

Introduction to quantitative methods–

SOS2160 Knowledge development in social work

Mats Eirik Lillehagen, Tirsdag 23. April 2024

Overview of the lecture

1. Quantitative methods in brief and why it is important
2. Basic concepts
3. Tools for the seminar – Reading tables and figures

1. Quantitative methods in brief, and why it is important

A world of numbers –



- More and more numbers
- What do we do?
- Becomming critical participants in society
- Big data and AI

The importance of quantitative lenses

- From chaos to meaning?
- A powerful tool
- A way to grasp the world



Statistics describe our (social work) world and contribute to solutions

- More and more (young) disabled people
- Immigration
- Elderly wave
- Technological advances
- Increasing inequality, poverty, health inequalities



Quantitative competence in education and professional practice



- *Theoretical knowledge*
- *Methodological knowledge*
- *Practical skills*
- *Ability to analyze social situations*
- *Critical reflection*
- *Courage to speak out about injustice*

Statistics (and our syllabus) in a bird's eye view



- Society seen from above
- What can be measured
- Contexts!
- Our example: What does a district look like from the height (and locally?).

2. Basic statistical concepts

The world of statistics - Basic classification

Descriptive statistics

**Generalization and
uncertainty**

Statistical descriptions – Two main parts

Simplify – Single variables
– Typical + spread

Relate – Context and
Correlations

What do you actually do when you use quantitative and statistical methods?

- Statistics are first of all about simplifying –
- Example – A resource-poor neighbourhood? Proportions of x,y,z? How do you make it measurable – quantify.
- Math: We take into account probability, randomness, etc. – but the objects of statistics have their own world and logic.
- We break down what are complicated and confusing phenomena into something that is described with a few terms and tools that are used again and again.
- Distribution is perhaps the main concept. Each value has a probability or proportion associated with it. Let's take a closer look.

Basic statistical concepts 1:

- *Units: the ones we want to say something about*
-
- *Variables: selected characteristics or characteristics of the entities that we want to say something about*
-
- *Values: how the properties are measured*



An important distinction - variable types

- **Categorical variables – take only certain values**
- **For example, country of origin, type of education, gender, social class.**
-
- **Continuous variables – things that can be counted – all values within possible extremes are included**
- **Age (0-122), temperature (-273.15->), height, house prices, population size, area per person**

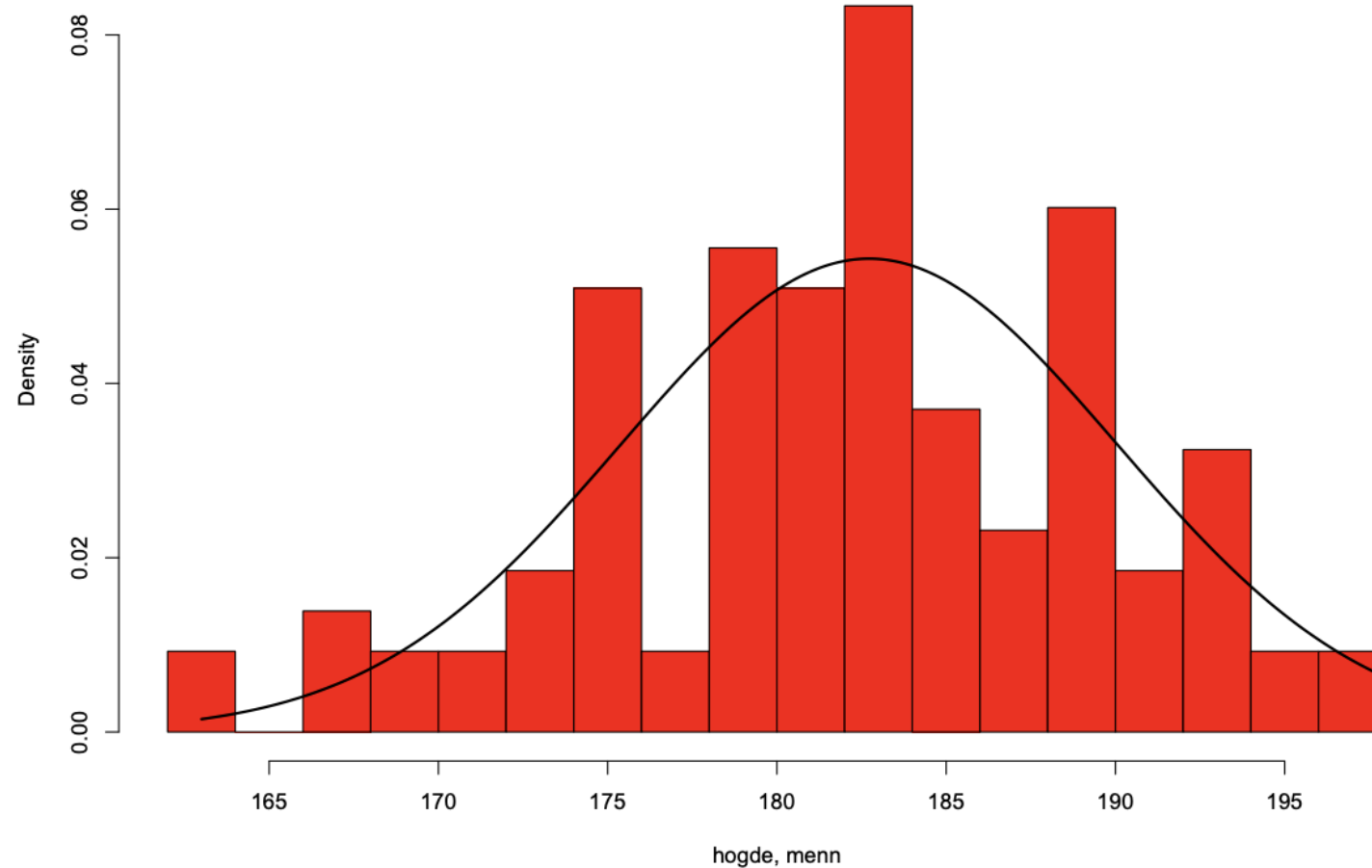
Basic Statistical Concepts 2: The Core Concepts

