

Графики, которые убеждают всех

Александр Богачев

Посвящается
родителям

УДК 311
ББК 60.6
Б73

Б73 Богачев, Александр Андреевич

Графики, которые убеждают всех / А. А. Богачев. —
Москва: Издательство АСТ, 2020. — 240 с. —
(БизнесНаставник).

978-5-17-127186-2

Первая книга об эффективном представлении данных от русскоязычного автора. Книга рассказывает, как подготовить данные к работе, как выбрать подходящий для своих данных график или диаграмму, как грамотно оформить график, чтобы он максимально доносил ваше сообщение, как распознать ситуацию, когда статистикой пытаются манипулировать.

Александр Богачев — один из ведущих в стране специалистов по визуализации данных, инфографике и дата-журналистике. Работал ведущим дизайнером в Студии инфографики сайта РИА.ру, руководителем отдела дата-спецпроектов в РБК, сотрудничал как информационный дизайнер и руководитель проектов с крупнейшими российскими компаниями. Преподает визуализацию данных в Высшей школе экономики.

УДК 311
ББК 60.6

978-5-17-127186-2

Все права защищены. Никакая часть данной книги не может быть воспроизведена в какой бы то ни было форме без письменного разрешения владельцев авторских прав.

© Богачев А. А., 2020
© Богачев А. А., иллюстрации, 2020
© ООО «Издательство АСТ», 2020

Оглавление

От автора.....	13
Введение	17
Глава 1. Пример было/стало. О процессе в общем.....	23
Было/стало	24
Этапы работы	33
Глава 2. Понимать свою задачу и аудиторию	35
Что еще будет влиять на вашу работу?	37
Главное во второй главе	42
Глава 3. Формулируем сообщение.....	43
Качество данных	48
Погружение в тему и контекст.....	51
Знать основы статистики	58
Быть аналитиком. Задавать вопросы.....	62
Как найти важное и интересное в вашем датасете?.....	68
Создание новых данных внутри датасета.....	69
Формулируем сообщение.....	79
Об объективности и непредвзятости.....	81
Главное в третьей главе	84

Глава 4. Когда визуализация данных не нужна 85

Просто число 86

Слово о таблицах..... 88

Если в вашей таблице есть числовые значения, как ее лучше оформить?..... 89

Таблицы как тепловые карты..... 95

Главное в четвертой главе 98

Глава 5. Как устроена визуализация данных, виды графиков 99

Суть визуализации данных..... 99

Эксперимент Сантьяго Ортиза 100

Как образуется визуализация 106

Типы сравнения..... 111

1. Сравнение 112

2. Сравнение части с целым..... 113

3. Временное сравнение 114

4. Распределение 116

5. Корреляция 117

6. Сравнение отклонения/разницы 120

7. Иерархия 120

8. Этапы/ процесс/ поток..... 120

9. Составление профилей..... 121

Какие графики лучше..... 123

Что еще влияет на выбор графика?..... 126

Уровень подготовленности аудитории 127

Формат..... 128

Условия использования 129

Техническая возможность.....	129
Необходимость обновления.....	130
Главное в пятой главе.....	130
Сравнение.....	131

Глава 6. Виды данных, как их визуализировать, и что делать в сложных случаях131

Вертикальные и горизонтальные столбики.....	132
Какая ширина должна быть между столбиками?.....	134
Чем заменить столбиковую диаграмму?.....	134
Часть и целое	136
Какие еще ошибки бывают при использовании круговой диаграммы?.....	138
Как лучше показывать покомпонентное сравнение?.....	139
Чем заменить круговую диаграмму?.....	140
Изменения во времени.....	143
Можно ли заменить линию столбиками?.....	143
Какую размерность выбрать для линейного графика? Можно ли обрезать шкалу?	144
Можно ли из линейного графика сделать площадной — закрасить его под линией?.....	147
Что делать, если данные отсутствуют?	148
Распределение.....	150
Разница со столбиковой диаграммой.....	151
Как выбирать интервалы?	152
Альтернативы гистограммам	153
Корреляция/взаимосвязь	156
Альтернативы графику рассеивания.....	158

Отклонение значений/разница	160
Что делать, если график выглядит перегруженным? Боремся с «лапшой»	169
Еще раз об алгоритме выбора графиков	176
Главное в шестой главе	178
Глава 7. Ошибки в визуализации данных	179
Приблизительность недопустима	180
Столбиковые диаграммы — всегда от нуля	181
Ошибки в круговой диаграмме	186
Нарушение непрерывности временной оси	190
О подписях значений на графике	197
Главное в седьмой главе	198
Глава 8. Как сделать график, наглядно доносящий ваше сообщение	199
Как это работает	200
Теория близости	203
Предвнимательные атрибуты	204
Акценты	206
Стандартные настройки Excel	208
Цвет. Основы	210
Характеристики цвета	212
Для чего нужен цвет?	213
Про цвет для категорий	214
Про цвет для количественных данных	215
Гармония между цветами	216
Что делать, если цветов слишком много?	218

