

Dataskеpticism

Mám data.

Data samotná jsou v pořádku.

**Data interpretuji do formálně
správné vizualizace.**

Něco je špatně, lidé se smějí.

Proč?



Otázky

Mám opravdu kompletní data?

Nepřehlížím něco?

**Nezapomněl jsem na nějaký
skrytý vliv?**

Nedělám nějakou naivní chybu?

1

ceteris paribus

**„Ostatní stejné“ — zafixujeme
všechny veličiny kromě těch,
které sledujeme.**

Vynikající v laboratorních podmínkách.

**V reálném světě princip ceteris paribus prakticky nefunguje.
Reálný svět je totiž komplexní systém.**

**subjektivní
poznání**

známé známé

data

známé neznámé
hypotéza

neznámé známé

lidé neznají

sami sebe

neznámé neznámé

!!!

V datech nemusí být všechno.

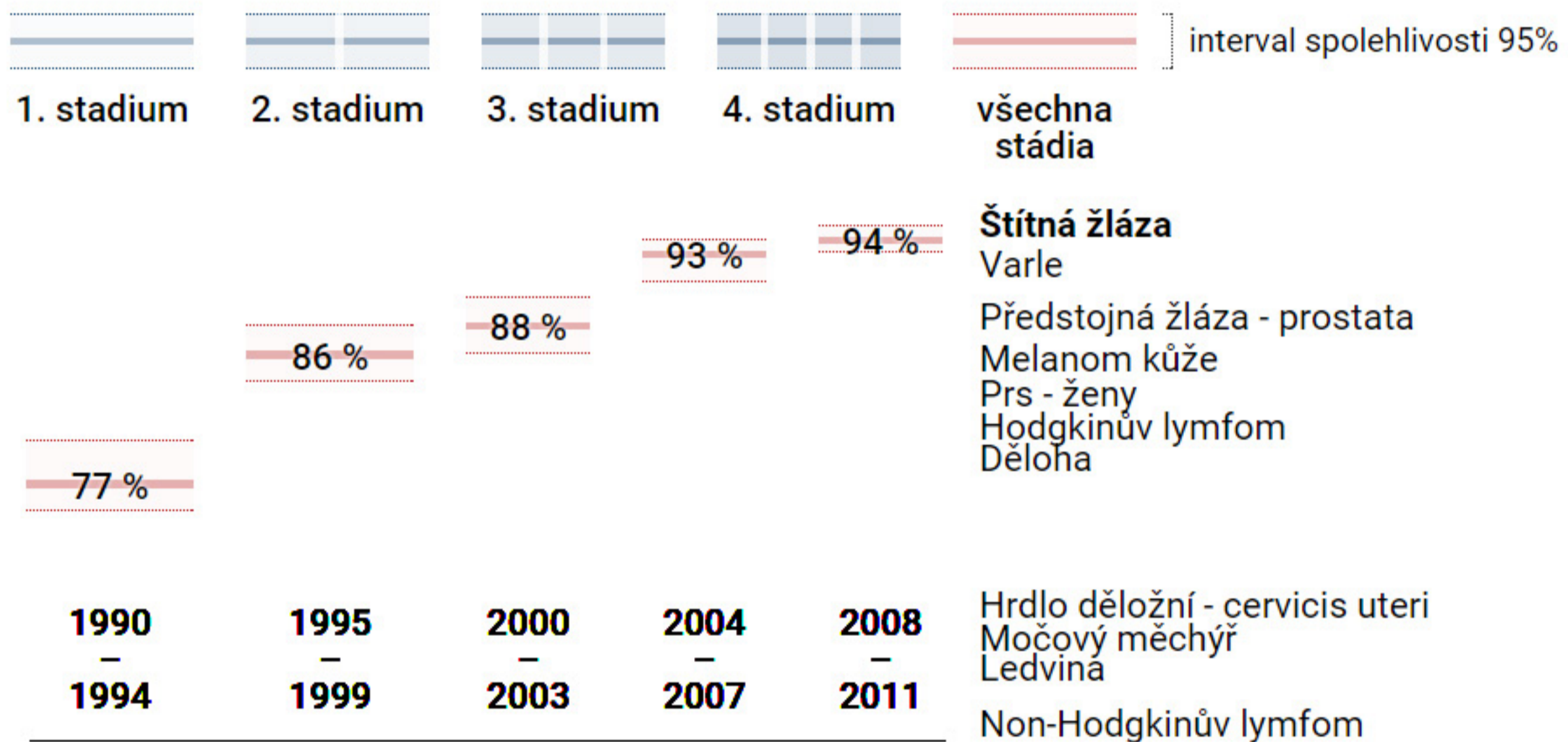
Vždy si uvědomte, že existují věci, jevy, skutečnosti, o kterých nevíte, že je nevíte. Mohou být (a často jsou) ve skutečnosti důležitější, než to co víte.



