

Варианты графического представления темпоральных изменений (на примере издания «ВЦИОМ новости»)

Variants of the graphical representation of temporal changes (using the example of the «VCIOM News» publication)

Дерябкина В.Д.

Студентка, ФГБОУ ВО «Санкт-Петербургский государственный университет»,
г. Санкт-Петербург
e-mail: varvara.deryabkina@yandex.ru

Deriabkina V.D.

Student, St. Petersburg State University, St. Petersburg
e-mail: varvara.deryabkina@yandex.ru

Аннотация

В процессе создания информационных текстов в СМИ авторы нередко используют темпоральные данные, полученные в ходе многочисленных исследований для визуального подтверждения суждений, изложенных в материалах. В статье описаны теоретические аспекты, связанные с иллюстрацией темпоральных изменений, а также рассмотрено применение этих правил на практике в действующем издании «ВЦИОМ новости».

Ключевые слова: темпоральность, темпоральные данные, визуализация, диаграммы, ВЦИОМ.

Abstract

In the process of creating informational texts in the media, authors often use temporal data obtained in the course of numerous studies to visually confirm the judgments expressed in the materials. The article describes the theoretical aspects related to the illustration of temporal changes, as well as the application of these rules in the current edition of VCIOM News.

Keywords: temporality, temporal data, visualization, diagrams, VCIOM.

Введение

Темпоральность – это понятие, связанное с временем или временными аспектами. Оно описывает характеристики, связанные со временем, последовательностью событий, длительностью или изменениями во времени. Термин "темпоральность" часто используется в философии, лингвистике, искусстве и других областях для обозначения временных аспектов явлений и процессов.

В широком смысле, темпоральные данные представляют собой произвольные данные, содержащие явную или неявную связь с определенными датами или промежутками времени. Это определение охватывает почти все виды данных и информации.

Темпоральность увязывает содержание материала с реальностью. Всякое издание СМИ имеет связь с категорией времени в нескольких аспектах одновременно. «Темпоральный характер трансляций динамических СМИ определяется способом освещения их категорий, являющихся общими для любых СМИ, в специфическом для них контексте – временном» [1].

В случаях, когда исследователи рассматривают DATA-журналистику как объект изучения, в контексте рассуждения о темпоральности могут иметь в виду непосредственно

данные по определенным показателям, собранные в течение конкретного времени. В таких ситуациях темпоральные данные становятся не только частью содержания текста, но и могут быть использованы в качестве иллюстрации. Для этого на основе темпоральных данных создают диаграммы, графики или иную визуализацию.

Визуализация данных представляет собой метод графического отображения информации и аналитики с использованием графиков, диаграмм, карт и дашбордов. В отличие от числовых значений, которые могут быть представлены в таблицах и документах, визуализация позволяет наглядно представить взаимосвязи между процессами, периодами роста или спада, а также зависимости между различными показателями. Визуальный формат информации помогает создать цельную картину предмета изучения.

В случаях, когда изобилие цифр может привести к потере смысла, рекомендуется использовать визуализацию для упрощения читателю анализа больших объемов данных. Этот инструмент позволяет систематизировать огромное количество информации, облегчает ее понимание и выделяет тренды.

«Цель визуализации: сделать данные наглядными для читателя. Отсюда задачи визуализации связаны непосредственно с работой с данными: показать, сравнить, организовать или связать их» [6, с. 5]. Прежде всего, любая отличная визуализация функциональна в том смысле, что она позволяет вам получить доступ или увидеть тенденции и закономерности [5].

Следующий параметр – эстетическая ценность. Очевидно, что визуализация должна быть привлекательной, потому что в противном случае читатели не остановятся и не будут взаимодействовать с ней или извлекать смысл из неё. Таким образом, эстетический уровень визуализации также чрезвычайно важен.

Далее, любая качественная визуализация должна быть содержательной, т.е. позволять читателю узнать что-то новое, удивительное, возможно, непредвиденное. Качественно выполненная визуализация раскрывает вещи, о которых читатель ранее не знал, меняет отношение к уже известному.

Так, хорошая визуализация должна соответствовать всем вышеперечисленным пунктам, чтобы оставаться наиболее полезной и присутствовать не только в качестве иллюстрации к тексту, но также и в качестве равноправного его дополнения [12].

Для того, чтобы составить, в первую очередь, функциональную визуализацию, нужно верно определить, какая графическая форма более подходит для представления данных. Существует множество способов визуализации динамики, которые используются в графическом представлении данных. Чаще всего применяются линейные, радиальные, столбиковые, бокс-плот, ленточные и фигурные диаграммы [10].

