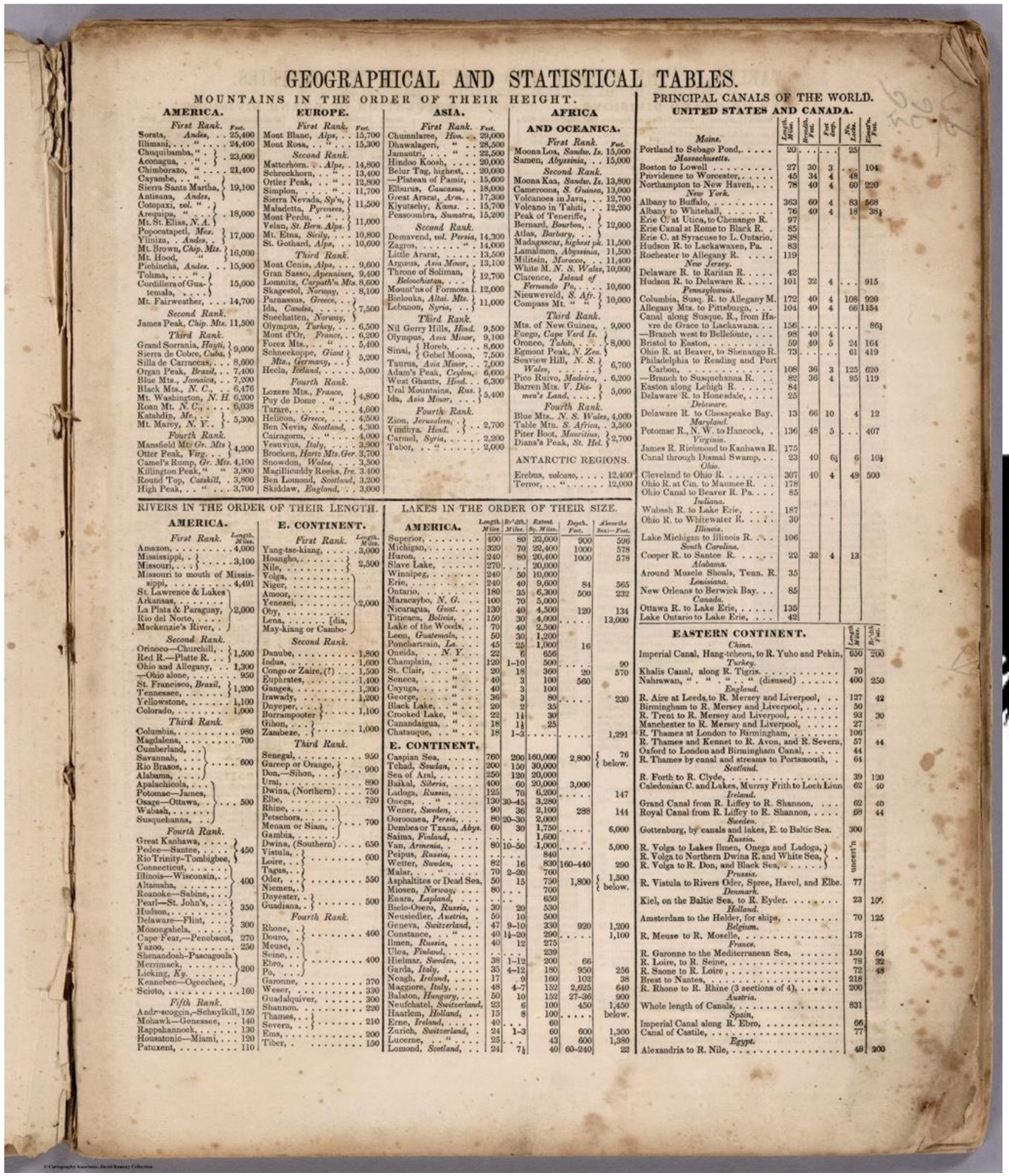
Øvinger

- Elektroniske øvinger:
  - Minst 4 (av 7) godkjente elektroniske øvinger fra Blokk 1.
  - Minst 3 (av 6) godkjente elektroniske øvinger fra Blokk 2.
- Skriftlige innleveringer:
  - Totalt minst 3 (av 6) skriftlige innleveringer må være godkjent.
  - Minst 1 (av 3) godkjente skriftlige innleveringer fra Blokk 1.
  - Minst 1 (av 3) godkjente skriftlige innleveringer fra Blokk 2.

tors-fredag

Motivasjon for grunnkurset i statistikk

Hva er statistikk?



GEOGRAPHICAL AND STATISTICAL TABLES.

MOUNTAINS IN THE ORDER OF THEIR HEIGHT.

AMERICA.

First Rank.	Feet.
Sorata, *Andes*,	25,400
Illimani, "	24,400
Chuquibamba, "	23,000
Aconagua, "	23,000
Chimborazo, "	21,400
Cayambe, "	19,100
Sierra Santa Martha, *Antisana, Andes*,	19,100
Cotopaxi, *vol.* "	18,000
Arequipa, "	18,000
Mt. St. Elias, *N. A.*	17,000
Popocatapeti, *Mex.*	17,000
Ylimiza, *Andes*,	16,000
Mt. Brown, *Chip. Mts.*	16,000
Mt. Hood, "	15,900
Pichincha, *Andes*,	15,000
Tolima, "	15,000
Cordillera of Guatemala,	14,700
Mt. Fairweather,	
Second Rank.	
James Peak, *Chip. Mts.*	11,500
Third Rank.	
Grand Sorrania, *Hayti*,	9,000
Sierra de Cobre, *Cuba*,	8,600
Silla de Carracas,	7,400
Organ Peak, *Brazil*,	7,200
Blue Mts., *Jamaica*,	7,200
Black Mts., *N. C.*,	6,476
Mt. Washington, *N. H.*	6,200
Ronn Mt. *N. C.*,	6,038
Katahdin, *Me.*,	5,300
Mt. Marcy, *N. Y.*,	5,300
Fourth Rank.	
Mansfield Mt. *Gr. Mts*	4,200
Otter Peak, *Virg.*,	4,200
Camel's Rump, *Gr. Mts.*	4,100
Killington Peak, "	3,900
Round Top, *Catskill*,	3,800
High Peak, "	3,700

EUROPE.

First Rank.	Feet.
Mont Blanc, *Alps*,	15,700
Mont Rosa, "	15,300
Second Rank.	
Matterhorn, *Alps*,	14,800
Schreckhorn, "	13,400
Ortler Peak, "	12,800
Simplon, "	11,700
Sierra Nevada, *Sp'n*,	11,500
Maladetta, *Pyrenees*,	
Mont Perdu, "	
Velan, *St. Bern. Alps*,	11,000
Mt. Etna, *Sicily*,	10,800
St. Gothard, *Alps*,	10,600
Third Rank.	
Mont Cenis, *Alps*,	9,600
Gran Sasso, *Apennines*,	9,400
Lomnitz, *Carpath'n Mts.*	8,600
Skagestol, *Norway*,	8,100
Parnassus, *Greece*,	
Ida, *Candia*,	7,500
Snechatten, *Norway*,	
Olympus, *Turkey*,	6,500
Mont d'Or, *France*,	6,200
Forez Mts., "	5,400
Schneekoppe, *Giant Mts., Germany*,	5,300
Hecla, *Iceland*,	5,000
Fourth Rank.	
Lozere Mts., *France*,	4,800
Puy de Dome "	
Tarare, "	4,600
Helicon, *Greece*,	4,500
Ben Nevis, *Scotland*,	4,300
Cairngorm, "	4,000
Vesuvius, *Italy*,	3,900
Brocken, *Harts Mts.Ger.*	3,700
Snowdon, *Wales*,	3,500
Magillicuddy Reeks, *Ire.*	3,400
Ben Lomond, *Scotland*,	3,200
Skiddaw, *England*,	3,000

ASIA.

First Rank.	Feet.
Chumularee, *Himalaya*,	29,000
Dhawalageri, "	28,500
Jamautri, "	22,500
Hindoo Koosh,	20,000
Belur Tag, highest,	20,000
—Plateau of Pamir,	15,600
Elburus, *Caucasus*,	18,000
Great Ararat, *Arm.*	17,300
Kiyutschy, *Kamt.*	15,700
Pessoombra, *Sumatra*,	15,200
Second Rank.	
Demavend, *vol. Persia*,	14,300
Zagros, "	14,000
Little Ararat,	13,500
Argaeus, *Asia Minor*,	13,100
Throne of Soliman, *Beloochistan*,	12,700
Mount'ns of Formosa I.	12,000
Bielouka, *Altai. Mts.*	11,000
Lebanon, *Syria*,	
Third Rank.	
Nil Gerry Hills, *Hind.*	9,500
Olympus, *Asia Minor*,	9,100
Horeb,	8,600
Sinai, Gebel Moosa,	7,500
Taurus, *Asia Minor*,	7,000
Adam's Peak, *Ceylon*,	6,600
West Ghauts, *Hind.*	6,300
Ural Mountains, *Rus.*	5,400
Ida, *Asia Minor*,	
Fourth Rank.	
Zion, *Jerusalem*,	2,700
Vindhya, *Hind.*	
Carmel, *Syria*,	2,200
Tabor, "	2,000

AFRICA AND OCEANICA.

First Rank.	Feet.
Moona Loa, *Sandw. Is.*	15,000
Samen, *Abyssinia*,	15,000
Second Rank.	
Moona Kaa, *Sandw. Is.*	13,800
Cameroons, *S. Guinea*,	13,000
Volcanoes in Java,	12,700
Volcano in Tahiti,	12,200
Peak of Teneriffe,	12,000
Bernard, *Bourbon*,	
Atlas, *Barbary*,	
Madagascar, *highest pk.*	11,500
Lamalmon, *Abyssinia*,	11,500
Militsin, *Morocco*,	11,400
White M. *N. S. Wales*,	10,000
Clarence, *Island of Fernando Po*,	10,600
Nieuweveld, *S. Afr.*	10,000
Compass Mt. "	
Third Rank.	
Mts. of New Guinea,	9,000
Fuego, *Cape Verd Is.*	
Oronoo, *Tahiti*,	8,000
Egmont Peak, *N. Zea.*	
Seaview Hill, *N. S. Wales*,	6,700
Pico Ruivo, *Madeira*,	6,200
Barren Mts. *V. Diemen's Land*,	5,000
Fourth Rank.	
Blue Mts., *N. S. Wales*,	4,000
Table Mtn. *S. Africa*,	3,500
Piter Boot, *Mauritius*,	2,700
Diana's Peak, *St. Hel.*	

ANTARCTIC REGIONS.

Erebus, *volcano*,	12,400
Terror, "	12,000

PRINCIPAL CANALS OF THE WORLD.

UNITED STATES AND CANADA.

	Length. Miles.	Breadth. Feet.	Feet deep.	No. Locks.	Breadth. Feet.
Maine.					
Portland to Sebago Pond,	20	. .	. .	25	
Massachusetts.					
Boston to Lowell	27	30	3	. .	104
Providence to Worcester,	45	34	4	48	
Northampton to New Haven,	78	40	4	60	220
New York.					
Albany to Buffalo,	363	60	4	83	568
Albany to Whitehall,	76	40	4	18	38½
Erie C. at Utica, to Chenango R.	97				
Erie Canal at Rome to Black R.	85				
Erie C. at Syracuse to L. Ontario,	38				
Hudson R. to Lackawaxen, Pa.	83				
Rochester to Allegany R.	119				
New Jersey.					
Delaware R. to Raritan R.	42				
Hudson R. to Delaware R.	101	32	4	. .	915
Pennsylvania.					
Columbia, Susq. R. to Allegany M.	172	40	4	108	920
Allegany Mts. to Pittsburgh,	104	40	4	66	1154
Canal along Susque. R., from Havre de Grace to Lackawana.	156	. .	. .	. .	86½
—Branch west to Bellefonte,	98	40	4		
Bristol to Easton,	59	40	5	24	164
Ohio R. at Beaver, to Shenango R.	73	. .	. .	61	419
Philadelphia to Reading and Port Carbon,	108	36	3	125	620
—Branch to Susquehanna R.	82	36	4	95	119
Easton along Lehigh R.	84				
Delaware R. to Honesdale,	25				
Deloware.					
Delaware R. to Chesapeake Bay.	13	66	10	4	12
Maryland.					
Potomac R., N. W. to Hancock,	136	48	5	. .	407
Virginia.					
James R. Richmond to Kanhawa R.	175				
Canal through Dismal Swamp,	23	40	6½	6	10½
Ohio.					
Cleveland to Ohio R.	307	40	4	49	500
Ohio R. at Cin. to Maumee R.	178				
Ohio Canal to Beaver R. Pa.	85				
Indiana.					
Wabash R. to Lake Erie,	187				
Ohio R. to Whitewater R.	30				
Illinois.					
Lake Michigan to Illinois R.	106				
South Carolina.					
Cooper R. to Santee R.	22	32	4	13	
Alabama.					
Around Muscle Shoals, Tenn. R.	35				
Louisiana.					
New Orleans to Berwick Bay.	85				
Canada.					
Ottawa R. to Lake Erie,	135				
Lake Ontario to Lake Erie,	42				

RIVERS IN THE ORDER OF THEIR LENGTH.

AMERICA.

