# Lecture 6

Scientific Methodology

# Today: Overview

- Some general principles
- Case study: IRT and our experiment
- Case study: Methods in mathematical research and the Four-colour Theorem
- Some techniques for analyzing data.

# The scientific attitude

We can characterize the scientific method by the attitudes of scientists. According to Merton the following should be the attitudes. It is five principles gathered under the acronym CUDOS:

- Communalism - knowledge should be accessible for all people.
- Universalism - everyone should have the right to contribute.
- Disinterestedness - science should be objective and not ruled by special interests.
- Originality - the results should be new.
- Skepticism - scientists should be open to criticism.

# What methods are used?

We can divide the methods into three groups:

- Deductive methods
- Methods for doing experiments
- Methods for reporting and publishing results.

# From lecture 3

- In the lecture we described the Hypothetico - Deductive Method.
- The HD Method is a tool for relating hypothesis to experiments and observations.
- In a way it is a more complex form of induction.
- It is commonly regarded as the fundamental method in science.

# General questions

- What do you want to do? What is your project?
- Why do you want to do it? Is it important? Is it interesting?
- How do you plan to do it? Which methods will you use?
- When do you plan to do it? How long time will it take?

# The subject

- It should be clearly stated.
- It should be significant. For instance, it should not be just a repetition of something already done.
- It should have clearly stated boundaries.
- It should be such that relevant data can be obtained.
- It should be such that significant conclusions can be drawn.

# The form of the project

- It can be in form of a question, for instance, is functional programming better than imperative programming?
- It can be in the form of an hypothesis, for instance, functional programming is better than imperative programming.

# The importance of being right

A famous mathematician once said that the most important thing is being right. You must have the talent for choosing hypotheses that are correct. You must have a sound intuition!

# Scientific method in project work

We can characterize the project work by dividing it into four phases:

- Preparing Analysis
- Finding hypotheses
- Synthesis of partial results
- Validation of results

# Analysis

The goal is to get an understanding of the problem/project. This understanding can involve the following steps:

- Describe the problem.
- Decide on a measure of success.
- Do studies on similar problems.
- Define goals.

# Hypothesis

Here we have to be creative and try to find hypotheses and possible solutions to problems. This includes:

- State the hypothesis/solution clearly.
- Find consequences of the hyptothesis/solution.
- Find criteria for judging if the hypothesis/solution is true/works.

# Synthesis

Here we test the hypothesis or implement and test the solution:

- If we have a solution to a problem we implement the solution.
- Do experiments for testing if the consequences of the hypothesis are true or if the solution works.
- Analyze the results.

# Validation

Here we evaluate the hypothesis/solution and the results of the experiments:

- Try to measure how well the experiments confirm/falsify the hypothesis or how well the solution works.
- Try to decide if the hypothesis is true or if the solution works.
- Do documentation by writing a rapport or scientific paper.
- Submit your results for criticism from colleagues or independent referees.

# Case study: Our IRT project

We will describe the project related to the group experiment on seminar 4.

# Analysis

We will describe the project related to the group experiment on seminar 4. Some years ago Johan read a rather detailed description of IRT  Earlier this fall we decided that we could actually run this technique as an experiment.

Of course, more studies could have been done, but ...

# Hypothesis

We decided to let the student do the test. That meant that we would have to have a suitable set of questions. We decided on a set of hard Swedish words.

Did we have any hypothesis? Perhaps not really. We wanted to see what would happen. But probably a hypothesis was that we would get a ranking of the students that would differ from the one we got by just counting the number of correct answers.

# Synthesis

We will describe the project related to the group experiment on seminar 4. We did the experiment in eight different groups. That gave us 114 results. We ran a computer program for computing the skill/difficulty levels.

It turned out that the new ranking actually was the same as the old one! This was surprising at first. But soon we were able to prove that this is the expected result.

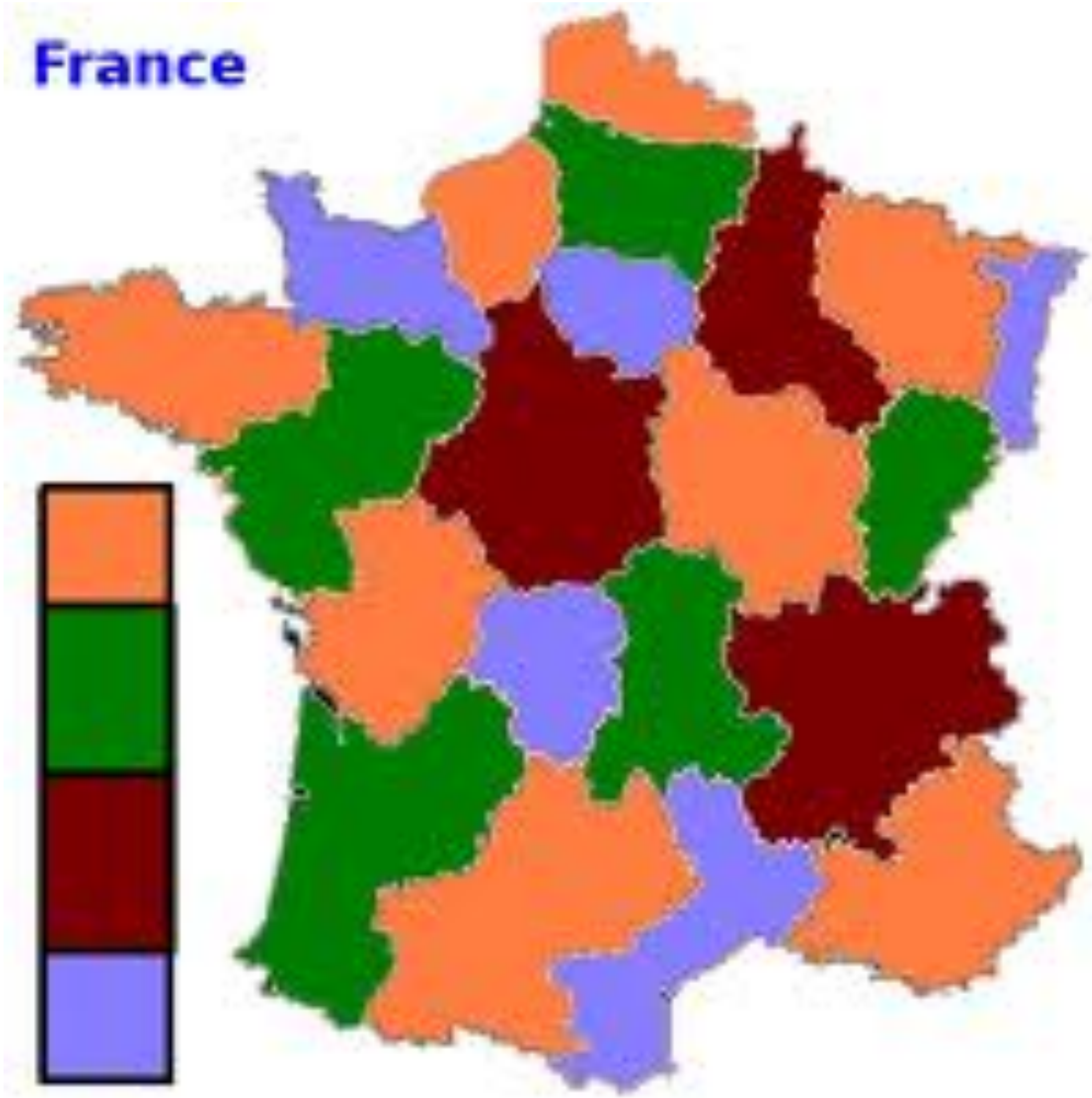
We used a slightly different model which gave us a different result.

