

Cēloņsakarību loma statistikas analīzē



Aldis Ērglis

Emergen Latvija

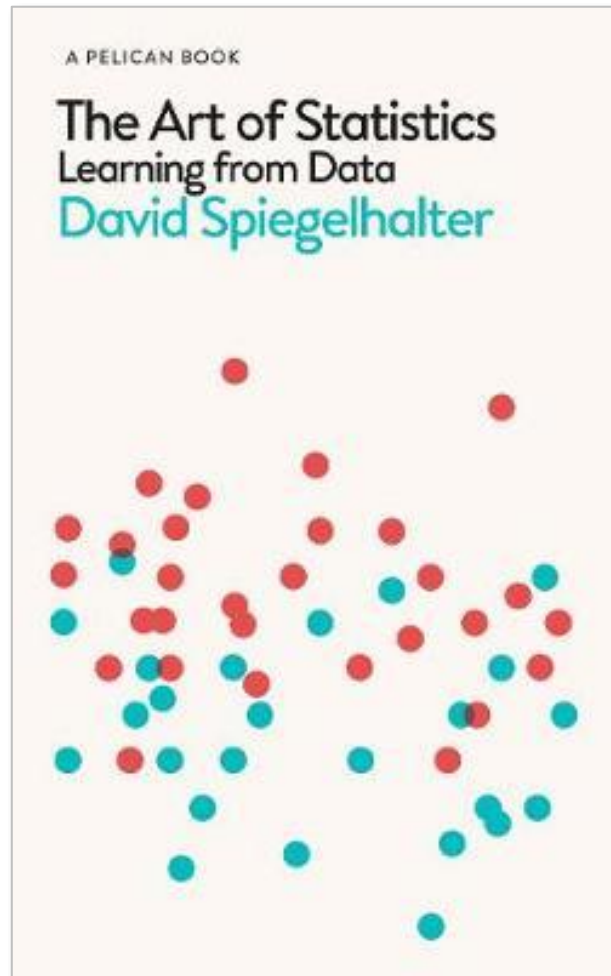
#STATISTIKAI100



Cēloņsakarību loma statistikas analīzē

Aldis Ērglis

The Art of Statistics: Learning from Data



- Statistikas pieeja spēlējusi nozīmīgu lomu pasaules zinātniskā izpratnē
- Lielo datu laikmetā statistikas zināšanas ir svarīgākas kā jebkad agrāk

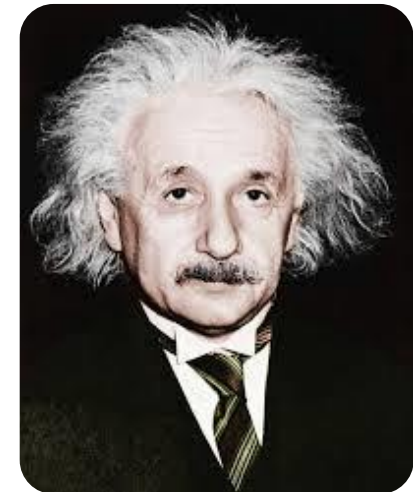
Machine Teaching (Mašin-mācīšana)

The screenshot shows the Microsoft Research ODP (Open Domain Parser) interface. The main window is titled 'ODP Ingredient CHANGE'. On the left, there is a 'Review Model' section with a bar chart showing the model's confidence in its predictions. The chart has a y-axis from 0.0 to 1.0 and a series of green and red bars. Below the chart are 'teach' and 'publish' buttons. The main 'Explore' section displays a recipe for 'Picnic' with various ingredients and preparation instructions. The recipe text is: 'Occasion: Picnic', 'Preparation Method: Combine', 'Product Type: Roasted Turkey Breast', 'Ingredients: 20 Each whole wheat pita, 10 Ounces spinach leaves, washed and dried, 1 Pound, 14 Ounces cucumbers, thinly sliced, 1 Pound, 4 Ounces red or yellow bell pepper, seeded and cut into strips, 10 Ounces alfalfa sprouts, washed and dried, 80 Slices tomato, (about 14 tomatoes), 1 Pound, 14 Ounces ranch dressing, prepared, 4 Pounds, 6 Ounces OVEN ROASTED TURKEY BREAST, shaved'. A text box at the bottom of the recipe text says 'like this, to increase the accuracy.' The 'Entities' section on the right lists categories: Ingredient, Amount, Unit, Food, Preparation. The 'Explain' section shows a list of entities: 2 pack, Alcohol, Breads, Cheese, Combiner Words, Dairy, End Parens, Fruits and Veg. The status bar at the bottom indicates 'Server: Connected | Session: Valid | Training: Completed | Training Accuracy: 87.207% (RM) | RETRAIN'.