First Rank.	Length. Miles.
Amazon,	4,000
Mississippi,	3,100
Missouri,	
Missouri to mouth of Mississippi,	4,491
St. Lawrence & Lakes	2,000
Arkansas,	
La Plata & Paraguay,	
Rio del Norte,	
Mackenzie's River,	
Second Rank.	
Orinoco—Churchill,	1,500
Red R.—Platte R.	
Ohio and Allegany,	1,300
—Ohio alone,	950
St. Francisco, *Brazil*,	1,200
Tennessee,	
Yellowstone,	1,100
Colorado,	1,000
Third Rank.	
Columbia,	980
Magdalena,	700
Cumberland,	
Savannah,	600
Rio Brasos,	
Alabama,	
Apalachicola,	
Potomac—James,	
Osage—Ottawa,	500
Wabash,	
Susquehanna,	
Fourth Rank.	
Great Kanhawa,	450
Pedee—Santee,	
Rio Trinity–Tombigbee,	
Connecticut,	
Illinois—Wisconsin,	400
Altamaha,	
Roanoke—Sabine,	
Pearl—St. John's,	350
Hudson,	
Delaware—Flint,	300
Monongahela,	
Cape Fear,—Penobscot,	270
Yazoo,	250
Shenandoah–Pascagoula	200
Merrimack,	
Licking, Ky.	
Kennebec—Ogeechee,	
Scioto,	160
Fifth Rank.	
Androscoggin,–Schuylkill,	150
Mohawk—Genessee,	140
Rappahannock,	130
Housatonic—Miami,	120
Patuxent,	110

E. CONTINENT.

First Rank.	Length. Miles.
Yang-tse-kiang,	3,000
Hoanghoo,	2,500
Nile,	
Volga,	
Niger,	
Amoor,	
Yeneseï,	2,000
Oby,	
Lena, [dia,	
May-kiang or Cambo-	
Second Rank.	
Danube,	1,800
Indus,	1,600
Congo or Zaire, (?)	1,500
Euphrates,	1,400
Ganges,	1,300
Irawady,	1,200
Dnyeper,	
Burrampooter	1,100
Gihon,	
Zambeze,	1,000
Third Rank.	
Senegal,	950
Gareep or Orange,	900
Don,—Sihon,	
Ural,	890
Dwina, (Northern)	750
Elbe,	720
Rhine,	
Petschora,	700
Menam or Siam,	
Gambia,	
Dwina, (Southern)	650
Vistula,	600
Loire,	
Tagus,	
Oder,	550
Niemen,	
Dnyester,	
Guadiana,	500
Fourth Rank.	
Rhone,	
Douro,	460
Meuse,	
Seine,	
Ebro,	400
Po,	
Garonne,	370
Weser,	330
Guadalquiver,	300
Shannon,	220
Thames,	
Severn,	210
Ems,	200
Tiber,	150

LAKES IN THE ORDER OF THEIR SIZE.

AMERICA.	Length. Miles.	Breadth. Miles.	Extent. Sq. Miles.	Depth. Feet.	Above the Sea.—Feet.
Superior,	400	80	32,000	900	596
Michigan,	320	70	22,400	1000	578
Huron,	240	80	20,400	1000	578
Slave Lake,	270	. . .	20,000		
Winnipeg,	240	50	10,000		
Erie,	240	40	9,600	84	565
Ontario,	180	35	6,300	500	232
Maracaybo, *N. G.*	100	70	5,000		
Nicaragua, *Guat.*	130	40	4,500	120	134
Titicaca, *Bolivia*,	150	30	4,000	. . . .	13,000
Lake of the Woods,	70	40	2,500		
Leon, *Guatemala*,	50	30	1,200		
Ponchartrain, *La.*	45	25	1,000	16	
Oneida, *N. Y.*	22	6	656		
Champlain, "	120	1–10	500	. . . .	90
St. Clair, "	20	18	360	20	570
Seneca, "	40	3	100	560	
Cayuga, "	40	3	100		
George, "	36	3	80	. . . .	230
Black Lake, "	20	2	35		
Crooked Lake, "	22	1½	30		
Canandaigua, "	18	1½	25		
Chatauque, "	18	1–3	. . . .	. . . .	1,291
E. CONTINENT.					
Caspian Sea,	760	200	160,000	2,800	{ 76 below.
Tchad, *Soudan*,	200	150	30,000		
Sea of Aral,	250	120	20,000		
Baikal, *Siberia*,	400	60	20,000	3,000	
Ladoga, *Russia*,	125	70	6,200	. . . .	147
Onega, "	130	30–45	3,280		
Wener, *Sweden*,	90	36	2,100	288	144
Ooroomea, *Persia*,	80	20–30	2,000		
Dembea or Tzana, *Abys.*	60	30	1,750	. . . .	6,000
Saima, *Finland*,	. .	. .	1,600		
Van, *Armenia*,	80	10–50	1,000	. . . .	5,000
Peipus, *Russia*,	. .	. .	840		
Wetter, *Sweden*,	82	16	830	160–440	290
Malar, "	70	2–20	760		
Asphaltites or Dead Sea,	50	15	750	1,800	{ 1,500 below.
Miosen, *Norway*,	80	. .	700		
Enara, *Lapland*,	. .	. .	630		
Bielo-Osero, *Russia*,	30	20	530		
Neusiedler, *Austria*,	50	10	500		
Geneva, *Switzerland*,	47	9–10	330	920	1,200
Constance, "	40	1½–20	290	. . . .	1,100
Ilmen, *Russia*,	40	12	275		
Ulea, *Finland*,	. .	. .	239		
Hielmar, *Sweden*,	38	1–12	200	66	
Garda, *Italy*,	35	4–12	180	950	256
Neagh, *Ireland*,	17	9	160	102	38
Maggiore, *Italy*,	48	4–7	152	2,625	640
Balaton, *Hungary*,	50	10	152	27–36	900
Neufchatel, *Switzerland*,	23	6	100	450	1,450
Haarlem, *Holland*,	15	8	100	. . . .	below.
Erne, *Ireland*,	40	. .	60		
Zurich, *Switzerland*,	24	1–3	60	600	1,300
Lucerne, "	25	. .	43	600	1,380
Lomond, *Scotland*,	24	7½	40	60–240	22

EASTERN CONTINENT.

	Length. Miles.	Breadth. Feet.
China.		
Imperial Canal, Hang-tcheou, to R. Yuho and Pekin,	650	200
Turkey.		
Khalis Canal, along R. Tigris.	70	
Nahrawan, " " " " (disused)	400	250
England.		
R. Aire at Leeds to R. Mersey and Liverpool,	127	42
Birmingham to R. Mersey and Liverpool,	50	
R. Trent to R. Mersey and Liverpool,	93	30
Manchester to R. Mersey and Liverpool,	27	
R. Thames at London to Birmingham,	106	
R. Thames and Kennet to R. Avon, and R. Severn,	57	44
Oxford to London and Birmingham Canal,	44	
R. Thames by canal and streams to Portsmouth,	64	
Scotland.		
R. Forth to R. Clyde,	39	120
Caledonian C. and Lakes, Murray Frith to Loch Linn	62	40
Ireland.		
Grand Canal from R. Liffey to R. Shannon,	62	40
Royal Canal from R. Liffey to R. Shannon,	68	44
Sweden.		
Gottenburg, by canals and lakes, E. to Baltic Sea.	300	
Russia.		
R. Volga to Lakes Ilmen, Onega and Ladoga, R. Volga to Northern Dwina R. and White Sea, R. Volga to R. Don, and Black Sea,	uncert'n	
Prussia.		
R. Vistula to Rivers Oder, Spree, Havel, and Elbe.	77	
Denmark.		
Kiel, on the Baltic Sea, to R. Eyder.	23	10?
Holland.		
Amsterdam to the Helder, for ships,	70	125
Belgium.		
R. Meuse to R. Moselle,	178	
France.		
R. Garonne to the Mediterranean Sea,	150	64
R. Loire, to R. Seine,	78	32
R. Seine to R. Loire,	72	48
Brest to Nantes,	218	
R. Rhone to R. Rhine (3 sections of 4),	200	
Austria.		
Whole length of Canals,	831	
Spain.		
Imperial Canal along R. Ebro,	66	
Canal of Castile,	77	
Egypt.		
Alexandria to R. Nile,	48	200





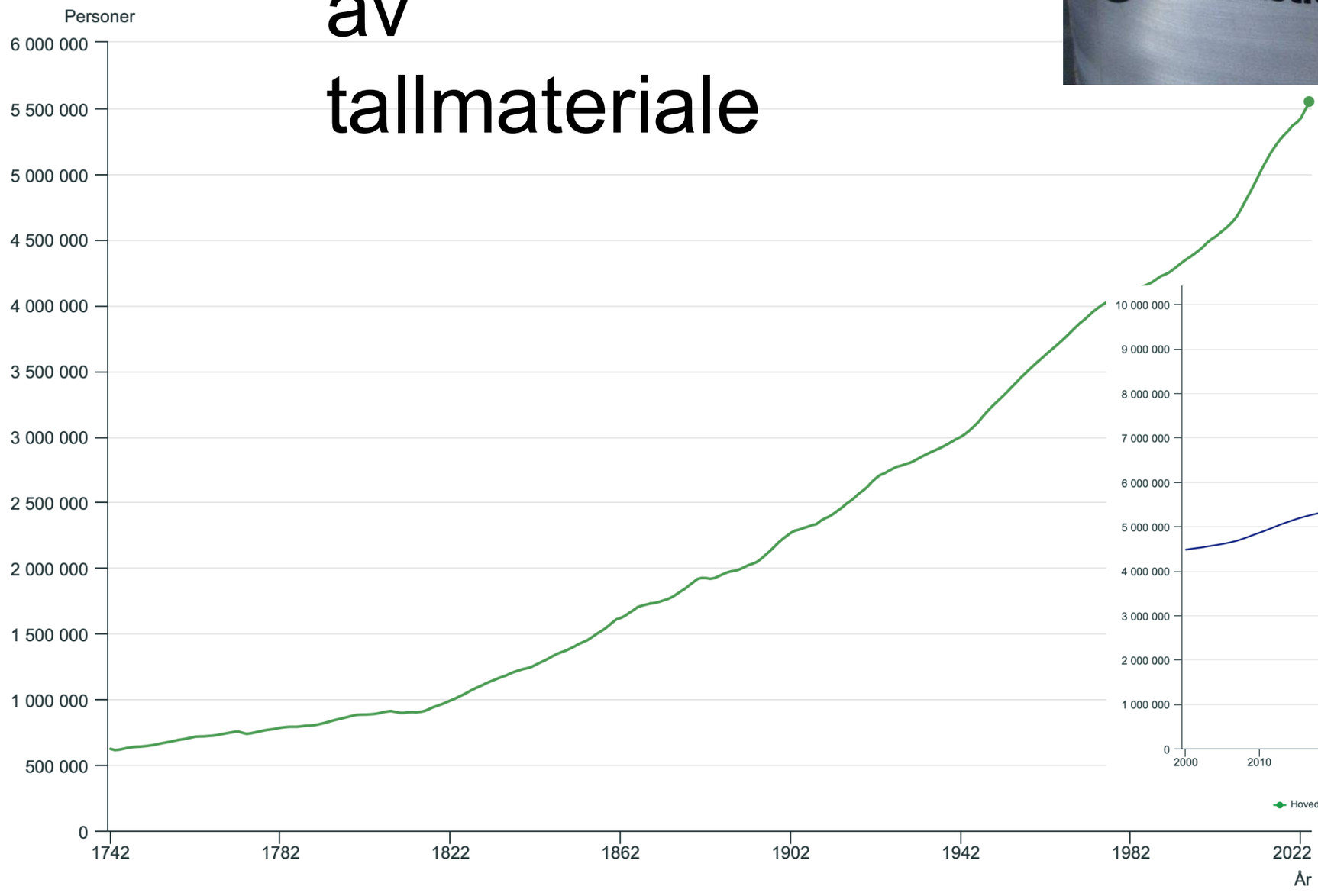


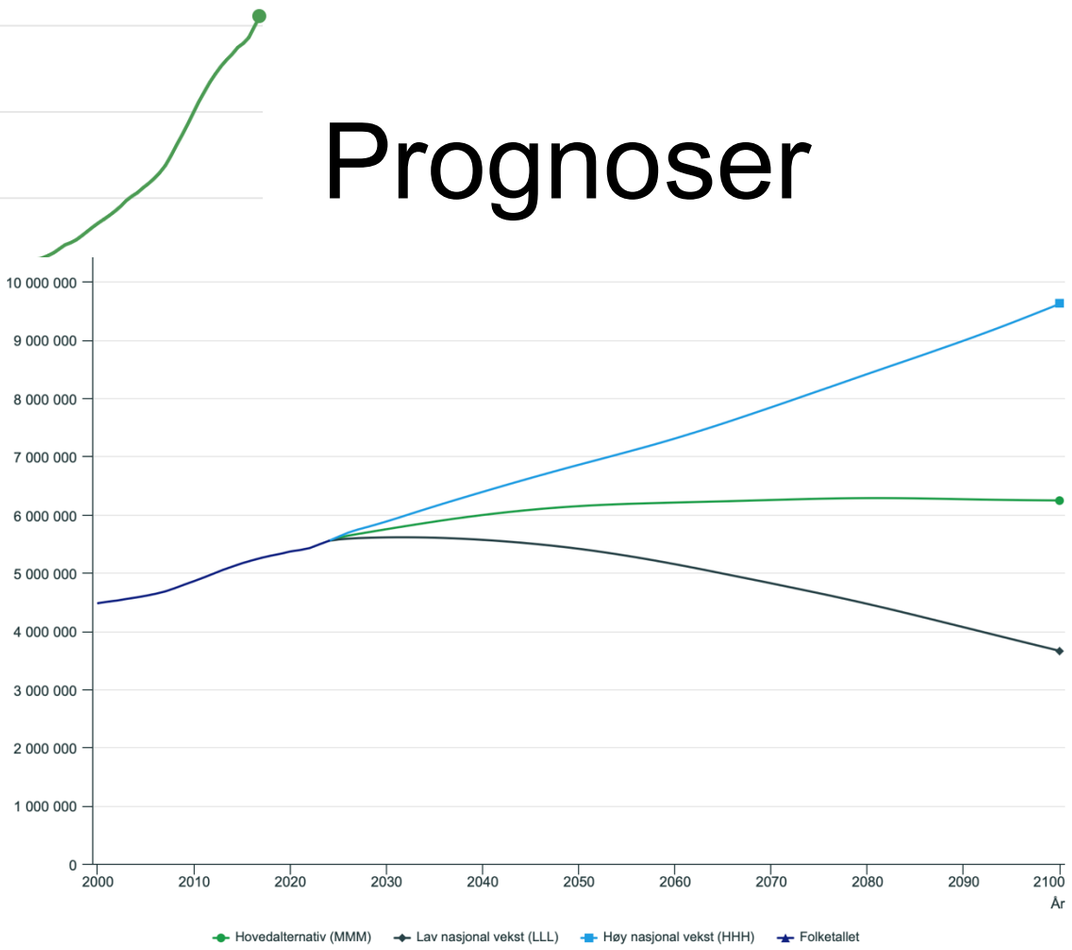
Hva er statistikk?