# Validation

This phase is still on-going. There are some questions we should consider now:

- Should we try more models and different analyses of the data? We have some ideas about this.
- Was the test situation adequate and significant?
- Repeat the test once more? Perhaps with other data?
- Write a paper and try to publish?

# A case study: The Four-colour Theorem

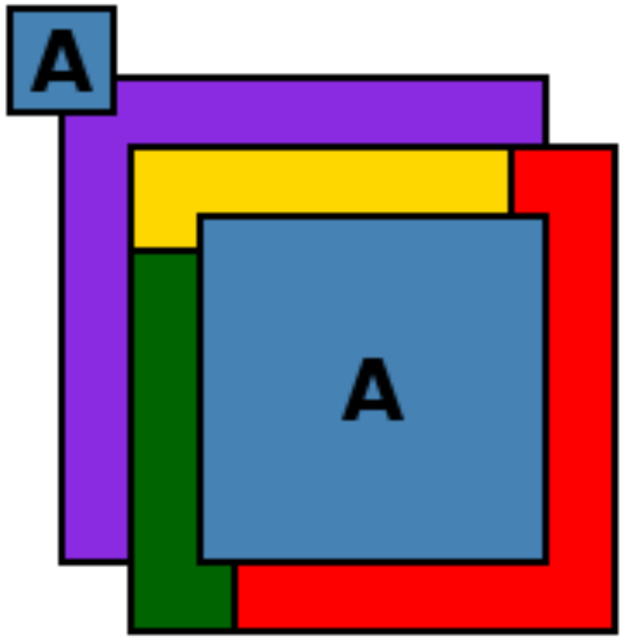


# The theorem

- Every (planar) map can be coloured with four colours. A colouring is required to be such that no neigbouring countries have the same colour.
- This theorem was conjectured in 1852 and finally proved in 1976.

# Is the theorem true?

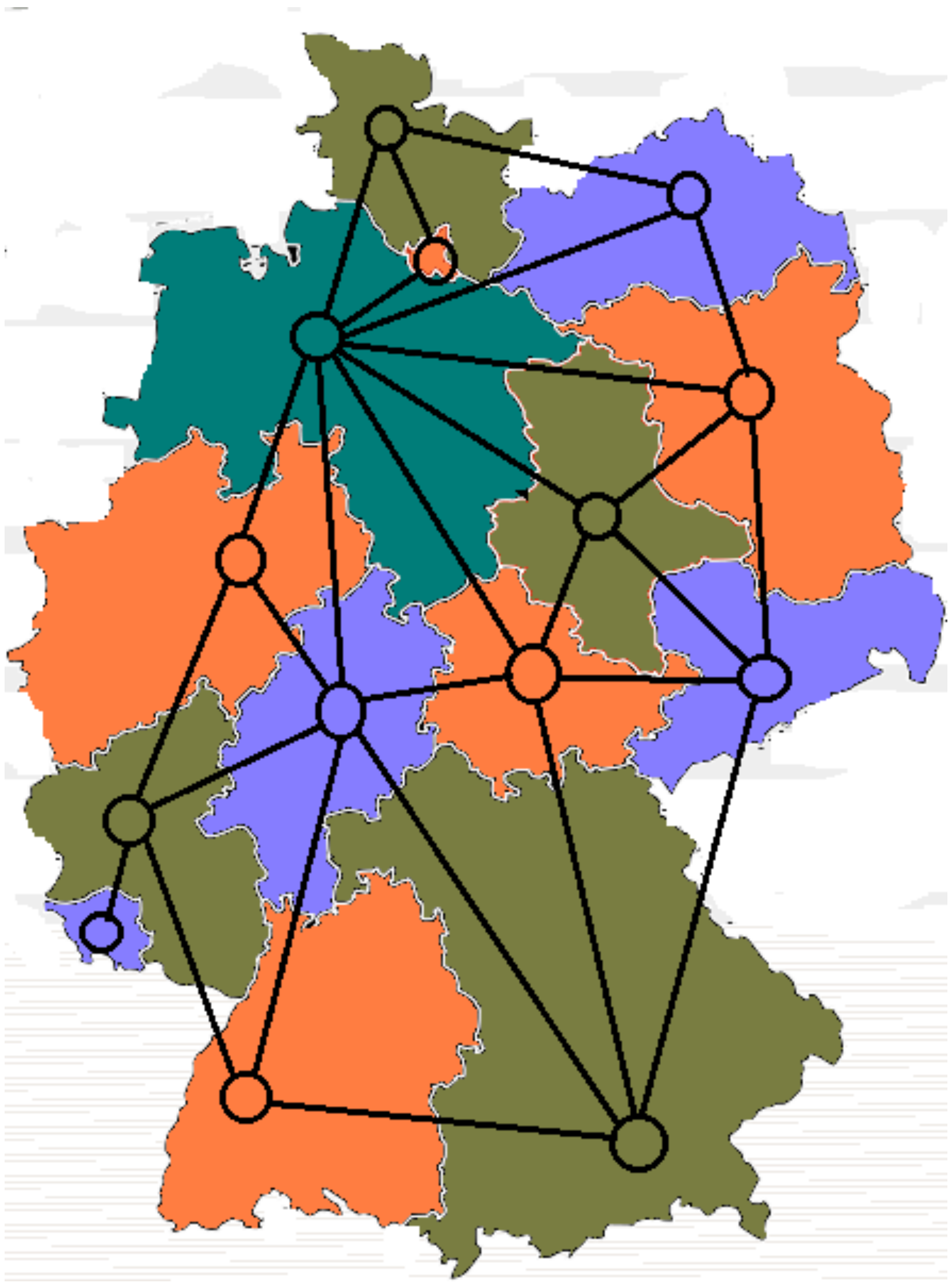
- If we look at a map we notice that two thing can complicate matters: Islands and lakes.
- More generally we see that non-connected countries will give us problems.
- It can be shown that if we allow non-connected countries we can find maps where the four-colour theorem is false.



