



gauss  
algorithmic

**Jak využívat umělou inteligenci ke zvyšování  
produktivity práce a efektivity využívání zdrojů**

- Pomáháme organizacím využívat bohatství skryté v datech - ***“We must know, we will know”***.
- Analýza (velkých) dat, strojové učení a prediktivní analytika.
- Vývoj nástrojů pro kybernetickou bezpečnost, monitoring produktivity a efektivity organizace.
- Behaviorální profilování lidí / zařízení, detekce anomálií, rozpoznávání obrazu.
- Návrh i realizace infrastruktur pro zpracování velkého množství dat v reálném čase.



**Carl Friedrich Gauss**

1777 - 1855

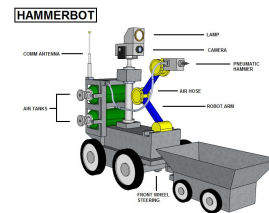
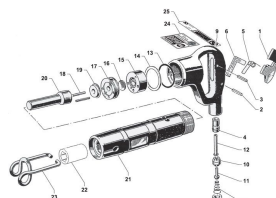
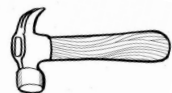
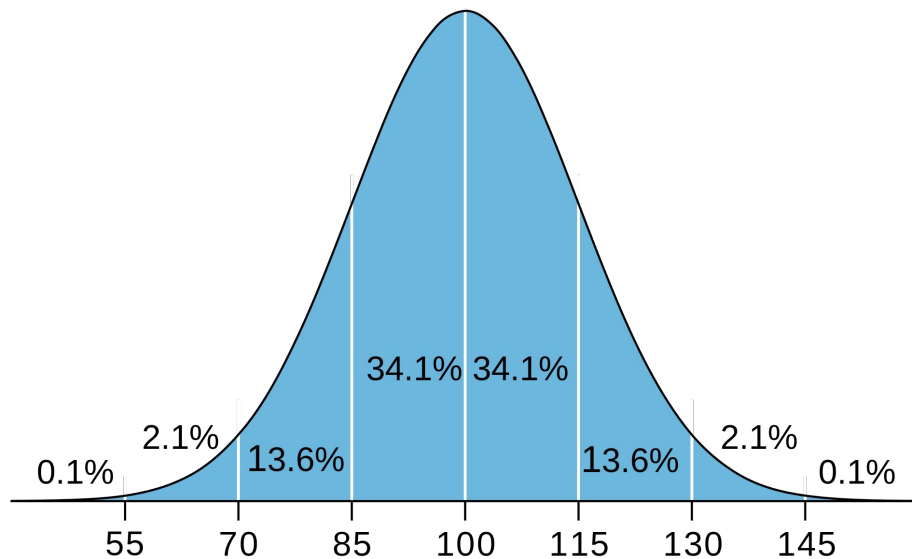
- Matematik / Statistik
- Fyzik / Astronom
- Otec, matka a strýc  
pojišťovací matematiky



**David Hilbert**

1862- 1943

- Matematik
- Optimista
- Wir müssen wissen.
- Wir werden wissen.



Složitost

Newtonovy zákony 1687	Boyle Mariottův zákon	Coulombův zákon Gaussův zákon elektrostatiky
Carnotův cyklus	Maxwellovy rovnice	Planckova konstanta
Fotoelektrický jev	Von Neumannova architektura	Tranzistor 1947

- Není problém přiznat, že inovaci nerozumíme, pokud chápeme její přínos a dokážeme ho **kontrolovat a řídit**.
- Nikdy nevíme vše. Manager je placen za rozhodování za nejistoty.



- Fyzikální podstata motivace zavádění průmyslových inovací je stále stejná.

