



UiT Norges arktiske universitet

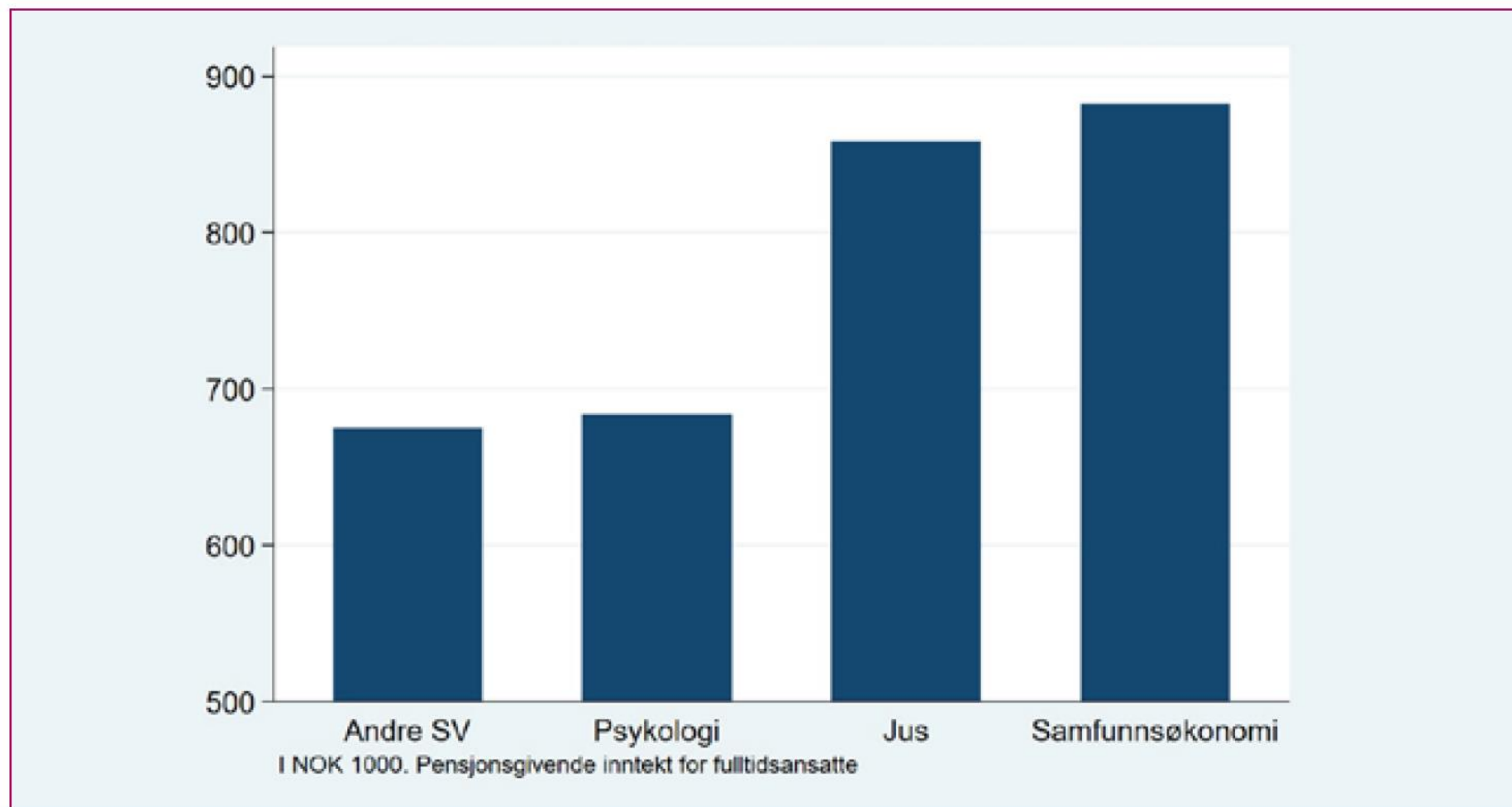
SOK-1004 Økonomiske emner med programmering

Forelesning 1 – Økonomisk etterforskning

Derek J. Clark

«Økonomer er bare opptatt av penger!»

- Nei, men



Figur 2: Hva de tjener i 2017. Gjennomsnitt.

