

«BI проект закончен – что вы забыли сделать?»

Работая над созданием хранилища данных, многие забывают о главной цели проекта: упростить аналитикам работу с информацией. На извлечение, очистку и трансформацию данных тратится столько сил и ресурсов, что проектировать красивые отчёты и создавать аналитический портал становится некогда.

«Мы сдали проект, объединили десять систем, три терабайта данных, в двухчасовое окно загружаем сто гигабайт... Аналитики? У них есть данные, пусть сами и рисуют свои отчёты».

Но эти отчёты и аналитический портал — единственная часть айсберга, которая видна пользователям. Помните аналогию Ральфа Кимбалла, представляющего аналитическую систему как ресторан? Выбирая, где вы будете ужинать, вы думаете не только о качестве продуктов и мастерстве шеф-повара. Вам важно, как выглядит

зал и сервируются блюда. Приглашая аналитиков создавать отчёты самостоятельно, вы говорите им: «Проходите на кухню, берите, что понравится, овощи мы нарезали».

5 способов испортить отчёт:

- сделать логотип компании как можно больше и контрастнее;
- использовать только яркие цвета и трёхмерные графики, лучше — с градиентной заливкой;
- включить анимацию, в том числе в таблицы и графики
- создать креативный график, с которым пользователь никогда не сталкивался раньше;
- использовать насыщенную фоновую картинку — космос, джунгли или рабочий стол с иконками.

Технически спроектировать хороший отчёт несложно. Почти все современные BI-инструменты позволяют создавать детализированные отчёты, не углубляясь в программирование. Основная проблема — **визуальное представление**.

Мы привыкли к тому, что графика, которая нас окружает, захватывает внимание и доносит готовую идею

максимально наглядно. Информационный дизайн деловых журналов и рекламы выглядит завораживающе просто, мы моментально считываем сообщение автора... и именно поэтому их приёмы нельзя использовать в аналитическом отчёте. То, что украшает таблицу в журнале, будет только раздражать бизнес-аналитика:

- яркие цвета, сложный фон и эффекты отвлекают внимание от собственно данных. Необходимость работать с таким графиком несколько раз в день ужасно утомляет;
- готовое «авторское» сообщение и любые акценты в графике мешают искать новые идеи и связи между данными. Упрощенная и тем более «игровая подача» не дают аналитику делать его работу: наблюдать и делать выводы.

Зоны внимания в отчёте: вы смотрите на центральный квадрат, если он отделен от остальных линией. Затем взгляд движется по часовой стрелке.



Главные правила визуальной подачи информации нарушаются почти в каждом отчёте, созданном непрофессионалом в области инфографики. Например, в левый верхний угол, самое «ценное» место на экране, традиционно ставят рубрикатор или логотип компании — а на их месте могли бы быть данные поважнее (по крайней мере, для аналитика). Или вешают в «шапку» отчёта картинку из фотобанка — группу людей, напряжённо смотрящих на пользователя. Человек подсознательно всегда выделяет направленный на него взгляд. И, читая этот отчёт, всё время фокусируется на

«шапку с глазами», отвлекаясь от данных.

70% информации, получаемой и анализируемой всеми органами чувств — визуальные данные. В человеческом мозге запущены мощнейшие механизмы обработки визуального потока, выделения точек внимания и сравнения объектов. Научное направление визуализации информации, изучающей эти механизмы, существует уже более 30 лет.

Основы теории визуализации, к сожалению, не преподают в университетах — за исключением, пожалуй, некоторых курсов статистики. Производители программного обеспечения не только не помогают пользователям создавать читаемые отчёты — часто они даже сами провоцируют их на ошибки. На презентациях продукта им показывают трёхмерные графики с тенями и градиентной заливкой, сочными глянцевыми цветами и кнопками меню, «которые хочется лизнуть». Как эффективно использовать визуализацию для анализа данных — не показывают. Пользователи некоторое время экспериментируют с разными формами представления и возвращаются к привычным стандартам.

Примеры визуализаций, которые мы видим, решают задачу «показать другим». Умение использовать визуализацию для анализа данных остается за кадром.

