

ВИЗУАЛИЗАЦИЯ ДАННЫХ

Князьков К.В.



Визуализация: что это?

- **Визуализация данных** – это визуальное представление данных в виде, который обеспечивает наиболее эффективную работу человека по их изучению.
- Визуализация = Искусство + Наука
- UPDATE: Визуализация = Визуальное отображение + Интеративность

ВИЗУАЛИЗАЦИЯ: ЗАЧЕМ?

- Она позволяет видеть то, что иначе сложно заметить.
- Она дает ответы на многие вопросы быстрее.
- Хорошая визуализация позволяет “исследовать данные”, поиграть с ними, выявляя интересные вещи, что особенно важно в журналистике, к примеру.
- Визуализация помогает справиться с возрастающей сложностью и разнообразием данных.
- Субъективное восприятие информации, доверие к информации выше, когда она подана визуально.

ПРИМЕР – ИСТОРИЧЕСКИЕ ДАННЫЕ

French invasion of Russia

From Wikipedia, the free encyclopedia

"Russian Campaign" redirects here. For other uses, see [Russian Campaign \(disambiguation\)](#).

"Patriotic War of 1812" redirects here. Not to be confused with the "Great Patriotic War", the Russian name for the Eastern European theater of World War II.

The **French Invasion of Russia**, known in Russia as the **Patriotic War of 1812** (Russian: Отечественная война 1812 года; Otechestvennaya Voyna 1812 Goda) and in France as the **Russian Campaign** (French: Campagne de Russie), began on 24 June 1812 when Napoleon's Grande Armée crossed the Neman River in an attempt to engage and defeat the Russian army.^[8] Napoleon hoped to compel Tsar Alexander I of Russia to cease trading with British merchants through proxies in an effort to pressure the United Kingdom to sue for peace.^[9] The official political aim of the campaign was to liberate Poland from the threat of Russia. Napoleon named the campaign the Second Polish War to gain favor with the Poles and provide a political pretense for his actions.^[10]

The Grande Armée was a very large force, numbering 680,000 soldiers (including 300,000 of French departments). Through a series of long marches Napoleon pushed the army rapidly through Western Russia in an attempt to bring the Russian army to battle, winning a number of minor engagements and a major battle at Smolensk in August. Napoleon hoped the battle would mean an end of the march into Russia, but the Russian army slipped away from the engagement and continued to retreat into Russia, while leaving Smolensk to burn.^[11] Plans Napoleon had made to quarter at Smolensk were abandoned, and he pressed his army on after the Russians.^[12]

As the Russian army fell back, Cossacks were given the task of burning villages, towns and crops.^[9] This was intended to deny the invaders the option of living off the land. These scorched-earth tactics greatly surprised and disturbed the French, as the willingness of the Russians to destroy their own territory and harm their own people was difficult for the French to comprehend.^[13] The actions forced the French to rely on a supply system that was incapable of feeding the large army in the field. Starvation and privation compelled French soldiers to leave their camps at night in search of food. These men were frequently confronted by parties of Cossacks, who captured or killed them.

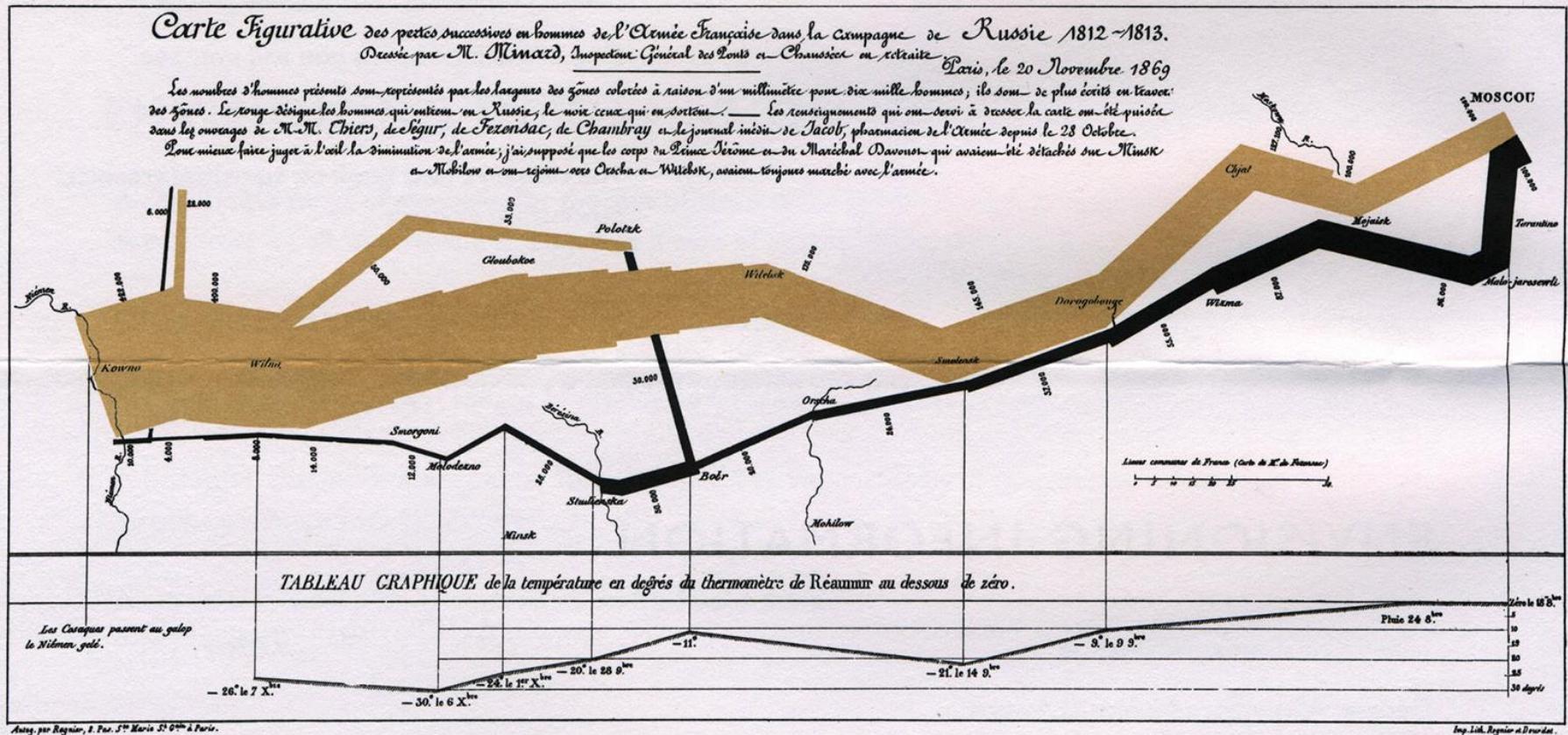
The Russian army retreated into Russia for almost three months. The continual retreat and the loss of lands to the French upset the Russian nobility. They pressured Alexander I to relieve the commander of the Russian army, Field Marshal Barclay. Alexander I complied, appointing an old veteran, Prince Mikhail Kutuzov, to take over command of the army.

On 7 September the French caught up with the Russian army which had dug itself in on hillsides before a small town called Borodino, seventy miles west of Moscow. The battle that followed was the largest and bloodiest single-day action of the Napoleonic Wars, involving more than 250,000 soldiers and resulting in 70,000 casualties. The French gained a victory, but at the cost of 49 general officers and thousands of men. The Russian army was able to extricate itself and withdrew the following day, leaving the French without the decisive victory Napoleon sought.^[14]

Napoleon entered Moscow a week later. In another turn of events the French found puzzling, there was no delegation to meet the Emperor. The

French invasion of Russia	
Part of Napoleonic Wars	
	
	
Clockwise from top left: The Battle of Borodino by Louis Lejeune; The Fire of Moscow by Albrecht Adam; Marshal Ney at the battle of Kaunas by Auguste Raffet; French retreat by Illarion Pryanishnikov.	
Date	24 June – 14 December 1812 (5 months, 2 weeks and 6 days)
Location	Russian Empire
Result	Decisive Russian victory ^[1] Destruction of French Allied Army Start of the War of the Sixth Coalition
Belligerents	
 French Empire	 Russian Empire
 Duchy of Warsaw	Formal ally:
 Napoleon's Italy	 Sweden

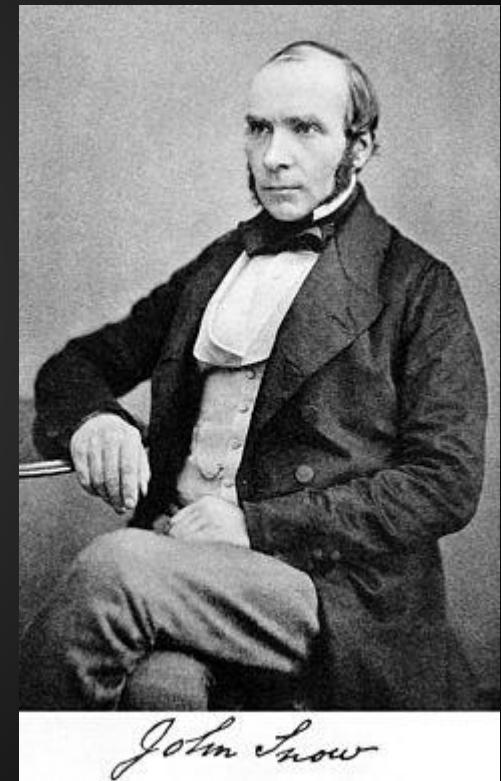
ПРИМЕР ВИЗУАЛИЗАЦИИ

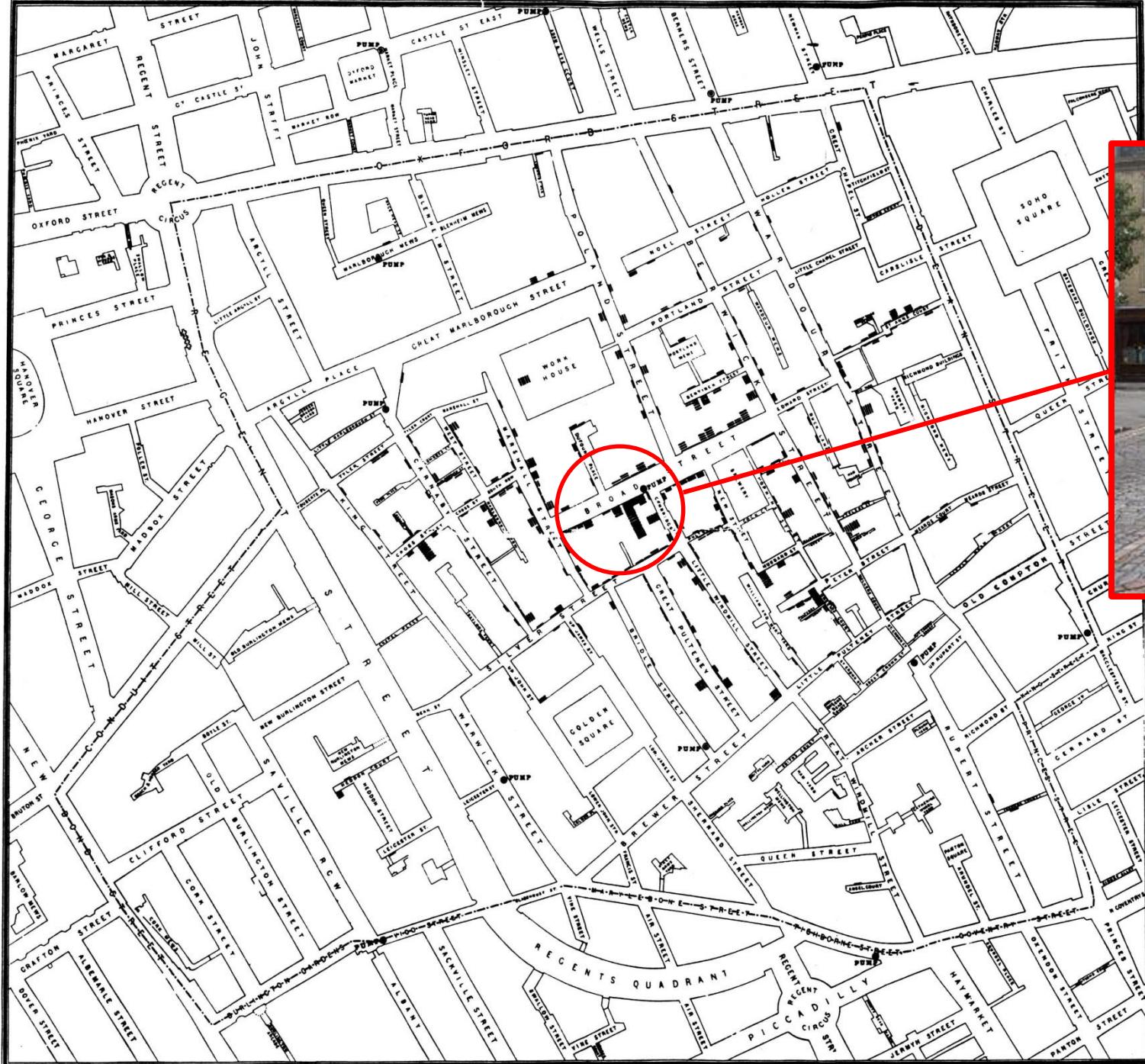


координаты – позиция на карте, ширина – количество войск, цвет – направление, X – время, Y – температура

ПРИМЕР – ЭПИДЕМИЯ ХОЛЕРЫ

- 1845 г. эпидемия холеры в центре Лондона
- На то время способ передачи болезни не был известен
- Джон Сноу провел исследование случаев заражения и определили причину





Визуализация: ЗАЧЕМ?

1. Эффективная передача знаний при общении (рассказ истории)
2. Анализ данных для выявления причин явлений и для понимания скрытых эффектов

Квартет Энскомба

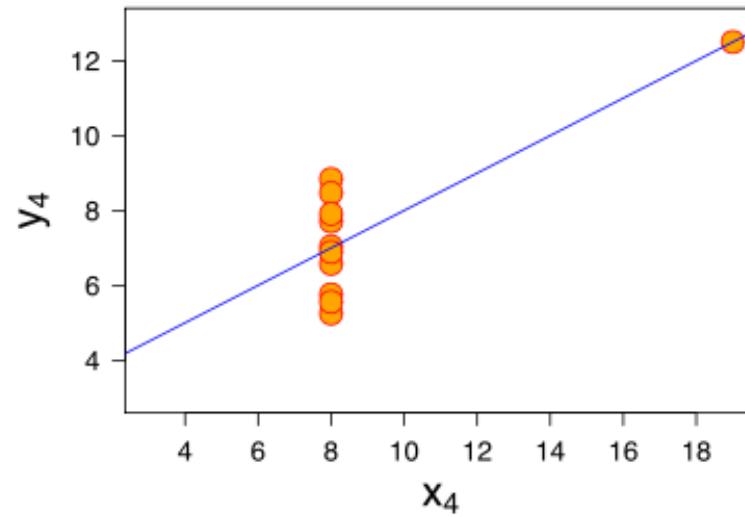
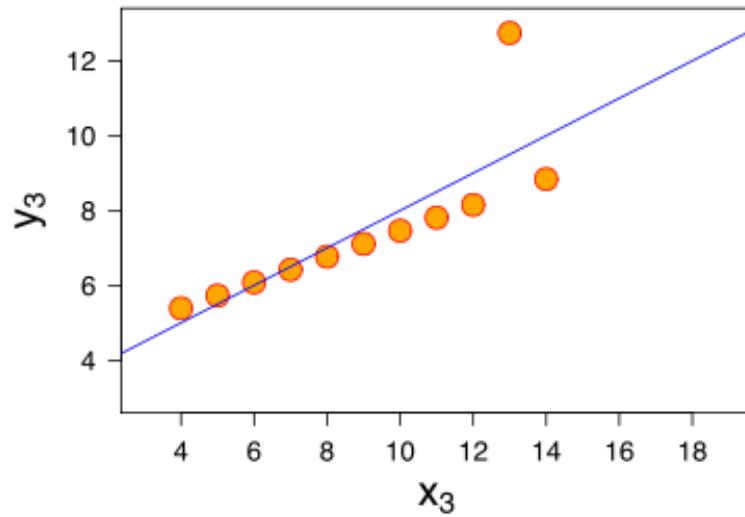
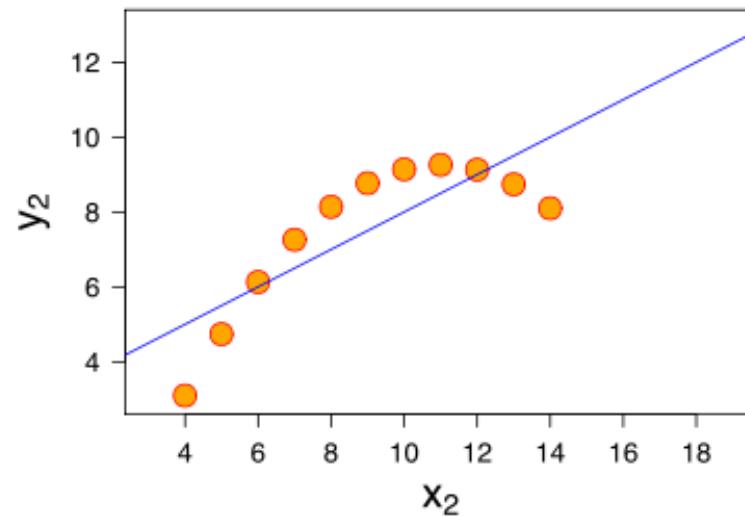
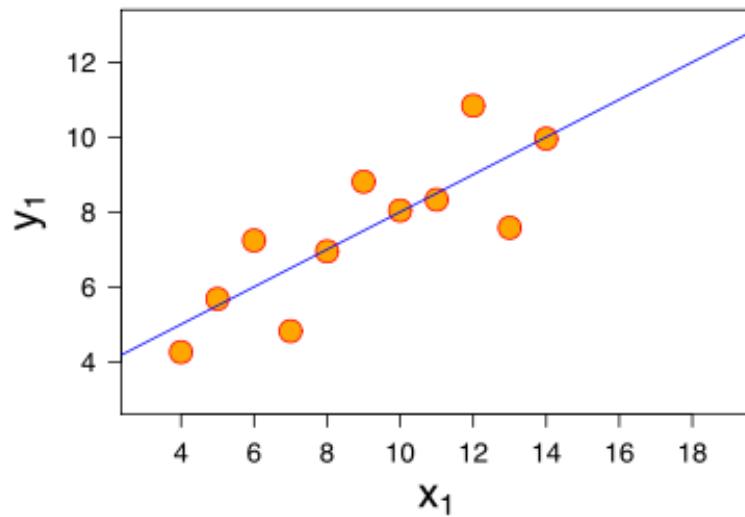
Квартет Энскомба

