**Верю не верю**

Аракелян Арсен

Чтобы разобраться в том, как и где искажается информация, мы решили чуть подробнее рассмотреть основные этапы сбора, анализа и визуализации данных. Выяснилось, что скомпрометировать данные можно на любом из них. Включая сбор данных, особенно если никакого сбора-то и нет, а информация взята…

**С потолка**

Такой наглый подход и раньше встречался редко, а сейчас его и вовсе можно заносить в Красную книгу, поскольку человека, который постоянно придумывает несуществующую статистику, слишком легко поймать за руку. Впрочем, за руку ловят не часто - утверждения, подкрепленные статистикой и отсылкой к авторитетам, по умолчанию считаются достоверными, и сомневаются в их справедливости лишь те, кто придерживается противоположного мнения и имеет на руках статистику, доказывающую обратное утверждение.

Поэтому больше всего придуманной статистики на темы, которые никому особенно не важны. Знаете ли вы, что, по данным Всемирной Организации здравоохранения, блондинки являются носителями вымирающего гена и в течение ближайших двухсот лет исчезнут с лица Земли? А знакомо ли вам утверждение, согласно которому человек должен выпивать ежедневно не меньше восьми стаканов воды? Или, возможно, вы слышали, что ежегодно полторы сотни человек погибает под ударами упавших с пальмы кокосовых орехов.

Если слышали, то имейте в виду, что ни одно из этих утверждений не соответствует действительности. ВОЗ ничего подобного о блондинках не говорила, исследований на эту тему не вела и даже публично опровергла эти заявления (так что у блондинок есть шанс на выживание); на опасность зрелых кокосов обратили внимание копирайтеры британской страховой компании, которые, конечно, никаких подсчетов не проводили, а байка про восемь стаканов воды появилась в результате неправильно понятой журналистом научной статьи, автор которой хоть и писал о злополучных стаканах, отдельно упоминал, что это общий объем жидкости, включающий в себя как чистую воду, так и воду, содержащуюся в продуктах питания (а там ее достаточно много).

Иногда придуманная статистика выбрасывается в народ не для того, чтобы убедить кого-то в своей правоте, а чтобы убедить аудиторию в неправоте оппонента. Джоэл Бест в книге "Наглая ложь и статистика" приводит пример, когда активисты атаковали Рональда Рейгана, утверждая, что в США полтора миллиона бездомных. Активисты придумали эти данные (отвечая на любую критику удивительно конструктивным образом: "Вы что, сомневаетесь, что бездомных мало? Вам что, наплевать на проблемы бездомных?"). Все понимали, что данные о бездомных взяты с потолка, но еще очень долго администрации приходилось убеждать общественность, что реальные показатели в десятки раз меньше - общественность была уверена, что правительство занижает невыгодные для себя показатели, ведь дыма без огня не бывает. И если полтора миллиона были придуманы, значит, в реальности как минимум миллион двести - и ни одним бездомным меньше.

**Британские ученые**

Феномен "британских ученых" все еще ждет своего Линнея. "Британские ученые" - один из главных источников недостоверной информации, в том числе и статистической (для российской прессы, в Британии они пользуются меньшей популярностью). Само словосочетание звучит достаточно серьезно, да и поди проверь, что думают британские ученые на самом деле. Кроме того, Великобритания большая страна, в ней много ученых - вполне возможно, что при определенных усилиях там можно найти ученого, который придерживается нужной точки зрения.

Одним из главных источников "научных новостей" с островов Туманного Альбиона является газета Daily Mail, чей авторитет в научном мире близок к авторитету в том же научном мире газеты "Огни Заполярья". Daily Mail - очень качественный таблоид, сотрудники которого умеют придумывать яркие заголовки и хорошие сюжеты. Ни проверкой полученных данных, ни, боже упаси, рецензированием статей они не занимаются - и даже из научных новостей выбирают только те, которые можно эффектно подать и продать целевой аудитории. А целевая аудитория Daily Mail - особенные люди, которые в Британии давно стали объектом злых шуток (у нас аналог подобрать довольно трудно - ну разве что читатели "Комсомольской правды" и "Экспресс-газеты"? - но что-то подсказывает, что 35 процентам населения, которые считают Евгения Петросяна своим любимым артистом, русский аналог Daily Mail пришелся бы по душе).

