

# R Notebooks og reproduserbarhet

## Innhold

<b>Innledning</b>	<b>1</b>
Litteraturgjennomgang . . . . .	1
Analyse . . . . .	2
Konklusjon . . . . .	3
Referanseliste . . . . .	3
Appendiks . . . . .	3

## Innledning

Vitenskapen går fremover på oppdagelser og funn som vi kan ha tillit til. En viktig møte å ha tillit til disse funnene er at det er mulig å reprodusere resultatene til noe andre har funnet. Ofte er det slik at det er problemer med reproduserbarhet på mange av forskningsartiklene som blir lagt ut i forskningstidsskrifter (McNutt, 2014).

Reproduserbarhet kan deles inn i tre ulike begreper. Det er Goodman et al. som har beskrevet begreperne, og har de har delt dem inn i **metode reproduserbarhet**, *resultat reproduserbarhet* og **robusthet og generaliserbart** (Goodman et al. (2016)).

- Moderne reproduserbarhet referer til å gi tilstrekkelig detalj om studieprosedyrer og data slik at de samme prosedyrene kan nøyaktig gjentas, både i teorien og i virkeligheten (Goodman et al. (2016)).
- Resultat reproduserbarhet referer til å oppnå de samme resultatene fra gjennomføringen av en uavhengig undersøkelse, dersom prosedyrer er så nært tilpasset det opprinnelige eksperimentet som mulig (Goodman et al. (2016)).
- Robusthet referer til stabiliteten i eksperimentelle konklusjoner som har variasjoner i enten antagelser eller eksperimentelle prosedyrer. Det er relatert til begrepet generaliserbarhet, som referer til vedvarheten av en effekt i innstillinger som er forskjellige fra et eksperimentelt rammeverk Goodman et al. (2016)

## Litteraturgjennomgang

Dewald et al. (1986) prøvde å replikere artikler de fant i “journal of money, credit and banking.” av 54 artikler greide de å replikere 2 stykker. Problemet var at de hadde for lite info om arbeidet bak funnene deres.

Etter disse funnene ble det positivt innstilt på at økonomiske artikler skulle lage arkiv så det gikk an og reprodusere og replikere artiklene. McCullough et al. (2008) prøvde å replikere 186 emperiske artikler, hvor kun 69 hadde arkiver. 7 av disse hadde ikke mulighet på grunn av datavare. resterende 62 kunne 14 replikeres, som er en forbedring fra 2 av 54. De fleste artikler hadde ikke arkiv på grunn av null kontroller av at reglene faktisk ble overholdt.

Gentleman and Temple Lang foreslår en metode og en basis infrastruktur for å brukes til offentlige dokumenter/artikler på en substansiell måte for å løse problemet med reproduserbarhet. Metoden egner seg best for artikler som krever bruk av databruk. De foreslo en måte der hvor du kombinerer tekst og kode i samme dokument. Disse dokumentene kan være selvstendige eller være et del av et kompendium. (Gentleman, 2005)

R-notebook kan også være et godt program hvor kode og tekst er sammenhandlet. Samt at du har muligheten for branches og master for å lagre arbeidet underveis hvor informasjon ikke blir mistet på veien.

Det finnes også andre grunner for at det kan være vanskelig å reproducere artikler. arbeidet utført kan være mer komplekst enn først antatt. Det kan være rapportert nøyte men det kan være en variabel om har blitt glemt, som gjør at det ikke kan bli replikert. Det er også mulighet for at det var kun flaks som påvirket utfallet. Det kan også være at forsøket er manipulert slik at det har en bias, som skaper ønsket resultat. (McNutt, 2014)

American Statistical Association og noen andre har foreslått for Marcia McNutt om å øke medlemmer for deres “Board of Reviewing Editors”. De skal hjelpe med å kontrollere forskere om at de driver skikkelig forskning, og opprettholde de høye standardene som kreves i forskning. Det nevnes også at det er forskere som har hovedansvaret for å opprettholde kvaliteten på forskningen selv. (McNutt, 2014)

Schmidt har fem funksjoner av replikasjon i sosiale studier.