**Neznámé neznámé může
být subjektivní.**

**To, že nevíme, že něco nevíme,
nutně neznamená, že to neví
někdo jiný.**

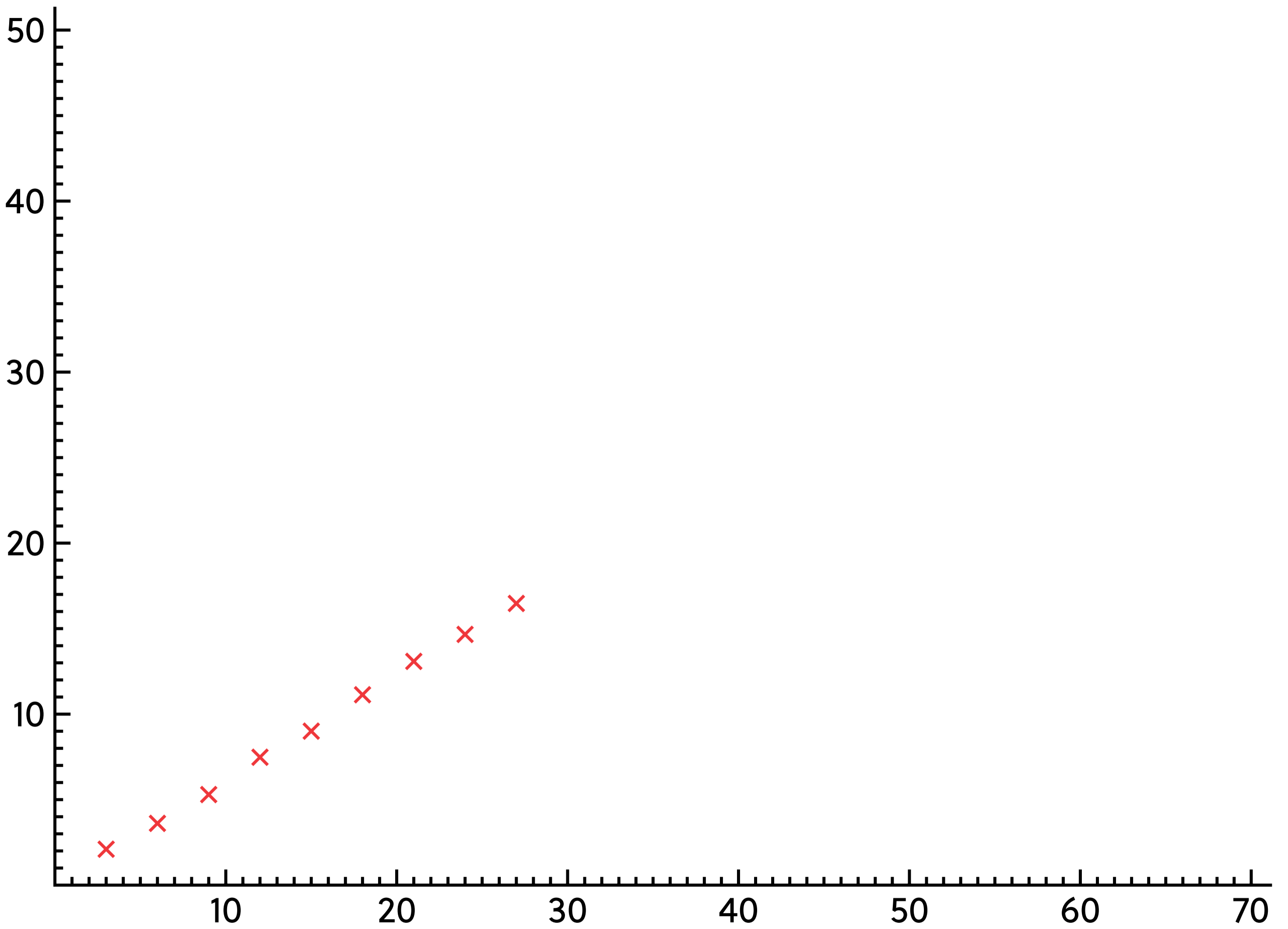
Příklad

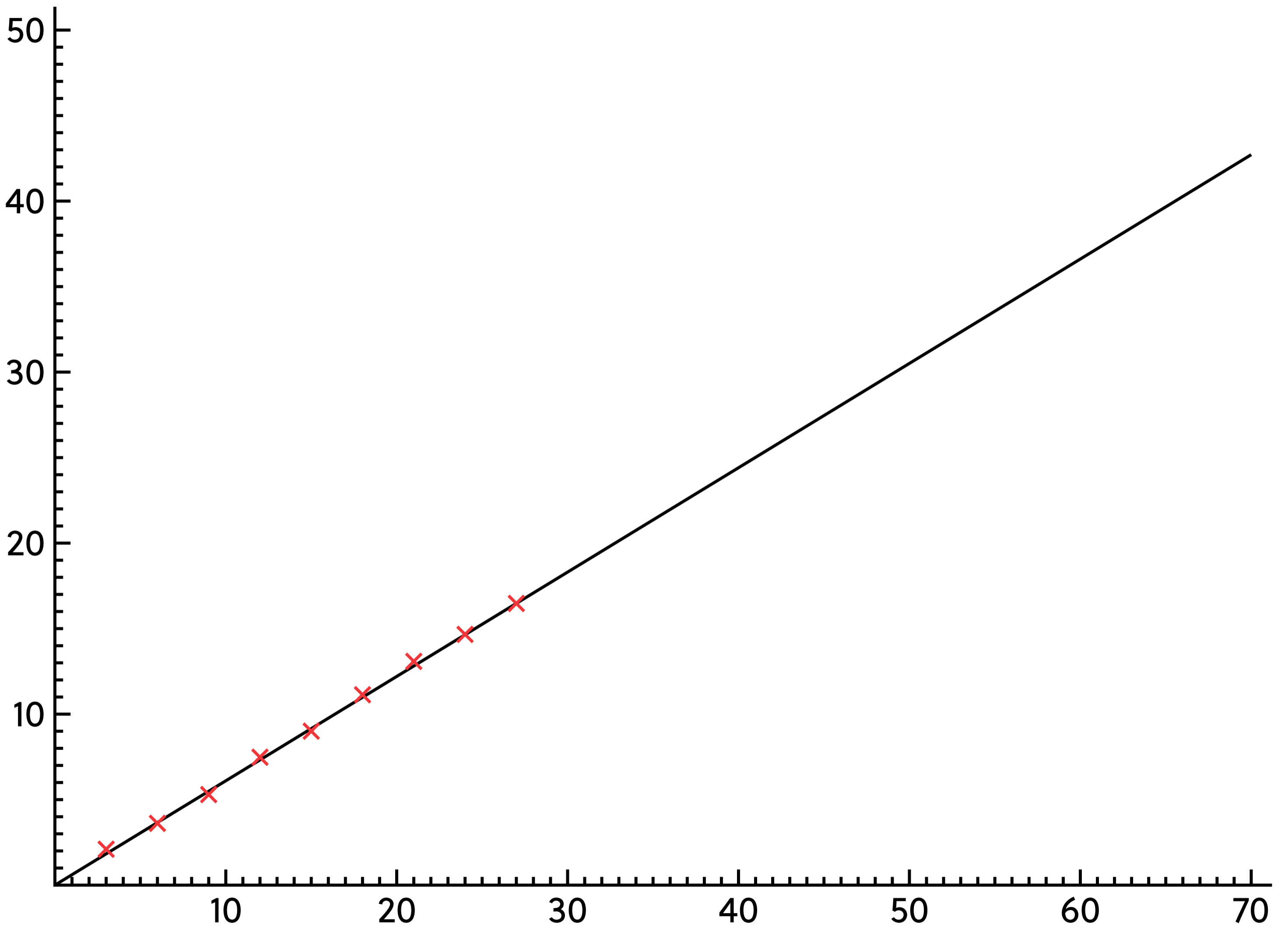


