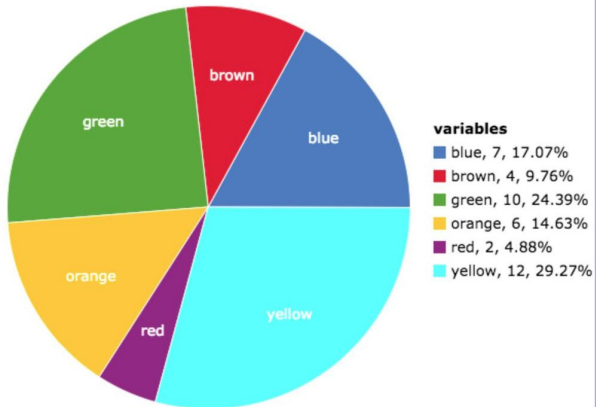


Анализ данных

Блок 4: Визуализация данных

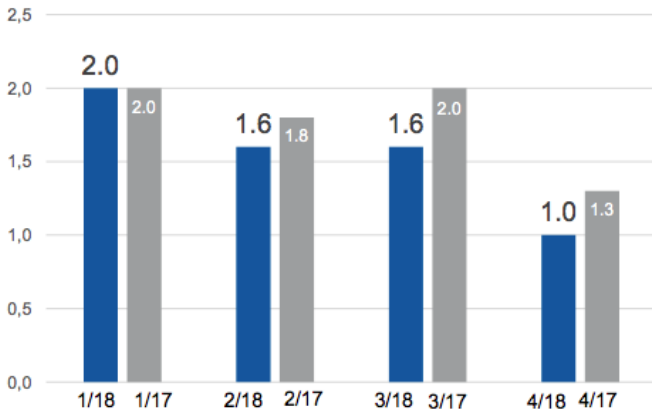
Елисеева Виталия, Небольсина Анастасия

1 августа 2020



EBIT by quarter

in € million



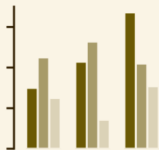


Section 1

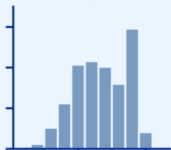
Как визуализировать различные виды данных

Как выбрать нужный тип графика?

