

FORBEREDELSE TIL GRUPPEEKSAMEN I MET4 VÅR 23

I dette dokumentet vil dere finne en innledning til problemstillingen som vi skal se på ved årets gruppeeksamen i MET4. Som vedlegg til dette dokumentet vil dere finne litt lesestoff som danner grunnlaget for problemstillingen, samt det eksakte datasettet som også blir utlevert med selve eksamensoppgavene 24. april 2023. Vi går også gjennom noen R-kommandoer som kan være nyttige for å jobbe med akkurat dette datasettet, samt en generell tilbakemelding på gruppeeksamen som ble gitt til hele kullet for noen år siden.

Introduksjon

Den 27. februar 2020 avsa Høyesterett endelig dom i et gruppesøksmål mot DNB Asset Management AS. Dommen markerte punktum for en lang prosess der Forbrukerrådet på vegne av 180 000 andelseiere i tre aksjefond forvaltet av DNB mente at forvaltningsgebyret som kundene betalte var for høyt, og de krevde prisavslag. DNB vant i Oslo Tingrett, mens saksøkerne fikk medhold i Borgarting Lagmannsrett. Høyesterett opprettholdt dommen fra Lagmannsretten, slik at DNB til slutt måtte tilbakebetale om lag 350 millioner kroner som andelseierene hadde betalt for mye i forvaltningsgebyr.

Bakgrunnen for søksmålet var at de aktuelle aksjefondene var markedsført som *aktive*. Det betyr at de ikke bare passivt følger markedet gjennom å etterligne for eksempel hovedindeksen på Oslo Børs, men at investeringene forvaltes aktivt i et forsøk på å oppnå bedre avkastning enn hovedindeksen, en jobb som de kan ta ekstra betalt for. Forbrukerrådet og andelseierene mente derimot at disse fondene *ikke* var aktivt forvaltet i praksis, men at de var alt for like hovedindeksen og passive fond – og at andelseierne dermed hadde betalt for en vare de ikke fikk.

Her kan vi lese hele dommen fra Høyesterett:

<https://lovdata.no/avgjorelse/hr-2020-475-a>

Her er en kort omtale av dommen fra NRK:

<https://www.nrk.no/norge/dnb-tapte-i-hoyesterett-1.14921899>

Tabell 1: Variabler i datasettet

Variabel	Beskrivelse
<code>month</code>	Måned for observasjon (1-12)
<code>year</code>	År for observasjon (2013-2017)
<code>date</code>	Dato for observasjon
<code>post_2014</code>	Dummyvariabel som tar verdien 1 for datoer etter 2014, 0 ellers
<code>fund_id</code>	Unik identifikasjon for hvert fond
<code>initial_closet</code>	Dummyvariabel som tar verdien 1 dersom fondet var et skapindeksfond ved starten av observasjonsperioden (<code>active_share</code> < 40), 0 ellers
<code>fund_size</code>	Kapital under forvaltning (millioner USD)
<code>fund_fee</code>	Månedlig forvaltningskostnad, gjort om til månedlig størrelse i prosent, oppdatert årlig. Det vil si: Kostnad inneværende måned som prosent av investert beløp
<code>net_flow</code>	Netto prosentvis vekst i kapital utenom avkastning
<code>fund_return</code>	Brutto (det vil si før omkostninger) prosentvis avkastning for fondet denne måneden
<code>alpha</code>	Differanse mellom fondsavkastning og benchmark (se under) før kostnader, i prosent månedlig
<code>active_share</code>	Mål på hvor mye fondet avviker fra benchmark i prosent