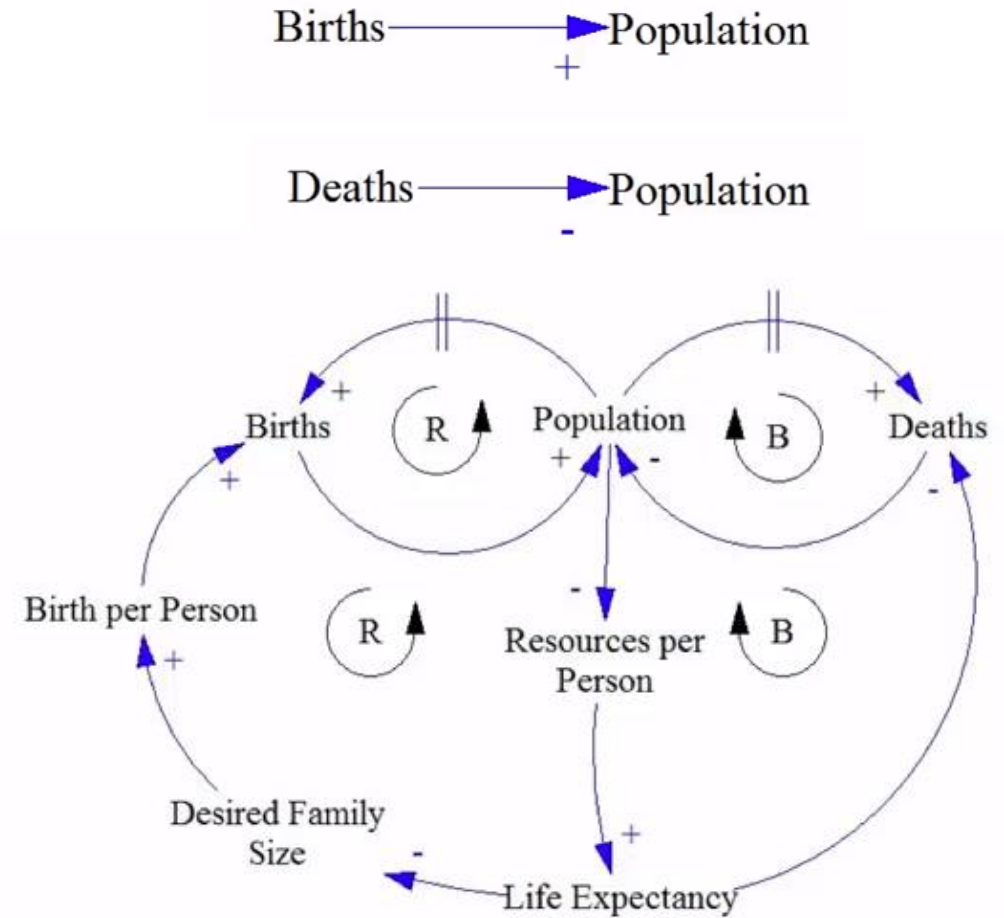
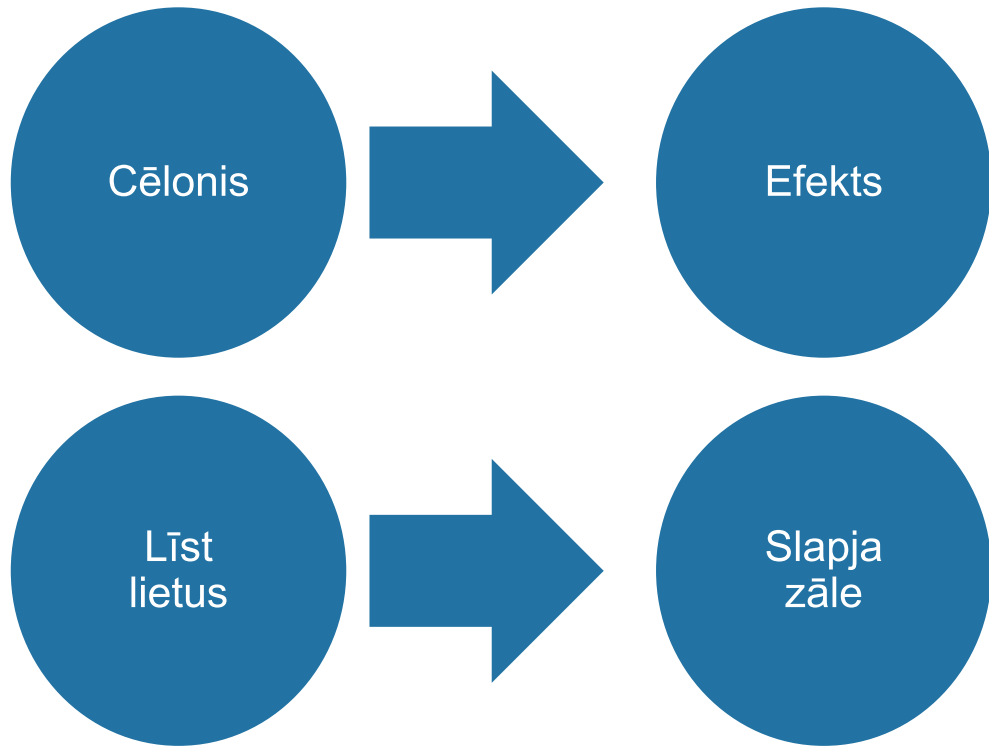
Prologs

Development of Western science is based on two great achievements: the invention of the formal logical system (in Euclidean geometry) by the Greek philosophers, **and the discovery of the possibility to find out causal relationships by systematic experiment (during the Renaissance).**

