

Ольга Базалева. Мастерство визуализации данных

Эта книга о том, как создавать профессиональную визуализацию с помощью диаграмм в Excel. Ведь качественная визуализация делает принятие решения в разы более быстрым и легким. Эта книга – практическое руководство на каждый день для тех, кто работает с числами и хочет упростить анализ информации. Книга содержит чек-листы и памятки, анализ реальных кейсов известных компаний, готовые решения, лайфхаки Excel и главу об оформлении таблиц.

Ольга Базалева. Мастерство визуализации данных. Как доносить идеи с помощью графиков и диаграмм. – СПб.: ООО «Диалектика», 2020. – 192 с.



Купить цифровую книгу в [ЛитРес](#), бумажную книгу в [Ozon](#) или [Лабиринте](#)

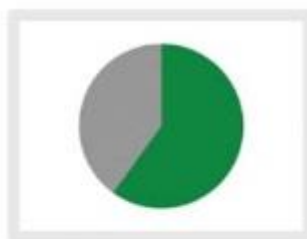
Глава 1. Немного теории

Есть всего четыре вида диаграмм, которые понятны всем и при этом подходят для большинства данных. Старайтесь использовать их.

ЛИНЕЙНАЯ ДИАГРАММА (ГРАФИК)



КРУГОВАЯ ДИАГРАММА



ЛИНЕЙЧАТАЯ ДИАГРАММА



СТОЛБЧАТАЯ ДИАГРАММА (ГИСТОГРАММА)



Рис. 1. Простые диаграммы

Когда не обойтись без более сложного вида диаграммы, обязательно давайте пояснения, как читать диаграмму, или проговаривайте это вслух, если вы выступаете с презентацией.

Построение диаграммы начинается с определения цели. Например, у вас есть данные по выручке компании и конкурентов за два года в разбивке по месяцам. Цели могут быть следующие:

1. Рассмотреть продажи компании.
2. Сравнить динамику продаж компании и конкурентов.
3. Выявить изменения выручки компании год на год.
4. Показать динамику долю компании и конкурентов.

Определившись с целью, можно выбрать диаграмму (по моему опыту, схема охватывает 90% всех аналитических задач):