NHH-professorene Petter Bjerksund og Trond Døskeland fungerte som ekspertvitner for Forbrukerrådet i denne prosessen. De har i kjølvannet av rettssaken skrevet en forskningsartikkel om temaet *skapindeksfond* sammen med doktorgradskandidatene André Sjuve og Andreas Ørpetveit. Denne artikkelen er enda ikke publisert i et vitenskapelig tidsskrift, men en arbeidsversjon er vedlagt eksamensoppgavene (og ligger for øvrig åpent tilgjengelig på nettet¹). Utgangspunktet for artikkelen er at de vil undersøke hvorvidt (såkalte) aktive fond i de skandinaviske landene endret forvaltningsstrategi i forbindelse med at myndighetene gjennomførte undersøkelser rundt praksisen i 2014 og 2015. Dette gjør de ved å sammenligne skandinaviske fond med ikke-skandinaviske fond (som ikke ble utsatt for slike undersøkelser i like stor grad). Gjennom en “diff-in-diff”-analyse påviser de at de skandinaviske fondene hadde en statistisk signifikant økning i andel aktiv forvaltning etter 2015 som følge av myndighetenes undersøkelser sammenlignet med ikke-skandinaviske fond.

Vi har tilgang til den *skandinaviske delen* av datasettet som ble brukt i artikkelen i forbindelse med denne eksamensoppgaven. Det betyr at vi ikke har anledning til å gjøre den samme diff-in-diff-analysen med påfølgende kausal inferens, men vi kan likevel gjøre mange interessante undersøkelser om hvorvidt fondene økte sin aktivitet akkurat i tidsrommet der de ble undersøkt av myndighetene.

Last datasettet inn i R-minnet ved hjelp av følgende kommando, der vi antar at filen ligger i arbeidsmappen din:

```
load("heks_v23.Rdata")
```

Datasettet `funds` inneholder månedlige observasjoner av en rekke skandinaviske fond fra og med 2013 til og med 2017. Variablene i datasettet er beskrevet i Tabell 1.

I tabellen viser *benchmark* til et passivt forvaltet fond som følger indeksen. En høy verdi av `active_share` betyr at fondet avviker mye fra benchmarken, og er dermed aktivt forvaltet. Vi benytter i utgangspunktet 40% som grenseverdi for hva vi regner som et aktivt forvaltet fond, men dette kan problematiseres i besvarelsen dersom dere ønsker det.

Alle prosenttall er faktiske prosenter og ikke andeler: dersom `fund_fee` tar verdien 0.0788 betaler kunden 0.0788% av kapitalen i forvaltningsgebyr den måneden. Dersom verdien på `fund_return` er 6.089, hadde fondet en avkastning på 6.089% den måneden. Og så videre.

Noen nyttige R-funksjoner

Datasettet er et paneldatasett med en rad for hver kombinasjon av måned-år og fond. Denne strukturen gjør at vi ganske sikkert får nytte av `filter()`-funksjonen som vi så i R-introduksjonen vår. Vi kan for eksempel lage et underutvalg med bare fondene som var skapindeksfond ved starten av perioden:

¹https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3635718

```
library(dplyr)
funds %>% filter(initial_closet == TRUE)
```

Vi kan også slenge på en `skapindekser <-` på starten av den linjen dersom vi ønsker å lagre dette datasettet for senere analyse. Vi kan også bruke flere kriterier inne i samme `filter()`-kommando. Hvis vi ønsker å hente ut alle observasjonene som var gjort i januar (`month == 1`) 2023 (`year == 2013`), kan vi gjøre slik:

```
funds %>% filter(month == 1, year == 2013)
```

Når vi jobber med dette datasettet kan vi være interessert i å regne ut gjennomsnittet av en variabel innad i grupper. For eksempel vil vi kanskje regne ut gjennomsnittlig avkastning for fondene som var lite aktive i starten av perioden for seg, og gjennomsnittlig avkastning for resten av fondene for seg. Da kan vi bruke to svært nyttige funksjoner i `dplyr`-pakken; `group_by()` og `summarise()`. Den første av disse funksjonene lager en gruppering i datasettet som alle påfølgende funksjoner i pipe-sekvensen vil respektere. `summarise()`-funksjonen er designet for å lage deskriptiv statistikk for et datasett. La oss begynne med å demonstrere hvordan den fungerer. La oss si at vi ønsker å regne ut gjennomsnittlig månentlig avkastning for alle fondene i datasettet vårt, i tillegg til standardavvik, minimumverdi og maksimumverdi. Vi kan selvsagt kjøre `mean(funds$fund_return, na.rm = TRUE)`, `sd(funds$fund_return, na.rm = TRUE)`, og så videre hver for seg (Merk at det er enkelte verdier som mangler i datasettet, noe som betyr at vi må legge til argumentet `na.rm = TRUE` i disse funksjonene for at de skal ignorere de manglede verdiene. Hvis vi ikke hadde gjort det ville gjennomsnittet, standardavviket, min og max selv blitt manglende verdier NA).

