

Негласные правила визуализации

(и когда их нарушать)

29 января 2020 г.

Кайзер Фунг

Визуализация данных похожа на решение пазла. Чтобы добиться успеха, вам следует знать некоторые вещи заранее: сцена, которая будет раскрыта по частям, — это история, которую вы хотите передать с помощью графика данных. Вам также нужно понимать, что у вас есть в распоряжении : части пазла — это набор данных перед вами. А затем ваша задача — собрать пазл или собрать элементы вашего графика, чтобы рассказать историю ваших данных.

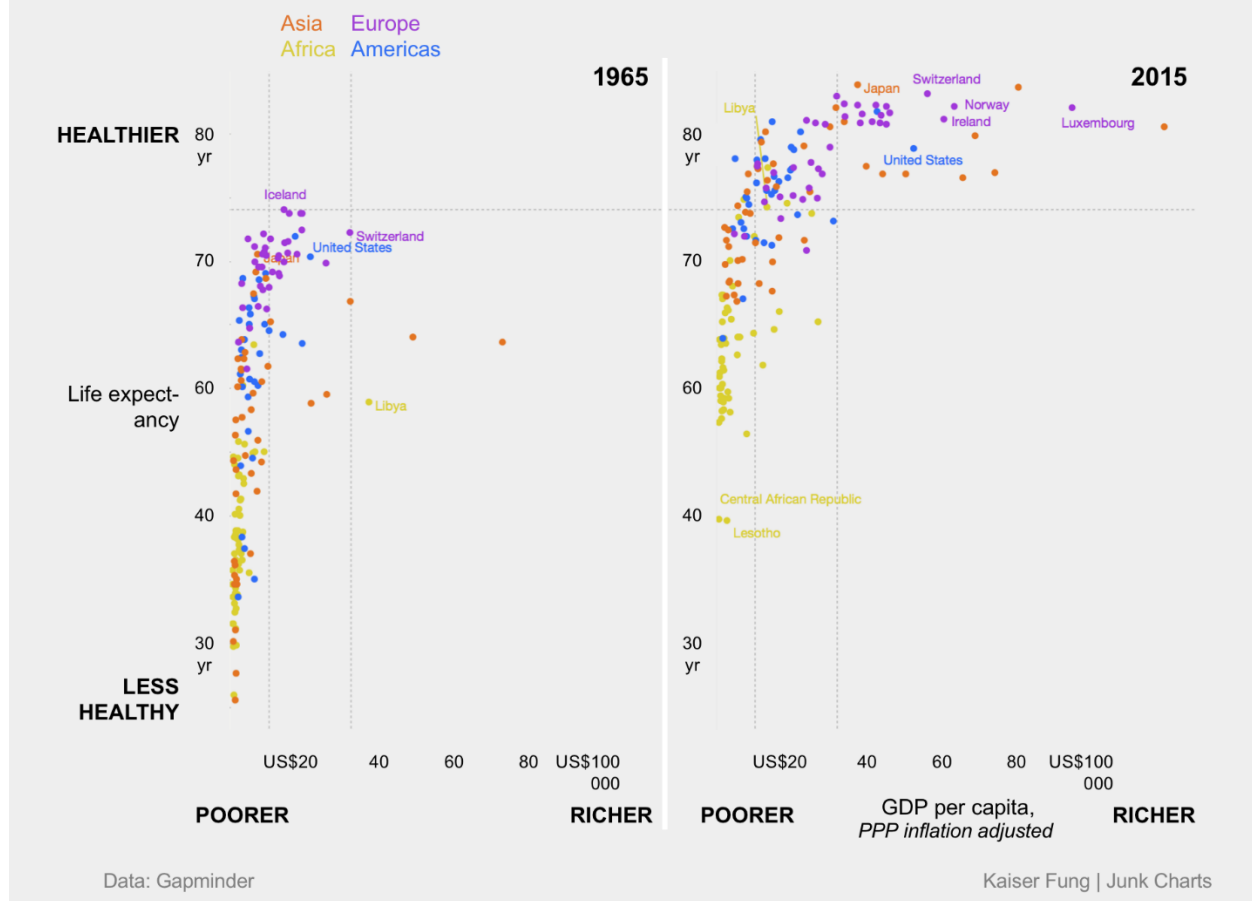
Визуализация данных — это выбор. Вместо слов мы выбираем изображения. Как гласит поговорка: одна картинка стоит тысячи слов. Воспринимаемая сила визуальной среды проистекает из ее эффективности и многомерности.

Рассмотрим следующий обзор состояния здоровья и благосостояния в мире, составленный на основе данных, собранных в рамках проекта Garminder :

Последние 50 лет (1965–2015) ознаменовались огромным прогрессом как в сфере здоровья, так и в сфере благосостояния стран. Рост средней продолжительности жизни был поистине впечатляющим. В 1965 году Исландия занимала первое место среди всех стран, где средний житель жил до 74 лет. К 2015 году почти вся Европа, большая часть Америки, большинство стран Азии и даже некоторые страны Африки достигли или превысили этот показатель.

За последние 50 лет большая часть мира также стала богаче, если измерять по ВВП на душу населения с поправкой на паритет покупательной способности и инфляцию. В 1965 году большинство стран имели доход ниже 5000 долларов США на человека; к 2015 году этот показатель вырос более чем до 10 000 долларов. Швейцария с доходом в среднем 32 000 долларов на человека в 1965 году оставалась одной из самых богатых стран Европы, однако к 2015 году её обогнали Ирландия, Норвегия и Люксембург. Многие африканские страны также стали богаче, за исключением Ливии, которая оказалась заметным исключением. Теперь давайте посмотрим на визуализацию этих данных в виде пары диаграмм рассеяния: сила визуального средства очевидна.

Tremendous global gains in health and wealth over 50 years



Изображение: Диаграмма, показывающая существенный рост здоровья и благосостояния во всем мире в период с 1965 по 2015 год.

Разум легко различает различные темы для разговора, подробно описанные выше. В тексте информация поступает по крупницам в предписанной последовательности. На картинках наши глаза блуждают, выискивая информацию по нескольким измерениям одновременно. Познание направляется элементами дизайна, такими как опорные линии, легенды, метки данных и аннотации. Следует отметить, что богатство визуальной среды позволяет выходить на поверхность сложным отношениям, которые при вербальном выражении приводят к длинным, перегруженным оговорками предложениям.

Эффективность и многомерность визуальной среды возникают из набора соглашений и правил, которые упорядочивают коммуникации между производителями визуализации данных и ее потребителями. Эти соглашения и правила часто негласны: это визуальный эквивалент высказывания «это само собой разумеется» .

Представьте себе, если бы эта глобальная диаграмма здоровья и благосостояния была снабжена дополнительным полем «Как читать эту диаграмму», как показано ниже:

How to Read this Chart

On this chart, you find average life expectancy, an indicator of a nation's health, depicted on the vertical axis; and income per capita, an indicator of a nation's wealth, measured by GDP per capita, PPP inflation adjusted, depicted on the horizontal axis. Life expectancy is measured in years, and shown in the range between 35 and 80 years, in increments of 10 years. Per-capita income in U.S. dollars is shown in the range between US\$0 and US\$120,000, with tick marks every US\$10,000 and labels every US\$20,000. Each country is represented by a dot in the two-dimensional surface: the health indicator of each country can be read off the vertical axis by eyeballing a horizontal line to it while the income indicator of each country can be similarly read off the horizontal axis by dropping a vertical line to it. For example, the United States in 1965 had average life expectancy of about 70 years, and income of about US\$20,000 per capita. Each country's dot is colored according to which continent it belongs to. Purple dots are European countries; blue dots, American; orange dots, Asian, and yellow dots, African. Two scatter plots are placed side by side. The left plot shows the data for 1965 while the right plot shows data for 2015. Comparing the positions of the dots across the two plots reveals the change in health and wealth in 50 years. Such comparisons can be made for individual countries, where labeled, and for continents, as blobs of dots. Two reference lines are drawn vertically at US\$10,000 of per-capita income, and US\$32,000, which was Switzerland's number for 1965, when it was the richest nation in Europe. Comparing the cluster of purple dots on the left and on the right of the US\$32,000 reference line shows the growth in income in Europe during the 50 years. Comparing the number of dots to the left and to the right of the US\$10,000 reference line reveals the tremendous growth in income across the world.

Стоп! Вы хотите накричать на меня: большинство этих слов не нужны. Ваше возражение обосновано. Включение в текст вставки «Как это читать» умаляет преимущества визуального носителя. Длинные инструкции устраняются, когда дизайнеры следуют определенным соглашениям и правилам, которые интуитивно понятны читателям. Это само собой разумеется.

В оставшейся части этого Long Read мы выделим основной набор соглашений и правил, которые должны направлять наше производство графики данных. Ссылки, перечисленные в конце, дают дальнейшее объяснение этих и других правил. Глава 1 книги информационного дизайнера Альберто Каиро « Как лгут диаграммы » весьма актуальна, поскольку он описывает, как потребители должны читать диаграммы с точки зрения производителя. Альберто подчеркивает возможное существование «ментальных моделей» визуализации данных, так что визуальные коммуникации успешны, когда модель дизайнера сходится с моделью читателя.

Большинство соглашений и правил в визуализации данных не являются уникальными — в некоторых случаях сосуществуют конкурирующие, противоречивые соглашения. Правила — например, то, как мы обращаемся с цветами — со временем развиваются по мере совершенствования инструментов. У каждого соглашения есть свое исключение: когда наш

дизайн намеренно противоречит правилу, мы обращаем внимание наших читателей на это отклонение, в том числе, когда это уместно, предоставляя поле «Как читать эту диаграмму».

