

VIZUÁLNA REPREZENTÁCIA INFORMÁCIÍ ČASŤ 2.

KRÁTKA REKAPITULÁCIA

OBJAVIŤ NEČAKANÉ, OVERIŤ OČAKÁVANÉ
KOMUNIKOVAŤ ZNÁME

VIZUÁLNY JAZYK PRE NUMERICKÉ INFO
Farby, dĺžky, pozícia, plocha

Mozog:
POROVNÁVA VEĽKOSTI, DETEGUJE VZORY

Vo vzoroch hľadáme:
VZORY (ZHLUKY, LÍNIE, TRENDY)
ANOMÁLIE (OUTLIERS)

ZÁKLADNÉ NÁSTROJE VIZUÁLNEJ ANALÝZY

SCATTERPLOT

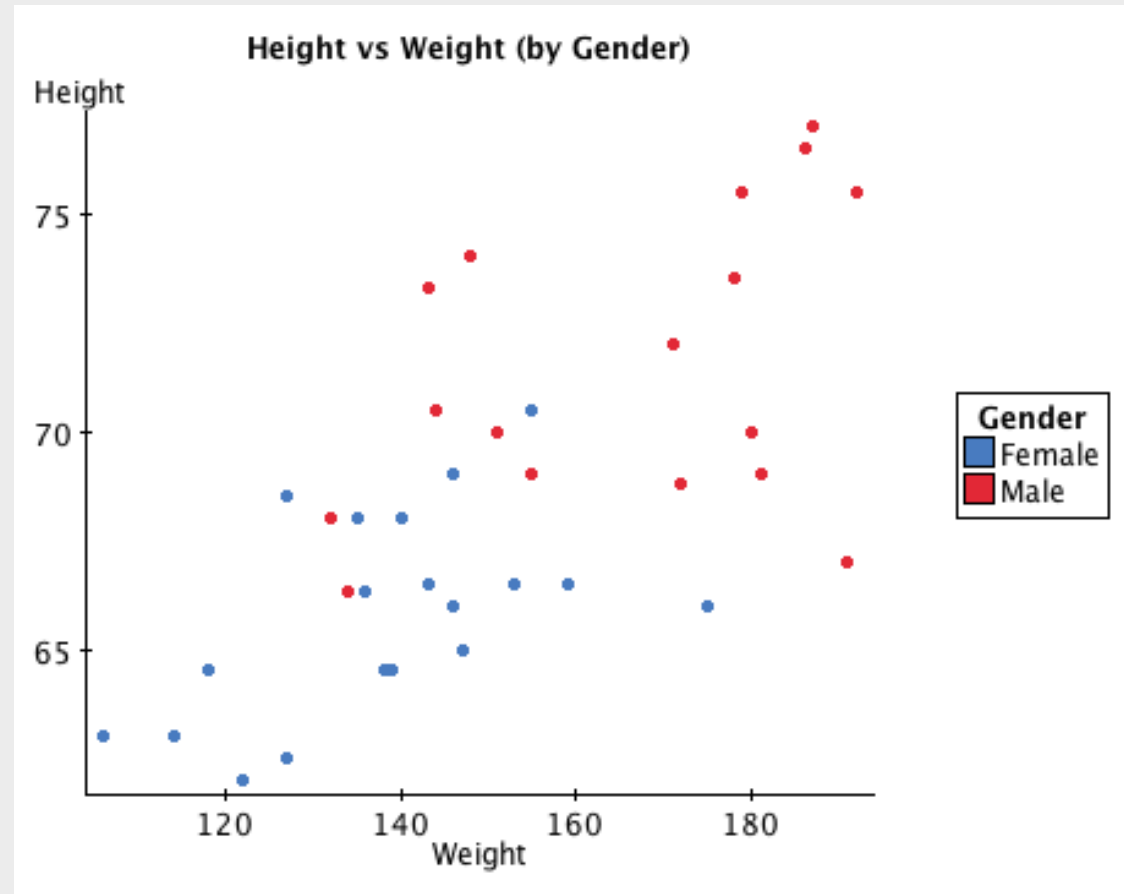
DVA ROZMERY, DVE VELIČINY

Hľadáme:
KORELÁCIE
ZHLUKY
ANOMÁLIE

ZÁVISLOSTI

Lineárne
Exponenciálne, ...

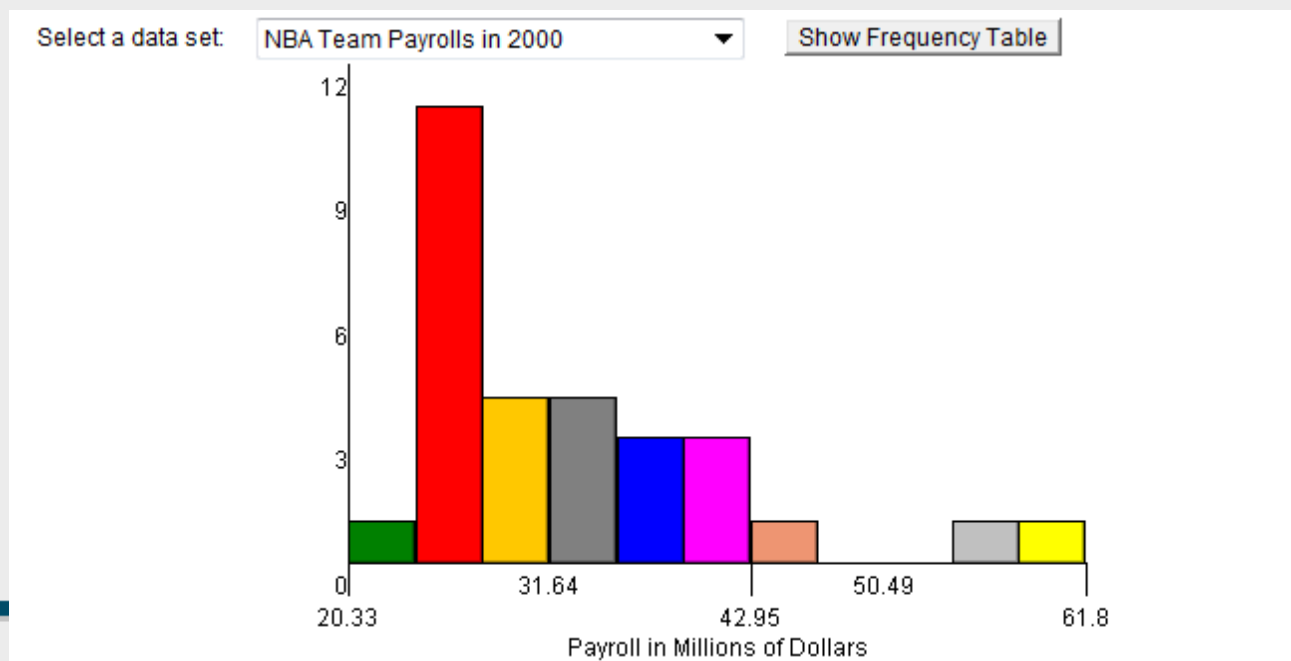
ĎALŠÍ ROZMER CEZ FARBU, TVAR



HISTOGRAM

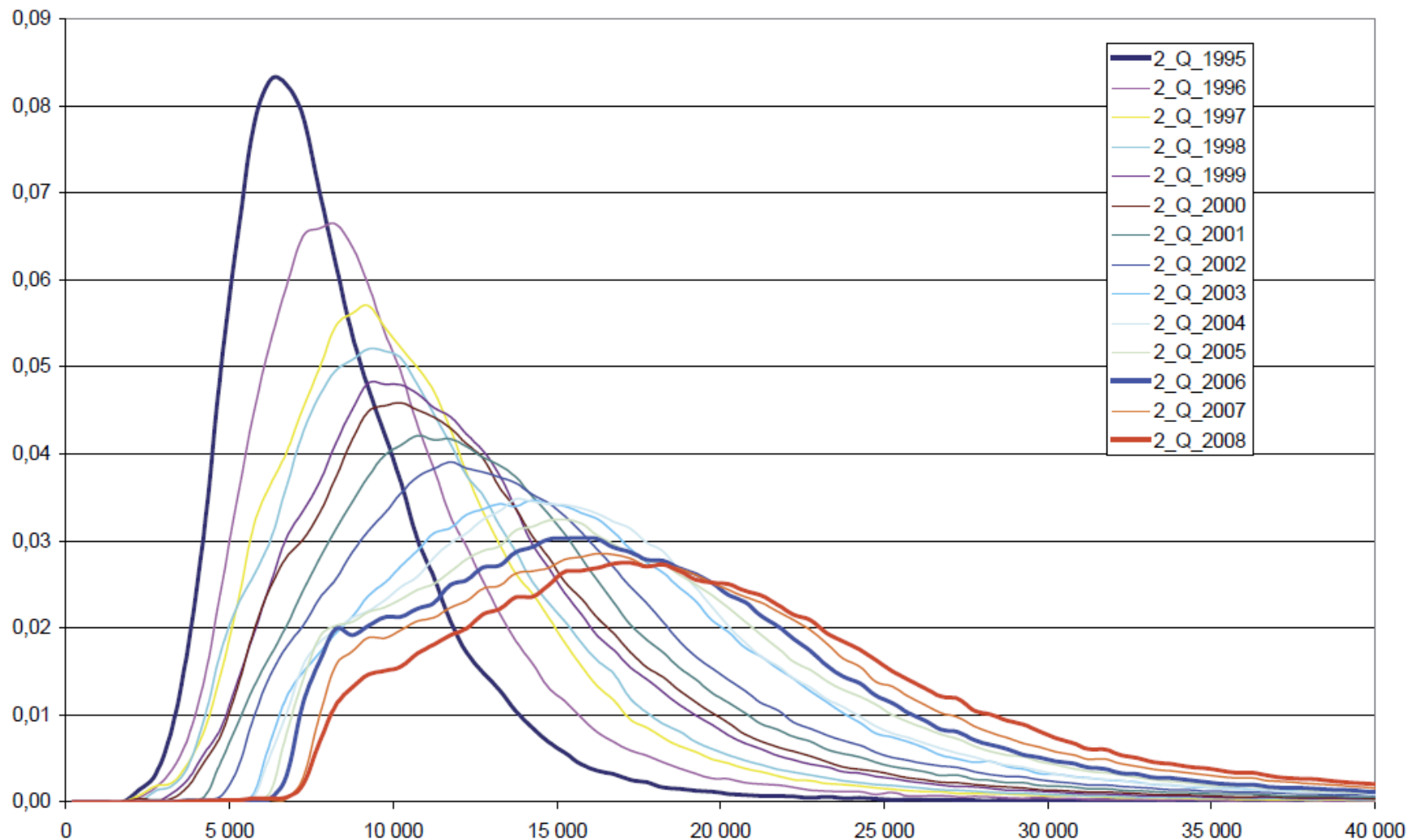
DISTRIBÚCIA HODNÔT NA JEDNEJ VELIČINE

1. Rozdeliť rozpätie hodnôt na intervaly
2. Pre každý interval spočítať, koľko prvkov má svoju hodnotu v tomto intervale („Početnosť“)
3. Absolútne (*počet*) / Relatívne (%) hodnoty