**"Независимые" лаборатории**

Не исключено, что в большинстве "независимых" лабораторий работают именно "британские ученые". Как бы то ни было, "независимые" лаборатории удивительно часто добиваются такого результата, какого ждет от них рекламодатель. И главный метод обеспечения нужных результатов - грамотный подбор группы, на которой будут производиться испытания. В идеале статистическая выборка должна быть достаточно широкой и составленной случайным образом, однако на практике это зачастую невыполнимо. Если, например, заказчика интересует, как отнесутся к его товару обеспеченные домохозяйки от 20 до 27 лет, то платить за привлечение к опросу пенсионеров он не будет, да и выборку в два десятка человек может посчитать вполне достаточной. На выходе получаем изящный пресс-релиз, из которого следует, что девяносто процентов женщин, принявших участие в исследовании эффективности крема от морщин, отметили, что после трех недель использования крема морщины значительно разгладились. Самое интересное, что все в этом пресс-релизе - чистая правда, хотя в двадцать лет морщин, наверное, не очень много.

Величина выборки во многом зависит от объекта исследований. Если, скажем, для исследования эффективности крема с лихвой хватило бы тысячи случайно выбранных женщин, то для проверки эффективности прививки от полиомиелита не хватило в свое время 1200 детей, половине которых была сделана прививка, а половине - нет. Собственно, сама проверка прошла вполне успешно - никто из привитых детей во время эпидемии полиомиелита не заболел. Подпортило картину то, что из непривитых детей тоже ни один не заболел. По той простой причине, что полиомиелит, слава богу, не простуда. И даже во время эпидемии вероятность заболеть полиомиелитом грозит очень малой доле населения. Детям повезло. Экспериментаторам, которые должны были изначально взять более широкую выборку, - не очень.

Порой источником нужных заказчику данных является не полувиртуальная независимая лаборатория, а вполне уважаемая научная организация. Позволить себе такое могут только очень богатые компании, заказывающие исследования сразу в нескольких институтах, а потом выбирающие те результаты, что не противоречат "политике партии". По очевидным причинам информации о таких сделках крайне мало, но то тут, то там, бывает, всплывет кусочек внутренней корпоративной переписки, из которой заинтересованные читатели могут узнать, например, что в начале 1990-х одна крупная табачная компания запретила публикацию заказанного ею же исследования, потому что по результатам этого исследования получалось, что на здоровье животных пассивное курение действует плохо. К счастью, у компании на руках было другое исследование, в котором говорилось, что в течение девяноста дней у наблюдаемых животных никаких серьезных отклонений замечено не было. Оно и пошло в ход.

Еще один пример, который можно здесь привести, не связан напрямую с независимыми лабораториями, а связан, скорее, с не очень добросовестной рекламой. Как и подавляющее большинство статистических подлогов, он крайне прост и формально подлогом не является, однако вводит потребителя в заблуждение. Это реклама паевых инвестиционных фондов, которые - если верить этой самой рекламе - являются крайне прибыльными предприятиями. На заинтересовавшегося будущими миллионами инвестора с удовольствием обрушат гору статистической информации, наглядно доказывающей, как мудра управляющая компания, как осторожно и вместе с тем эффективно она распределяет средства, как выросли активы ее клиентов за последний год и так далее. И все это правда. Все так и есть. С той лишь разницей, что таких фондов у компании множество, рекламирует она, разумеется, самые успешные и никаких гарантий, что активы будут расти и дальше, конечно же, не дает. Если бы компанию интересовала не прибыль, а репутация, она, возможно, обратила бы внимание потенциального инвестора на то, что потеряла деньги на многих своих начинаниях, а также на тот простой факт, что предсказать рост или падение в долгосрочном периоде попросту невозможно. Но в компании, которые заботятся о репутации в ущерб прибыли, много денег не принесут, поэтому финансисты делают вид, что вырастили целый лес, даже если из всех саженцев прижился лишь один. Умному и этого достаточно, а у дурака все равно кто-нибудь деньги вытянет - так лучше уж мы, чем какие-нибудь мошенники. И ведь не придерешься ни к чему. Все честно.