Статистические свойства

I		II		III		IV	
x	y	x	y	x	y	x	y
10.0	8.04	10.0	9.14	10.0	7.46	8.0	6.58
8.0	6.95	8.0	8.14	8.0	6.77	8.0	5.76
13.0	7.58	13.0	8.74	13.0	12.74	8.0	7.71
9.0	8.81	9.0	8.77	9.0	7.11	8.0	8.84
11.0	8.33	11.0	9.26	11.0	7.81	8.0	8.47
14.0	9.96	14.0	8.10	14.0	8.84	8.0	7.04
6.0	7.24	6.0	6.13	6.0	6.08	8.0	5.25
4.0	4.26	4.0	3.10	4.0	5.39	19.0	12.50
12.0	10.84	12.0	9.13	12.0	8.15	8.0	5.56
7.0	4.82	7.0	7.26	7.0	6.42	8.0	7.91
5.0	5.68	5.0	4.74	5.0	5.73	8.0	6.89

Свойство	Значение
Среднее X	9
Дисперсия X	11
Среднее Y	7.50
Дисперсия Y	4.122 или 4.127
Корелляция между X и Y	0.816
Линейная регрессия	$y = 3.00 + 0.500x$

Квартет Энскомба



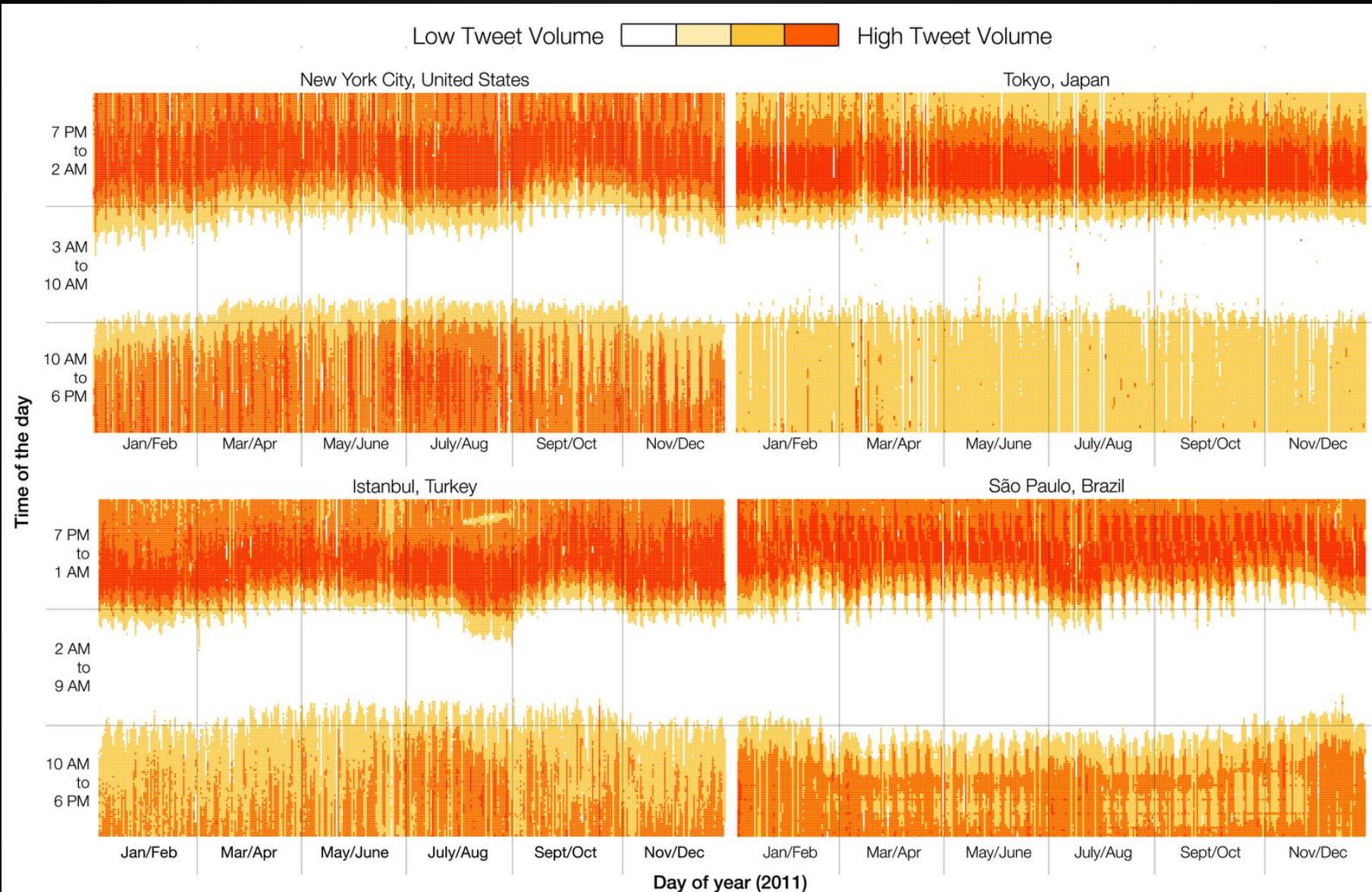
ПРИМЕРЫ «ХОРОШЕЙ» ВИЗУАЛИЗАЦИИ



ПЕРИОДИЧЕСКАЯ СИСТЕМА ХИМИЧЕСКИХ ЭЛЕМЕНТОВ Д. И. МЕНДЕЛЕЕВА

ПЕРИОДЫ	ГРУППЫ ЭЛЕМЕНТОВ																		VIII	B		
	A I	B A II	B A III	B A IV	B A V	B A VI	B A VII	B A														
1	(H)															H	1.00794 Hydrogenium Водород	He	2.002602 Helium Гелий			
2	Li Lithium Литий	6.941	Be Beryllium Бериллий	9.0122	B Boron Бор	10.811	C Carbonium Углерод	12.011	N Nitrogenium Азот	14.007	O Oxygenium Кислород	15.999	F Fluorium Фтор	18.998	Ne Neon Неон	20.179	Ar Argon Аргон	39.948 Argon Аргон	Символ элемента Относительная атомная масса Порядковый номер			
3	Na Natrium Натрий	22.99	Mg Magnesium Магний	24.305	Al Aluminum Алюминий	26.9815 ²	Si Silicium Кремний	28.086	P Phosphorus Фосфор	30.974	S Sulfur Сера	32.066	Cl Chlorium Хлор	35.453	Ar Argon Аргон	39.948			Название элемента Распределение электронов на энергетических уровнях			
4	K Kalium Калий	39.098	Ca Calcium Кальций	40.08	Sc Scandium Скандий	44.956	Ti Titanium Титан	47.90	V Vanadium Ванадий	50.941	Cr Chromium Хром	54.938	Mn Manganese Марганец	55.847	Fe Ferrum Железо	58.933	Co Cobaltum Кобальт	58.70 Ni Nickolum Никель				
	Cu Cuprum Медь	63.546	Zn Zincum Цинк	65.39	Ga Gallium Галий	69.72	Ge Germanium Германий	72.59	As Arsenicum Мышьяк	74.992	Se Selenium Селен	78.96	Br Bromum Бром	79.904	Kr Krypton Криптон	83.80						
5	Rb Rubidium Рубидий	85.468	Sr Strontium Стронций	87.62	Y Yttrium Иттрий	88.906	Zr Zirconium Цирконий	91.22	Nb Niobium Ниобий	92.906	Mo Molybdaenum Молибден	95.94	Tc Technetium Технеций	97.91	Ru Ruthenium Рутений	102.906	Rh Rhodium Родий	106.4 Pd Palladium Палладий				
	Ag Argentum Серебро	107.868	Cd Cadmium Кадмий	112.41	In Indium Индий	114.82	Sn Stannum Олово	118.71	Sb Stibium Сурьма	121.75	Te Tellurium Теллур	127.60	I Iodum Иод	126.9045	Xe Xenon Ксенон	131.29						
6	Cs Cesium Цезий	132.905	Ba Barium Барий	137.33	La* Lanthanum Лантан	138.9055	Hf Hafnium Гафний	178.49	Ta Tantalum Тантал	180.9479	W Wolframium Вольфрам	183.85	Re Rhenium Рений	186.207	Os Osmium Осмий	192.22	Ir Iridium Иридий	195.08 Pt Platinum Платина				
	Au Aurum Золото	196.967	Hg Hydrargyrum Ртуть	200.59	Tl Thallium Таллий	204.38	Pb Plumbum Свинец	207.19	Bi Bismuthum Висмут	208.980	Po Polonium Полоний	209.98	At Astatum Астат	209.99	Rn Radon Радон	[222]						
7	Fr Francium Франций	[223]	Ra Radium Радий	[226]	Ac** Actinium Актиний	[227]	Rf Rutherfordium Феरзерфордий	[261]	Db Dubnium Дубний	[262]	Sg Seaborgium Сиборгий	[263]	Tb Terbium Тербий	[262]	Dy Dysprosium Диспрозий	[262]	Ho Holmium Гольмий	[262]	Er Erbium Эрбий	[262]	Yb Ytterbium Иттербий	[269]
	FORMУЛЫ ВЫСШИХ ОКСИДОВ		R ₂ O		RO		R ₂ O ₃		RO ₂		R ₂ O ₅		RO ₃		R ₂ O ₇				RO ₄			
	ФОРМУЛЫ ЛЕТЧИХ ОДНОРОДНЫХ СОЕДИНЕНИЙ								RH ₄		RH ₃		RH ₂		RH							
ЛАНТАНОИДЫ*		58 Ce Сериум Церий	59 Pr Прасеодим Празеодим	60 Nd Нодим Неодим	61 Pm Прометий Прометий	62 Sm Самарий Самарий	63 Eu Европий Европий	64 Gd Гадолиний Гадолиний	65 Tb Тербий Тербий	66 Dy Диспрозий Диспрозий	67 Ho Холмий Гольмий	68 Er Эрбий Эрбий	69 Tm Титаниум Титаниум	70 Yb Иттербий Иттербий	71 Lu Лютениум Лютениум							
АКТИНОИДЫ**		80 Th Thorium Торий	81 Pa Protactinium Протактиний	82 U Уран Уран	83 Np Neptunium Нептуний	84 Pu Plutonium Плутоний	85 Am Americium Америций	86 Cm Curium Кюрий	87 Bk Berkrium Берклий	88 Cf Californium Калифорний	89 Es Einsteinium Энштейн	90 Fm Fermium Фермий	91 Md Mendelevium Менделевий	92 No Nobelium Нобелий	93 Lr Lawrencium Лауренсий							