Vi kan også få ut en fin tabell med den deskriptive statistikken ved hjelp av `summarise()`

```
funds %>% summarise(Gjennomsnitt = mean(fund_return, na.rm = TRUE),
                    Standardavvik = sd(fund_return, na.rm = TRUE),
                    Minimum = min(fund_return, na.rm = TRUE),
                    Maksimum = max(fund_return, na.rm = TRUE))
```

```
# A tibble: 1 x 4
  Gjennomsnitt Standardavvik Minimum Maksimum
    <dbl>         <dbl>    <dbl>    <dbl>
1      1.48         3.44   -13.0     15.7
```

Men la oss gjøre dette mer interessant. Vi ønsker å ta for oss kun observasjoner gjort i 2013, og så ønsker vi *en linje med deskriptiv statistikk for hver fondskategori*, en linje for de som hadde lavt aktivitetsnivå ved starten av perioden, og en annen linje for de som ikke var i denne kategorien. Dette kan vi nok klare manuelt ved å lage en tabell som over for hver kategori

(ved hjelp av `filter()`, for eksempel), men det er mye enklere hvis vi bare bruker `group_by()` til å fortelle at vi ønsker deskriptiv statistikk splittet opp på kategori:

```
funds %>%
  filter(year == 2013) %>%
  group_by(initial_closet) %>%
  summarise(Gjennomsnitt = mean(fund_return, na.rm = TRUE),
            Standardavvik = sd(fund_return, na.rm = TRUE),
            Minimum = min(fund_return, na.rm = TRUE),
            Maksimum = max(fund_return, na.rm = TRUE))
```

```
# A tibble: 2 x 5
  initial_closet Gjennomsnitt Standardavvik Minimum Maksimum
  <lgl>          <dbl>          <dbl>    <dbl>    <dbl>
1 FALSE          2.44          3.06    -6.39     11.3
2 TRUE           2.09          2.96    -5.62      8.20
```

Den første funksjonen i sekvensen over er grei, den bare filtrer ut observasjonene som ikke er gjort i 2013. Den neste funksjonen *grupperer* observasjonene etter verdien av variabelen `initial_closet`. Denne funksjonen forandrer ikke på selve innholdet i datasettet, den bare legger til metainformasjon om at alle funksjoner i pipe-sekvensen som kommer etterpå skal respekterer gruppetilhørigheten som den definerer. Med andre ord: alt som skjer i pipe-sekvensen etter `group_by(initial_closet)` skal skje gruppevis innen hver kategori. Deretter bare slenger vi på hva vi ønsker å gjøre, i dette tilfellet å lage et sammendrag av variabelen `fund_return` som over, men nå med en linje per kategori.

Vi kan gruppere på flere variabler også. Vi kan for eksempel se for oss et plott av *gjennomsnittlig månedlig avkastning per år, for hver kategori*. Her har vi to grupperingsvariabler (`initial_closet` og `year`), men det er ikke noe problem:

```
funds %>%
  group_by(initial_closet, year) %>%
  summarise(Gjennomsnitt = mean(fund_return, na.rm = TRUE),
            Standardavvik = sd(fund_return, na.rm = TRUE),
            Minimum = min(fund_return, na.rm = TRUE),
            Maksimum = max(fund_return, na.rm = TRUE))
```