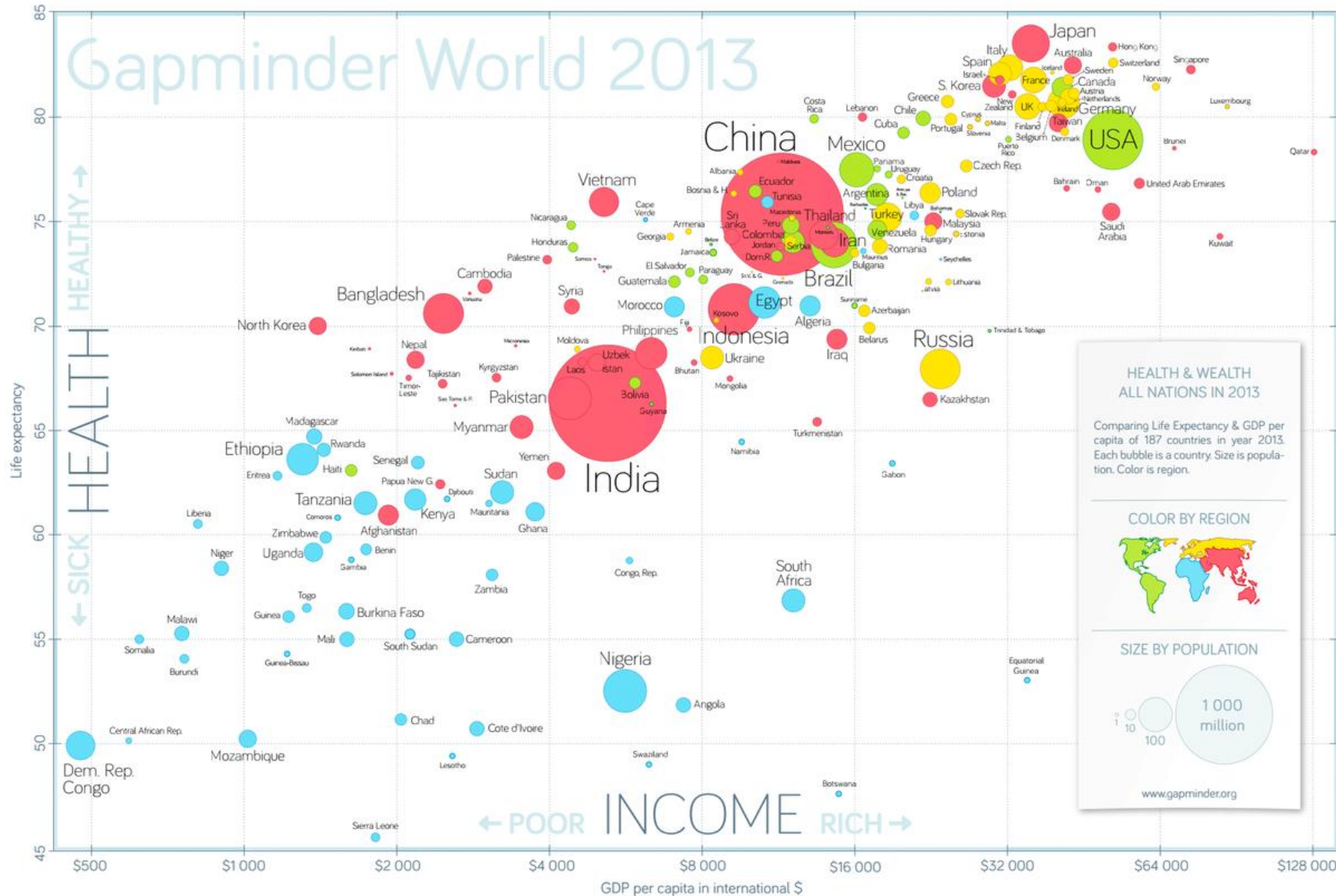


Albert Einstein
(1953)

Kas ir cēloņsakarības?



Vai ir atkarība starp dzīves ilgumu un ienākumiem?

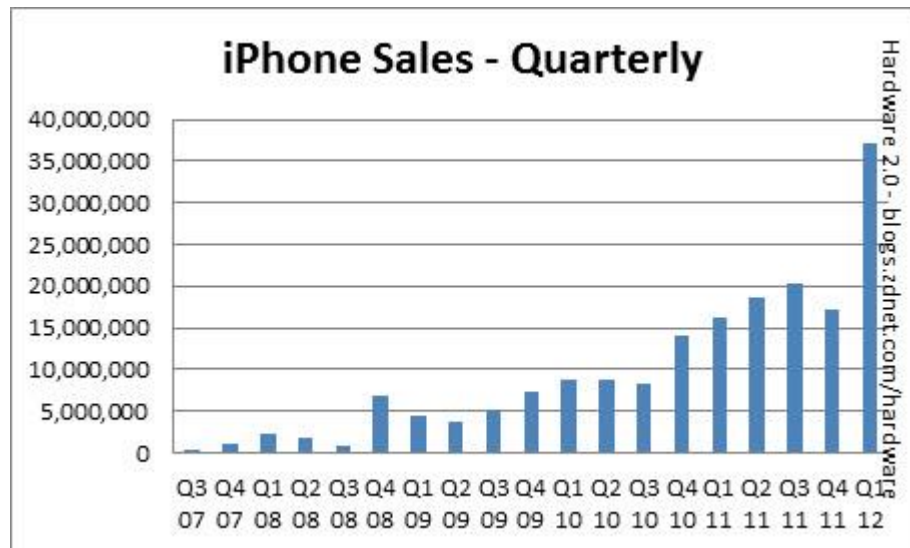


GAPMINDER

<https://www.gapminder.org>

Datu analīzes mērķis ir labāki lēmumi

Pieejamie dati ierobežo domāšanu, mēs domājam par to, par ko mums ir dati



Intuīcija (*nevaram izskaidrot*)



Dati + ??? = Lēmums



Cēloņsakarības (*varam izskaidrot*)

Mēs veidojam vairāk atskaišu kā jebkad

The image displays a screenshot of the Power BI Desktop interface on the left, showing a dashboard with various visualizations: a sunburst chart, a stacked bar chart, a horizontal bar chart, a line chart, and a bubble chart. A 'Visuals' pane is open, showing a list of visual types and their filters. On the right, 'The Visuals Reference' guide is shown, titled 'THE VISUALS REFERENCE - FOR MICROSOFT POWER BI - SEP. 2018'. The guide is organized into categories: COMPARISON, CHANGE OVER TIME, RANKING, SPATIAL, FLOW, DISTRIBUTION, CORRELATION, SINGLE, FILTER, NARRATIVE, and MISCELLANEOUS. Each category contains a grid of visual icons. A legend at the bottom of the guide explains the icons: a green square for 'Recommended', a yellow square for 'There is a better alternative', a grey square for 'Don't use in the category', a blue square for 'Custom visual', a purple square for 'Certified visual', and a red square for 'Required'. The footer of the guide includes the text '2018 © SQLBI AND ORNL ARE TRADEMARKS OF SQLBI CORP. | POWER BI IS A TRADEMARK OF MICROSOFT CORP.'

Kas ir efektīva atskaite?