# A more exact formulation

- We require that the countries must be simply connected and have boundaries that are sufficiently simple.
- Furthermore, two countries meeting just in one point are not to be considered as neighbours.
- These complications make it natural to study the *dual* graph-form: Every planar graph can be coulored with four colours (in the normal node-colouring sense).

# The dual graph

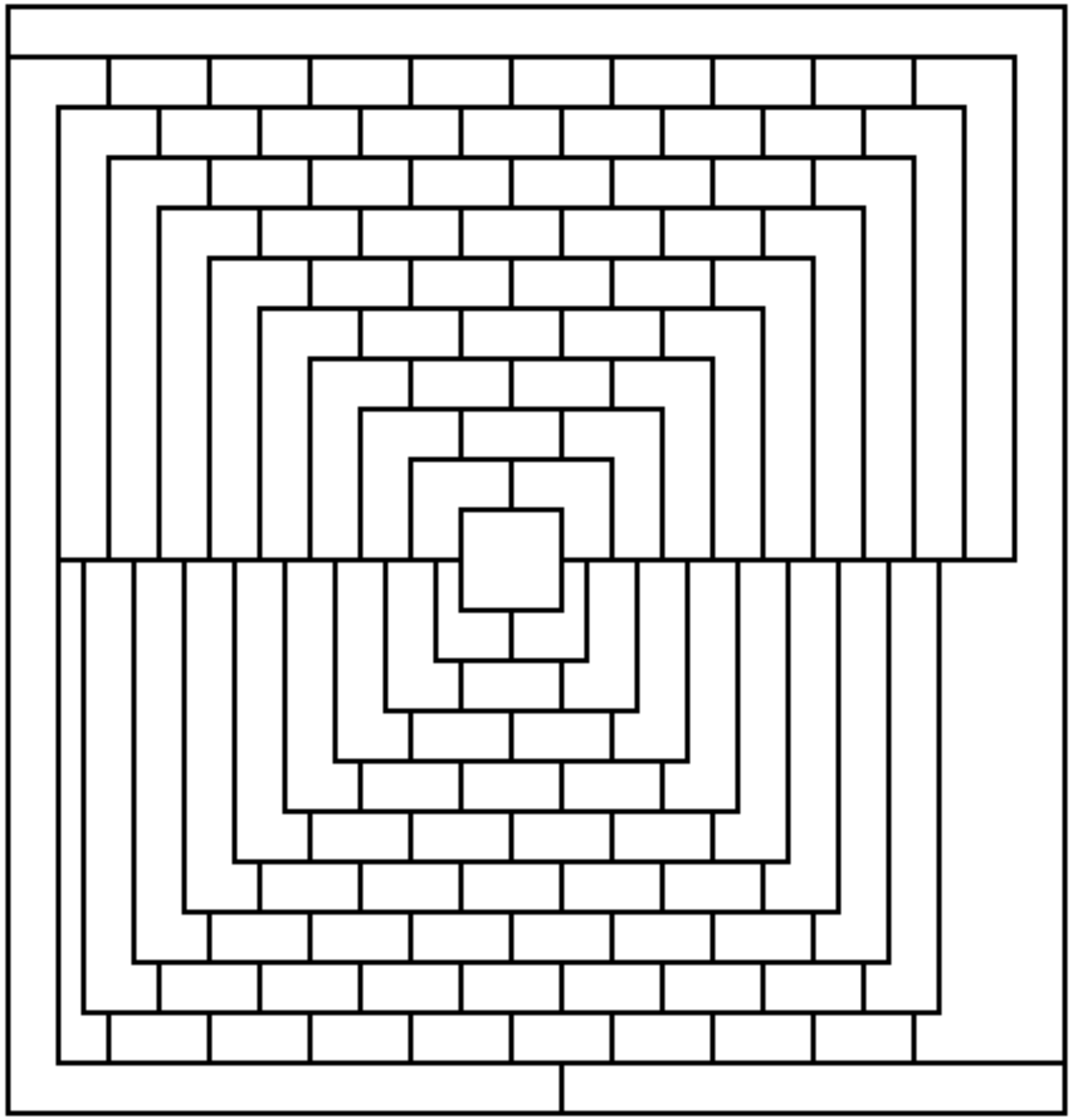


# Method comments 1

- It is important to get an exact formulation of the problem as soon as possible.
- A problem can often be expressed in different forms. Even if the forms are equivalent, one of them can be easier to work with than the other.