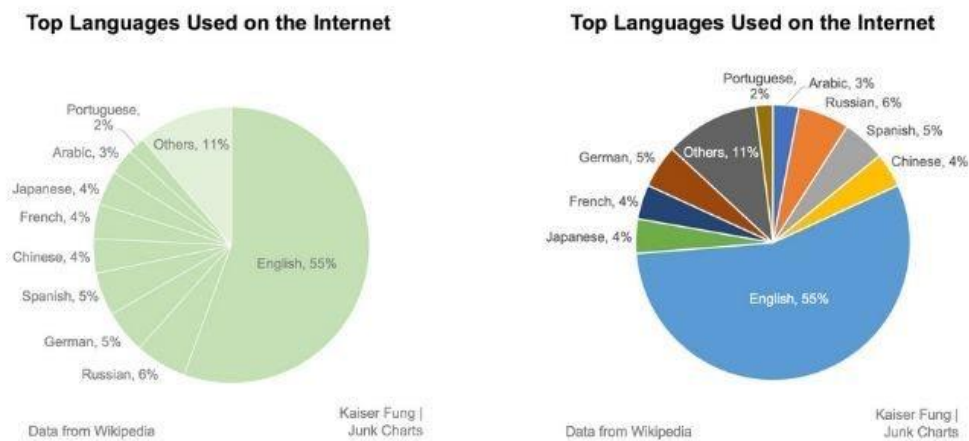
В своей книге «Грамматика графики» Леланд Уилкинсон различает эстетику — кодирование данных в геометрические объекты — от руководств, которые помогают пониманию. Следуя этому различию, я организую соглашения и правила визуализации данных в две группы. В каждом разделе два отображения одних и тех же данных сопоставлены, одно соответствует, а другое отклоняется от выделенного соглашения или правила, чтобы раскрыть обоснование этих лучших практик.

Конвенции по эстетике

Круговые диаграммы

Круговая диаграмма продолжает существовать, несмотря на то, что является наиболее критикуемой экспертами по визуализации данных формой диаграммы. Некоторые круговые диаграммы пригодны для использования, если они следуют соответствующим соглашениям.

Давайте рассмотрим пример, взятый из заметки разработчика визуализации данных Ксана Грегга . Эти две круговые диаграммы отображают языки, используемые в Интернете:



Пример первый: конформная круговая диаграмма (слева). Пример второй: расходящаяся круговая диаграмма (справа).

В первом примере читателям сообщается, что английский язык используется более чем в половине интернета, в то время как на каждый из шести других языков, от русского до японского, приходится около пяти процентов.

При построении этой круговой диаграммы я следовал ряду соглашений: а) использовать разумное количество срезов, объединяя второстепенные категории при необходимости б) упорядочить срезы по размеру от самого большого к самому маленькому с) поместить срез «Другое» в конец последовательности, независимо от схемы упорядочивания д)

расположить первый и самый большой срез напротив верхнего вертикального радиуса и расположить остальные срезы по часовой стрелке е) изменять цвета только в том случае, если цвета кодируют данные. В этом случае я использовал более светлый оттенок для среза «Другое», сигнализируя о том, что он один состоит из нескольких языков и что это наименее важный срез на диаграмме.

«Большинство соглашений и правил визуализации данных не являются уникальными — в некоторых случаях сосуществуют конкурирующие, противоречивые соглашения.»

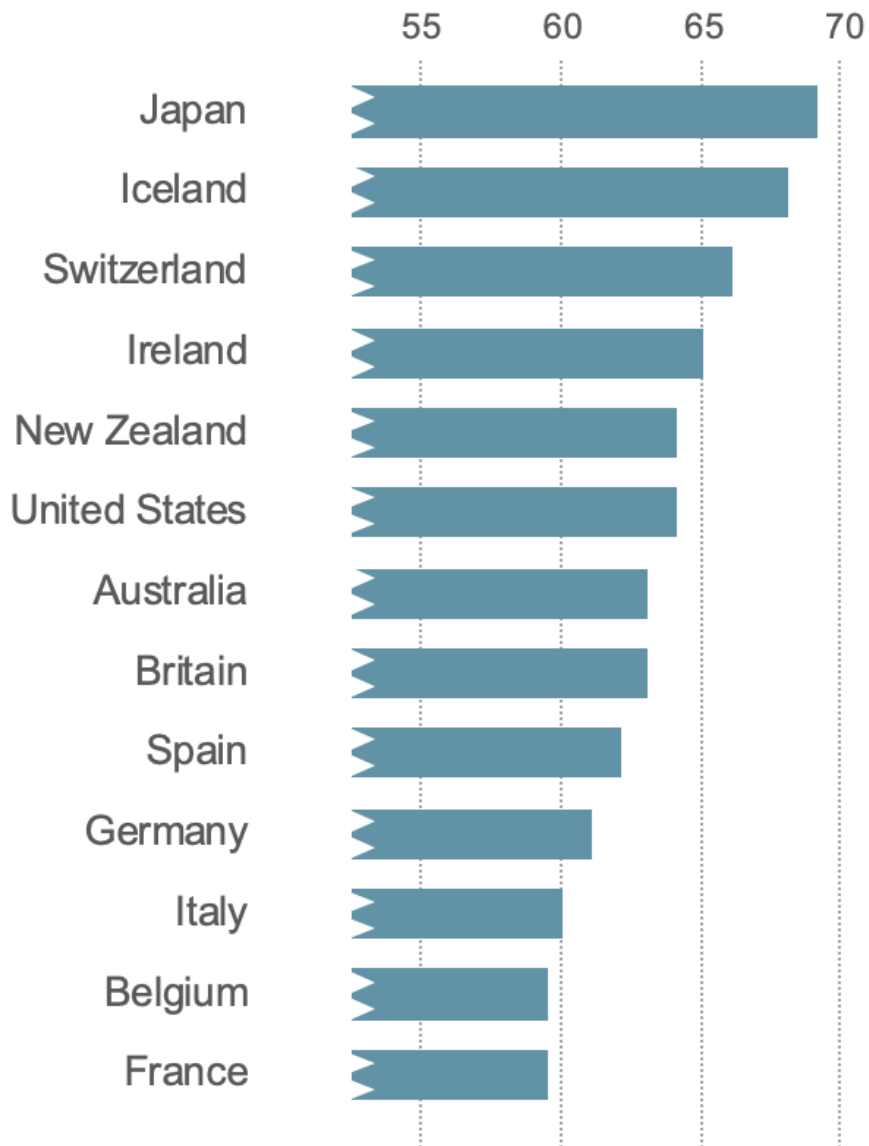
Эти правила негласны. Дизайнер вызывает их молча, а читатель применяет их интуитивно. Когда такие правила игнорируются, требуется больше времени, чтобы переварить круговую диаграмму. Взгляните на пример расхождения, в котором самый большой сектор круга расположен под случайным углом, другие секторы идут в случайном порядке, и каждому сектору назначается произвольный цвет. Когда создатель диаграммы отклоняется от соглашений, читатель должен потратить время, чтобы выяснить логику дизайна.

Гистограммы

Основное соглашение на столбчатой диаграмме (и, соответственно, на столбчатой диаграмме) — это правило «начать с нуля», которое гласит, что нижний предел оси значений должен быть установлен на ноль. Наш следующий пример, адаптированный из **The Economist**, — это образец, который не следует этому соглашению.

Working culture

Effective retirement age for men



Adapted from: The Economist, 26 Nov 2005.

Kaiser Fung | Junk Charts

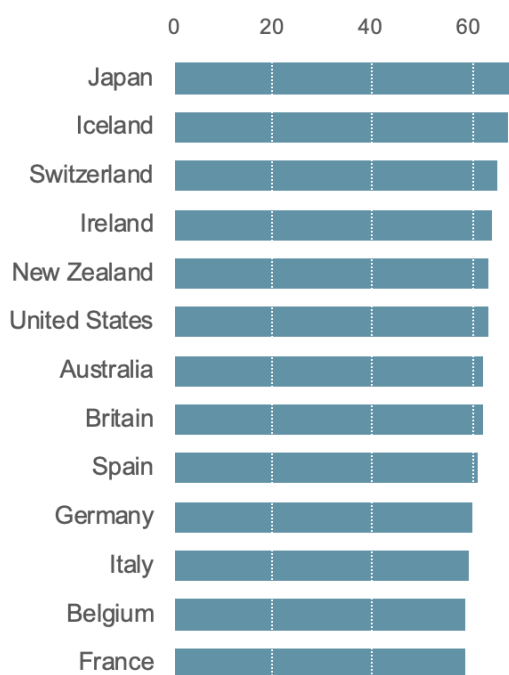
Пример первый: расходящаяся столбчатая диаграмма

На этой диаграмме читатель понимает, что пенсионный возраст в Швейцарии в два раза больше, чем во Франции, поскольку швейцарская полоса в два раза шире французской. Последняя строка не может быть правдой, и она не является правдой: швейцарский пенсионный возраст всего на 10% выше французского. Когда ось значений расширяется до нуля, как в нашем примере соответствия, соотношение ширин полос восстанавливается до соотношения данных.