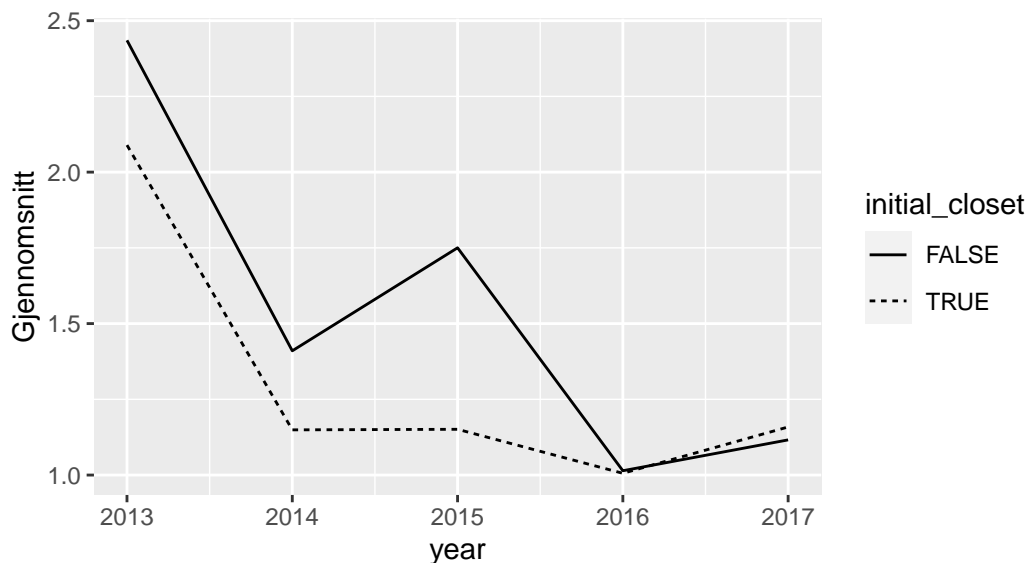
```
# A tibble: 10 x 6
# Groups:   initial_closet [2]
  initial_closet year Gjennomsnitt Standardavvik Minimum Maksimum
  <lgl>          <dbl>          <dbl>          <dbl>    <dbl>    <dbl>
```

1	FALSE	2013	2.44	3.06	-6.39	11.3
2	FALSE	2014	1.41	2.71	-7.04	10.7
3	FALSE	2015	1.75	4.56	-10.2	13.1
4	FALSE	2016	1.01	3.93	-13.0	15.7
5	FALSE	2017	1.12	2.86	-8.56	10.9
6	TRUE	2013	2.09	2.96	-5.62	8.20
7	TRUE	2014	1.15	2.29	-4.06	7.88
8	TRUE	2015	1.15	4.19	-7.50	10.5
9	TRUE	2016	1.01	3.51	-8.74	8.32
10	TRUE	2017	1.16	2.52	-7.14	7.56

Da får vi en lengre data frame med et gjennomsnitt per år, per kategori. Denne kan også sendes direkte videre til `ggplot`, der vi kan bruke for eksempel `linetype` eller `color` til å skille mellom kategoriene:

```
library(ggplot2)

funds %>%
  group_by(initial_closet, year) %>%
  summarise(Gjennomsnitt = mean(fund_return, na.rm = TRUE),
            Standardavvik = sd(fund_return, na.rm = TRUE),
            Minimum = min(fund_return, na.rm = TRUE),
            Maksimum = max(fund_return, na.rm = TRUE)) %>%
  ggplot(aes(x = year, y = Gjennomsnitt, linetype = initial_closet)) +
  geom_line()
```



Hvis vi ønsker å gjøre klassiske hypotesetester på dette datasettet så kan det være nyttig å huske kommandoen `pull` for å gjøre om en data frame med en kolonne til en vektor. Avkastningen til fondene i januar 2023 kan hentes ut som følger:

```
funds %>%  
  filter(month == 1, year == 2013) %>%  
  select(fund_return) %>%  
  pull
```

