

Данные и чувство прекрасного



Аня Малькова
ЭФ МГУ'17
IBS DHR

План

- Почему визуализировать полезно
- Визуализация и хорошие манеры
- Основные типы визуализаций
- Как это выглядит на практике

Почему визуализировать полезно

756395068473

658663037576

860372658602

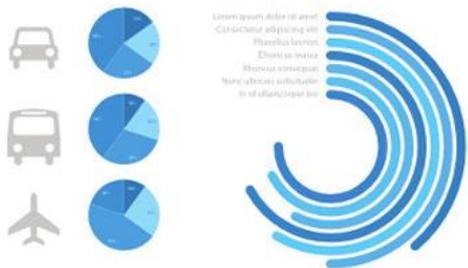
846589107830

756**3**9506847**3**

65866**303**7576

860**3**72658602

8465891078**30**



Ut enim ipsam dolor sit amet
 Consectetur adipiscing elit
 Phasellus laoreet
 Etiam sit amet
 Morbi ut ipsum
 Nam ultrices euismod
 In id ultrices sit

Analytics

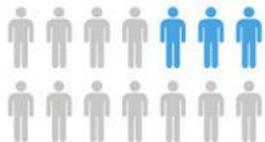


Ut enim ipsam dolor sit amet, consectetur adipiscing
 elit. Phasellus laoreet rhoncus massa, rhoncus
 consequat. Donec tempus ac. Nam ultrices euismod
 euismod. In id ultrices sit

World map (political)



Statistics



Визуализация и хорошие манеры

Визуализация и хорошие манеры

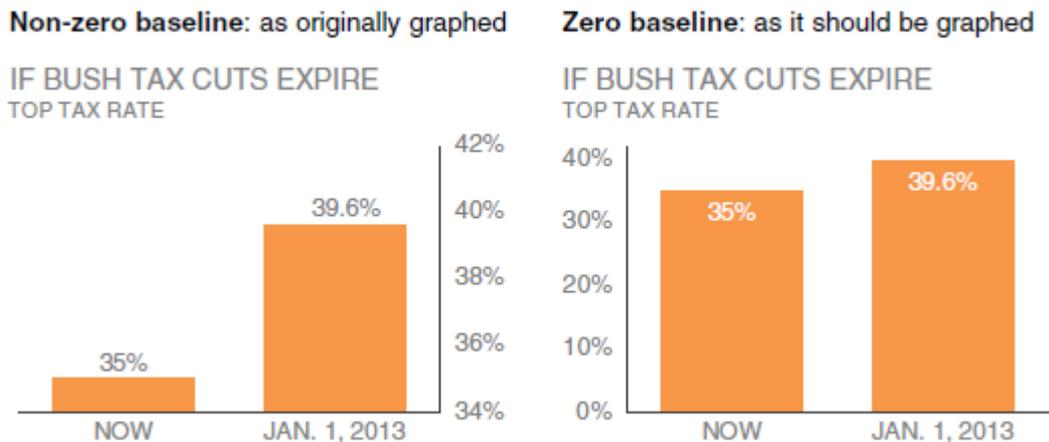
У графика всегда должны быть:

- Название
- Легенда
- Подписанные оси
- Ссылка, если график не ваш

Визуализация и хорошие манеры

Данные не должны искажаться при визуализации

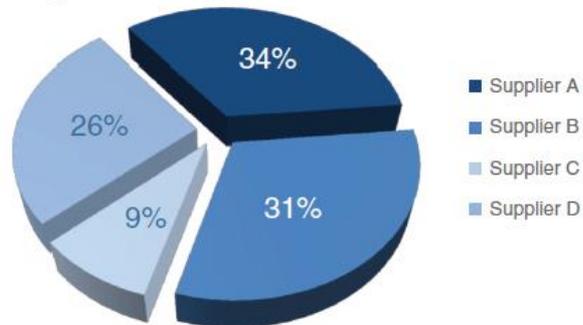
- гистограммы рисуются от 0
- логарифмические шкалы подписываются