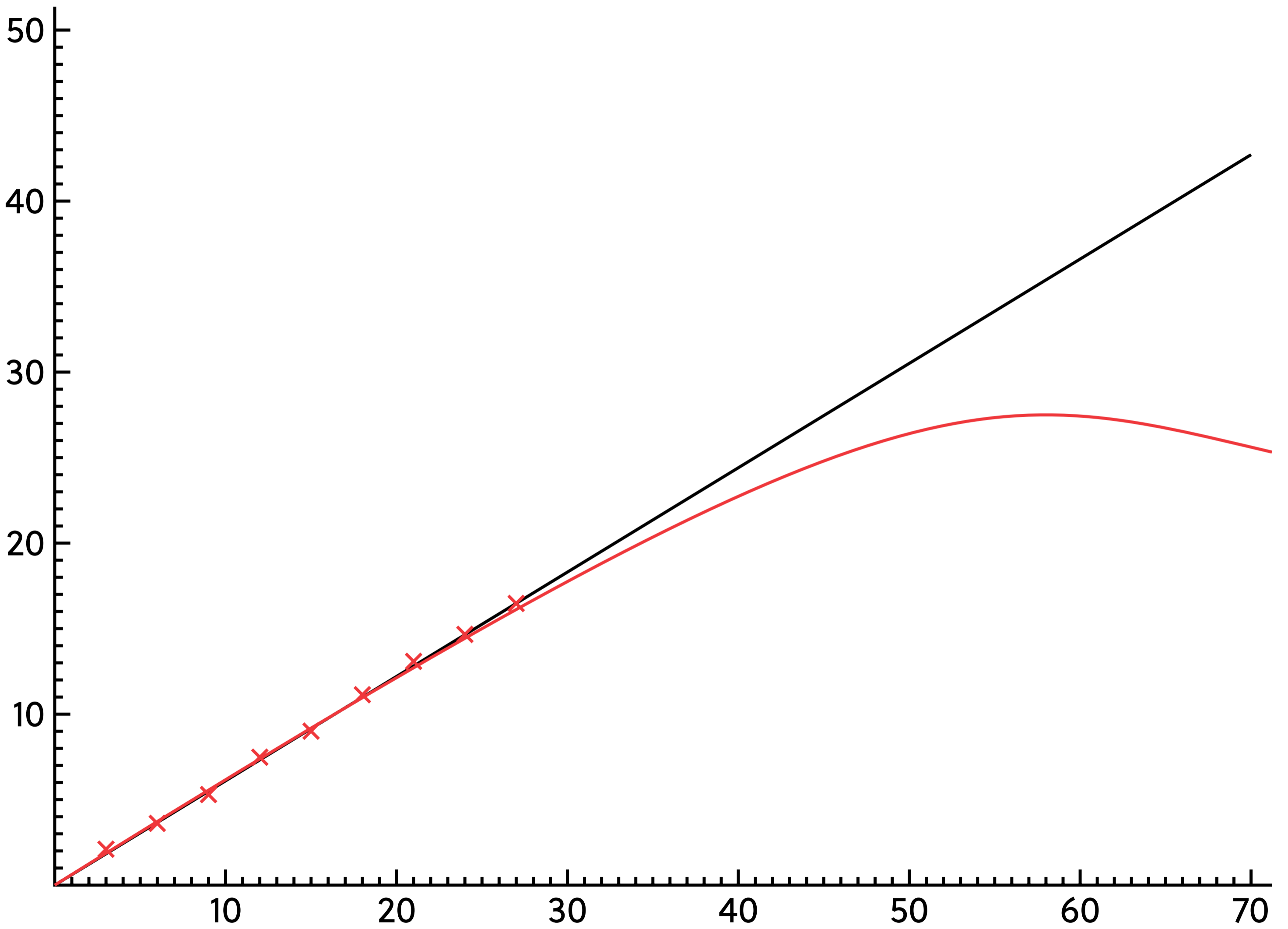
→ America's epidemic of unnecessary care