HISTOGRAM

Rozdělení četností mezd v čase – do 40 000 Kč



BOXPLOT

ZOBRAZENIE ŠTATISTICKÝCH VELIČÍN

Maximálna hodnota

Horný kvartil

Medián

Spodný kvartil

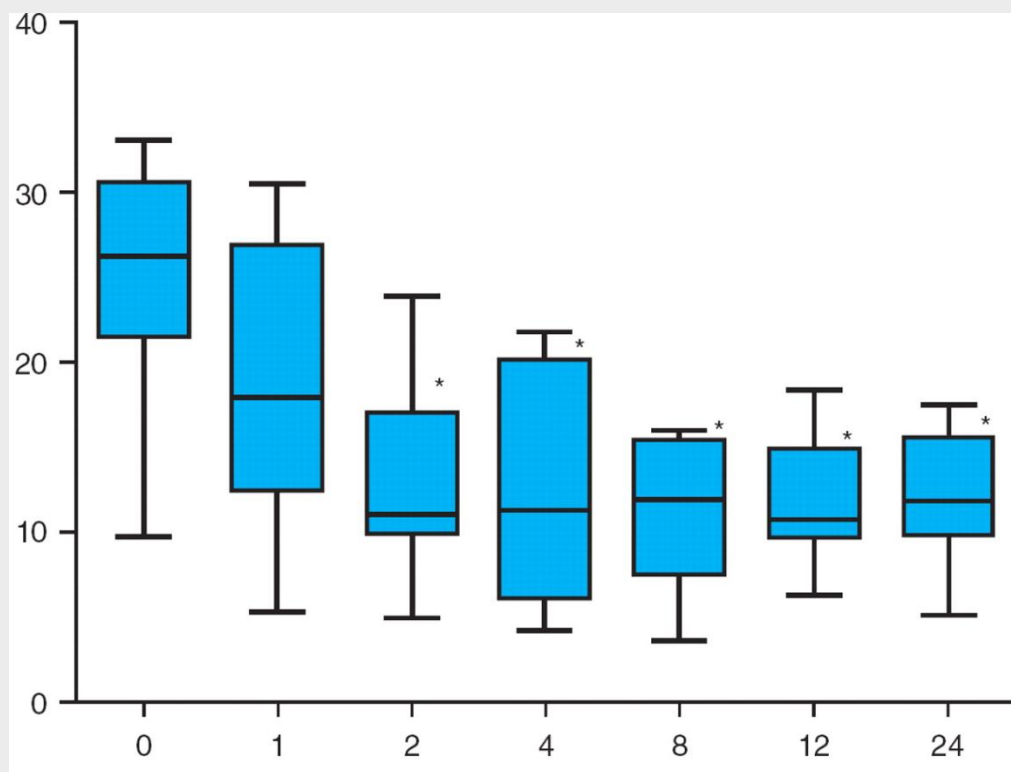
Minimálna hodnota

Aby sme videli:

ROZDELENIE

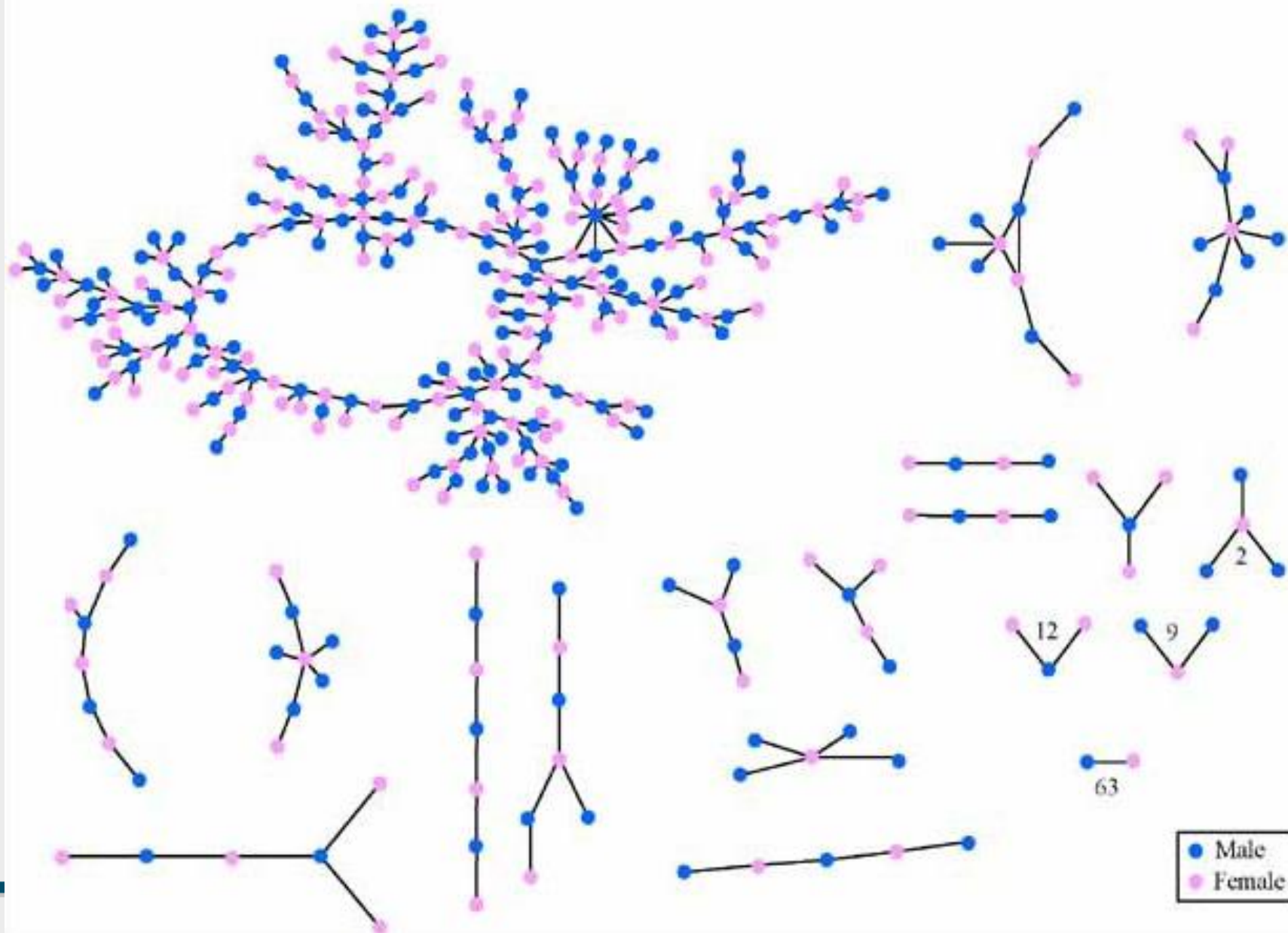
HODNÔT,

SPOĽAHLIVOSŤ ODHADU



REPREZENTÁCIA VZŤAHOV

REPREZENTÁCIA VZŤAHOV



GULIČKY A ČIARKY (NODE-LINK DIAGRAM)

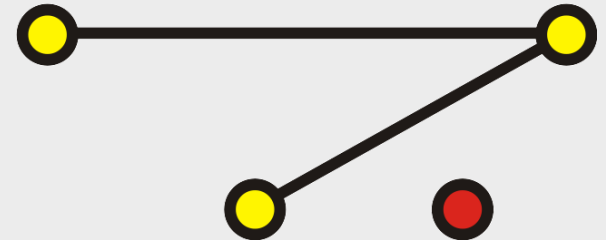
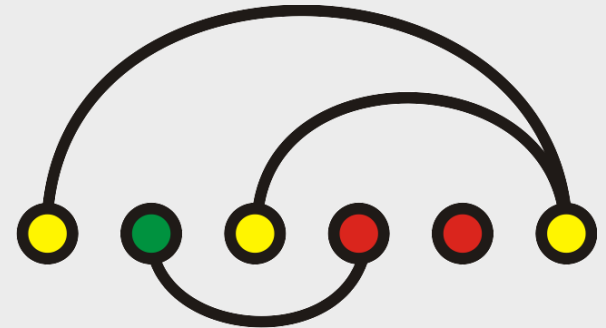
ENTITY → VRCHOLY
VZŤAHY → HRANY