Визуализация и хорошие манеры

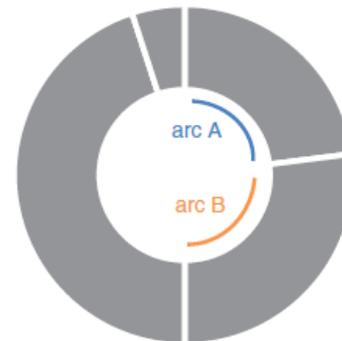
3D графики – плохая идея

Supplier Market Share



Круговые/donut диаграммы -
тоже (увы)

The donut chart



Про цвет

- Не больше 5 цветов на одном графике
- Оттенки одного цвета – лучший выбор
- Всё незначительное: единицы измерения, пояснения под графиком можно сделать серым
- В большинстве случаев белый фон выигрывает

Про цвет

Country Level Sales Rank Top 5 Drugs

Rainbow distribution in color indicates sales rank in given country from #1 (red) to #10 or higher (dark purple)

Country	A	B	C	D	E
AUS	1	2	3	6	7
BRA	1	3	4	5	6
CAN	2	3	6	12	8
CHI	1	2	8	4	7
FRA	3	2	4	8	10
GER	3	1	6	5	4
IND	4	1	8	10	5
ITA	2	4	10	9	8
MEX	1	5	4	6	3
RUS	4	3	7	9	12
SPA	2	3	4	5	11
TUR	7	2	3	4	8
UK	1	2	3	6	7
US	1	2	4	3	5

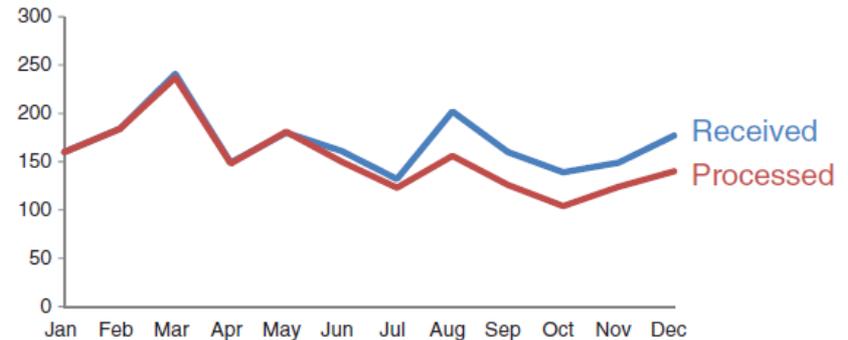
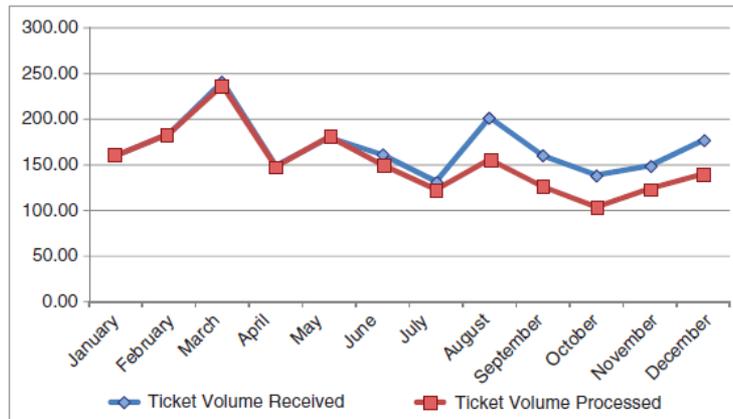
Top 5 drugs: country-level sales rank

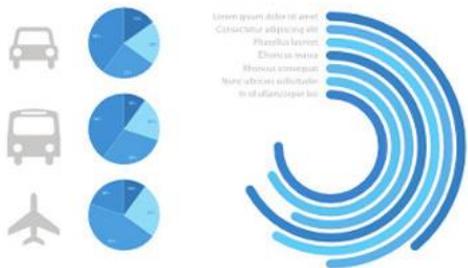
RANK **1** 2 3 4 5+

COUNTRY DRUG	A	B	C	D	E
	Australia	1	2	3	6
Brazil	1	3	4	5	6
Canada	2	3	6	12	8
China	1	2	8	4	7
France	3	2	4	8	10
Germany	3	1	6	5	4
India	4	1	8	10	5
Italy	2	4	10	9	8
Mexico	1	5	4	6	3
Russia	4	3	7	9	12
Spain	2	3	4	5	11
Turkey	7	2	3	4	8
United Kingdom	1	2	3	6	7
United States	1	2	4	3	5