- We therefore want to say something about properties that vary, e.g. Crime level, education level or house prices. Variables.
-
- Statistics are basically very simple, and mainly say two things.
- 1: Something about the distribution of these variables
- 2: Possibly something about how the distributions are connected -

Let's take a closer look at the core concepts – Distribution of Continuous Variables

- How do you describe a distribution?
- Basic: state the probability of each value – throwing dice.
- You say something about where the center of gravity is – what is typical
- And you say something about the spread – how much variation do you find?

Image example - a typical distribution - height (Norwegian men, military service recruitment)



In other words, we captured the continuous variable using

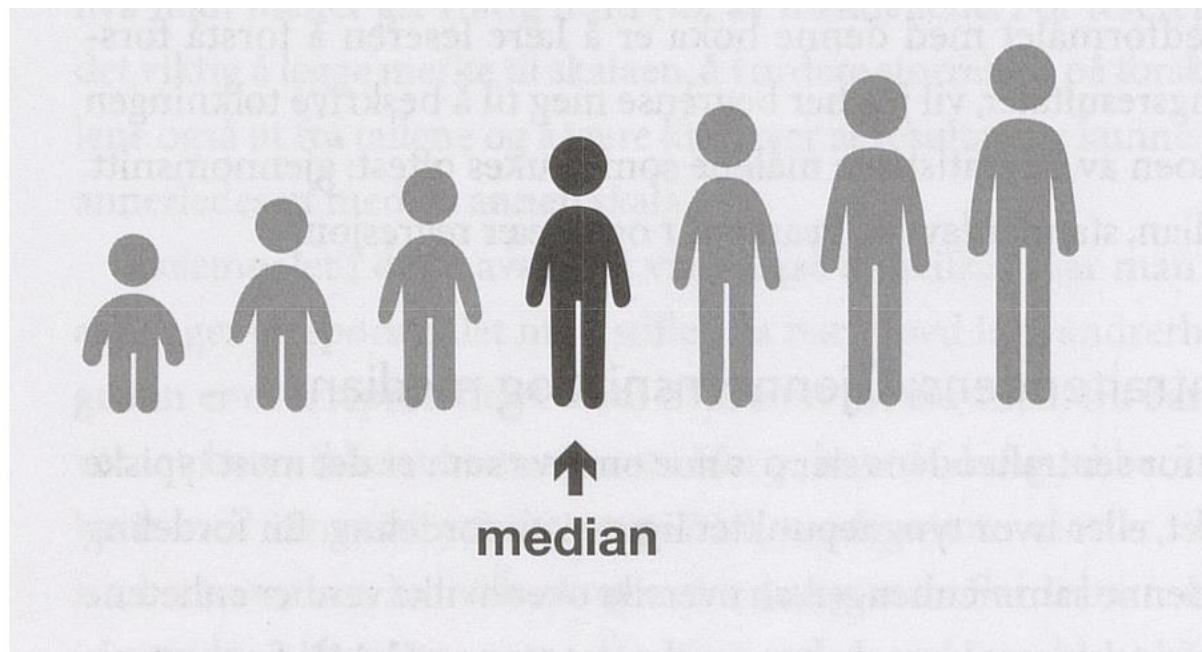
Central tendency

Spread

Two measures of central tendency - Average

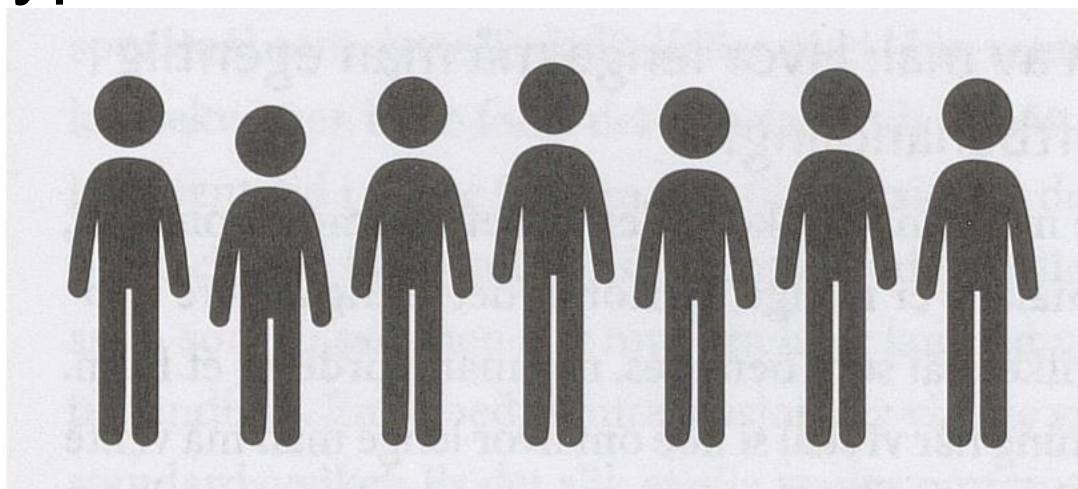
- The most common way to calculate the central tendency/the typical.
- Average age in a neighborhood = add up the ages of everyone and divide by the total sum of the ages.
- What is important to remember – the average can be unduly influenced up by extreme values – such as income and the extremely rich.

Two measures of central tendency - Median

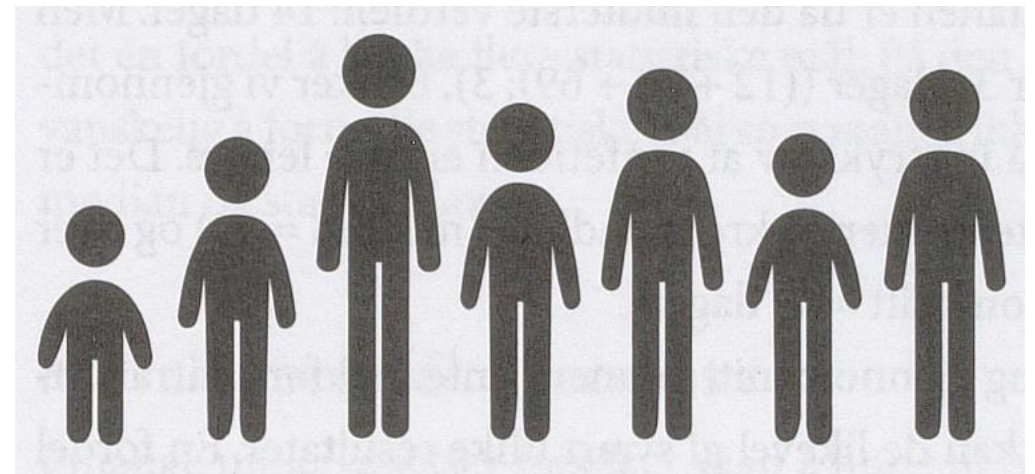


Measures of spread – variance and standard deviation

- The basic idea is again quite simple – it's about what is typical deviation from the mean – how dispersed are the values around typical.