VEĽA ZÁLEŽÍ
NA POZÍCII VRCHOLOV

Zoskupenia, štruktúry, úrovne...
Zoradenie

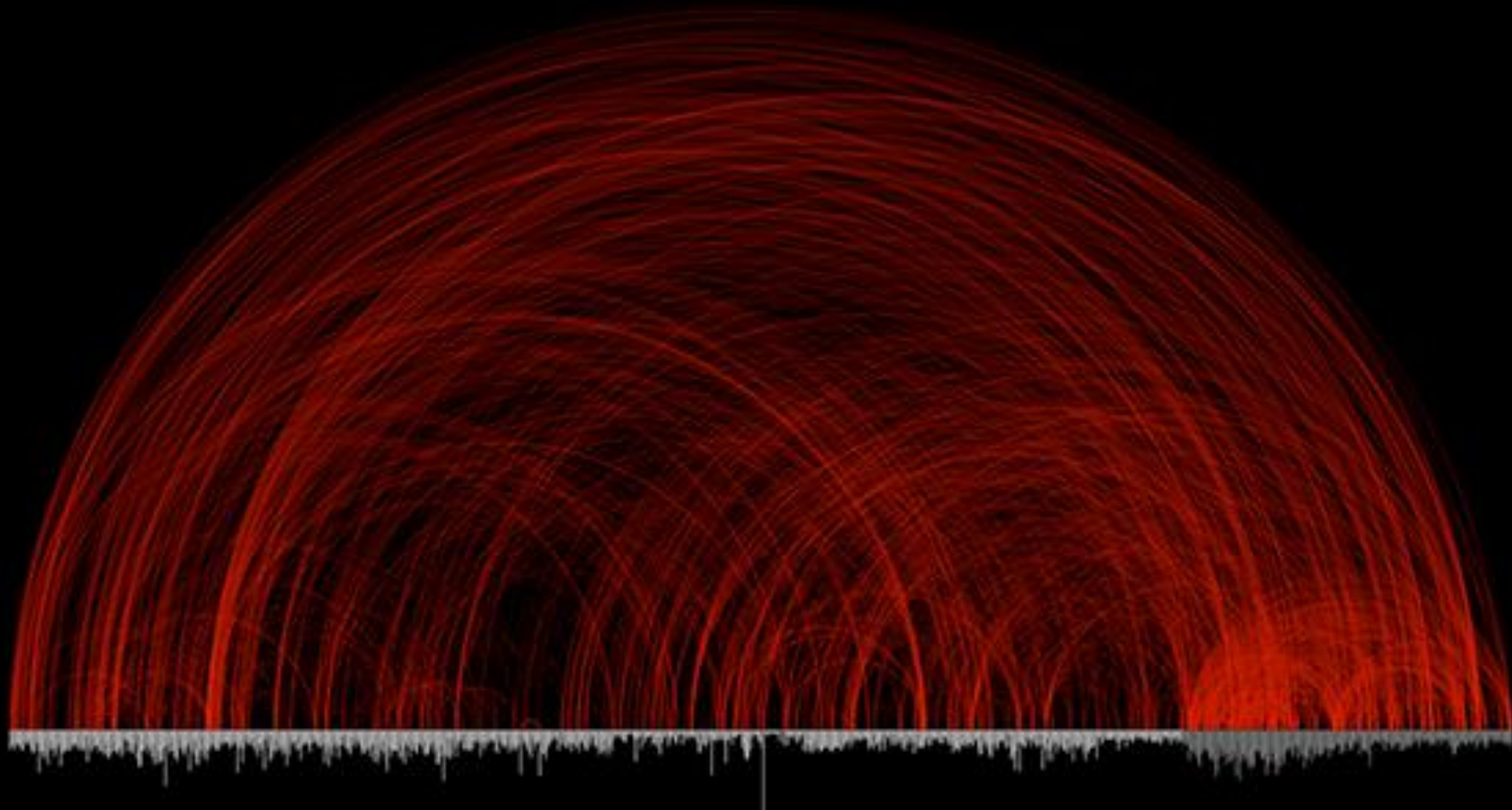
TVAR HRÁN

Napr. orientácia vzťahu



OBLÚKOVÝ (ARC) DIAGRAM

Contradictions in the Bible



ARC DIAGRAMY

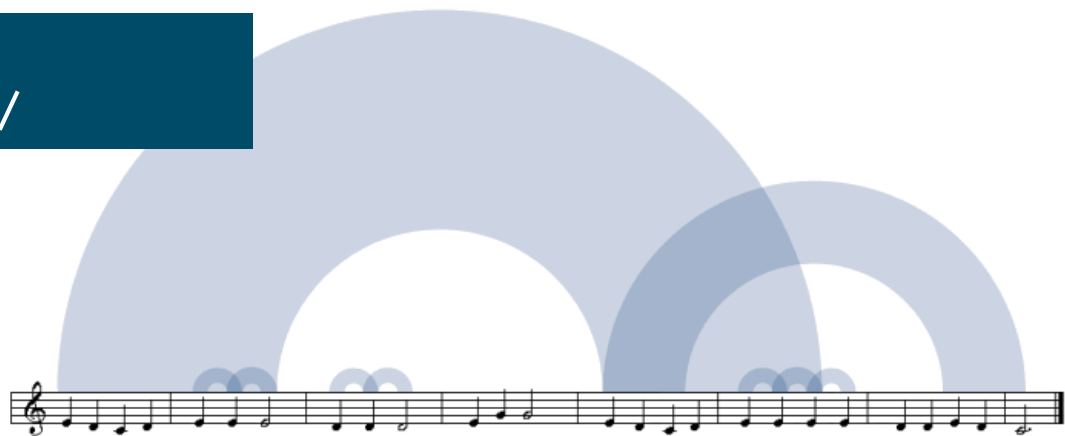
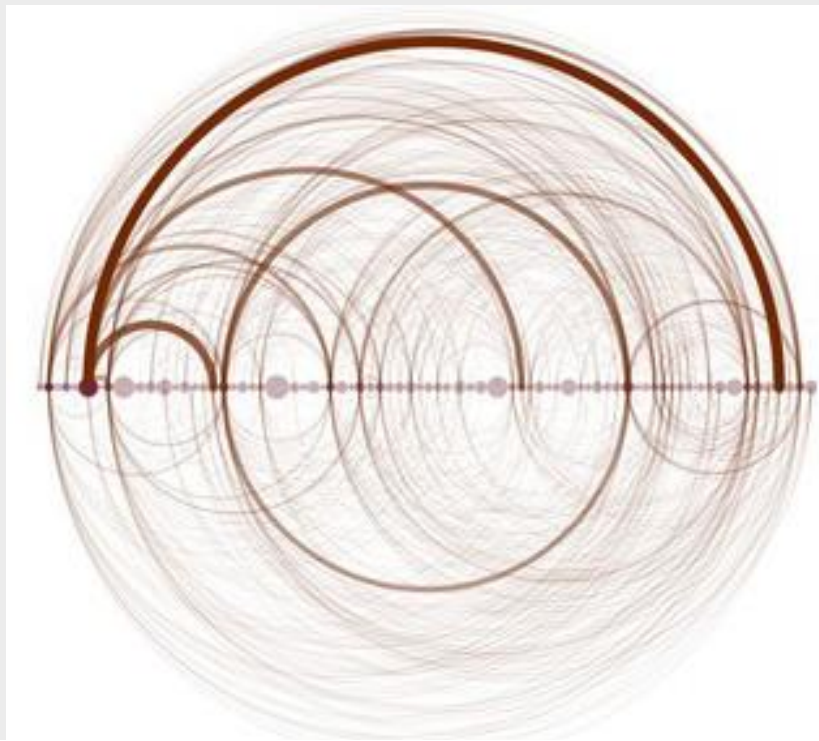
LINEÁRNE ZORADENIE

ORIENTÁCIA OBLÚKOV

Horná / Dolná

INTERVALY

Martin Wattenberg: The shape of song
<http://www.turbulence.org/Works/song/>



ČO SA DÁ NÁJSŤ V SIETI

ZHLUKY UZLOV

Na základe pozícií
(takmer) Kompletné podgrafy

VÝZNAMNÉ MIESTA V SIETI

Artikulácie, mosty, hviezdice

REŤAZE

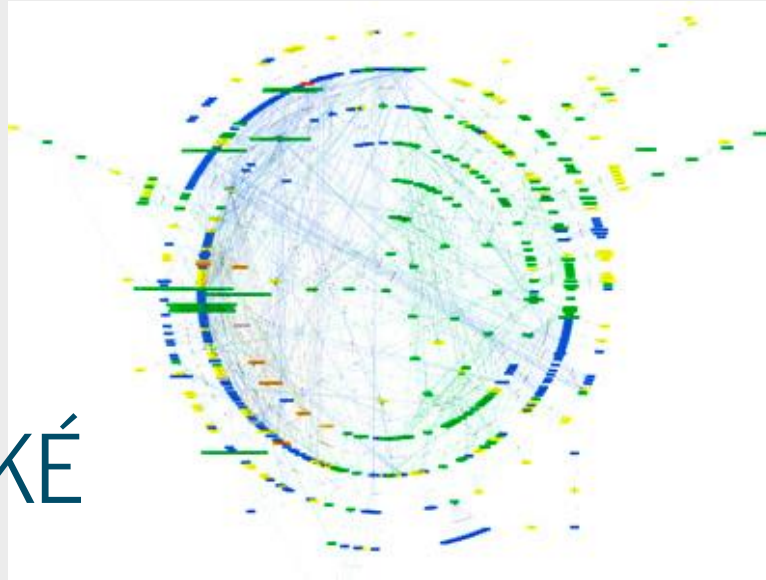
Podgraf plný artikulácií



RÔZNE DRUHY POZICIOVANIA

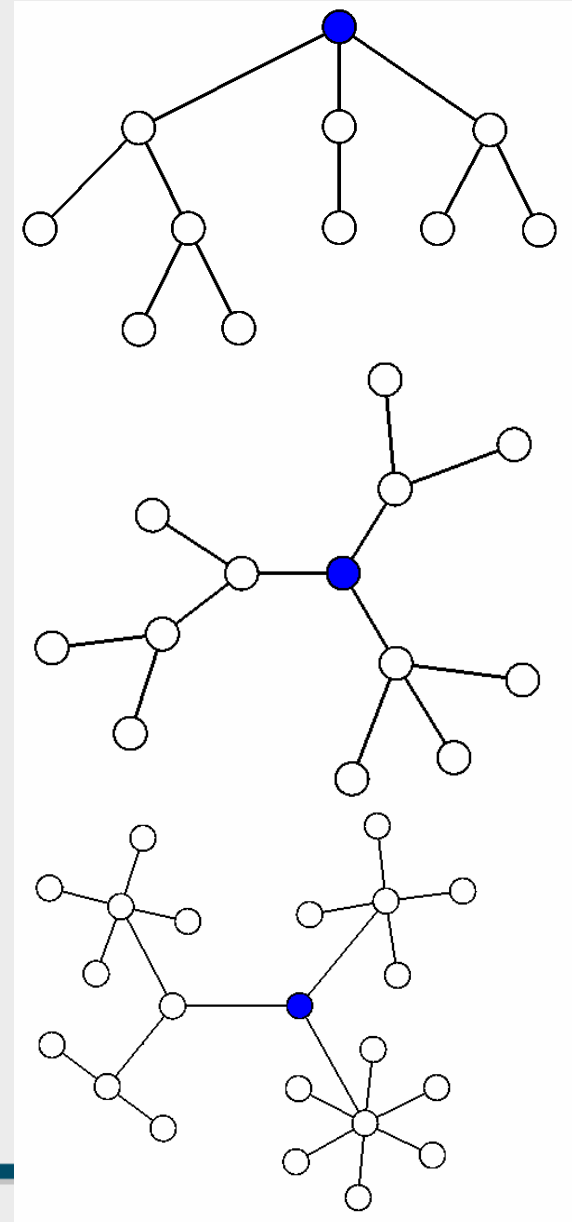
NÁHODNÉ
ČASOVÉ
LINEÁRNE
KRUHOVÉ
GEOGRAFICKÉ

...

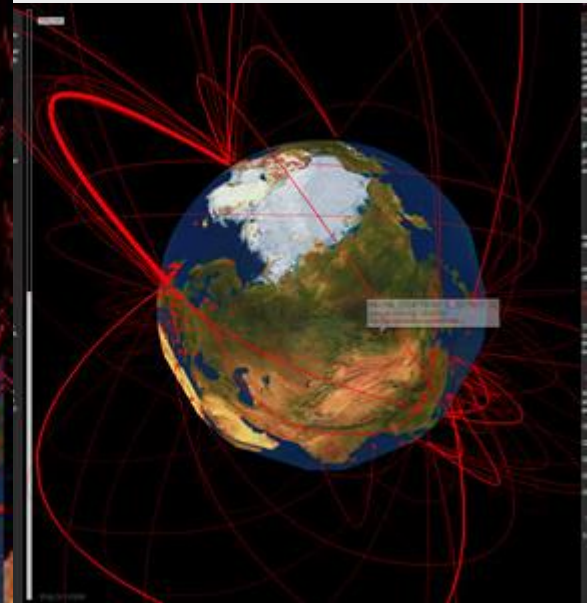


STROMY:

Hierarchické
Radiálne
Balónové...