Про избыточность

- Убираем направляющие
- Убираем метки на графиках
- Переносим легенду на сам график





Ut enim deserit amet
Consectetur adipiscing elit
Phasellus laoreet
Etiam sit amet
Morbi euismod
Nunc ultrices euismod
In id ultricesper leo

Analytics

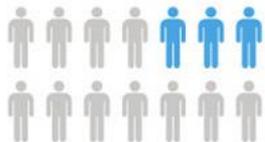


Ut enim deserit amet, consectetur adipiscing
elit. Phasellus laoreet Phasellus massa, Phasellus
consequat enim tempus ac. Nunc ultrices euismod
euismod. In id ultricesper leo.

World map (political)



Statistics



Основные типы визуализаций

Просто текст / карточка / KPI chart

20%

of children had a
traditional stay-at-home mom
in 2012, compared to 41% in 1970

HR-бюджет/ сотрудника (руб.)

151,56K +26.05%

Факт (vs. План)

▲ План

ФОТ / сотрудника (руб.)

76,2K -0.46%

Факт (vs. План)

▼ План

KPI Chart

Apples

2016 » Mar » 31

210

Revenue



Oranges

2016 » Mar » 31

-10

Revenue



Большая цифра – отличный вариант, если нужно показать 1 или 2 значения

Таблица

Heavy borders

Group	Metric A	Metric B	Metric C
Group 1	\$X.X	Y%	Z,ZZZ
Group 2	\$X.X	Y%	Z,ZZZ
Group 3	\$X.X	Y%	Z,ZZZ
Group 4	\$X.X	Y%	Z,ZZZ
Group 5	\$X.X	Y%	Z,ZZZ

Light borders

Group	Metric A	Metric B	Metric C
Group 1	\$X.X	Y%	Z,ZZZ
Group 2	\$X.X	Y%	Z,ZZZ
Group 3	\$X.X	Y%	Z,ZZZ
Group 4	\$X.X	Y%	Z,ZZZ
Group 5	\$X.X	Y%	Z,ZZZ

Minimal borders

Group	Metric A	Metric B	Metric C
Group 1	\$X.X	Y%	Z,ZZZ
Group 2	\$X.X	Y%	Z,ZZZ
Group 3	\$X.X	Y%	Z,ZZZ
Group 4	\$X.X	Y%	Z,ZZZ
Group 5	\$X.X	Y%	Z,ZZZ

Таблица полезна, когда много данных в разных единицах измерения
В остальных случаях более наглядна графика

Heatmap

Table

	A	B	C
Category 1	15%	22%	42%
Category 2	40%	36%	20%
Category 3	35%	17%	34%
Category 4	30%	29%	26%
Category 5	55%	30%	58%
Category 6	11%	25%	49%

Heatmap

LOW-HIGH

	A	B	C
Category 1	15%	22%	42%
Category 2	40%	36%	20%
Category 3	35%	17%	34%
Category 4	30%	29%	26%
Category 5	55%	30%	58%
Category 6	11%	25%	49%