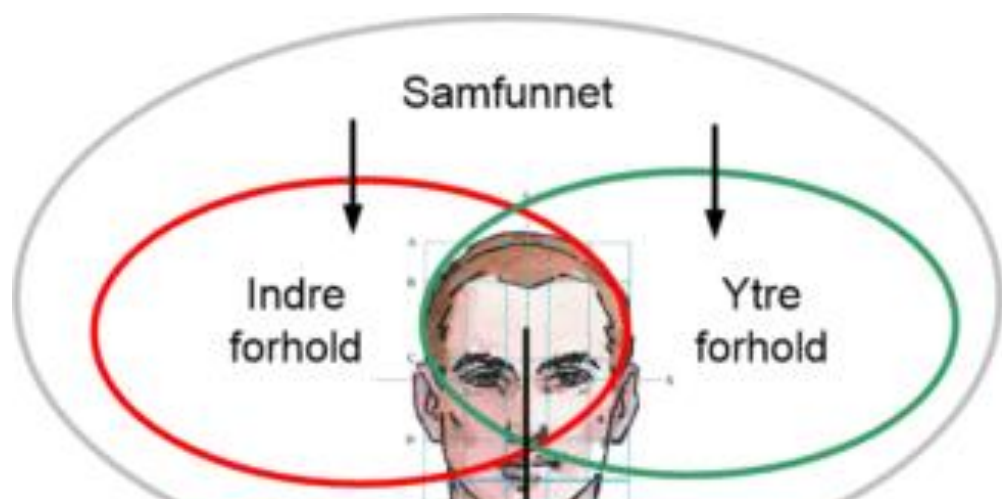
«From zero to hero?»



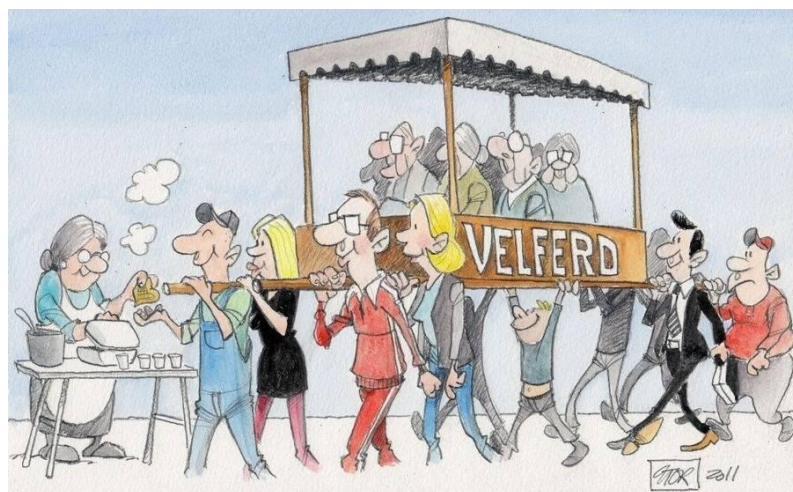
"If you knew that economics was the 'dismal science,' why did you become an economist?"



Det handler om mennesker.....



og



«Best» mulig fordeling av samfunnets ressurser?



Økonomisk «etterforskning»

Spesifiser
problemstillingen

Forenkle
virkelighet med
en modell

Innhente data

Kjør analyse

Evaluer og
presenter
resultatene

Foreslå tiltak

Evaluer tiltakene

Problemstillingen

- Avgrenset tema
- Et spørsmål eller påstand
- Utforskende og presis



Modell av virkelighet

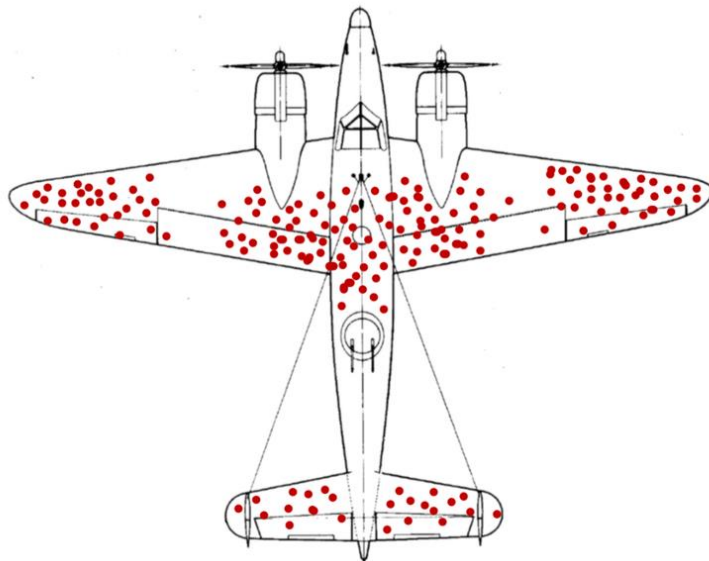
- T-bane i Oslo: modell og virkelighet
- London underground



Innhente data

- Forstå hva du har (og hva du mangler)

Section of plane	Bullet holes per square foot
Engine	1.11
Fuselage	1.73
Fuel system	1.55
Rest of the plane	1.8