Внимательный читатель замечает, что дизайнер примера 1 поместил символ разрыва на левом краю каждого столбца, сигнализируя, что его ширина усечена (более чем наполовину). Таким образом, создатель сознательно бросает вызов эстетической конвенции столбчатых диаграмм. Признание не исправляет искажение, вызванное усечением, что приводит к вероятной неверной интерпретации.

Working culture

Effective retirement age for men



Data: The Economist, 26 Nov 2005.

Kaiser Fung | Junk Charts

Пример два: соответствующая столбчатая диаграмма

Конечно, обновленная столбчатая диаграмма все еще неадекватна. Более эффективное отображение достигается путем переключения на точечный график, как показано в примере три. Другой эффективный вариант отображения — сосредоточиться на разрывах

между фактическим и официальным пенсионным возрастом, как показано в примере четыре. Оба эти дизайна работают вокруг правила «начать с нуля».

Working culture

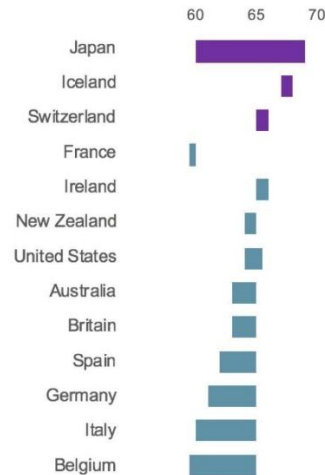
Effective retirement age for men



Data: The Economist, 26 Nov 2005.
Kaiser Fung | Junk Charts

Working culture

Japanese men work about 10 years **after** reaching retirement age on average. In most countries, men retire **before** the retirement age.

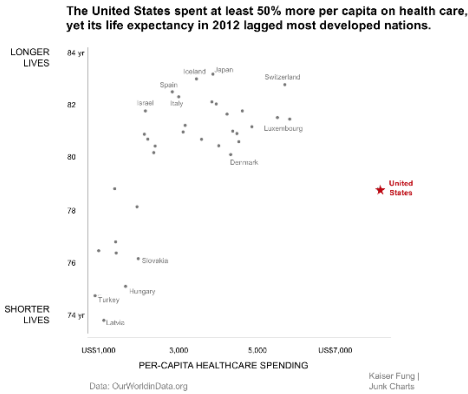


Data: The Economist, 26 Nov 2005.
Kaiser Fung | Junk Charts

Пример третий: точечный график (слева). Пример четвертый: отображение пробелов (справа).

Диаграммы рассеяния

Диаграмма рассеяния отображает каждую единицу данных как точку на поверхности, охватываемой двумя осями. Горизонтальное (x) и вертикальное (y) положение точки кодирует две переменные. Форма облака точек визуализирует природу корреляции между двумя переменными. Наслаждайтесь великолепной диаграммой рассеяния в нашем следующем примере, которая выделяет Соединенные Штаты как страну-выброс, в которой чрезмерные расходы на здравоохранение не привели к ожидаемому росту продолжительности жизни.



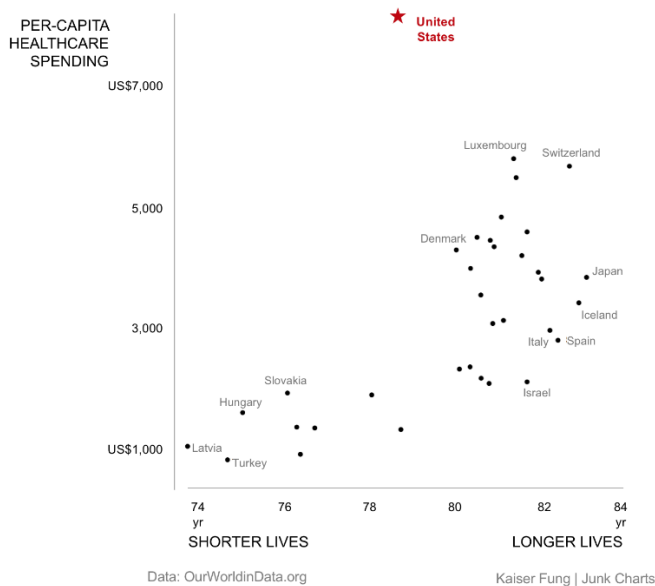
Пример первый: конформная диаграмма рассеяния

Соглашение определяет, какую переменную размещать на какой оси. В этом примере расходы на здравоохранение на душу населения считаются движущей силой результатов в области здравоохранения. По соглашению расходы на здравоохранение (объясняющая переменная) кодируются как x , а ожидаемая продолжительность жизни (результат) как y .

Для сравнения, в нашем расходящемся примере оси x и y поменялись местами. Его визуальная форма — отражение тех же данных по диагонали 45 градусов.

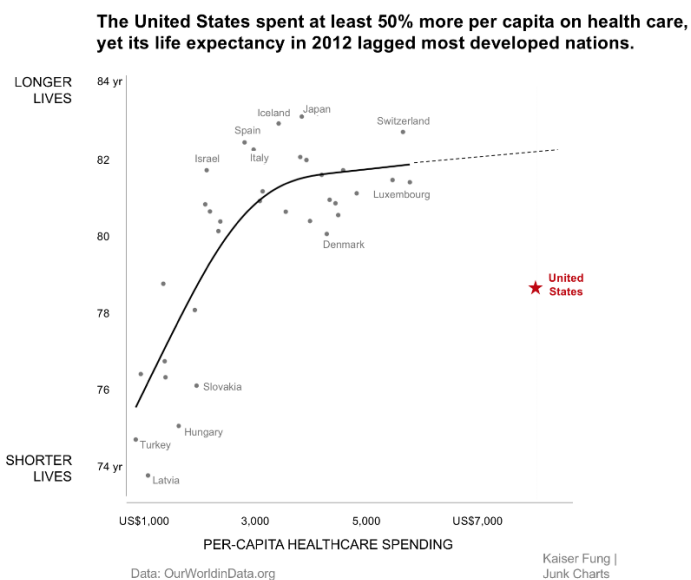
Поскольку дизайн узурпирует условности, многие читатели, особенно те, кто имеет подготовку в областях STEM, отреагируют с замешательством и даже раздражением. Хотя для этого правила нет четкого императива дизайна, преобладает сильное научное обоснование.

The United States spent at least 50% more per capita on health care, yet its life expectancy in 2012 lagged most developed nations.



Пример два: расходящаяся диаграмма рассеяния

Стандартным дополнением к диаграмме рассеяния является линия регрессии (также ошибочно называемая «линией тренда» ведущей на рынке программой для работы с электронными таблицами Excel). Регрессионный анализ количественно определяет корреляцию между двумя переменными, отображаемыми диаграммой рассеяния. Линия регрессии создается таким образом, чтобы минимизировать среднее расстояние между линией и облаком точек. Наша следующая диаграмма рассеяния включает линию регрессии.



Пример первый: соответствующий график рассеяния с линией регрессии

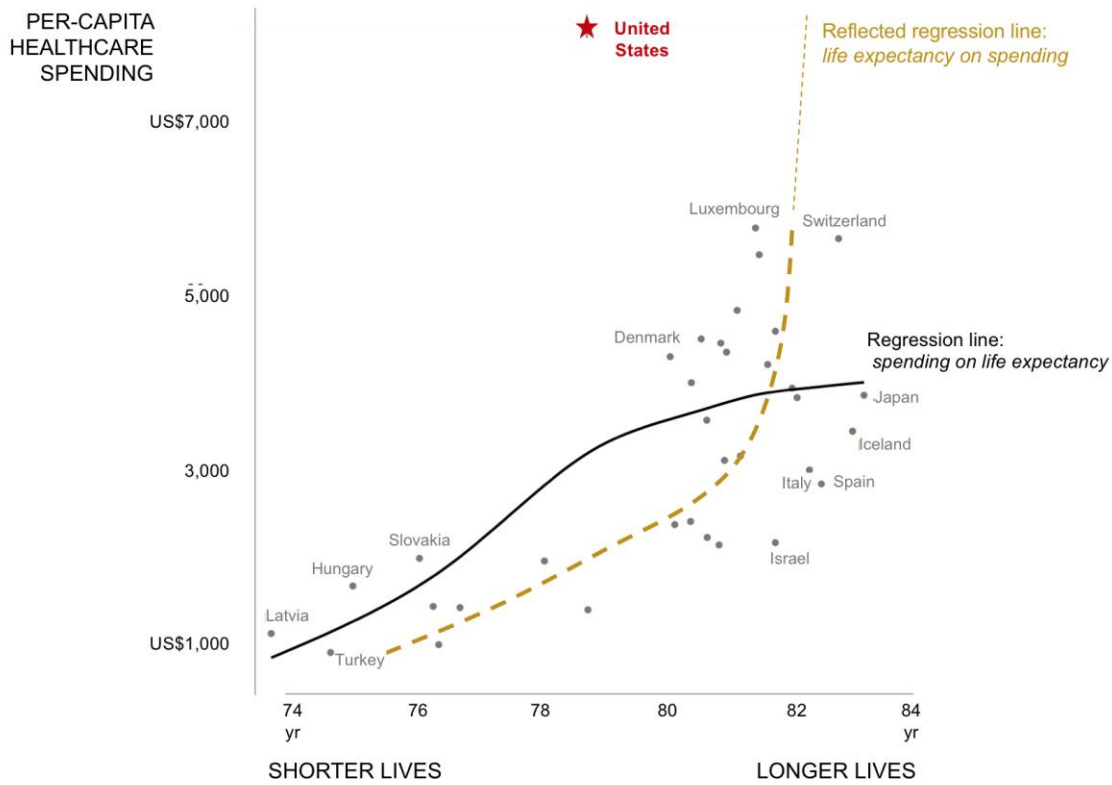
Самое важное, что расстояние между заданной точкой и линией регрессии измеряется по вертикали, а не по горизонтали. Это вертикальное разделение также является подсказкой, по которой читатель узнает ключевое сообщение графика: американцы должны были бы наслаждаться ожидаемой продолжительностью жизни более 82 лет, учитывая их уровень расходов, если бы дополнительные расходы трансформировались в дополнительные годы жизни с той же скоростью, что и в других странах.