GEOGRAFICKÉ ROZLOŽENIE UZLOV



<http://bestiario.org/research/citydistances/>

SILOVÉ ROZLOŽENIE (FORCE-DIRECTED)

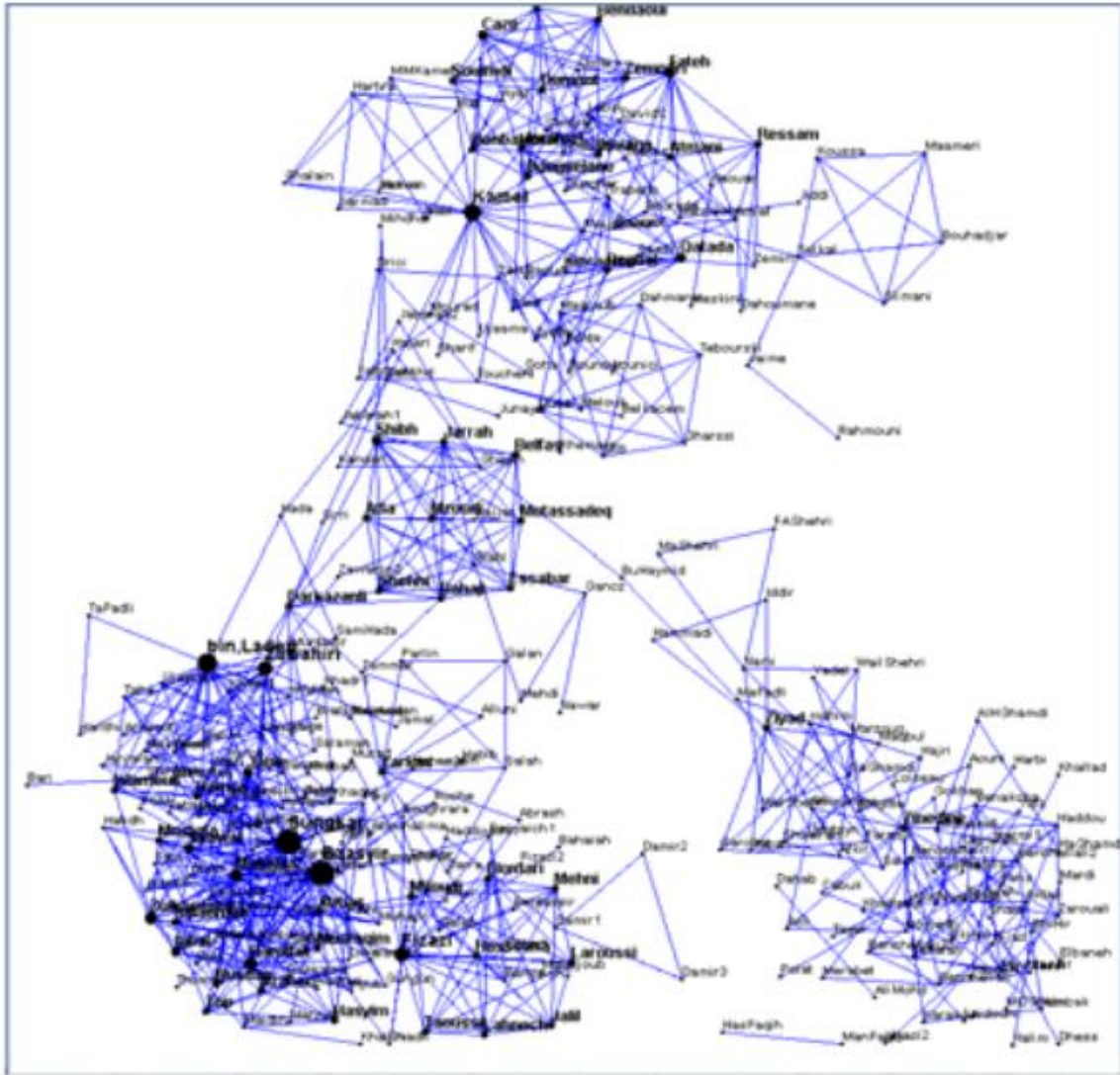
NA ZÁKLADE PODOBNOSTI

Podobné = priťahujú sa

Nepodobné = odpudzujú sa



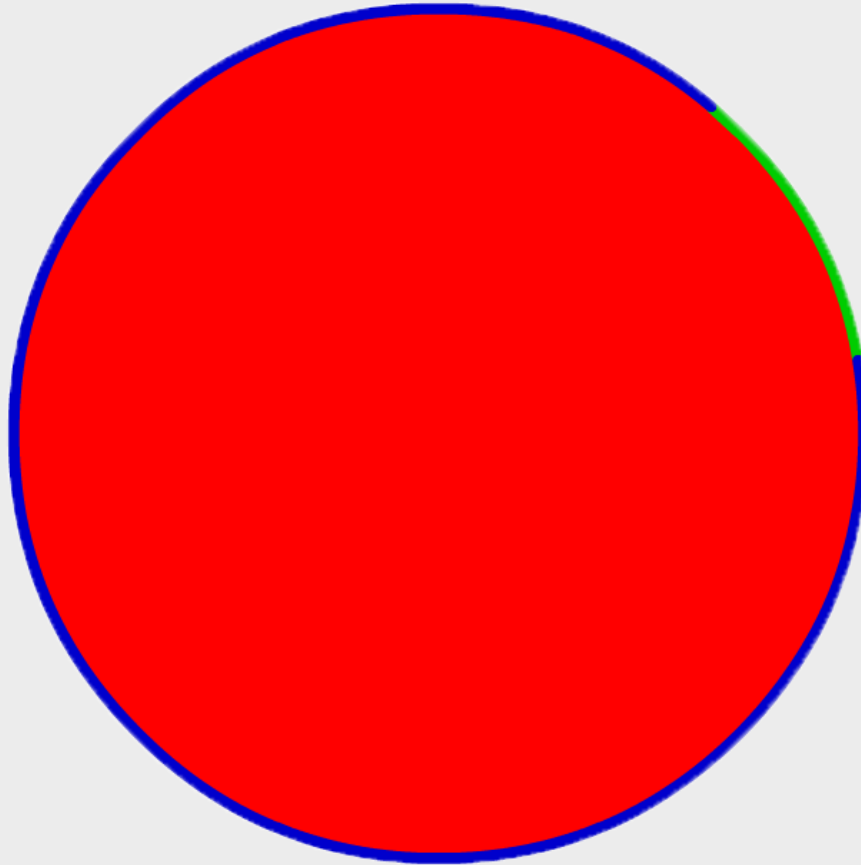
SOCIÁLNE SIETE



Analyzing the Terrorist Social Networks with Visualization Tools

Yang & Liu & Sageman

POZOR NA POČET DÁT

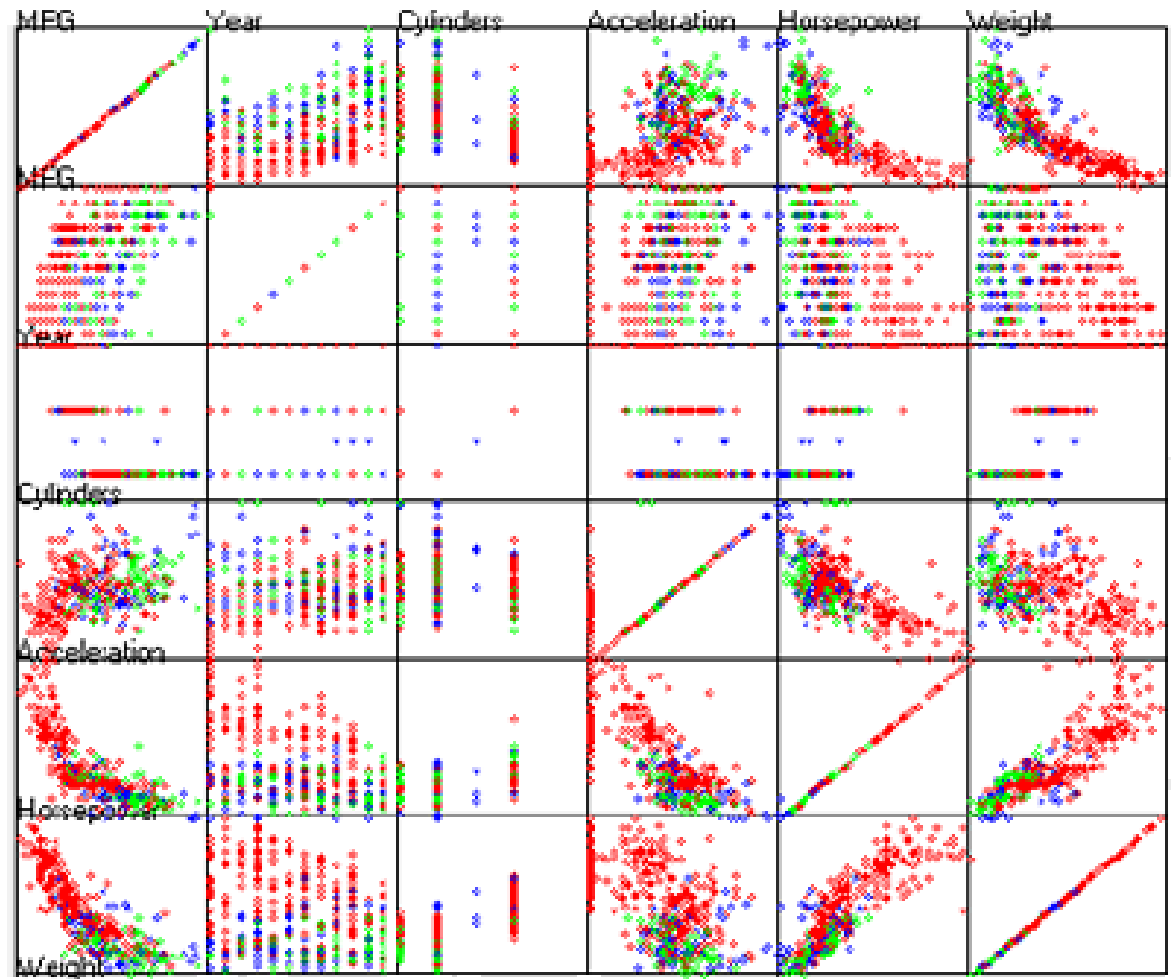


Databáza IMDB: 250.000 filmov vs. 900.000 hercov

**ČO S VIACERÝMI
ROZMERMI?**

VIAC MENEJ-ROZMERNÝCH GRAFOV

Matica
scatterplotov



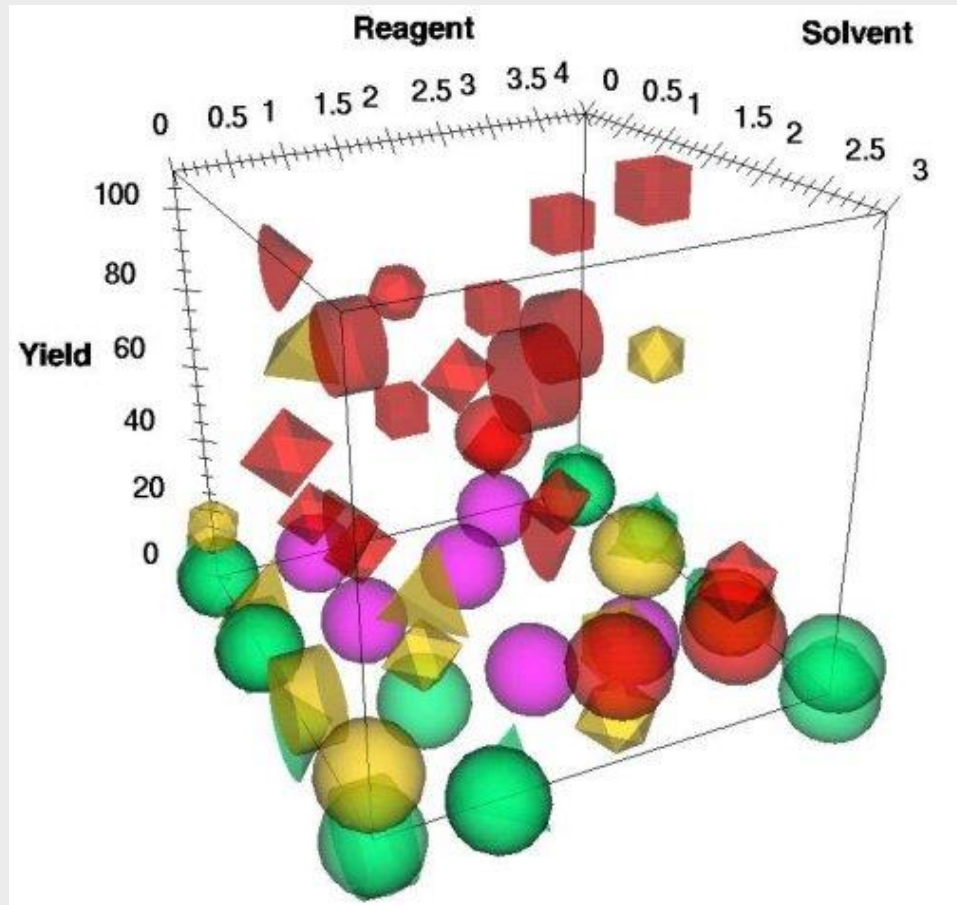
PRIDAŤ VIAC ROZMEROV DO GRAFU

TRETÍ GEOMETRICKÝ ROZMER (3D)

FARBA

TVAR

...