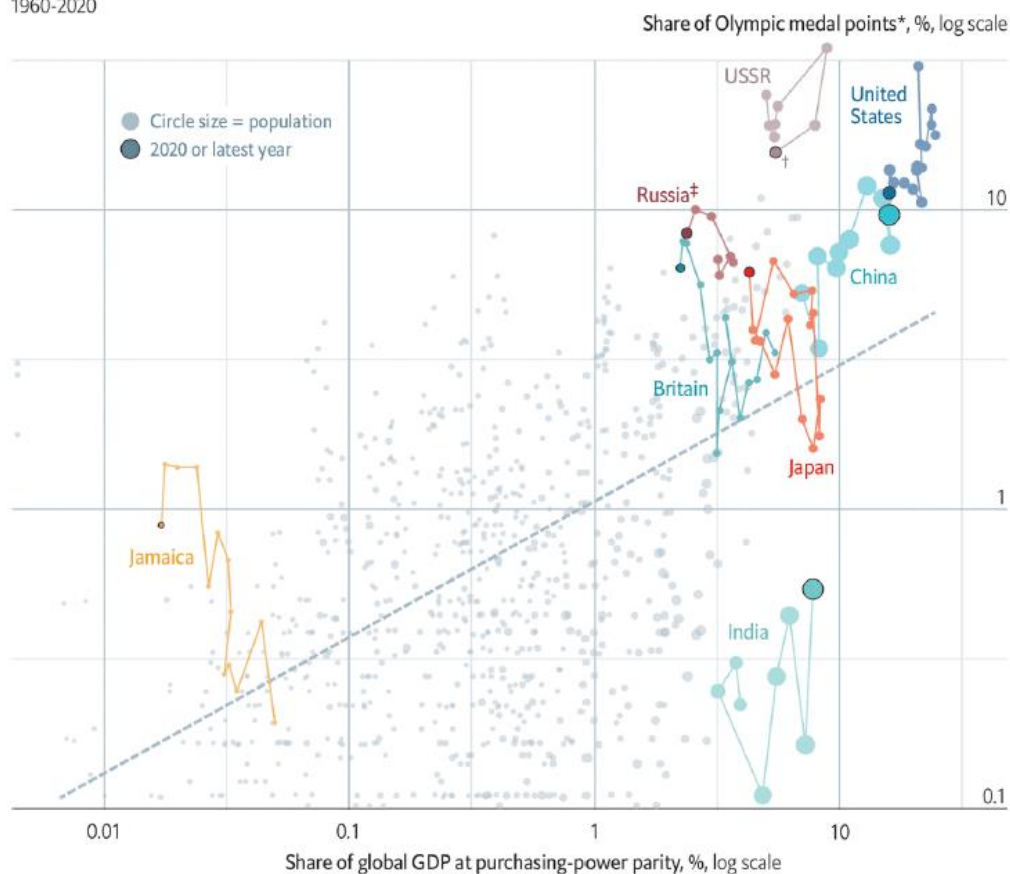
Abraham Wald



Økonometrisk analyse

Born to run

Economic output v Olympic medals
1960-2020



*Weighted medals tally, where: gold=3 points; silver=2 points; bronze=1 point

†1992, newly-independent USSR states competed as a single team ‡Russian Olympic Committee in 2020

Sources: International Olympic Committee; Olympedia.org; Maddison Project, University of Groningen; *The Economist*

The Economist

$\% \text{ medaljepoeng} = f(\% \text{BNP}, \text{befolkning}, \text{helse}, \text{utdanning})$

‘GDP alone explains 55% of the variation in Olympic medals won since 1960’. *The Economist* 8.8.21.