Oppsamling
av
tallmateriale



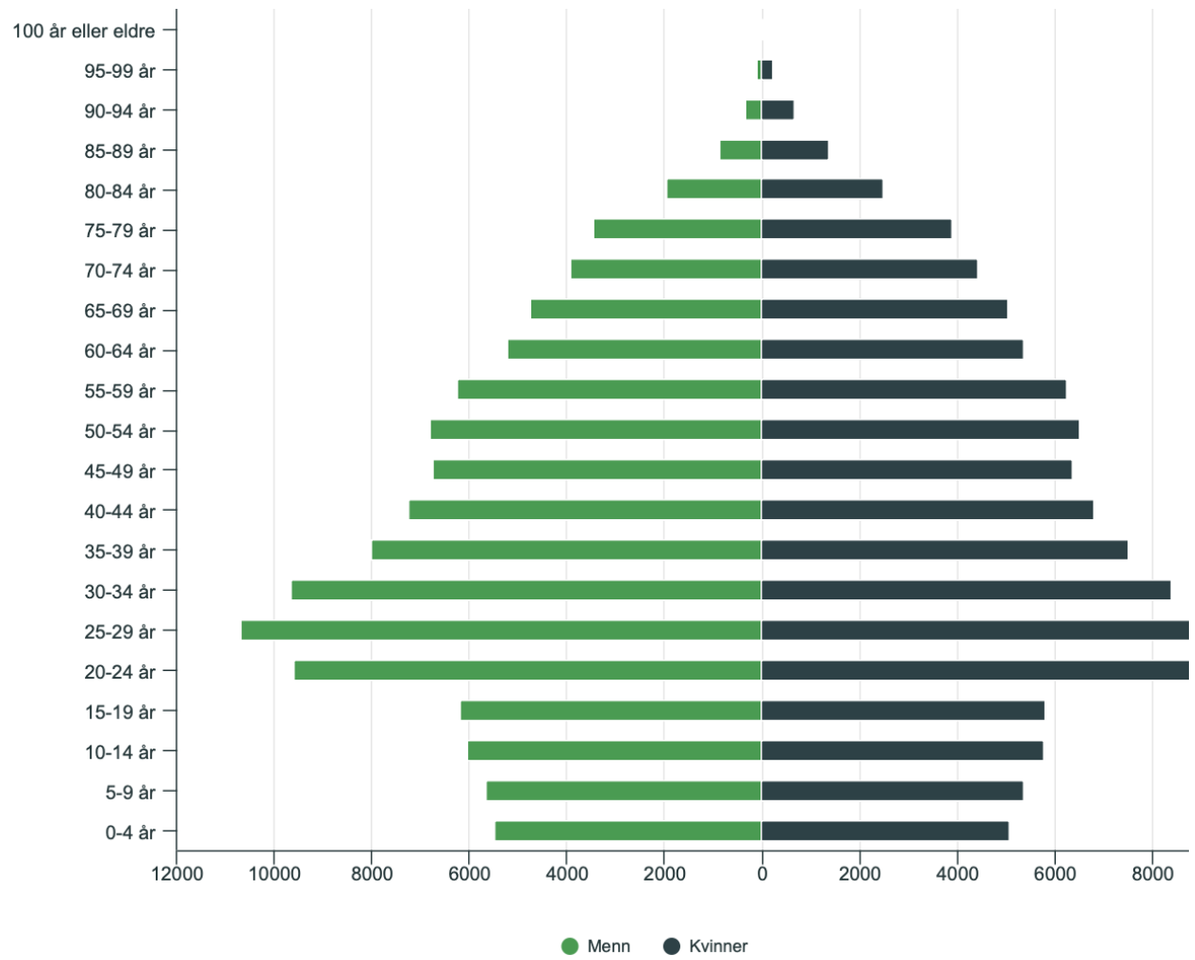
Prognoser



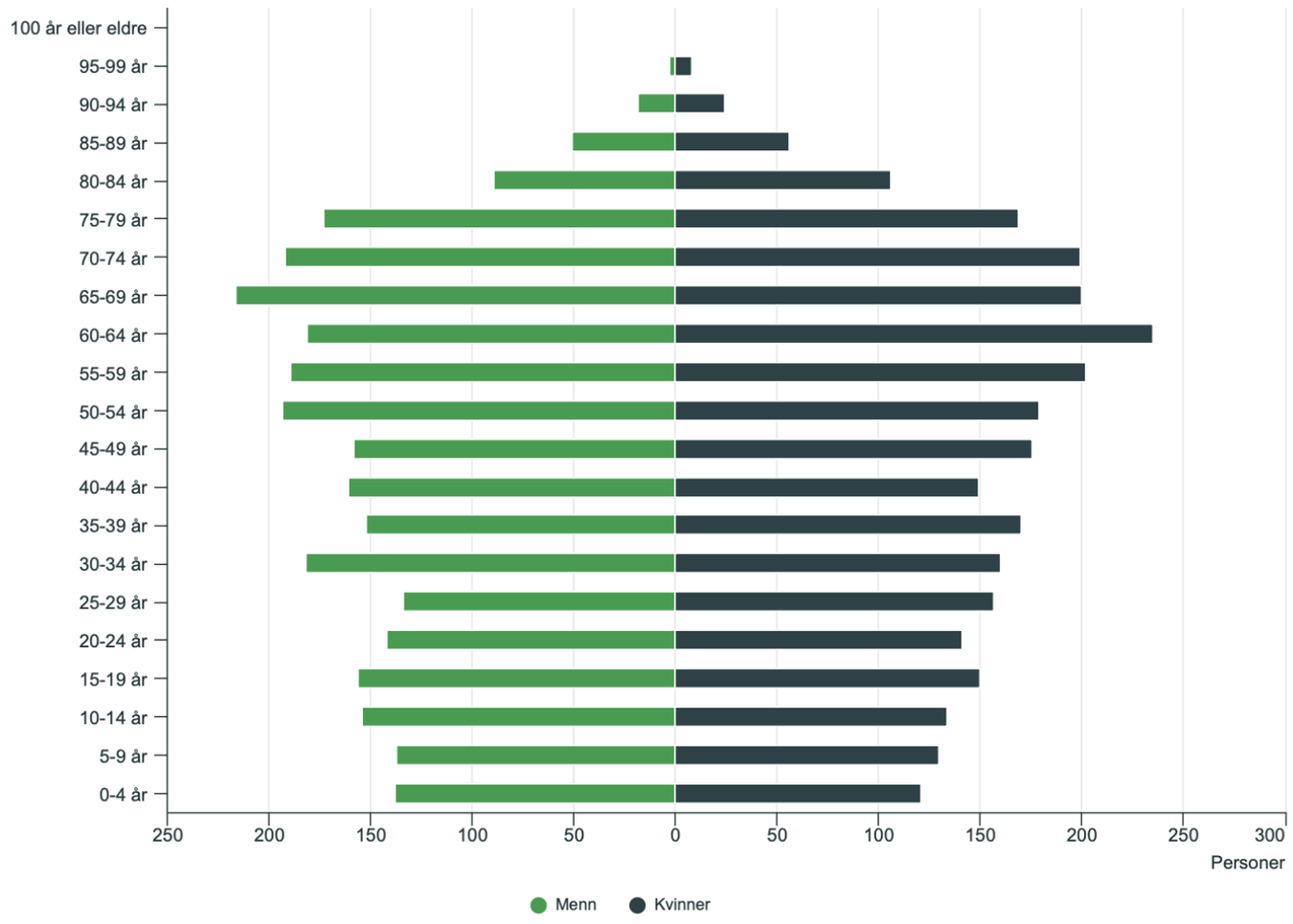


Hva er statistikk?

Trondheim

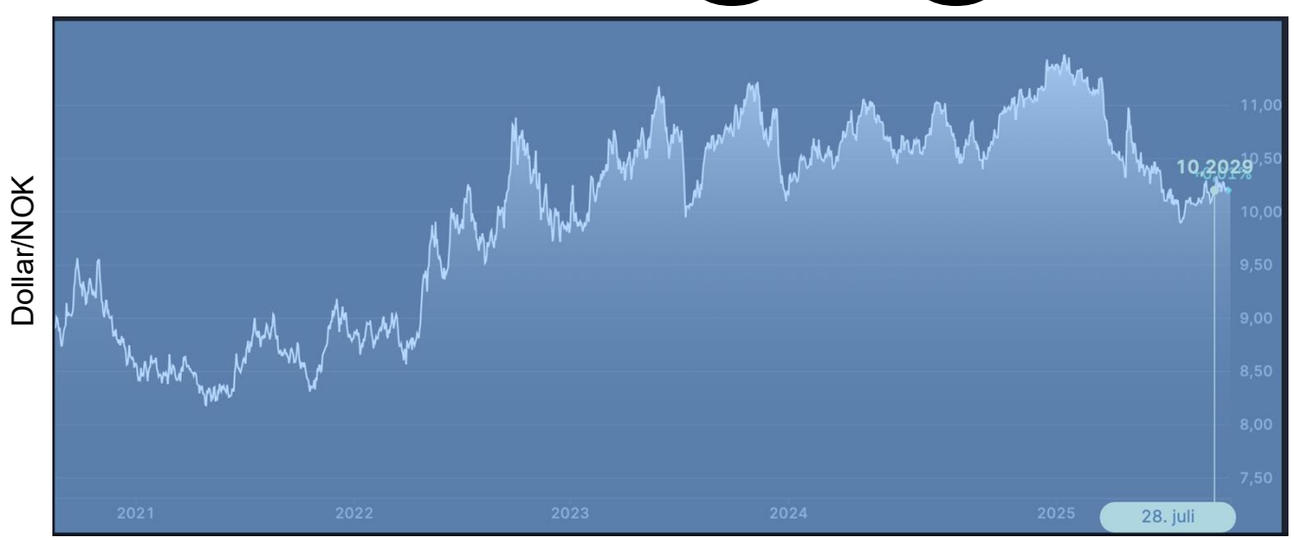


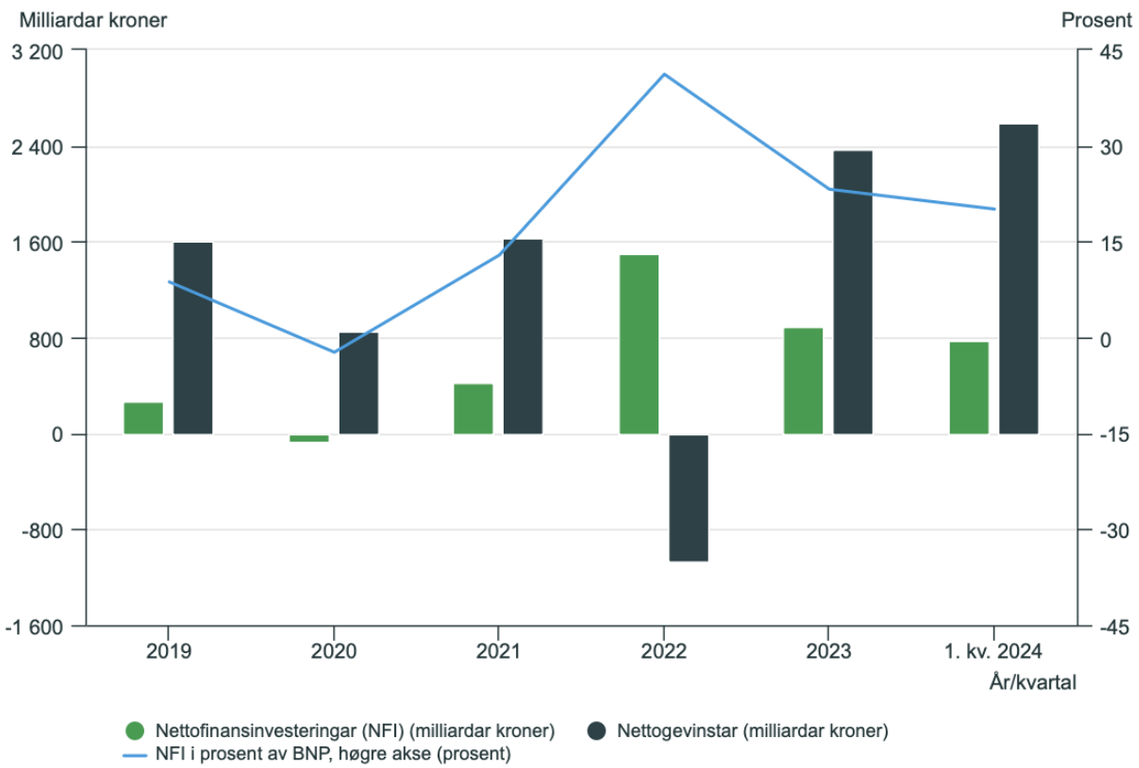
Røros

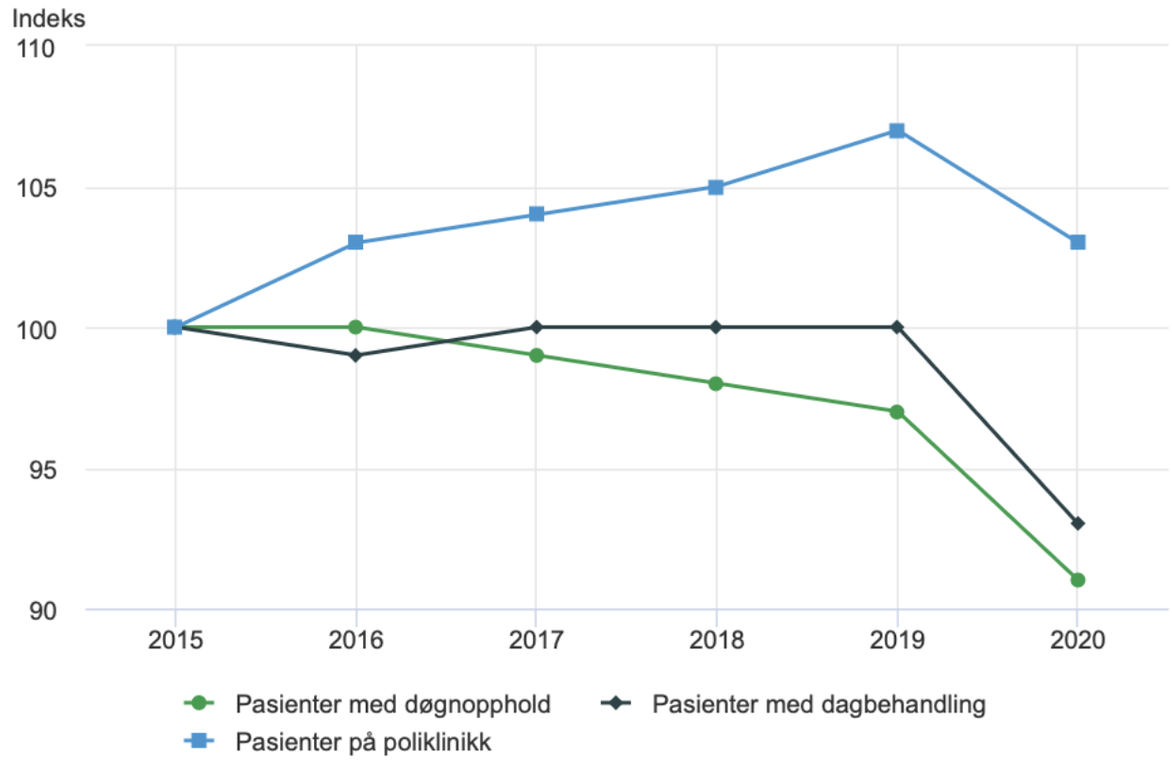


Alderspyramider – nyttig for

Tall og modeller for beslutningsgrunnlag







Hvordan oppsummere tall?