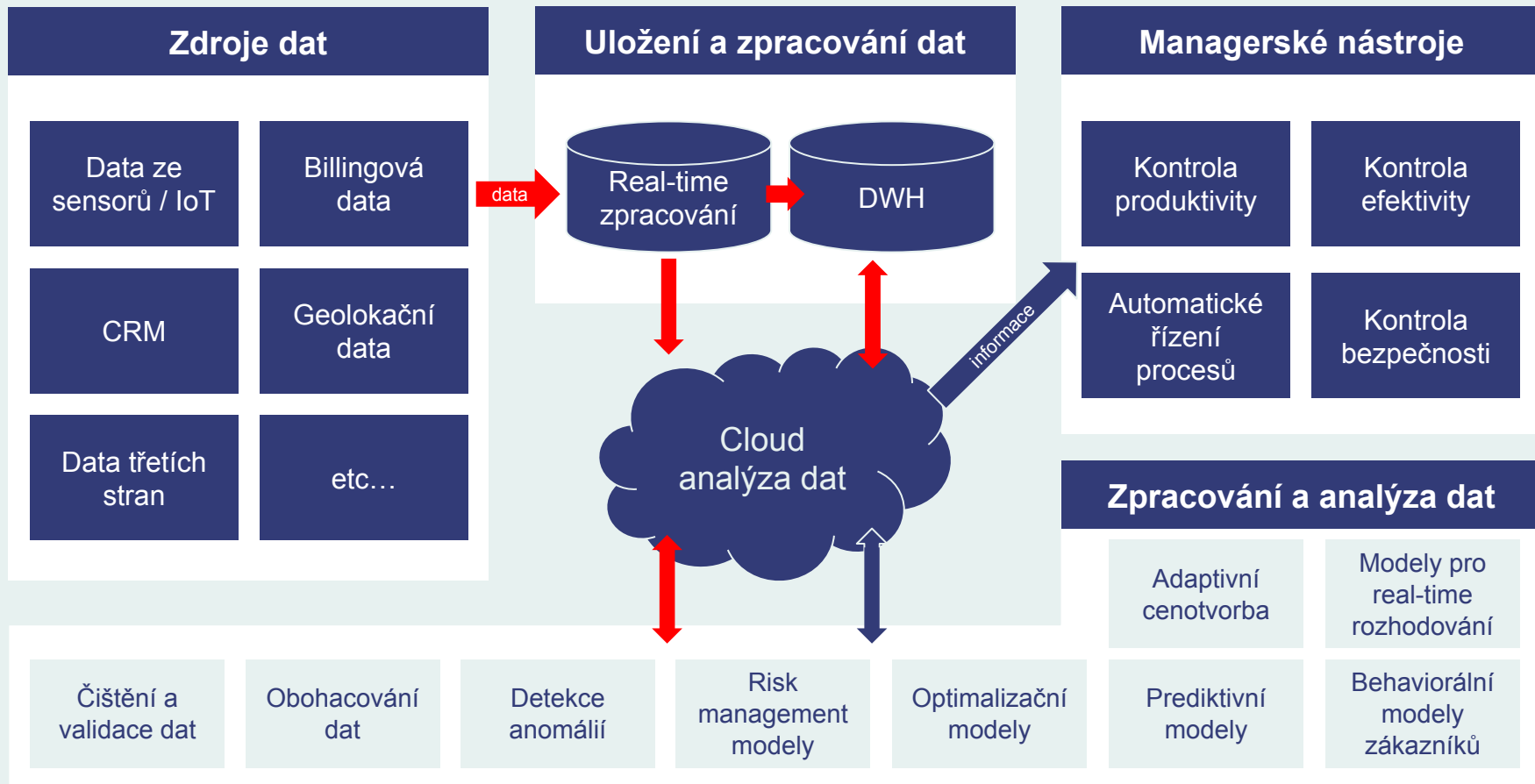
$$E_k \quad M_k \quad T_k \quad P \uparrow$$

- Při konstantní spotřebě energie a materiálu a ve stále stejném čase vyprodukovat více.

## Možnosti zvyšování produktivity a efektivity v organizaci:

1. **Větší / lepší kladivo.**
2. **Lepší informace a kontrola → sběr dat a jejich analýza.**
3. **Změna paradigmatu → umělá inteligence.**

Nezapomínejme na bod 2. v očekávání příchodu bodu 3.