REDUKCIA POČTU ROZMEROV

$$ND \rightarrow PD, P < N$$

MÔŽEME DOSTAŤ AJ ÚPLNE NOVÉ ROZMERY,
KTORÉ NEMAJÚ JASNÝ VÝZNAM

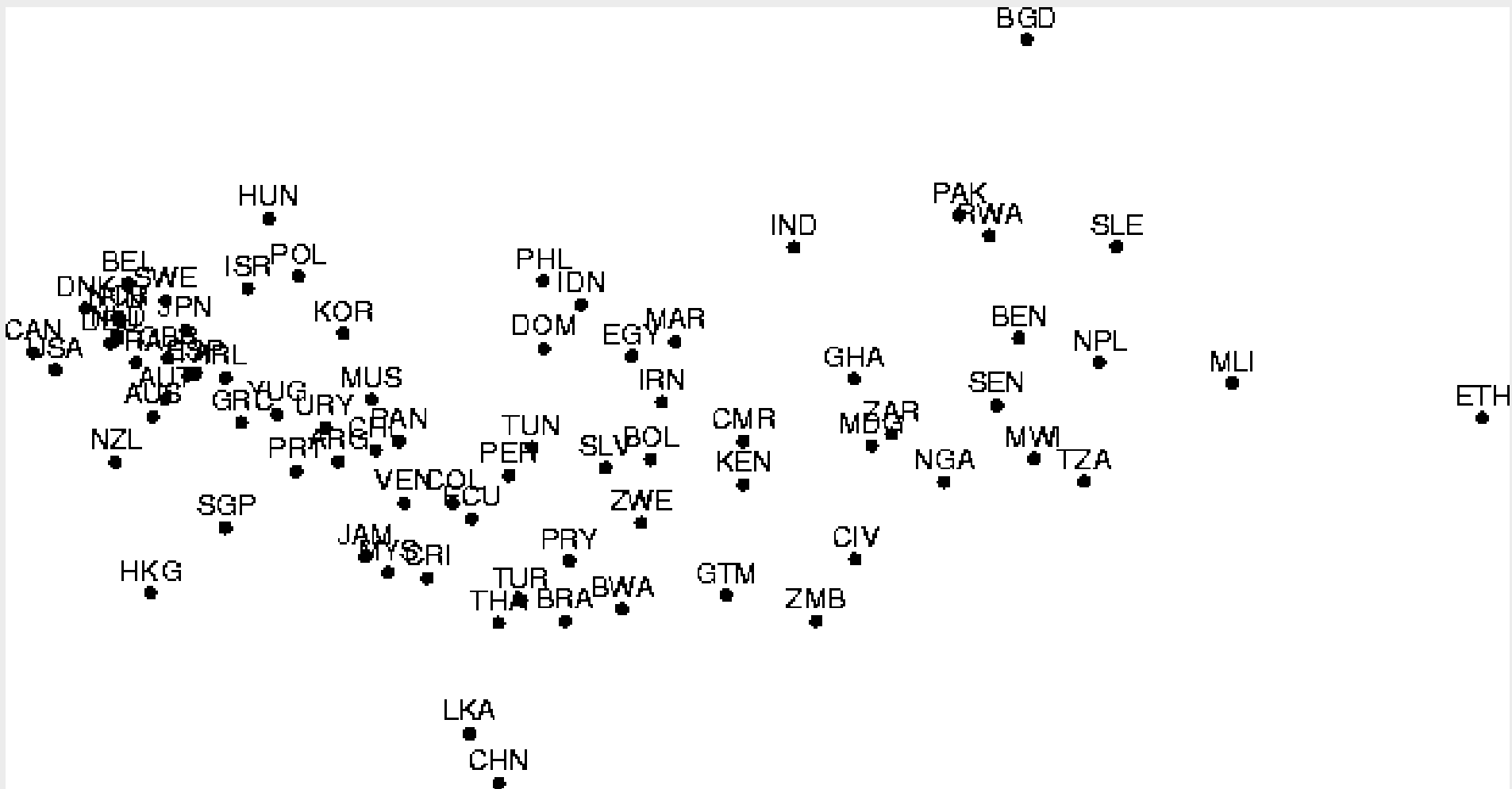
PODSTATNÉ JE ZACHOVAŤ

Vzťahy - zoskupenia, trendy, anomálie

Topológiu

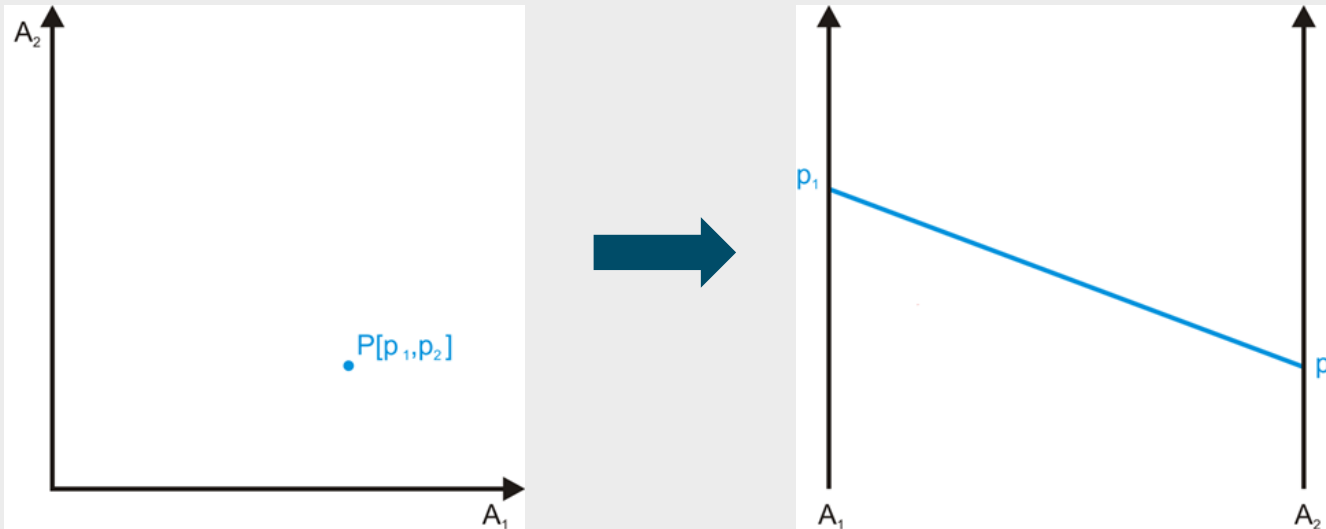
Relatívne vzdialenosti a miery

FORCE-DIRECTED LAYOUT (FDL)