- Steg 1: **gjennomsnittet (mean)**.

Angir senteret som observasjonene (tallene) er spredt rundt.

10 data:

$$m=\bar{x}=\frac{1}{10}\sum_{i=1}^{10} x(i)$$

Gjennomsnitt, Hollywood-filmer

		Gjennomsnitt	
Sjanger	**Antall**	Budsjett (millioner $)	Billettsalg (millioner $)
Action	32	89,63	249,05
Animation	12	114,92	286,58
Comedy	27	38,50	107,53
Drama	21	25,33	44,63
Fantasy	2	62,60	664,72
Horror	17	25,79	73,23
Romance	11	38,40	135,95
Thriller	13	30,79	86,91
Alle	135	53,48	150,74

- **Median = den midterste observasjonen**
- En annen måte å angi senteret som observasjonene er spredt rundt

F.eks. en arbeidsplass med 10 ansatte

Alder: 20 21 21 22 22 22 23 23 66 67

Gjennomsnitt: 31

Median: 22

Empirisk varians

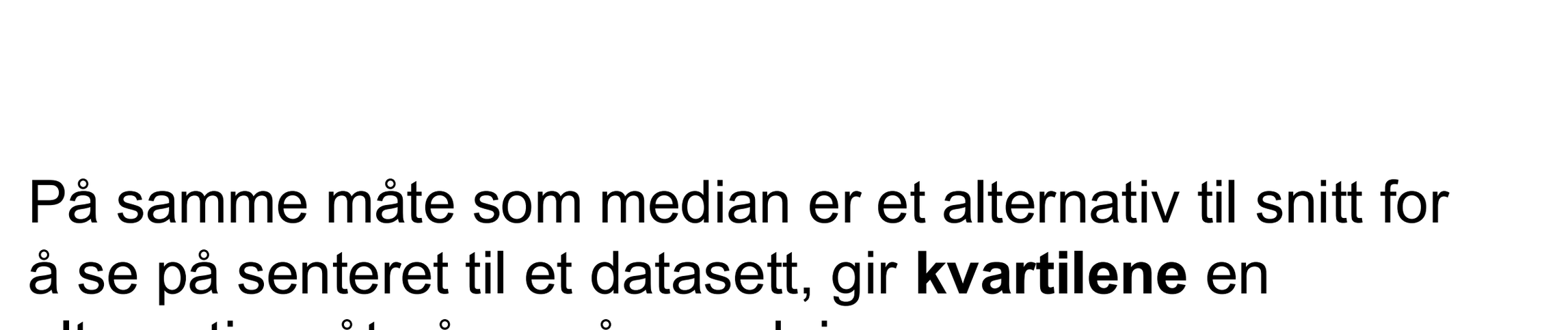
- Den **empiriske variansen** er et mål på hvor mye spredning det er rundt gjennomsnittet i et datasett

For observasjonene $x_1, x_2, ..., x_{10}$ så er den empiriske variansen:

$$s^2 = \frac{1}{10-1} \sum_{i=1}^{10} (x_i - \bar{x})^2$$

s kalles standardavvik.

- På samme måte som median er et alternativ til snitt for å se på senteret til et datasett, gir **kvartilene** en alternativ måte å se på spredninga



$$\mathrm{IQR} = Q_3 - Q_1 = 9 - 3 = 6$$

Levetiden til pattedyr

Table 2.14 *Longevity of mammals*

Species	Longevity	Species	Longevity	Species	Longevity
Baboon	20	Elephant	40	Mouse	3
Black bear	18	Elk	15	Opossum	1
Grizzly bear	25	Fox	7	Pig	10
Polar bear	20	Giraffe	10	Puma	12
Beaver	5	Goat	8	Rabbit	5
Buffalo	15	Gorilla	20	Rhinoceros	15
Camel	12	Guinea pig	4	Sea lion	12
Cat	12	Hippopotamus	25	Sheep	12
Chimpanzee	20	Horse	20	Squirrel	10
Chipmunk	6	Kangaroo	7	Tiger	16
Cow	15	Leopard	12	Wolf	5
Deer	8	Lion	15	Zebra	15
Dog	12	Monkey	15		
Donkey	12	Moose	12		

Table 2.14

© John Wiley & Sons, Inc. All rights reserved.

Levetiden til patte



© Britta Kasholm-Tengve/iStockphoto

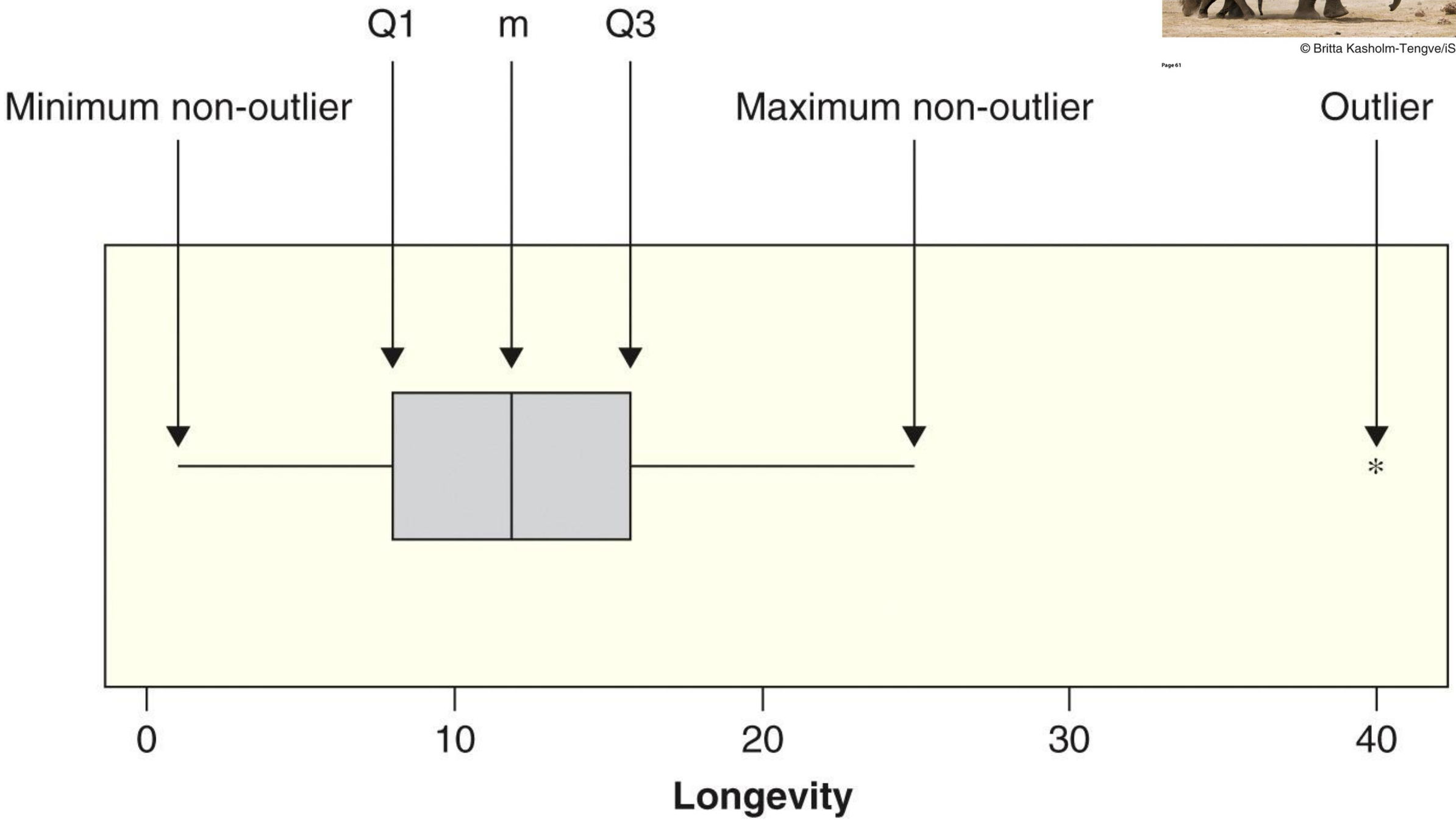


Figure 2.31

© John Wiley & Sons, Inc. All rights reserved.

The New York Times

Technology

For Today's Graduate, Just One Word: Statistics

TWITTER

LINKEDIN

COMMENTS (58)

E-MAIL

PRINT

REPRINTS

SHARE

THE SESSIONS
COMING SOON

MOUNTAIN VIEW, Calif. — At Harvard, Carrie Grimes majored in anthropology and archaeology and ventured to places like Honduras, where she studied Mayan settlement patterns by mapping where artifacts were found. But she was drawn to what she calls "all the computer and math stuff" that was part of the job.

Enlarge This Image



Thor Swift for The New York Times

Carrie Grimes, senior staff engineer at Google, uses statistical analysis of data to help improve the company's search engine.

"People think of field archaeology as Indiana Jones, but much of what you really do is data analysis," she said.

Now Ms. Grimes does a different kind of digging. She works at Google, where she uses statistical analysis of mounds of data to come up with ways to improve its search engine.

Ms. Grimes is an Internet-age statistician, one of many who are changing the image of the profession as a place for dronish number nerds. They are finding themselves increasingly in demand — and even cool.

"I keep saying that the sexy job in the next 10 years will be statisticians," said Hal Varian, chief economist at Google. "And I'm not kidding."

The rising stature of statisticians, who can earn $125,000 at top companies in their first year after getting a doctorate, is a byproduct of the recent explosion of digital data. In field after field, computing and the Web are

Multimedia



Graphic

Data Sleuths in an Internet Age

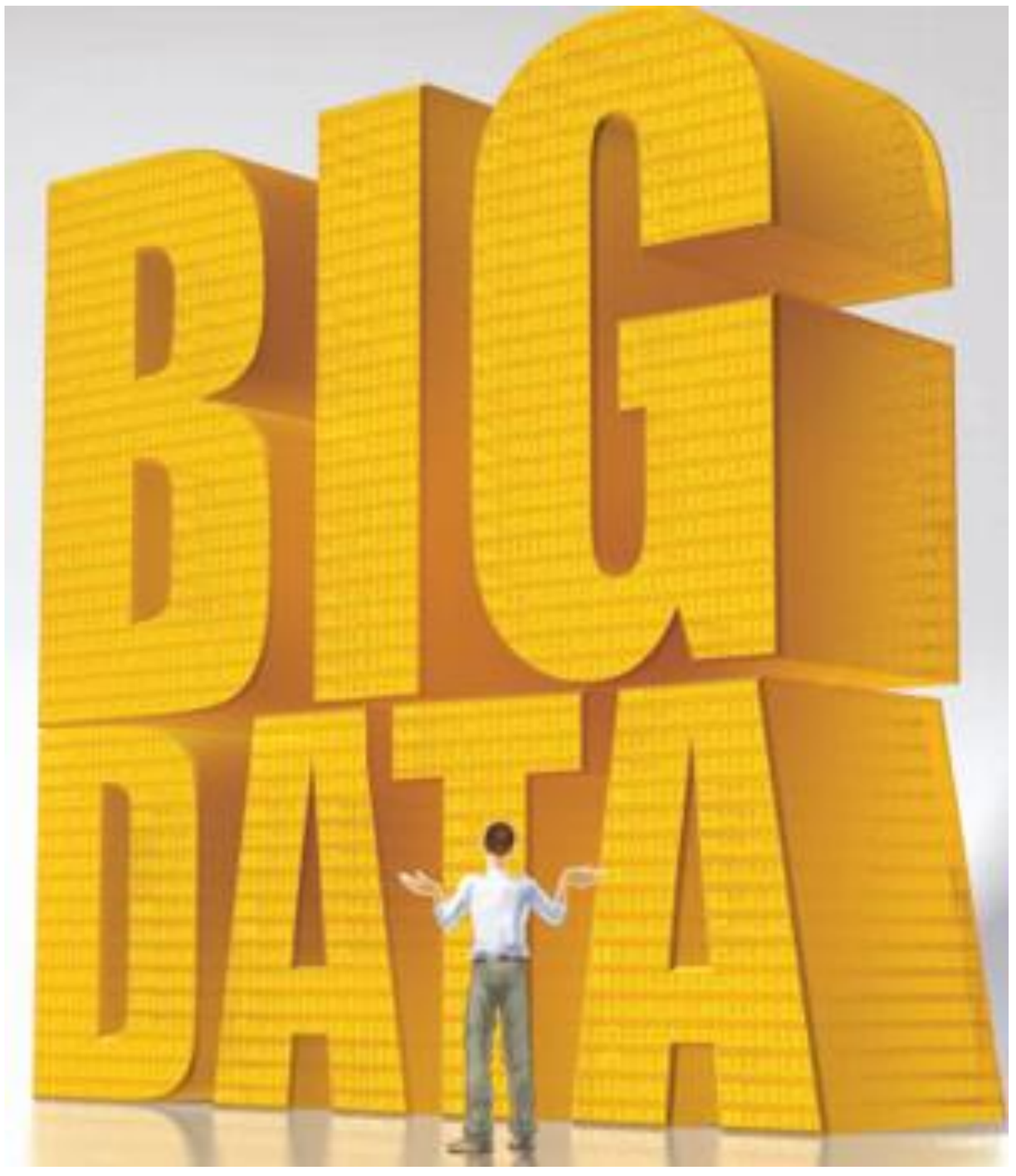
"I keep saying that the sexy job in the next years will be statisticians," says chief economist at Google.