Grouped Bars



Histogram



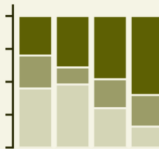
Pie Chart



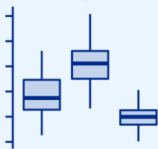
Multiple Pie Charts



Stacked Bars



Boxplots



Scatterplot



Choropleth



Столбчатая диаграмма: средние значения для качественных/количественных признаков

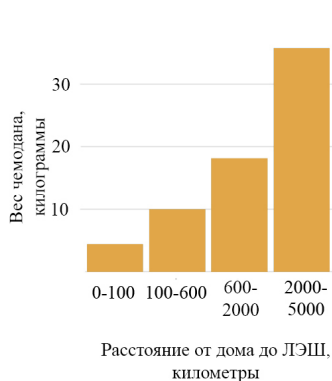


Figure: Что влияет на вес чемодана

Столбчатая диаграмма с качественными: попробуйте группировку

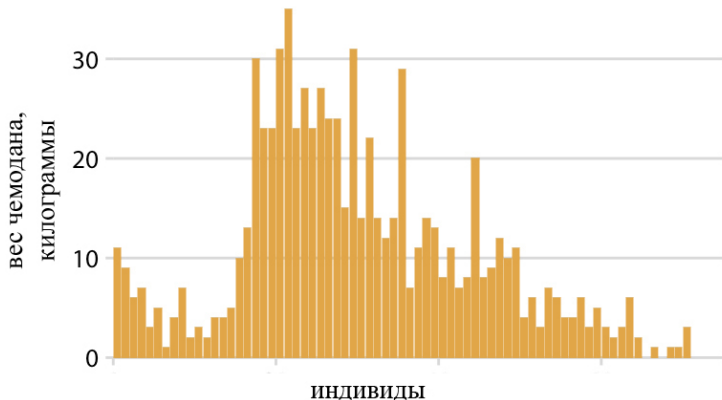


Figure: Что влияет на вес чемодана

Гистограмма: частоты для качественных/количественных признаков

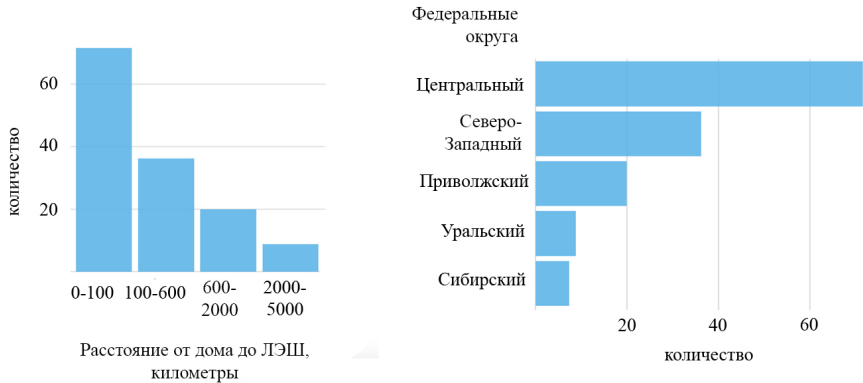


Figure: Что влияет на вес чемодана

Гистограмма с количественными: попробуйте несколько длин интервалов

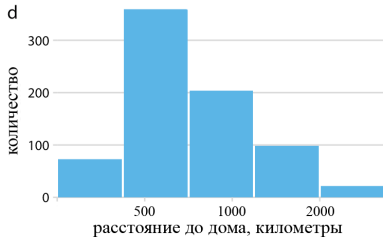
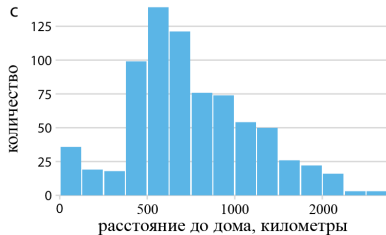
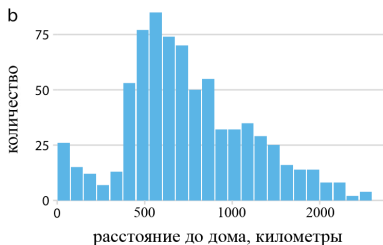


Figure: Что влияет на вес чемодана

Гистограмма с накоплением

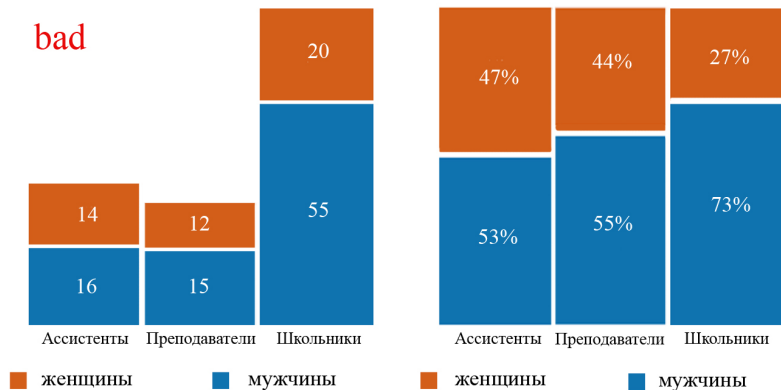


Figure: Гендерный состав ЛЭШ

Age pyramid

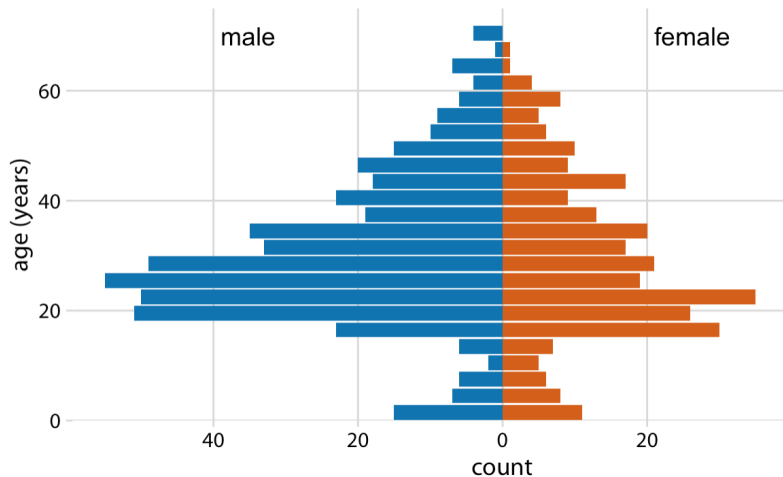


Figure: The age distributions of male and female Titanic passengers visualized as an age pyramid