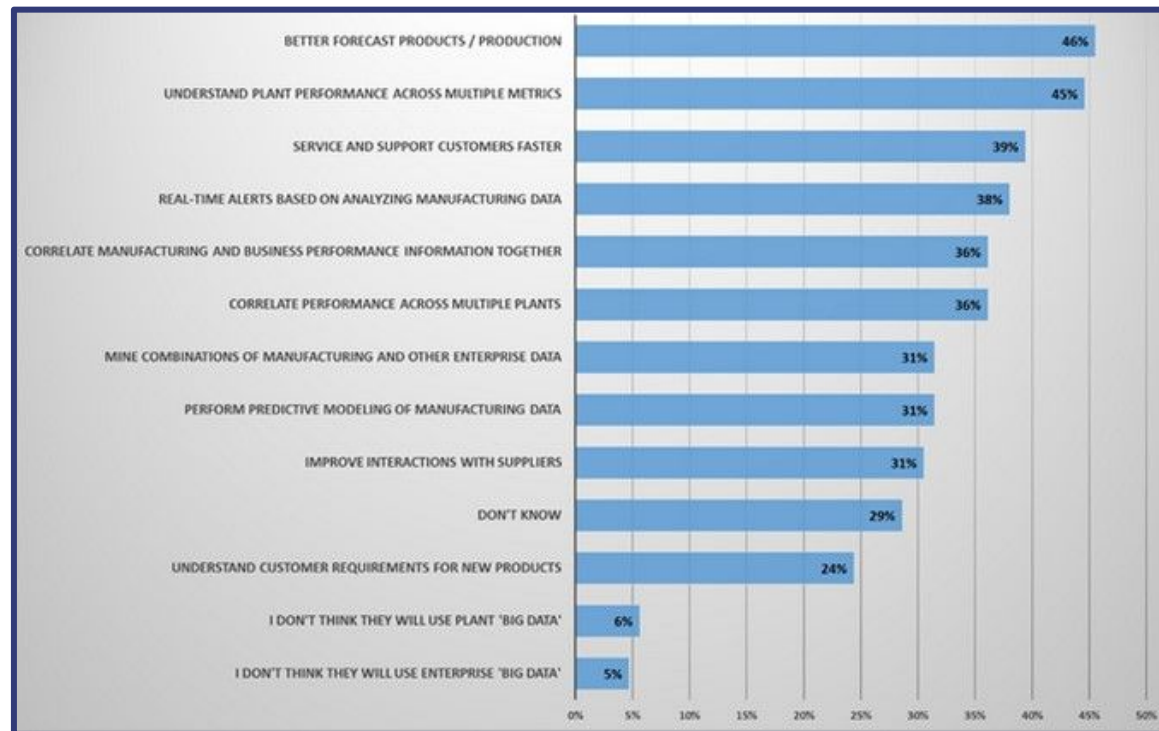
- Co právě dělá stroj/pracovník?
- Jaká je spotřeba energie?
- Jak dlouho trvá start stroje?
- Jaká je spotřeba materiálu?
- Jak dlouho trvá výrobní úkon?
- Jak se vyvíjí trh / kolik vyrábět?
- Jaká jsou rizika a jak je minimalizovat?
- Jaké jsou požadavky klientů?
- Jaké jsou nejčastější reklamace?
- Jaké jsou důvody vzniku poruch?
- Jaká je nejvyšší možná cena?
- Jak si vede konkurence?
- etc...

otázky

odpovědi

# DATA

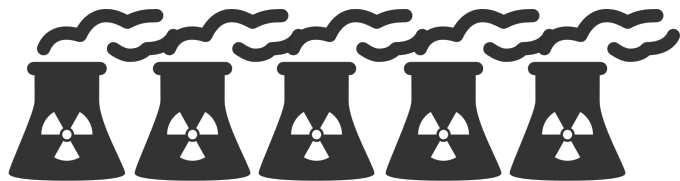
**Vše měříme, ukládáme data a myslíme si, že tak známe otázky na všechny odpovědi.**





## 12 000 0000

instalovaných fyzických serverů jen v USA



## 77 miliard kWh

ročně spotřebují produkci 5 dukovanských elektráren

**30 % serverů v průběhu posledních 6 měsíců vůbec nic nespočítalo.**

**Pouze 10 % - 15 % servery spotřebované elektřiny je využíváno na výpočty.**

## Co to znamená?

- Organizace kupují zbytečnou IT infrastrukturu a hromadí na ní data.
- Neumí se ptát a nacházet v datech informace.
- Nepoužívaná data jsou pouze náklad (HW + energie).
- Management poslouchá DWH analytiky zaseknuté v 90. letech.
- Management nemůže dobře řídit digitální organizaci, pokud nemá informace pro kontrolu procesů, lidí a jejich činnosti v reálném čase.

**Co nekontrolujete, to neřídíte.**

## Zdroje dat

Strojově  
generovaná  
data

Data ze  
sensorů

Geolokační  
data, etc..

Lidmi  
generovaná  
data

Návštěvnost  
webu

CRM, etc...

Teplota 1

Teplota 2

Tlak + místo

Prohlédnutá  
stránka A

Prohlédnutá  
stránka B

Útrata za  
měsíc

Nové faktory, které mění  
přístup, technologie i  
software využívány k  
extrakci informací z dat.

### Čas

Záleží na časových intervalech mezi jednotlivými událostmi

DTW a SQL  
nevyhovují

### Pořadí

Záleží na pořadí, ve kterém k události nastaly

Tradiční  
statistické  
metody nestačí

### Vzory

Důležité jsou typické vzory chování

Klasická datová  
infrastruktura  
nevyhovuje

Místo DTW Data Lake,  
místo SQL, prediktivní  
analýtika (algoritmy)

- Upravit datovou infrastrukturu tak, aby byla schopná real-time zpracování velkého množství dat (1000 dotazů za sekundu).
- Adoptovat technologie Apache Hadoop, Hive, **Spark**, Kafka, atd..
- Implementovat technologie a nastavit procesy pro rychlou reakci na základě z dat zjištěných informací.
- Jsou potřeba analytici, kteří ovládají jazyk Python / Scala / Java.
- Naučit se pracovat jako vědec - nebát se experimentů a selhání.