Фон	219
О шрифтах.....	219
Что нельзя делать со шрифтом.....	222
Что можно и нужно делать со шрифтом.....	224
Микротипографика	225
Тексты и аннотации.....	226
Аннотации.....	230
Источники	231
Главное в восьмой главе	232
Главное в книге	234
В качестве завершения.....	235
Список рекомендуемой литературы.....	237

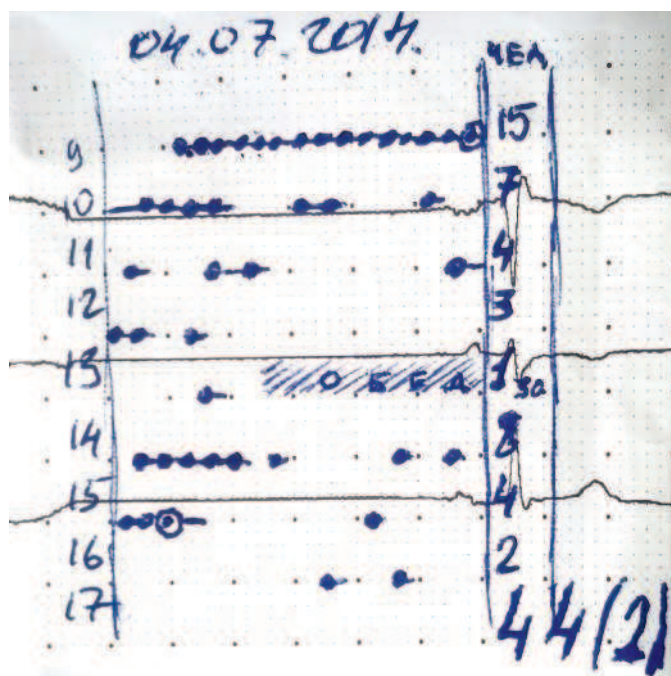
От автора

Я думаю, инфографика и визуализация данных завораживали меня с раннего детства. У нас дома были детские энциклопедии издательства «Росмэн». Обычно они полны инфографики: вулканы в разрезе, схема Солнечной системы и тому подобное.

Подростком я увлекался футболом и футбольной статистикой. Интернет тогда был совсем не развит, и мы с братом создавали собственные энциклопедии футбольных чемпионатов, вели турнирные таблицы, рисовали схемы игр на вылет, считали голы, пасы и желтые карточки. Потом эта система перешла и на результаты компьютерных игр.

Когда в 17 лет я поступил в медицинскую академию, то плотно познакомился с одним из наиболее старых жанров инфографики — атласом анатомии. После окончания ВУЗа я стал работать врачом. Казалось, инфографика забыта. Но и во время работы врачом я создавал базы данных пациентов и рисовал схемы лечения.

Вот, например, результаты одного дня профилактических осмотров. Осмотры отображены на пленке ЭКГ (сама ЭКГ никакого отношения к визуализации данных не имеет). Это результаты самого легкого дня — последнего. В первый день было 130 осмотров, во второй — 112, в третий — 88. А в последний день, изображенный на рисунке, — всего 44.



Точка — начало приема. Точка с кружком — прием с ЭКГ. В предыдущие дни я тоже фиксировал тех, кто приходил. Но оптимальную форму визуализации нашел только в последний день.

В 27 лет я ушел из медицины, поменял ряд других специальностей и окончательно остановился на инфографике и визуализации данных. Год в Санкт-Петербурге руководил мини-студией, специализирующейся на медицинской инфографике. Затем переехал в Москву и стал работать редактором в Студии инфографики сайта РИА.ру. За 2,5 года работы там выпустил больше 50 проектов, завоевал вместе с коллегами несколько международных премий, отточил навыки информационного дизайнера и столкнулся со всем спектром задач по визуализации данных. Во время работы в Студии у меня было правило: «При создании каждого нового проекта осваивать один новый инструмент или прием». Это позволило в короткие сроки овладеть массой различных техник и подходов к визуализации данных.

Мы делали в РИА масштабные проекты, но меня все больше стали привлекать простые графики и диаграммы — оказалось, что они не так просты. Чем больше я погружался в это, тем чаще замечал неудачные, перегруженные, непонятные визуализации данных вокруг. Тогда я создал в Telegram канал «Чартомойка» (chart по-английски — диаграмма), где стал проводить разборы попадавших мне графиков.

После ухода из РИА я консультировал бизнес, маркетинг, медиа и некоммерческие организации в отношении графического представления данных и презентаций. Работал с крупными компаниями над проектами, связанными с инфографикой и визуализацией данных.