АКТИВНОСТЬ ПОЛЬЗОВАТЕЛЕЙ TWITTER

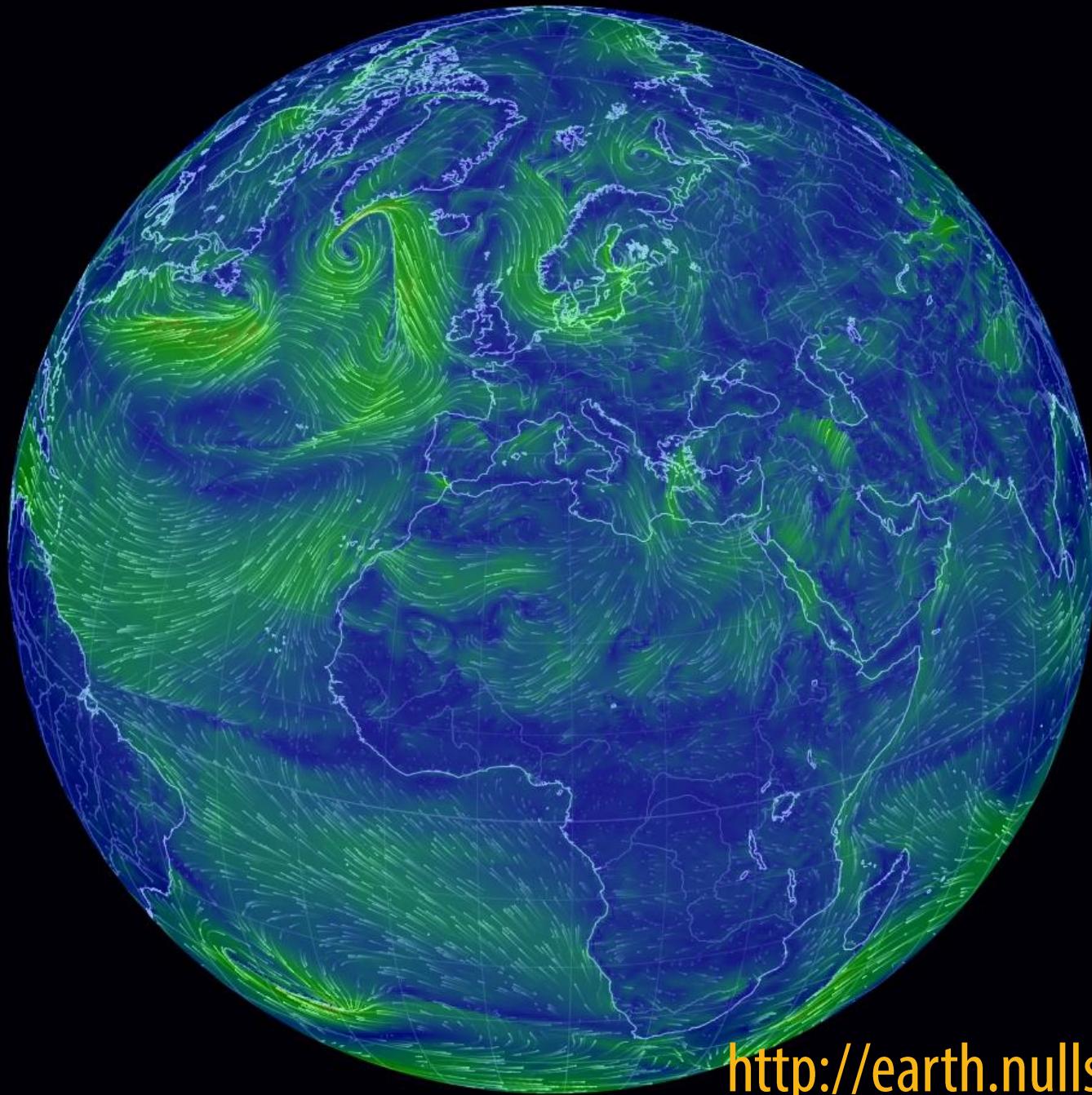


THE OPTE PROJECTS

<http://www.opte.org/>

<https://blog.twitter.com/2012/studying-rapidly-evolving-user-interests>

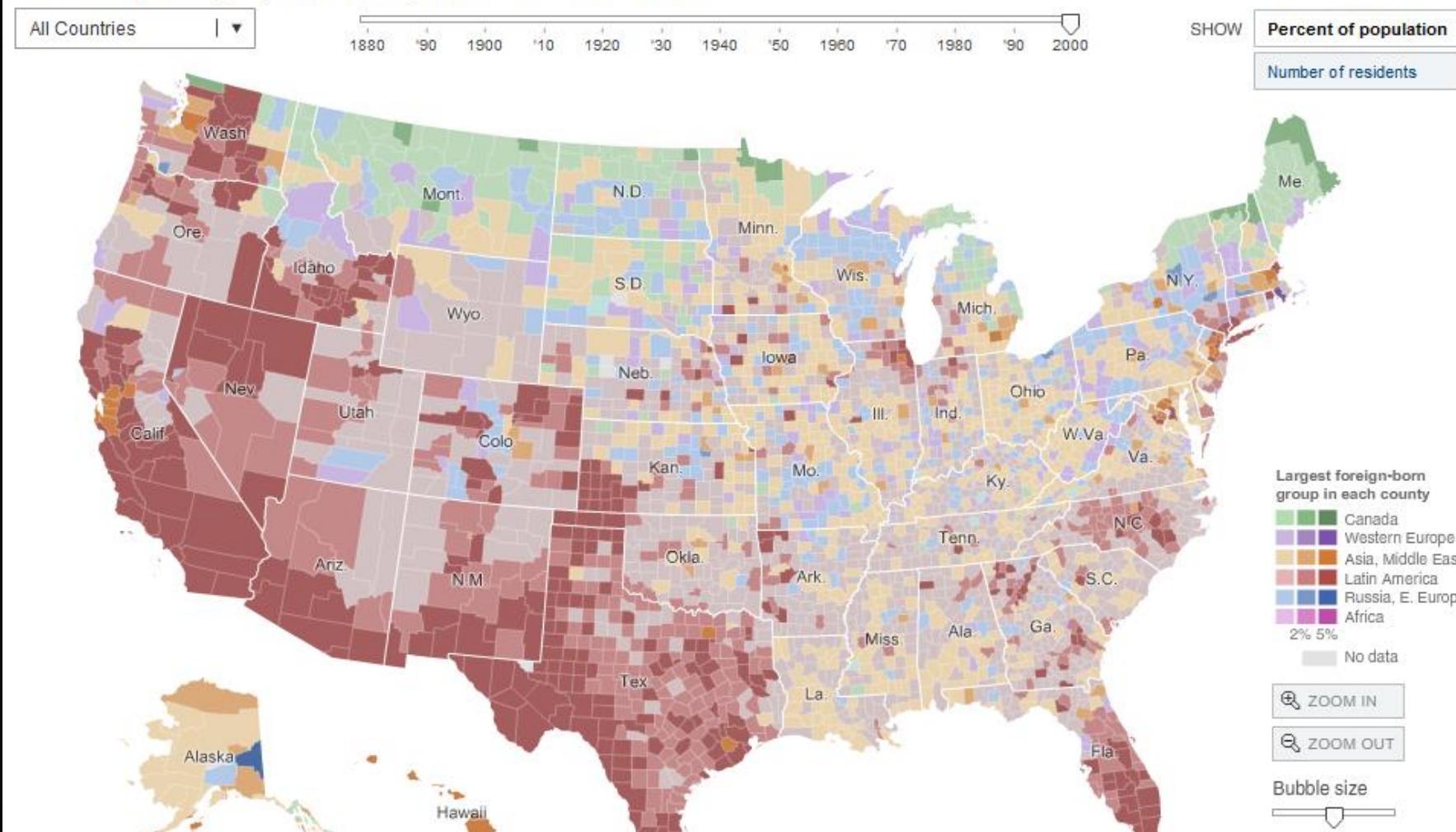
<https://blog.twitter.com/201>



<http://earth.nullschool.net>

Immigration Explorer

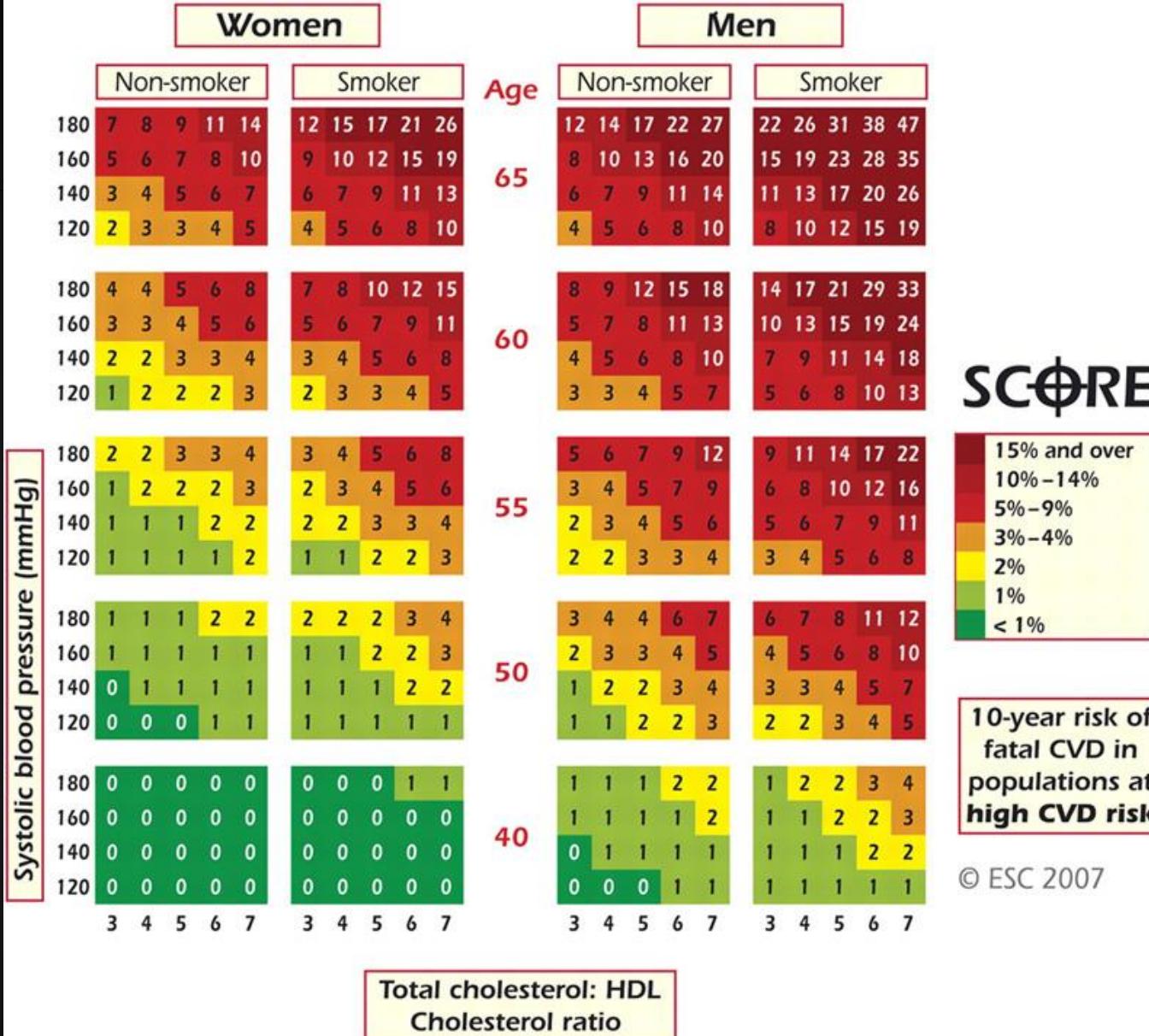
Select a foreign-born group to see how they settled across the United States.



Note: Due to limitations in the Census data, foreign-born populations are not available in all areas for all years.

http://www.nytimes.com/interactive/2009/03/10/us/20090310-immigration-explorer.html?_r=0

SCORE



Визуализация риска заболеть сердечно-сосудистыми заболеваниями в течение 10 лет в зависимости от пола, возраста, артериального давления, уровня холестерина и курения

ПРИМЕРЫ ПЛОХОЙ ВИЗУАЛИЗАЦИИ



Anatomy of a Winning TED Talk

1%

Sophisticated Visual Aids

We're not sure who puts the D in TED—most of the best presentations favor tepid PowerPoint slide shows (sorry, Brené Brown), Pictionary-quality drawings (really, Simon Sinek?), or no props at all.

5%

Opening Joke

Remember the one about the shoe salesmen who went to Africa in the 1900s? That's how Benjamin Zander opened his talk—which turned out to be about classical music.

5%

Spontaneous Moment

Don't overprepare. Tease the guy in the front row ("You could light up a village with this guy's eyes"). Command the stagehand who handles the human brain you brought.

5%

Statement of Utter Certainty

People come for answers—give 'em what they want, as Shawn Achor did: "By training your brain ... we can reverse the formula for happiness and success."

12%

Snappy Refrain

The TED equivalent of "I have a dream." Example: "People don't buy what you do; they buy why you do it." Repeat 7x.

23%

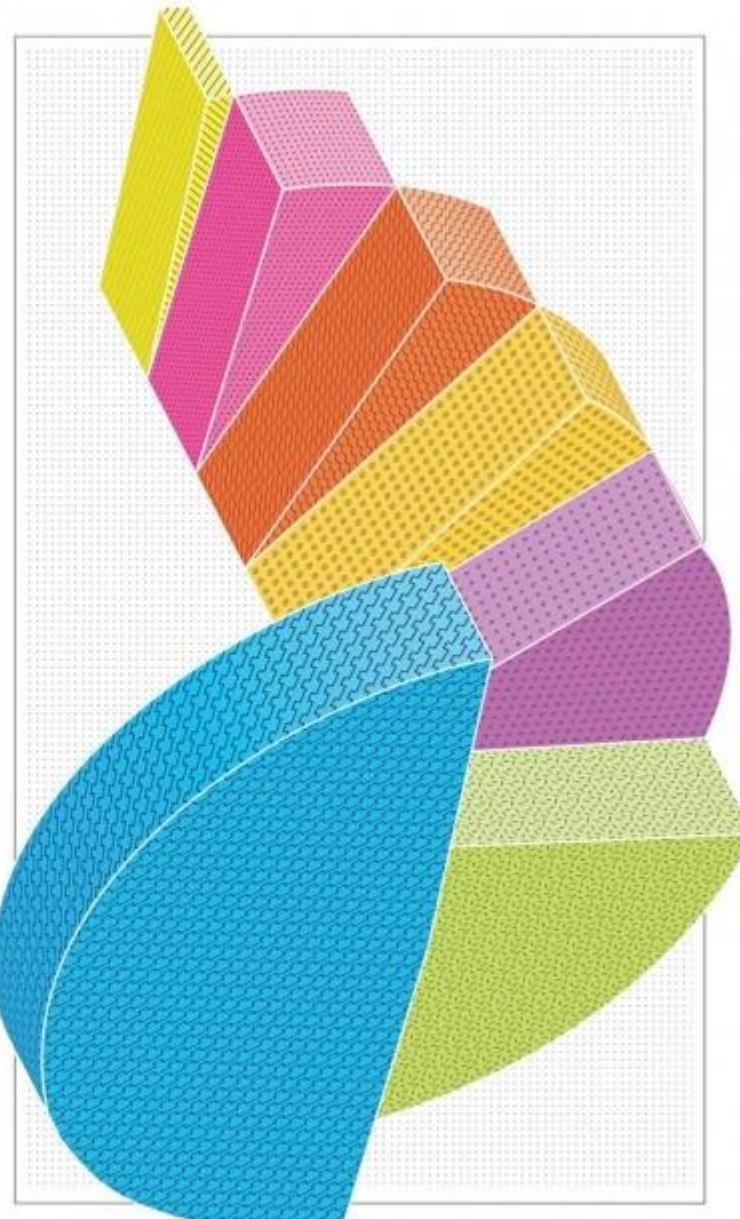
Personal Failure

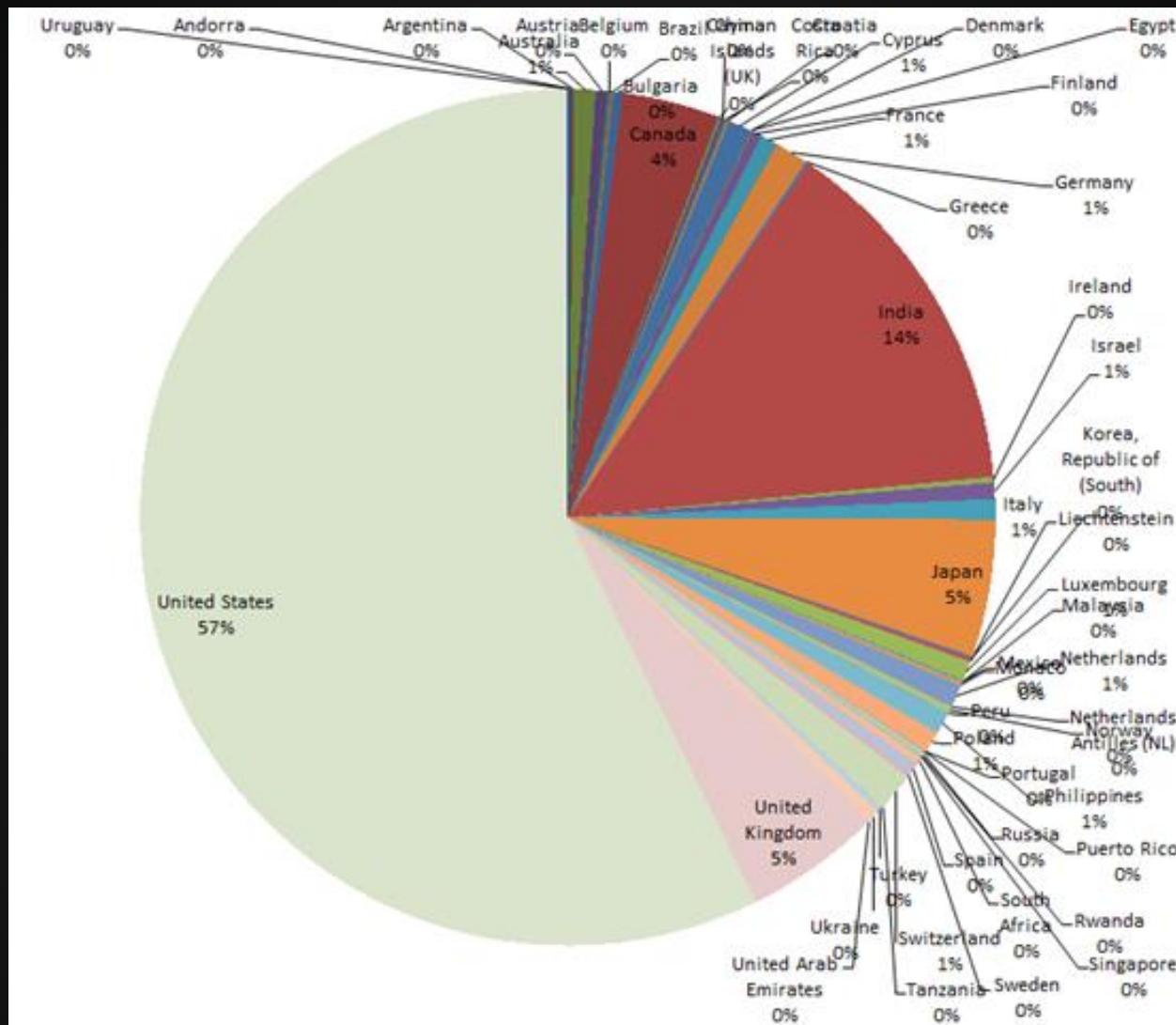
Be relatable. We want to know about that nervous breakdown. Or at least the time you didn't fit in at summer camp.

49%

Contrarian Thesis

Wait a sec—we should be playing more videogames? The more choices we have, the worse off we are? TED is where conventional wisdom goes to die.



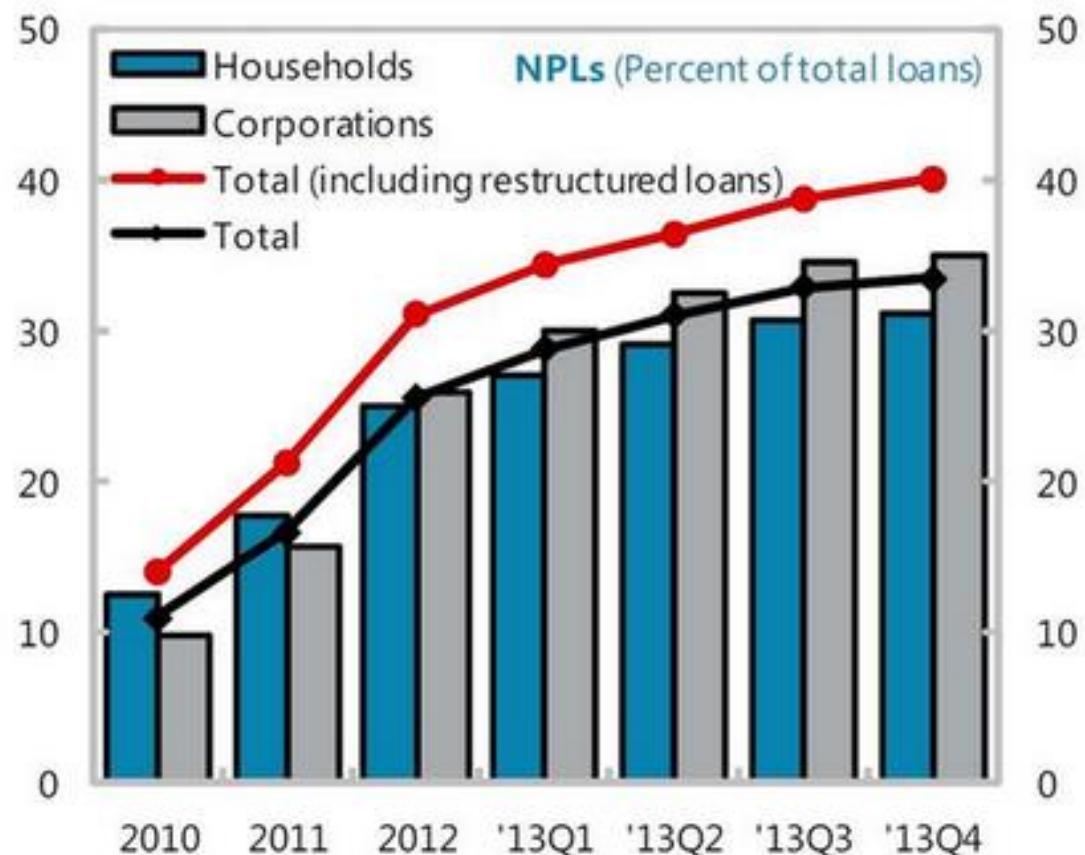


<http://www.infragistics.com/community/blogs/ux/archive/2012/10/24/the-road-to-hell-is-paved-with-data-visualization.aspx>



GreekFire23 @GreekFire23 · Jun 11

Spot the improving economy in Greece looking at their **non-performing loan** chart: pic.twitter.com/mYdjFTsRj9



Sources: Bank of Greece; and IMF staff calculations.