Sběrem dat o provozu, klientech, dodavatelích a ostatních stakeholderech dostáváme hroznou, nechutnou, a děsivou hromadu dat, kterou nelze zpracovat běžnými metodami (popisná statistika, SQL, atd.).

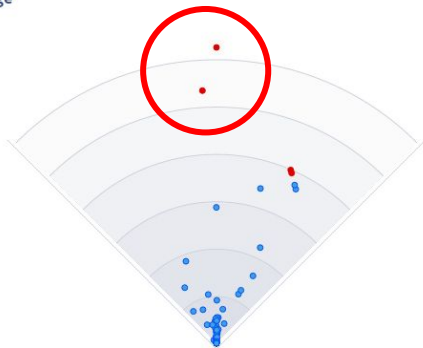
**Jak zpracovat získaná data v reálném čase, automatizovat řízení organizace a být včas upozorněn na sníženou efektivitu či nízkou produktivitu?**



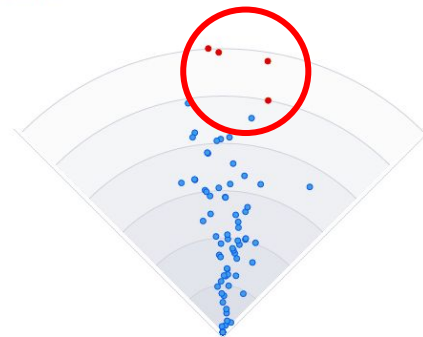
**Využít možnosti strojového zpracování dat, strojového učení, algoritmů pro detekci anomálií, prediktivní analytiky.**

Anomalies, from **26 Jan, 2015** to **01 Feb, 2015**, All

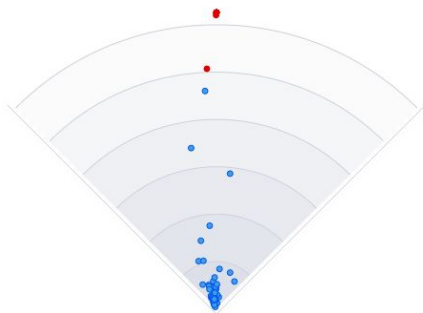
Web usage



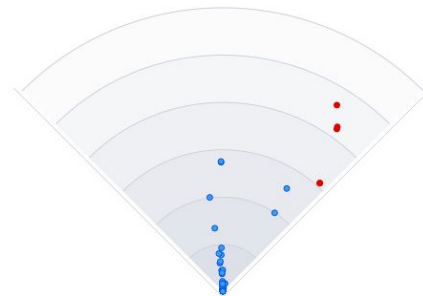
Software usage



Computer usage



Printer usage





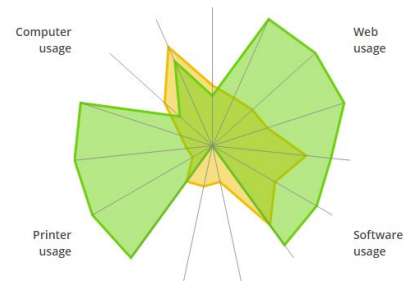
Karel Sedláček Th.D., marketing department, from 26 Jan, 2015 to 01 Feb, 2015