Low variance – fairly similar heights



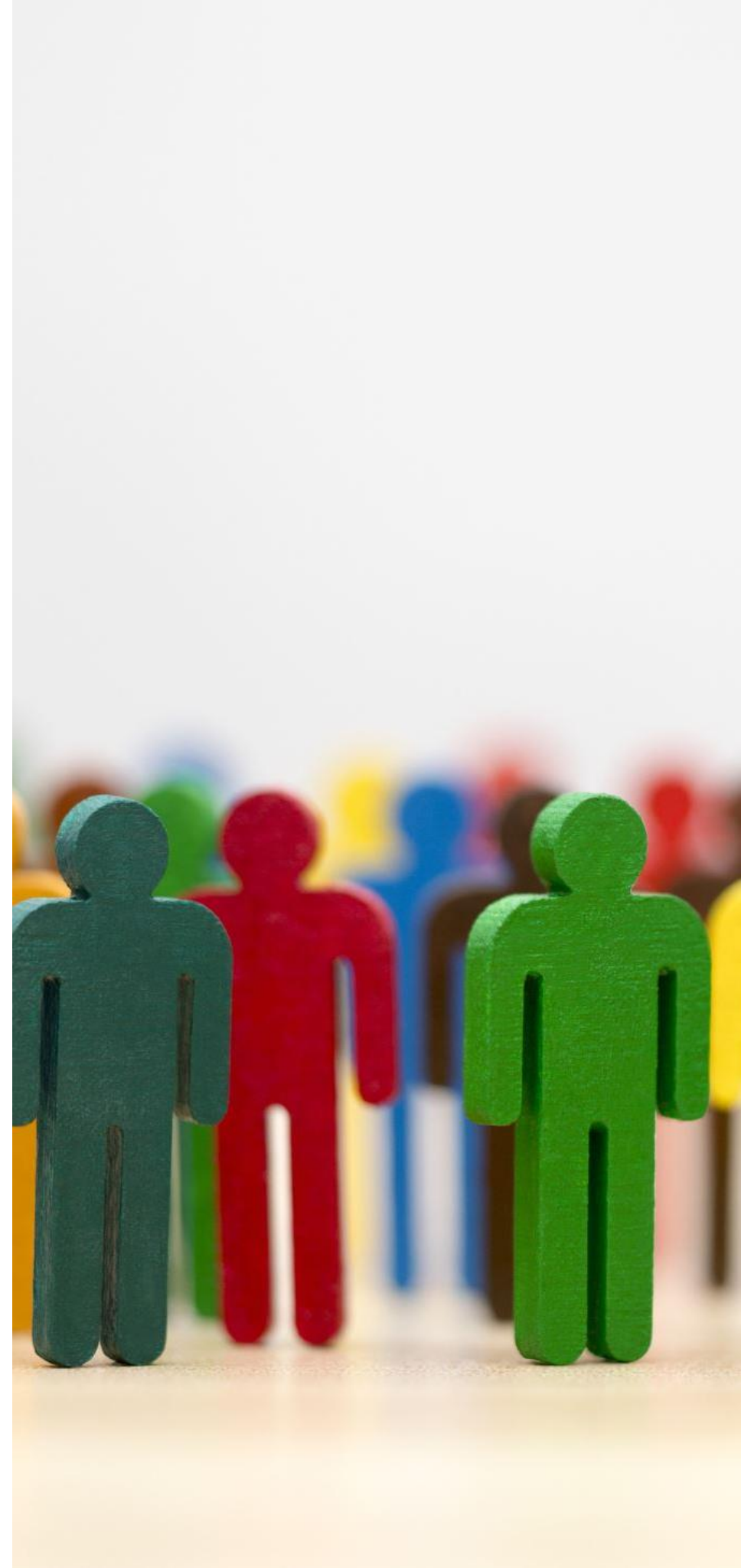
Here there is greater variation

Short recap – Description of continuous variables such as height and income

- Statistical descriptions of measurable quantities are very simple
- The typical value (mean, median) + a spread (variance, standard deviation).
- This illustrates the simplification and abstraction that is typical for quantitative methods – in order to look only at the most important aspects of a phenomenon – boil it down.