Atskaite - informācija



Darbinieki



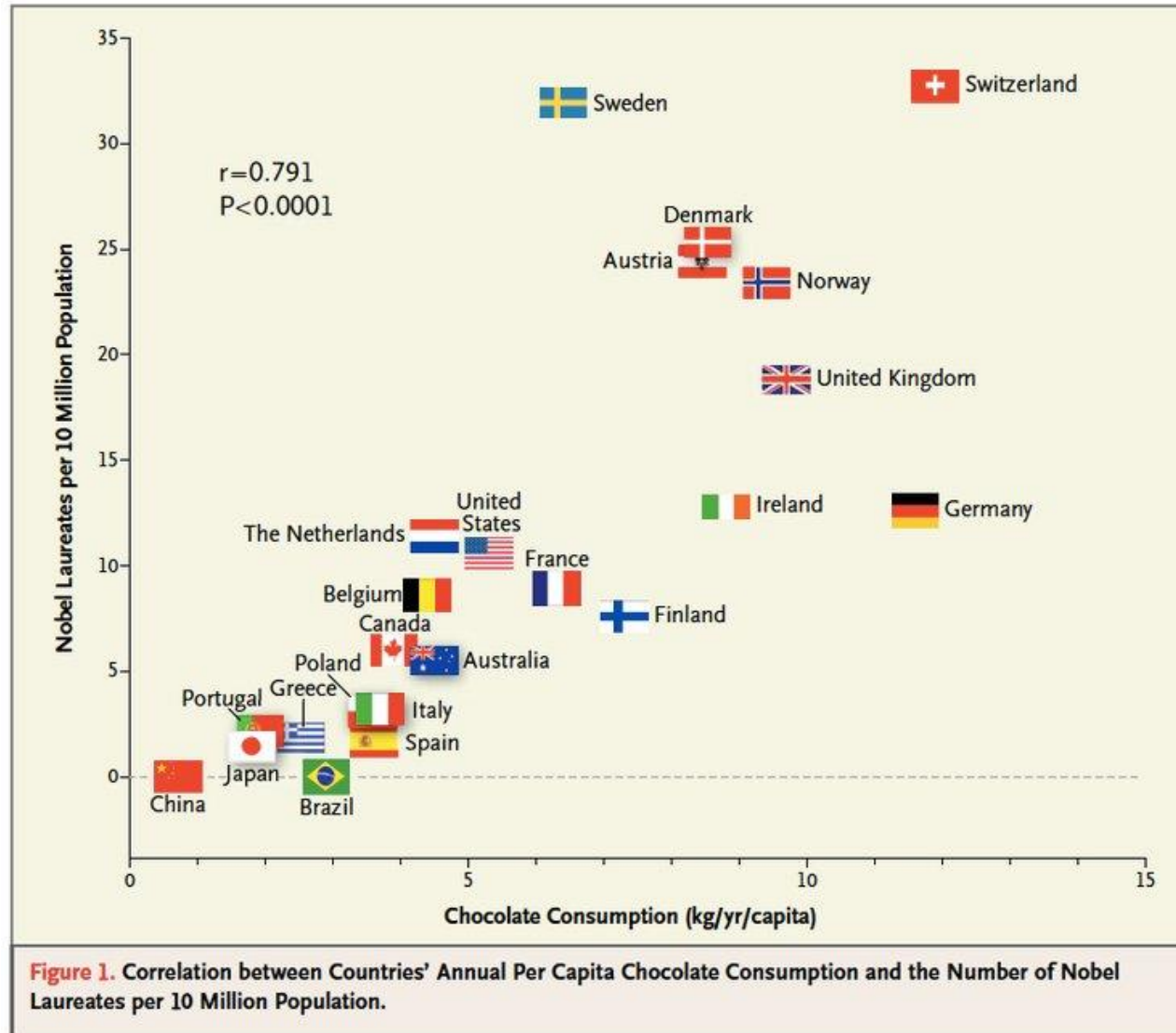
Savādāka rīcība



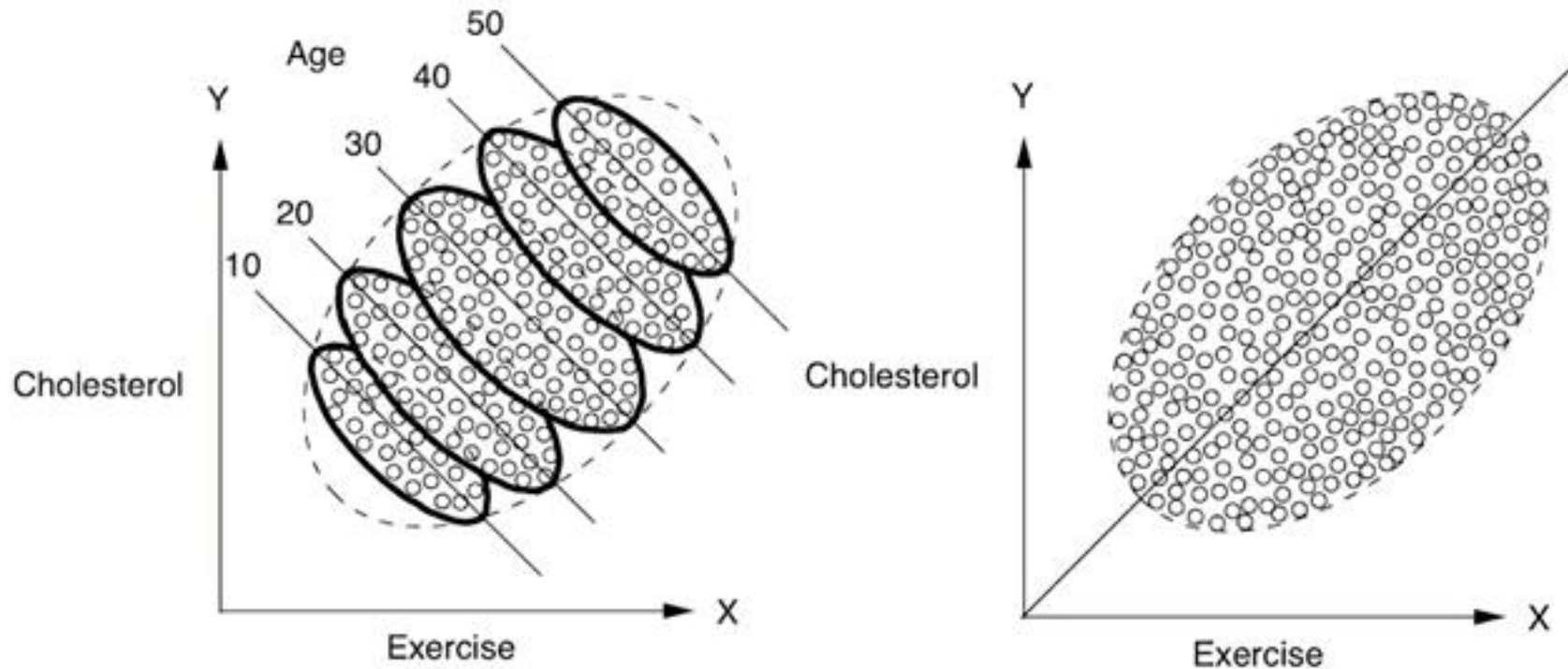


Svarīgākais jautājums datu
analīzē – KĀPĒC?