Боксплот: распределение количественного признака

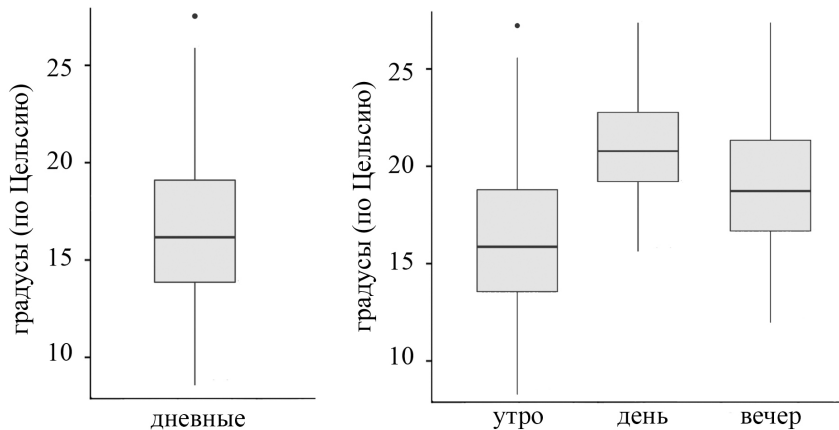


Figure: Средняя температура воздуха

Боксплот: распределение количественного признака

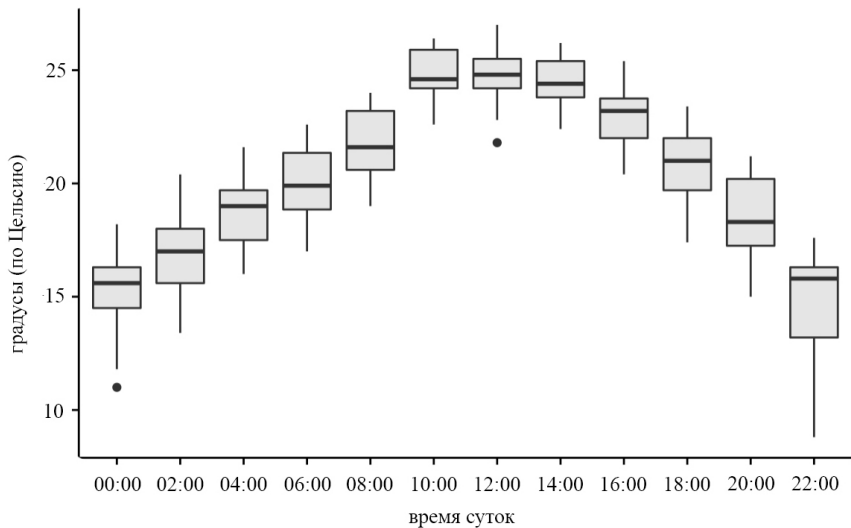


Figure: Средняя температура воздуха

Pie-chart

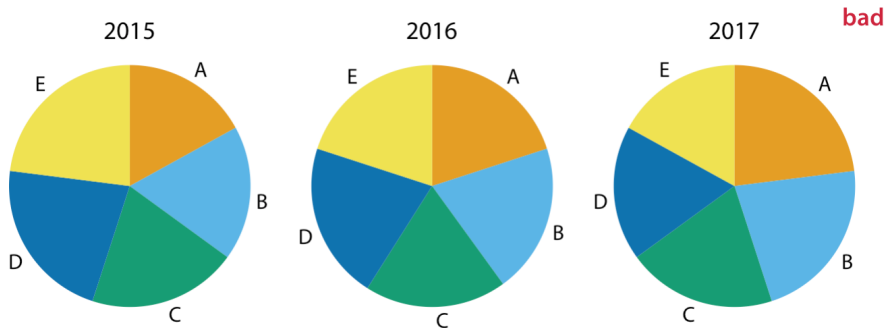
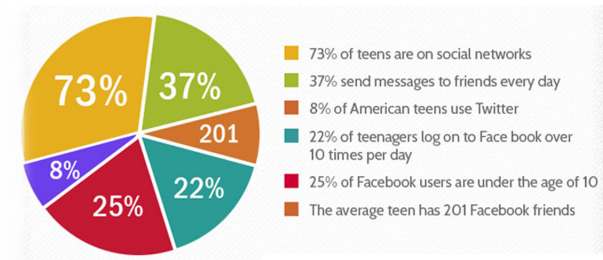
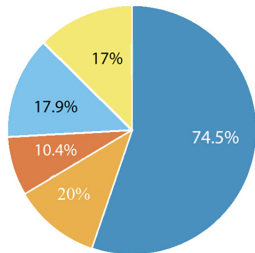
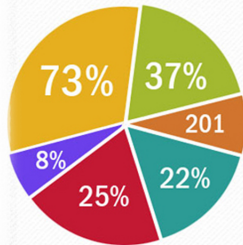
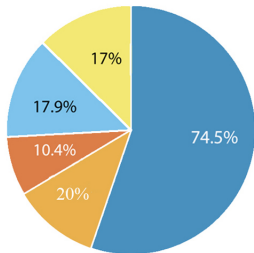