Now I will introduce measures of the distribution of categorical variables

- Here we are not talking about central tendency and dispersion in the same way, but the same basic idea – each number gets its share.
- We can talk about the most common value, but it is typical to refer to proportions divided into other categories, which can be measured via fractions or percentages.
- The proportion of women in the population is around 50% or $1/2$ (somewhat lower). The proportion of women increases with increasing levels of education, for example.
- We will look at distributions across neighbourhoods and sub-areas, but first.



The last core concept – correlations or correlation – part 2 of descriptive statistics

Simplify – Single variables
– Typical + spread

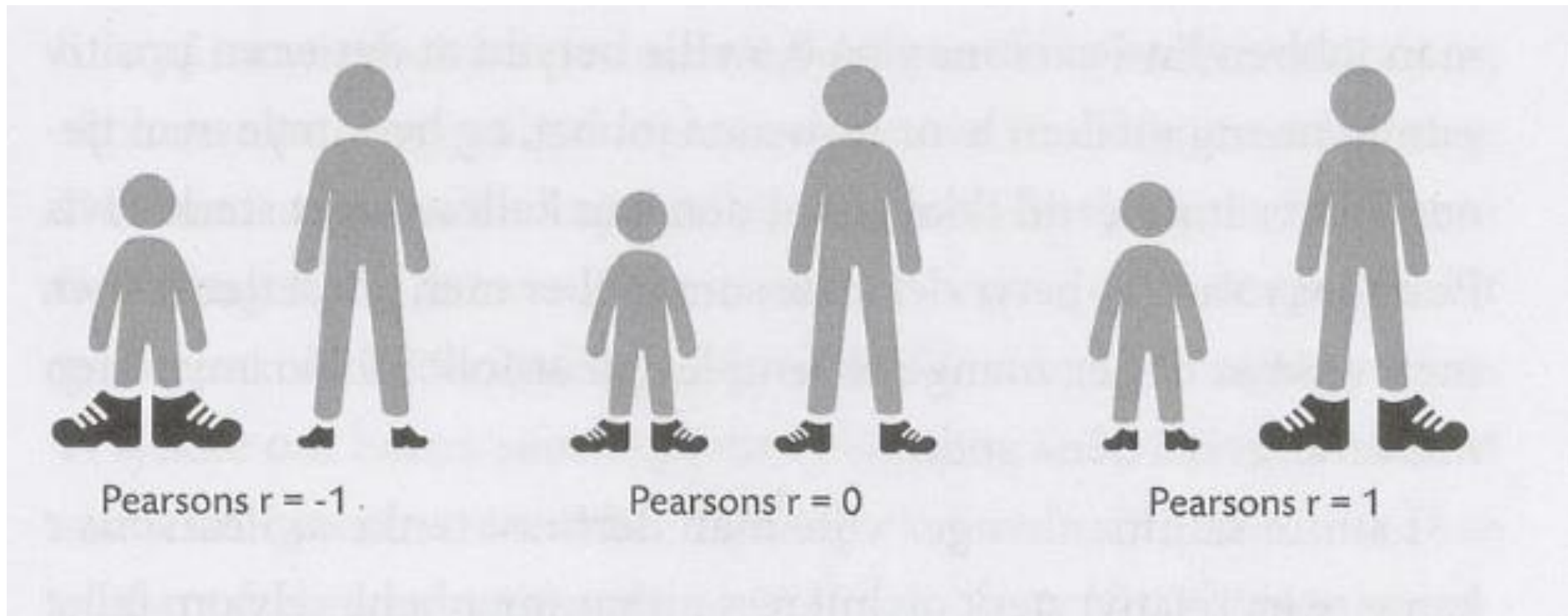
Relate – Relationships
and Correlations

What is meant by correlation?

- Putting a number on covariation – often measured via Pearson's R, see textbook – do variables move in step or out of step?
- Again, the idea is actually quite simple – do high values of one variable correspond to low values of another? If you are tall, do you also have bigger shoes? Being less tall also implies having a smaller shoe-size? Or are there no such patterns?
- Let's look at the different possibilities



Examples of correlation patterns:



More relevant examples.