Перестановка осей x и y не отражает линию регрессии (как это происходит с точками). Ведь то, что минимизирует вертикальные расстояния между точками и линией, не минимизирует горизонтальные расстояния. Как показано в примере два ниже, где я поменял местами оси примера один выше, линия регрессии x на y не совпадает с отраженной линией регрессии y на x.

Примечательно, что эта конвенция не диктует, какая переменная должна быть объясняющей переменной, а какая — переменной результата. После того, как дизайнер определит эти роли, конвенция определяет, какая переменная будет назначена на какую ось. То есть, пример два уместен, если ожидаемая продолжительность жизни предлагается

в качестве объяснения изменчивости расходов на здравоохранение. Подробнее об этой теме читайте в моем блоге .

Regressing spend on life expectancy does not equal the reflection of regressing life expectancy on spend.



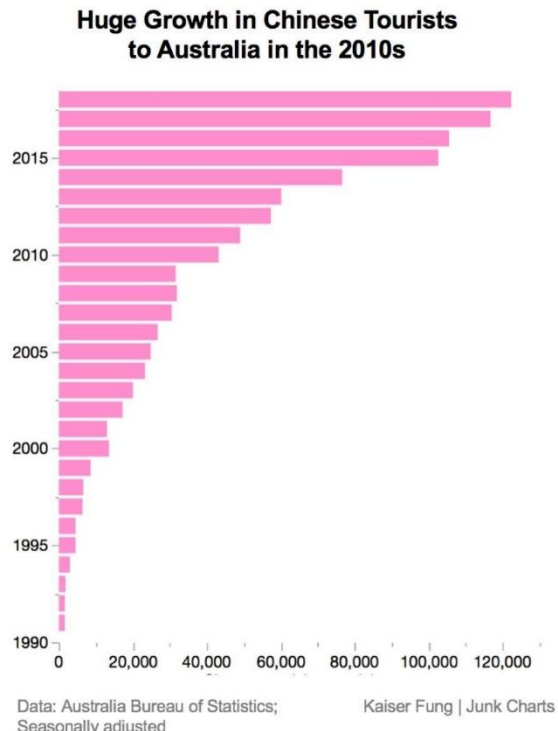
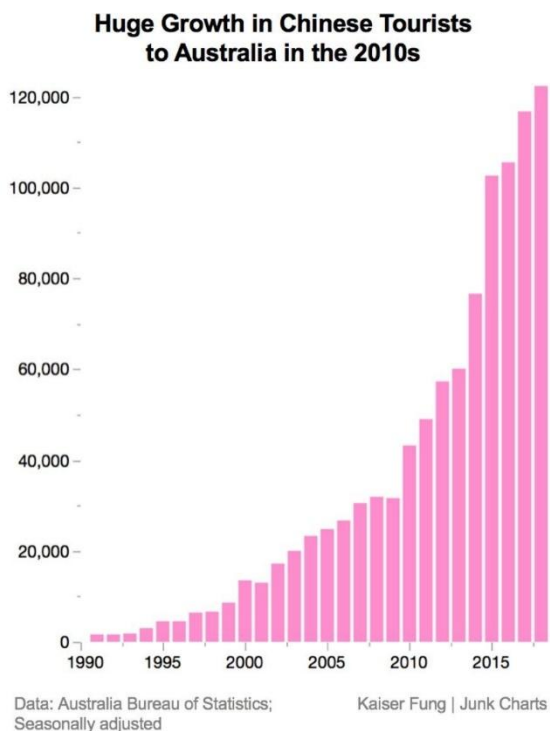
Data: OurWorldinData.org

Kaiser Fung | Junk Charts

Пример два: расходящаяся диаграмма рассеяния с линией регрессии и наложение отраженной линии регрессии, если поменять оси местами.

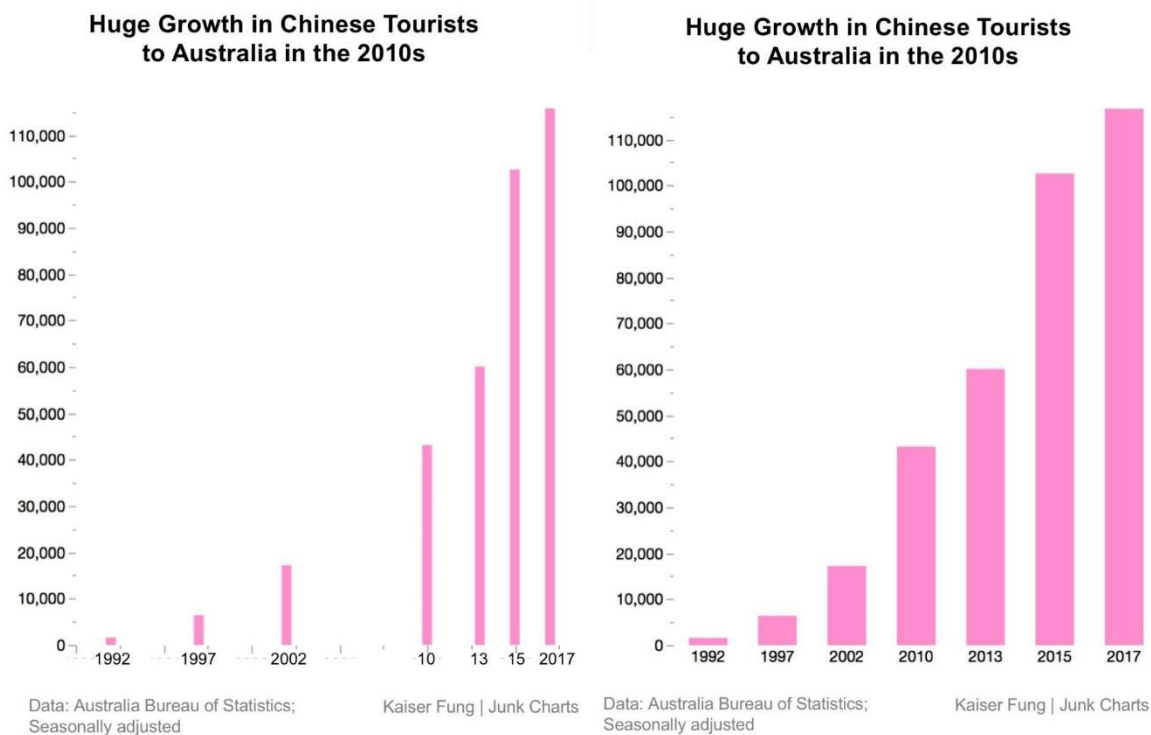
Графики временных рядов

Естественное место для переменной времени, такой как годы, месяцы и даты, находится на горизонтальной оси. По соглашению время идет слева направо (замените справа налево в странах с направлением справа налево (RTL)). Следующая пара диаграмм показывает быстрый рост числа китайских туристов, посещающих Австралию. Сравните пример один, в котором время идет слева направо, с примером два, в котором время идет снизу вверх. Соглашение слева направо является негласным правилом, общим для производителей и потребителей графики данных в культурах, где читают слева направо. Отклонение от этого правила всегда замедляет познание.



Пример первый: график согласованного временного ряда (слева). Пример второй: график расходящегося временного ряда (справа).

Еще одно правило для диаграмм временных рядов — пропорциональный интервал. Когда данные собираются с неравными интервалами, деления на оси времени должны имитировать неравномерность. В противном случае диаграмма искажает темпы роста. В другом расходящемся примере ниже тенденция роста выглядит линейной, а не «хоккейной клюшкой», артефактом применения равномерного интервала к неравномерно распределенным данным.



Пример три: еще один график согласованного временного ряда (слева). Пример четыре: еще один график расходящегося временного ряда (справа).

Цветовая кодировка

В эпоху социальных сетей цвет стал любимым дополнением к любой графике данных. Вот несколько основных правил, регулирующих применение цвета:

а) Установите ограничение на количество цветов. Как предлагает Дона Вонг в *The Wall Street Journal Guide to Information Graphics*, «впускайте цвета изящно, как если бы вы принимали родственников в своем доме».

б) Тот же цвет, те же данные; разница в цвете должна отражать разницу в данных. Это правило дисквалифицирует произвольное назначение цветов.