Scatterplot



Источник: www.gapminder.org

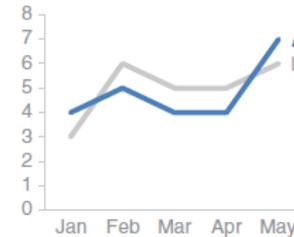
Графики / Line graphs

Используются в основном для временных рядов

Single series



Two series



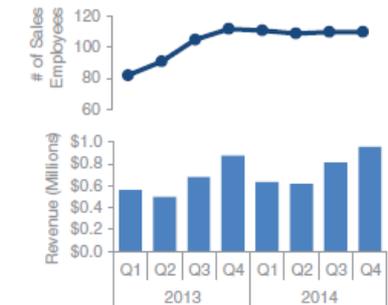
Multiple series



Alternative 1: label directly



Alternative 2: pull apart vertically

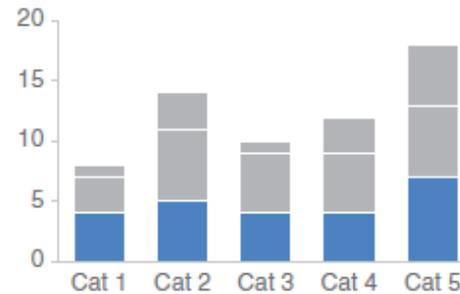


Дополнительные оси сложно считываются

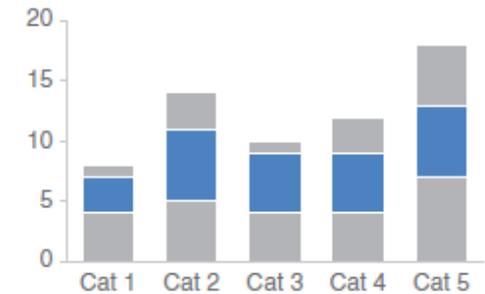
Гистограммы

- Вертикальные
- Горизонтальные
- С накоплением
- Приведённые к 100%

Comparing **these** is easy



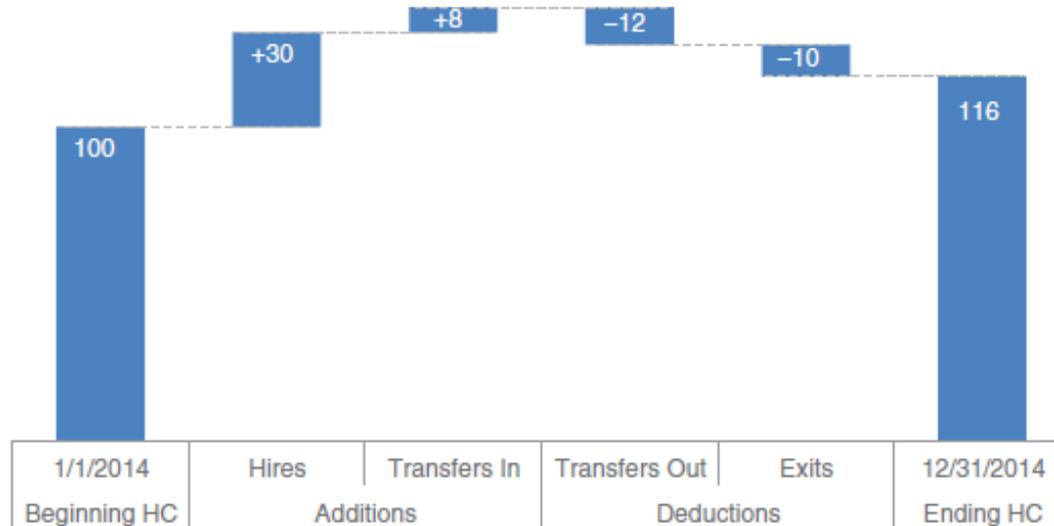
Comparing **these** is hard



Waterfall chart

2014 Headcount math

Though more employees transferred out of the team than transferred in, aggressive hiring means overall headcount (HC) increased 16% over the course of the year.