Overview Description Alerts

## Highlights

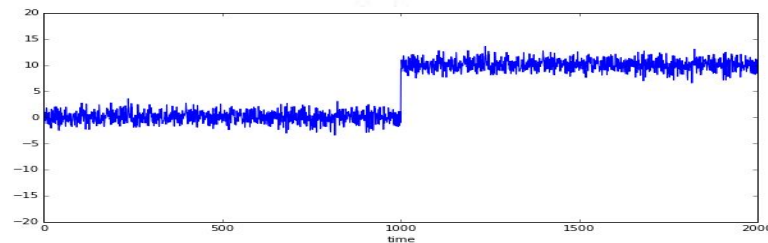
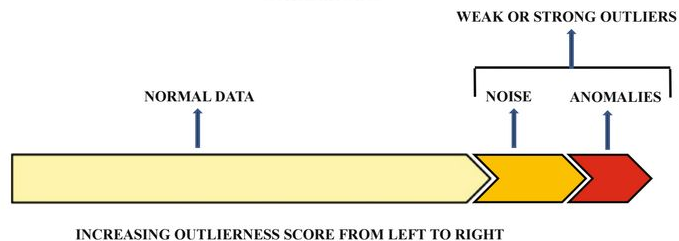
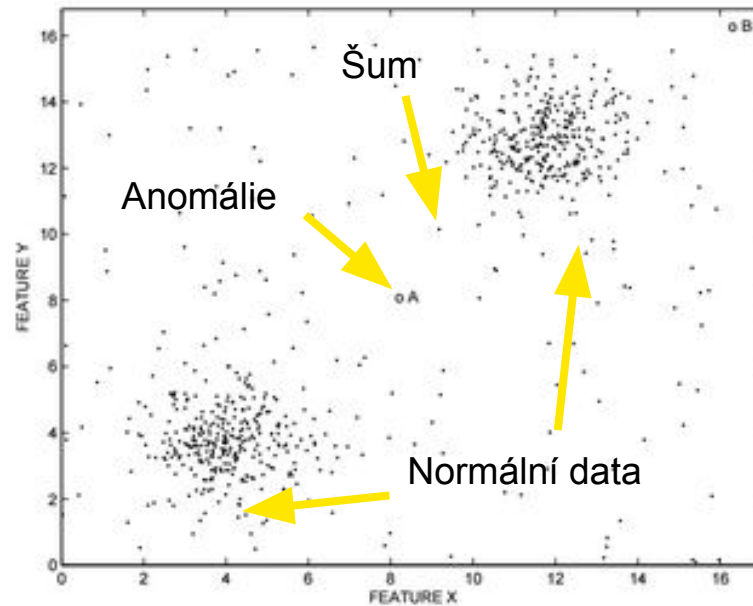
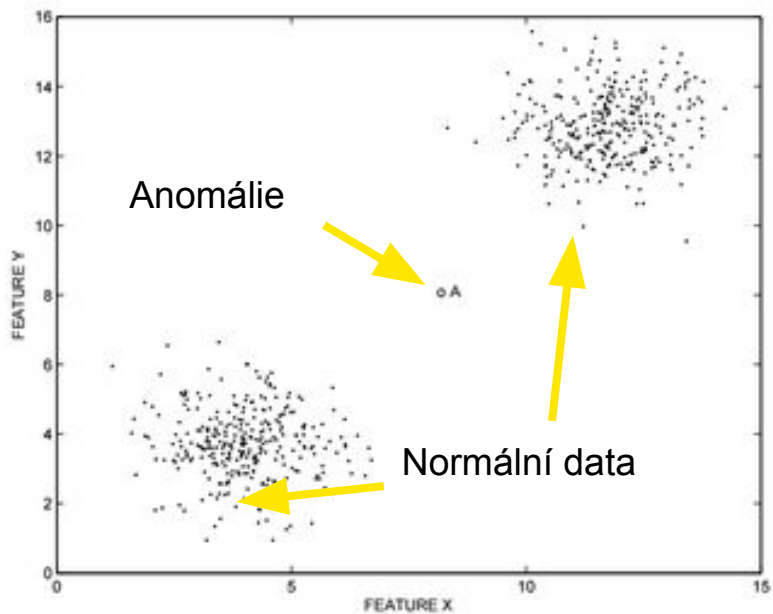
<b>56%</b> Avg productivity	<b>11.00 - 12.00</b> Peak productivity
<b>5h 07m</b> Daily avg software time	<b>2h 16m</b> Daily avg web time
<b>5h 08m</b> Avg active computer usage time	<b>LVV-CVM2KYRLVL</b> Primary computer
<b>859 / 0</b> Pages printed	<b>\\LVV-CVM2KYRLVL\HP LaserJet 1200 Series PCL 6</b> Primary printer