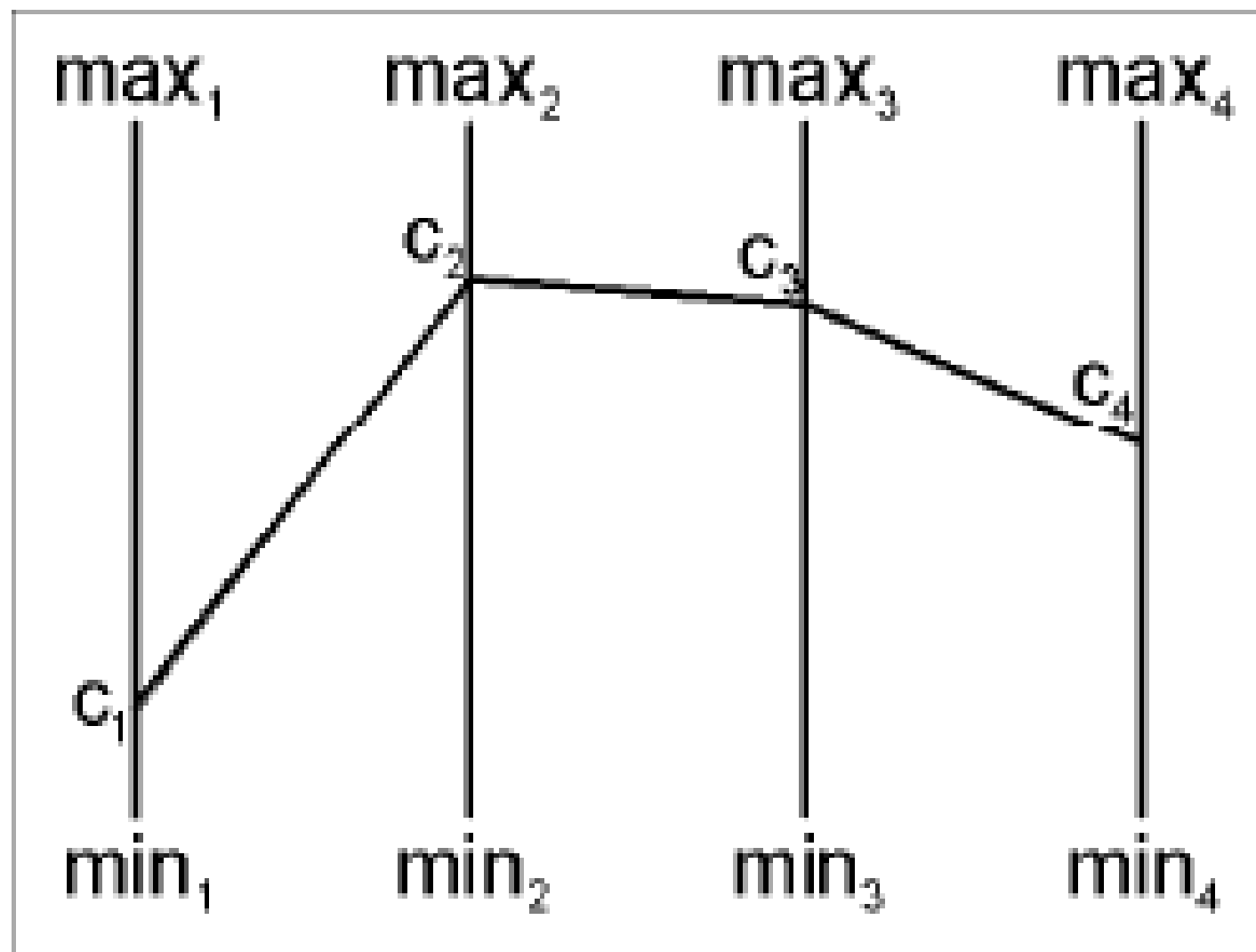
PARALELNÉ SÚRADNICE

SÚRADNICOVÉ OSI NIE SÚ KOLMÉ
ALE ROVNOBEŽNÉ
NAMIESTO BODU JE ČIARA



TAKTO MÔŽEME UMIESTNIŤ VEĽA OSÍ VEDĽA
SEBA DO ROVINY

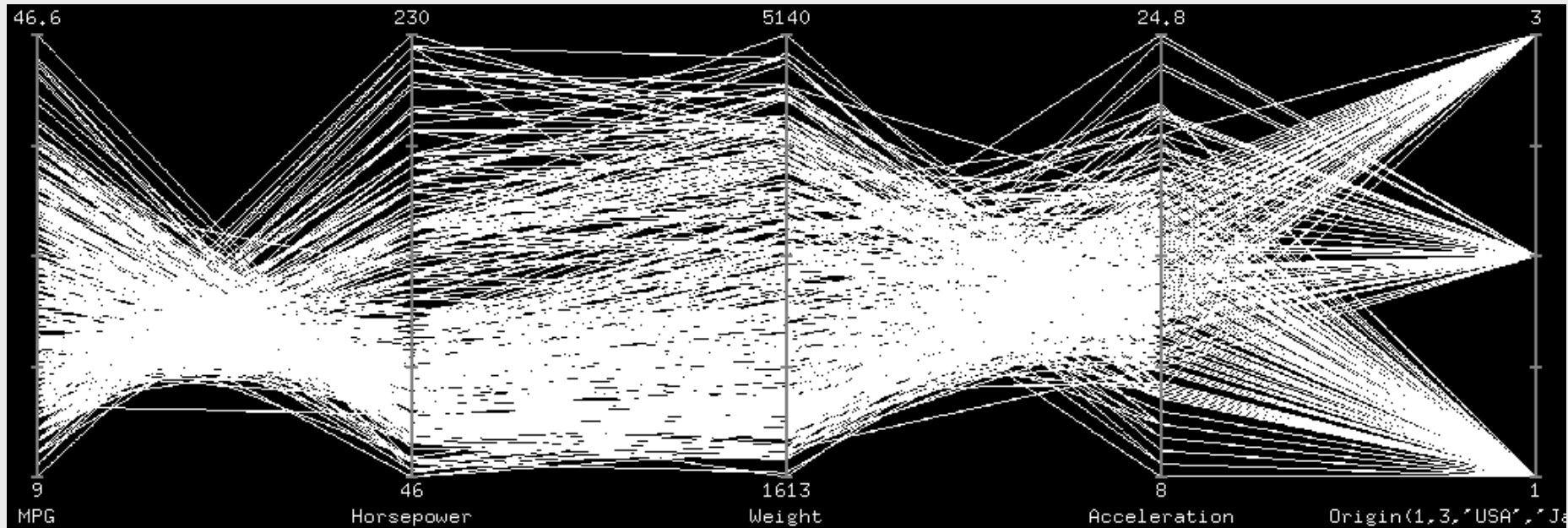
ŠTVOR-ROZMERNÉ DÁTA



PÄŤ-ROZMERNÉ DÁTA

ÚDAJE O AUTÁCH

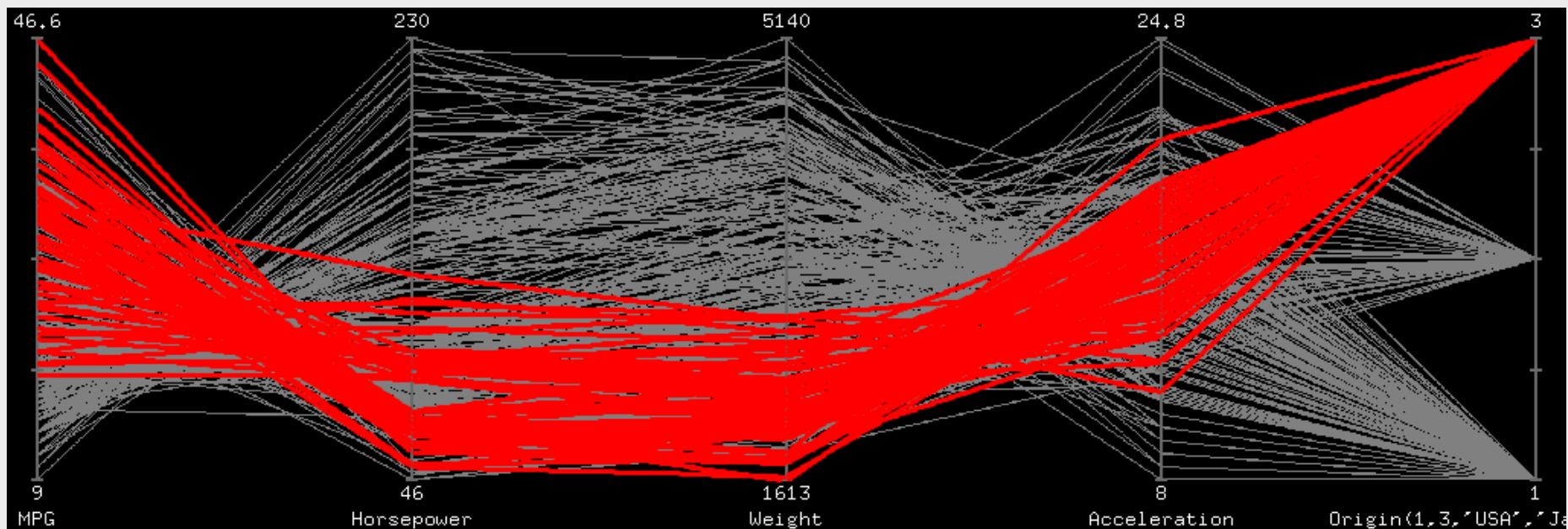
dojazd, výkon, hmotnosť, zrýchlenie, krajina



ZHLUKY DÁT

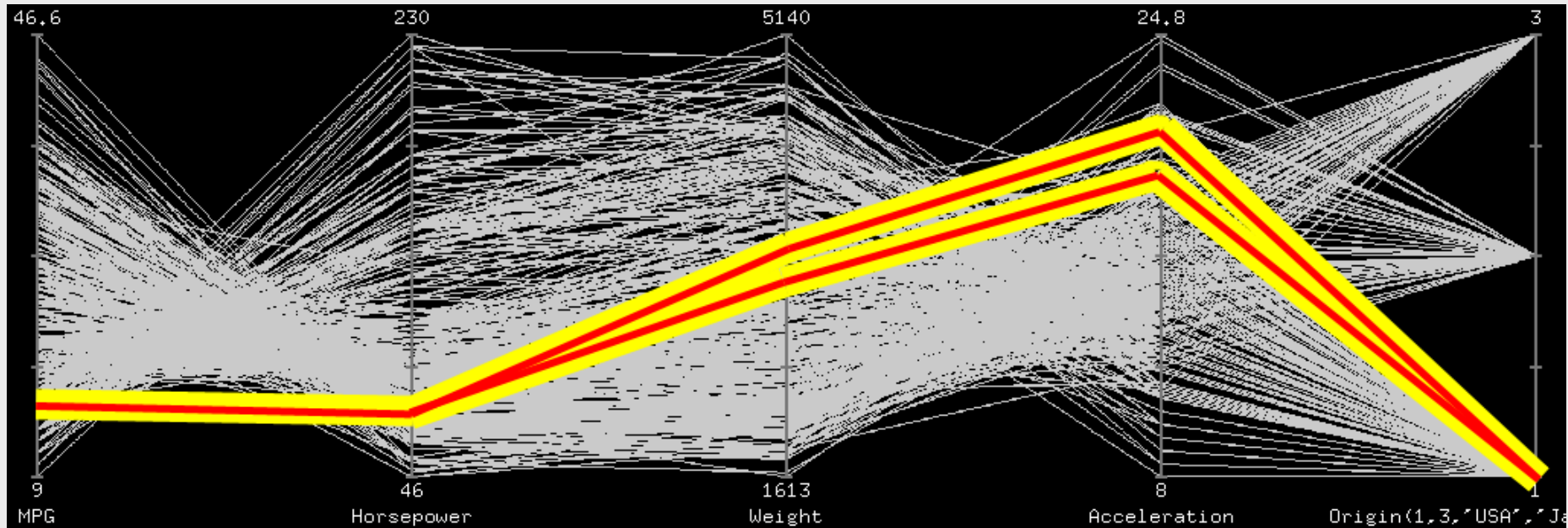
OZNAČENÉ BOLI EURÓPSKE AUTÁ

Majú nízky výkon, sú ľahké, majú dobrý dojazd



ANOMÁLIE V DÁTACH

VÝKON-DOJAZD SÚ NEGATÍVNE KORELOVANÉ
ALE NAŠLI SME ANOMÁLIU



Dve autá, ktoré sú slabé a majú vysokú spotrebu

ČASOVÉ ZÁVISLOSTI

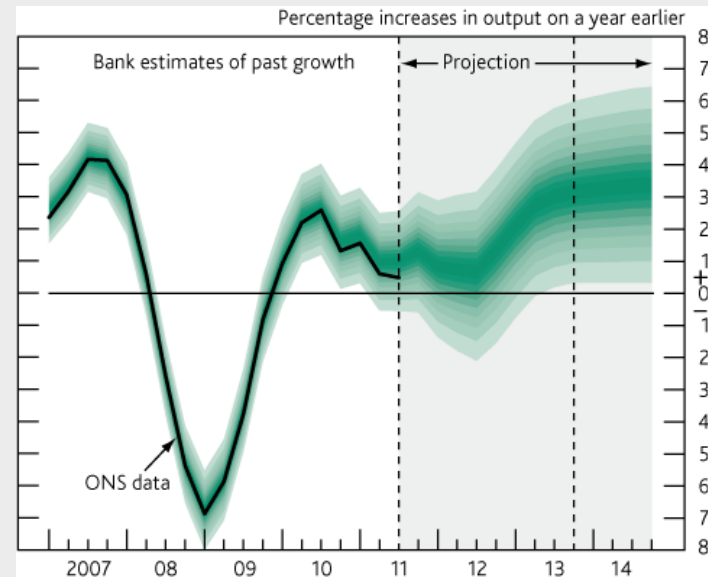
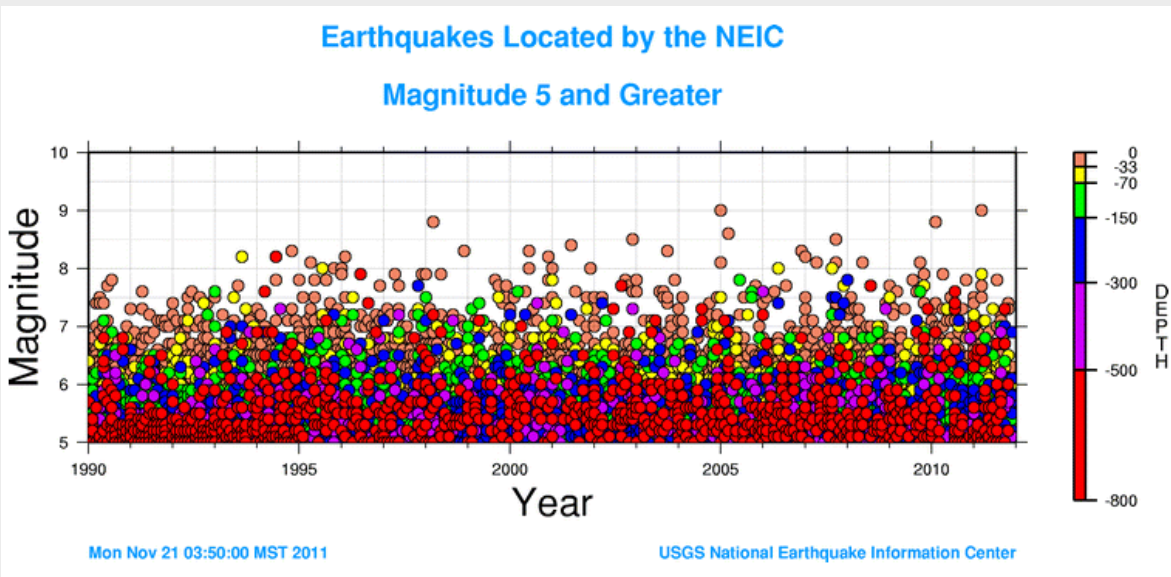
ČO SA DÁ NÁJŠŤ V ČASOVÝCH DÁTACH?