"And I'm not kidding."



“We’re rapidly entering a world where everything can be monitored and measured. But the big problem is going to be the ability of humans to use, analyze and make sense of the data.”
–MIT





Nye muligheter fra data og modeller…

Hvilken film skal vi se?
Hvilken bok?
Hvilken sang?
Co-pilot,
språkmodeller,
Bildegenerering,







Statistisk maskinoversettelse

Maskinoversettelse

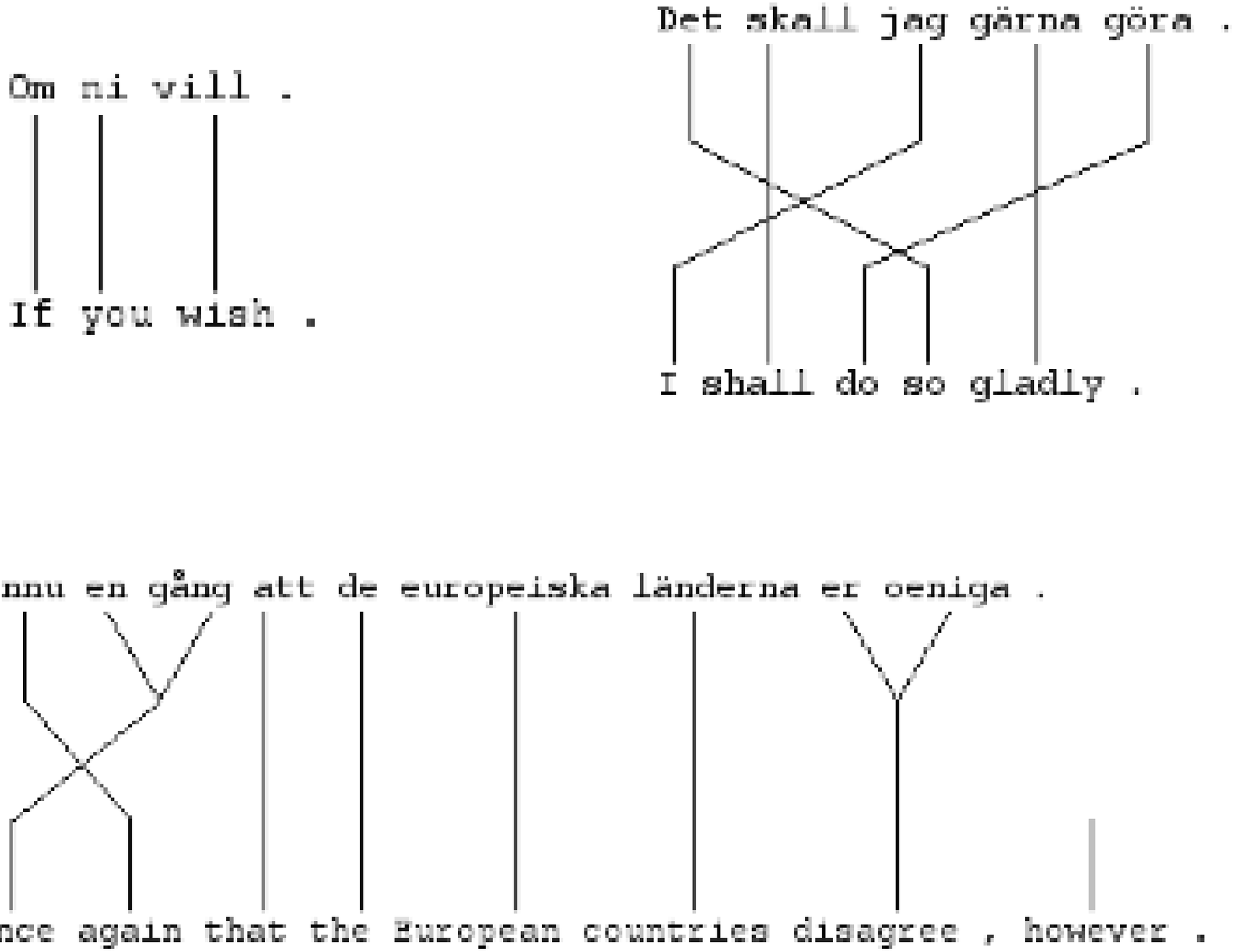
- Maskinoversettelse er automatisk oversettelse av tekst fra et språk til et annet
- Første forsøk på 1950-tallet, mellom engelsk og russisk
  - De første systemene var regelbaserte, og ganske dårlige
- På 1990-tallet begynte særlig IBM å forske på statistiske rammeverk
  - De siste årene har dette arbeidet ført til systemer som fungerer (særlig pga Google)
- før: mennesker prøvde å hardkode regler for alle mulige språk og variasjoner
- nå: maskiner prøver å lære sammenhenger mellom språk av seg selv

Parallell tekst

1 Resumption of the session
2 I declare resumed the session of the European Parliament adjourned on Friday 17 December 1999, and I would like once again to wish you a happy new year in the hope that you enjoyed a pleasant festive period .
3 Although , as you will have seen , the dreaded ' millennium bug ' failed to materialise , still the people in a number of countries suffered a series of natural disasters that truly were dreadful .
4 You have requested a debate on this subject in the course of the next few days , during this part - session .
5 In the meantime , I should like to observe a minute ' s silence , as a number of Members have requested , on behalf of all the victims concerned , particularly those of the terrible storms , in the various countries of the European Union .
6 Please rise , then , for this minute ' s silence .
7 (The House rose and observed a minute ' s silence)
8 Madam President , on a point of order .
9 You will be aware from the press and television that there have been a number of bomb explosions and killings in Sri Lanka .
10 One of the people assassinated very recently in Sri Lanka was Mr Kumar Ponnambalam , who had visited the European Parliament just a few months ago .
11 Would it be appropriate for you , Madam President , to write a letter to the Sri Lankan President expressing Parliament 's regret at his and the other violent deaths in Sri Lanka and urging her to do everything she possibly can to seek a peaceful reconciliation to a very difficult situation ?
12 Yes , Mr Evans , I feel an initiative of the type you have just suggested would be entirely appropriate .

1 Återupptagande av sessionen
2 Jag förklarar Europaparlamentets session återupptagen efter avbrottet den 17 december . Jag vill på nytt önska er ett gott nytt år och jag hoppas att ni haft en trevlig semester .
3 Som ni kunnat konstatera ägde "␣den␣stora␣år␣2000␣-␣buggen␣" aldrig rum . Däremot har invånarna i ett antal av våra medlemsländer drabbats av naturkatastrofer som verkligen varit förskräckliga .
4 Ni har begärt en debatt i ämnet under sammanträdesperiodens kommande dagar .
5 Till dess vill jag att vi , som ett antal kolleger begärt , håller en tyst minut för offren för bl.a. stormarna i de länder i Europeiska unionen som drabbats .
6 Jag ber er resa er för en tyst minut .
7 (Parlamentet höll en tyst minut) .
8 Fru talman ! Det gäller en ordningsfråga .
9 Ni känner till från media att det skett en rad bombexplosioner och mord i Sri Lanka .
10 En av de personer som mycket nyligen mördades i Sri Lanka var Kumar Ponnambalam , som besökte Europaparlamentet för bara några månader sedan .
11 Skulle det vara möjligt för er , fru talman , att skriva ett brev till den srilankesiska presidenten i vilket parlamentets beklagande uttrycks över hans och de övriga brutala dödsfallen i Sri Lanka och uppmanar henne att göra allt som står i hennes makt för att få en fredlig lösning på en mycket komplicerad situation ?
12 Ja , herr Evans , jag tror att ett initiativ i den riktning ni just föreslagit skulle vara mycket lämpligt .

Parallell tekst



Oversettelse ved Bayes regel

$$p(e \mid f) = \frac{p(f \mid e)\, p(e)}{p(f)}$$

Deler opp et vanskelig problem i to lettere deler:

- $P(\mathbf{f} \mid \mathbf{e})$ korrekthet (mening) : **oversettelsesmodellen**
- $P(\mathbf{e})$ velformethet (grammatikk) : **språkmodellen**

Mye data og bedre statistiske metoder gjør dette mye bedre enn før!


灭火瓶
HAND GRENADE


路安
滑全
TO TAKE NOTICE OF SAFE
THE SLIPPERY ARE VERY CRAFTY

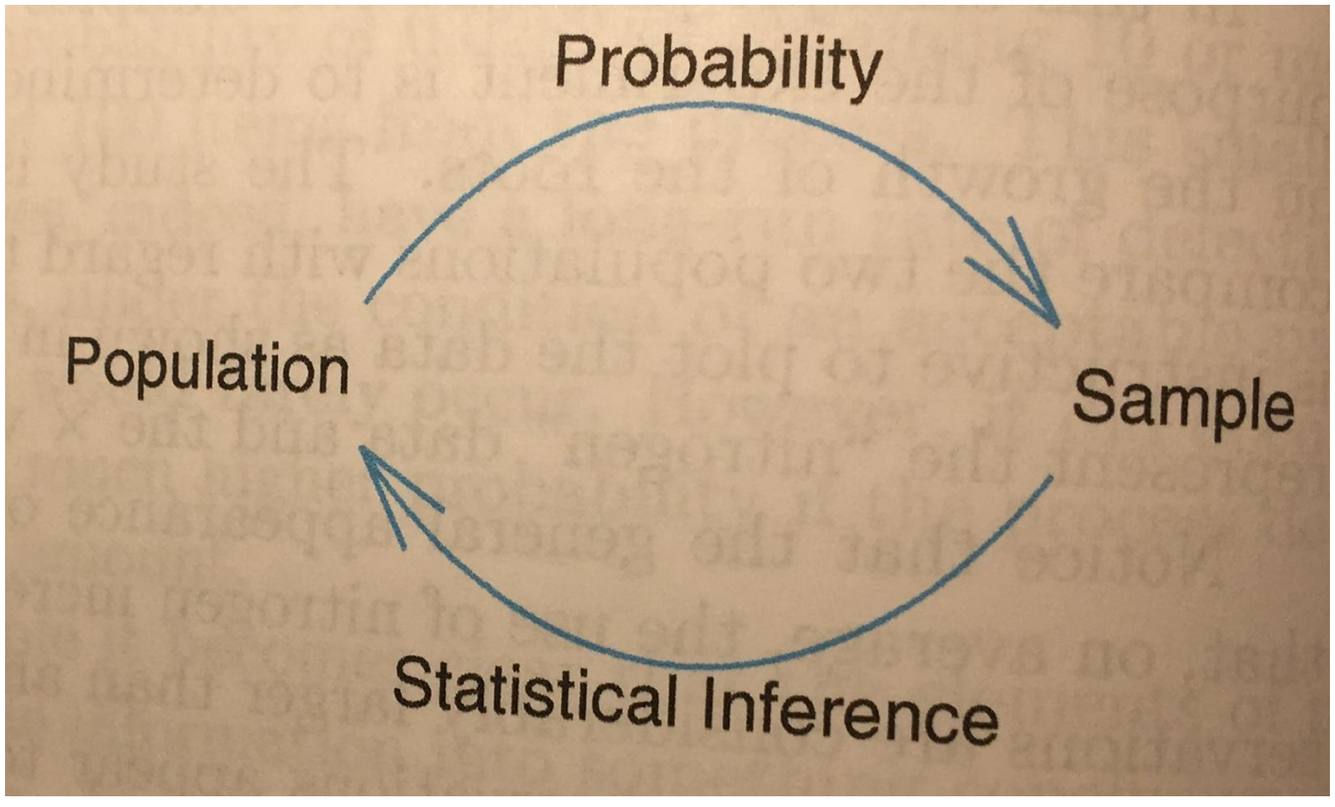

Qualifcation
WASH BY HAND 30°C
NO CHLORINE
NO WRING
IN COLESLAW
IRON 120°C
JIAN FANG ZHI HUO JIN PAO ZAI SH
LI GUO JIU


WARNING
THE "UNICO" TICKET IS NOT ISSUED ON THE TRAIN AND PASSENGERS ARE NOT ALLOWED ON TRAIN WHITHOUT IT.
WHEN THE TICKET OFFICE IS CLOSE, PASSENGERS CAN BUY THE TICKET FROM ANOTHER OUTLET.
THE "UNICO" TICKET, IS ONLY VALID IF OBLITERATED BEFORE GETTING ON THE TRAIN.
PASSENGERS WHO HAVE NOT OBLITERATED THEIR TICKETS WILL BE CONSIDERED WITHOUT T CKETS AND OBLIGED TO PAY PENALTY.


