The Norwegian Citizen Survey: A Visual Overview

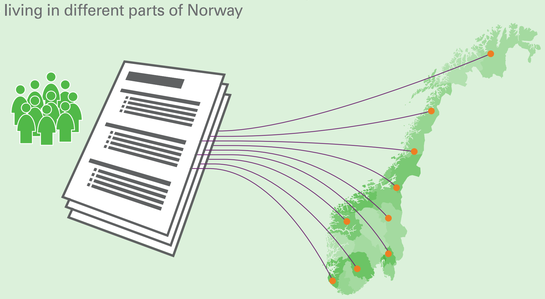
Agency for Public Management and eGovernment (DIFI)

DATA GATHERING

Part 1: Citizens of Norway

The general survey

The first survey asked 30.000 Norwegian citizens for their perceptions of:
- the largest public services in Norway
- the welfare state and democracy in Norway
- living in different parts of Norway



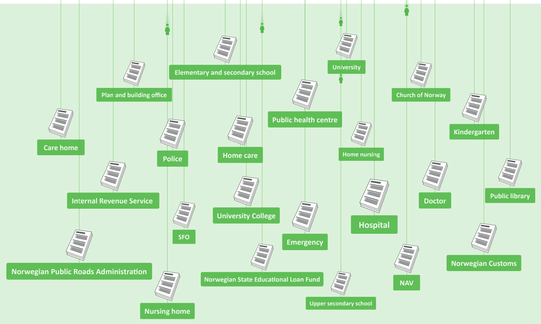
Perception of services indepentent of use

The respondent hasn't necessarily experience with the services he or she evaluate

Representative, but weighting is necessary

Use variable as weight for each respondent. Respondene x weight. Challenge: using combined weights (i.e. men + leaders). Algorithm.

Respondents who have used any of the public services are asked to complete an additional, more specific survey.



Part 2: Users of public services

23 follow-up surveys

Based on experience

Additional surveys were Norwegians' perceptions of services they actually have used (based on experience)

Not representative, no weighting

The data is not representative (not enough respondents), and therefore not weighted. It is possible to see what specific people answered in part 1, and compare the data.

Compare what people answered in part 1 and part 2 (regarding the same service)

DATA PROPERTIES

The Data

The format of the data

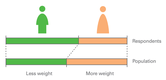
380 Questions

11.000 Respondents

Weighting

Respondents vs population

How many is the person answering for?



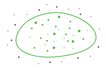
Challenge: data quality

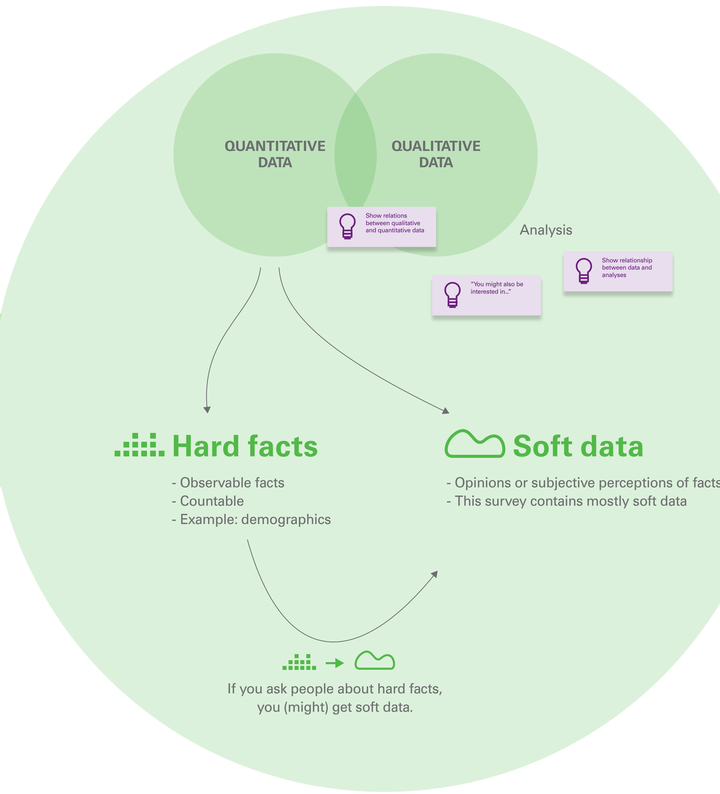
All numbers look solid, but there are uncertainties, and margins of error.

Show uncertainty / solidity of the data visually

Sample of population

When filtering, the samle size decreases. We need a large (enough) number of samples to get representative data. There are





Statistical issues

Significance

Is the difference a result of chance or not?

Is there a statistically measurable difference?

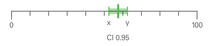
In statistics, a result is called statistically significant if it has been predicted as unlikely to have occurred by chance alone, according to a pre-determined threshold probability, the significance level.

Combining dispersion and significance in one visualization would be innovative.

Confidence Interval (CI)

How confident are we that the result would be the same for the rest of the population? What is the margin of error for the value?

A confidence interval is a range around a measurement that conveys how precise the measurement is. Example, with a confidence interval of 0.95, there is a 95 % chance that all other measurements will be between value x and y.



CI 0.95

Dispersion

To what degree do the respondents agree or not?

The most widely used indicator of dispersion is the standard deviation which, in a nutshell, is based on the deviation of each score from the mean.

Low standard deviation

VS

High standard deviation

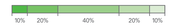
Standard deviation

The standard deviation provides us with a measure of just how spread out the scores are: a high standard deviation means the scores are widely spread; a low standard deviation means they're bunched up closely on either side of the mean.

DATA VISUALIZATION

Typical ways of presenting Likert scale data

Distribution (%)



Advantage: shows distribution
Disadvantage: complex, hard to compare with other

Score (0-100)

Advantage: easy to compare
Disadvantage: does not show distribution (to what degree people agree).
"Don't know" responses are removed.

Let users switch between distribution and scores themselves. Some people want to see both.

Plot every answer on a scale, in order to see both distribution and score..

Show standard deviation visually, combined with score. This would describe whether people agree or not.

Visualize what people agree on, and what they don't agree on.

Are there any contrasts in the data?

Large vs small municipalities

Eksempel:
Hjemmesykepleie
Omsorgsbolig
Plan- og bygningskontor

Users vs other

Innbyggere vs brukere av en tjeneste

2010-2013

Å forstå hvordan skjemaer skal fylles ut
Mulighet for selv å utføre oppgaver over internett

What's interesting to visualize?

- Compare different topics, see the questions underneath
- Compare different regions and different services
- Time series 2010-2013-2015 (when there is more data), as entry to more data
- Navigation between views
- Showing the complexity of the data is in itself valuable
- Let people filter by what they think is important (like OECD), i.e. libraries.
- Use the visualization not only for presenting data, but for analyzing it as well
- Compare data from my selection with the overall data (like in Medarbeiderundersøkelsen)

Connect to other data sources

KOSTRA
Kommune Stat Raportering
http://www.ssb.no/offentlig-sektor/kostra/

StatRes
Skatteetaten
http://www.ssb.no/133500/statres-skatt-hovedtall
vanskelig innleding ift. Innbyggerundersøkelsen.

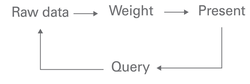
SSB
Statistisk Sentralbyrå
http://www.ssb.no/133500/statres-skatt-hovedtall

Åpent API

Technical reference: 'Medarbeider-undersøkelsen'

Data handling

This was the best way to do it, Rune (Leikanger) spent lots of time on this.
Datahotellet + High Charts



Threshold value

> 30 respondents

DIFI does not show results if data is based on less than 30 respondents