ČASOVÉ TRENDY (NÁRAST, POKLES...)

PREDPOVEDE

ČASOVÉ VZORY (CYKLY, OPAKOVANIA)

ČASOVÉ KORELÁCIE



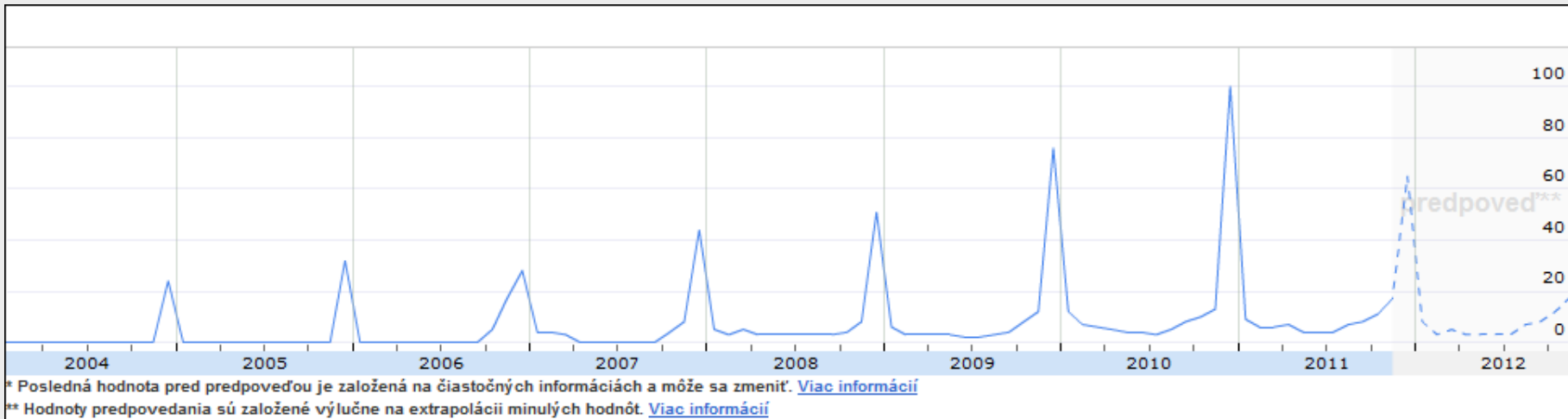
DETEKCIA UDALOSTI

UDALOSŤ - KEĎ DÔJDE K ZMENE V TRENDĎ



Google Insights pre termín „Pracovné miesta“

CYKlickÉ TRENDY, PERIODICKÉ UDALOSTI



Google Insights pre termíny „Futbal” a „Kapustnica”

DLHODOBÉ AJ LOKÁLNE SPRÁVANIE



Google Insights pre termíny „Playstation” a „XBOX”

CYKLIKÉ UDALOSTI V ŠPIRÁLOVOM GRAFE

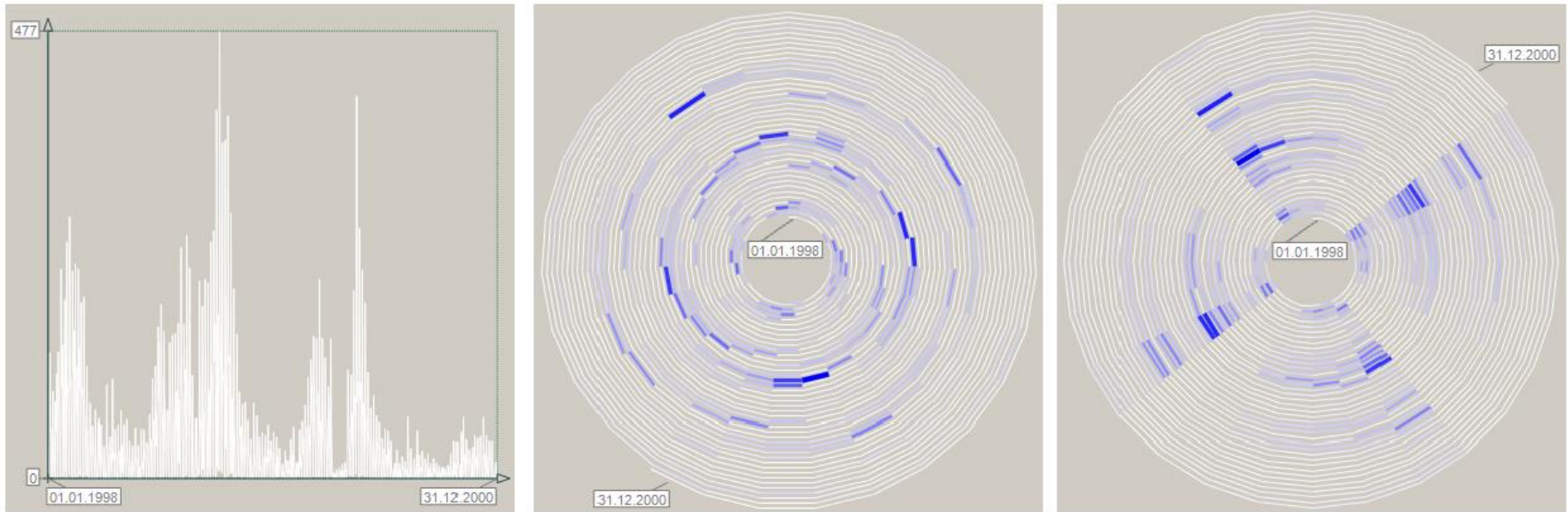
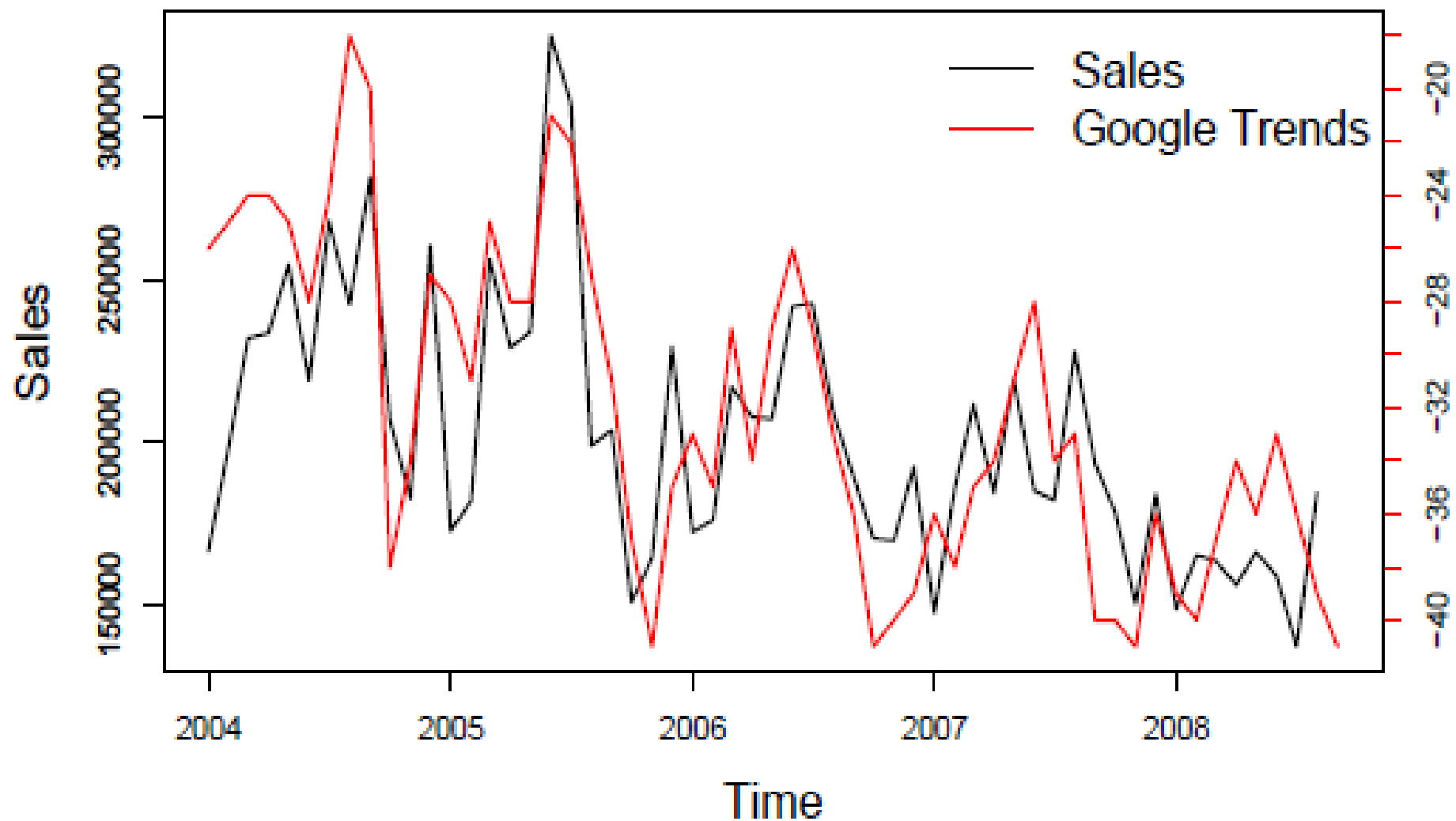


Fig. 1. Different visual representations of a time-oriented dataset describing the number of influenza cases over a period of three years – left: Time series plot (periodic pattern is difficult to discern), center: SpiralGraph encoding 27 days per cycle (improperly parameterized – periodic pattern is hard to see), right: SpiralGraph encoding 28 days per cycle (properly parameterized – periodic pattern stands out).