Funksjon 1. kontrollering av informasjonshenting, for å forsikre seg om at informasjonen i deres artikkel ikke er generert av flaks eller manipulert. Funksjon 2. kontrollering for “internal validity” —som er, en forsikring av de eksperimentelle resultatene er en valid test av hypotesen. som ikke reflekterer reflect utilsiktet feil i utførelsen. Funksjon 3. Kontrollere for svindel. Funksjon 4. Ta i bruk generaliserbarhet. Funksjon 5. Ta i bruk verifisering av den underliggende hypotesen. (Fidler og Wilcox (2018))

## Analyse

DataCite er et stort register hvor man har prøvd å samle vitenskapelig data slik at det ikke bare blir liggende rundt hos de ulike tidsskriftene (Brace (2009)). Målet er å gjøre det lettere for andre å få tak i dataene.

Det å bruke “code chunks” eller “text chunk” kan være med på å gjøre det lettere å reproducere eller replikere dataene til andre og sine egne data. Det kan gjøre at de som skal jobbe videre med det får en større forståelse over hvordan resultatene har blitt til. Noe vi tror kan være et problem er hvis det kun legges til en av “chunkene”, som kan gjøre at de som skal reproducere eller replikere ser hvordan de skal gjøre det, men at de ikke klarer å overføre det til sine egne data. Det beste vil da være å ha både koding og tekstforklaring.

Hvis det er lettere for andre å reproducere eller replikere det arbeidet du har gjort, så vil det nok være lettere å få anerkjennelse for arbeidet. Det vil være mulig å videreutvikle det du har startet, og det blir ikke bare et uforståelig arbeid som ingen kan ha utbytte av.

I dagens samfunn hvor verden bare blir mindre og man får tak i nesten all forskning på internett, blir det nok et større fokus på å reproducere og replikere andre sitt arbeid med sine egne data og eventuelt videreutvikle arbeidet. Det ser dermed ut til at dagens økte interesse for reproduserbarhet ikke kun er et blaff, men at det til og med vil øke i tiden fremover. Det er også et økt fokus på høyere utdanning i flere land, noe som igjen vil føre til at flere ønsker å jobbe med å produsere, reproducere og replikere forskningsartikler.

Det som vil bli viktig fremover for forskningen er å bygge vaner med å dokumentere forskningen fra starten, slik at det ligger som automatikk i arbeidsmetoden. Får vi innarbeidet rutiner til forskere fremover vil dette bli automatisert også skape gode arbeidsarkiv og kompendium, noe som vil føre til bedre reproduserbarhet for forskere i fremtiden som vil replikere forsøkene.

Svindel og manipulering av forskningsartikler vil også bli redusert med hyppigere kontroller, noe som vil bidra til at seriøse forskere vil bli mer anerkjent for sitt ærlige arbeid. Noe som vil skape bedre fundament for fremtidige forskere sitt arbeid, når man kan være mer sikker på at forskningsartikler har større sannsynlighet for å ha troverdige og ærlige utfall.

Mer som jobber med arkiv og kompendium vil også forbedre oversikten av informasjon, som igjen vil effektivisere lete tid og unødvendige forsøk på hypoteser som allerede har blitt utarbeidet utallige ganger fra før med samme resultat. Dette vil spare enorme ressurser for forskere og institutter som ikke trenger å teste unødvendige hypoteser.