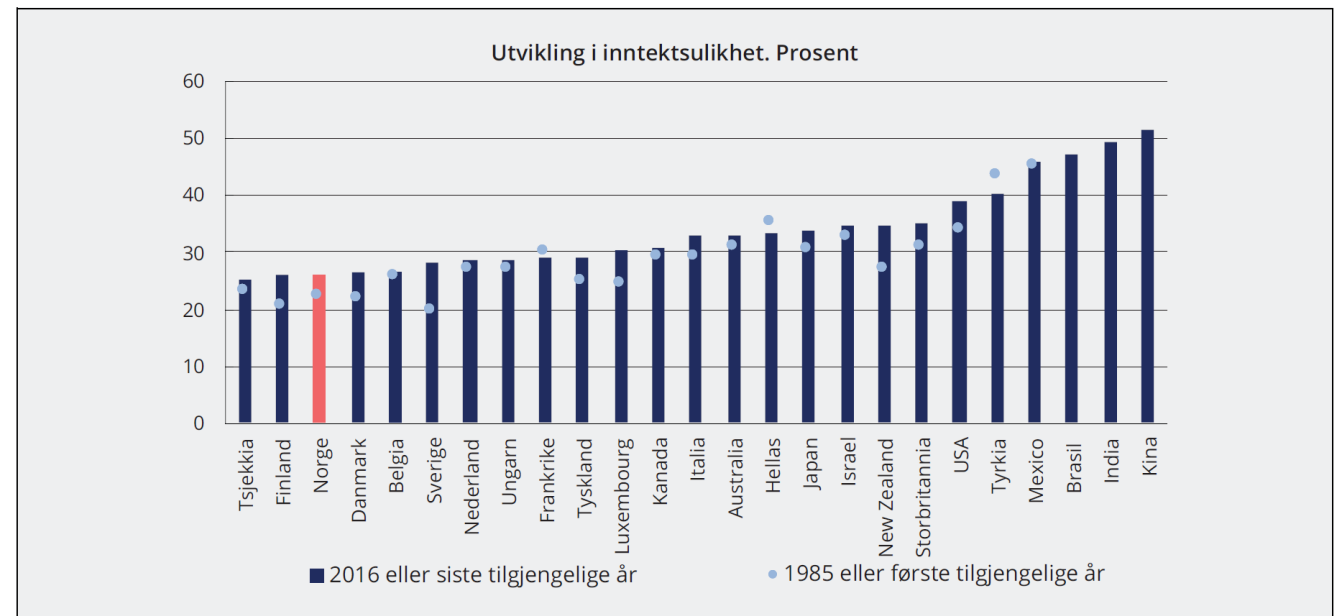
Presentasjon av resultater

- Datavitenskap gir oss mange muligheter:
 - Reiser på London undergrunn
 - Eksempel fra Gapminder
 - Måling av ulikhet:

28

Meld. St. 13
Muligheter for alle

2018–2019



Figur 2.5 Gini-indeks for inntekt etter skatt. 1985¹ og 2016 eller siste tilgjengelige år. OECD-skala

Er resultatene fornuftige?