Figure: Процент участников ЛЭШ из 5 наиболее крупных федеральных округов

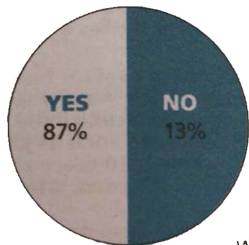
Pie-chart



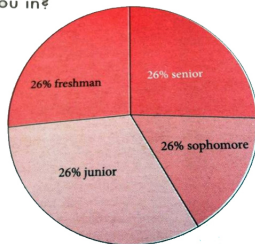
Pie-chart



- 73% of teens are on social networks
- 37% send messages to friends every day
- 8% of American teens use Twitter
- 22% of teenagers log on to Face book over 10 times per day
- 25% of Facebook users are under the age of 10
- The average teen has 201 Facebook friends



What grade are you in?



Would you pay more for your milk to help dairy farmers?

Скаттерплот: Зависимость между количественными признаками

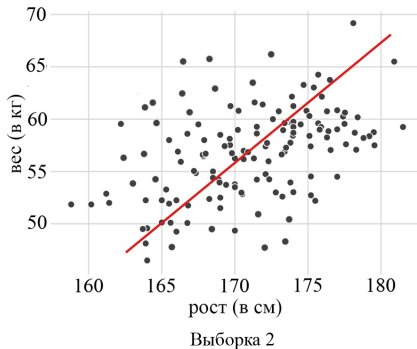
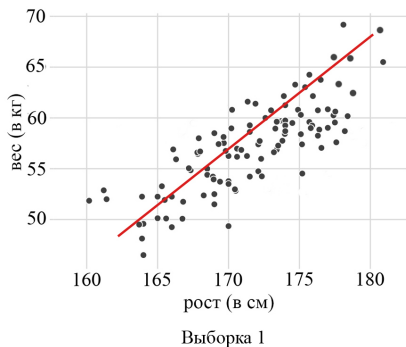


Figure: Рост и вес школьников на ЛЭШ 2018

Скаттерплот: деление на подвыборки

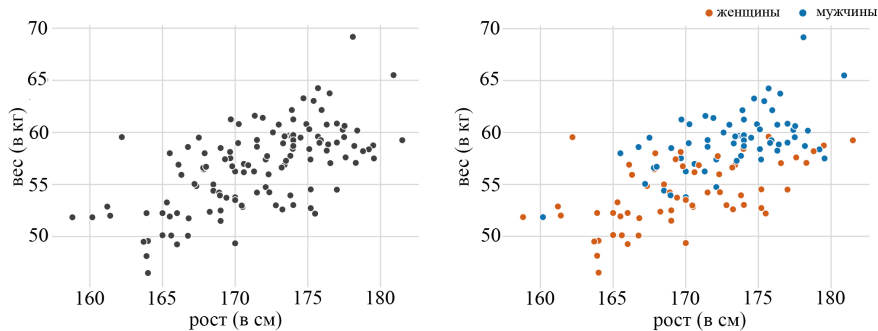


Figure: Рост и вес школьников на ЛЭШ 2018

Скаттерплот: выбросы

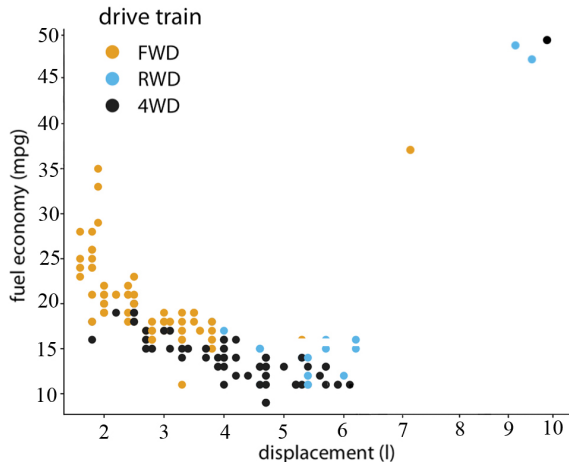


Figure: City fuel economy versus engine displacement, for popular cars released between 1999 and 2008