**Государственные агентства**

На самом деле, данные государственных агентств - это, пожалуй, одни из самых чистых и лучших данных, которые может получить исследователь. Тем не менее государственные службы тоже не являются источником истины в последней инстанции, поскольку фиксируют не сами события, а бюрократическую реакцию на них. Для примера возьмем две цитаты, которые, дословно передавая одну и ту же по сути статистическую информацию, по смыслу противоречат друг другу. Первая взята из журнала "В мире науки" (июнь 2006 года).

По данным Федеральной службы государственной статистики, российские семьи распадаются все реже, однако количество браков с каждым годом сокращается. Согласно последним подсчетам социологов, в 2004 г. в России было зарегистрировано 635 тыс. разводов. По сравнению с предыдущими годами эта цифра внушает оптимизм: так, в 2003 г. расторгли брак 798 тыс. пар, в 2002 г. - 853 тыс. Однако меньшее количество разводов приходится и на меньшее количество браков: 979 тыс. в 2004 г. против 1 млн. 91 тыс. в 2003 г. и 1 млн. 19 тыс. - в 2002 г.”. Вторая - из журнала "Финанс" (#44, декабрь 2005 года).

За первые девять месяцев 2005 года количество разводов и браков, по данным Росстата, заметно изменилось, причем в противоположных направлениях. По сравнению с прошлым годом россияне стали жениться на 7% чаще, а разводиться - на 6% реже. Началось это не вчера. Статистика показывает, что брак становится все крепче уже третий год подряд. По сравнению с 2002 годом количество разводов сократилось почти в полтора раза и за первые девять месяцев составило примерно 442 тыс., в то время как семей было создано более 800 тыс. И если в 2002 году на каждые сто браков приходилось 84 развода, то в 2005-м (по итогам девяти месяцев) - только 55”.

В цитатах есть странности большие и небольшие. В первой цитате почему-то приводится откровенно устаревшая статистика - и это небольшая странность, которую можно объяснить недосмотром редактора, писавшего врезку второпях (например, не успел человек найти более актуальной информации - бывает). А вот то, что в статье, из которой взята вторая цитата, факт уменьшения количества заключаемых браков в 2002–2004 годах опущен как незначительный, - уже большая странность. Это часто встречающийся прием: если нужно показать уверенный рост чего-либо, внимание читателя фокусируется на тщательно выбранном периоде времени, где рост действительно наблюдается. К слову, поскольку развод является производной функцией брака, интересно посмотреть, не вызвано ли уменьшение количества разводов тем простым фактом, что начали распадаться браки, заключенные в "неурожайные" для оформления брака годы, но такой статистики, к сожалению, нет, как нет и достоверных сведений о том, сколько способен продержаться среднестатистический брак, заканчивающийся разводом.

Впрочем, речь о другом. А что, вообще говоря, означают эти числа? Насколько они соотносятся с реальностью, данной нам в ощущениях? Да не очень-то. Эта статистика собрана на основе актов гражданского состояния, и, как следствие, в нее вошли только те пары, которые оформили свои отношения официально. Люди, живущие в гражданском браке, в официальную статистику не попали, потому что их учитывать трудно, точнее дорого. Меж тем они живут вместе, ведут общее хозяйство, растят детей (таких пар, правда, не очень много - обычно люди оформляют отношения перед рождением ребенка, - но тоже есть), расходятся, наконец. Но в этой статистике их нет, хотя если бы государственное агентство принципиально игнорировало людей с рыжим цветом волос, "погрешность" измерения была бы куда меньше. Собственно, к агентству претензий меньше всего - оно считает то, что может посчитать. Другое дело, что эти данные интерпретируются так, будто являются отражением реальной ситуации, хотя никаких оснований предполагать такое нет.