- Correlations are quite rarely 1 or -1 (perfect)
- Are Hermansen (2020) shows that the average correlation between grade point average and crime in neighbourhoods is - minus 0.25.
- Education level and poverty in neighbourhoods are often linked – consistent with the negative correlation documented here.
- Another example is house prices and standard of living. Higher house prices are typically associated with a higher standard of living (R is positive).

What do correlations mean?

- Naturalism and legality? Causality?
- More or less general relationships.
- Statistical relationships can be causal or non-causal
- Some important concepts here are confounding and reverse causality – read more in chapter 4 of Fekjær's book.

When are causal inferences reasonable?, some practical examples – Neighborhoods.

- All these connections can and should be discussed – But some are easier to determine than others?
- Level of education -> level of poverty?
- House prices <-> living conditions ?



**Before we move on
to the last part – To
say something about
everyone from
someone**

The desire to say something as certain
as possible about the population.

Requires the selection to be
random/straight extraction. (Ungdata.

Any information about everyone

In our survey, for example, it becomes
difficult.

Statistical margins of error in general
you will learn more about later.

Opinion polls – The Labour Party has
lost 0.5% in the last poll, but the margin
of error is 1%. Can you say something
credible?

Actually not that difficult to get an overview

- The basic concepts— units, variables, and values
- Descriptive statistics
- Distributions - Central tendency + spread + proportions
- Covariation – correlation and causality

- Inferential statistics – generalization and margins of error

3. Reading tables and figures

The importance of these skills

- Tables and figures based on statistics are everywhere.
- With a little practice, you learn to extract the information you are interested in.
- Important in studies and in working life and as a citizen.
- It can even be fun – almost no matter what you're interested in.

Let's look at a relevant example

- Tomorrow (seminar) and In connection with the fieldwork, we will collect information about districts in Oslo.
- <https://bydelsfakta.oslo.kommune.no/>
- Here you will find a lot of interesting information –
- Several tables as on the right
- Here we get information on the percentage of individuals who have completed upper secondary school (VGS).

Ikke fullført videregående skole

Andel Antall **Historisk (andel)** H     VALG

Personer 21–29 år som ikke har fullført videregående skole i Bydel St. Hanshaugen

Geografi	2018	2019	2020	2021	2022
Hammersborg	19,1	20,5	19,9	19,5	20,8
Bislett	20,5	18,9	20,8	21,0	18,6
Ila	19,6	19,5	18,3	18,9	18,1
Fagerborg	18,0	17,8	16,0	17,6	15,9
Lindern	23,9	21,1	21,6	21,9	21,9
Bydel St. Hanshaugen	20,1	19,6	19,5	19,8	19,2
Oslo i alt	25,0	24,4	23,8	23,3	22,4

Getting started with the table

- How do you start?
- Some info on proportions
- Make a plan! A lot of numbers.
- Look at the description/headline.
- Find out the unit, variables, values.
- Start by thinking 2-dimensionally – cross-tabulation.

Ikke fullført videregående skole

Andel Antall **Historisk (andel)** H     VALG

Personer 21–29 år som ikke har fullført videregående skole i Bydel St. Hanshaugen

Geografi	2018	2019	2020	2021	2022
Hammersborg	19,1	20,5	19,9	19,5	20,8
Bislett	20,5	18,9	20,8	21,0	18,6
Ila	19,6	19,5	18,3	18,9	18,1
Fagerborg	18,0	17,8	16,0	17,6	15,9
Lindern	23,9	21,1	21,6	21,9	21,9
Bydel St. Hanshaugen	20,1	19,6	19,5	19,8	19,2
Oslo i alt	25,0	24,4	23,8	23,3	22,4

Some questions to ask

- Start by looking at differences within a given year – lowest percentage completed?
- The highest proportion?
- Then:
- Development over time? Any clear patterns?