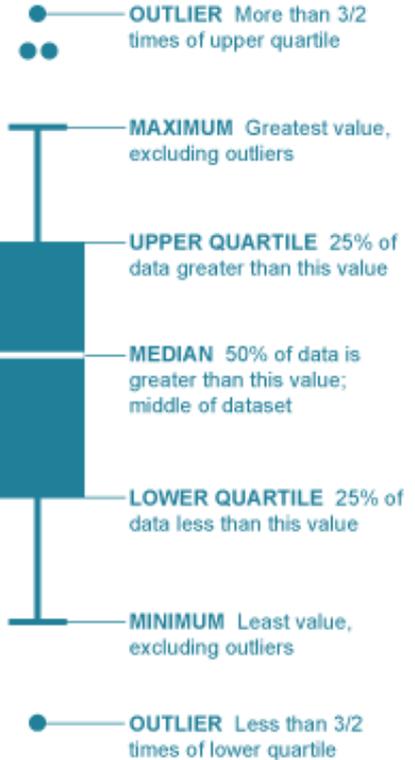
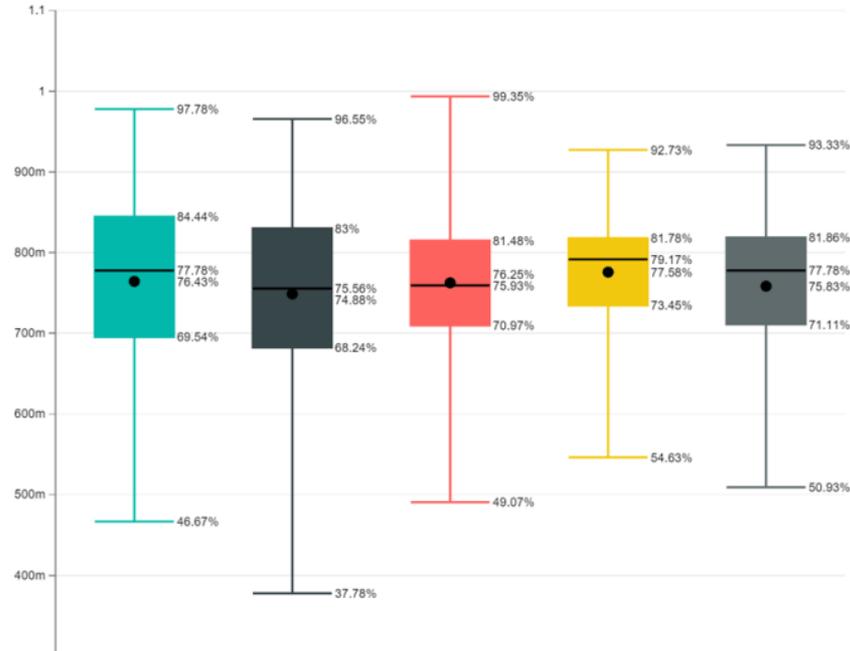
Используем, когда нужно показать составляющие одного целого

Hint: в Excel рисуется как гистограмма с накоплением с нижними блоками цвета фона

Ящик с усами / boxplot

Grade by Subject and Full Name

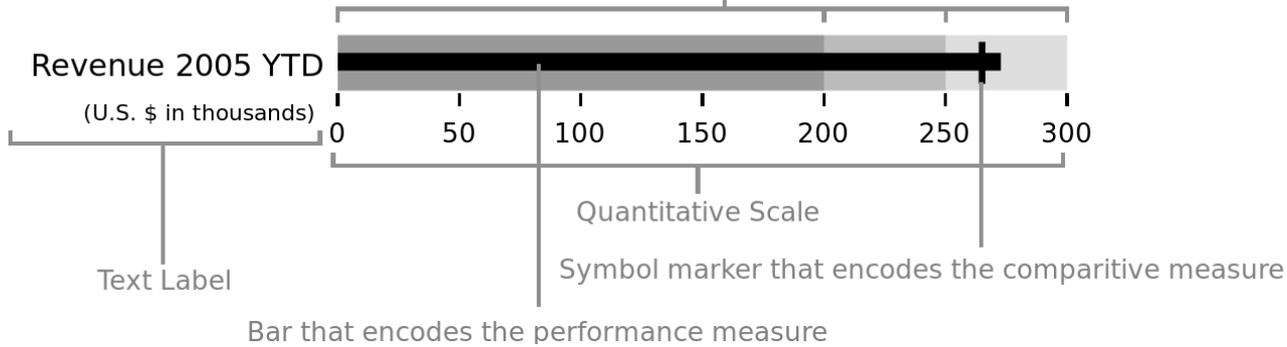
Subject ● Languages ● Life Skills ● Math & Comp Sci ● Science ● Social Studies

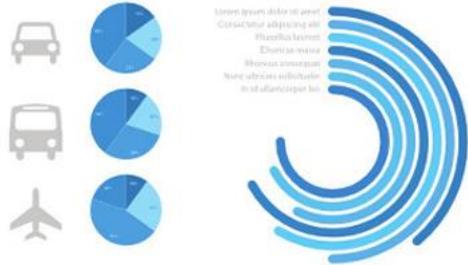


Тяжело считывается большинством людей

Bulletgraph

Background fill colors that encode qualitative ranges such as bad, satisfactory, and good





Ut enim deserit et amet. Consectetur adipiscing elit. Phasellus lacus. Etiam in massa. Moris ut porttitor. Nam ultrices euismod. In id ultricesque leo.