זהירות! מים עמוקים
סמוך לחוף
احذروا مياه عميقه بجانب الشاطى
noun area bathing
adjective forbidden

Sannsynlighet og statistikk

Kap 1, Walpole, Myers, Myers og Ye



Sannsynlighet og statistikk

Kobling muliggjør

- Usikkerhetskvantifisering
- Beslutninger under usikkerhet

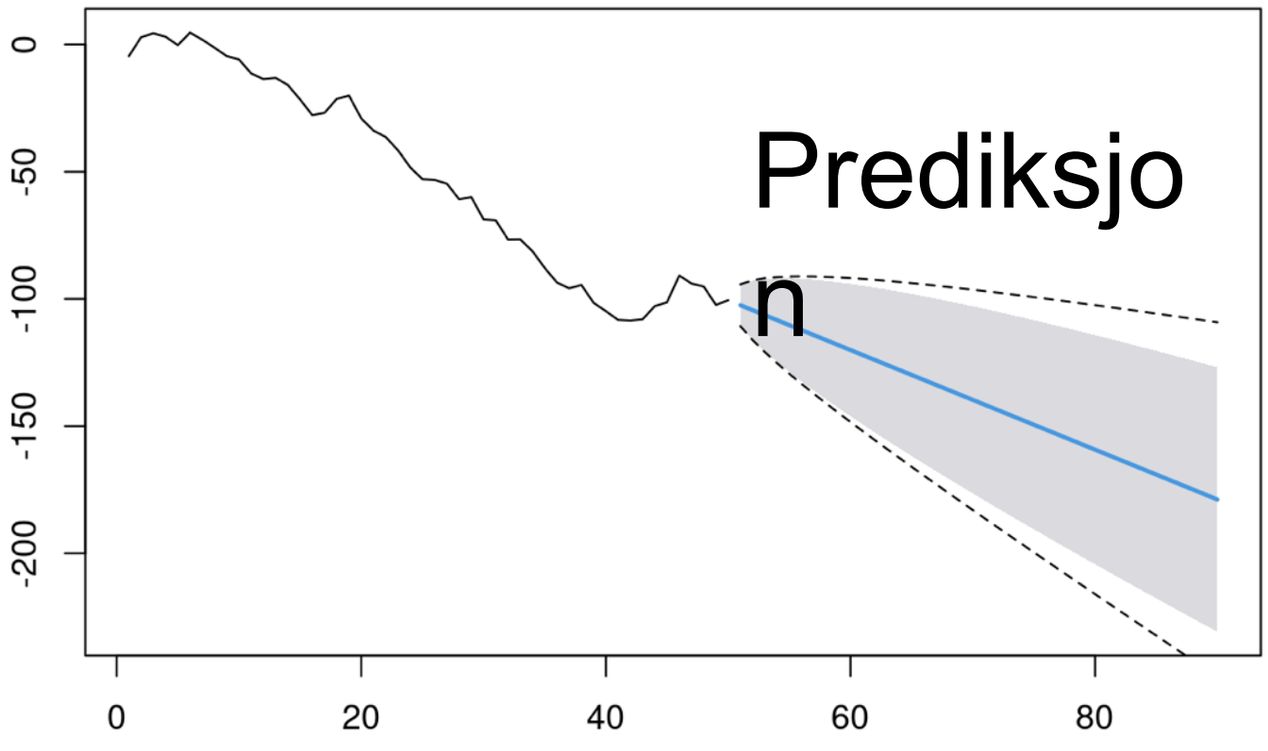


Table 1.1: Data Set for Example 1.2

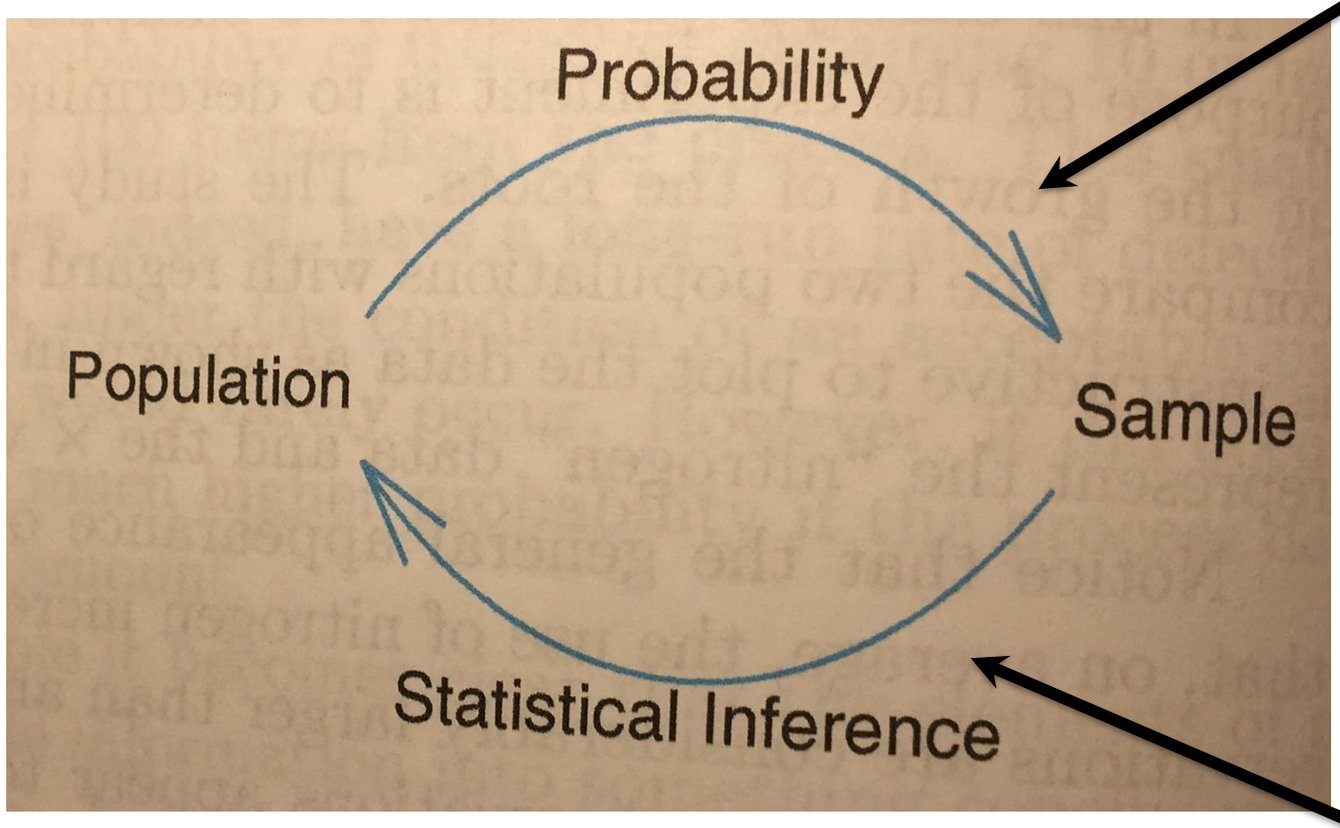
No Nitrogen	Nitrogen
0.32	0.26
0.53	0.43
0.28	0.47
0.37	0.49
0.47	0.52
0.43	0.75
0.36	0.79
0.42	0.86
0.38	0.62
0.43	0.46

Er den ene gruppen ‘bedre’ enn de
andre?

Det kan se slik ut, men det kan og
være et utslag av tilfeldig variasjon

Sannsynlighet og statistikk

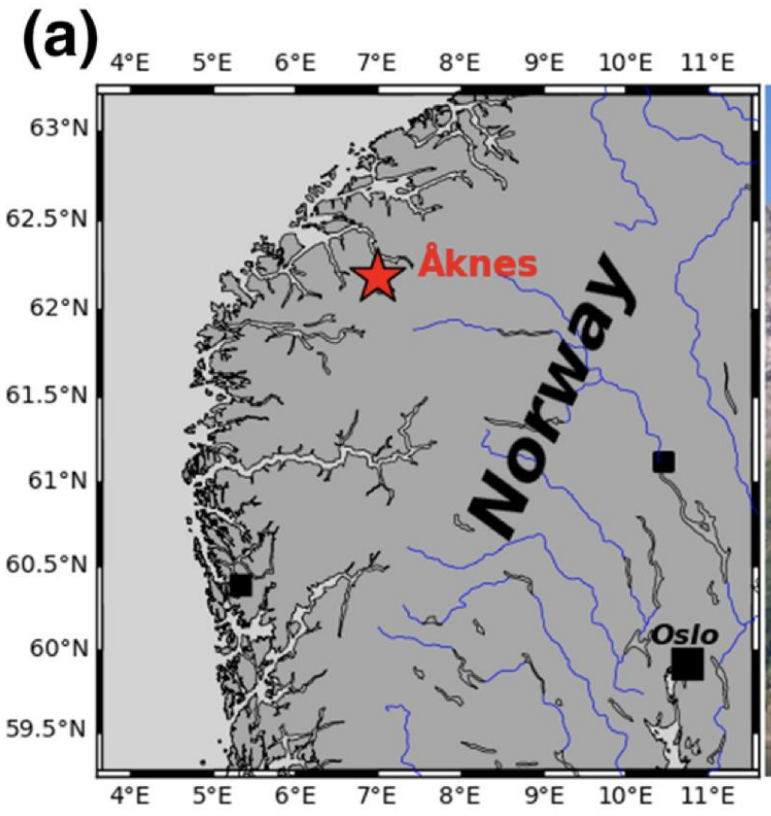
Første del
(6 uker).

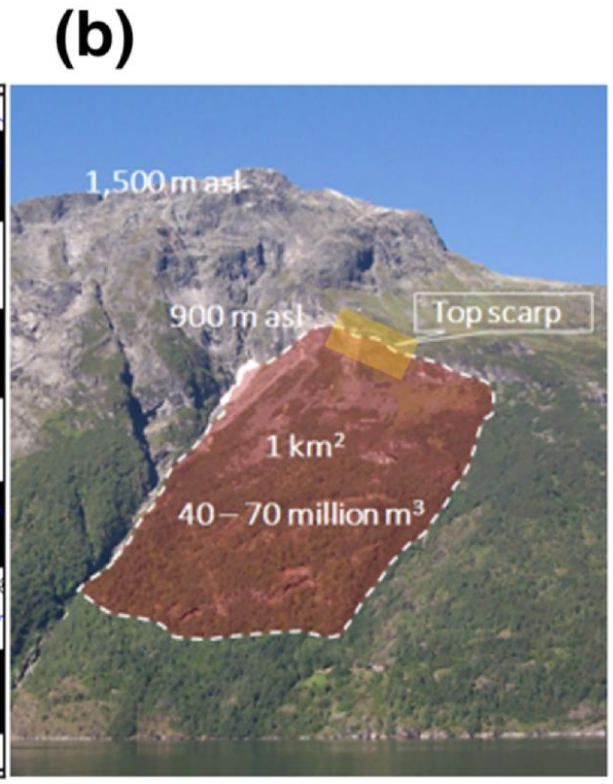


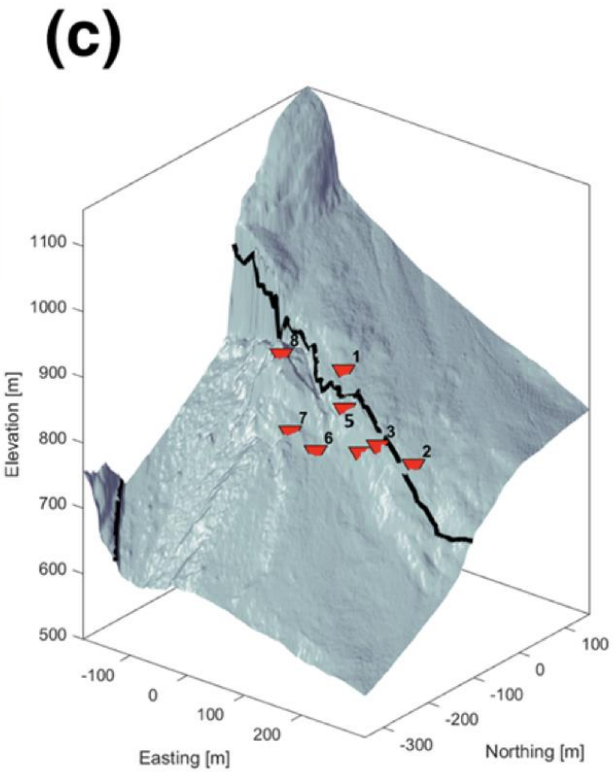
Siste del
(6 uker).