Korelācijas nav cēloņsakarība

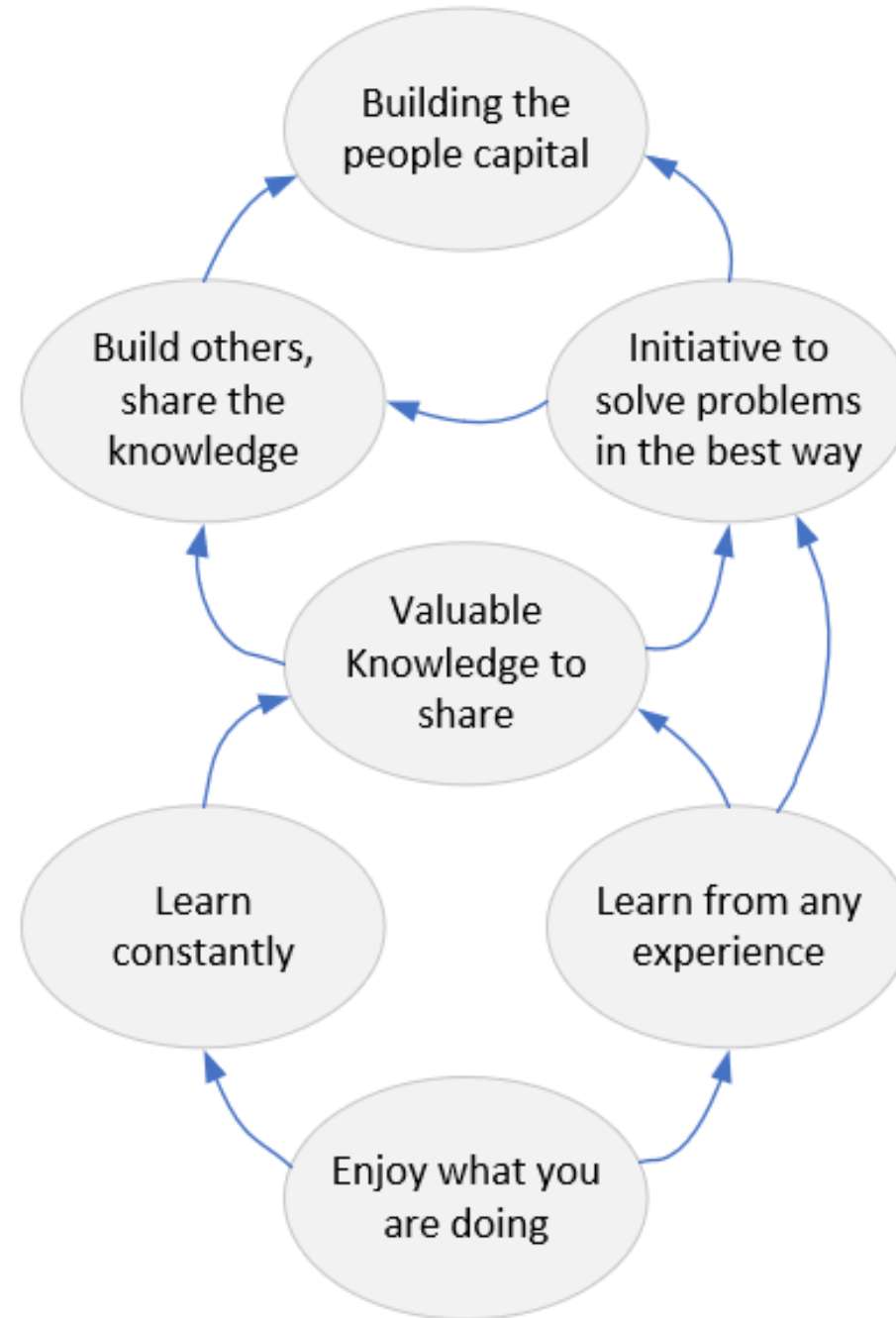


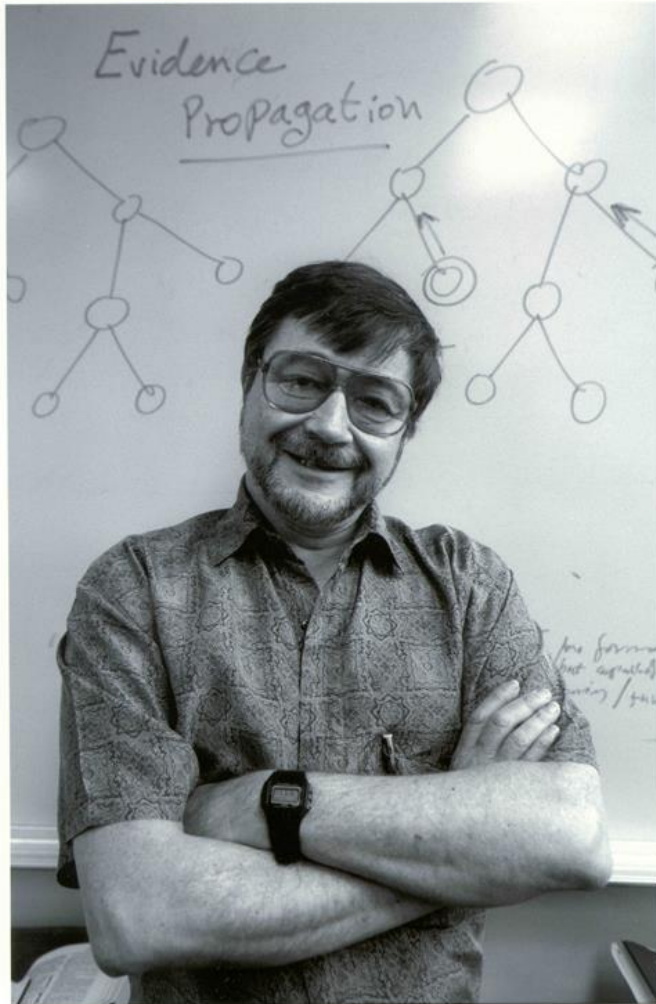
Kāpēc secinājumi var būt nepareizi?