nature publishing group

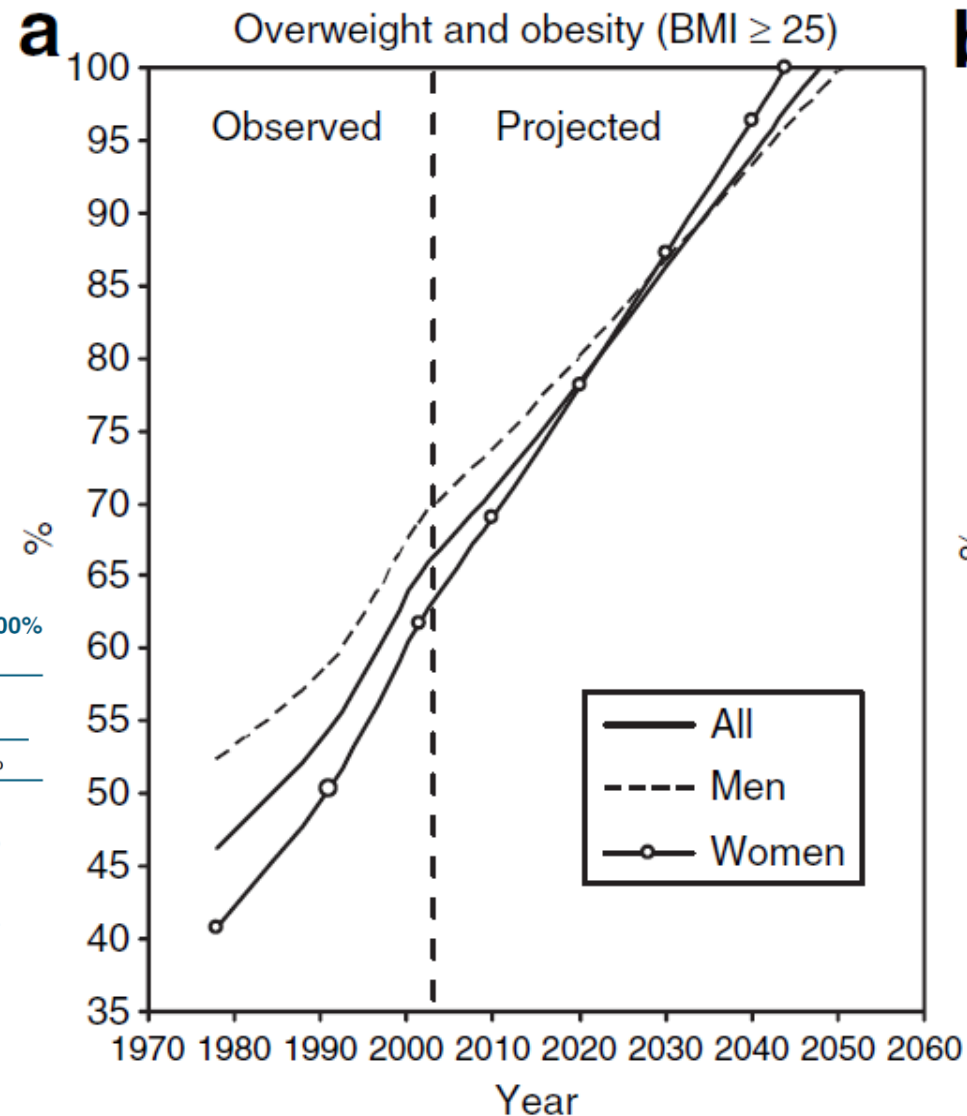
ARTICLES
EPIDEMIOLOGY

Will All Americans Become Overweight or Obese? Estimating the Progression and Cost of the US Obesity Epidemic

Youfa Wang¹, May A. Beydoun¹, Lan Liang², Benjamin Caballero¹ and Shiriki K. Kumanyika³

Table 2 Future projections:^a time when the prevalence of overweight or obesity among US adults will reach 80, 85, 90, and 100% and prevalence of overweight (BMI ≥ 95th percentile^b) among US children will reach 30, 40, and 50%

	Gender	Ethnicity	Year when the prevalence will reach		
			80%	90%	100%
US adults	All	All	2022	2035	2048
	Men	All	2020	2035	2051
	Women	All	2022	2033	2044



Foreslå og evaluer tiltak

Har tiltaket ønsket effekt?



www.alamy.com - EY0W5B



Datavitenenskap (fra kollega Prof. Øystein Myrland)

Volume 455 Issue 7209, 4 September 2008



Feb 27th 2010 edition >



The data deluge | Feb 27th 2010 | The Economist

What is data science?

The future belongs to the companies and people that turn data into products.

By Mike Loukides. June 2, 2010

We've all heard it: according to Hal Varian, statistics is the next sexy job. Five years ago, in What is Web 2.0, Tim O'Reilly said that "data is the next Intel Inside." But what does that statement mean? Why do we suddenly care about statistics and about data?

In this post, I examine the many sides of data science — the technologies, the companies and the unique skill sets.

The web is full of "data-driven apps." Almost any e-commerce application is a data-driven application. There's a database behind a web front end, and middleware that talks to a number of other databases and data services (credit card processing companies, banks, and so on). But merely using data isn't really what we mean by "data science." A data application acquires its value from the data itself, and creates more data as a result. It's not just an application with data; it's a data product. Data science enables the creation of data products.



Maps are data made into a product.
(source: New York Public Library)

The Age of Big Data

By STEVE LOHR

Published: February 11, 2012

The New York Times
Sunday Review



[Click Here To Access!](#)

POPULAR SCIENCE

THE
FUTURE
NOW

THE CONTROL CENTERS

Using Data to Feed the World,
Solve Cold Cases, Battle Malware,
Predict Our Fate »12

OFFICER ALGORITHM

Can a Crime Be Prevented
Before It Begins? »38

NEW WAYS OF SEEING

A Gallery of
Extraordinary
Infographics »41

PLUS

Juan Enriquez
Reprograms Life
»20

James Gleick
Unsplits the Bit
»58

AND

Lawrence
Weschler
Questions the
Cloud
»76

SPECIAL ISSUE

DATA IS POWER

HOW INFORMATION
IS DRIVING
THE FUTURE



Data Science som vitenskap [De Veaux et al. (2017)]

“science of planning for, acquisition, management, analysis of, and inference from data”

“... emphasis on using data to describe the world”

“Data science is inherently interdisciplinary. Working with data requires the mastery of a variety of skills and concepts, including many traditionally associated with the fields of statistics, computer science, and mathematics.”

“Data at the Core”

“Analytical (Computational and Statistical) Thinking”

Ikke alltid nødvendig å være “big” for å være relevant

Forbes

Apr 30, 2014, 9:25am EDT

How Big Data Can Predict The Wine Of The Century



SAP BRANDVOICE | Paid Program
Innovation

By James Marland, Vice President, Network Growth, Ariba – an SAP Company

While Moneyball, and Brad Pitt's good looks, became the face that launched a thousand big data blog posts, I've often thought about other examples, which might appeal to those who prefer to pour a glass a wine versus pore over box scores and Hadoop clusters. This begs the question: Can Big Data help me find a good bottle of wine?

According to the book “[Supercrunchers](#)”, the answer is yes. Imagine trying to determine if 2013 will be a good year for cabernet because you want to invest in wine futures or want to place an early order for a few cases of the good stuff from your wine merchant. The usual approach is to ask a wine connoisseur who has decades of experience and uses the “swish and spit” technique to expose complex wine flavors.