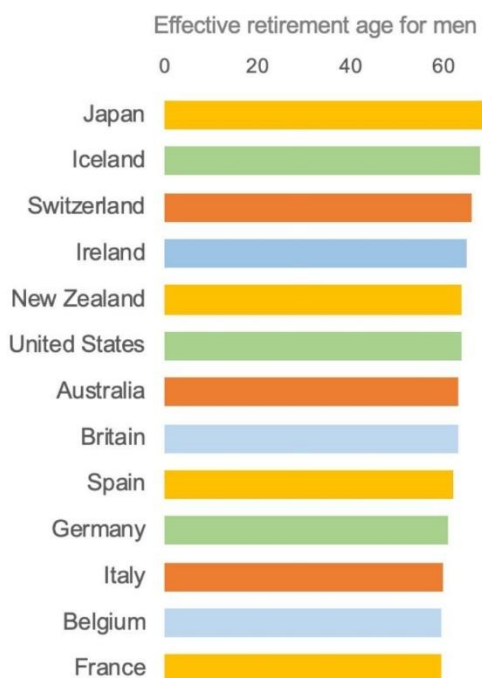
в) Используйте определенные пары цветов с осторожностью, поскольку они нагружены смыслом. В бизнес-сообществе черный цвет является положительным, а красный — отрицательным, но в некоторых культурах черный цвет является зловещим, а красный — благоприятным. Для тепловых карт красный цвет является горячим, а синий — холодным,

в то время как в политике США пара цветов красный-синий обозначает две основные политические партии. Как я отметил в начале, условности иногда конфликтуют.

г) Многие авторы рекомендуют делать диаграммы удобными для читателей, страдающих дальтонизмом, например, просматривая версию в оттенках серого.

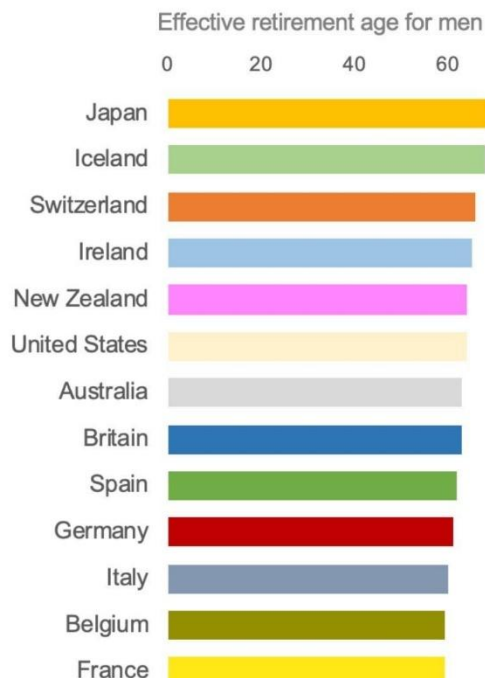
Давайте рассмотрим два варианта нашей столбчатой диаграммы, которая показывает пенсионный возраст в 13 странах. Слева столбцам присвоены бессмысленные цвета, что расходится с правилом b. Читатель путается из-за ложного сигнала, бесплодно ища данные за цветовой схемой. Справа дизайн использует уникальный цвет для каждого столбца, что противоречит правилу a. Здесь цветовая палитра становится отвлекающим фактором, снижая скорость понимания.

Meaningless colours



Data: The Economist, 26 Nov 2005.
Kaiser Fung | Junk Charts

Too many colours

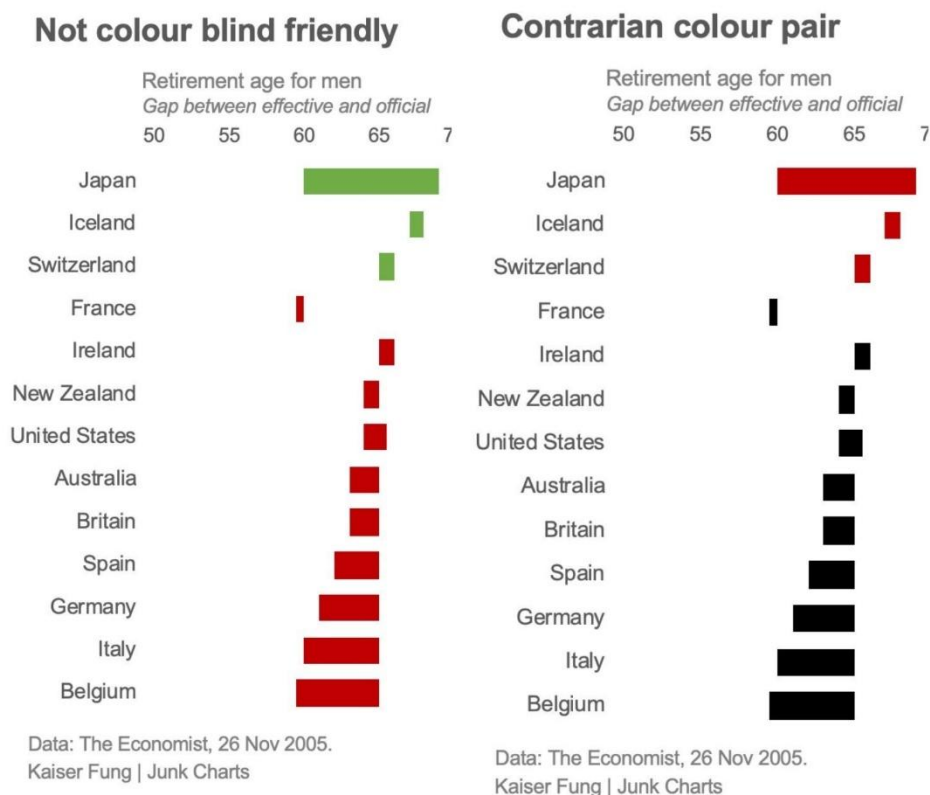


Data: The Economist, 26 Nov 2005.
Kaiser Fung | Junk Charts

Пример первый: расходящиеся цветовые соглашения

Два графика из примера два получены из столбчатой диаграммы, которая отображает разрывы в пенсионном возрасте. По выбору дизайнера положительный разрыв означает,

что фактический пенсионный возраст превышает официальный пенсионный возраст. Диаграмма слева применяет зеленый цвет к положительным разрывам, а красный — к отрицательным. Правило d советует не объединять красные и зеленые оттенки, потому что читатель, страдающий дальтонизмом, не сможет их различить. Диаграмма справа кодирует положительные разрывы красным цветом, а отрицательные — черным. Такой выбор цветов сбивает с толку из-за традиции, особенно популярной в бизнесе, использовать красные чернила для отрицательных чисел.



Пример два: более расходящиеся цветовые соглашения

Соглашения о путеводителях

Дизайнеры диаграмм добавляют направляющие, такие как легенды, оси, линии сетки и метки, с явной целью ускорения познания. Как указали Эдвард Тафти и другие эксперты, такие направляющие иногда дают обратный эффект, если они плохо реализованы. В ответ был разработан большой набор соглашений и правил.

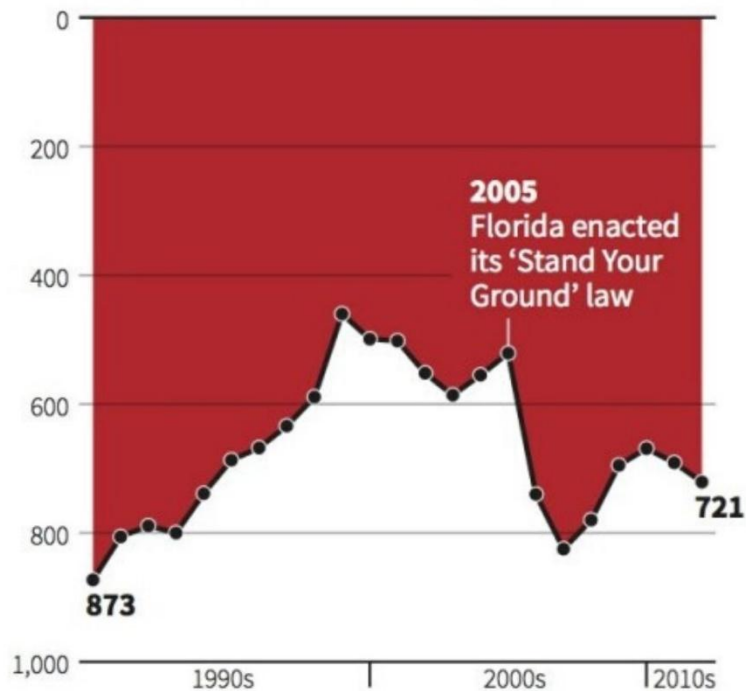
Топоры

Само собой разумеется, что оси имеют канонические направления. На вертикальной оси большие значения располагаются над меньшими, а на горизонтальной оси большие значения располагаются справа от меньших (за исключением стран RTL). Узурпация этих правил приводит к бессмысленным диаграммам.

В 2014 году агентство Reuters опубликовало следующую линейную диаграмму, которая сразу же вызвала бурю твитов в сообществе специалистов по визуализации данных.

Gun deaths in Florida

Number of murders committed using firearms



Source: Florida Department of Law Enforcement

C. Chan 16/02/2014

REUTERS

Изображение: Диаграмма, показывающая влияние закона «О стой на земле» во Флориде на количество смертей от огнестрельного оружия.