[1]	6.08939	5.54667	7.18305	6.12294	5.47210	6.52256	5.80180	5.94033
[9]	5.51255	4.85156	6.10399	5.54887	6.10279	6.04491	5.53272	6.49608
[17]	7.10685	5.81972	10.14133	5.75186	6.46599	4.24375	5.53660	5.71664
[25]	6.60428	3.10281	5.61963	4.36545	4.50213	6.16042	5.89403	4.76213
[33]	5.80116	5.53534	5.72908	5.20981	4.06396	5.12779	7.20667	3.26091
[41]	4.55055	5.87303	5.88760	1.52417	4.94444	2.42509	4.32543	3.44571
[49]	5.87371	6.16577	5.09902	7.09127	5.10143	5.01370	4.77189	3.55176
[57]	4.74691	4.61307	4.92057	4.73598	3.78062	NA	5.87097	7.00593
[65]	4.48932	2.32374	5.14663	4.65610	5.98860	5.36109	4.96101	4.28597
[73]	NA	NA	4.60967	6.18913	7.24428	5.55551	5.20687	4.56720
[81]	5.35004	5.06137	4.61310	4.46410	5.56712	4.95552	4.92543	5.79747
[89]	5.54670	5.41083	5.74418	5.61896	5.85997	6.04999	5.31770	7.54199
[97]	5.15574	5.78428	4.91416	4.97420	0.45170	5.75134	5.64653	5.63309
[105]	5.29144	5.60102	5.61317	5.14367	6.01479	5.38108	4.40538	3.86358
[113]	3.52754	6.64148	5.06090	5.00243	4.30982	4.31450	5.06239	5.53415
[121]	5.49241	2.41586	5.03570	3.50932	7.00832	4.63206	5.28512	3.81065
[129]	1.51875	1.25117	5.95658	5.22654	3.49870	4.81226	3.69903	5.38911
[137]	6.08309	4.65675	1.87828	4.98343	4.03348	6.21717	NA	3.63968
[145]	5.31726	4.86316	9.41173	6.22785	4.55453	3.97263		

(Her ser vi også at det er noen manglende verdier av denne variabelen).

Det kan også være aktuelt å *pare* observasjoner for å gjøre en statistisk test. Vi kan for eksempel hente ut avkastningen til alle fondene i januar 2023 som over:

Vi kan så gjøre det samme for å hente ut avkastningen i den siste måneden i observasjonsperioden (desember 2017). Men finnes det en måte å sette disse to data framene sammen igjen der vi er sikker på at observasjonene fra det samme fondet havner i samme rad? Ja, da kan vi bruke en funksjon som heter `full_join()`, også den i `dplyr`-pakken. La oss først lagre de to datasettene hver for seg:

```
df1 <-  
funds %>%
```

```

filter(month == 1, year == 2013) %>%
select(fund_id, fund_return)

df2 <-
  funds %>%
  filter(month == 12, year == 2017) %>%
  select(fund_id, fund_return)

```

Vi kan så *join*e disse sammen igjen ved å be `full_join()` bruke variabelen `fund_id` for å matche:

```

df3 <- full_join(df1, df2, by = "fund_id")
df3

```

```

# A tibble: 150 x 3
  fund_id      fund_return.x fund_return.y
  <chr>          <dbl>          <dbl>
1 F000YY123NZG      6.09            1.52
2 F00WA4731SLZ      5.55            1.44
3 F000DZ1937WFW      7.18           -1.40
4 F00SB5595TYG      6.12            0.135
5 F00VYD547UM4      5.47            1.74
6 F00EGM2027NWG      6.52             NA
7 F00FUY68PBG       5.80            0.649
8 F00BDY11505KF      5.94            0.911
9 F00SCK1491GFY      5.51            1.48
10 F00IEL651CF4      4.85            1.17
# ... with 140 more rows
# i Use `print(n = ...)` to see more rows

```

Her ser vi de to månedentlige avkastningene for hvert fond under hvert sitt variabelnavn – `full_join()` la til `.x` og `.y` for å skille dem fra hverandre.

Retteskjema

Vi gjør oppmerksom på at vi kommer til å bruke retteskjemaet som er lagt ut i fillageret på Canvas vil bli brukt til å bedømme besvarelsene.