# True or false?

- When we face a conjecture we have to guess if it is true or not.
- If we think it is true we try to prove it.
- If we think it is false we try to find a counter-example.



A counterexample?

# A proof?

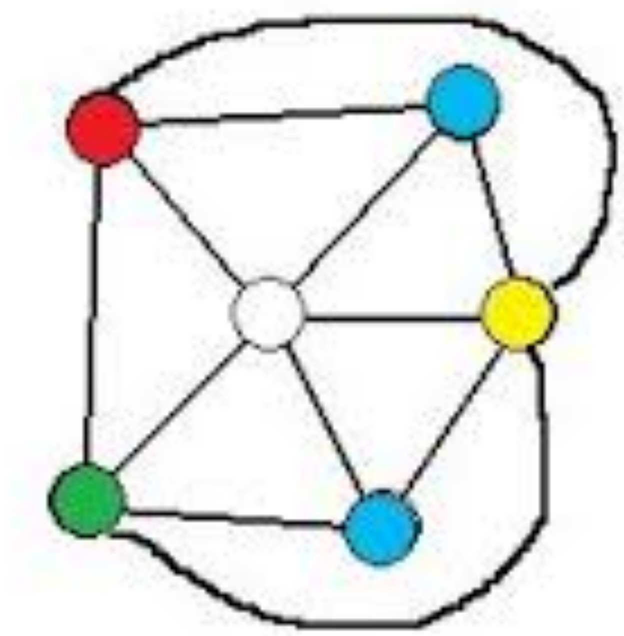
- Most people believed that the FC-Theorem was true. But how can we prove it?
- One attempt is to try to find an algorithm which actually colours any map with no more than four colours.
- But then we have to prove that the algorithm always manage to do this.
- We could try to find some more complicated existence-proof of a four-colouring.
- We could use mathematical induction.

# Kempe's "proof"

- In 1879 Sir Alfred Kempe managed to "prove" the FC-Theorem.
- He had a very good idea which used induction.
- He observed that all maps must contain at least one country surrounded with no more than five countries.

# Details

- In the dual form we must have at least one node with degree no more than five.
- Remove the node and colour the rest of the graph with four colours.
- If, necessary, re-colour the graph so that no more than three colours are used around the start-node.
- Kempe "showed" that this can always be done.
- So then we can colour our graph with four colours!



# Not so!

- In fact, the re-colouring which Kempe described does not work.
- This error was undiscovered for ten years!
- The error was then spotted by Heawood.

# Method comments 2

- If a proof is erroneous, it means that there is a counterexample.
- Counterexamples come in two forms:
- Global counterexample - An example which shows that the statement in the theorem is false.
- Local counterexample - An example which shows that a step in the proof is incorrect.
- Kempe's proof fell due to a local counterexample (of course).

# Algorithms

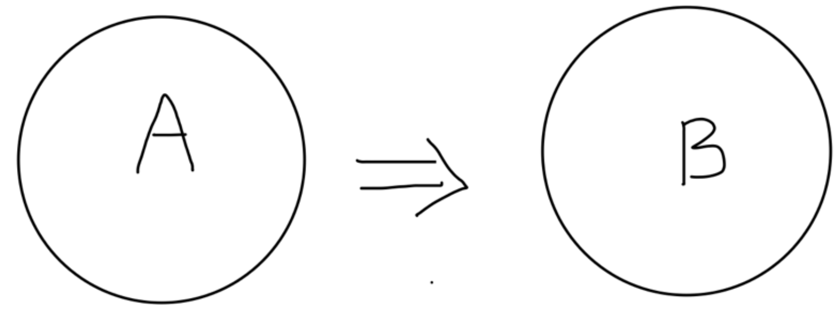
- We can apply the same reasoning to the correctness of algorithms.
- An algorithm takes an input and is supposed to deliver an output of a certain kind.
- An FC-algorithm take a plane graph as input and outputs a FC.
- We can speak of two kinds of counterexamples against the correctness of the algorithm:
- Global counterexample - An example which gives output on the wrong form.
- Local counterexample - An example which makes a certain step in the algorithm impossible to perform.

# Method comments 3

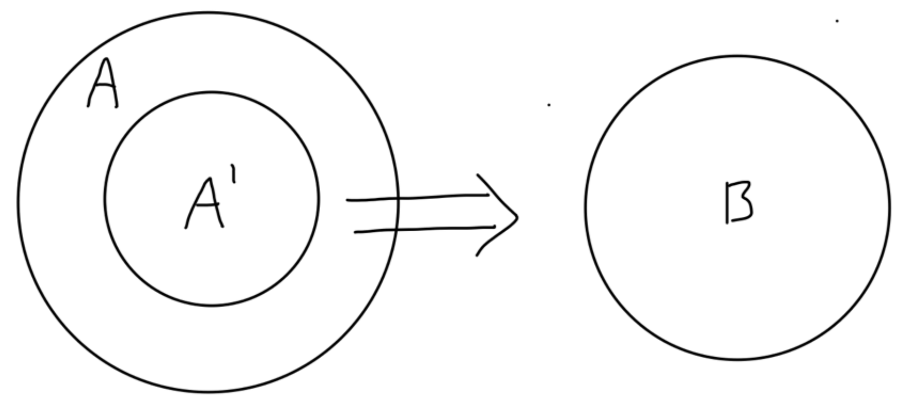
- Let us assume that we have a theorem of the form A => B. (For instance, A: A graph is plane B: The graph can be coloured with four colours.)
- We can *weaken* the theorem by replacing A or B with other statements. The weaker theorem can perhaps be proved.
- 1. Assume A' => A. Then A' => B is a *weakened* form of the theorem.
- 2. Assume B => B'. Then A => B' is a *weakened* form of the theorem.