Скаттерплот: наложение точек

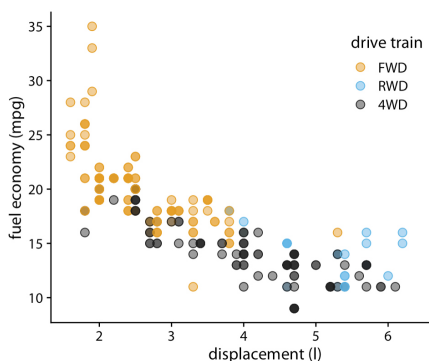
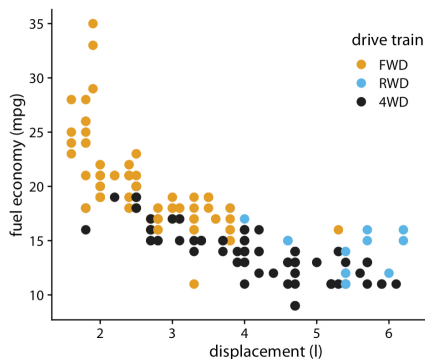


Figure: City fuel economy versus engine displacement, for popular cars released between 1999 and 2008

Таблицы

а

Rank	Title	Amount
1	<i>Star Wars: The Last Jedi</i>	\$71,565,498
2	<i>Jumanji: Welcome to the Jungle</i>	\$36,169,328
3	<i>Pitch Perfect 3</i>	\$19,928,525
4	<i>The Greatest Showman</i>	\$8,805,843
5	<i>Ferdinand</i>	\$7,316,746

Таблицы

a

Rank	Title	Amount
1	<i>Star Wars: The Last Jedi</i>	\$71,565,498
2	<i>Jumanji: Welcome to the Jungle</i>	\$36,169,328
3	<i>Pitch Perfect 3</i>	\$19,928,525
4	<i>The Greatest Showman</i>	\$8,805,843
5	<i>Ferdinand</i>	\$7,316,746

b

Rank	Title	Amount
1	<i>Star Wars: The Last Jedi</i>	\$71,565,498
2	<i>Jumanji: Welcome to the Jungle</i>	\$36,169,328
3	<i>Pitch Perfect 3</i>	\$19,928,525
4	<i>The Greatest Showman</i>	\$8,805,843
5	<i>Ferdinand</i>	\$7,316,746

Таблицы

a

ugly

Rank	Title	Amount
1	<i>Star Wars: The Last Jedi</i>	\$71,565,498
2	<i>Jumanji: Welcome to the Jungle</i>	\$36,169,328
3	<i>Pitch Perfect 3</i>	\$19,928,525
4	<i>The Greatest Showman</i>	\$8,805,843
5	<i>Ferdinand</i>	\$7,316,746

b

ugly

Rank	Title	Amount
1	<i>Star Wars: The Last Jedi</i>	\$71,565,498
2	<i>Jumanji: Welcome to the Jungle</i>	\$36,169,328
3	<i>Pitch Perfect 3</i>	\$19,928,525
4	<i>The Greatest Showman</i>	\$8,805,843
5	<i>Ferdinand</i>	\$7,316,746

c

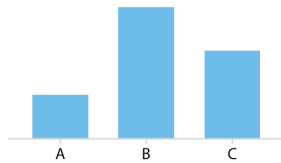
Rank	Title	Amount
1	<i>Star Wars: The Last Jedi</i>	\$71,565,498
2	<i>Jumanji: Welcome to the Jungle</i>	\$36,169,328
3	<i>Pitch Perfect 3</i>	\$19,928,525
4	<i>The Greatest Showman</i>	\$8,805,843
5	<i>Ferdinand</i>	\$7,316,746

Section 2

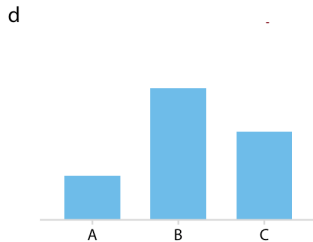
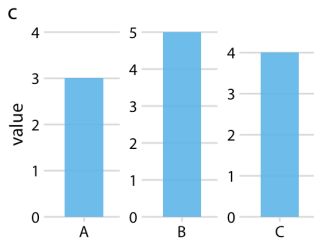
Красивая и эффективная визуализация данных

Основные ошибки – в общем

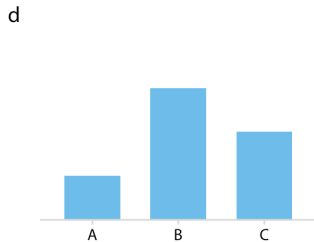
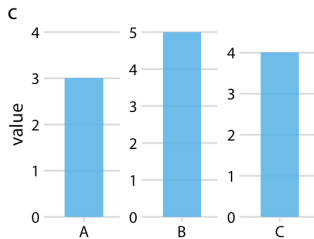
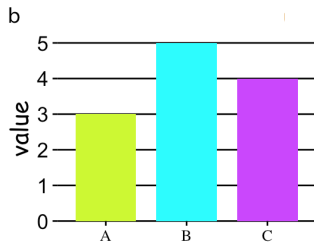
d