## Behaviour anomaly detection



## Web usage

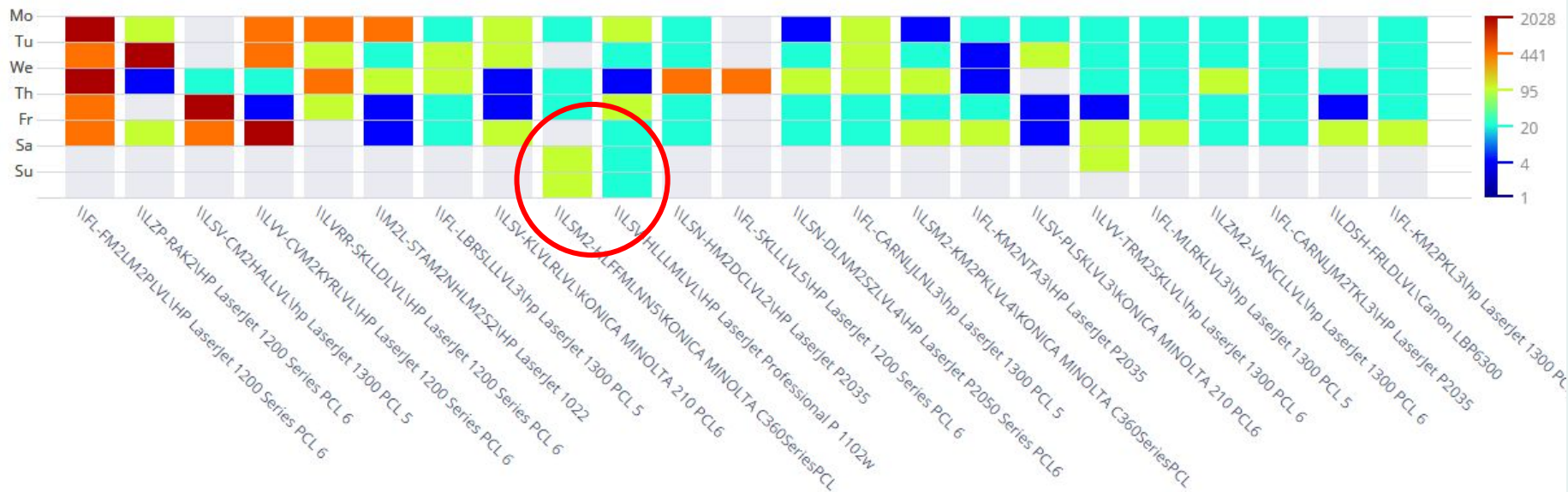




# Jednoduchý příklad z kancelářské praxe - vizualizace

Printers usage, from **12 Jan, 2015** to **18 Jan, 2015**, All

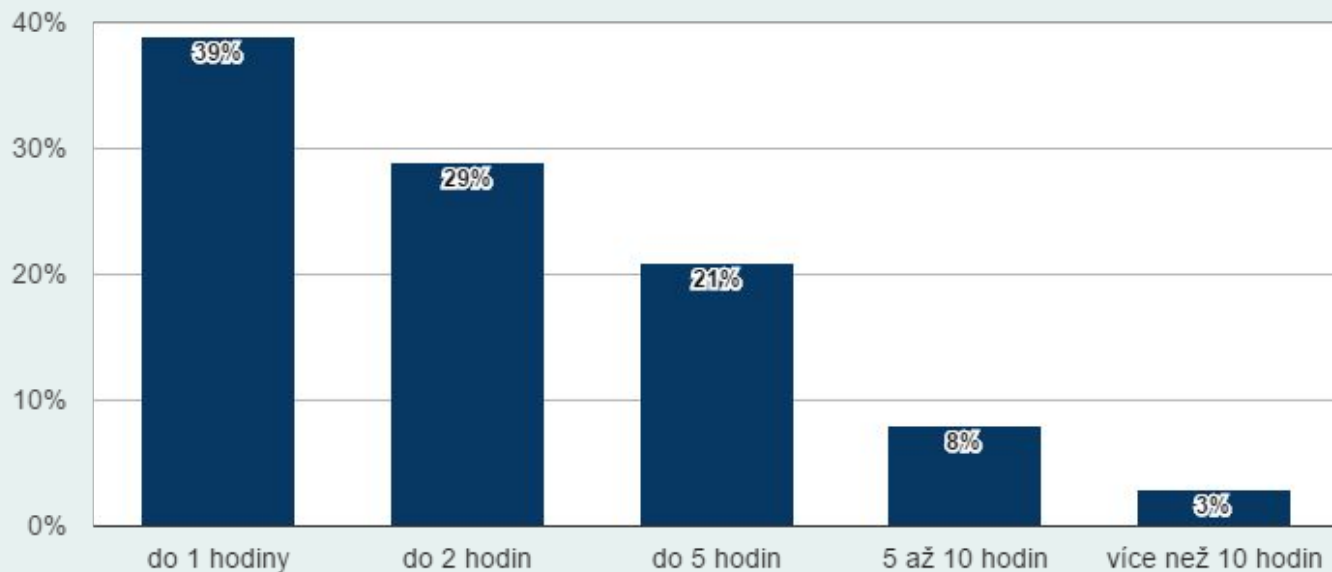
Usage by day



- Spotřeba elektřiny na zbytečně běžících počítačích a strojích.
- Plýtvání materiálem.
- Zneužívání firemních zařízení k osobním účelům.
- Neefektivní nákupy zdánlivě potřebných, ale nevyužívaných zařízení.
- Nutno počítat s možností krádeže dat (digitální firma - digitální zločin).
- Pracovní návyky vedoucí k neefektivnímu využívání času.