Еще один небольшой пример связан с уменьшением количества самоубийств в Российской Федерации. Очень здравая критика статистики по самоубийствам вообще изложена у Джоэла Беста в "Наглой лжи и статистике", но основное положение можно кратко пересказать и в журнальной статье, благо оно формулируется просто: мы не знаем, сколько людей действительно убивает себя. Вопреки распространенному мнению многие (я боюсь говорить "большинство", пусть будет "многие") самоубийцы не оставляют записок, и галочка в графе "самоубийство" зачастую не более чем предположение. Порой вполне обоснованное. Мы вполне можем предположить, что человек, съевший упаковку снотворного, собирался покончить жизнь самоубийством. Или, например, если он написал на асфальте "В моей смерти прошу винить Клаву К.", а потом выбросился из окна - это тоже похоже на самоубийство. Или, допустим, он ехал ночью по МКАДу со скоростью 200 км/час и попытался въехать в заграждение - это самоубийство или нет?

Во всех гипотетических случаях, кроме, пожалуй, второго, мы можем иметь дело как с самоубийством, так и с несчастным случаем или даже с убийством. Таким образом, в основе этих статистических данных лежат даже не акты гражданского состояния, которые пусть и неполно отражают реальность, но, по крайней мере, адекватно отображают какую-то часть ее, а суждение неизвестного нам человека, принимающего решение на месте. Мотивы этого человека нам тоже, по большому счету, неизвестны. Возможно, в течение нескольких лет подряд он записывал в самоубийства все сомнительные случаи, кроме тех, когда не было никаких сомнений, что расправиться с собой самостоятельно жертва не могла. Возможно, его начальнику несколько месяцев назад позвонили и мягко намекнули, что для областного центра такое высокое количество самоубийств попросту неприлично. Мог человек выпасть из окна в результате несчастного случая? Конечно, мог. А если мог - значит, выпал.

На всякий случай повторюсь, что у нас нет никакой достоверной информации о том, почему самоубийств в России стало меньше (и, разумеется, все предположения о звонках или намеренной маскировке убийств под самоубийства не более чем гипотетические примеры - не более того). Но достоверной статистики по самоубийствам в России нет. Она могла бы быть - пусть не абсолютно, но хотя бы относительно достоверной, когда можно рассуждать о темпах роста или падения количества самоубийств, не привязываясь к конкретным цифрам, - но для этого нужна гарантия, что правила игры на перепутье не меняются.

Если рассуждения о статистике самоубийств не кажутся вам убедительными, взгляните на динамику статистики преступлений в Российской Федерации и, если удастся, поговорите с людьми, которые пытались обратиться в милицию с трудно раскрываемым делом. Впрочем, эта статистика вполне достоверна в относительном смысле - если считать, что сопротивление дежурного, не желающего открывать дело, представляет собой постоянную величину и с течением времени не меняется.

**Добрые люди**

Добрые люди могут сильно отличаться друг от друга, но объединяет их одно - желание доказать свою точку зрения любым способом. И статистика им только помогает. Для начала возьмем относительно безобидный пример социальной рекламы, которая эффективно эксплуатирует как данные официальной статистики, так и распространенные заблуждения. Руководствуясь исключительно благими намерениями.

Многие, наверное, видели ролик, в котором говорится о том, что на российских дорогах ежегодно гибнет около 35 тысяч человек, то есть население небольшого города. Поскольку всем нам уже уши прожужжали про пьяных водителей, логично предположить, что в подавляющем большинстве смертей виноваты именно они. Однако знакомство с цифрами показывает, что, во-первых, за последние семь лет в статистике ГИБДД больше 35 тысяч погибших было только в 2003 году; во-вторых, в 2006 году (а более свежей годовой статистики по понятным причинам не существует) погибло не 35 тысяч человек, а 32 724 человека, и в-третьих, по вине водителей, находившихся за рулем в нетрезвом состоянии, погибло 2 673 человека (что, конечно, тоже очень много, но все же меньше десяти процентов от общего числа погибших). При этом пятая часть всех ДТП произошла из-за нарушений правил дорожного движения пешеходами, которых погибло больше семи тысяч человек.