Ikke fullført videregående skole

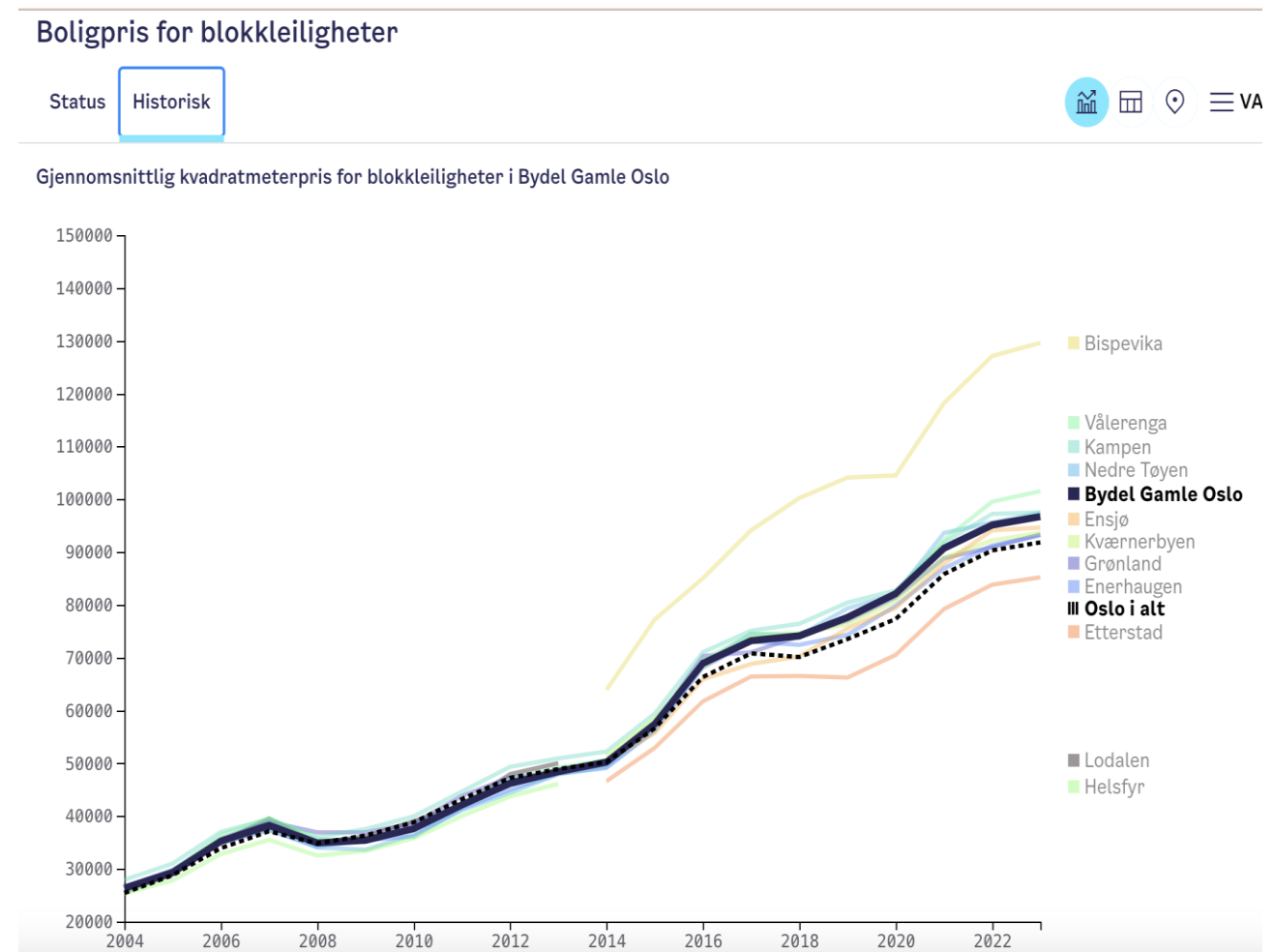
Andel Antall **Historisk (andel)** H     VALG

Personer 21–29 år som ikke har fullført videregående skole i Bydel St. Hanshaugen

Geografi	2018	2019	2020	2021	2022
Hammersborg	19,1	20,5	19,9	19,5	20,8
Bislett	20,5	18,9	20,8	21,0	18,6
Ila	19,6	19,5	18,3	18,9	18,1
Fagerborg	18,0	17,8	16,0	17,6	15,9
Lindern	23,9	21,1	21,6	21,9	21,9
Bydel St. Hanshaugen	20,1	19,6	19,5	19,8	19,2
Oslo i alt	25,0	24,4	23,8	23,3	22,4