2

predikce

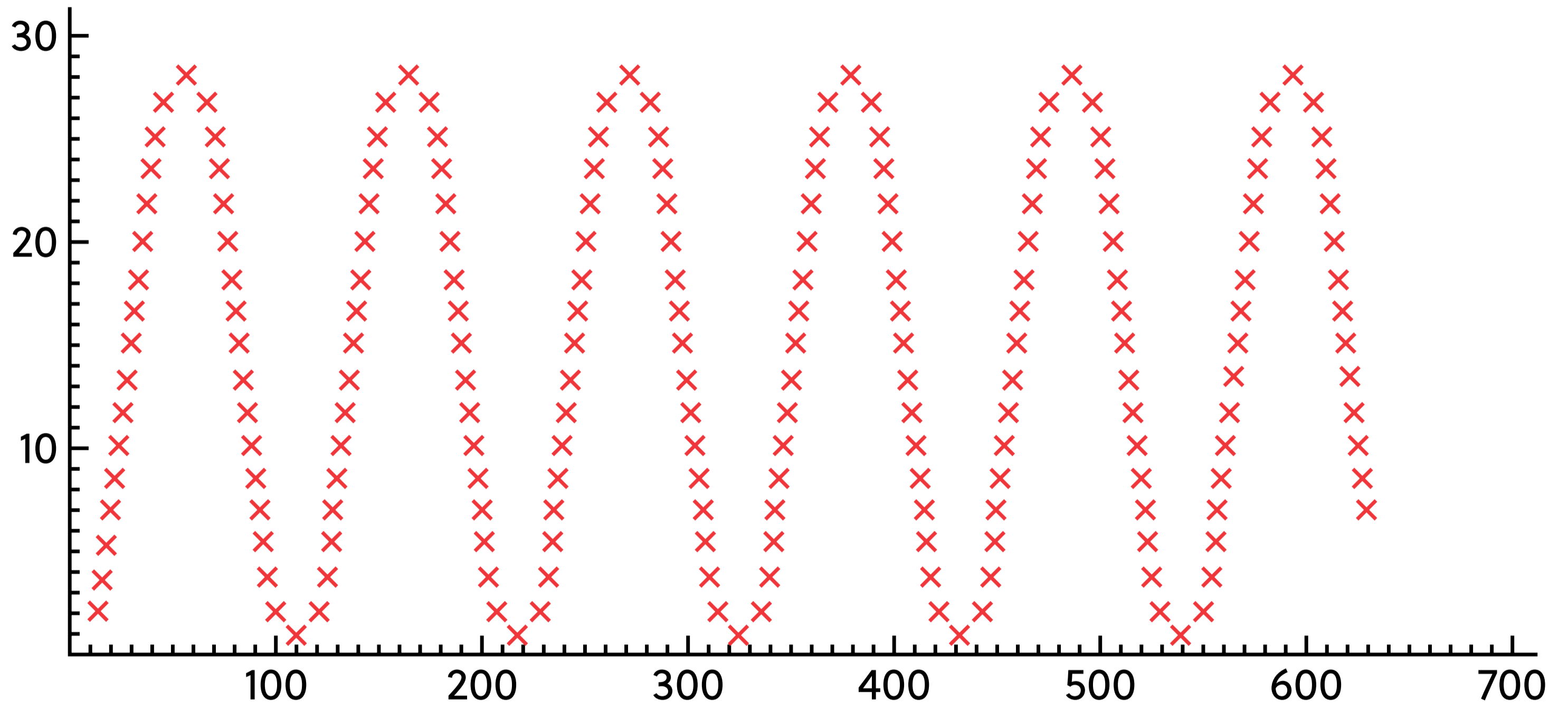






**Nesnažte se předpovídat
(extrapolovat) na základě
historických dat.**

**Techniky na předpovídání
existují, ale vyžadují úplně
jiný přístup.**

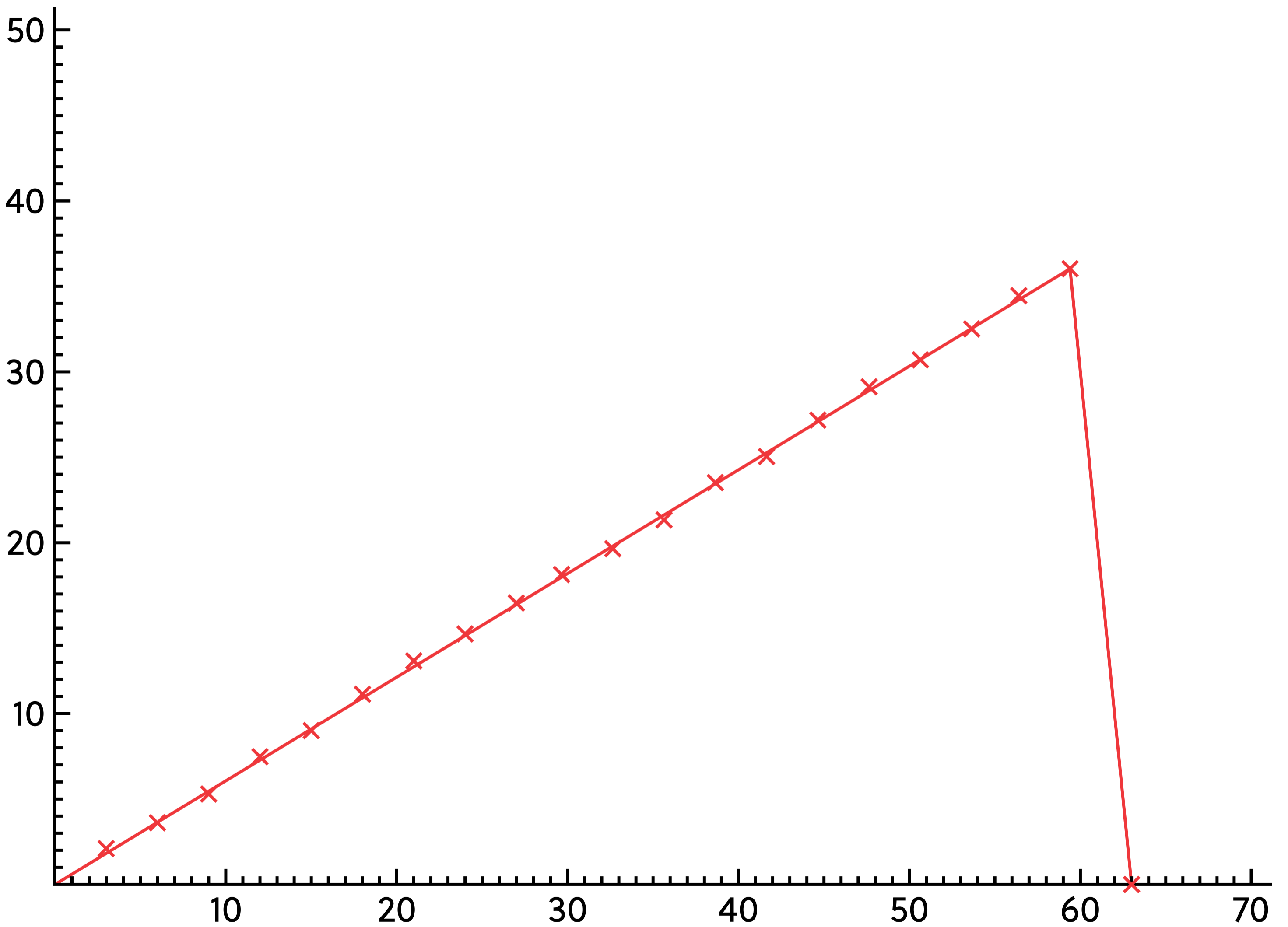


„Nejužitečnější problém“

Abychom poznali jaké rozložení data mají, musíme mít dostatek datových bodů.

Abychom poznali, kolik je dostatek datových bodů, musíme znát rozložení dat.

**konfirmační
zkreslení**



→ **Confirmation bias**

8

Pokladník charitativní společnosti si v červenci udělal tabulku, aby viděl, kolik peněz přibývá na konto každý měsíc. Řekl si: „Věřím, že to půjde tak dál,“ a dopsal do tabulky další tři čísla. Jaká?

měsíc	květen	červen	červenec	srpen	září	říjen
stav konta [v milionech Kč]	1,5	1,7	1,9			

klam narativity

Klam narativity se týká naší omezené schopnosti podívat se na sekvenci faktů a nesplétat přitom ihned nějaké vysvětlení nebo — což je víceméně totéž — je nevnímat jako logický řetězec se vztahem příčiny a následku. Vysvětlení fakta stmeluje. Lépe se však pamatuje; dávají větší smysl. Taková náchylnost nám však může uškodit ve chvíli, kdy v nás vytváří dojem, že jsme daným faktům porozuměli.

— Nassim Taleb, Černá labuť

korelace



kauzalita

Vědci zjistili, že...

US spending on science, space, and technology correlates with Suicides by hanging, strangulation and suffocation