**Zaměstnanci v USA stráví kontrolováním e-mailu v průměru 6.3 hodiny denně.**

**64 %** zaměstnanců navštěvuje denně weby  
nesouvisející s náplní jejich práce.



čas strávený zaměstnanci týdně na nepracovních webech

Firmu se 100 zaměstnanci to ročně pouze na mzdě přijde na\*:

**2 000 000Kč**

ČR zaměstnává cca 68 072 úředníků, takže ji to ročně stojí\*:

**1 300 000 000Kč**

\*při průměrné měsíční mzdě 25 000Kč

## Bezpečnost

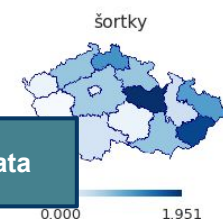
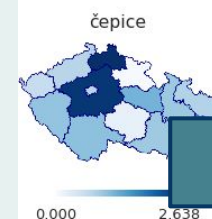
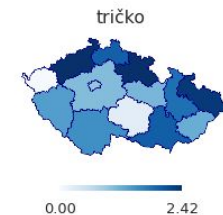
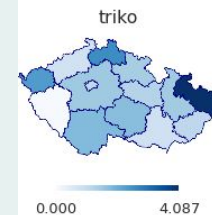
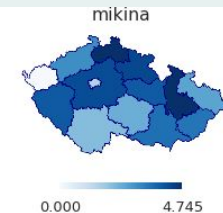
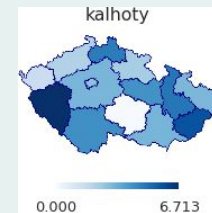
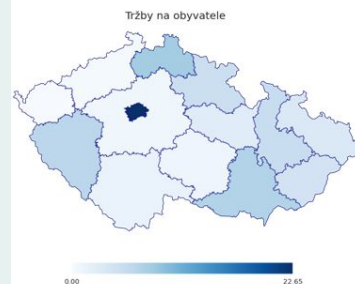
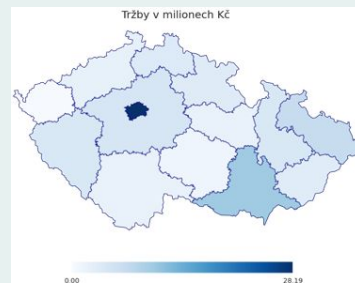
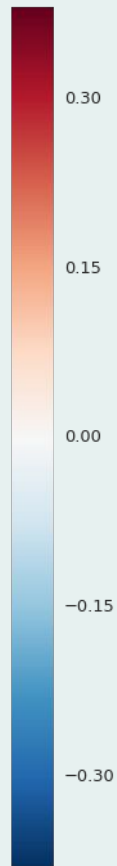
- Provádí někdo aktivity, které mohou poškodit organizaci?
- Mění někdo významně své standardní chování?
- Dochází k úniku informací? atd..

## Produktivita práce

- Dělají pracovníci to co mají?
- Využívají prostředky organizace nebo je zneužívají?
- Kteří pracovníci a oddělení jsou efektivnější... atd.

# Gauss Algorithmic - Další příklady z praxe

	Maximální teplota	Minimální teplota	Srážky	Sněhová pokrývka
set	-0.016	-0.0071	0.091	-0.001
klíčenka	-0.23	-0.23	-0.076	0.014
triko	0.024	0.042	0.11	-0.07
šaty	0.15	0.16	0.22	-0.096
3/4 kalhoty	0.16	0.15	-0.05	-0.04
obuv	0.11	0.13	0.21	-0.16
kraťasy	0.17	0.18	0.14	-0.11
ponožky	-0.14	-0.13	0.084	0.028
šortky	0.2	0.19	0.038	0.12
bunda	-0.17	-0.16	0.056	0.087
tričko kr. rukáv	-0.06	-0.06	0.11	0.035
sukně	0.24	0.25	0.25	-0.039
příbor	0.049	0.016	0.053	-0.058
taška	-0.22	-0.2	0.037	-0.092
kalhoty	-0.046	-0.032	0.12	0.021
lezečky	0.016	-0.0051	0.11	-0.14
karabina	0.36	0.25	-0.023	-0.093
batoh	-0.21	-0.19	0.0086	-0.055
čelovka	-0.14	-0.13	0.05	-0.12
prací prostředek	-0.055	-0.037	0.16	0.027
funkční tričko	-0.061	-0.058	0.095	0.047
spodní prádlo	-0.13	-0.12	0.06	0.2
sandály	0.079	0.091	0.13	-0.1
nůž	-0.011	-0.0046	0.0093	-0.011
peněženka	-0.21	-0.2	-0.15	0.22
stan	0.22	0.19	-0.19	-0.18
košile	0.23	0.23	0.25	-0.13
rukavice	-0.12	-0.092	0.18	0.028
impregnace	-0.073	-0.073	0.022	-0.082
ručník	-0.12	-0.12	-0.077	-0.081
tričko dl. rukáv	-0.38	-0.38	0.0073	0.38
boty	0.073	0.078	0.12	-0.066
karimatka	0.067	0.077	0.091	-0.064
čepice	-0.12	-0.11	0.1	-0.0026
pouzdro	0.011	0.0094	0.34	-0.062
spací pytel	0.18			-0.11
vesta	-0.11			-0.01
mikina	-0.054			-0.017
tričko	-0.07			-0.009