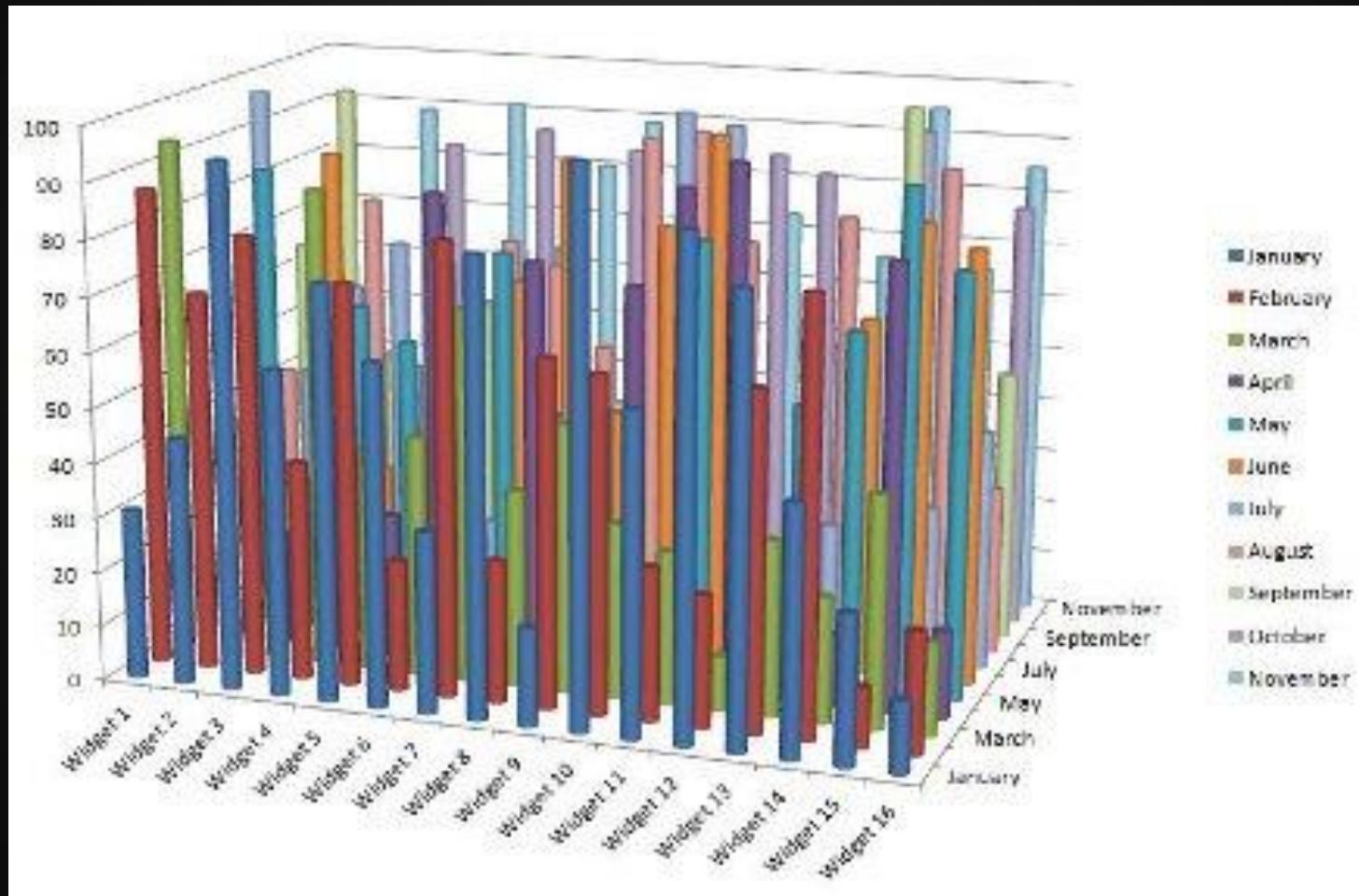
RETWEETS

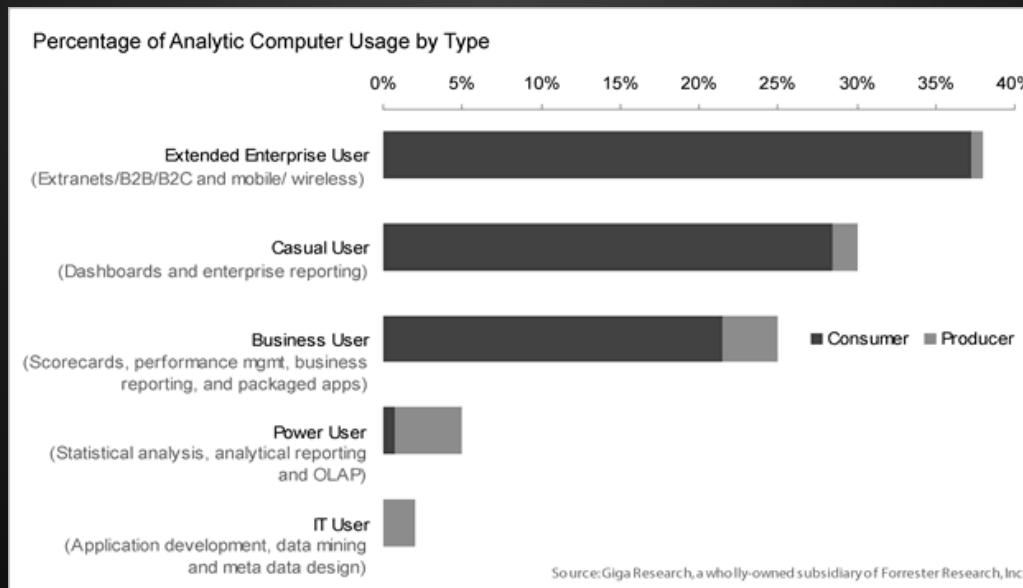
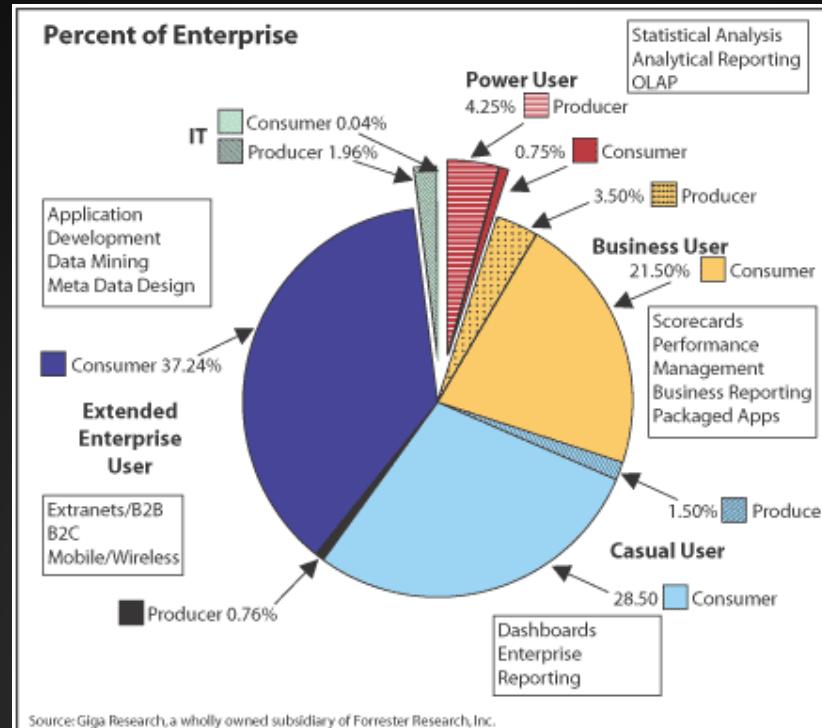
5



4:39 AM - 11 Jun 2014 · Details

Flag media





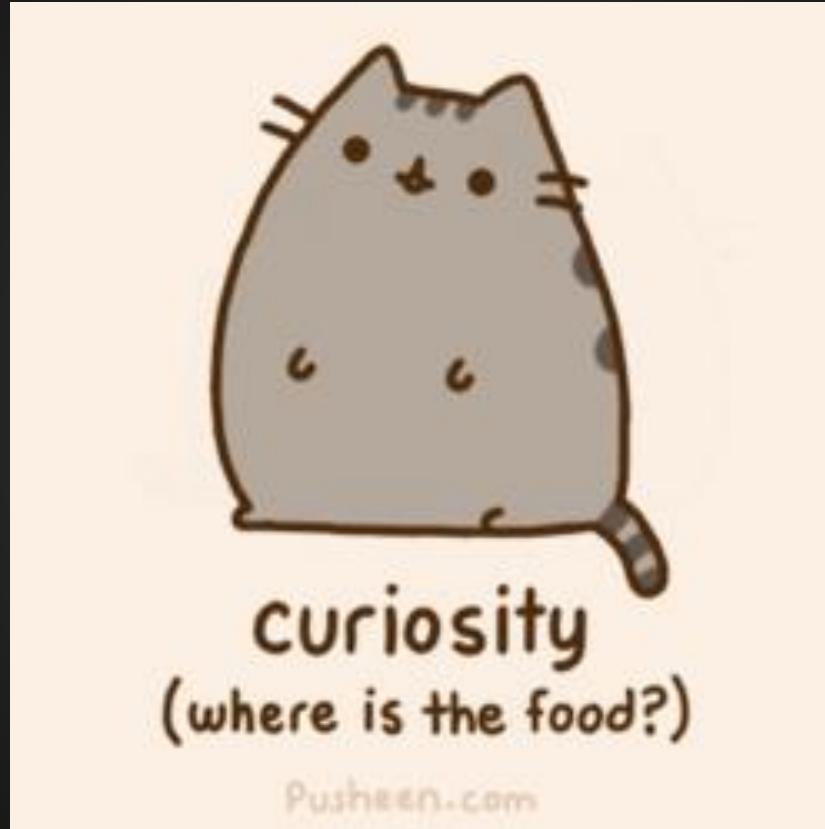
ЕЩЕ ПРИМЕРЫ

<http://viz.wtf/>

What quality problems have you noticed on graphs submitted to your journal? (by Editors of 25 Scientific Journals)

1. Lines too thin (or thick)
2. Symbols too small or difficult to distinguish
3. Lettering too small
4. Too many shades of gray on bars
5. Use of gray symbols or lines
6. Font difficult to read
7. Too many patterns on bars
8. Use of both different symbols and different lines
9. Redundant title printed on graph
10. Use of three-dimensional bars for only two variables
11. Key outside the graph

КАК ЖЕ ВИЗУАЛИЗИРОВАТЬ ПРАВИЛЬНО?



Graphical displays should (by Edward Tufte)

- show the data;
- induce the viewer to think about the substance rather than about methodology, graphic design, the technology of graphic production or something else;
- avoid distorting what the data have to say;
- present many numbers in a small space;
- make large data sets coherent;
- encourage the eye to compare different pieces of data;
- reveal the data at several levels of detail, from a broad overview to the fine structure;
- serve a reasonably clear purpose: description, exploration, tabulation or decoration;
- be closely integrated with the statistical and verbal descriptions of a data set.

Six fundamental principles of analytical design

(by Edward Tufte)

1. Show comparisons, contrasts, differences.
2. Show causality, mechanism, explanation, systematic structure.
3. Show multivariate data; that is, show more than 1 or 2 variables.
4. Completely integrate words, numbers, images, diagrams.
5. Thoroughly describe the evidence. Provide a detailed title, indicate the authors and sponsors, document the data sources, show complete measurement scales, point out relevant issues.
6. Analytical presentations ultimately stand or fall depending on the quality, relevance, and integrity of their content.

ВЫБОР ГРАФИКА ПО STEPHEN FEW

- **Time-series:** A single variable is captured over a period of time, such as the unemployment rate over a 10-year period. A line chart may be used to demonstrate the trend.
- **Ranking:** Categorical subdivisions are ranked in ascending or descending order, such as a ranking of sales performance (*the measure*) by sales persons (*the category*, with each sales person a *categorical subdivision*) during a single period. A bar chart may be used to show the comparison across the sales persons.
- **Part-to-whole:** Categorical subdivisions are measured as a ratio to the whole (i.e., a percentage out of 100%). A pie chart or bar chart can show the comparison of ratios, such as the market share represented by competitors in a market.
- **Deviation:** Categorical subdivisions are compared against a reference, such as a comparison of actual vs. budget expenses for several departments of a business for a given time period. A bar chart can show comparison of the actual versus the reference amount.
- **Frequency distribution:** Shows the number of observations of a particular variable for given intervals, such as the number of years in which the stock market return is between intervals such as 0-10%, 11-20%, etc. A histogram, a type of bar chart, may be used for this analysis.
- **Correlation:** Comparison between observations represented by two variables (X,Y) to determine if they tend to move in the same or opposite directions. For example, plotting unemployment (X) and inflation (Y) for a sample of months. A scatter plot is typically used for this message.
- **Nominal comparison:** Comparing categorical subdivisions in no particular order, such as the sales volume by product code. A bar chart may be used for this comparison.
- **Geographic or geospatial:** Comparison of a variable across a map or layout, such as the unemployment rate by state or the number of persons on the various floors of a building. A cartogram is a typical graphic used

		Value Encoding		Objects	Data
		Lines	Bars	Boxes	
Featured Subdivisions	1. Time	Value is measured over time (presently, monthly, etc.)	Value is ordered by size (increasing, decreasing)	Value is a part (percentage) of a whole and is related to other parts of total	Value is a whole (e.g., market share, total sales)
	2. Ranking	Value is a list of items where the rank order matters (e.g., top 10)	Value is ordered by size (increasing, decreasing)	Value is a part (percentage) of a whole and is related to other parts of total	Value is a whole (e.g., market share, total sales)
Part-to-Whole	3. Part-to-Whole	Value is a part (percentage) of a whole and is related to other parts of total	Value is a whole (e.g., market share, total sales)	Value is a part (percentage) of a whole and is related to other parts of total	Value is a whole (e.g., market share, total sales)
	4. Deviation	The difference between two values is more important than the absolute value of either and the direction of the difference	Value is a frequency distribution (e.g., histogram)	Value is a scatter plot, to determine if there is a correlation between two variables	Value is a pie chart, to determine if there is a deviation from the expected distribution
Comparison	5. Comparison	Comparison of two paired sets of data, such as extreme heights and weights of individuals to determine if there is a significant difference	Value is a histogram of two sets of data	Value is a scatter plot	Value is a scatter plot
	6. Nominal Comparison	Comparison of two or more categories for a set of variables, such as product models, or countries	Value is a bar chart of variables displayed on a map	Value is a scatter plot of variables displayed on a map	Value is a bar chart of variables displayed on a map
Geographic or Geospatial	7. Geographic	Value is measured across a geographic area, such as the unemployment rate by state	Value is a bar chart of variables displayed on a map	Value is a scatter plot of variables displayed on a map	Value is a bar chart of variables displayed on a map
	8. Geospatial	Value is measured across a layout, such as the number of persons on the various floors of a building	Value is a scatter plot of variables displayed on a map	Value is a scatter plot of variables displayed on a map	Value is a scatter plot of variables displayed on a map

Derived from the book
Show Me the Numbers
© Stephen Few 2004-2014

http://www.perceptualedge.com/articles/misc/Graph_Selection_Matrix.pdf

ЦЕЛИ ВИЗУАЛИЗАЦИИ И ТИПЫ ВИЗУАЛИЗАЦИИ

1. Сравнение данных

- Bar chart
- Line chart
- Time series chart
- Area chart

3. Распределение данных

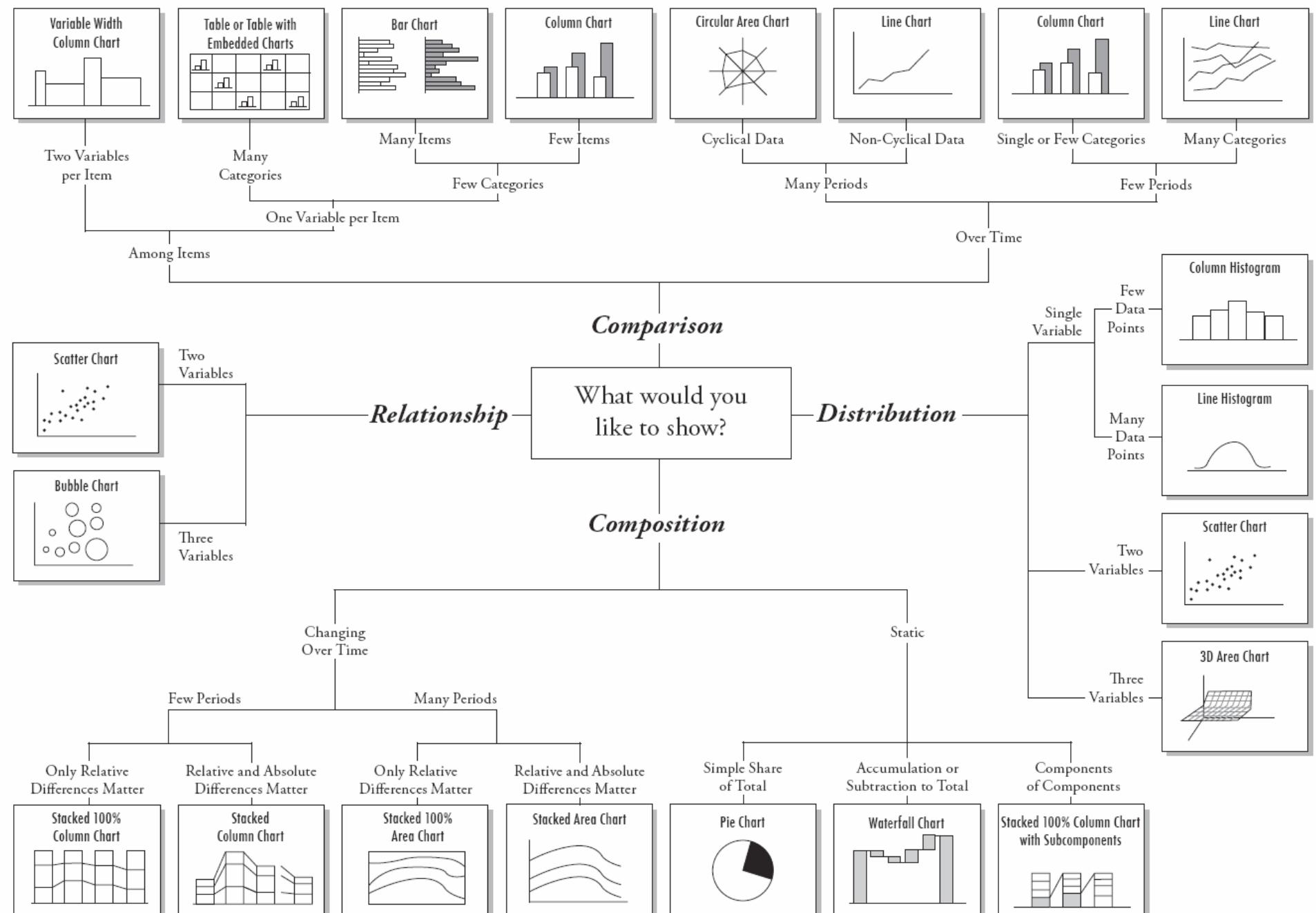
- Bar chart, Histogram
- Line chart, density
- Scatter plot chart

2. Отношения между данными

- Scatter plot
- Circular network
- Network
- Map

4. Композиция данных

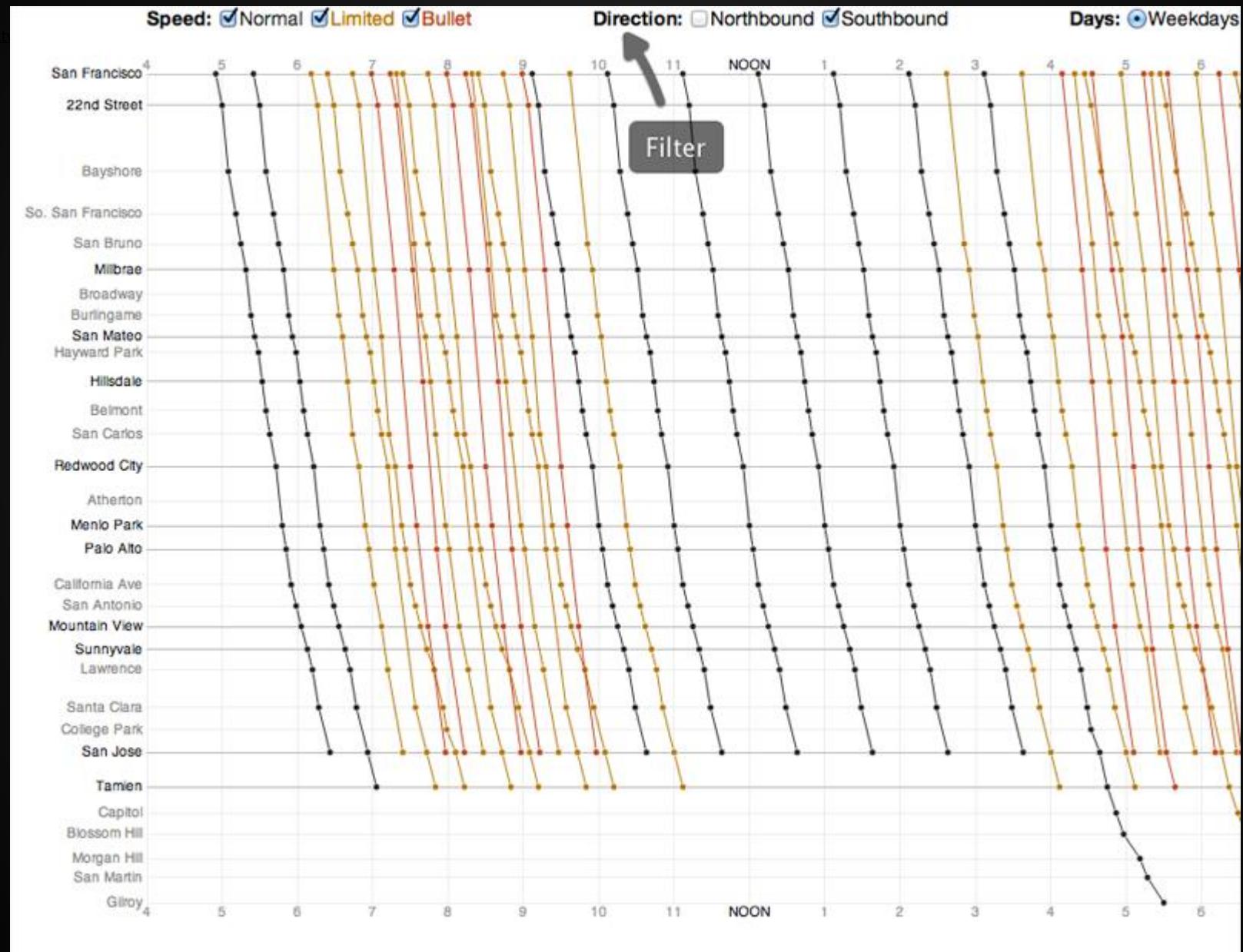
- Tag cloud
- Pie chart
- Percent Stacked bar chart