Из социальной рекламы следует, что виновато в смертях несоблюдение правил ПДД, причем несоблюдение именно водителями. По-человечески сей призыв понятен, и, наверное, можно подобрать статистику, доказывающую именно это. Но куда деть семь тысяч пешеходов, которые, на свою беду, не смотрят на дорогу? Как объяснить необычайно высокую смертность из-за аварий в Московской области по сравнению с Москвой (тут может быть несколько предположений, но хотелось бы, конечно, не предположений, а ответов)? Что, например, делать с тем фактом, что в Москве занудное соблюдение правил ПДД наверняка приведет к аварии в первую же неделю эксперимента?

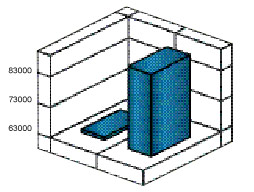
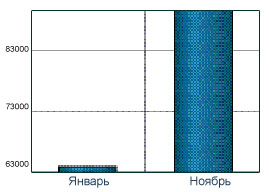
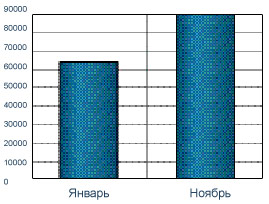
Мы не призываем не соблюдать правила. Но опубликованная статистика действительно порождает больше вопросов, чем ответов, призыв же соблюдать правила - правильный по сути - вряд ли является достаточным, чтобы значительно снизить смертность на дорогах. А правильно поставленные вопросы уж точно не помешали бы безопасности движения.

Да, откуда, кстати, взялись еще почти 2,5 тысячи человек? Они взялись из предположения, что статистика ГИБДД неполна, потому что в число погибших в ДТП включаются лишь те люди, что умерли в первые семь дней после аварии. Те, кто в этот срок, простите, не уложился, в статистику не входят. Не очень, правда, понятно, почему авторы лозунга решили округлить именно до 35 тысяч - с тем же успехом они могли сказать, что на дорогах погибает сорок тысяч человек в год. Все равно никаких способов проверить это утверждение нет - и тем не менее оно может быть правдой.

Копаясь в выложенной в общий доступ статистике ГИБДД, я, к своему удивлению, не обнаружил столь усердно рекламируемой отдельными сенаторами статистики по ДТП с участием праворульных автомобилей. Если она и существует, то простым смертным доступна лишь со слов членов Совета Федерации, утверждающих, что в 2004 году в ДТП с праворульными автомобилями погиб 2 741 человек.

У меня нет никаких сомнений, что члены Совета Федерации, пытаясь запретить ввоз и эксплуатацию автомобилей с правым рулем, руководствовались исключительно благими намерениями, но даже если принять на веру, что статистика такая существует, остается очень много вопросов. Например, по чьей вине были совершены эти ДТП (мы уже видели выше, что в значительной части ДТП виновны пешеходы, которых ни один гаишник, кстати говоря, не трогает)? Как распределились эти аварии по регионам? Например, в Дальневосточном округе попасть в ДТП на леворульном автомобиле затруднительно, так как леворульный автомобиль нужно еще найти. Фиксируют ли сотрудники ГИБДД при оформлении протокола, с правым или левым рулем был пострадавший автомобиль, и если да, как давно они это делают? Можно ли, сравнив количество леворульных и праворульных автомобилей в Москве (если есть такая статистика) или в стране (если есть такая статистика) с количеством ДТП, в которых виноваты водители этих автомобилей (есть такая статистика), сделать вывод, что праворульные автомобили опаснее леворульных или наоборот?

Случай с антипропагандой праворульных автомобилей - чистейшей воды подмена понятий, когда под видом ответа на один вопрос собеседнику дается ответ на вопрос, который мало того что не задавался, так еще никому особенно и не интересен. При этом совершенно неважно, насколько объективна приведенная в ответе статистика.