How an Economist Cracked the Wine Business by Predicting Prices With 90% Accuracy

Ashenfelter could've brought down an empire using a simple math formula



Nabil Alouani

Follow

Dec 23, 2020 - 5 min read



Illustration created by the author from a [Twitter post](#)

Orley Ashenfelter's adventure started with a revelation at a wine store. While holding a bottle he couldn't recognize, Ashenfelter knew he had two options:

1. Roll another \$20 dice at the risk of a disappointing taste.
2. Spend half his salary to guarantee a fine drink.

Predicting the Quality and Prices of Bordeaux Wine

Orley Ashenfelter Author Notes

The Economic Journal, Volume 118, Issue 529, 1 June 2008, Pages F174–F184, <https://doi-org.mime.uit.no/10.1111/j.1468-0297.2008.02148.x>

Published: 29 May 2008

Table 2 Regressions of Log Wine Price on Climate Variables

Independent variables	(1)	(2)
Age of vintage	0.0354 (0.0137)	0.0238
Average temperature over growing season (April–September)	–	0.6160
Rain in August	–	–0.00386
Rain in the months preceding the vintage (October–March)	–	0.00117
Average temperature in September	–	–
R-squared	0.212	0.828
Root mean squared error	0.575	0.287

Notes. All regressions are of the (logarithm of) the price of different vintages of a portfolio of Bordeaux chateau wines on climate variables, using as data the vintages of 1952–80, excluding the 1954 and 1956 vintages, which are now rarely sold; all regressions contain an intercept, which is not reported. Standard errors are in parentheses.

Samfunnsøkonomi med datavitenskap



Rstudio IDE (Integrated development environment)



- Vi kommer til å bruke Rstudio i dette kurset
- Dette må lastes ned på din datamaskin
- Datamaskin bør du ta med deg til alle forelesninger/seminarer
- Nedlasting (før forelesning 2):
 - Trinn 1. Last ned den siste versjonen av R fra <https://www.r-project.org/>. Velg CRAN mirror 0-Cloud (øverst på nedlastningssiden).
 - Trinn 2. Last ned gratis versjonen av RStudio Desktop fra <https://www.rstudio.com/products/rstudio/download/>