Generell tilbakemelding etter Gruppeeksamen V21

Vi gav en generell tilbakemelding til kullet som tok gruppeeksamen i vårsemesteret 2021. Det kan være elementer der som er til hjelp også for dere inn mot eksamen. Tilbakemeldingen er gitt under:

Hei alle. Da er sensuren for den gruppebaserte hjemmeeksamen falt, og karakterene er publisert. Nivået ligger omtrent der det pleier med et snitt litt under B, som på alle måter må sies å være et godt resultat! En annen ting som er akkurat som det pleier er at ikke alle er fornøyd med karakteren sin. Vi har mottatt ganske mange forespørsler om begrunnelse, og de skal vi svare på i tur og orden, men jeg tenkte her å gi en mer generell tilbakemelding som alle kan dra nytte av i videre studier.

Forståelse av den spesifikke situasjonen:

Noe av det vi er mest fornøyd med i årets gruppeeksamen er at dere i stor grad klarte å vise forståelse for den konkrete situasjonen. Begrepet “customer analytics” er svært viktig i store deler av det private næringslivet, og det er på mange måter et åpent spørsmål hvordan vi kan dra nytte av kundedata for å forstå hvorfor kunder sier opp avtalene sine, og forhindre at det skjer. Dette handler om mer enn å bare sette opp en logistisk regresjon og se hvor godt den predikerer, det handler også om å forstå at det finnes viktige og ikke-trivielle avveininger av tilfredshet og kostnader knyttet til en kampanje rettet mot risikokunder. Videre fikk vi kanskje erfare at statistikken i dette tilfellet ikke gav oss en vidundermedisin for prediksjon, og at det ikke er lett å slå den mest naive løsningen: å bare la det være og ikke ringe til noen. Mange grupper viste god forståelse for dette gjennom nøktern analyse av de statistiske resultatene, og har blitt belønnet for det gjennom karakteren. Andre klarte i mindre grad å vise selvstendige vurderinger, og virket i større grad å gå på autopilot. Det kan være med å forklare en karakter som kanskje ikke stod til forventningene.

Kildekritikk:

Svært mange grupper presenterte følgende formulering av den logistiske regresjonsmodellen (eller en variant): $\log(y_i) = \beta_0 + \beta_1 X_1 + \dots + \beta_k X_k$. Dette er jo *galt*, og vi stusset lenge på hvor dette egentlig kom fra. Vi saumfarte videoer og forelesningsnotater i jakten etter synderen uten å finne en forklaring på at så mange grupper gjorde den identiske feilen. Besynderlig. Til slutt fant vi ut av det: eksempelbesvarelsen som fikk A har gjort akkurat det samme (rett før tabell 3). Ved nærmere ettersyn så viser det seg at flere grupper har lagt seg meget tett opp mot den besvarelsen, noe som kanskje er fristende gitt at eksamensoppgavene hadde store fellestrekk. Dette fremstår som svært lite sjarmerende for sensor, spesielt når man regelrett kopierer en feil. Hvis du føler deg truffet her, så vil jeg anbefale å ta med seg følgende læringspunkter i videre studier:

- Vi kommer ikke til å kjøre plagiatsaker på dette. Denne spesifikke feilen har heller ikke blitt brukt som eneste begrunnelse for å vippe noen ned fra en karakter til en annen. På samme måte som før er det totalinntrykket som til syvende og sist bestemmer karakteren.
- Det er ikke en unnskyldning at dette stod i en A-besvarelse; A betyr ikke perfekt, og i dette tilfellet kan det hende at feilen enten gikk under radaren til sensoren, eller så ble det satt på kontoen for “blingser” som vi selvsagt tillater i ellers sterke besvarelser.
- Vær kritisk til det du leser! Dobbelsjekk! Det er forskjell på kilder! Det finnes feil over alt! Dersom du skriver noe galt er det du som får svi for det, i alle fall dersom du ikke oppgir kilde.