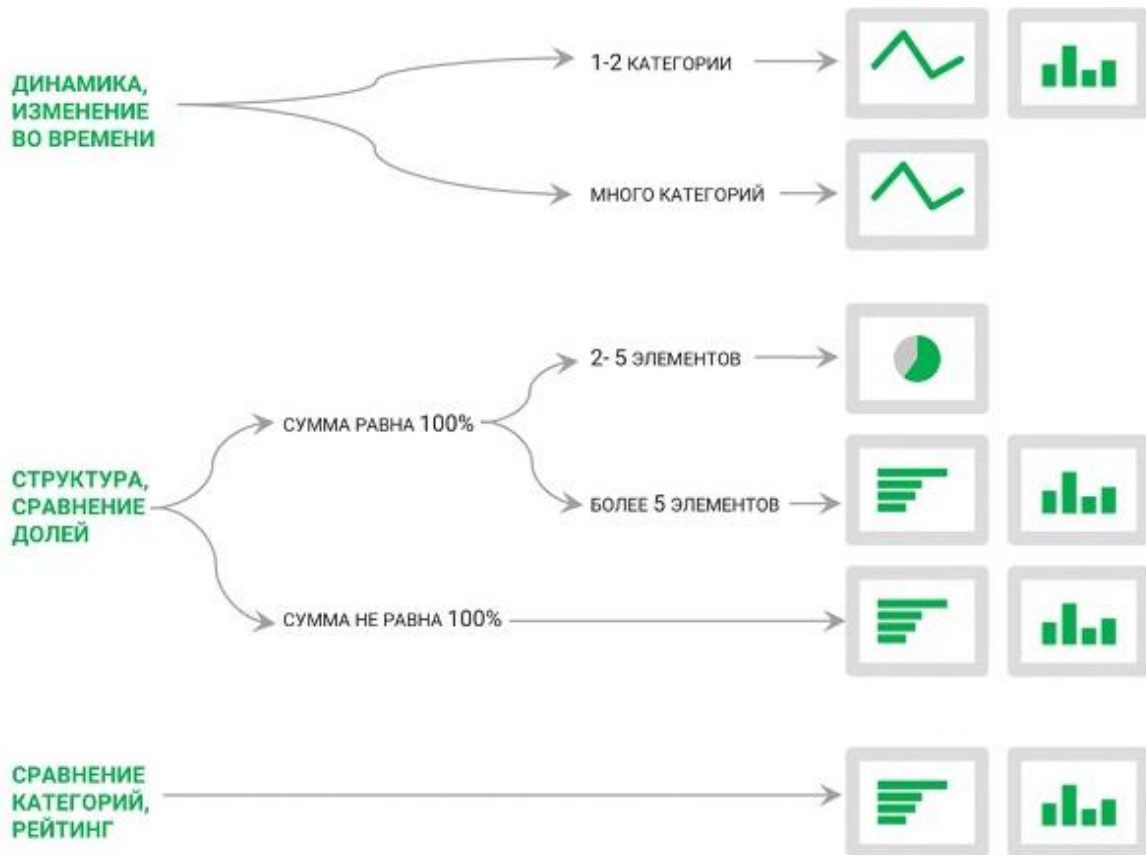


Рис. 2. Алгоритм выбора простых диаграмм

Глава 2. Одиннадцать правил визуализации данных

Не пытайтесь уместить на одну диаграмму всю имеющуюся информацию. Визуальный ряд не должен быть перегружен сложными и многоярусными диаграммами. Когда необходимо визуализировать много разных типов данных и категорий, целесообразнее разделить диаграмму на несколько частей. Например, если на линейном графике больше четырех-пяти линий или на столбиковой диаграмме больше двух категорий, не стоит уместать их на одном графике.

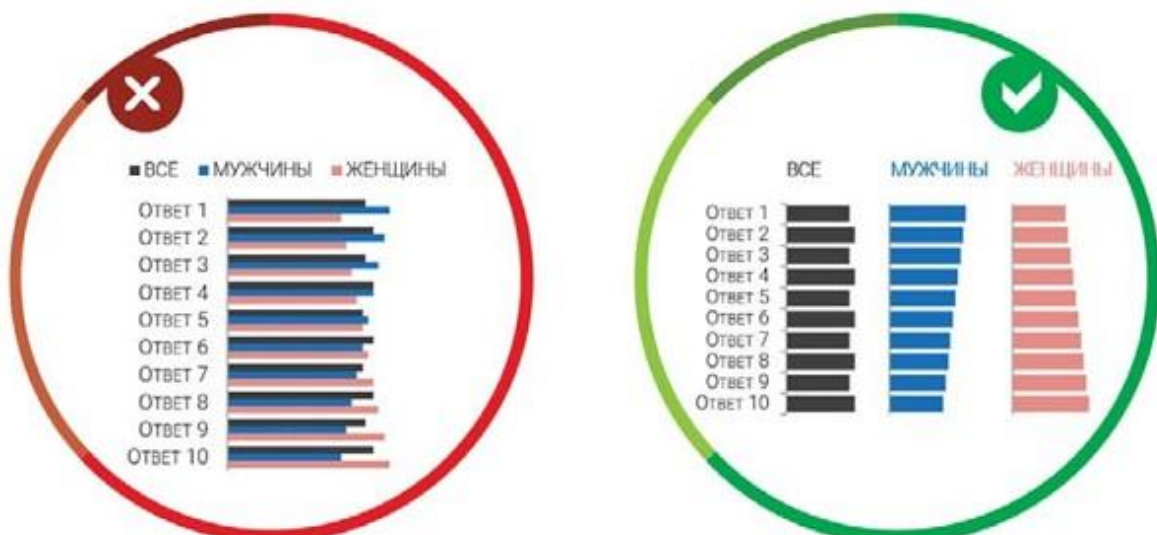


Рис. 3. Если данных много, разделите их

Числа должны быть с разделителями разрядов и без лишних знаков после запятой. У диаграммы должны быть название и полная легенда. Есть несколько основных категорий, которые у нас всегда ассоциируются с определенным цветом.

- Положительные и отрицательные значения – зеленый и красный.
- Да/нет, согласен/не согласен – зеленый и красный.
- Мужчины и женщины – голубой и розовый.
- Прочее/другое/остальное/нет ответа/затрудняюсь ответить – серый цвет.

Для сравнения показателей текущего года и прошедшего сделайте прошедший год более бледным. Оба года покажите в оттенках одного цвета.

Три кита удачной визуализации:

- Акцент на главной идее, которая может быть прописана в названии, показана с помощью цвета или добавлена как комментарий на график.
- Цвет помогает акцентировать главную идею, сделать легенду понятной, структурировать информацию, отделить менее значимые элементы.
- Чем проще, тем лучше. Для восприятия диаграммы не должны требовать долгого чтения сопроводительного текста. Диаграммы должны заменять текст.

Глава 3. Анализ ошибок на примере реальных кейсов

Семь главных ошибок: неверно выбран тип графика, плохо подобраны цвета, есть лишние элементы, не указана единица измерения, неудобный формат больших чисел — ненужные знаки после запятой, нет разделителей разрядов, в круговой диаграмме много секторов, в круговой диаграмме сумма секторов не равна 100%.