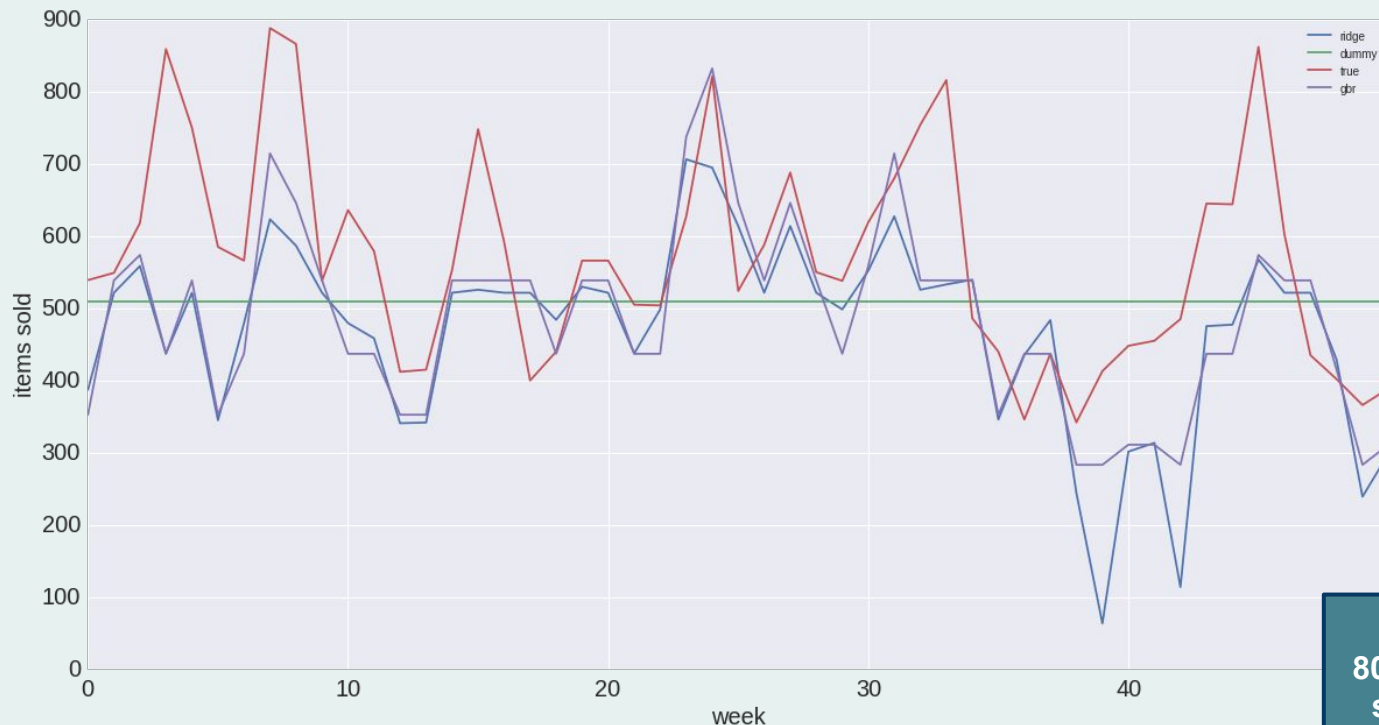
**Ekonomická data**

**Geodata**

**Počasí**



- Prediktivní model poptávky ve zbožových skupinách při známé cenové hladině.



80 % pravděpodobnost  
správné předpovědi.



gauss  
algorithmic

Děkuji za pozornost.

Čas na Vaše dotazy.