Проблема среднего

Ремарка

Просто, чтобы расставить точки над i. Мы исходим из того, что если ФЦП повышения безопасности движения поможет спасти жизнь хотя бы одному человеку, то она уже хороша. Но сам механизм социальной деформации вполне адекватных статистических данных любопытен. Тем более, что встречается довольно часто. Так, несколько лет назад в США договорились до того, что ожирение ежегодно уносит больше жизней, чем курение (что, вообще говоря, не совсем так).

Термин "среднестатистический" уже давно используется так широко, что и термином быть перестал. Впрочем, широкое использование не делает это слово яснее. В первую очередь потому, что непонятно, что именно понимать под средним вообще.

Классический пример (который приводит Хафф и почти все его последователи; правда, по-моему, нечто подобное я слышал еще в школе, хотя мои учителя Хаффа, конечно, не читали) звучит примерно так: в некой организации работают 100 человек, из них 95 получают, скажем, 40 тысяч рублей в месяц, четыре менеджера получают 100 тысяч рублей в месяц, а владелец бизнеса, он же генеральный директор, получает пять миллионов, потому что у него большая семья. Вопрос: какова средняя зарплата в этой организации?

Попытка рассчитать среднее арифметическое приводит нас к бессмысленной, в общем-то, зарплате в 92 тысячи рублей. Этот показатель, конечно, близок к зарплате менеджера, но вряд ли 95 рабочих сочтут такой подсчет справедливым. И понятно, что погоду портит директор, который со своей семьей выбивается из распределения. Поэтому для таких случаев обычно используется медиана - выбирается срединное значение (то есть, в нашем случае, сорок тысяч рублей). Разумеется, это тоже неидеальное решение, поскольку ни зарплаты менеджеров, ни зарплата директора в нем не учтены, однако оно куда ближе к реальному положению дел, чем предыдущий вариант.

Это совсем не высшая математика, в этом способен разобраться даже не самый способный третьеклассник. Однако когда мы сталкиваемся не с реальной статистикой, а скорее с ее пересказами, то очень часто понимаем, что у нас нет вообще никакой информации о том, как считались средние показатели (и почему).

Художник

И вот эти неполные, а зачастую дополненные предрассудками статистиков и многочисленных пересказчиков информации (каждый из которых на своем этапе может внести в данные ошибки), попадают наконец к художнику. У художника задача простая. Он должен взять табличку в Excel и сделать из нее хорошую картинку.

Для простоты давайте пока считать, что художник честный и действительно хочет сделать хорошую иллюстрацию, руководствуясь при этом своими представлениями о хорошем. И данные у него неплохие, бывают и хуже.

Гипотетический пример:

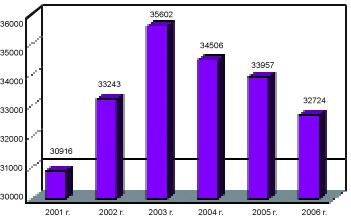
Журнал "Компьютерра" в городе N покупает сто человек. Тридцать пять студентов, тридцать руководителей, пятнадцать сотрудников компьютерных компаний, десять домохозяек, три биолога, четыре врача, два олигарха и кондуктор трамвая.

Логика диктует художнику, что мелкими показателями можно пренебречь. На диаграмме они будут выглядеть некрасиво. Поэтому художник, нисколько не сомневаясь в правильности выбранного решения, меняет статистику так, чтобы она укладывалась в картинку. И теперь "Компьютерру" в городе N покупают студенты, руководители, сотрудники и таинственное Другое, в которое художник запихнул и олигархов, и биологов, и врачей, и даже кондуктора.

Теперь предположим, что у нашего художника гибкая мораль, и главный редактор "Компьютерры" поставил перед ним сложную задачу: нарисовать такую иллюстрацию, из которой бы следовало, что за последние девять месяцев тираж журнала вырос в двадцать раз (на самом деле он вырос чуть меньше, чем в полтора раза, поэтому задача на первый взгляд кажется нетривиальной).