Рис. 4. Пример улучшения диаграммы

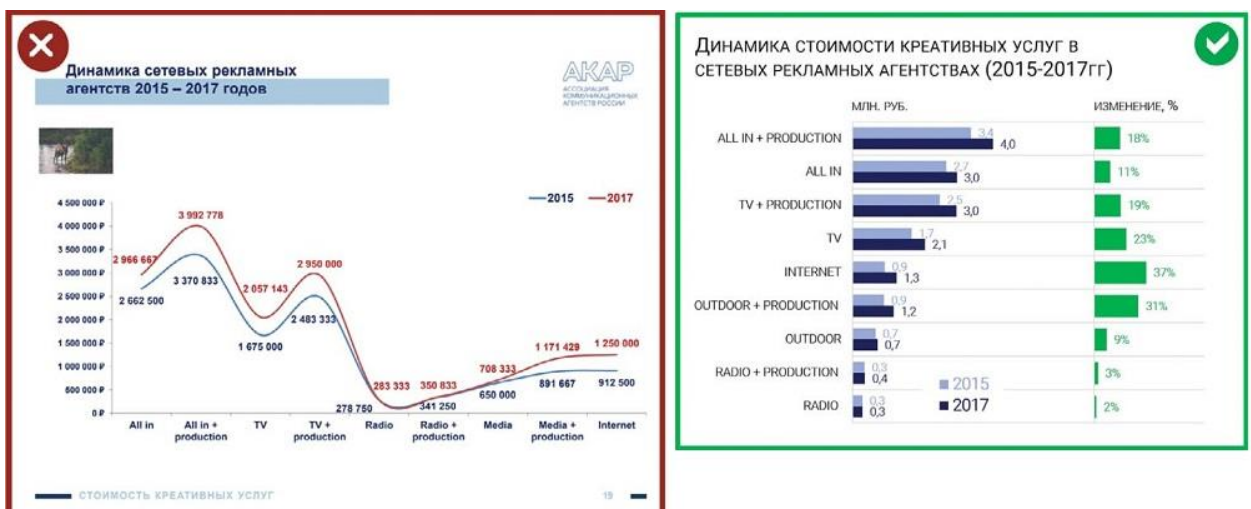


Рис. 5. Еще один пример

Глава 4. Приемы визуализации

Приемы использования цвета помогают читателю быстрее считывать идею ваших графиков. С помощью цвета можно показать положительные и отрицательные значения, ответы Да-Нет-Не знаю, максимум и минимум, прогнозные значения, текущий и прошлый период; расставить акценты; отделить менее значимую информацию; подчеркнуть фирменный стиль.



Рис. 6. Пример выделения отрицательных значений

Чтобы понятно показать на графике текущий и прошлый годы, достаточно сделать текущий год ярче, а прошлый – бледнее:



Рис. 7. Выделение данных за текущий и прошлый год

Помимо цвета, есть и другие способы сделать вашу визуализацию более наглядной и понятной: настройте на диаграмме перекрытие столбцов, уменьшите пространство между столбцами, используйте гладкий график на основе скользящего среднего, сделайте одну подпись для нескольких диаграмм, воспользуйтесь вспомогательной осью.

Если нужно сравнить показатели текущего года с показателями прошедшего, информация считывается легче, когда текущий год выводится на первый план, а прошлый остается на заднем:



Рис. 8. Сравнение показателей прошлого и текущего годов: а) стандартная диаграмма Excel, б) диаграмма с перекрытием столбцов

Чтобы добиться эффекта перекрытия, щелкните правой кнопкой мыши на любом столбце. Выберите *Формат ряда данных...* В группе *Параметры ряда* установите *Перекрытие рядов 100%*.