## Konklusjon

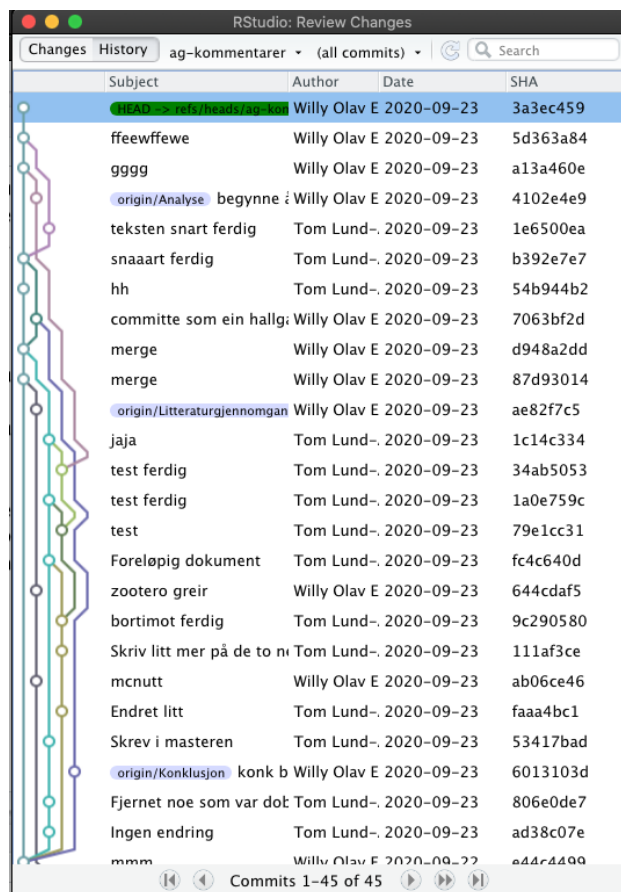
Det vi har tatt for oss i denne oppgaven er ulike eksempler, hvor det har blitt testet artikler for reproduserbarhet og replikeringsevne. Vi ser at det finnes en effekt for å skape arkiver med code chunks, tekst chunks, og compendium. For å ha ønsket effekt må det ha et system hvor det blir kontrollert og opprettholdt.

Analysen viser oss at forskere gjerne vil kunne skape reproduserbare artikler. Det som er problemet er at det er dyrt og tidkrevende og lagre alt på veien. Det kan også vise til feil de har gjort under prosessen. Analysen viser oss at forskerene må nærmest bli tvunget til dette og skape rutiner for ønsket effekt.

## Referanseliste

- Brase, J. (2009). DataCite - A Global Registration Agency for Research Data. *2009 Fourth International Conference on Cooperation and Promotion of Information Resources in Science and Technology*, 257–261. <https://doi.org/10.1109/COINFO.2009.66>
- Dewald, W. G., Thursby, J. G., og Anderson, R. G. (1986). Replication in Empirical Economics: The Journal of Money, Credit and Banking Project. *The American Economic Review*, 76(4), 587–603.
- Fidler, F., og Wilcox, J. (2018). Reproducibility of Scientific Results. I E. N. Zalta (Red.), *The Stanford Encyclopedia of Philosophy* (Winter 2018). Metaphysics Research Lab, Stanford University.
- Gentleman, R. (2005). Reproducible Research: A Bioinformatics Case Study. *Statistical Applications in Genetics and Molecular Biology*, 4(1). <https://doi.org/10.2202/1544-6115.1034>
- Goodman, S. N., Fanelli, D., og Ioannidis, J. P. A. (2016). What Does Research Reproducibility Mean? *Science Translational Medicine*, 8(341), 341ps12–341ps12. <https://doi.org/10.1126/scitranslmed.aaf5027>
- McCullough, B. D., McGeary, K. A., og Harrison, T. D. (2008). Do Economics Journal Archives Promote Replicable Research? *Canadian Journal of Economics/Revue canadienne d'économique*, 41(4), 1406–1420. <https://doi.org/10.1111/j.1540-5982.2008.00509.x>
- McNutt, M. (2014). Reproducibility. *Science*, 343(6168), 229–229. <https://doi.org/10.1126/science.1250475>

## Appendiks



Figur 1: Commit historie.