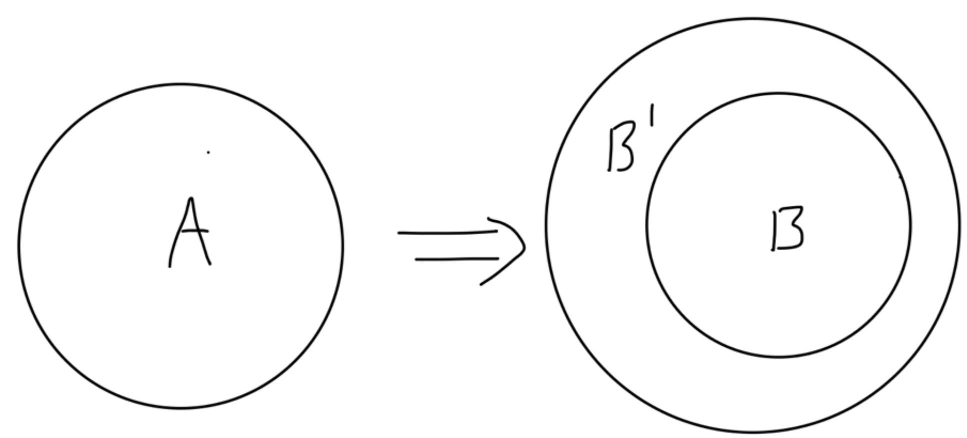
Weakening

The original theorem: A => B



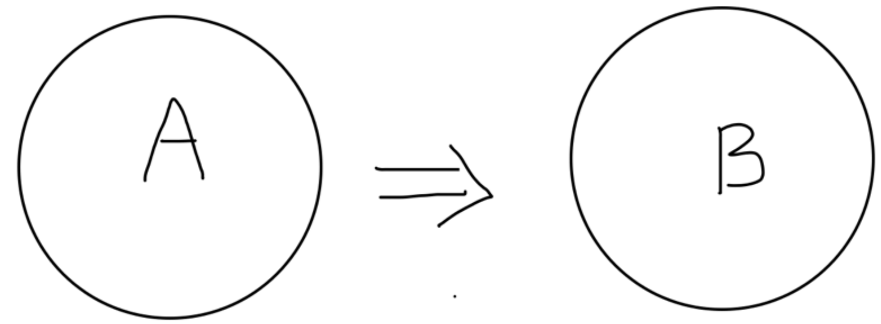

Weaker forms:



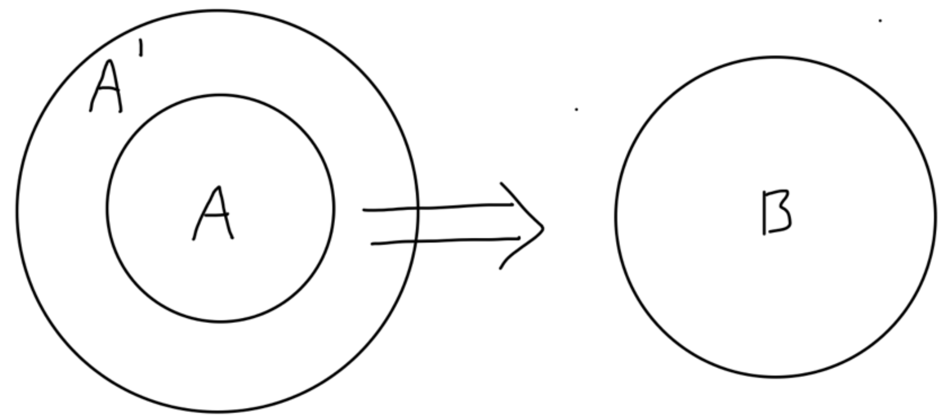



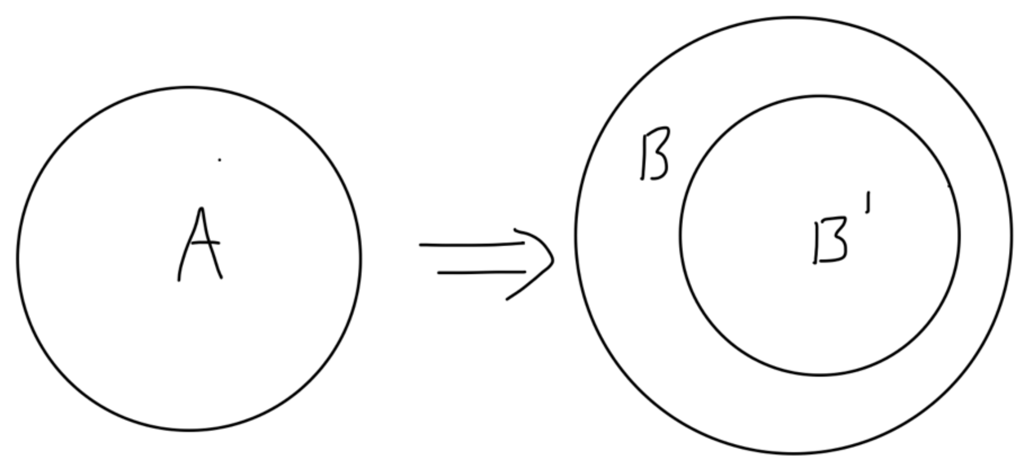
# Strengthening

The original theorem: A => B




Stronger forms:



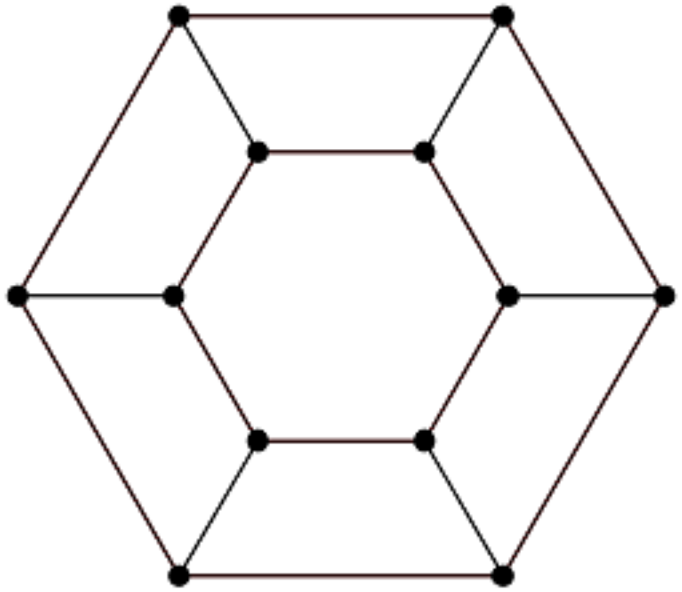



# The Five-colour Theorem

- In 1890 Heawood used Kempe's technique and proved that every plane graph can be coloured with no more than five colours.
- It is obviously a weakening of the FC-Theorem.
- Heawood's proof shows that an erroneous proof (Kempe's) can still be useful.