- Lær deg bedre sitatskikk. En ting er å bli “inspirert”, men i min bunke var det flere besvarelser som absolutt burde referert til eksempelbesvarelsen, og som kanskje hadde måttet betale en mye høyere pris i et masterkurs for mangel på referanse. Ta dette som et vennlig råd på veien videre.

Deskriptiv statistikk:

Den deskriptive statistikken var under pari i år. Det var mange grupper som ikke egentlig forholdt seg til resten av oppgaven, som handlet om klassifisering. Da forventer vi for eksempel tydelige oversikter om hvordan variablene fordeler seg i de to klassene. Det var også muligheter til å finne noen rariteter i datasettet (i.e. negative priser) som burde bli tatt hånd om eller i det minste kommentert. En stargazer-tabell over alle variablene og et histogram eller to over noen av variablene fremstår som helt isolert fra resten av oppgaven hvis ikke er noen videre diskusjon som er relevant for det som kommer etterpå. Da har man misset poenget med deskriptiv statistikk, og må naturlig nok betale for det.

”Vi hadde alt rett, men fikk bare B (eller C)”:

Dere er nå i alle fall over halvveis i bacheloren, og noen er allerede ferdige og skal videre på master. Enten vi liker det eller ikke, så må vi innse at det er på høy tid å oppdatere hvilke kriterier som gjelder for å hevde seg i toppen av karakterhierarkiet. MET4 er i grunn et godt bilde på en viktig overgang: vi jobber i større grad med *virkelige* problemer. Desverre har virkelige problemer en lei tendens til å være mye *vanskeligere*, og kanskje ikke på den måten at vi alltid trenger kompliserte modeller og lange formler. Nei, det er bare slik at den virkelige verden er mye mer nyansert enn fasiten bakerst i læreboken. Det finnes mange måter å angripe et problem på, men kanskje er det ingen måte som løser problemet fullstendig og for evig og alltid. Det er alltid noen nyanser som vi ikke klarer å ta hensyn til, eller noen avveininger som gjør at vi kan komme et stykke på vei langs en dimensjon, men kanskje på bekostning av at et annet problem blir verre å hanskes med. Med andre ord: det handler ikke bare om å regne riktig (selv om det er viktig, for all del), men om å jobbe seg frem til et svar som er så godt som mulig, som hjelper oss best mulig, og der vi samtidig har kontroll på hva vi ikke har klart å løse og hvilke konsekvenser det har.

Løsningsforslaget er ikke fullstendig. Det mangler fullstendig prosaen, argumentasjon for og mot, tolkning og diskusjon som trengs for å nå helt opp. Det inneholder noen kodelinjer som kan brukes som grunnlag for diskusjonen.

Universitets- og Høgskolerådet har generelle beskrivelser av karakterer for [BØA-studiet \(klikk for link\)](#) som kan være nyttige å kikke på. Blant annet ser vi der at man skal vise stor grad av selvstendighet for å få en A, og for B skal selvstendigheten og vurderingsevnen være meget god. De som lykkes godt i denne eksamensformen klarer nettopp dette, og det er klart at listen for de høyeste karakterene ligger høyere enn at man ikke har gjort direkte feil et sted. Dette er helt i tråd med [Bloom’s Taxonomy \(klik for link\)](#) som viser at “remember” og “apply” er lavere læringsnivå enn for eksempel “analyze” og “evaluate”. Jeg tror at ganske mange med stor fordel bør rekalkibrere hva som trengs for karakteruttelling i de videre studiene.

Noen leverer på et helt vilt høyt nivå

Jeg må avslutte med skryt! Det er noen grupper som leverer besvarelser som er helt eksepsjonelt gode! Vi har flere eksempler på besvarelser som analyserer problemstillingen på en nær profesjonell måte, som kommer frem til innsikter og konklusjoner basert på datasettet som er nye og genuint interesserante, og som viser mye kreativitet og initiativ. Det er åpenbart at mange har brukt tiden godt gjennom semesteret og opparbeidet seg en sterk statistisk intuisjon som har kommet godt med på eksamen. Toppnivået på NHH er ulikt alt jeg har sett som ekstern sensor ved andre institusjoner her i landet.