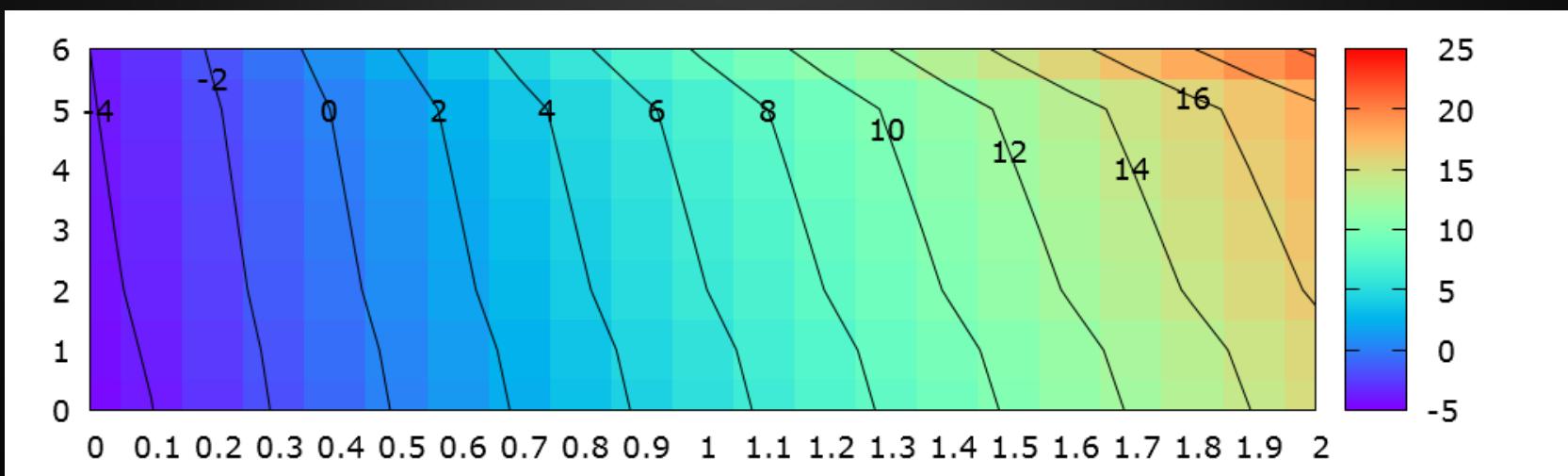
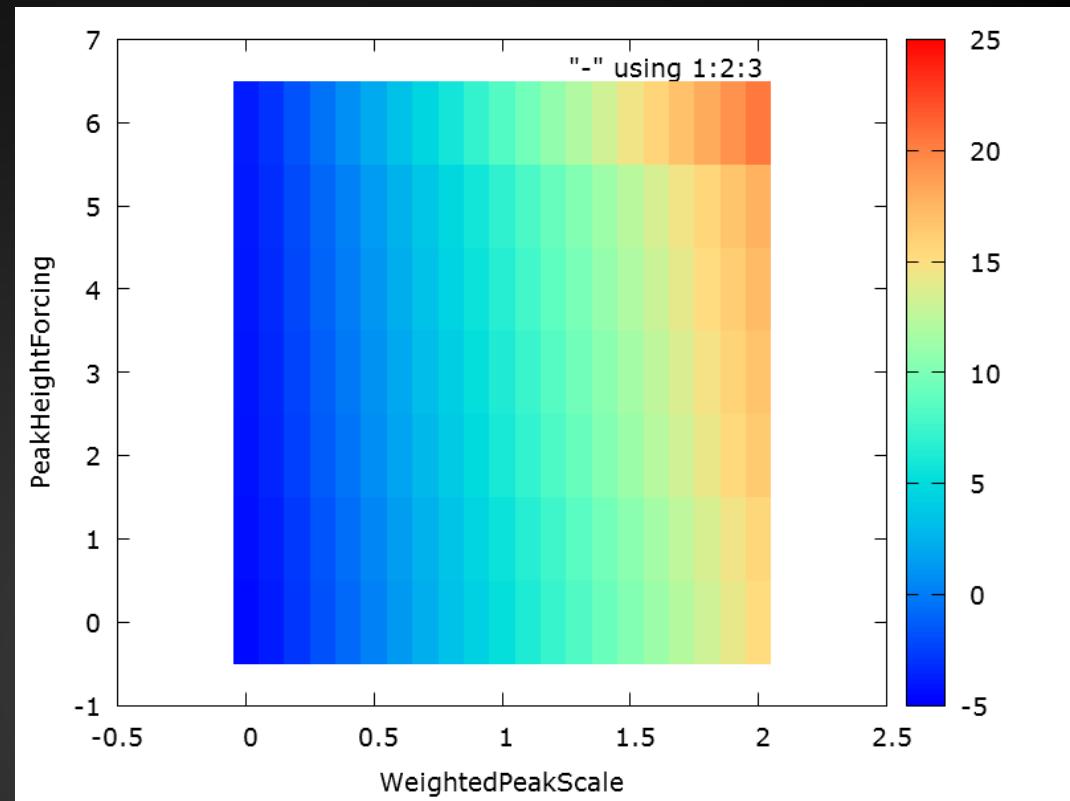
ВЫБОР ТИПА ВИЗУАЛИЗАЦИИ

Цель визуализации/ Тип данных	Отношения в данных	Распределение данных	Сравнение данных	Композиция данных
Непрерывные числовые	line area scatter bubble	scatter bubble	line area radar	stacked line full stacked line stacked area full stacked area
Непрерывные временные	line area radar scatter bubble		time line gantt waterfall radar	gantt stacked line full stacked line stacked area full stacked area
Дискретные		bar scatter bubble	bar pie doughnut	pie doughnut stacked bar full stacked bar
Географические	map line area	map scatter	map bar	map stacked bar full stacked bar
Логические		tree mental tree		tree map

TIME SERIES CHART

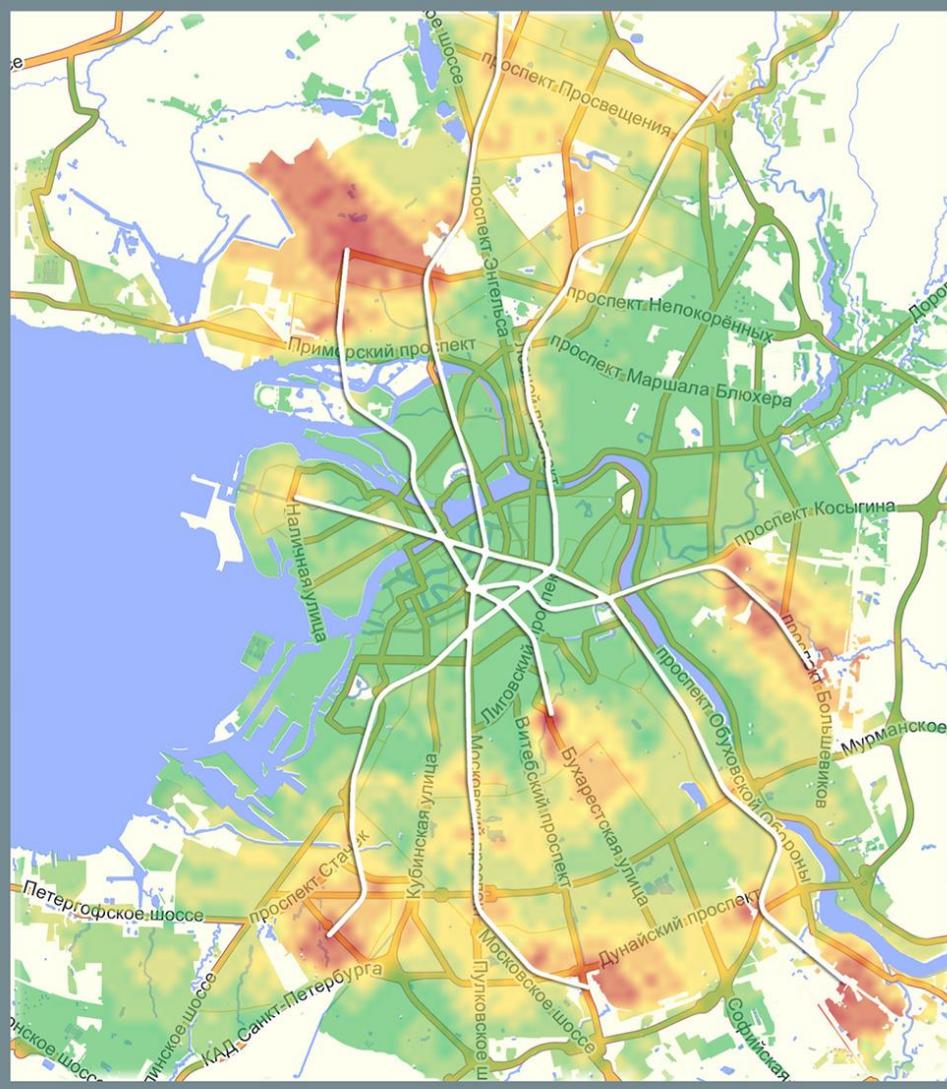


HEATMAP



9-10 ЧАСОВ УТРА, В ЦЕНТР

HEATMAP



На автомобиле быстрее,
чем на общественном транспорте

Однако

Общественным транспортом
быстрее, чем на автомобиле

на 30%

на 20%

на 10%

на 10%

на 20%

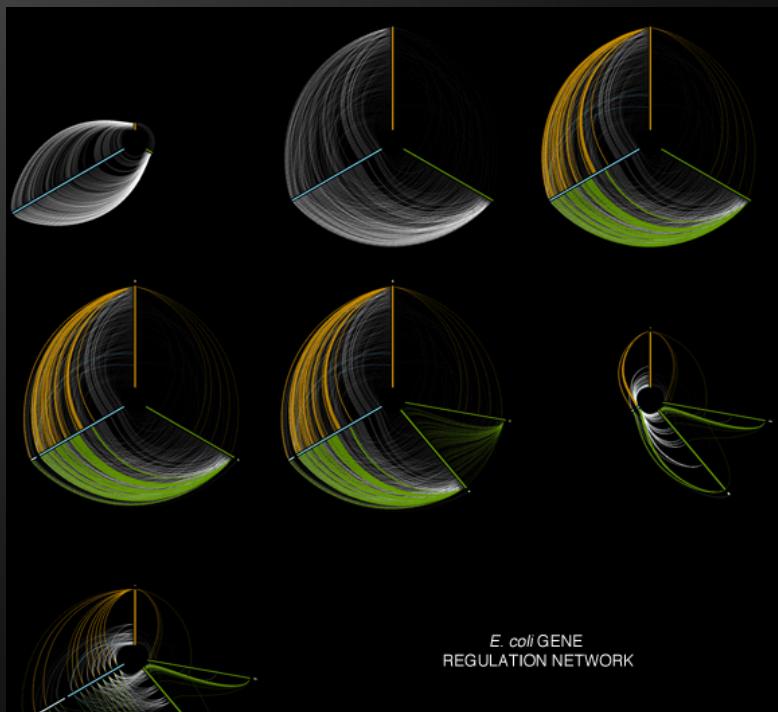
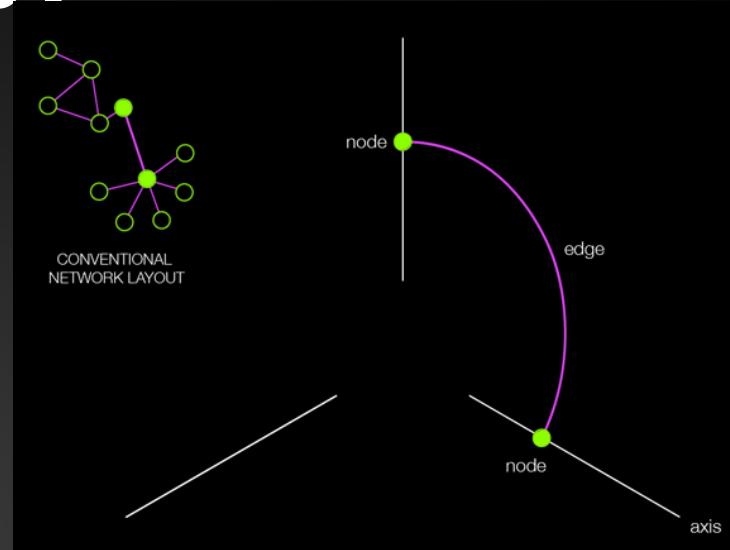
на 30%

TAG CLOUD



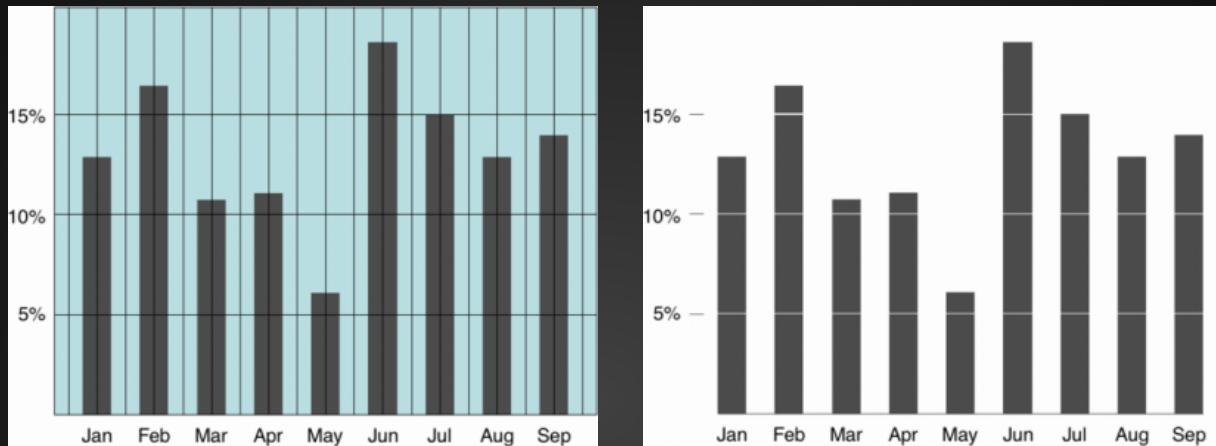
HIVE PLOT

- <http://www.hiveplot.net/>



DATA-INK RATIO

Соотношение данных и чернил



$$\text{Data-ink ratio} = \frac{\text{Data-ink}}{\text{Total ink used to print the graphic}}$$

= proportion of a graphic's ink devoted to the non-redundant display of data-information

= $1.0 - \text{proportion of a graphic that can be erased}$

Remove to improve (the **data-ink** ratio)

Created by Darkhorse Analytics

www.darkhorseanalytics.com

<https://speakerdeck.com/cherdarchuk/remove-to-improve-the-data-ink-ratio>

Remove to improve the **pie chart** edition

Created by Darkhorse Analytics

www.darkhorseanalytics.com

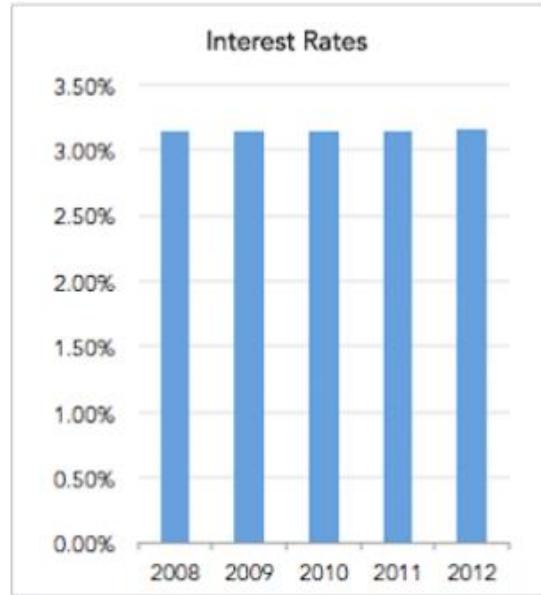
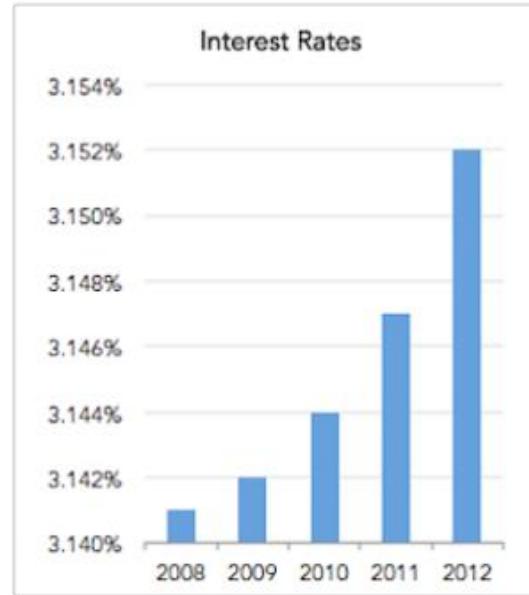
<https://speakerdeck.com/cherdarchuk/data-looks-better-naked-pie-chart-edition>

Remove
to improve
the **data tables** edition

LIE FACTOR

$$\text{Lie Factor} = \frac{\text{size of effect shown in graphic}}{\text{size of effect in data}}$$
$$\text{size of effect} = \frac{|\text{second value} - \text{first value}|}{\text{first value}}$$

Same Data, Different Y-Axis



Инструменты



Выбор цветовой палитры

Color Scheme Designer is now Paletton! Check out Paletton.com

Color Scheme Designer → [paletton](http://paletton.com)

2002-2010 © Petr Stanicek • v3.51 • [Blog & News](#)

Undo Redo Random Colorblind Color space Export Help

CREATE FREE STUNNING WEBSITES GO >

mono complement triad tetrad analogic accented analogic

Hue: 37° Angle: 37°

warm cold

R: 100 % G: 51 % B: 0 %

RGB: FF8300

Preview Light page example Dark page example

Show sample text

Do you like this app?
[Donate](#)

The screenshot shows the Color Scheme Designer application interface. At the top, there's a navigation bar with links for Undo, Redo, Random, Colorblind, Color space, Export, and Help. Below the navigation is a color palette strip with four squares: orange, blue, green, and brown. To the right of the strip is a button that says "CREATE FREE STUNNING WEBSITES" and "GO >". Below the palette strip is a text field with the Scheme ID: 0A31-w0w0w0w0. On the left side, there's a large circular color wheel with a central orange circle. The wheel is divided into warm colors (top half) and cold colors (bottom half). A specific color point is highlighted with a dot and labeled with its RGB values: R: 100%, G: 51%, B: 0%. The color's hex code, FF8300, is also displayed. Below the color wheel are tabs for "Hues", "Adjust Scheme", and "Color List". On the right side, there are three preview sections: "Preview", "Light page example", and "Dark page example". Below these is a "Show sample text" button. At the bottom left, there's a question "Do you like this app?" followed by a "Donate" button.