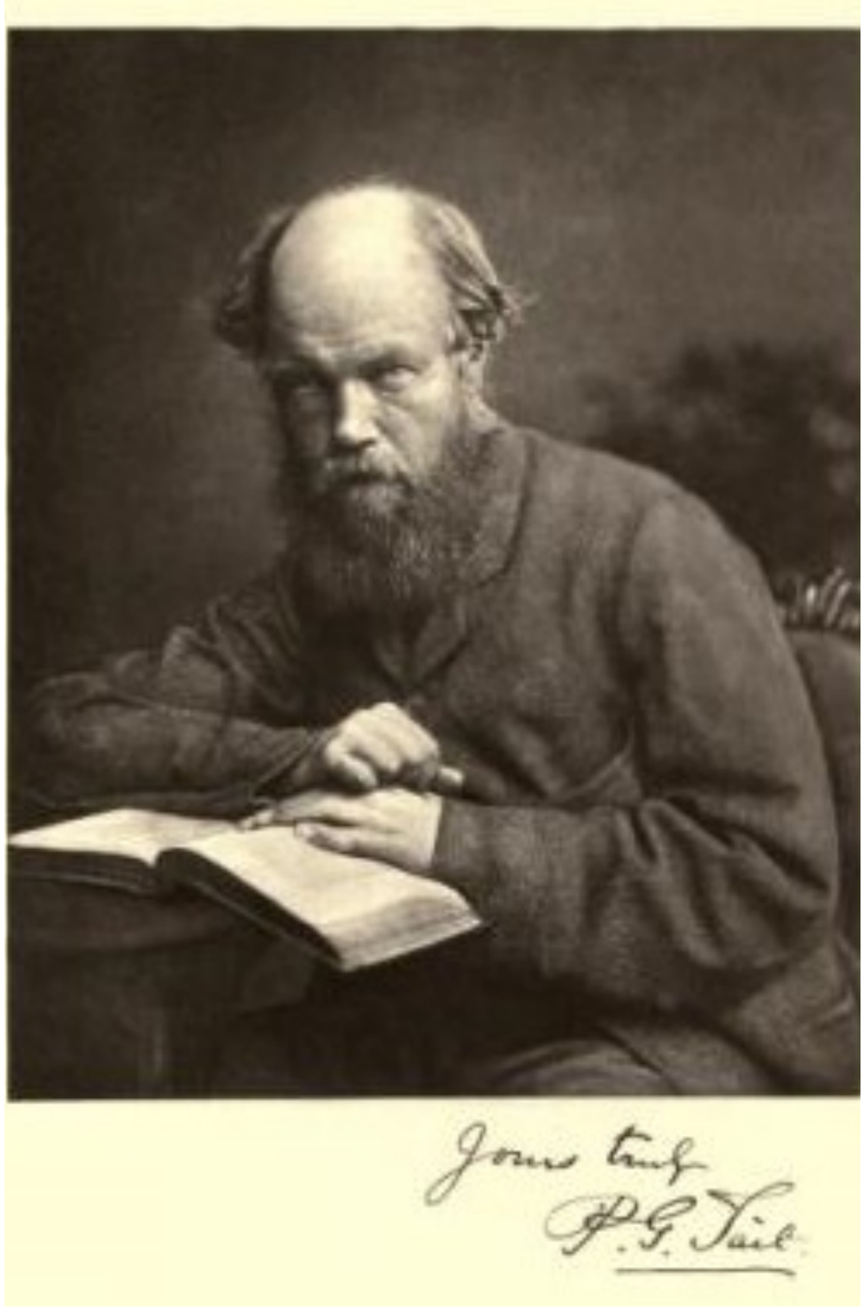
# Another weakening

- Even before Kempe's proof it was known that it is enough to prove the FC-Theorem for *cubic* maps.
- Cubic maps - Maps where all nodes have degree three.



# A reduction

- Tait managed to show that if we can show that every cubic map has a Hamiltonian Cycle, then the FC-Theorem must be true.
- But it turned out that there are (global) counterexamples to this statement, i.e. the existence of Hamiltonian Cycles.



# A new idea: Edge-colourings

- Given a graph we can colour its edges. We say that a colouring is correct if any edges with a common node is coloured with different colours.
- Vizing's theorem: If N is the minimal number of colours needed to colour the graph G and D is the maximal node-degree in G, then N is either D or D+1.
- Tait showed that the FC-Theorem is true if and only if every plane bridgeless cubic graph can be edge-coloured with three colours.

# Method comments 4

- We can speak about different problems. Informally we can say that Problem 2 is weaker than Problem 1 if a solution to Problem 1 would give us a solution to Problem 2.
- In the same way is Problem 1 stronger than Problem 2.
- And if a solution to any of the problems would give us a solution to the other one, we say that the problems are equivalent.

# A comparison with Complexity Theory

- In complexity theory we have the notation ≤ where Problem 2 ≤ Problem 1 means that there is a polynomial time reduction from Problem 2 to Problem 1.
- In our more general discussion we do not have a *formal* definition of reductions in this sense.

# What we have seen this far

- The problem of proving FCT for maps is equivalent to proving FCT to graphs.
- Heawood solved the weaker problem of proving that every plane graph can be 5-coloured.
- It was shown that FCT can be reduced to the (apparently weaker) problem of proving that every plane cubic map is three-colourable.
- Tait showed that FCT could be reduced to the problem of proving that every plane cubic graph has a Hamiltonian Cycle.
- Tait showed that FCT is equivalent to the problem of proving that every plane cubic graph can be edge-three-coloured.