«Чаще всего во временных рядах, или в темпоральных данных, люди стараются разглядеть тенденции» [13, с. 106]. В связи с этим именно графики, верным образом построенные в соответствии с темпоральными данными, позволяют увеличить уровень доверия к тексту в СМИ, а также ускорить процесс восприятия информации из текста. Читателю не нужно будет долго изучать данные в текстовом формате, будет достаточно лишь взглянуть на график или диаграмму.

В контексте их объема, количественные временные данные могут быть классифицированы на дискретные и непрерывные. Определение категории, к которой относится имеющаяся информация (дискретные или непрерывные данные), позволит определить наиболее подходящий способ их визуализации.

Дискретные данные представлены конкретными значениями, относящимися к определенным моментам или периодам времени, и имеют конечное число значений. В отличие от этого, непрерывные данные, такие как температура, могут быть измерены в любой момент времени на протяжении любого интервала, и постоянно изменяются.

От того, какие данные, дискретные или непрерывные, необходимо визуализировать, будет зависеть и выбор диаграммы.

В случае с дискретными данными есть три типа диаграмм: столбчатые, штабельные и точечные.

Столбчатые диаграммы

Столбчатые диаграммы являются одним из наиболее распространенных видов диаграмм и широко применяются для визуализации различных типов данных, включая временные ряды. Структура столбчатой диаграммы состоит из оси времени, оси значений и отдельных столбцов.

Ось времени представляет собой горизонтальную ось, также известную как ось абсцисс или ось X, на которой указываются моменты времени, расположенные в хронологическом порядке. Для временных данных эти моменты времени могут быть представлены в виде рядов месяцев, лет или других единиц времени. Ширина столбцов и расстояние между ними обычно не несут никакой значимой информации.

На вертикальной оси значений, также известной как ось ординат или ось Y, наносится линейная шкала диаграммы. Единицы на этой оси размещены на равном расстоянии друг от друга по всей длине оси, чтобы предотвратить неправильное чтение графика. Высота столбцов определяется относительно оси значений и визуально отображает его содержание. Чем меньше значение, тем ниже будет столбец, соответственно, чем больше значение, тем выше будет столбец.

Во многих средах значение по умолчанию для точки отсчета на оси ординат является минимальным значением в наборе данных. Однако это может привести к неправильному восприятию информации из диаграммы. Поэтому ось значений всегда должна начинаться с нуля, чтобы правильно отражать соотношение величин в столбчатой диаграмме.

Штабельные столбчатые диаграммы

Конфигурация штабельных столбчатых диаграмм (нормированных столбчатых диаграмм) аналогична обычным столбчатым диаграммам, с той разницей, что прямоугольники укладываются друг на друга в стык. Штабельные столбчатые диаграммы применяются в случаях, когда в списке данных имеется несколько подкатегорий и их сумма имеет значение.

Подобно обычным столбчатым диаграммам, штабельные диаграммы используются для визуализации как темпоральных, так и категорийных данных. Высота столбца в данном случае представляет суммарное значение для каждого месяца, а высота стека соответствует значению каждой подкатегории.

Диаграмма рассеяния

Диаграмма рассеяния, также известная как диаграмма в виде точек на плоскости, является распространенным инструментом для иллюстрации взаимосвязи между двумя переменными, в том числе в контексте темпоральных данных. В этом типе диаграммы время обычно представлено по горизонтальной оси, в то время как значения или измерения отображаются по вертикальной оси [11].

Использование точек вместо столбцов может быть предпочтительнее в некоторых случаях. Точки занимают меньше места на графике и, за счет отсутствия столбцов, создают ощущение непрерывного потока данных, который переходит от одной позиции к другой. В отличие от столбчатых диаграмм, где высота столбца служит визуальной подсказкой, диаграмма рассеяния использует местоположение точек для этой цели. Каждая точка имеет свои координаты по оси x и y, и читатель может сравнивать точки во времени на основе их расположения на графике.

Визуализация непрерывных временных рядов в научном стиле имеет сходства с визуализацией дискретных данных. В обоих случаях присутствуют ограниченные точки ввода данных, даже если набор данных является непрерывным. Структура визуализации будет одинакова, но разница заключается в описываемых явлениях в материальном мире.

Непрерывные данные сообщают постоянно изменяющиеся явления, и их визуализация должна отображать эту непрерывность.

«Часто необходимо сравнить большое количество временных рядов одновременно. Для этого используются горизонтальные графики (Horizon Graphs) как метод увеличения плотности записи данных из временных рядов для предварительного просмотра» [9, с. 142].