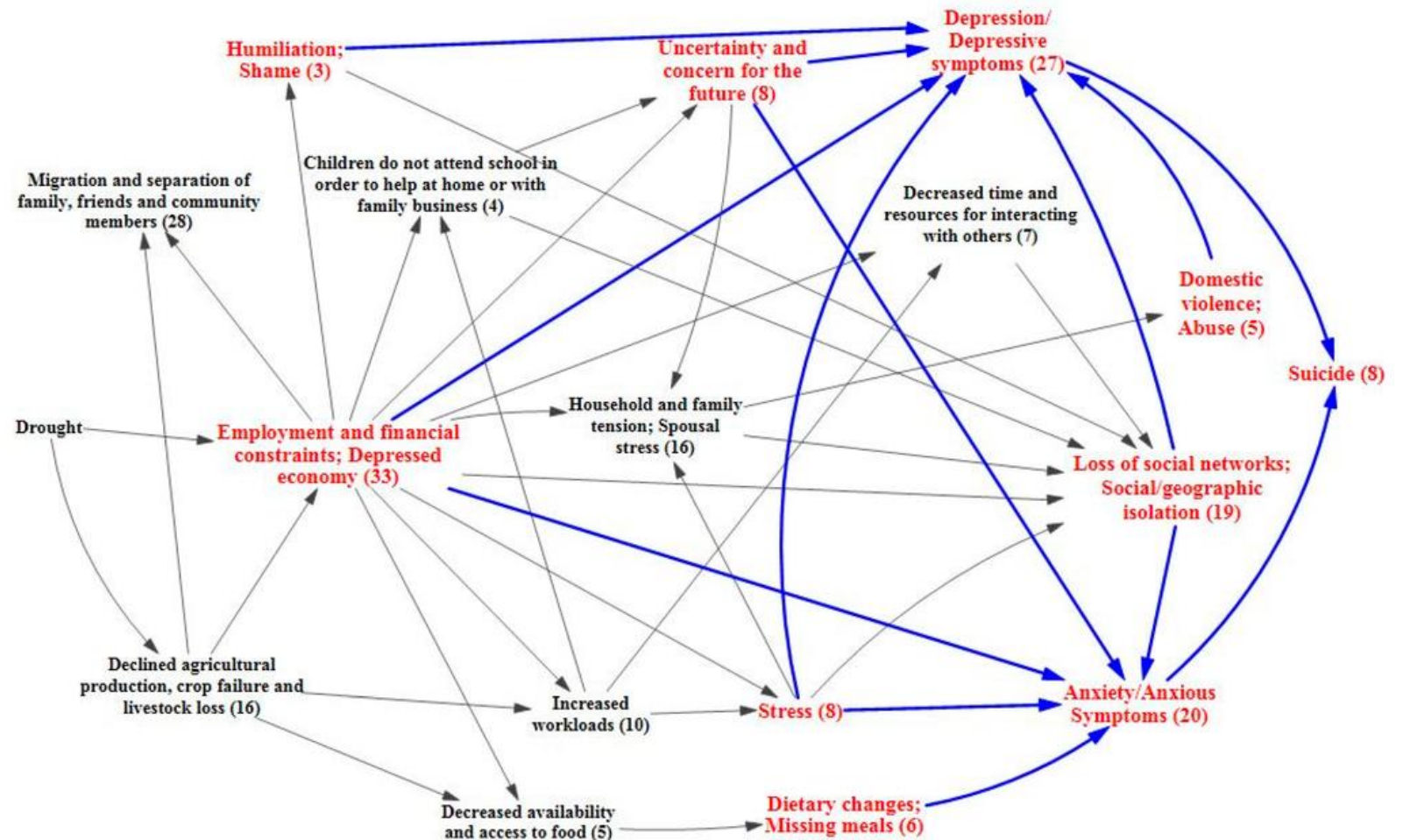
Exercise is helpful in every age group but harmful for a typical person.

Kas ietekmē organizācijas cilvēcisko kapitālu

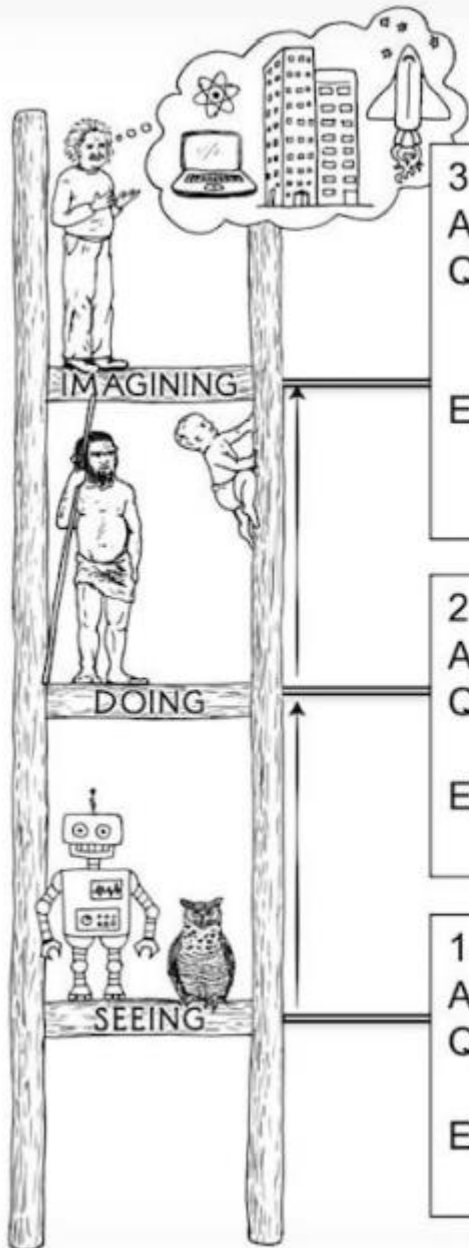




• Judea Perl



3-LEVEL HIERARCHY



3. COUNTERFACTUALS

ACTIVITY: Imagining, Retrospection, Understanding

QUESTIONS: *What if I had done . . . ? Why?*
(Was it X that caused Y? What if X had not occurred? What if I had acted differently?)

EXAMPLES: Was it the aspirin that stopped my headache?
Would Kennedy be alive if Oswald had not killed him? What if I had not smoked the last 2 years?

2. INTERVENTION

ACTIVITY: Doing, Intervening

QUESTIONS: *What if I do . . . ? How?*
(What would Y be if I do X?)

EXAMPLES: If I take aspirin, will my headache be cured?
What if we ban cigarettes?