Глава 5. Готовые решения для различных бизнес-данных

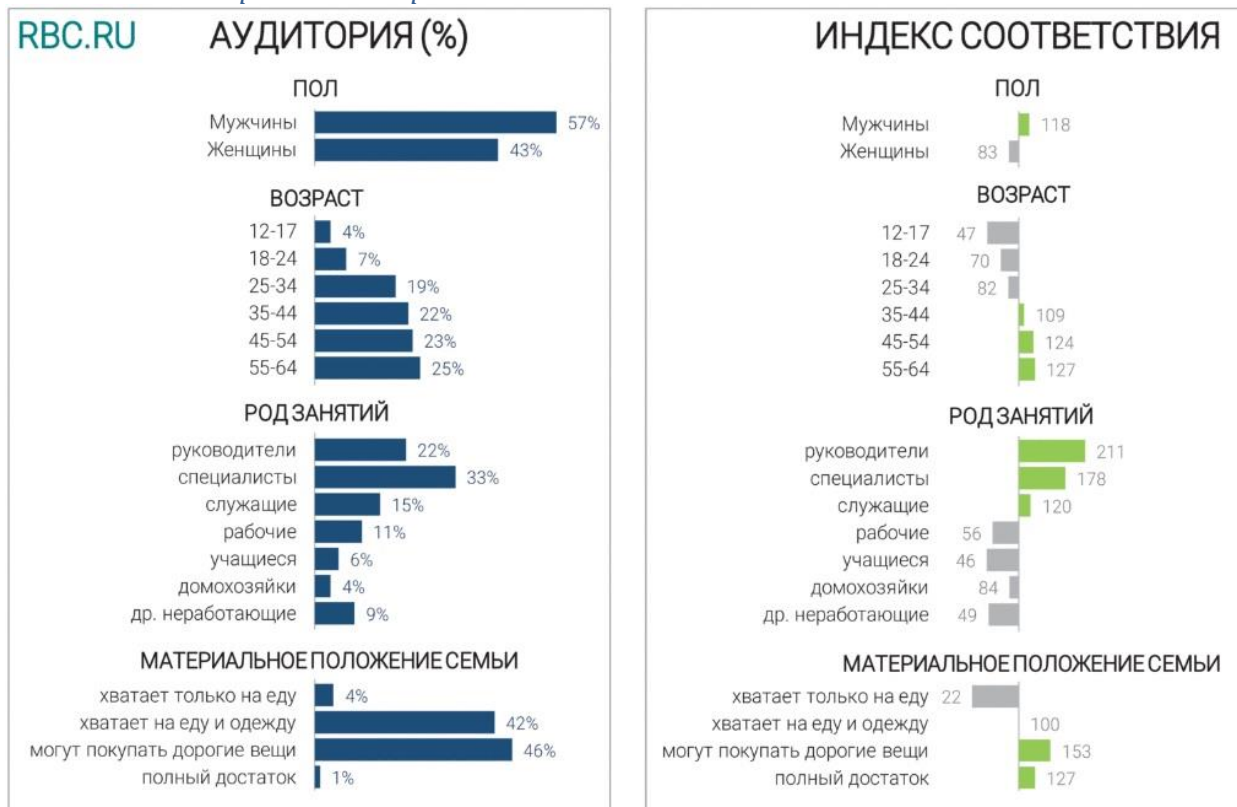


Рис. 9. Целевая аудитория на примере сайта РБК

Индекс соответствия (Affinity index) показывает, насколько данной группе свойственно посещение сайта. Если индекс больше 100, значит, свойственно, а если меньше – нет. Индекс превышает 100,

когда доля этой аудитории на сайте выше, чем в населении России в целом. Например, в населении всего 48% мужчин, а на РБК доля мужчин равна 57%. Индекс будет равен 118 (считается он по формуле $57/48 * 100$).

Глава 6. Как оформлять таблицы

Если ряды и столбцы данных хорошо структурированы, не используйте сетку. Если она всё же нужна, сделайте линии сетки светлыми. Числа должны быть с разделителями разрядов и без лишних знаков после запятой. Если применяете автоматическую поочередную заливку строк, используйте неяркие цвета. Не стоит перебарщивать, иначе все внимание будет направлено на форму, а не на содержание. Логика размещения данных в таблице может быть разная, но она должна быть. Обычно данные принято располагать от больших значений к меньшим. Выделяйте отрицательные значения красным. Если в таблице присутствует много нулевых значений или ошибок (#Н/Д, #ДЕЛ/0!), окрасьте их в бледный серый цвет (с помощью пользовательского формата чисел или условного форматирования). Выделяйте шапку и итоги. Помогайте цветом и отступами понять иерархическую структуру данных:

МАГАЗИНЫ	I КВАРТАЛ	II КВАРТАЛ
Доходы	3 500 000 р.	5 500 000 р.
Магазин1	2 000 000 р.	3 000 000 р.
Магазин2	1 500 000 р.	2 500 000 р.
Расходы	700 000 р.	1 100 000 р.
Магазин1	400 000 р.	600 000 р.
Магазин2	300 000 р.	500 000 р.
Прибыль	2 800 000 р.	4 400 000 р.
Магазин1	1 600 000 р.	2 400 000 р.
Магазин2	1 200 000 р.	2 000 000 р.

МАГАЗИНЫ	I КВАРТАЛ	II КВАРТАЛ
Доходы	3 500 000 р.	5 500 000 р.
Магазин1	2 000 000 р.	3 000 000 р.
Магазин2	1 500 000 р.	2 500 000 р.
Расходы	700 000 р.	1 100 000 р.
Магазин1	400 000 р.	600 000 р.
Магазин2	300 000 р.	500 000 р.
Прибыль	2 800 000 р.	4 400 000 р.
Магазин1	1 600 000 р.	2 400 000 р.
Магазин2	1 200 000 р.	2 000 000 р.

Рис. 10. Выделение иерархической структуры таблицы