Выбор цветовой палитры

Number of data classes: 3 i

Nature of your data: i

sequential diverging qualitative

Pick a color scheme:

Multi-hue: i Single hue: i

Only show: i

colorblind safe
 print friendly
 photocopy safe

Context: i

roads
 cities
 borders

Background: i

solid color
 terrain

color transparency

how to use | updates | downloads | credits

COLORBREWER 2.0
color advice for cartography

3-class YIOrBr

EXPORT

HEX

#ffff7bc
#fec44f
#d95f0e

© Cynthia Brewer, Mark Harrower and The Pennsylvania State University
Support
[Back to Flash version](#)
[Back to ColorBrewer 1.0](#)

axismaps

PYTHON MATPLOTLIB



- Множество типов
графиков и богатая
библиотека примеров

<http://matplotlib.org/>

GNUPLOT

Плюсы

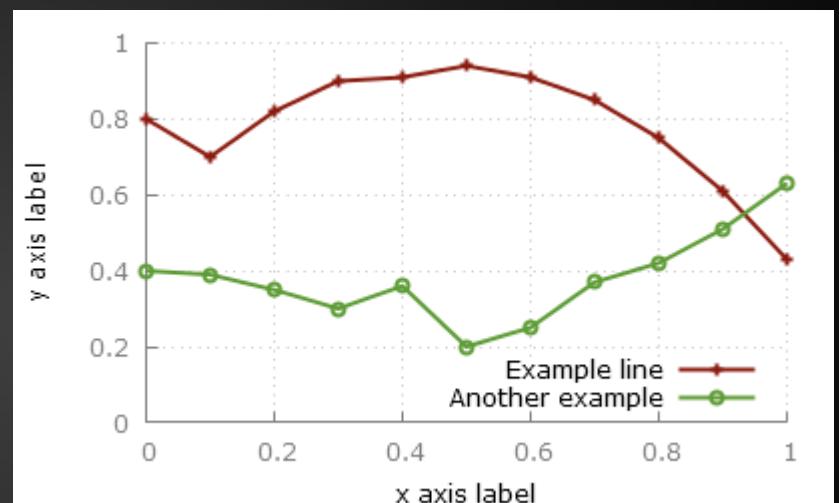
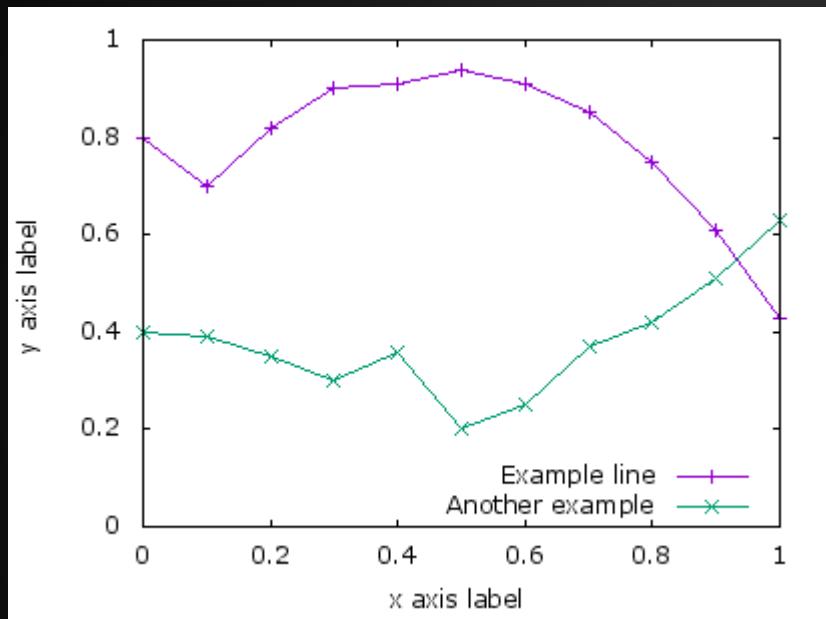
- DSL
- Несложный синтаксис, нюансы быстро гуглятся
- Концептуальное отличие от библиотек GPPL – регулирует только уровень представления
- Нативная обработка дат

Минусы

- Потенциальная ограниченность синтаксиса
- Отсутствует GIS
- Слабая поддержка статистической обработки

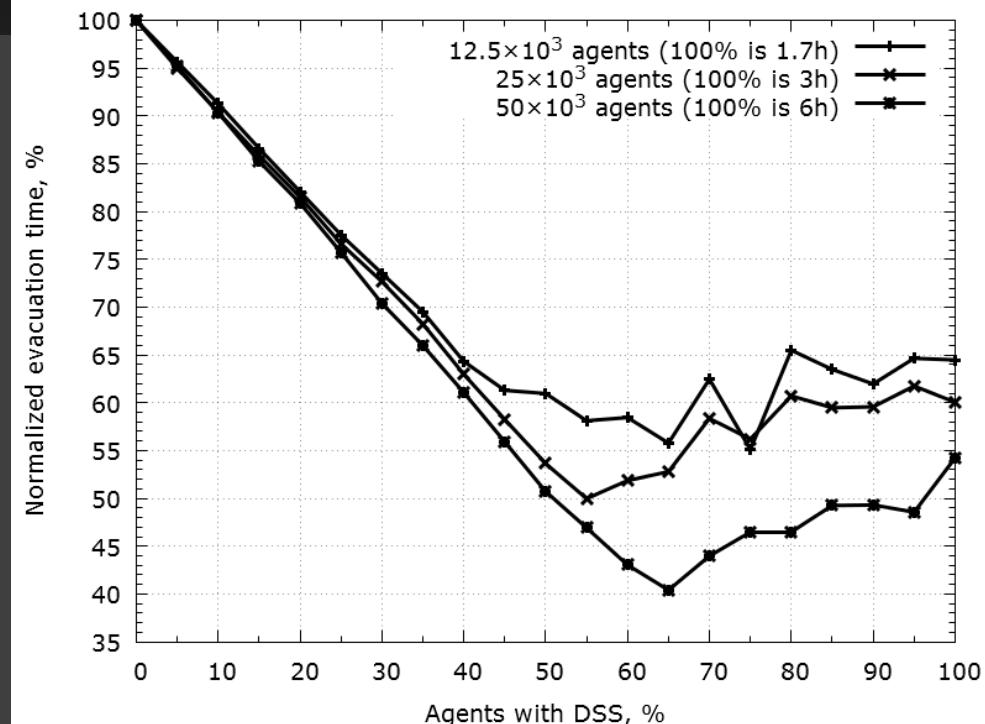
<http://www.gnuplot.info/>

GNUPLOT SINCE 1986 :)



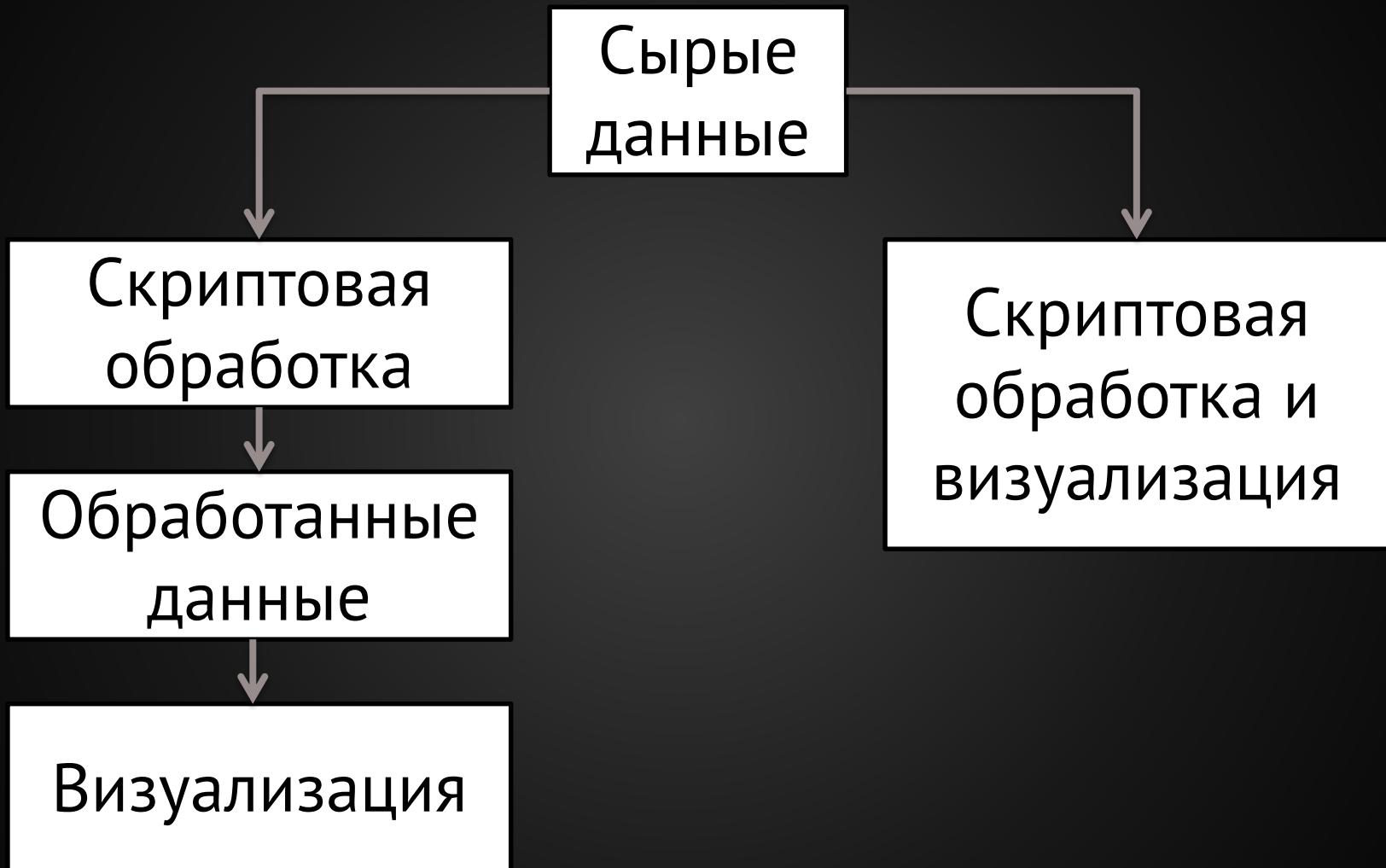
Анатомия скрипта GNUplot

```
1 set terminal pngcairo dashed size 800,600 \
2   enhanced monochrome font 'Verdana,14'
3
4 set autoscale
5 set xtic auto
6 set ytic auto
7
8 set ytic 35,5,100
9 set xtic 0,10,100
10
11 set yrange [35:100]
12 set xrange [0:100]
13
14 set mxtics 2
15 set mytics 5
16
17 set xlabel "Agents with DSS, %"
18 set ylabel "Normalized evacuation time, %"
19
20 set grid ytics lt 0 lw 1 lc rgb "#bbbbbb"
21 set grid xtics lt 0 lw 1 lc rgb "#bbbbbb"
22
23 set datafile separator ";"
24
25 set output 'xy.png'
26 plot "xy.12.csv" using 1:(100.0 * $2/101.5) with linespoints lw 3 lt 2 lc "black" \
27   pt 1 ps 1.5 title "12.5×103 agents (100% is 1.7h)", \
28   "xy.25.csv" using 1:(100.0 * $2/180.5) with linespoints lw 3 lt 1 lc "black" \
29   pt 2 ps 1.5 title "25×103 agents (100% is 3h)", \
30   "xy.50.csv" using 1:(100.0 * $2/362.9) with linespoints lw 3 lt 3 lc "black" \
31   pt 3 ps 1.5 title "50×103 agents (100% is 6h)"
32
33 end
```

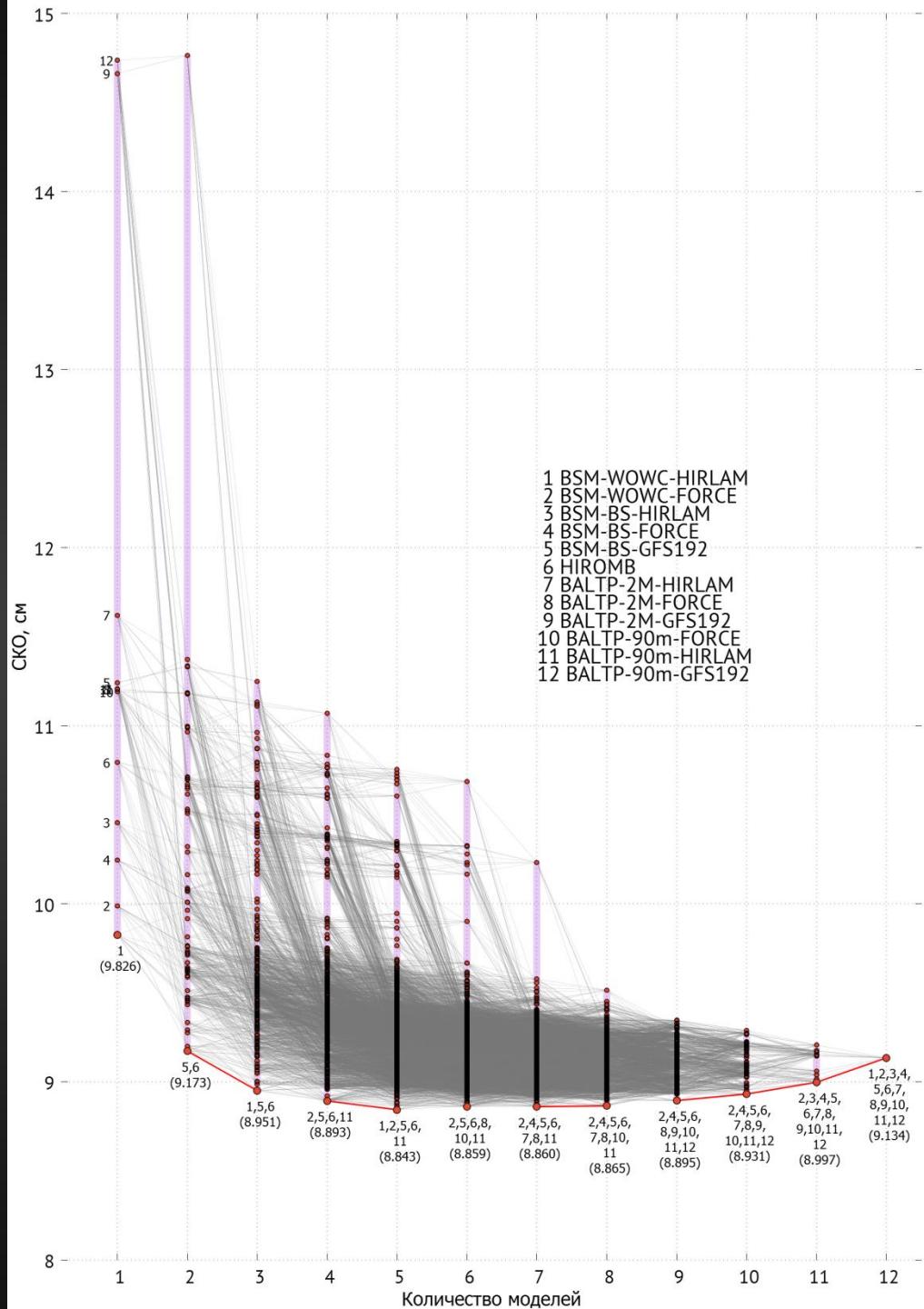
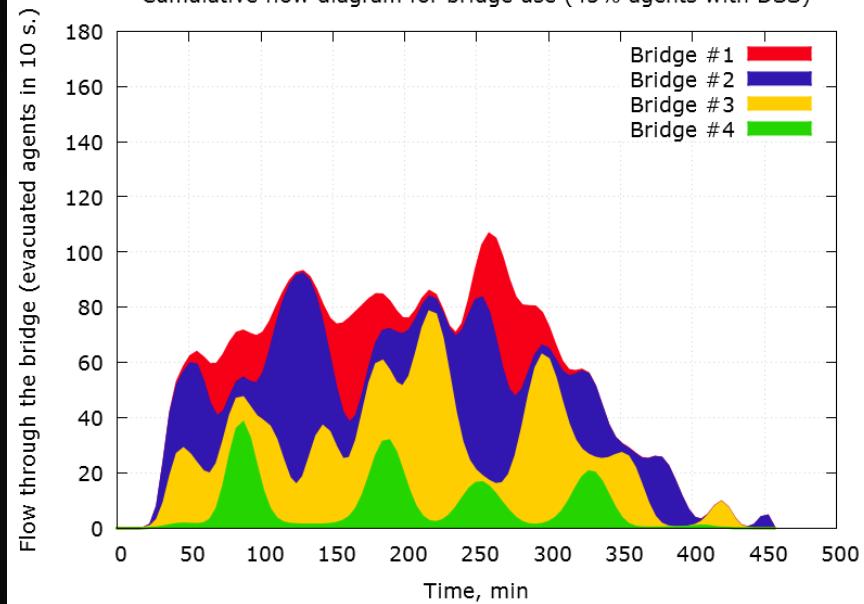


- Терминал, шрифт
- Оси
- Сетка
- Данные как из файла так и из скрипта
- Выходной файл
- Построение графиков
- End <!>