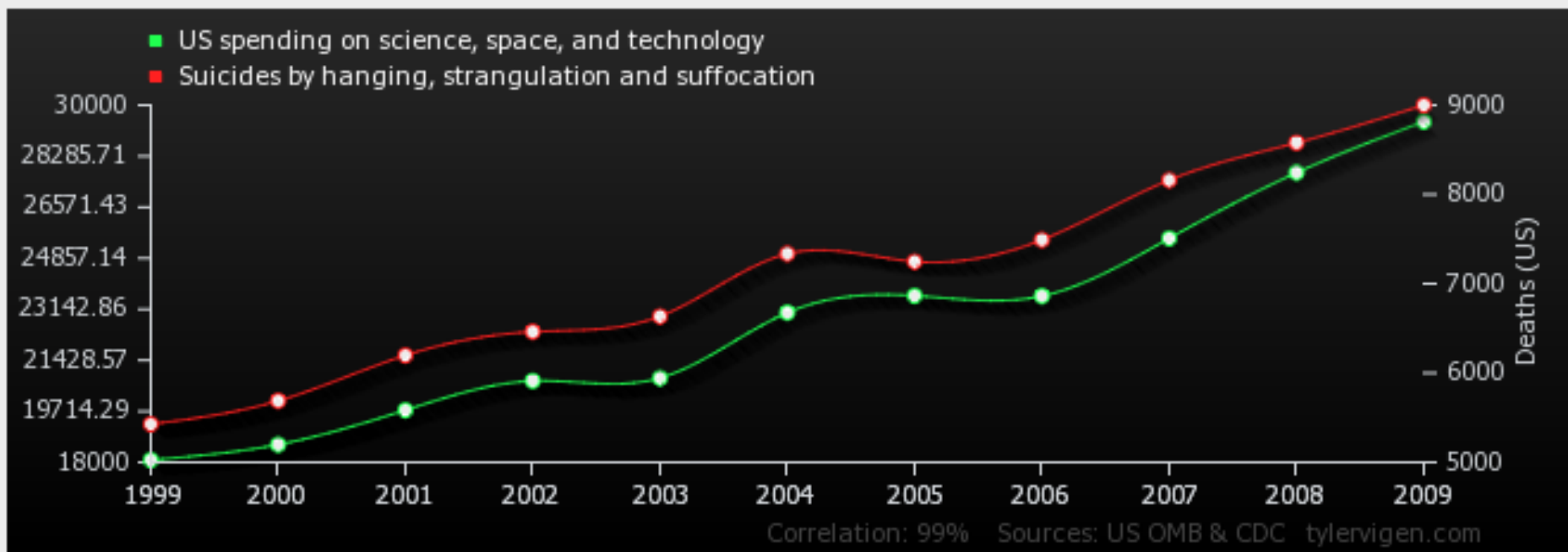
VZŤAHY MEDZI VELIČINAMI

Chevrolet



KORELÁCIA NIE JE KAUZALITA

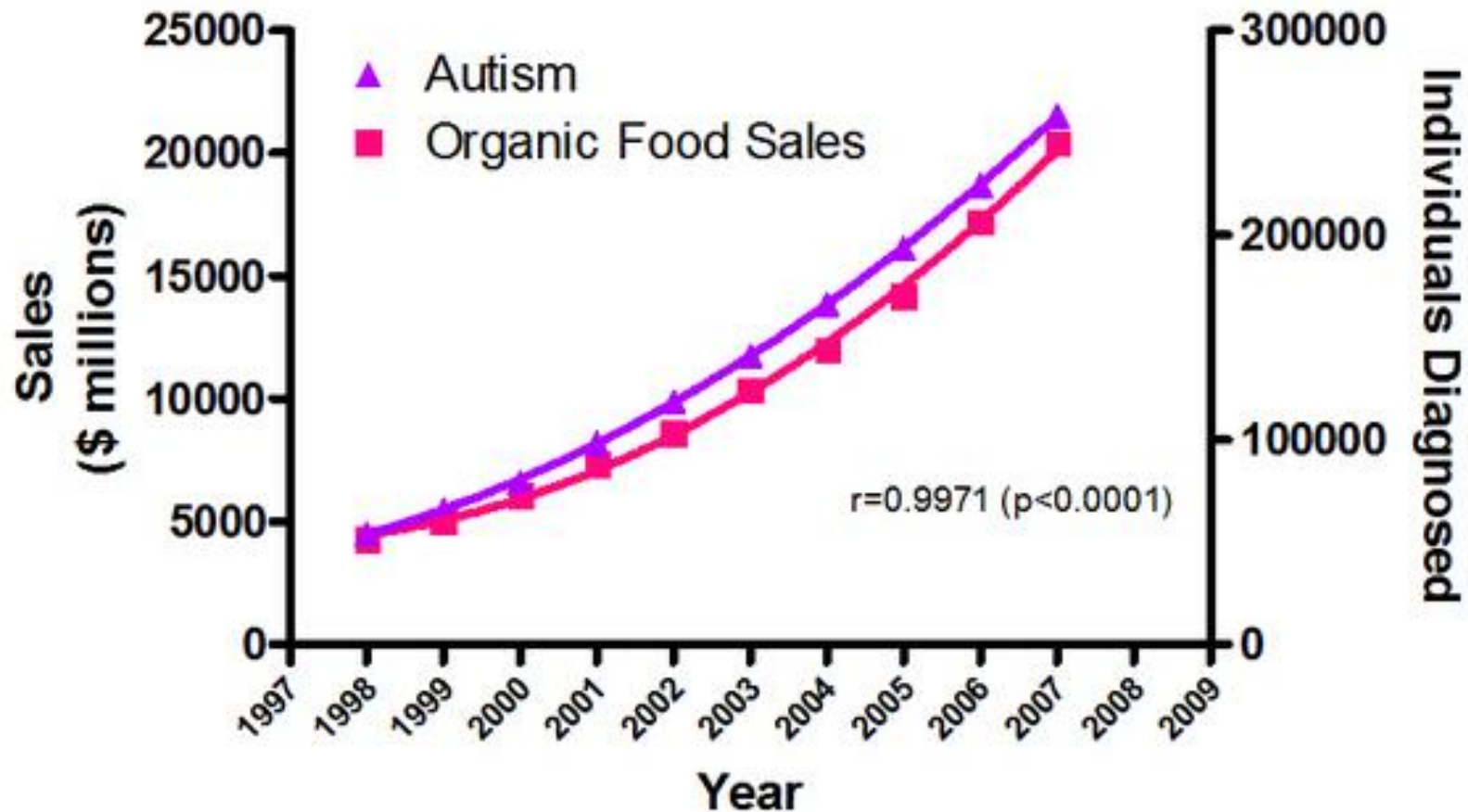
US spending on science, space, and technology
correlates with
Suicides by hanging, strangulation and suffocation



Upload this chart to imgur

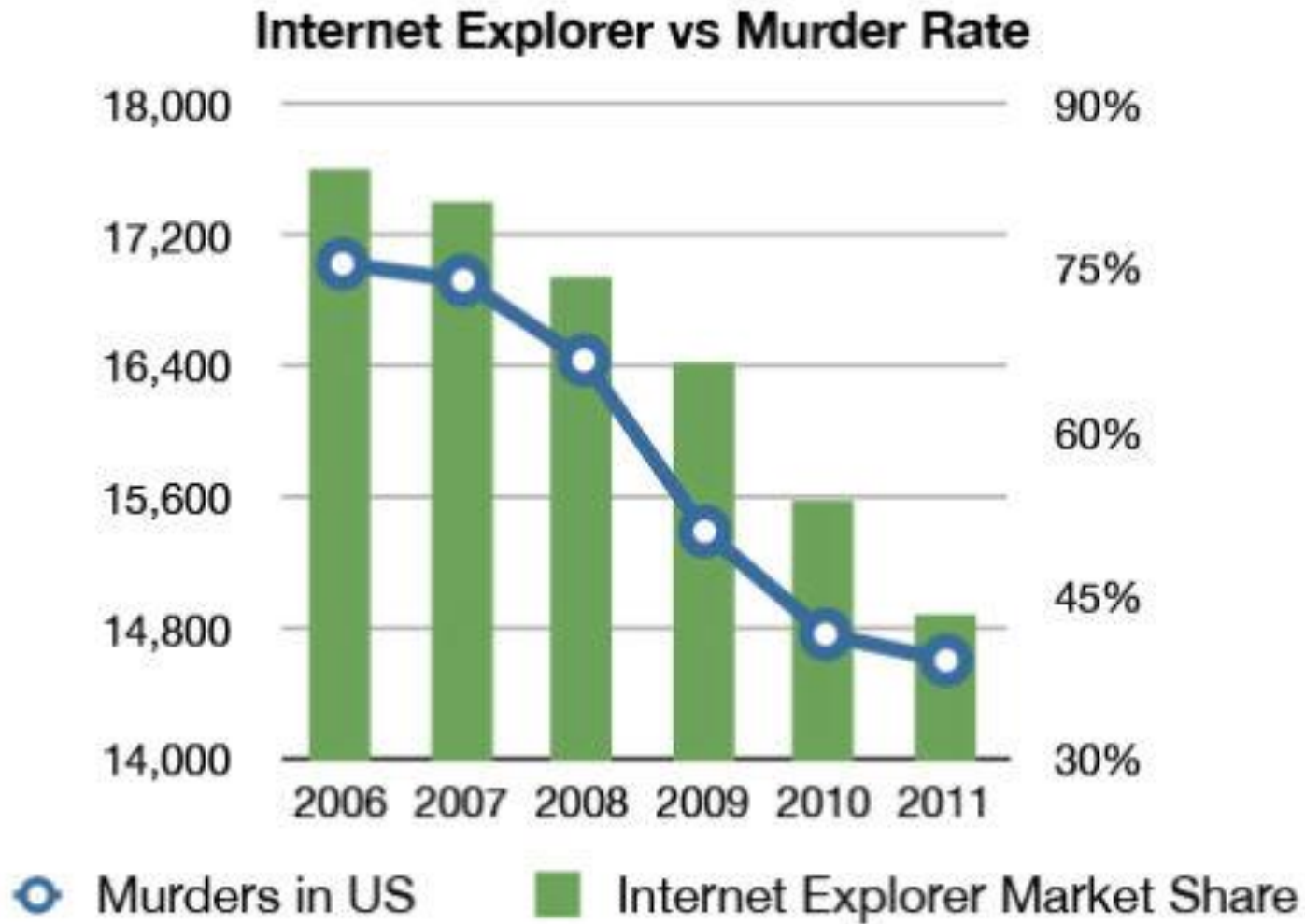
Spurious Correlations , <http://www.tylervigen.com/>

KORELÁCIA NIE JE KAUZALITA



Sources: Organic Trade Association, 2011 Organic Industry Survey; U.S. Department of Education, Office of Special Education Programs, Data Analysis System (DANS), OMB# 1820-0043: "Children with Disabilities Receiving Special Education Under Part B of the Individuals with Disabilities Education Act"

ALE NIEKEDY MOŽNO ... 😊



<http://gizmodo.com/5977989/internet-explorer-vs-murder-rate-will-be-your-favorite-chart-today>

ZÁVER

VIDIEŤ A VEDIET'

DÁTA → SÚVISLOSTI → POZNANIE

13.58801,11.0956
9.878124,12.55918
14.65398,10.18044
15.22057,9.993003
13.16265,12.96319
13.80617,8.889046
13.10214,10.56081
11.00403,11.86713
15.15475,11.70451
11.12639,9.643859
11.7104,13.59038
12.34107,11.48214
10.58664,11.86681
14.56957,10.57711
16.56765,14.32268
9.511439,10.67225
13.33699,10.70585



VIZUALIZÁCIA POMÁHA ODHALIŤ TIETO
SÚVISLOSTI ALEBO OVERIŤ SI POZNANIE