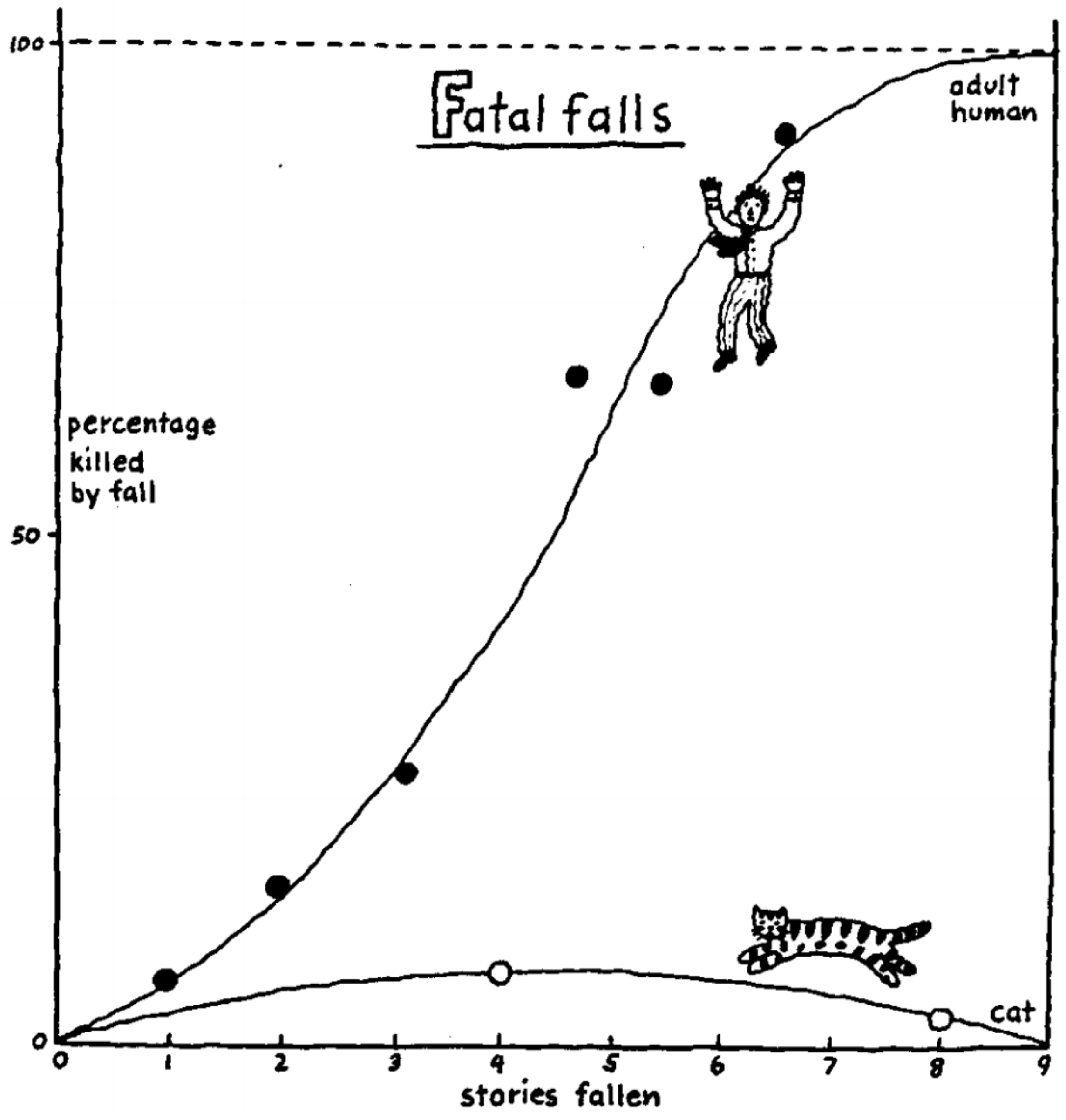
tylervigen.com

→ **Spurious Correlations**

**problém
přeživších**



In a study performed in 1987 it was reported that cats who fall from less than six storeys, and are still alive, have greater injuries than cats who fall from higher than six storeys. It has been proposed that this might happen because cats reach terminal velocity after righting themselves at about five storeys, and after this point they relax, leading to less severe injuries in cats who have fallen from six or more storeys.



Another possible explanation for this phenomenon would be survivorship bias. Cats that die in falls are less likely to be brought to a veterinarian than injured cats, and thus many of the cats killed in falls from higher buildings are not reported in studies of the subject.

— [Wikipedia: Survivorship bias](#)