1. ASSOCIATION

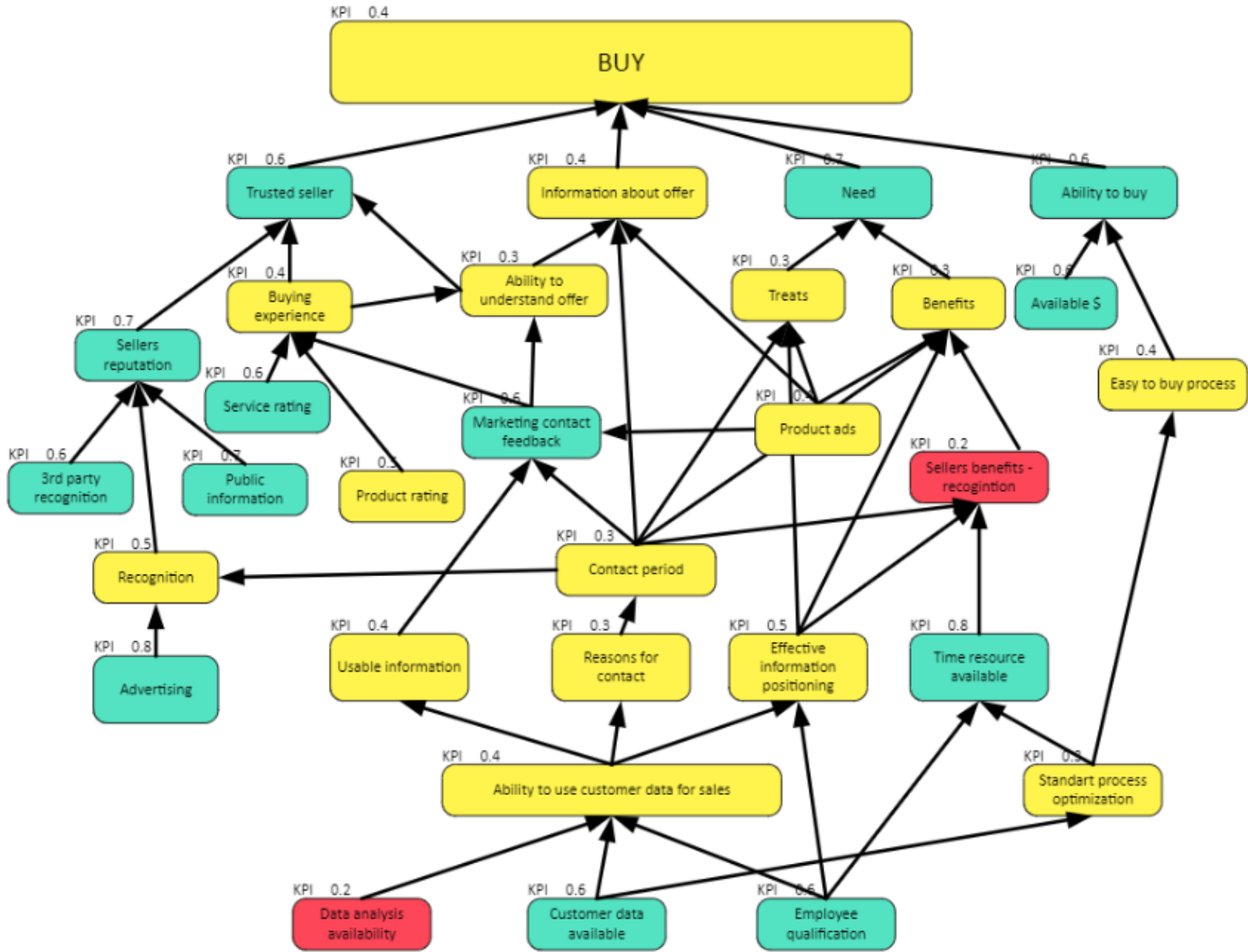
ACTIVITY: Seeing, Observing

QUESTIONS: *What if I see . . . ?*
(How would seeing X change my belief in Y?)

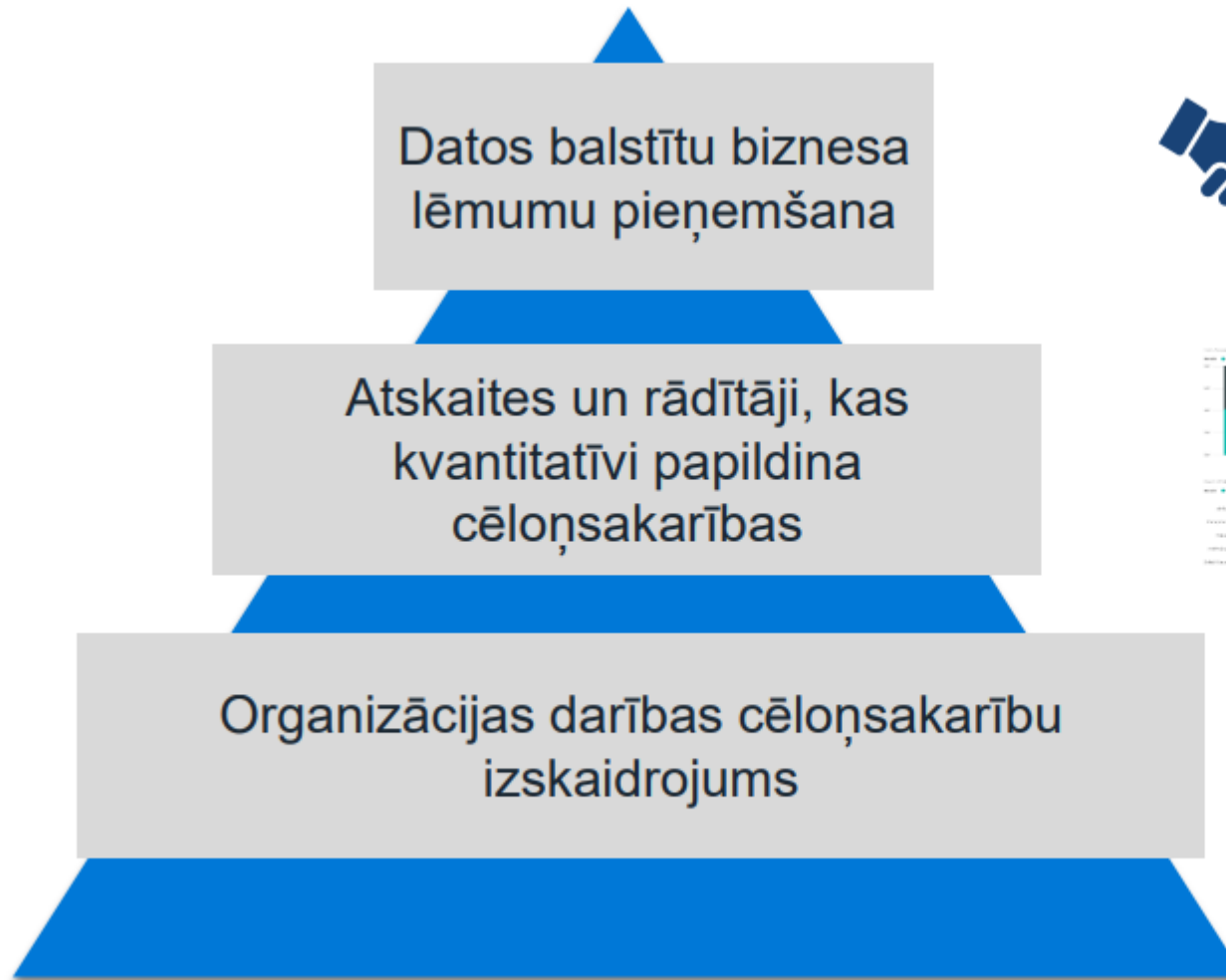
EXAMPLES: What does a symptom tell me about a disease?
What does a survey tell us about the election results?