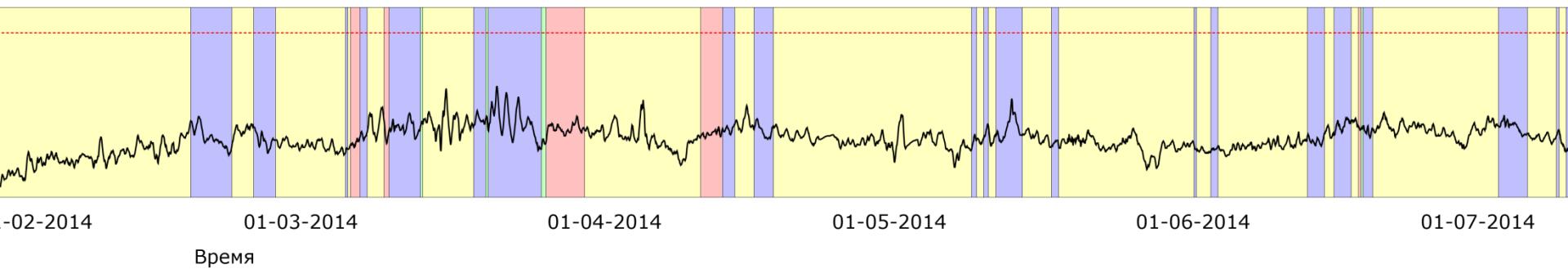
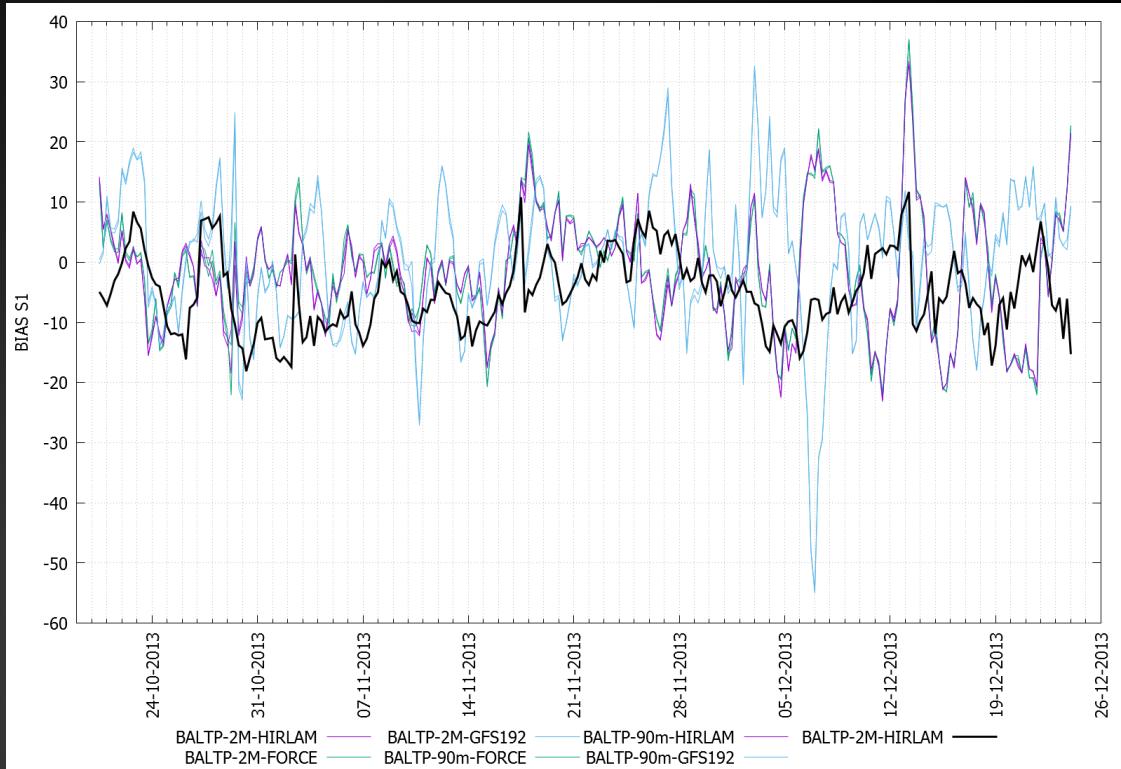
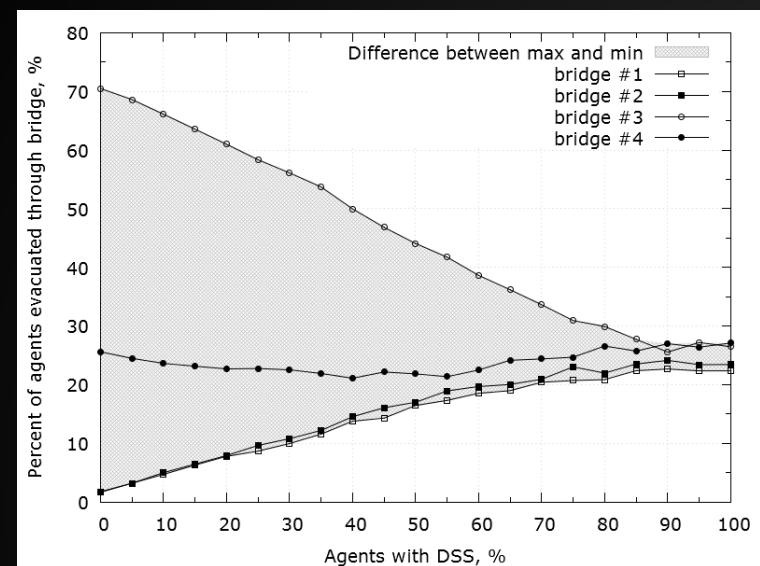
VISUALIZATION WORKFLOW



GNUPLOT



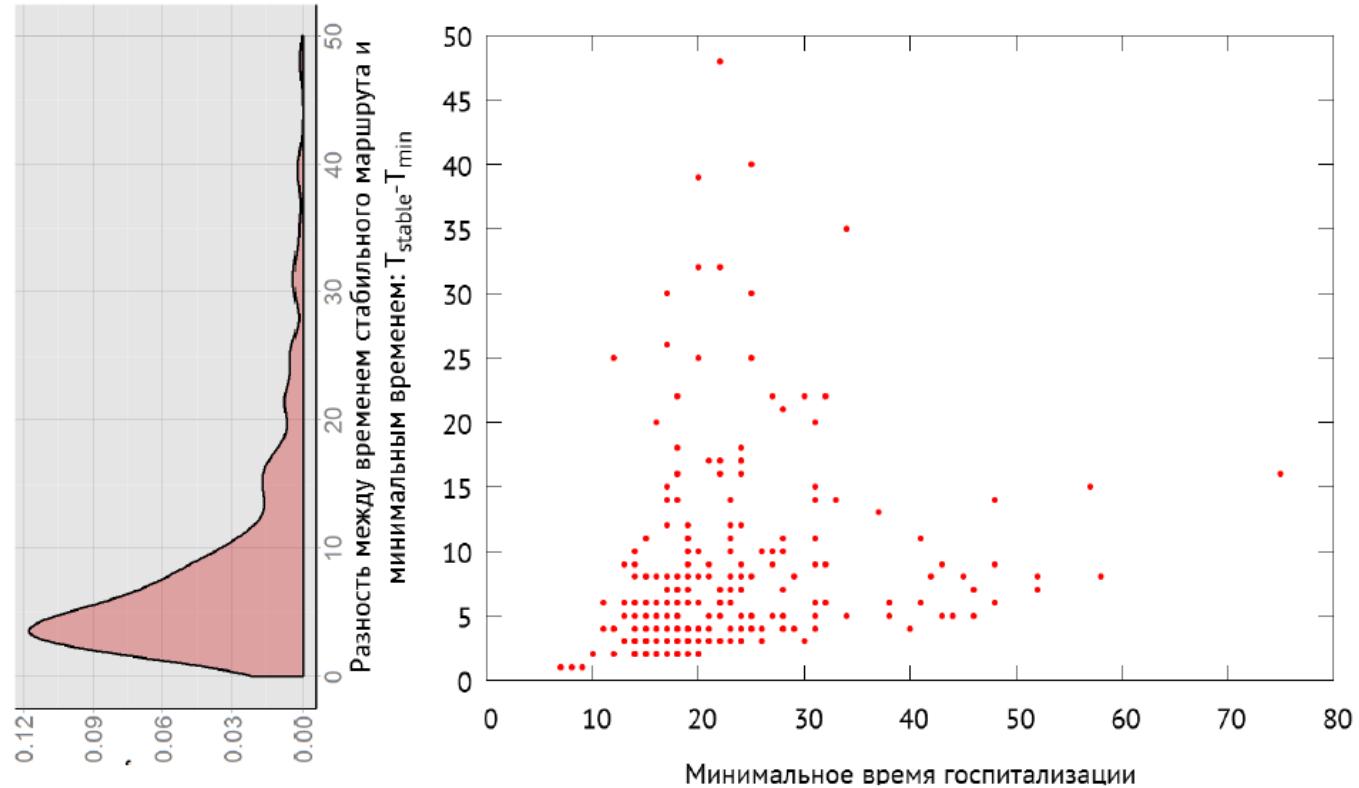
GNUPLOT



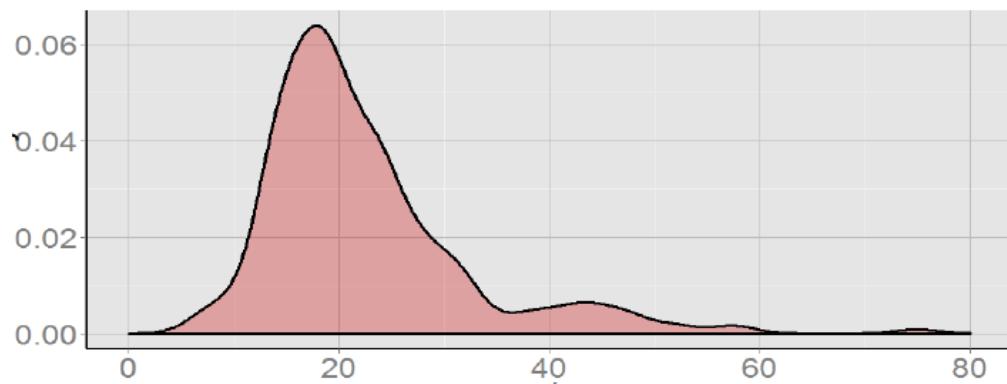
a_0
 $a_0 + a_1 * B$

$a_0 + a_2 * H$
 $a_0 + a_1 * B + a_2 * H$

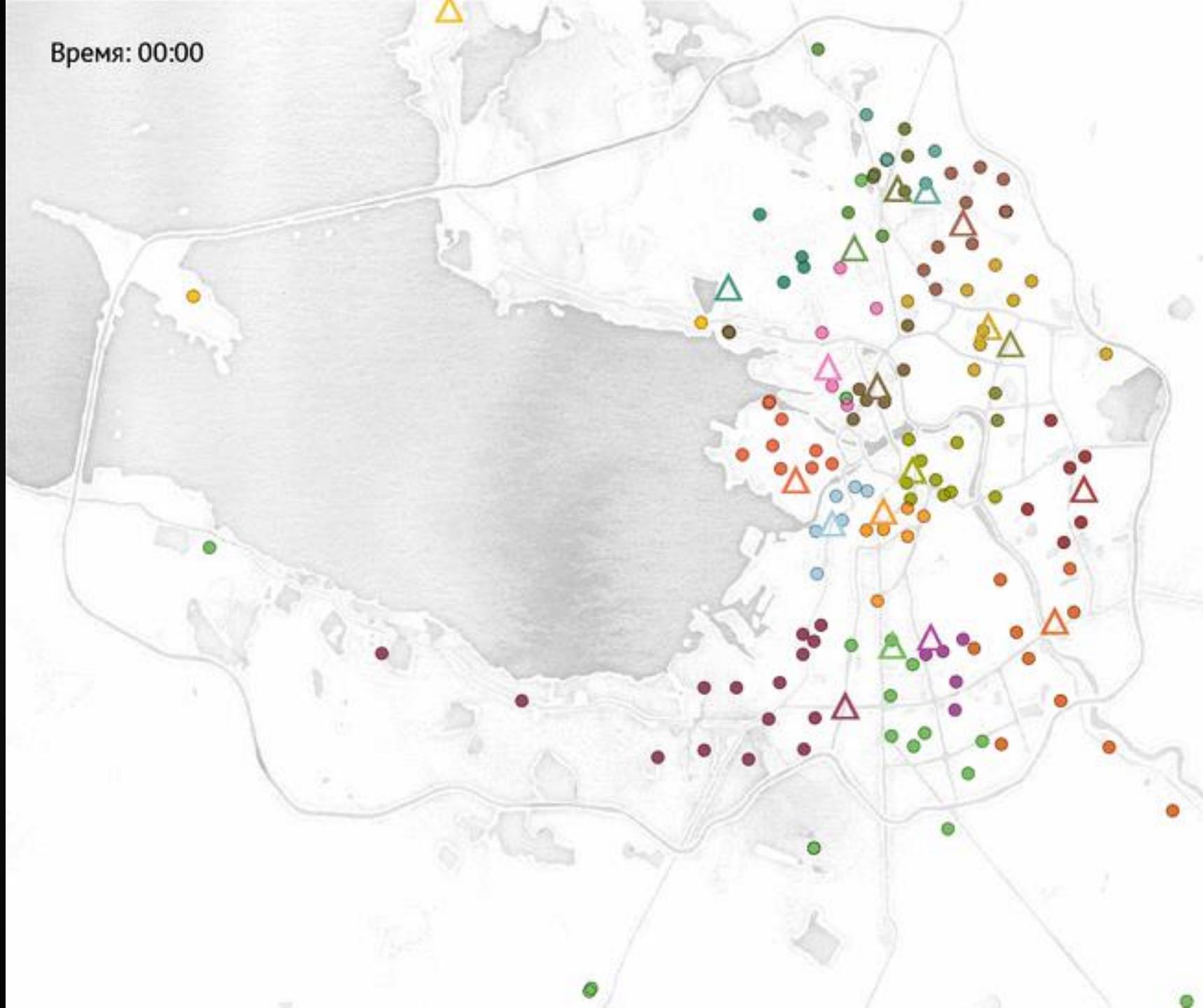
Случаи смены стабильного маршрута

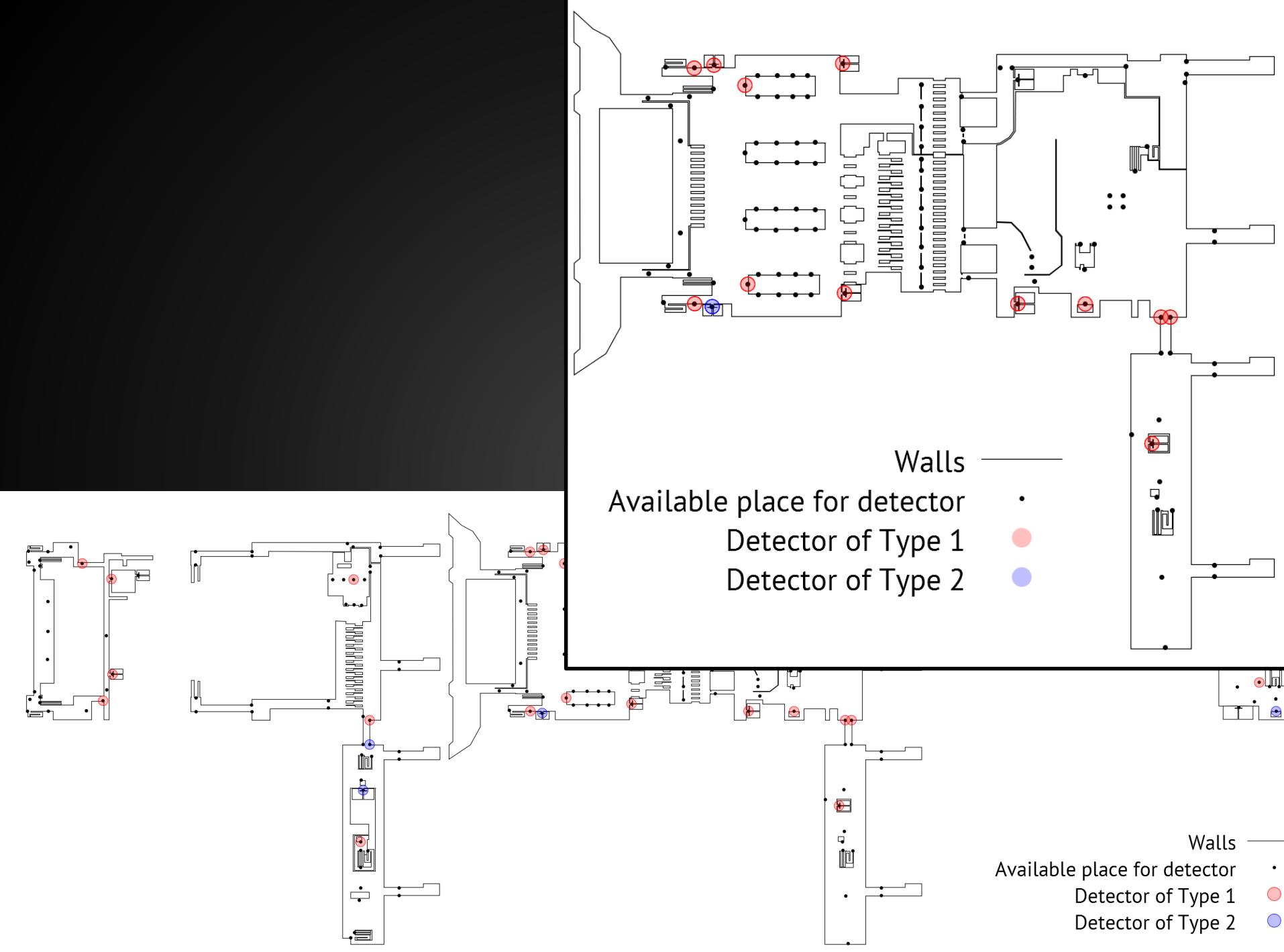


Минимальное время госпитализации



Время: 00:00



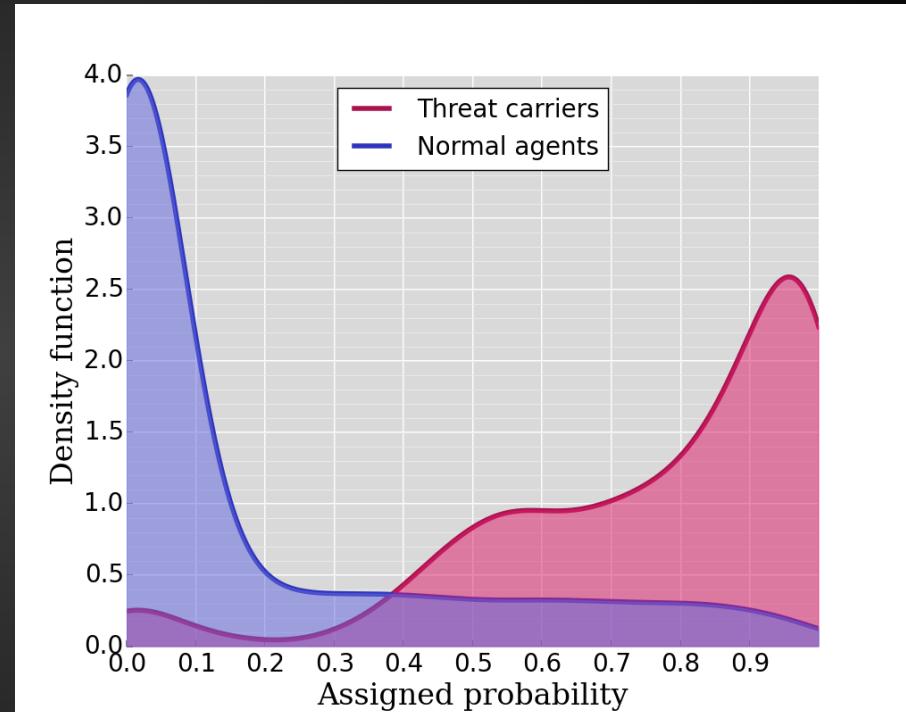
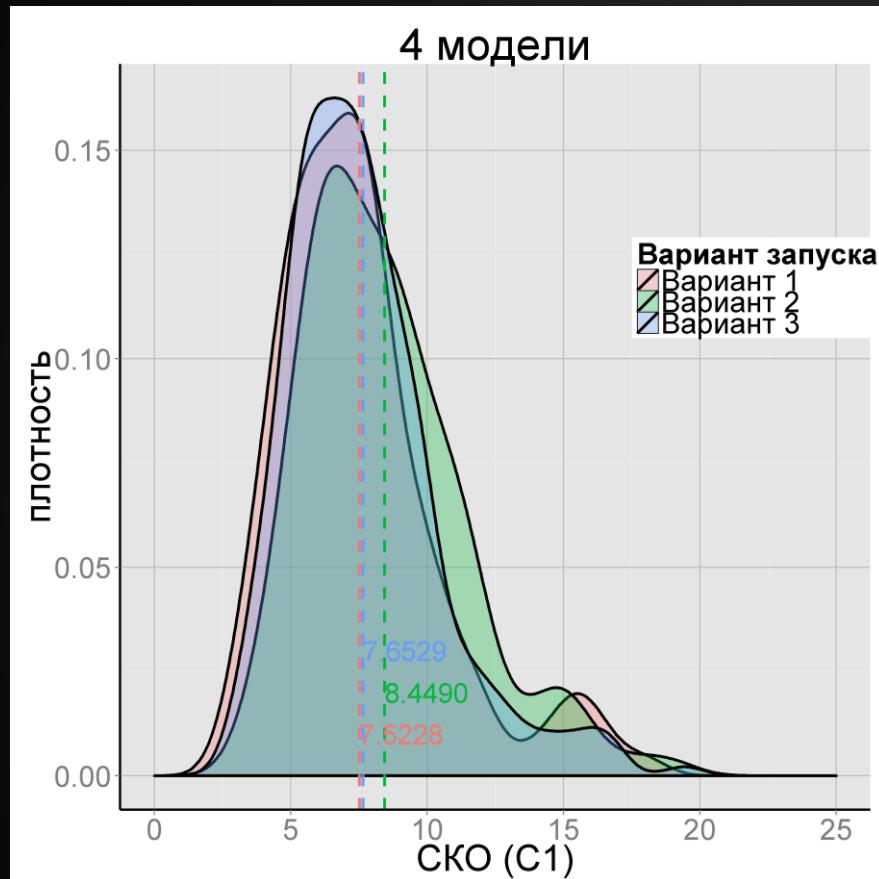