# Turning to harder problems

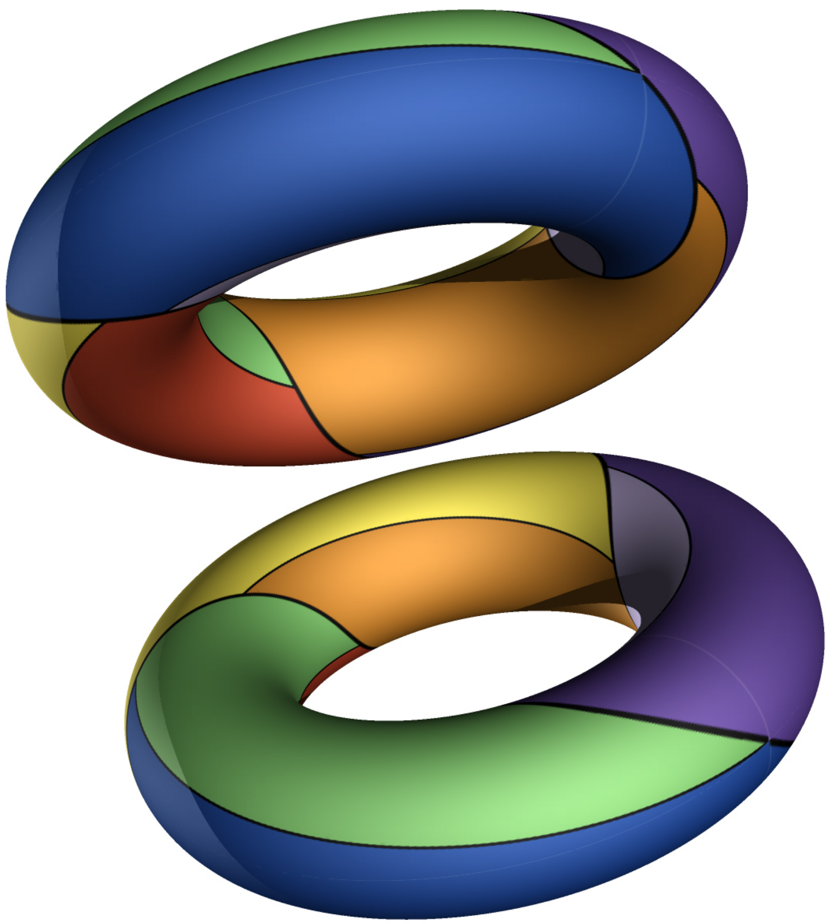
- It turned out that the FCT remained unproved despite all these promising approaches.
- What one could do then is to try to solve a harder problem.

# Chromatic polynomials

- The mathematician Birkhoff tried to solve an apparently harder problem. He wanted to decide in how many ways an arbitrary graph G can be coloured with x colours.
- It turns out that the answer is a polynomial P(G,x), a so called *chromatic polynomial.*
- Birkhoff tried to show that for all planar graphs G we have $P(G,4) > 0$. But he didn't succeed.

# Other types of maps

- Instead of plane maps we can consider maps on other bodies.
- For instance, on a torus it is quite easy to show that seven colours always suffice but not six colours.
- In fact, we can show variants of the FCT for all bodies except for spheres (which are equivalent to planes).



# Method comments 5

- We have seen several promising attempts to prove the FCT. Eventually, none of them gave a proof.
- Nevertheless we see that trying to solve a problem can lead to other interesting problems and solutions to them.

# The proof of the Four-colour Theorem

- The path towards the proof of the FCT starts with a return to Kempe's failed proof from 1879. The proof uses ideas that Kempe had.
- The proof uses induction over the size of the graph. Then we observe that a planar graph must have a set of *unavoidable* subgraphs.
- Then we prove that the subgraphs are *reducable*. This means that if the rest of the graph can be four-coloured, then this colouring can be extended to the subgraph with some minor changes to the original colouring.
- Kempe found the a simple unavoidable subgraph in form of a node with degree at most five. But he failed to prove that the subgraph is reducable (it is not).
- Appel and Haken had the idea that they should try to find more complicated unavoidable subgraphs.

# A computer proof

- Appel and Haken managed to find a set of 1936 *together unavoidable* subgraph. (That means that in any planar graph at least one of the subgraphs must occur.)
- But in order to prove that the subgraphs were *reducable* they had to rely on a computer program to find the re-colouring strategies.
- The proof became much debated and criticized. It opened for a discussion of what a proof really is or should be.

# Method comments 6

- So eventually the original idea by Kempe was triumphant.
- But in 1890 there was probably no easy way to see this.
- It was when all other strategies had failed that the return to the original idea seemed attractive.
- So sometimes a failed proof can be resurrected.

# Quantitative Data Analysis

We briefly describe some statistical methods for analysis of data. The methods are *parametrical,* i.e. we make assumptions about the distributions of the stochastic variables we measure. Two methods you should know are:

- Hypothesis Testing
- Maximum Likelihood-Method
- Linear regression and correlation

# Hypothesis Testing

Let us assume that we have a hypothesis we want to test. We compare it to a zero-hypothesis H0. We design a test which gives us a value t. We define a set C such that we reject H0 if t is a member of C. (That means that we accept H.) In that case we say that the test is significant at level a if the probability that t belongs to C is less than or equal to a, given the assumption that H0 is true. The probability a is usually small, like 0.05, 0.01 or 0.001.