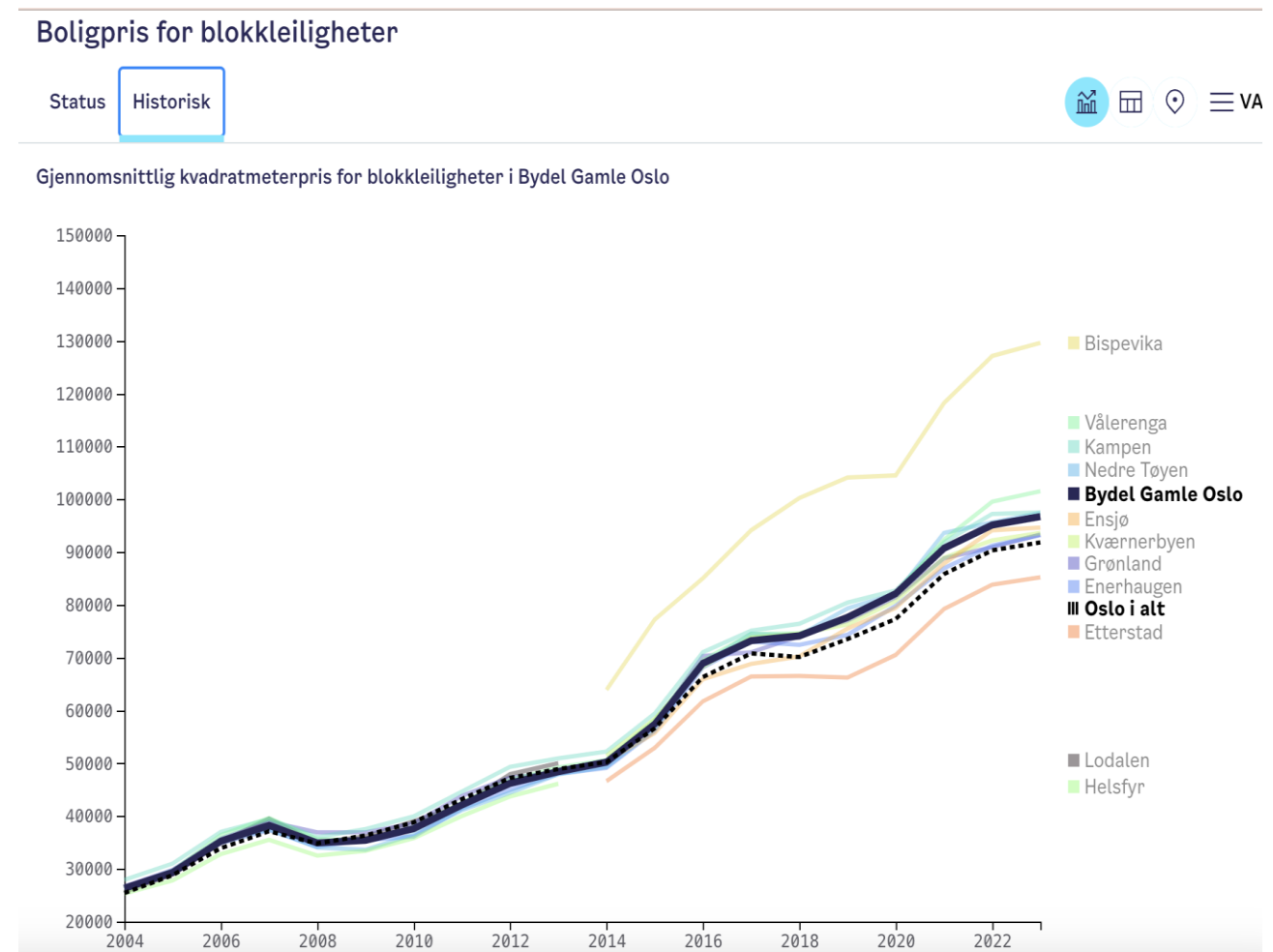
Reading typical statistical figures

- First: look at the title and description –
- Average here, but not spread
- Axes – house price and year.



Reading typical statistical figures

- Look for "trends" – is actually it going up, down, or are we seeing «random walks»?
- Where (geographically) is the trend taking place?
- Think through the meaning of the axis here – a statistical lie?
- See Fekjær's book for more details.



Thank you, see you tomorrow.

- Now you should be equipped with the basic concepts you need to get started with the tasks tomorrow and to get more benefit from Fekjær's book.
- The exercises will help you elaborate on the concepts even more.