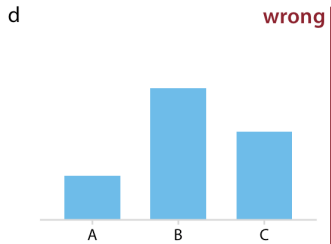
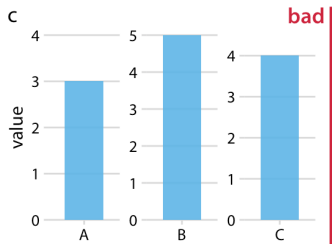
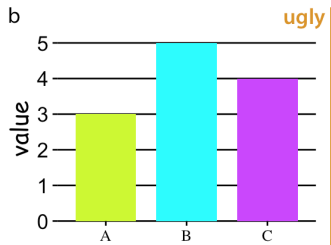
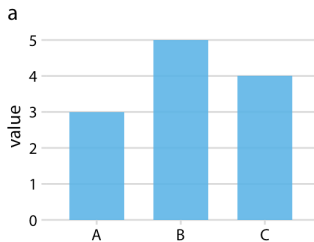
Основные ошибки – в общем



Основные ошибки – в общем



Основные ошибки – в общем



О легенде

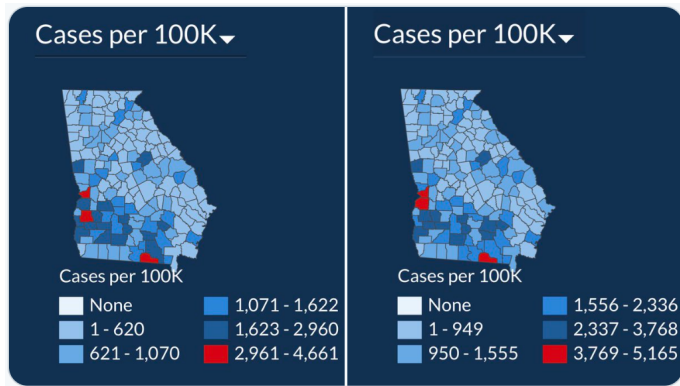


Figure: Total number of COVID19 cases in Georgia, US: July 2 vs. July 17

Несколько графиков

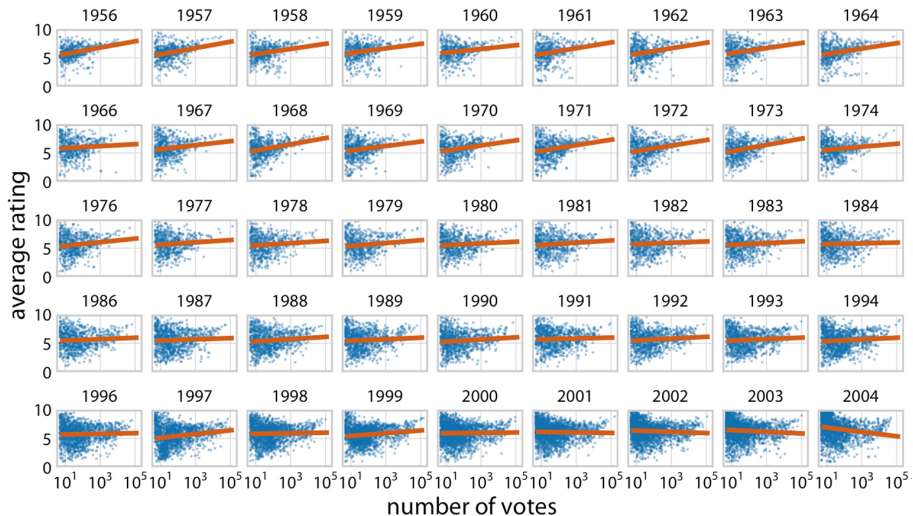
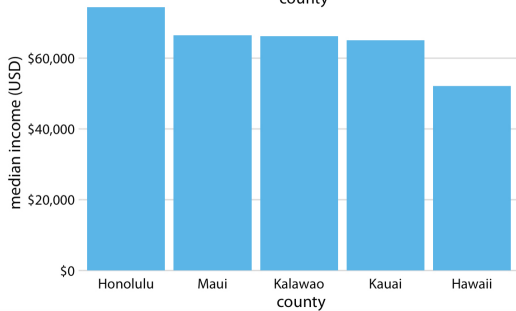
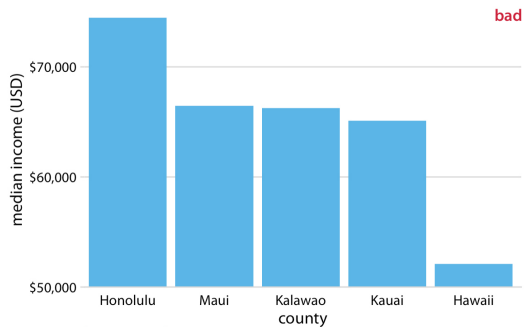