R



- Популярный открытый пакет статистической обработки данных
- Построение графиков и другой визуализации встроенными функциями и с помощью сторонних библиотек:
 - Ggplot2

GGPLOT (R)



RSTUDIO SHINY (R)

Shiny by RStudio

Search

OVERVIEW

TUTORIAL

ARTICLES

GALLERY

REFERENCE

DEPLOY

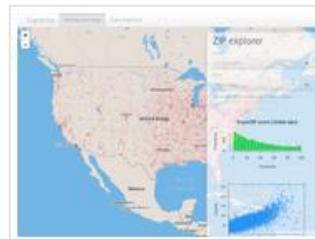
HELP

Gallery

This gallery contains useful examples to learn from. Visit the [Shiny User Showcase](#) to see an inspiring set of sophisticated apps.

Interactive visualizations

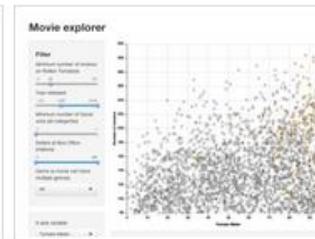
Shiny is designed for fully interactive visualization, using JavaScript libraries like [d3](#), [Leaflet](#), and [Google Charts](#).



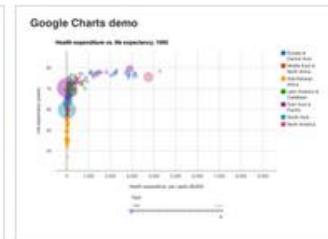
SuperZip example



Bus dashboard



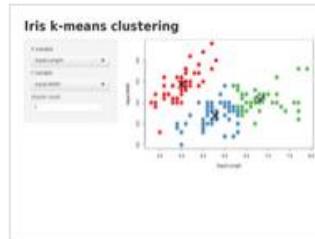
Movie explorer



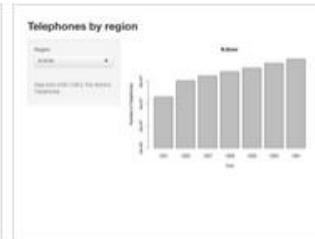
Google Charts

Start simple

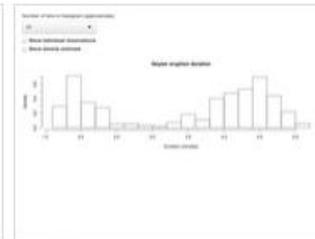
If you're new to Shiny, these simple but complete applications are designed for you to study.



Kmeans example



Telephones by region



Faithful



Word cloud

<http://shiny.rstudio.com/gallery/>

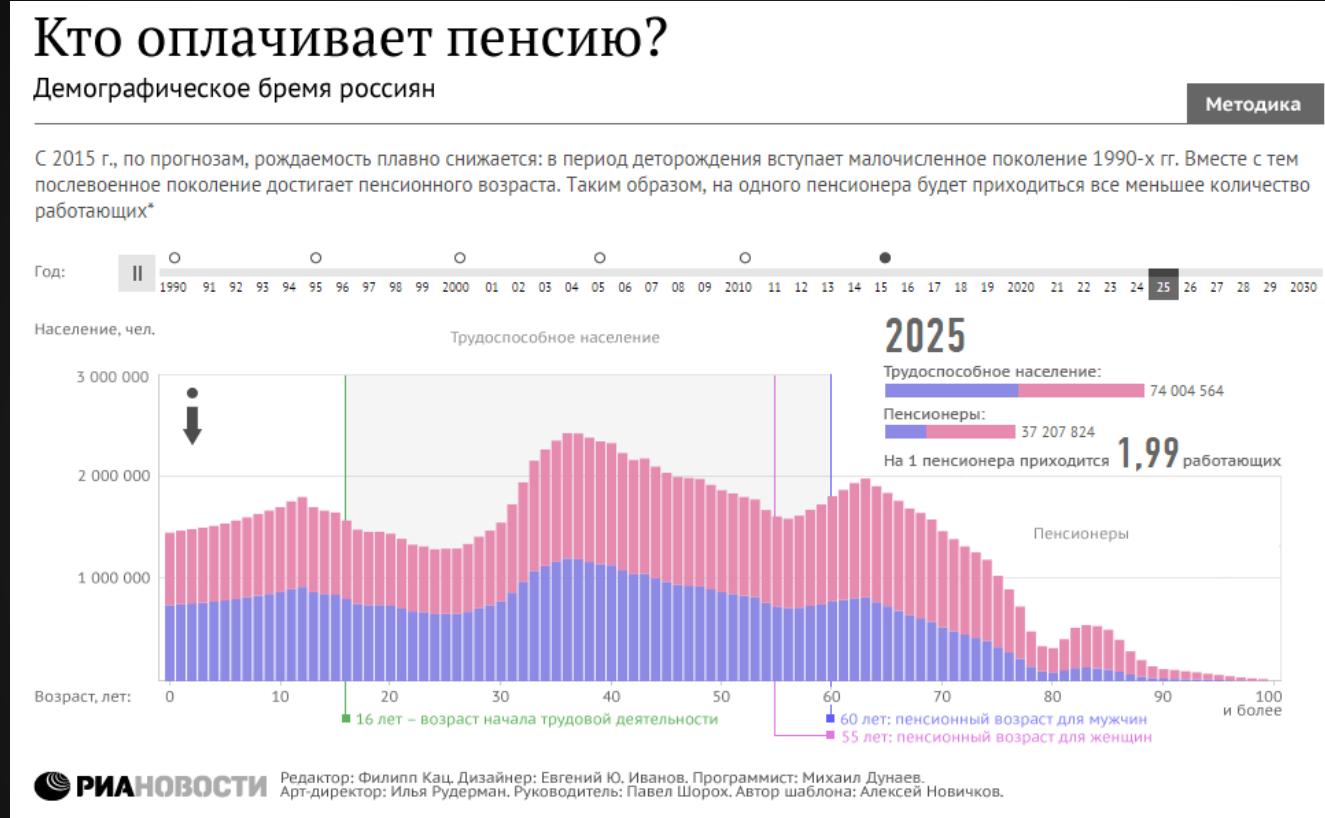
DB Data-Driven Documents



D3.js is a JavaScript library for manipulating documents based on data. D3 helps you bring data to life using HTML, SVG, and CSS. D3's emphasis on web standards gives you the full capabilities of modern browsers without tying yourself to a proprietary framework, combining powerful visualization components and a data-driven approach to DOM manipulation.

[See more examples.](#)

ПРИМЕРЫ (D3JS)



<http://ria.ru/infografika/20131108/975609928.html>

<http://blocks.org/mbostock>

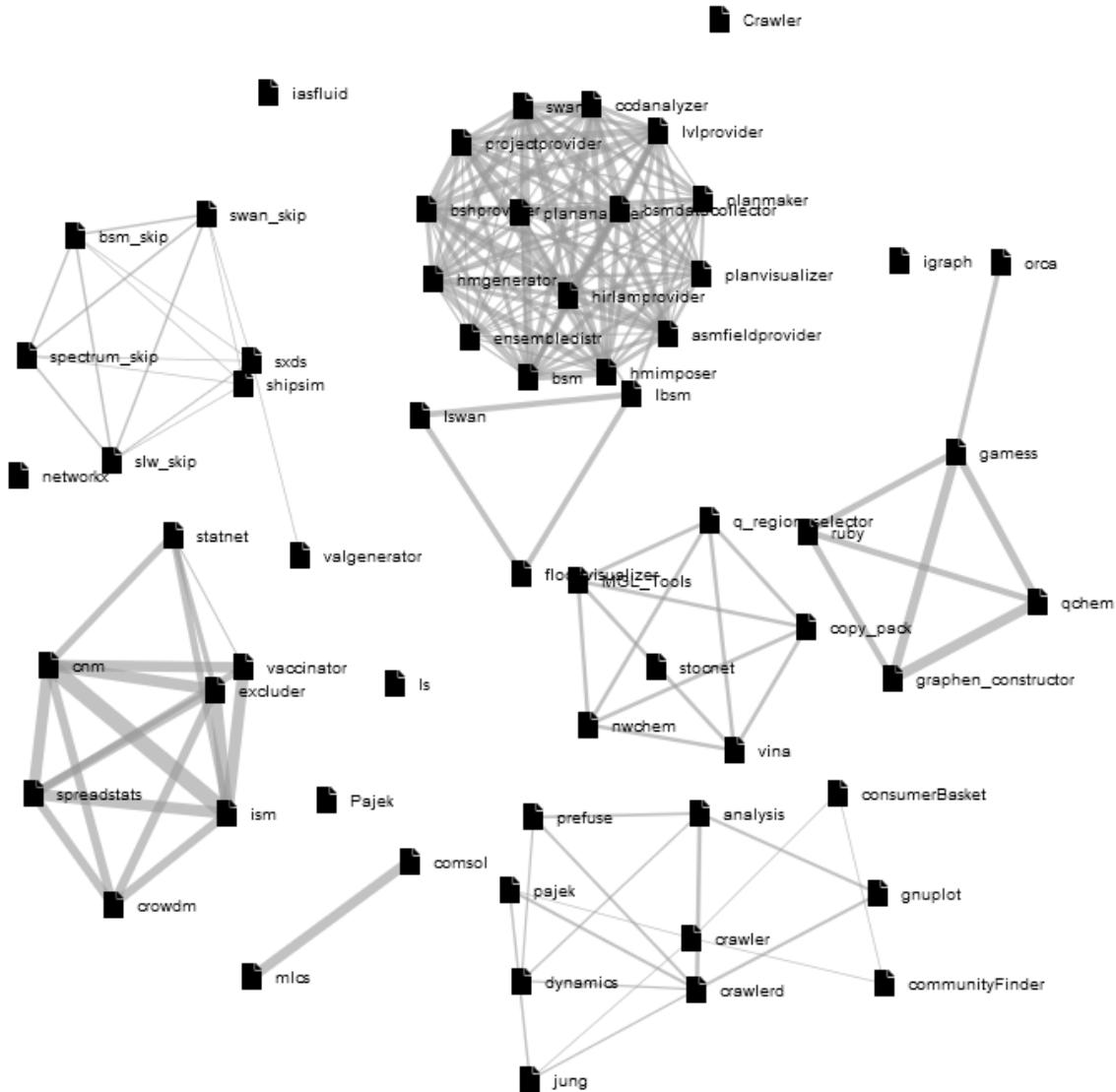
<https://github.com/mbostock/d3/wiki/Gallery>

D3JS

ПРИМЕР

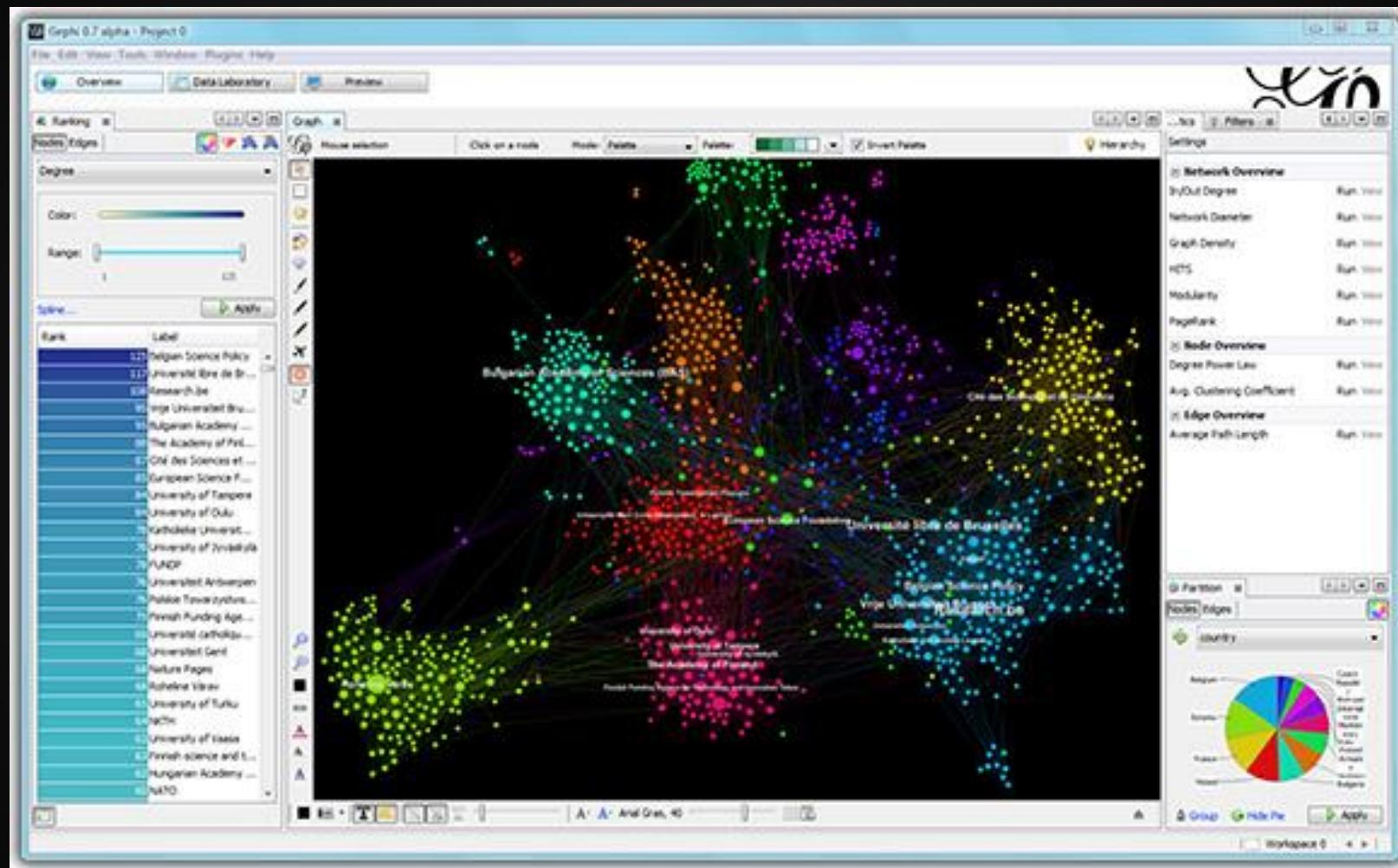
Анализ связей между пакетами

Анализируется совместное присутствие пакетов в композитных приложениях.



GEPHI

<https://gephi.github.io/>



<http://gephi.github.io/>

Источники и ссылки

- **Книги**
 - The Visual Display of Quantitative Information (Edward R. Tufte)
<http://www.amazon.com/The-Visual-Display-Quantitative-Information/dp/0961392142>
 - Semiology of Graphics: Diagrams, Networks, Maps (Jacques Bertin)
<http://www.amazon.com/Semiology-Graphics-Diagrams-Networks-Maps/dp/0299090604>
- **Инструменты**
 - Список рекомендуемых инструментов <http://selection.datavisualization.ch/>
 - Огромный каталог инструментов <http://www.visualisingdata.com/index.php/resources/>
- **Типы диаграмм**
 - Каталог типов диаграмм <http://www.datavizcatalogue.com/>
 - Типы визуализации
http://www.visual-literacy.org/periodic_table/periodic_table.html#
- **Разбор ошибок**
 - <http://cs.colby.edu/courses/S15/cs251/goodbad.php>
 - How to Lie with Data Visualization (Ravi Parikh)
<http://data.heapanalytics.com/how-to-lie-with-data-visualization/>
 - Неплохой разбор ошибок <http://www.perceptualedge.com/examples.php>
 - Каталог плохих визуализаций <http://flowingdata.com/category/visualization/ugly-visualization/>
- **Примеры визуализации**
 - Сайт о визуализации данных, содержит множество интересных статей о визуализации и инструментах
<http://datavisualization.ch/>
 - Примеры интерактивной визуализации на основе данных Twitter
<https://interactive.twitter.com/>
 - Примеры простой визуализации на R, Python, Gnuplot
<https://github.com/ze0n/scientific-code-samples>
- **Другие презентации по теме**
 - <http://www.slideshare.net/idigdata/data-visualization-best-practices-2013>
 - http://www.slideshare.net/SeanKandel/20131024-big-datavisualization?qid=8a3e771f-730c-4898-90cd-e0a9d1c2d84e&v=default&b=&from_search=34
 - http://www.slideshare.net/kristw/data-visualization-a-quick-tour-for-data-science-enthusiasts?qid=89a6d03a-9f6e-4c3e-afe8-c379ddbaf1bc&v=qf1&b=&from_search=1