Andre muligheter

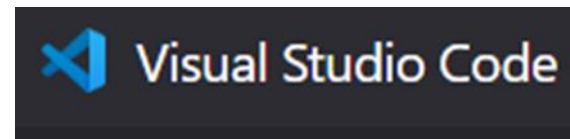
- <https://jupyter.uit.no/hub/login>



- Web-basert
- R og Python som programmeringsspråk
- Markdown som tekstspråk

- <https://code.visualstudio.com/>

- Krever nedlasting

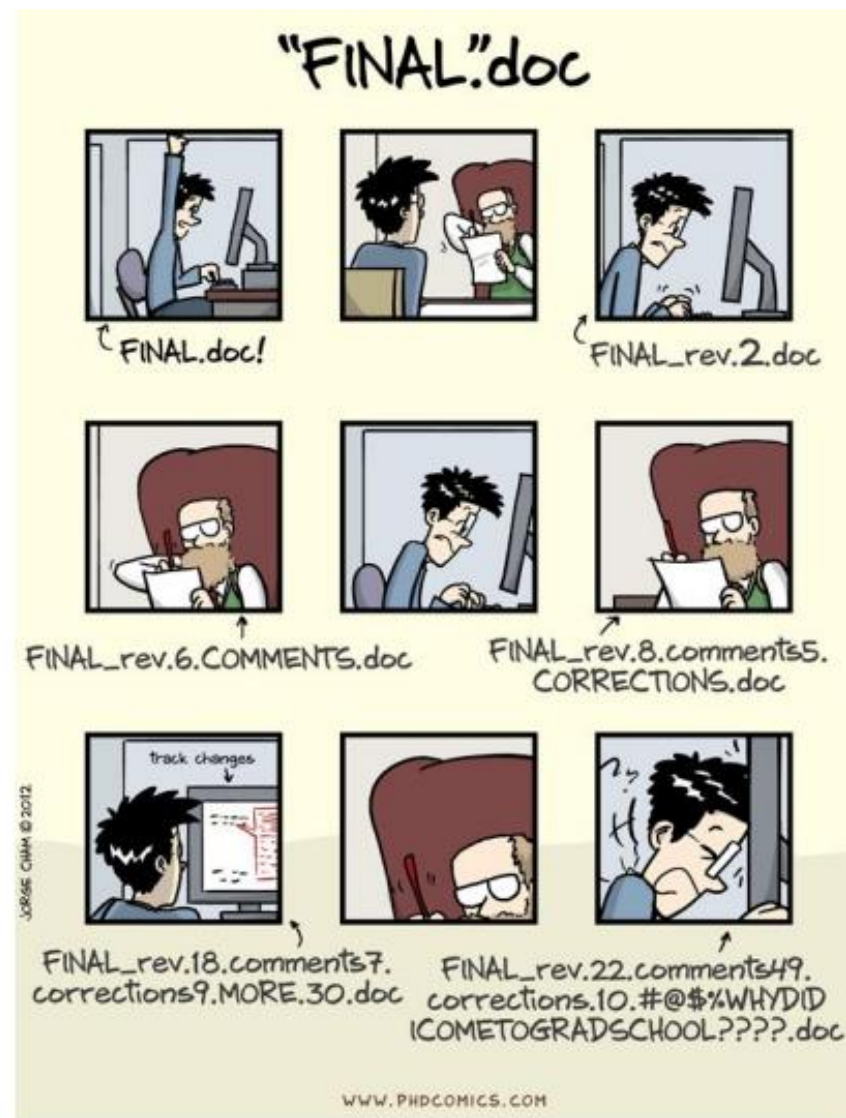


«Holde styr på alt»

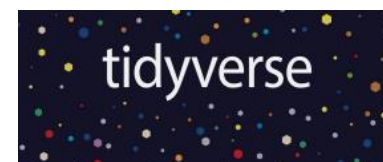
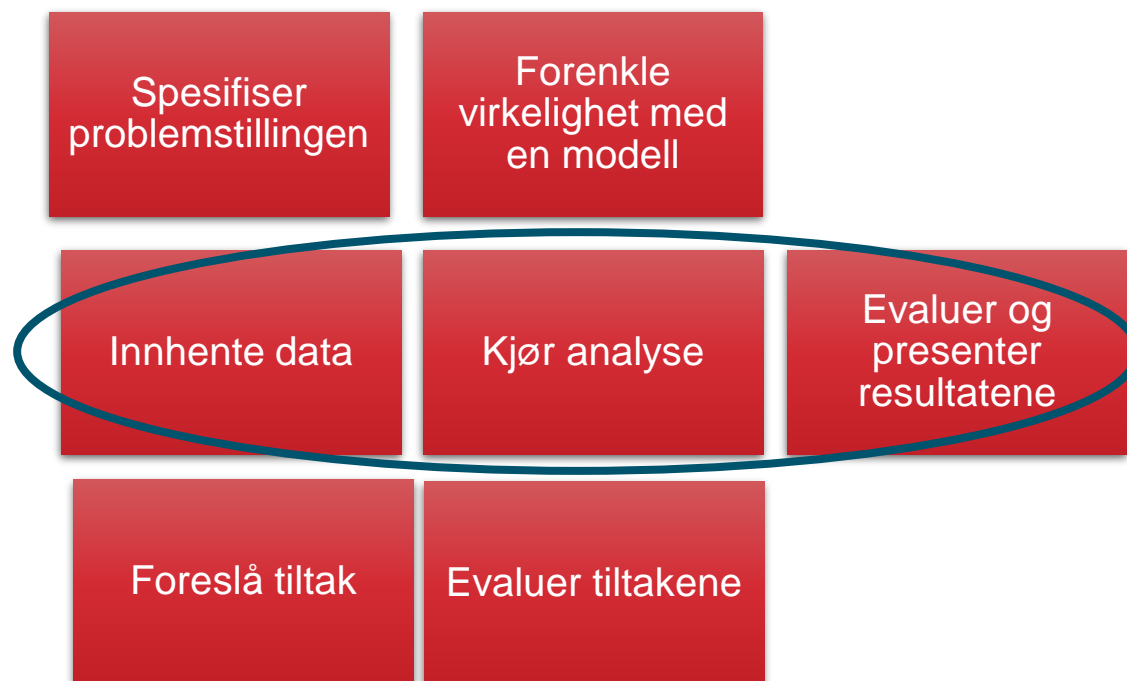
Versjonskontroll



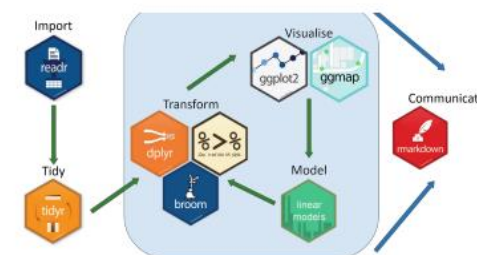
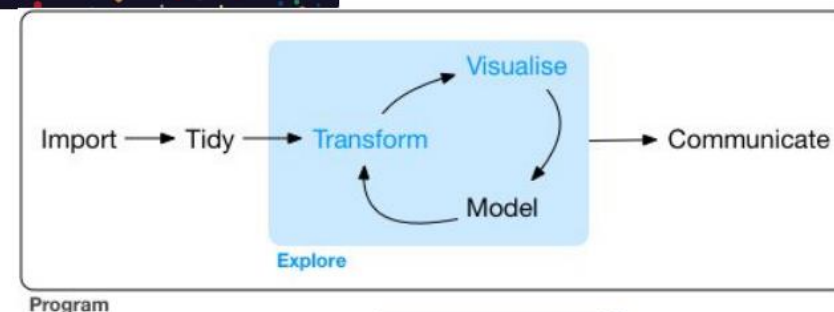
Lag en konto på <https://github.com/>