Figure: Average movie rankings versus number of votes, for movies from 1946 to 2005.

Пропорциональность, размерность осей



Как сортировать оси

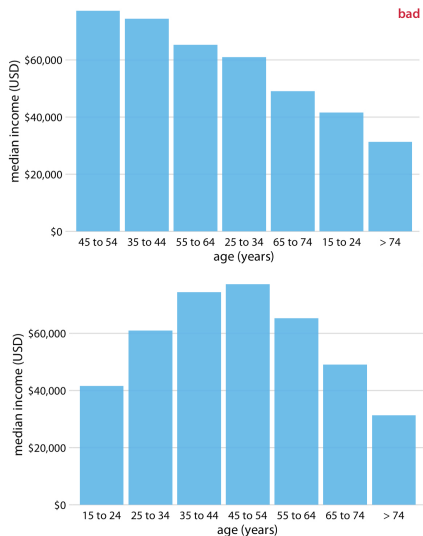


Figure: 2016 median U.S. annual household income versus age group

Как сортировать оси (пример из реальной жизни)

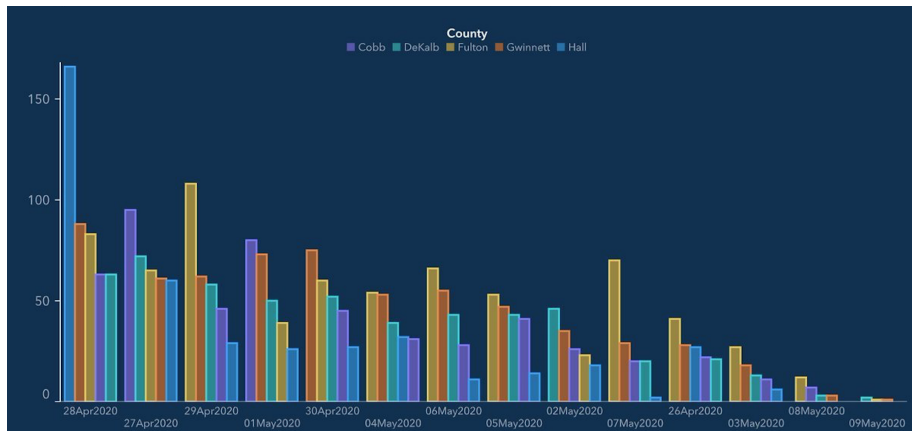


Figure: Top 5 counties with the Greatest number of COVID19 cases

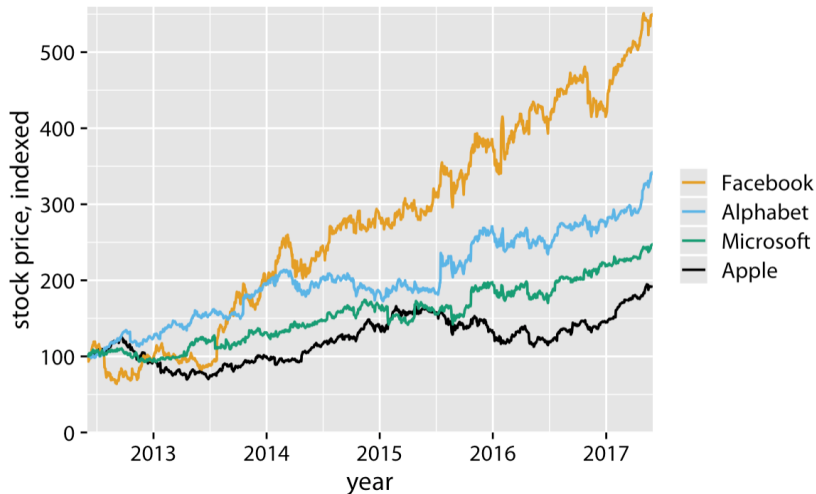


Figure: Stock price over time for four major tech companies.