Analytics



Ut enim deserit et amet. Consectetur adipiscing elit. Phasellus lacus. Etiam in massa. Moris ut porttitor. Nam ultrices euismod. In id ultricesque leo.

World map (political)



Statistics



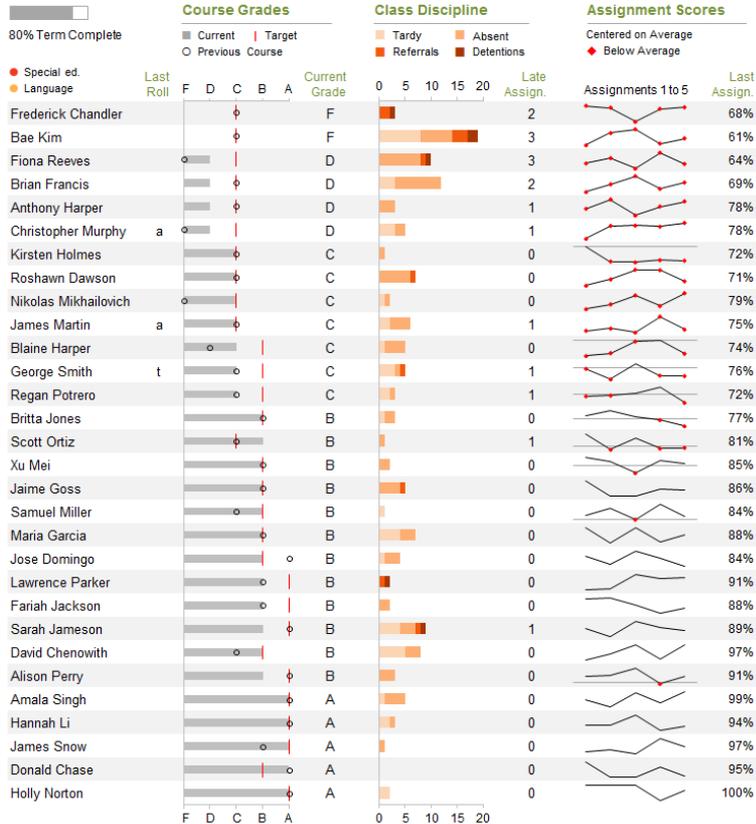
Как это выглядит на практике

Dashboard

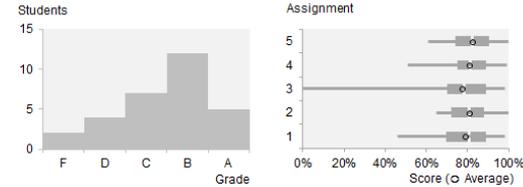
May 1, 2012
Tuesday

Grade 10 Algebra Course

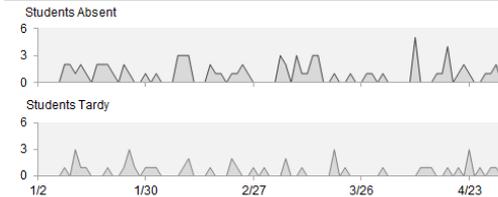
Note: All scores are expressed as percentage of points earned out of the total points possible. [HELP](#)



Grade and Assignment Score Distribution



Attendance (excluding weekends)



Standardized Math Assessment Score Distribution



Standardized Math Assessment Median Score

Current Class School	Other Class School	Other Class District	District
79.0%	77.4%	74.2%	71.9%

Что почитать

- Few, Stephen «Information Dashboard Design: Displaying Data for At-a-glance Monitoring» Analytics Press, 2013
- Nussbaumer Knaflic, Cole «Storytelling with Data: A Data Visualization Guide for Business Professionals» Wiley, 2015.
- Tufte, Edward R. «The Visual Display of Quantitative Information» Graphics Press, 2001