Noen ingeniør-anvendelser av sannsynlighetsregning og statistikk

Klassifisering av ulike typer ras ved Åknes



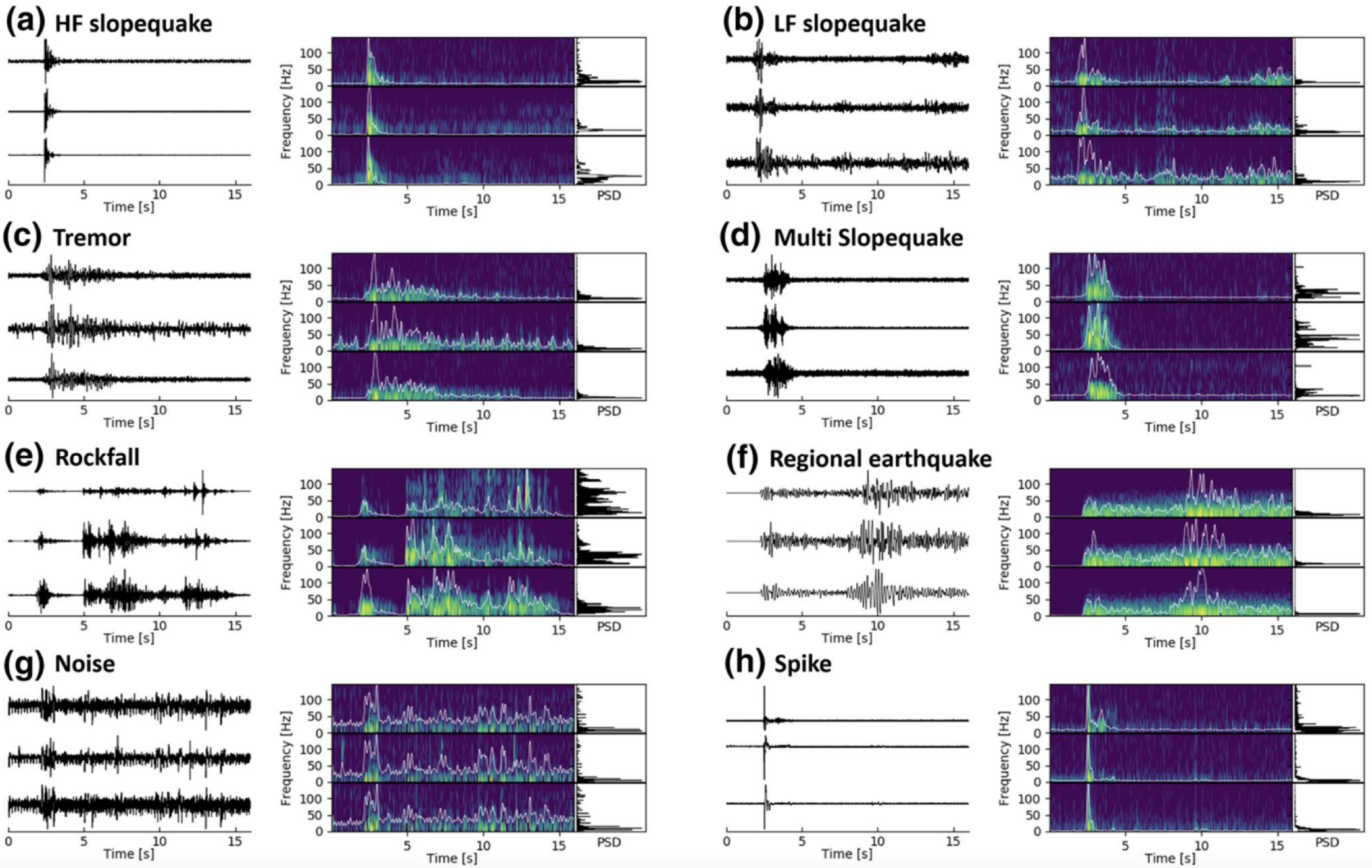




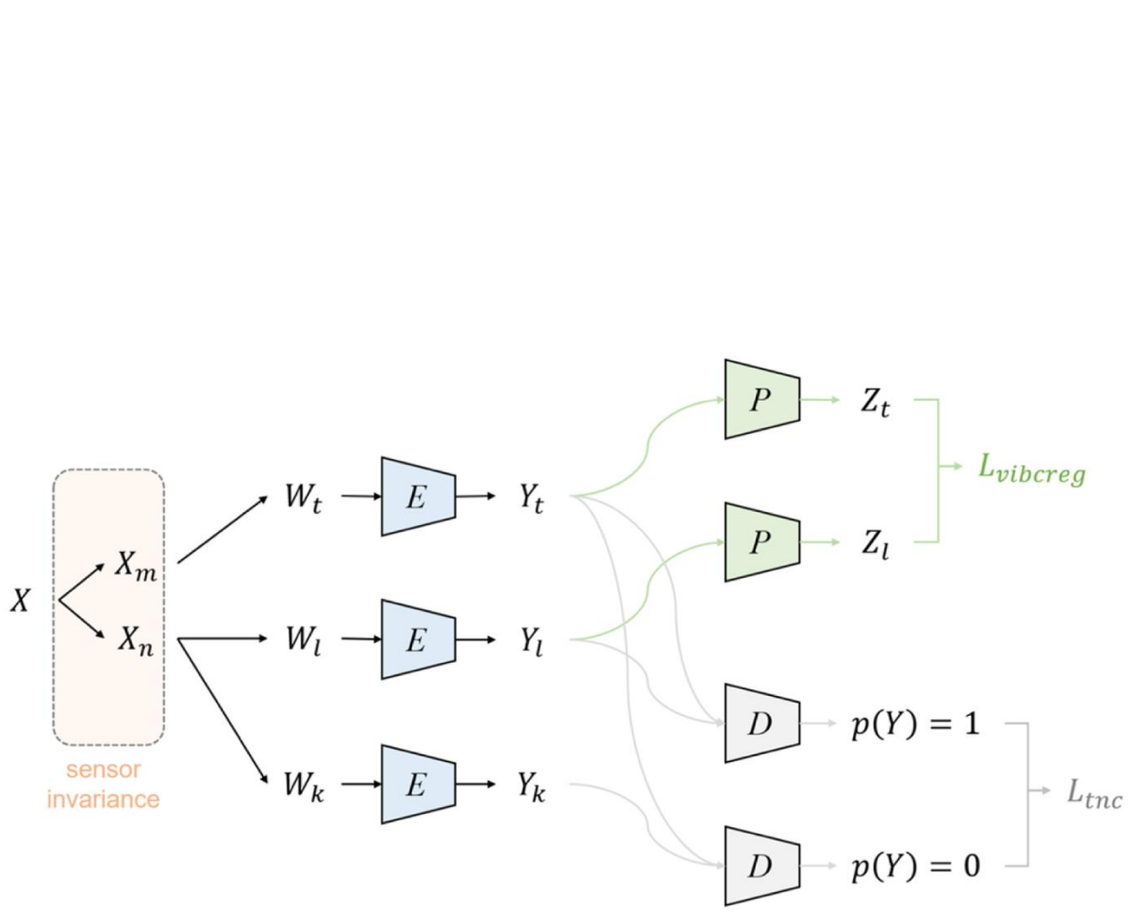


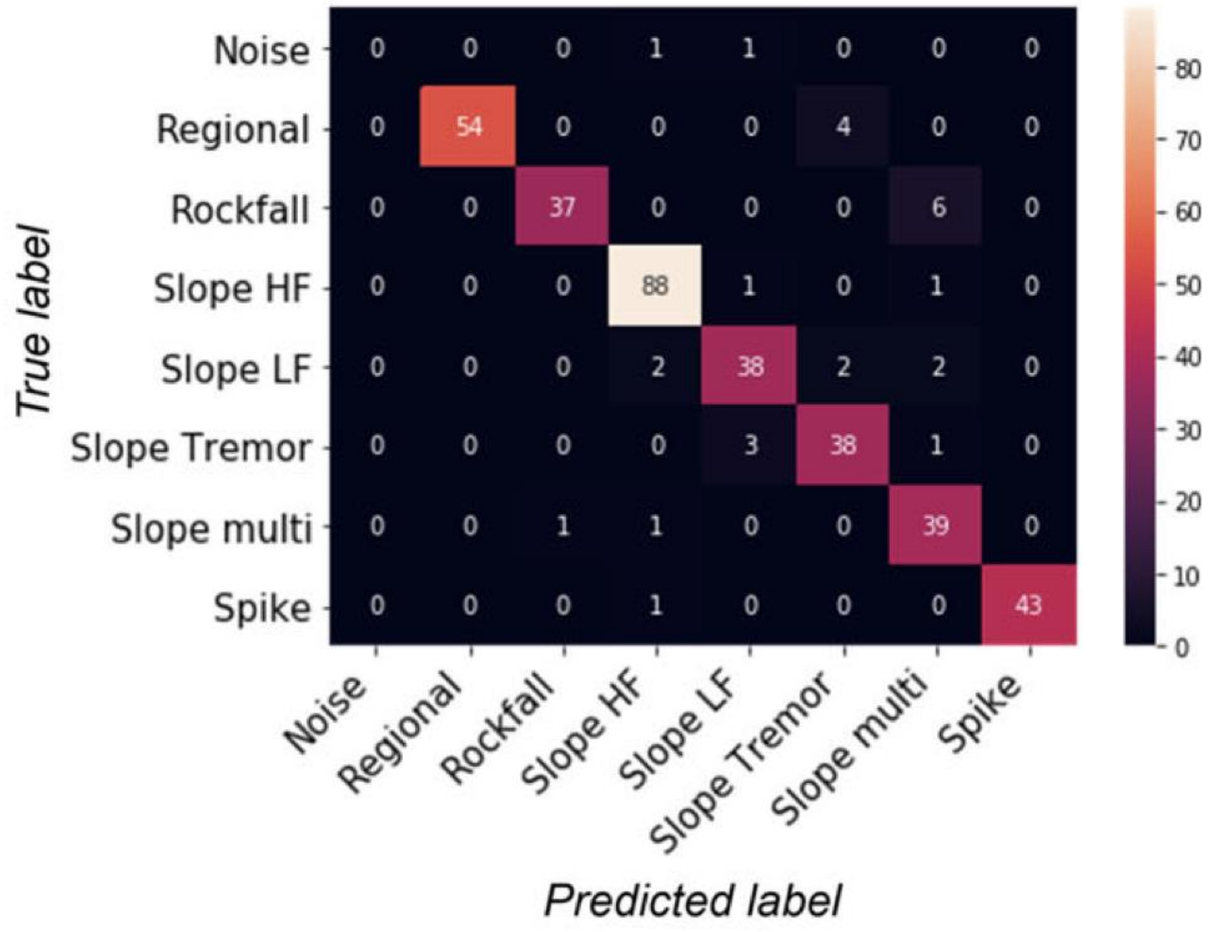
Bølgen

Klassifisering av ulike typer ras ved Åknes



Self-supervised neural networks klassifikasjon



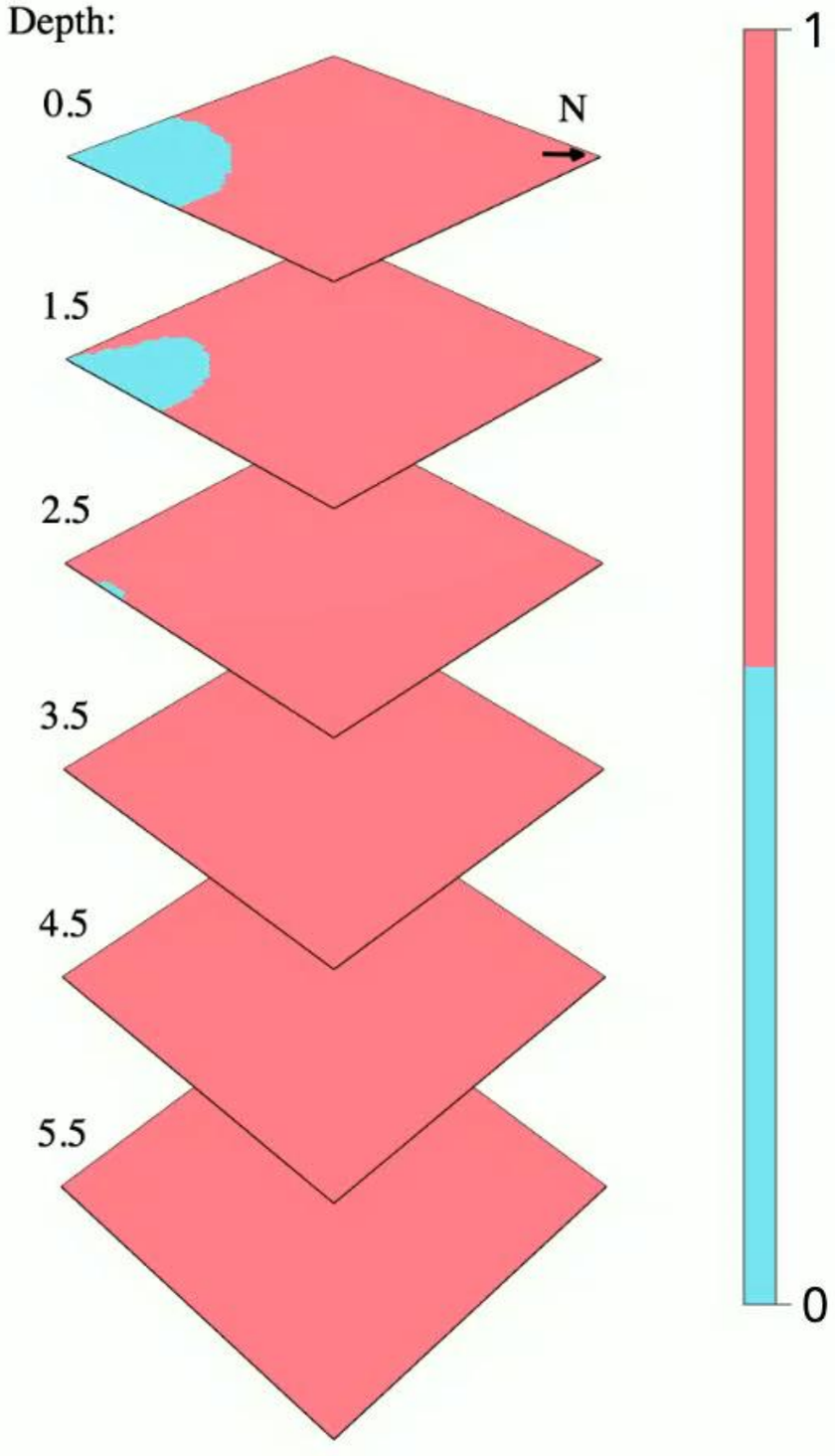


Monitoring fjord-river

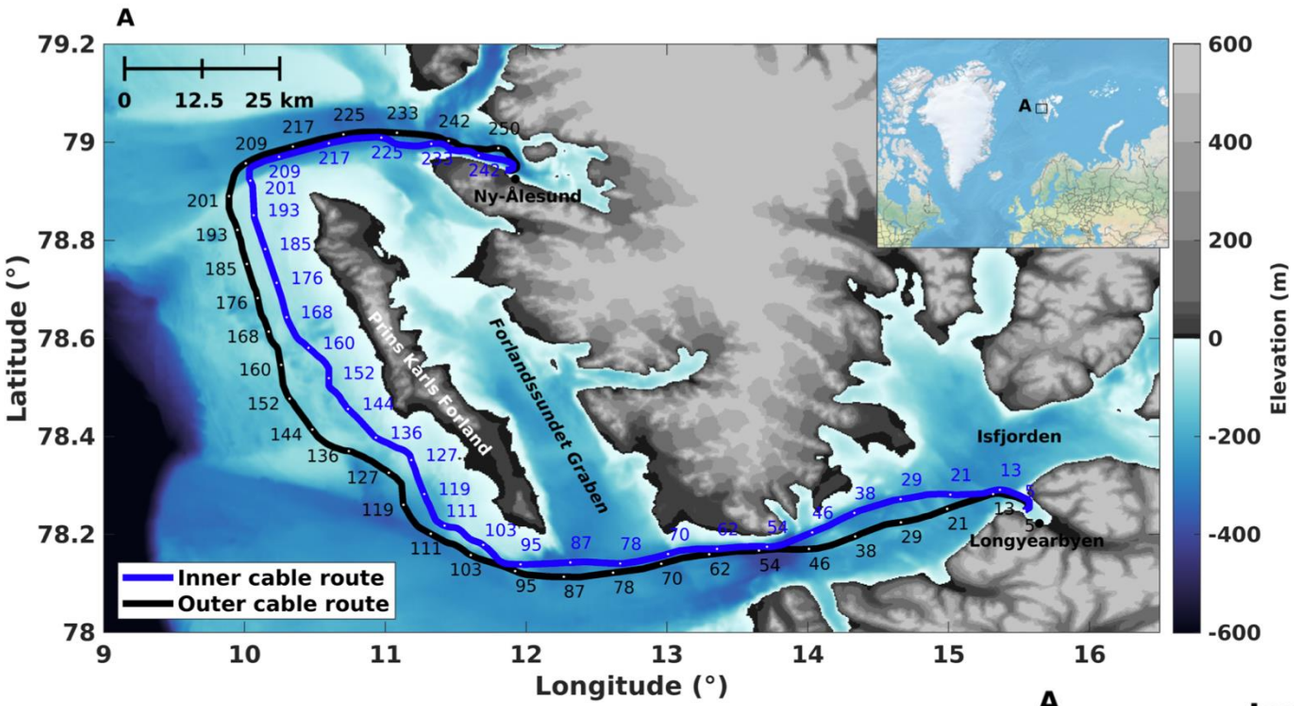
3D model tilpasset fra fysikk.