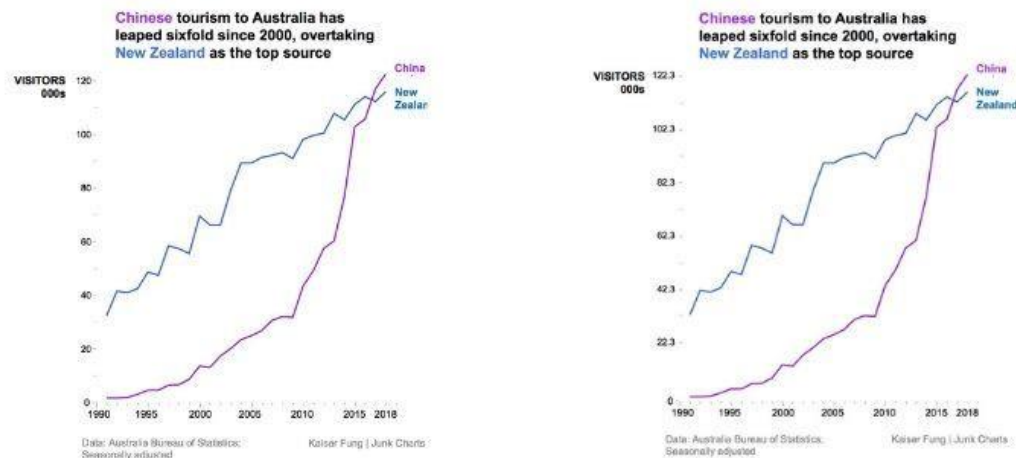
Закон Stand Your Ground, легализующий применение смертоносной силы для самозащиты, как ожидалось, должен был усугубить насилие с применением огнестрельного оружия, и тем не менее эта диаграмма отобразила тенденцию к снижению после его принятия в 2005

году. Обнаружив инверсию вертикальной оси, читатели поняли, что их интуиция «чем ниже, тем меньше» была неуместна. Реакция была уничтожающей. Профессор колледжа пожаловался: «Это настолько глубоко вводит в заблуждение, что мне противно показывать это вашим глазам». Этот шквал твитов показывает, почему дизайнеры должны следовать правилам, если только нет веской причины не делать этого.

Если задействовано измерение времени, принято размещать время на горизонтальной оси. В диаграмме рассеяния переменная результата должна быть закодирована на вертикальной оси, а объясняющая переменная — на горизонтальной оси.

Два других негласных правила — о пределах и делениях — определяют дизайн осей. Разумные пределы выбираются для удаления избыточного пустого пространства с поверхности построения. Деления должны попадать на легко интерпретируемые приращения и значения; например, последовательность [0, 20, 40, 60, ..., 120] вместо [2, 22, 42, 62, ..., 122] или, что еще хуже, [2.3, 22.3, 42.3, 62.3, ..., 122.3].

Следующая пара диаграмм идентична, за исключением меток осей. Они обе передают сообщение о том, что число китайских туристов, въезжающих в Австралию, превысило число туристов из Новой Зеландии с 2017 года. Более точные метки во втором примере сложнее понять.



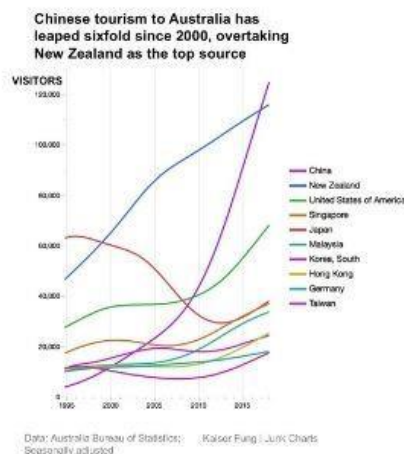
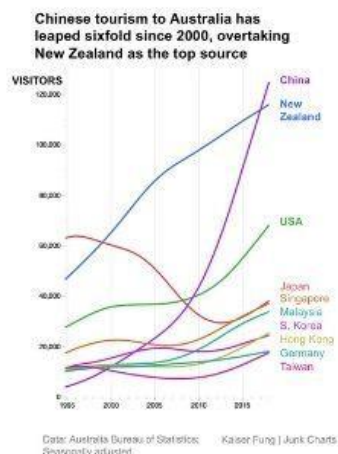
Пример первый: совпадающие оси (слева). Пример второй: расходящиеся оси (справа).

Легенды

Почти все диаграммы включают легенду. Цветовая легенда обычно встречается на линейных диаграммах, столбчатых диаграммах, круговых диаграммах, пузырьковых диаграммах и т. д. Первое правило для легенд — не использовать легенду, если возможны прямые подписи.

На линейной диаграмме с пучком линий обычно предпочтительнее размещать метки рядом с линиями, а не внутри поля легенды. Само собой разумеется, что цвета в легенде должны соответствовать цветам на самой диаграмме один к одному, и что порядок появления должен имитировать порядок на диаграмме. Популярное программное обеспечение, такое как Excel, часто нарушает это правило, показывая обратный порядок внутри поля легенды, когда элементы появляются на основной диаграмме.

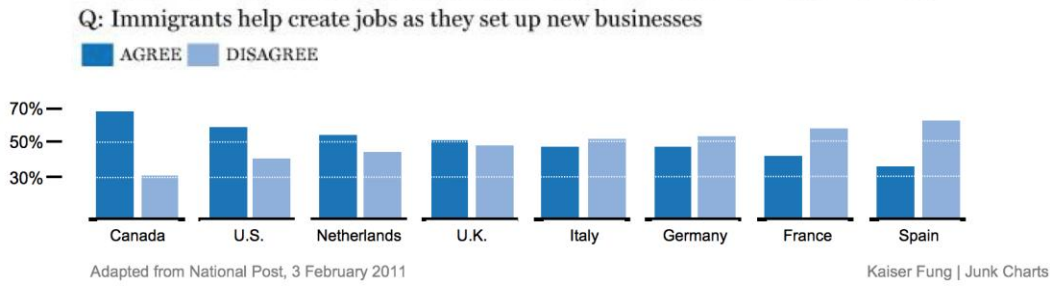
На этих графиках, демонстрирующих поразительный рост числа китайских туристов, посещающих Австралию, метки линий соответствуют рейтингу количества туристов в 2018 году, последнем году с данными. Пример первый, с прямой маркировкой, уменьшает количество покачиваний головой, чтобы связать легенду с линией.



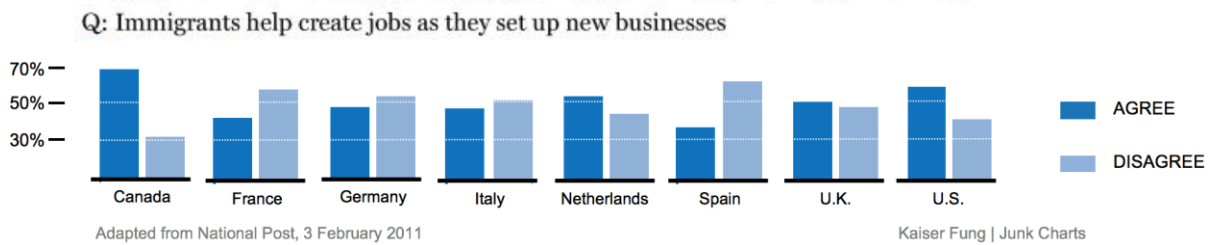
Пример первый: совпадающие легенды (слева). Пример второй: расходящиеся легенды (справа).

Для столбчатых или линейчатых диаграмм, если невозможно избежать поля легенды, принято размещать его над диаграммой под заголовками, поскольку читатели полагаются на информацию для интерпретации графика. Порядок категорий должен отражать ориентацию данных.

National Post отобразил результаты опроса об отношении к иммигрантам в серии парных столбчатых диаграмм, одна из которых воспроизведена в примере три. Эти диаграммы принимают соглашения о порядке стран в соответствии с долей респондентов, которые согласились с каждым утверждением, с цветовой легендой, размещенной сверху под названием диаграммы. Напротив, в примере четыре используется алфавитный порядок стран и правосторонняя легенда, что значительно усложняет познание.

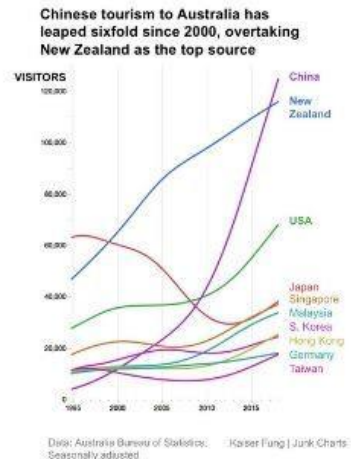
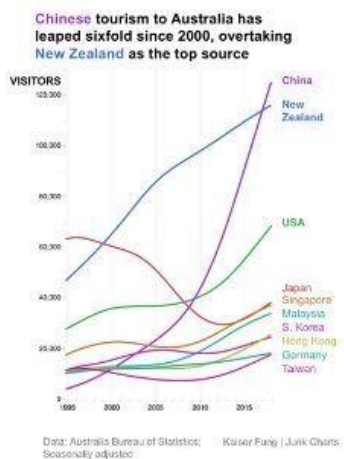


Пример третий: соответствие легендам



Пример четвертый: расходящиеся легенды

Новая традиция — встраивать легенду в заголовки или подзаголовки диаграмм. Применительно к графике данных по туризму в Австралии цветной текст в примере пять направляет взгляд читателя на ключевые страны — Китай и Новую Зеландию. Пример шесть требует от читателя немного больше усилий, чтобы связать заголовок диаграммы и метки линий.