Ale Diagoras, zvaný „neznaboh“, když přišel na Samothraku a jakýsi přítel se ho tázal, nepoznává-li on, který přičítá bohům nezájem o věci lidské, z tolika obrazů, kolik lidí uniklo slibem prudké bouři a šťastně dorazilo do přístavu, odpověděl: „To se tak dělá, nikde však nejsou namalováni, kteří ztroskotali a na moří zahynuli.“

— Cicero, O přirozenosti bohů

big Dada

FEB 16, 2012 @ 11:02 AM 2,850,601 VIEWS

How Target Figured Out A Teen Girl Was Pregnant Before Her Father Did



Kashmir Hill, FORBES STAFF

Welcome to The Not-So Private Parts where technology & privacy collide

[FOLLOW ON FORBES \(2081\)](#)



Opinions expressed by Forbes Contributors are their own.

FULL BIO ▾



TARGET

Target has got you in its aim

Every time you go shopping, you share intimate details about your consumption patterns with retailers. And many of those retailers are studying

→ **Big data: are we making a big mistake?**

Co si odnést

Nezapomeňte na to, že existuje svět i mimo vaše data.

Náhoda může hrát mnohem větší roli než si myslíte.

Kam dál?

Nassim Nicholas Taleb

— **Zrádná nahodilost**

— **Černá labuť**

— **Antifragilita**