Это делается очень просто. Во-первых, нужно выбрать правильную точку отсчета. Если взять за основу диаграмму, отражающую реальное положение вещей, то по ней, конечно, никакого двадцатикратного увеличения не получишь. Но стоит поднять ось абсцисс до уровня 63 тысяч и обрезать ушедшие в минус элементы, как положение моментально меняется. Визуально мы воспринимаем эту картинку как рост с 1 тысячи до 27 тысяч - налицо увеличение в 27 раз. Теперь, чтобы ось ординат с неудобными для нас показателями (ну в самом деле, 63 тысячи на нуле - кому это может быть интересно, это же так, детали?) меньше бросалась в глаза, мы можем сделать еще две подтяжечки (тут все зависит от наглости художника и отчаянности поставленной задачи). Во-первых, мы можем просто убрать все смущающие надписи - но это грубо. Более распространенный метод заключается в развороте графика: столбики с показателями переводятся в 3D, причем ближний к нам становится еще больше за счет перспективы, а мешающие восприятию подписи хоть и не убираются совсем, но в глаза бросаются куда меньше.

Чтобы посмотреть, как это применяется в реальной жизни, достаточно взглянуть на диаграмму 4. Это строгая деловая графика, здесь нет как такового разворота к зрителю, да и подписи все на месте, но сдвиг оси и 3D присутствуют.



В защиту сотрудников ГИБДД можно сказать, что, скорее всего, трехмерность графика не намеренная - парадокс в том, что несмотря на дружный хор специалистов по дизайну, которые предлагают не увлекаться 3D-графиками, наиболее популярные средства построения графиков фактически провоцируют пользователей на использование 3D. Делать 3D-графики несложно, воздействуют они эффективно, а то, что реальность отображают "творчески" - так это проблемы реальности.

С художником связана и проблема площадей. Очень часто для визуализации статистики удобнее применять не диаграмму, а специально созданную иллюстрацию, на которой увеличение какого-либо показателя демонстрируется увеличением площади. Но рисует художник на глазок, и, как правило, соотношение выдерживается нечетко (если нужно показать рост показателя, художник ошибается в большую сторону, если падение - в меньшую; другими словами, художник почти всегда рисует большую площадь, чем нужно).

В заключение скажем, что все перечисленные художественные приемы не порочны сами по себе. Смещение оси, например, нередко используется для того, чтобы показать изменение показателя (если использовать реальные масштабы, читатель просто ничего не заметит). Для демонстрации изменения ежедневных биржевых индексов смещение оси обычное дело. Но именно эти приемы чаще всего используются для искажения информации, которая уже могла пройти через несколько раундов искажения на этапах сбора и интерпретации.

**Вместо резюме**

Картина получается безрадостная. У читателя, разглядывающего трехмерный график в газете, есть куча шансов нарваться на фальшивку, даже если участники цепочки - кристальной чистоты люди. Начинается все с неполной статистики, которая во многих случаях и не может быть полной. Неполные сырые данные дополняются предубеждениями интерпретатора и анализируются в удобном для него ключе. Затем этот блок данных попадает к художнику, который проводит еще пару хирургических операций, а затем - рисует приблизительный образ в очередной раз измененной информации так, как он его понял.

И это при том, что мы рассказали только о верхушке айсберга - из-за недостатка места в статью не вошли такие замечательные приемы, как искусственная корреляция, когда два схожих внешне, но не связанных между собой тренда объединяются, ошибки программного обеспечения (один из самых эффектных графиков, доказывающих глобальное потепление, на самом деле является результатом грубейшей компьютерной ошибки, исправление которой приводит к графику совершенно другой формы) и многое-многое другое.

Адекватная статистика существует, но, как любой качественный товар, стоит дорого. И как любой, даже самый дорогой товар, поставляется as is - ни стоимость отчетов, ни имя поставщика информации гарантией качества не являются, пресловутый человеческий фактор и сложность сбора статистических данных все портят. Но у покупателей есть хоть какой-то шанс.

Что же касается тех, кто ратует за свободу информации… Нам приходится есть, что дают. И платить за якобы бесплатную информацию либо собственным временем, оценивая, как ее могли исказить в процессе подгонки в удобоваримую форму, либо собственными решениями, которые будут приняты на ее основе.