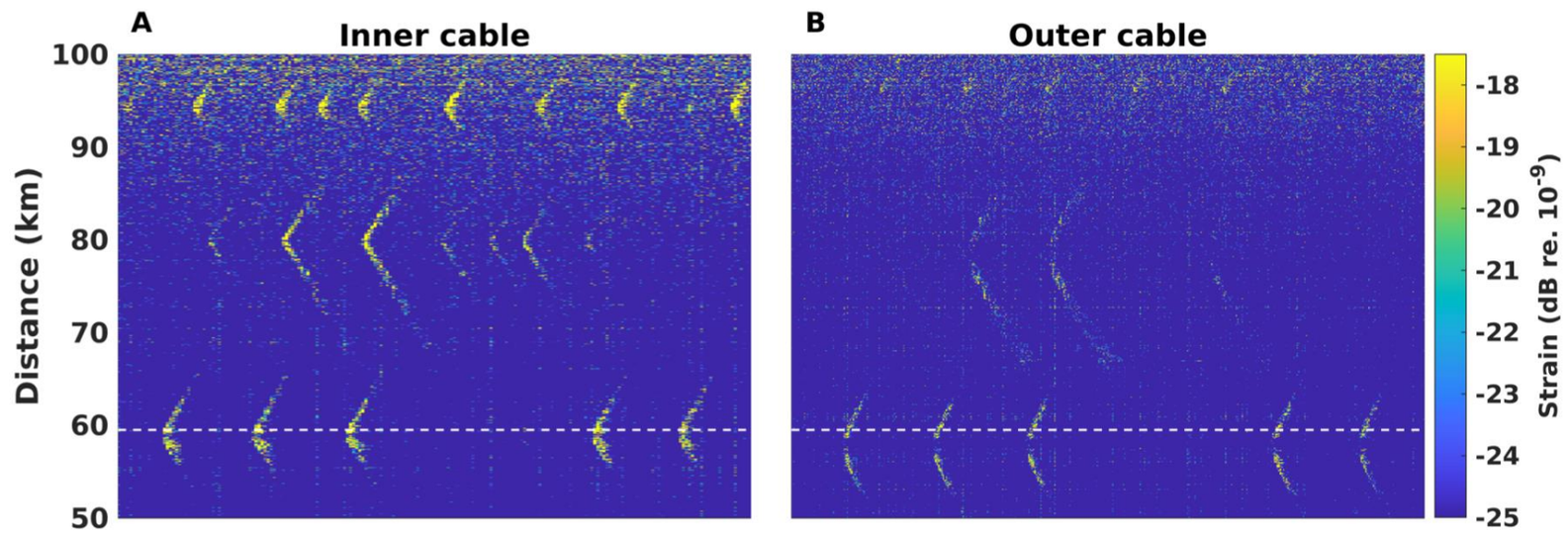
i) Modellen (Gaussisk prosess) oppdates med in-situ AUV data.
ii) Målet er at AUV-en skal kartlegge havvariasjon.

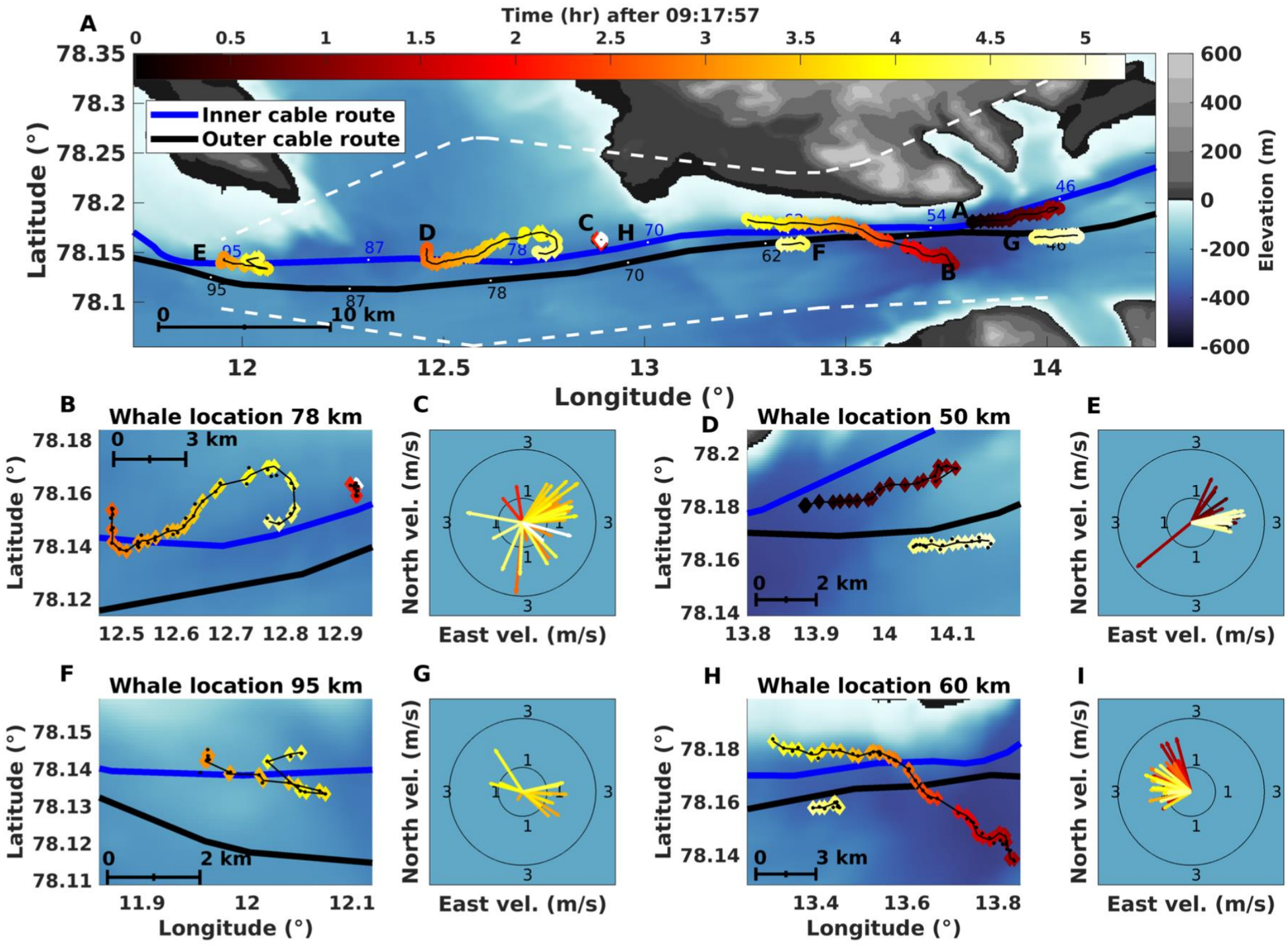




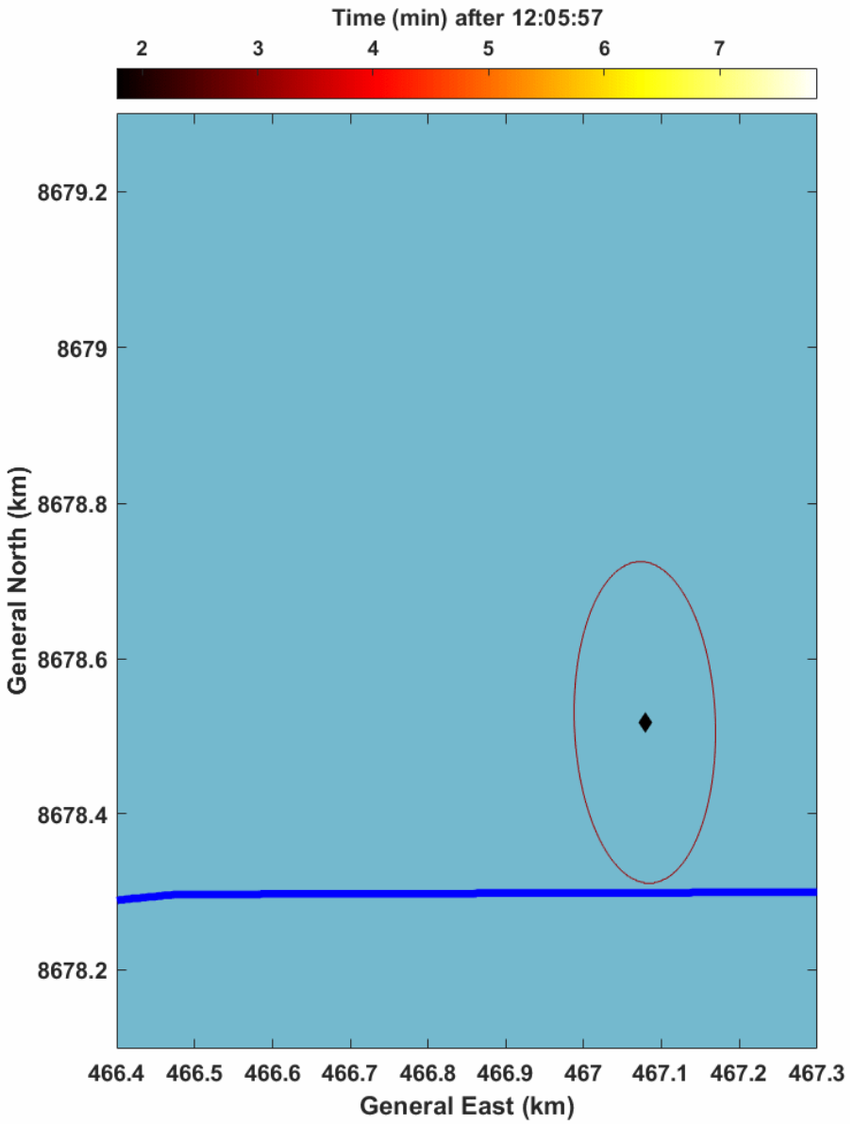
Fiberoptiske kabler og tracking av hvaler





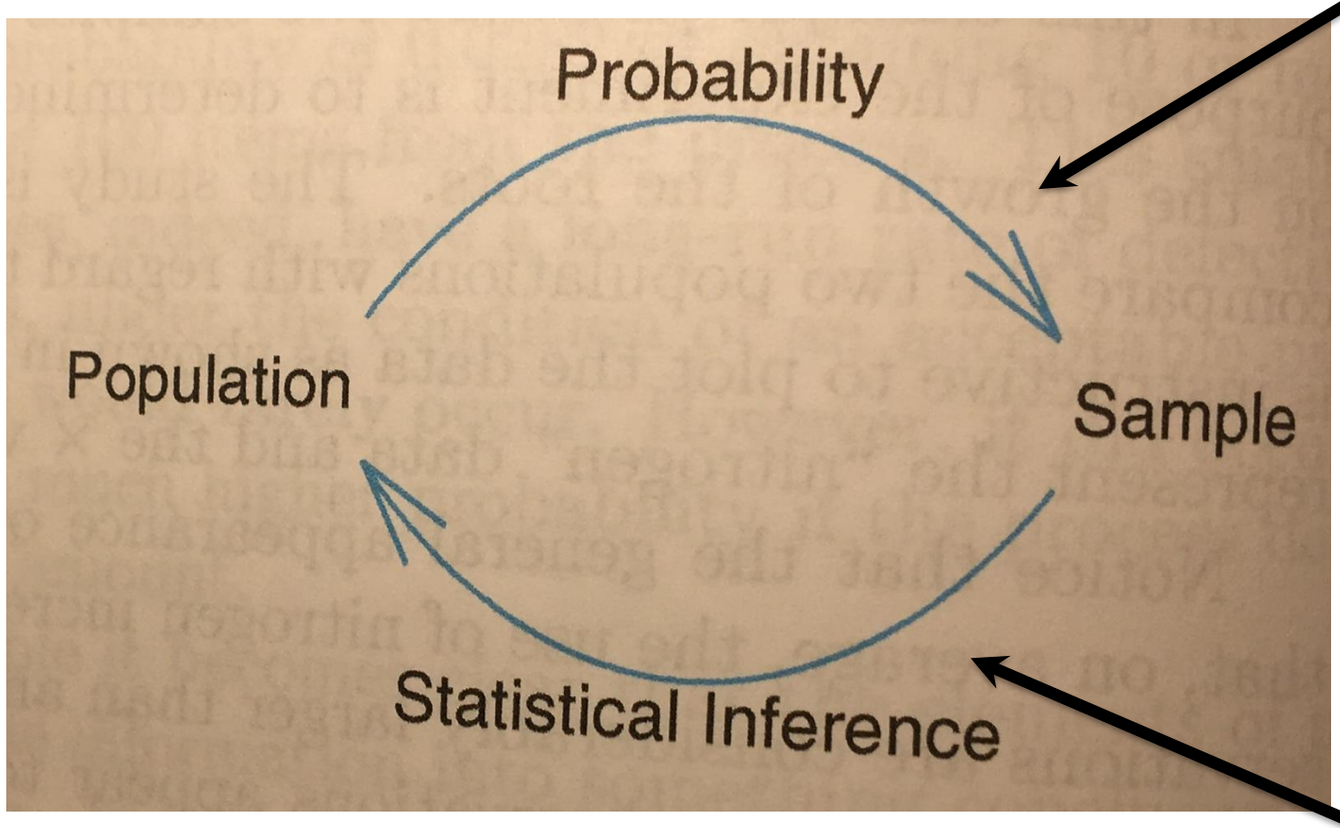

A
Time (hr) after 09:17:57
0
0.5
1
1.5
2
2.5
3
3.5
4
4.5
5
Inner cable route
Outer cable route
Latitude (°)
78.35
78.3
78.25
78.2
78.15
78.1
Elevation (m)
600
400
200
0
-200
-400
-600
E
D
C
H
F
A
G
B
95
87
78
70
62
54
46
0
10 km
12
12.5
13
13.5
14
Longitude (°)
B
Whale location 78 km
0
3 km
78.18
78.16
78.14
78.12
12.5
12.6
12.7
12.8
12.9
C
North vel. (m/s)
East vel. (m/s)
D
Whale location 50 km
78.2
78.18
78.16
78.14
0
2 km
13.8
13.9
14
14.1
E
F
Whale location 95 km
78.15
78.14
78.13
78.12
78.11
0
2 km
11.9
12
12.1
Longitude (°)
G
H
Whale location 60 km
78.18
78.16
78.14
0
3 km
13.4
13.6
13.8
Longitude (°)
I

Zoom-in på en hval over tid



Sannsynlighet og statistikk

Første del
(6 uker).



Siste del
(6 uker).