Økonomi & Datavitenskap

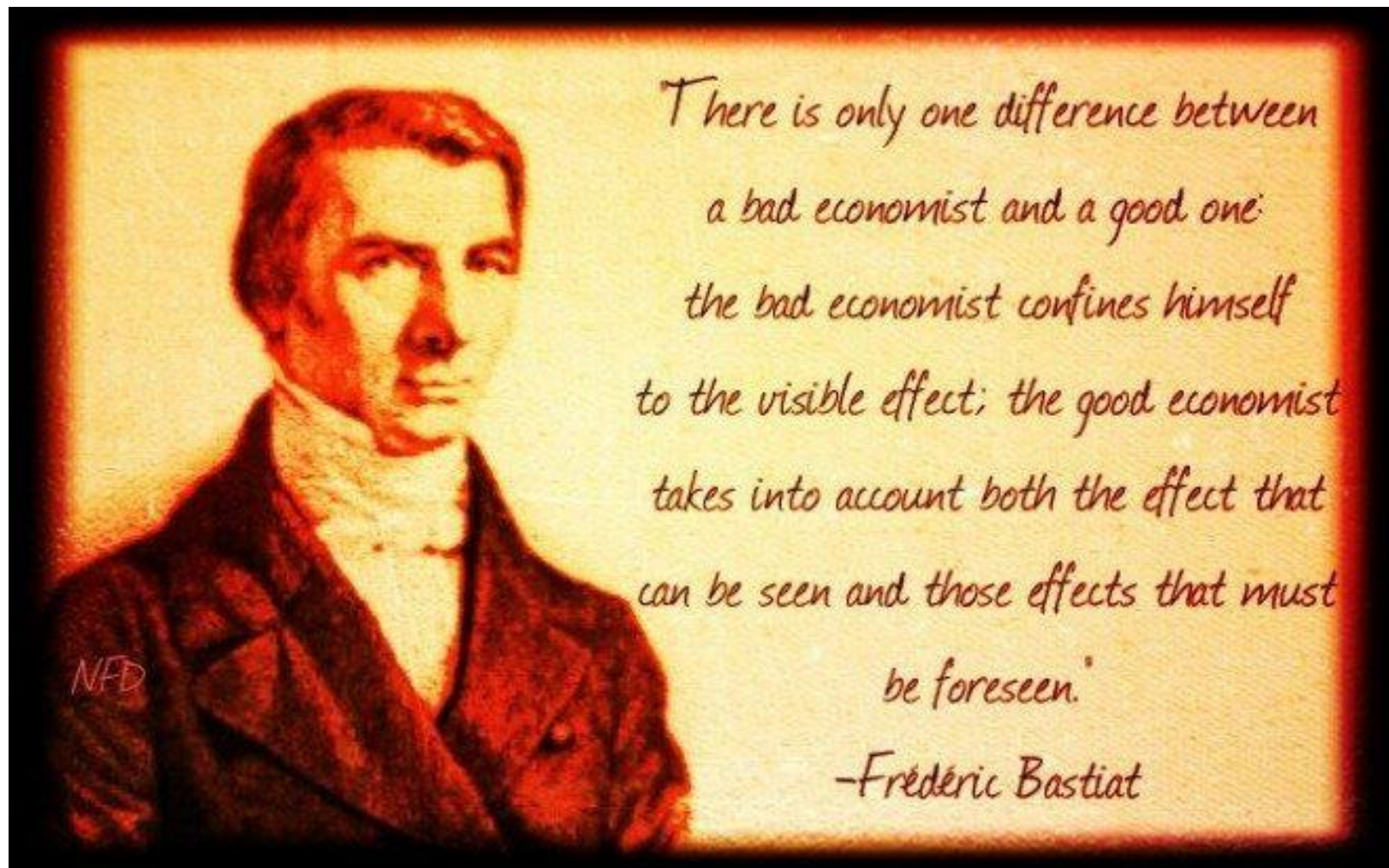


Hadley Wickham's model



(Eksempel fra R)

Begge deler er viktig – teori og empiri



*There is only one difference between
a bad economist and a good one:
the bad economist confines himself
to the visible effect; the good economist
takes into account both the effect that
can be seen and those effects that must
be foreseen."*

-Frédéric Bastiat