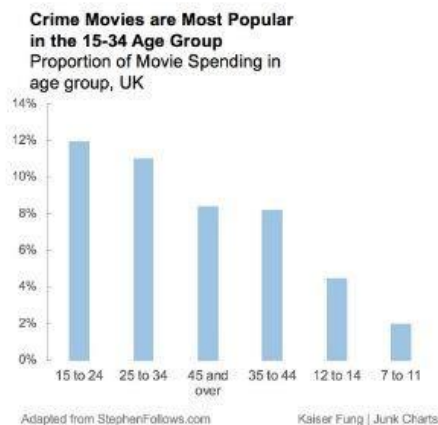
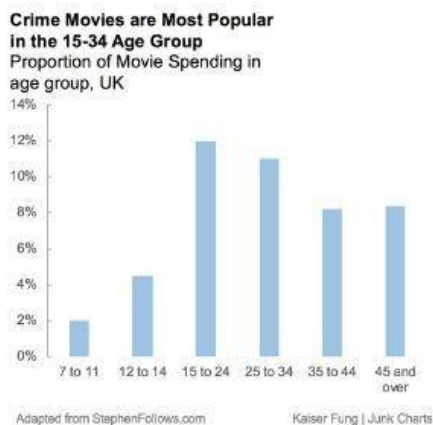


Пример пятый: совпадающие легенды (слева). Пример шестой: расходящиеся легенды (справа).

Заказ

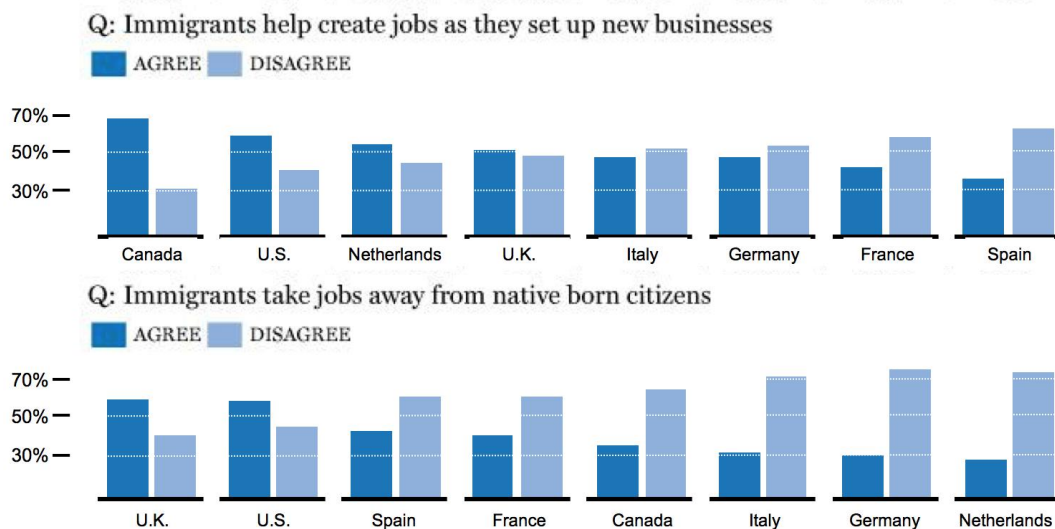
То, как элементы упорядочены на диаграмме, имеет огромное влияние на понимание читателя. Несмотря на пристрастие программного обеспечения к алфавитной схеме, это редко бывает правильным выбором. Говард Уэйнер, автор *Visual Revelations* и других книг по визуализации данных, высмеял это как «Алабама первая!» (Алабама — первый штат в алфавитном порядке.) Соглашение требует использования естественного порядка, когда он доступен. Временные переменные, возрастные группы, группы доходов, уровни образования и т. д. — все имеют естественный порядок.

Пример один показывает относительную популярность криминальных фильмов в разных возрастных группах в Соединенном Королевстве, составленную исследователем Стивеном Фоллоузом. Возрастные группы представлены в естественном порядке, от самых молодых к самым старшим. Сортировка по значению, как показано в примере два, не очень хорошо работает с данными, имеющими естественный порядок, поскольку глаза прыгают, чтобы восстановить последовательность.



Пример первый: согласованный порядок (слева). Пример второй: расходящийся порядок (справа).

При создании панели графиков правило заключается в том, чтобы сохранять одинаковый порядок значений на всех диаграммах. Пара графиков в примере три, адаптированная из ранее цитируемого исследования отношения к иммиграции *The National Post*, иллюстрирует, почему изменение порядка стран от диаграммы к диаграмме затрудняет возможность читателя сравнивать ответы на два вопроса опроса.



Adapted from National Post, 3 February 2011

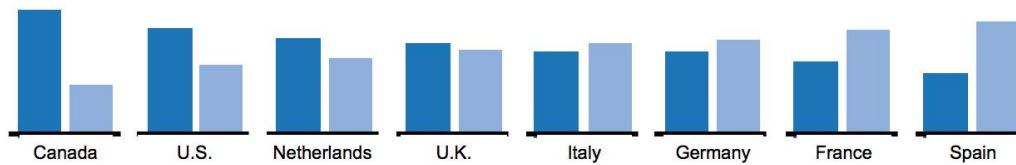
Kaiser Fung | Junk Charts

Пример третий: расходящийся порядок

Мы должны зафиксировать порядок стран по всей панели, как показано в примере 4. Страны располагаются слева направо по убыванию доли респондентов, согласившихся с первым утверждением.

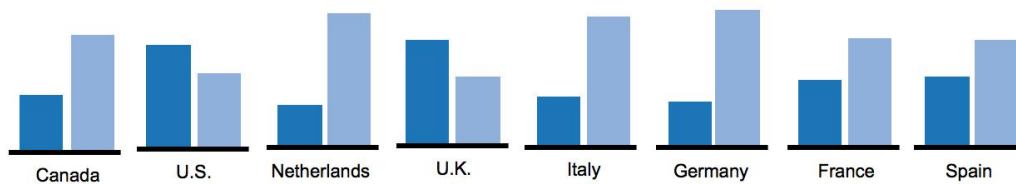
Q: Immigrants help create jobs as they set up new businesses

AGREE DISAGREE



Q: Immigrants take jobs away from native born citizens

AGREE DISAGREE



Adapted from National Post, 3 February 2011

Kaiser Fung | Junk Charts

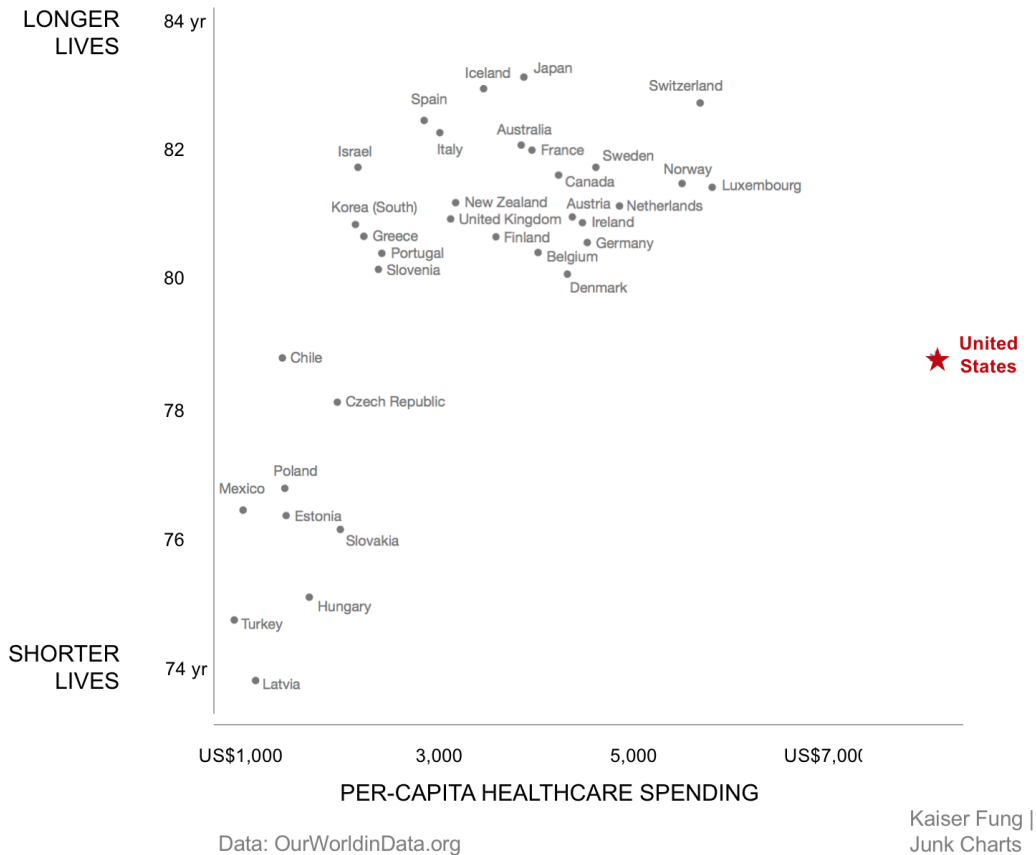
Пример четвертый: соответствие приказу

Аннотация

Текст, используемый экономно, дополняет визуальный опыт. Многие авторы рекомендуют использовать информативные заголовки диаграмм. Дизайнер должен заменить заголовки диаграмм по умолчанию, назначенные графическим программным обеспечением, которые обычно формируются путем объединения заголовков осей. Другое правило — объяснять все аббревиатуры и жаргон. Также принято включать источник(и) данных в нижний колонтитул.

При маркировке данных правило заключается в том, чтобы маркировать элементы, которые являются ключевыми частями истории. Не маркируйте все. Метки, по сути, дают читателям подсказки относительно наиболее значимых элементов. Пример 1 воспроизводит более раннюю диаграмму, исследующую эффективность расходов на здравоохранение, с полным набором меток стран. Слишком много меток претендуют на внимание читателя.