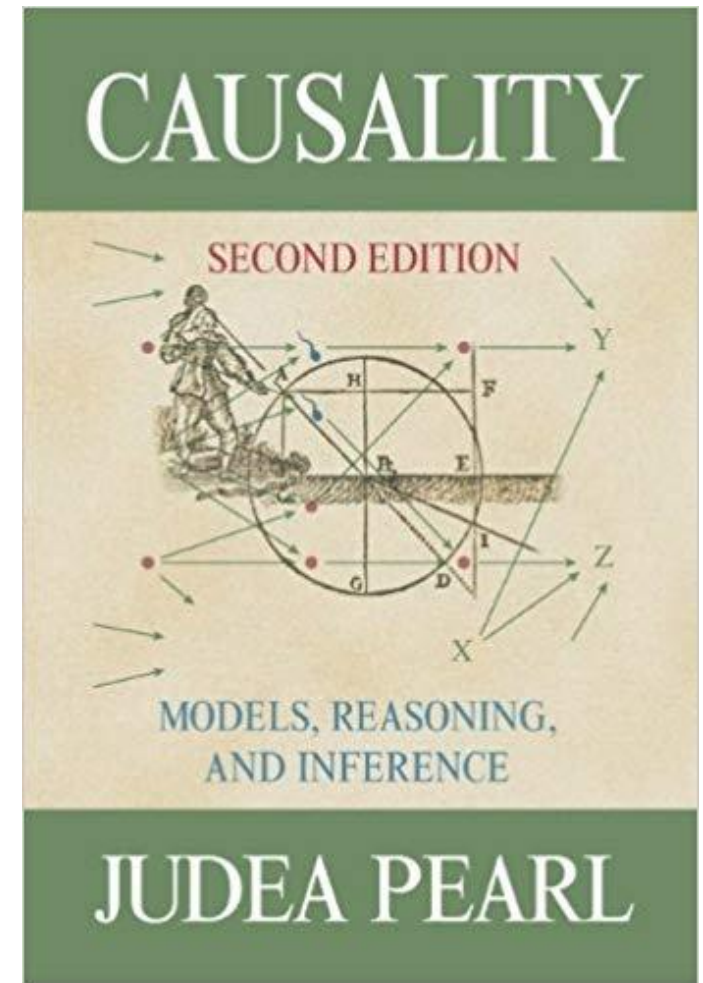
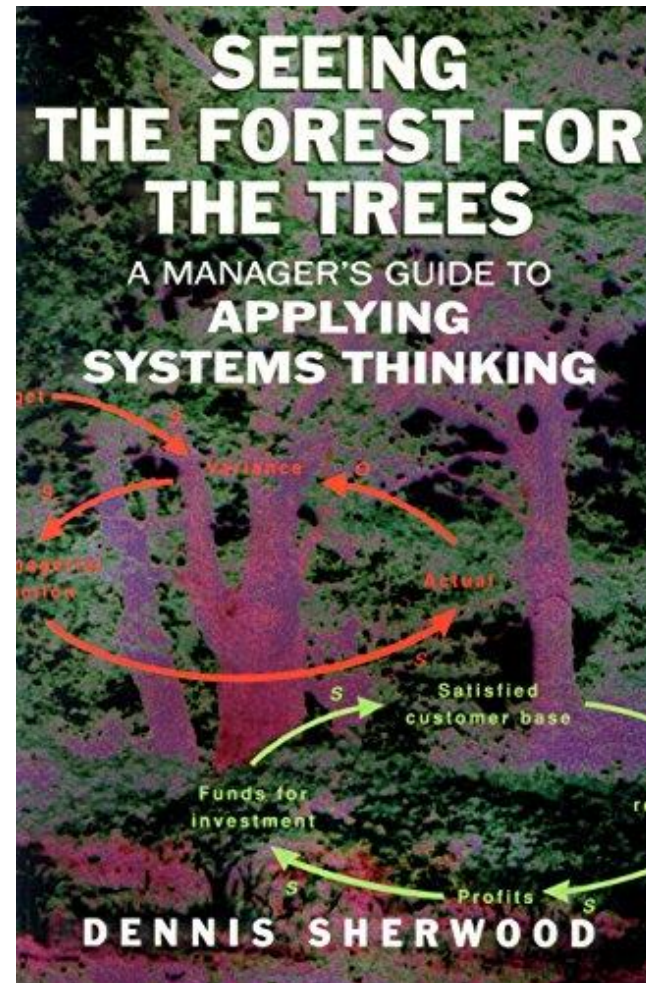
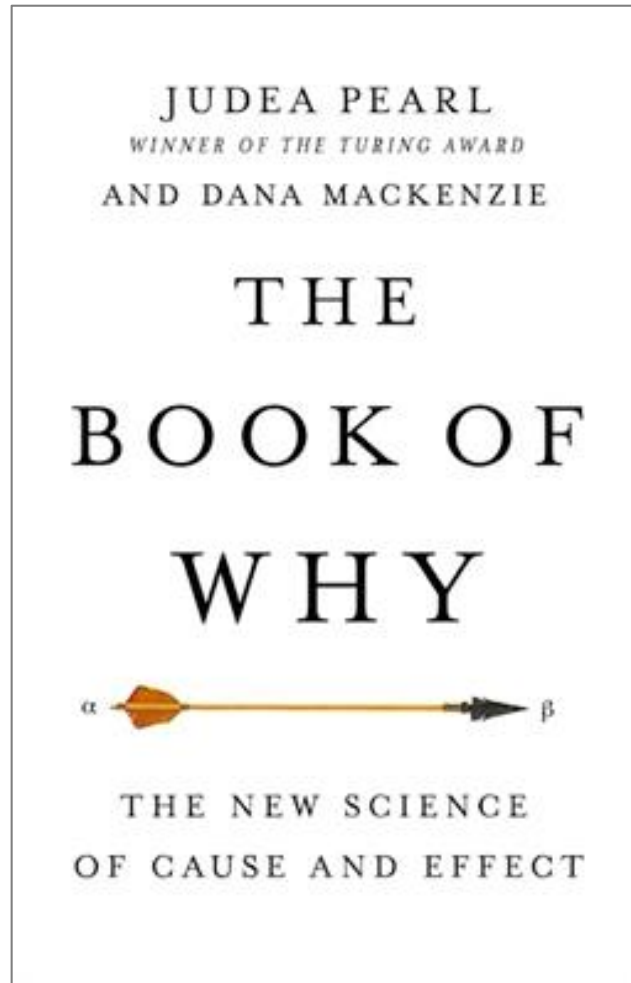
Cēloņsakarības –
veids kā domā tikai
cilvēks

Datos var atrast
abildes uz visiem
jautājumiem, izņemot
jautājumu: Kāpēc?



Priekšnoteikumi datos balstītu lēmumu pieņemšanai







Paldies

Aldis.erglis@emergn.com