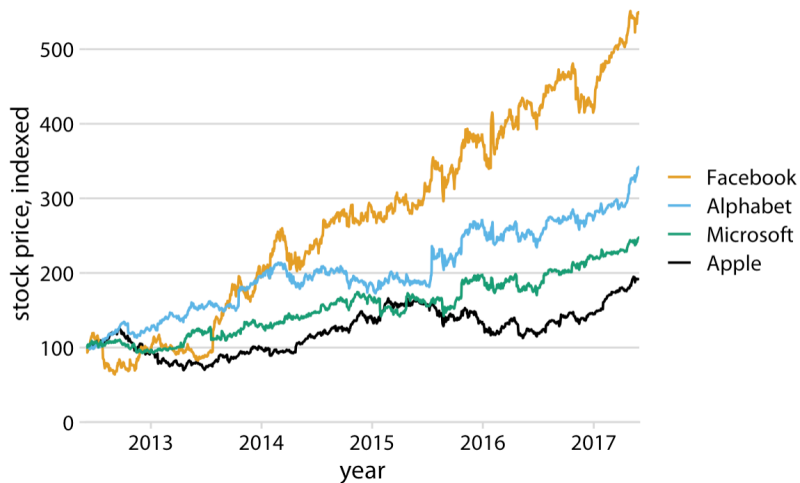
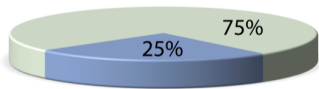


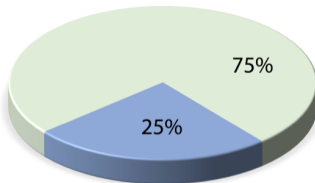
Figure: Stock price over time for four major tech companies.

3D, тени, объём – зло!

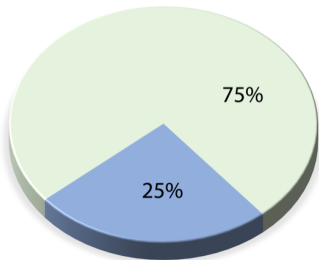
a



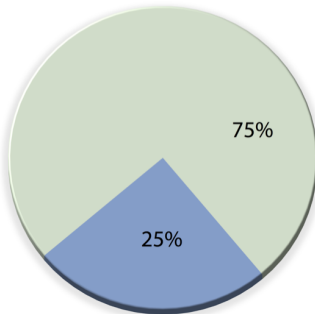
b



c



d



3D, тени, объём – зло!

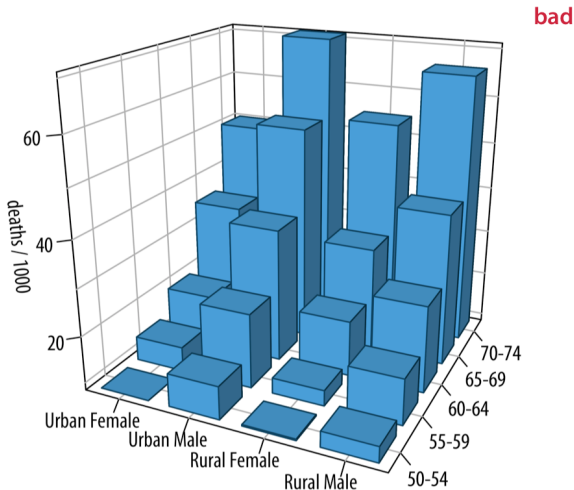


Figure: Mortality rates in Virginia in 1940.

Про инфографику: бесполезная карта

Total Military Spending of Iran and Russia



Source: [World Bank](#). As of 2013

Summarize

- График – самостоятельная смысловая единица;
- Белый фон, чёрные оси и подписи;
- Ранговые переменные отсортированы по рангу;
- Верно выбран масштаб осей, размерность единиц измерения;
- Не забыть подписать: название (короткое, но информативное), легенду (цвета, единицы измерения), оси;
- Цвета или почему не стоит делать радугу;
- Пропорции или как выбрать шкалу;
- 3D, тени, анимация, фигурные столбики – зло.

Что почитать дальше?

- Источник иллюстраций: Fundamentals of Data Visualization by Claus O. Wilke (<https://serialmentor.com/dataviz/>)
- Источник найденных в интернете уродливых графиков: https://twitter.com/Chart_Crime
- Джин Желязны "Говори на языке диаграмм"
- Tufte, E. R. Envisioning Information
- Tufte, E. R. The Visual Display of Quantitative Information
- Инфостиль: Максим Ильяхов и Людмила Сарычева «Пиши, сокращай»