The United States spent at least 50% more per capita on health care, yet its life expectancy in 2012 lagged most developed nations.



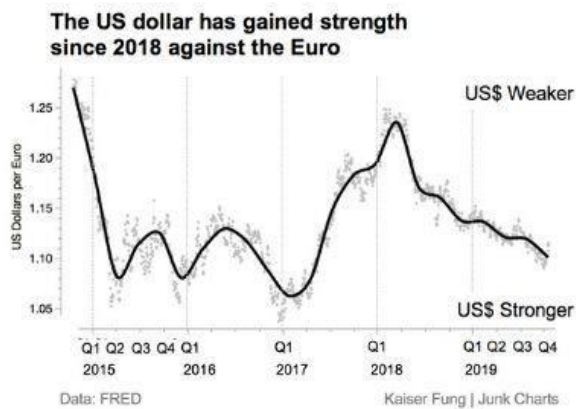
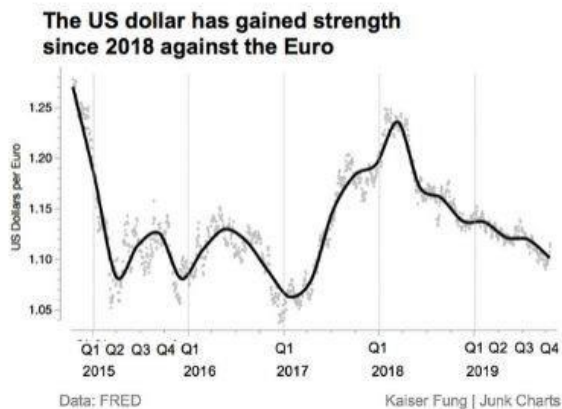
Пример первый: расходящиеся этикетки

Когда следует игнорировать условности

Соглашение возникает, когда множество практиков соглашаются с мудростью элемента дизайна. Некоторые правила имеют когнитивные обоснования, подкрепленные научным экспериментом, который, по-видимому, не является ни достаточным, ни необходимым для их популяризации. Исследователи Билл Кливленд и Роберт Косара провели некоторые из таких исследований. Но почти у каждого соглашения есть исключения. Мой совет: подумайте дважды, прежде чем нарушить правило, но не думайте дважды, если это необходимо.

Я сталкивался со многими примерами диаграмм, в которых та или иная условность оправданно отбрасывается для улучшения понимания. Позвольте мне закончить этот длинный текст примером, в котором нарушение правил приносит дивиденды.

Представьте себе, что вы используете пример один, чтобы донести до американских читателей сообщение о том, что доллар США укрепляется по отношению к евро с 2018 года. Визуальное впечатление от линии тренда, идущей вниз, противоречит сообщению об укреплении. Поскольку обменный курс выражается как количество долларов США за один евро, чем меньше это число, тем сильнее доллар США. Одним из решений этой визуальной проблемы является размещение аннотаций, как в примере два.



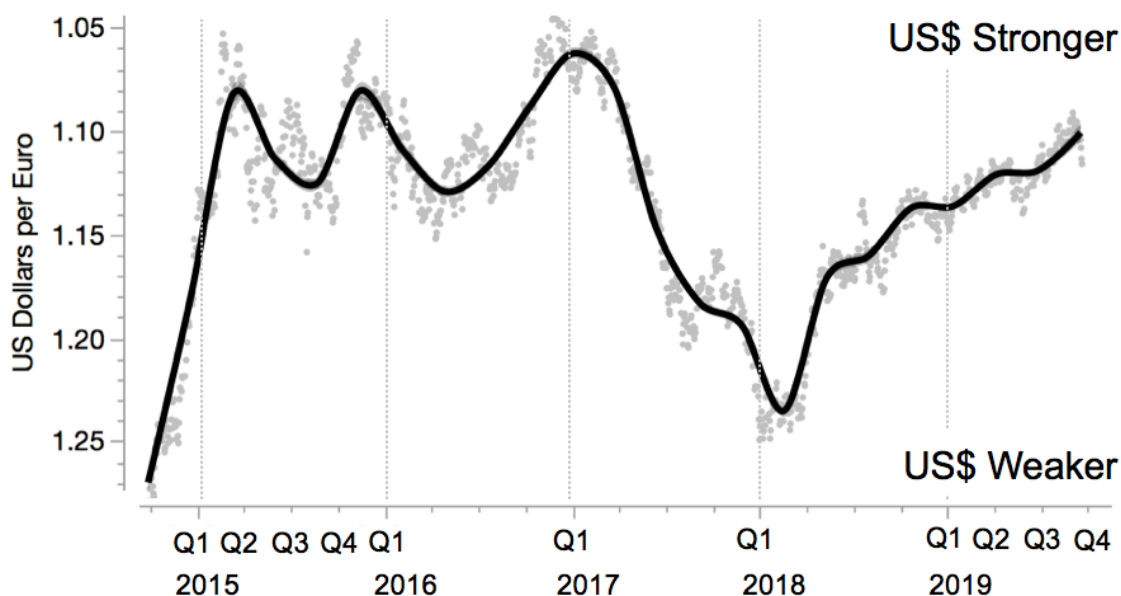
Пример первый: диаграмма без аннотации (слева). Пример второй: диаграмма с аннотацией (справа).

Такая аннотация просто устанавливает ворота для головоломки, которую должен решить читатель. Почему нижняя линия представляет более сильный доллар США?

В этой ситуации дизайнер может также нарушить правило осей, инвертировав вертикальную ось. Пример три идентичен нашему второму примеру, за исключением инверсии оси (и последующего переворачивания меток).

The US dollar has gained strength since 2018 against the Euro

Note vertical axis is inverted: stronger US\$ means fewer US\$ per Euro



Data: FRED

Kaiser Fung | Junk Charts

Пример третий: расходящаяся диаграмма с перевернутой осью.

Другой способ добиться этого эффекта — инвертировать соотношение обменного курса, выразив его как количество евро за один доллар США. Это решение не понравится финансовому сообществу, привыкшему смотреть на соотношение доллара США к евро.

Заключение

Визуальная среда отлично справляется с передачей большого объема информации в нескольких измерениях эффективно. Такая эффективность основана на наборе негласных правил и соглашений, неявно разделяемых производителями и потребителями графики данных. В этом длинном чтении мы рассмотрели выборку основных соглашений, охватывающих как эстетику, так и руководства по диаграммам. Дизайнеры визуализации данных могут использовать эти соглашения для упрощения своей графики, удаляя ненужные объяснения. Осознание этих негласных правил помогает избежать непреднамеренного недопонимания. Как и в случае с любым визуальным дизайном, в

зависимости от вашего конкретного приложения и аудитории, иногда может быть благоразумно бросить вызов соглашению. Наконец, подумайте дважды, прежде чем нарушить правило, но не думайте дважды, если это необходимо.

Обзор конвенций

Конвенции по эстетике

Круговые диаграммы

- ✓ Используйте разумное количество ломтиков.
- ✓ Объединить второстепенные категории в один раздел «Другое»
- ✓ Упорядочить ломтики по размеру от большего к меньшему
- ✓ Поместите фрагмент «Другое» в конец последовательности, независимо от порядка.
- ✓ Расположите первый и самый большой ломтик напротив верхнего вертикального радиуса.
- ✓ Расположить ломтики по часовой стрелке.
- ✓ Изменяйте цвета только в том случае, если цвета кодируют данные.

Гистограммы

- ✓ Начальная ось значений с нуля

Диаграммы рассеяния

- ✓ Поместите объясняющую переменную на горизонтальную ось
- ✓ Поместите переменную результата на вертикальную ось
- ✓ При добавлении линии регрессии назначьте переменную результата

Графики временных рядов

- ✓ По горизонтальной оси откладываем время.
- ✓ Время течет слева направо
- ✓ Если временные интервалы неравномерны, то и деления должны быть неравномерными.

Цветовая кодировка

- ✓ Ограничьте общее количество цветов
- ✓ Разница в цвете должна отражать разницу в данных.
- ✓ Используйте определенные пары цветов с осторожностью
- ✓ Сделайте диаграммы удобными для читателей с дальтонизмом

Соглашения о путеводителях

Топоры

- ✓ Используйте канонические направления (т.е. большие значения справа от меньших значений)
- ✓ Время идет по горизонтальной оси
- ✓ Поместите переменную результата на вертикальную ось
- ✓ Выберите ограничения, чтобы удалить лишнее пустое пространство
- ✓ Отметки должны приходиться на легко интерпретируемые приращения и значения.

Легенды

- ✓ По возможности используйте прямые метки.
- ✓ Цвета в легенде должны соответствовать цветам на диаграмме один к одному.
- ✓ Цвета в легенде должны быть представлены в том же порядке, в котором они указаны на диаграмме.
- ✓ Разместите легенду сверху под заголовком 5. Вставьте легенду в заголовки или подзаголовки диаграммы

Заказ

- ✓ Расположите значения в естественном порядке, когда это возможно.
- ✓ Избегайте алфавитного порядка по умолчанию, если это не оправдано контекстом.
- ✓ Сохраняйте одинаковый порядок на всех графиках в панели диаграмм.

Аннотация

- ✓ Используйте информативные заголовки диаграмм
- ✓ Объясните все сокращения и жаргонизмы.
- ✓ Включите источник данных в нижний колонтитул
- ✓ Маркируйте только ключевые элементы, а не все элементы