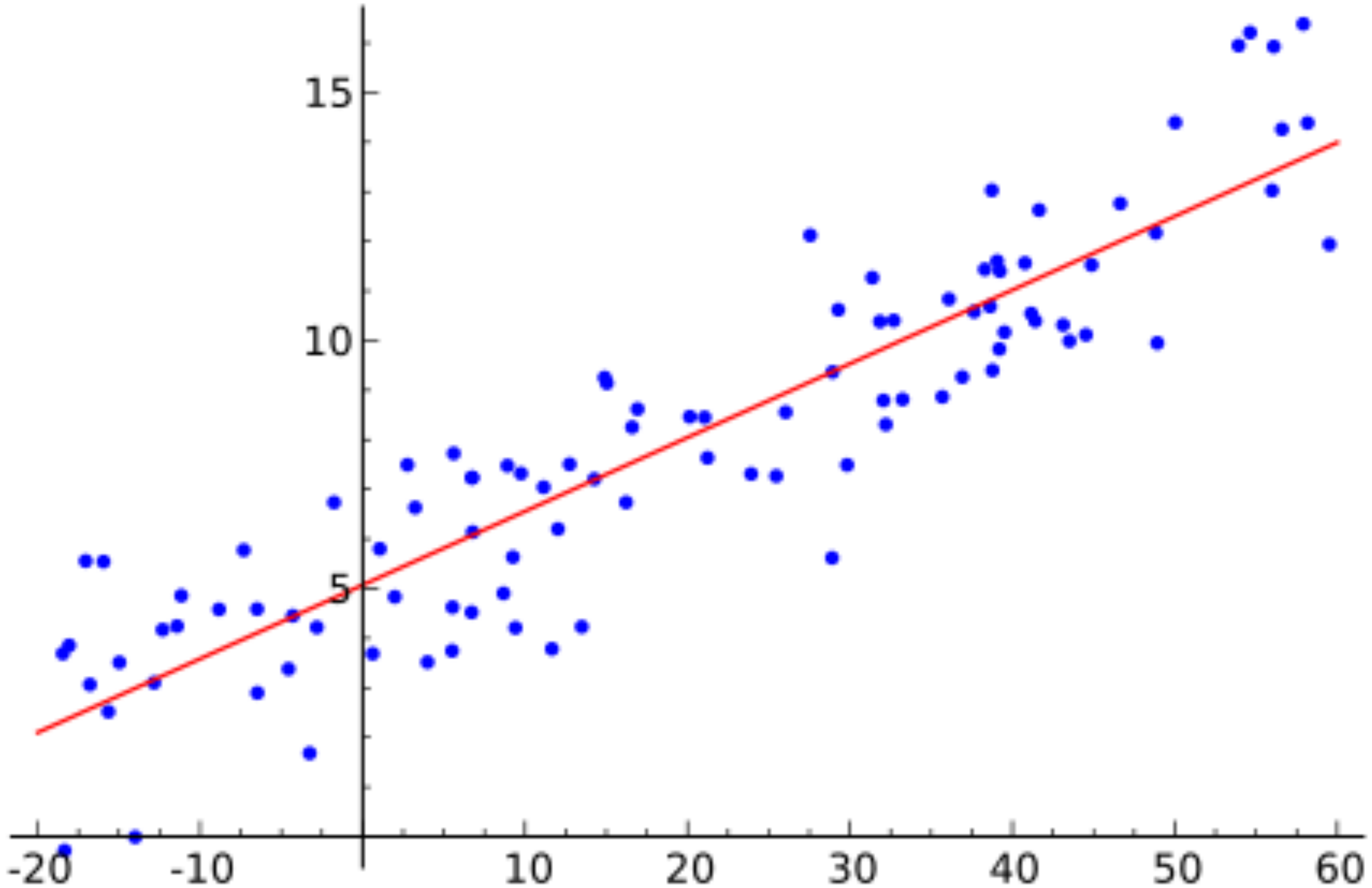
# The Maximum Likelihood-method

The method can be illustrated with two special cases:

1. Let us assume that we want to find the value of a parameter f. We do an experiment which gives us a value t. We then (analytically) find the value f0 for the parameter which maximizes the probability P(t l f = f0). We then say that f has the value f0.

2. Let us assume that we have made an observation E. We have different hypotheses H1, H2, ... , Hn which could be possible causes for E. We then chose the hypothesis Hi which maximizes P(E l Hi) and say that Hi is the cause for E.

# Linear regression



# Correlation

When data look like these in the graph they are obviously correlated. The regression line can be found with the Least Square-Method. If we want a measure to tell us how strong the correlation is we can use the *correlation coefficient*. It is computed in the following way:

$$C_{xy} = \frac{1}{n-1}(\sum x_i y_i - \frac{1}{n}(\sum x_i)(\sum y_i))$$

$$s_x = \sqrt{\frac{1}{n-1}(\sum x_i^2 - \frac{1}{n}\sum x_i)}$$

$$s_y = \sqrt{\frac{1}{n-1}(\sum y_i^2 - \frac{1}{n}\sum y_i)}$$

$$r = \frac{C_{xy}}{s_x s_y}$$

# Data

En del i forskningsprocessen är att hantera data av olika slag. Det visar sig att man kan beskriva data på flera olika sätt.

- Beskrivning efter abstraktionsnivå
- Uppdelning i primära och sekundära data
- Kvantitativa och kvalitativa data
- Mätning efter olika typer av skala

# Primära och sekundära data

- Primära data - Är direkta mätningar eller observationer av något. Kan också vara rapporter av personer som själva varot med om något
- Sekundära data - Är ofta sammanställning av primära data som har behandlats i någon form. De kan t.ex. förekomma i rapporter, artiklar eller böcker

# Kvantitativa och kvalitativa data

- Kvantitativa data - Sådana data som ges i form av tal
- Kvalitativa data - Data som inte enkelt kan ges i form av tal. Det kan vara åsikter, berättelser eller beskrivningar av situationer

# Samla primära data

Mängden metoder är stor. Vi kan dock nämna tre huvudområden

- Statistiska undersökningar. Sampling
- Intervjumetoder. Detta kommer att tas upp i en senare föreläsning
- Experiment. Några generella metoder för experiment är att man jämte den riktiga experimentgruppen också har en kontrollgrupp. Man bör också om möjligt använda så kallade dubbelblinda test

# Tre speciella analysmetoder

- Content analysis - Vi räknar helt enkelt förekomster av något t.ex. ord eller typer av bilder i dokument och använder förekomsterna som indikatorer
- Data mining - Vi använder program för att hitta mönster i data
- Metaanalys - Vi gör en analys av flera andra analyser samtidigt och försöker se mönster i dem