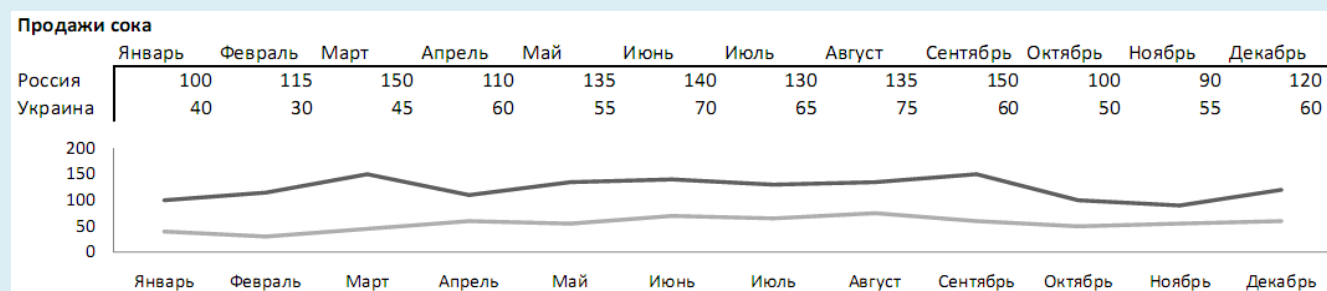


Жёсткий формат этих стандартов часто вызывает рабочие споры между аналитиками и разработчиками отчётов. Многомерная модель данных часто «уплощается» по просьбе клиента: например, ему нужны длинные «боковики» и бесконечные «шапки».

При обработке визуальной информации наша «быстрая» память может удерживать только 3-7 объектов одновременно. Цифра — один из таких объектов, именно поэтому нам так тяжело читать «непривычные» отчёты. Аналитик, работающий с одними и теми же таблицами много лет, мельком взглянув на гигантскую, запутанную для стороннего наблюдателя форму, легко может найти несоответствие в данных или «ненормальные» цифры. Его взгляд просто «натренирован» — именно поэтому некоторые клиенты требуют подгонять данные под формат.

Но вместо ежемесячных цифр вы можете нарисовать линию, показывающую годовые продажи продукта — и заменить тем самым двенадцать объектов внимания на один. Такая визуализация может сильно упростить работу аналитика, чьё восприятие не «заточено» на определённую форму отчёта.

В таблице представлены одни и те же данные. 2 линии дают гораздо более наглядную картину, чем 24 цифры.



Найдите пятёрки в этих таблицах. Анализировать данные проще, если они выделены, но слишком много акцентов ведут к хаосу

Найдите все пятёрки

1	9	0	6	5	6	4	2	3	6	7	8
3	5	1	0	8	1	6	4	8	3	0	4
9	10	4	2	4	3	1	6	9	10	5	1

Признак – размер

1	9	0	6	5	6	4	2	3	6	7	8
3	5	1	0	8	1	6	4	8	3	0	4
9	10	4	2	4	3	1	6	9	10	5	1

Много признаков

1	9	0	6	5	6	4	2	3	6	7	8
3	5	1	0	8	1	6	4	8	3	0	4
9	10	4	2	4	3	1	6	9	10	5	1

Простой пример: три таблицы, в каждой из которых нужно найти пятёрки. Чтобы ваш взгляд смог выделить предмет, он должен обладать набором признаков — определённым положением в пространстве, размером, цветом, наклоном, формой и проч. Если сочетать эти признаки, игнорируя правила восприятия, вы превратите любую таблицу в хаос.

И, оформляя продукт, не забывайте о тех, для кого вы его делаете. Разработчикам отчётов важно понимать, как именно люди используют результаты их труда. На одном из проектов нашей команде

повезло сидеть ровно за спинами аналитиков. Полдня мы подглядывали за тем, как люди используют наши отчеты — и этого оказалось достаточно, чтобы понять: как именно расположить фильтры, какие значения выбрать «по умолчанию», по каким признакам позволить сортировать данные. Мы увидели мелочи, на которые никто не обращает внимания при разработке и тестировании: «В отчете же есть нужный фильтр — ну и что, что к нему немного неудобно тянуться». И сократили время работы с отчетом в несколько раз. Было: «Запустил отчет, нашёл нужные фильтры, перевыбрал значения, получил результат». Стало: «Запустил отчет, (если нужно, перевыбрал фильтр), получил результат».

Такие «командировки в бизнес» — мощный стимул для разработчика. Приятно работать, если отчёты перестают быть «очередной простыней» и становятся ответами на вопросы.

5 книг по визуальному анализу данных

- [Stephen Few, Information Dashboard Design: The Effective Visual Communication of Data](#)
- [Stephen Few, Now You See It: Simple Visualization Techniques for Quantitative Analysis](#)
- [Edward Tufte, The visual Display of Quantitative Information](#)
- [Colin Ware, Visual Thinking for Design](#)
- [Dan Roam, The Back of the Napkin: Solving Problems and Selling Ideas with Pictures](#)