Диаграмма временных рядов

Диаграмма временных рядов, похожа на точечную диаграмму, за исключением того, что точки соединены между собой. Однако иногда точки могут быть невидимыми [2]. Диаграмма имеет точки с x- и y-координатами и соединительные линии, которые помогают выявить тенденции в данных. Часто полезно начинать ось значений с нуля, чтобы избежать искажений шкалы. Длина горизонтальной оси также может влиять на отображение тенденций. Если она слишком длинная, паттерны могут не быть замечены читателем.

Ступенчатый график является более сложной версией линейного графика. Одним из недостатков стандартного линейного графика является его предположение о стабильном движении от точки А к точке В. Однако в некоторых случаях после длительного времени в неизменном состоянии может произойти резкий скачок вверх или вниз. Например, процентные ставки могут оставаться на одном уровне в течение многих месяцев, а затем резко упасть за один день.

Ось значений на ступенчатом графике показывает шкалу значений, начиная с нуля. На оси времени представлены месяцы, в течение которых происходило действие. В такой визуализации появляются новые сегменты, включая скачки – мгновенные переходы к следующим значениям и плоские сегменты – линии, которые показывают отсутствие изменений в определенный период времени.

Вместо прямого соединения точек А и В, линия на ступенчатом графике остается на уровне одного значения до возникновения изменения. Затем в точке изменения линия прыгает вверх или вниз к следующему значению. В результате получается последовательность ступеней.

Диаграмма-кривая

Диаграмма-кривая – это графическое представление данных, которое может быть полезно при работе с большими объемами информации или в случае наличия шума в данных. В таких ситуациях может быть сложно заметить тенденции и паттерны, поэтому применяется метод оценки линии тренда. Цель этого метода - нарисовать линию, которая проходит через максимальное количество точек и минимизирует суммарное расстояние от точек до этой линии.

Однако, если тенденция данных нелинейна, нет смысла подгонять прямую линию к ним. Вместо этого, можно использовать статистический метод, разработанный Уильямом Кливлендом и Сьюзен Дэвлин, известный как "локально взвешенное сглаживание диаграммы рассеяния" (LOWESS) [7]. Этот метод позволяет подогнать данные более точно.

LOWESS начинает с исходных данных и делает небольшие срезы. На каждом срезе определяется наименьшая степень чтобы лучше соответствовать данным из этого среза.

Таким образом, LOWESS движется по данным, создавая серию миниатюрных кривых, которые затем объединяются в одну большую. Эта сглаженная кривая, полученная с помощью LOWESS, называется лессовой кривой. Она значительно упрощает общее восприятие тенденции за длительное время, проиллюстрированное в графике [14].

Существуют определенные эмпирические правила, которым нужно следовать, чтобы выбрать наилучшую графическую форму для представления данных. В первую очередь, тип визуализации зависит от издания, в котором она будет опубликована [8]. На выбор влияет как тип СМИ: сайт, газета, телевизионный канал, так и целевая аудитория. Также визуализация зависит от круга вопросов, на которые она должна помочь ответить.

В случае, если читатель захочет сравнить показатели – для этого подойдет линейная диаграмма. Для визуализации географических данных стоит прибегнуть к созданию картодиаграммы. При создании графики недопустимо опираться исключительно на авторское видение.

Главная цель визуализации данных – обеспечение доступа к тенденциям и закономерностям в данных. В качественной диаграмме или инфографике достаточно взглянуть на графическую форму без необходимости считывать каждое значение данных.

В выборе наиболее подходящего формата визуализации данных следует опираться на шкалу, разработанную исследователями Уильямом Кливлендом и Робертом МакГиллом в 1984 г. [15]. Они опубликовали научную статью, в которой проанализировали различные способы представления данных, классифицировали и организовали их в соответствии с тем, насколько точны суждения, на основе которых создают график.

В соответствии с их классификацией, если цель графического изображения состоит в том, чтобы позволить читателям точно сравнить разные цифры, стоит отдать предпочтение гистограмме, пузырьковому или точечному графику. Это три графические формы, которые позволяют делать более точные выводы. В ситуации, когда нужно дать краткий обзор данных, а не обеспечивать точное сравнение (например, при подведении промежуточных итогов по регионам), авторы рекомендуют уменьшить масштаб шкалы.

В вопросе изучения графического представления данных полезно обращаться к ВЦИОМ – Всероссийский центр изучения общественного мнения, старейшая и наиболее популярная компания по проведению социологических опросов и сбору общественного мнения в России.

На сайте издания «ВЦИОМ новости» регулярно публикуют итоги масштабных опросов, в которых можно проследить темпоральные изменения в различных сферах жизни. Наиболее популярными рубриками, в которых авторы прибегают к приведению данных социологических опросов, являются «Рейтинги», «Аналитические обзоры», «Аналитические доклады», «Презентации» и «Инфографика». Однако данные именно о темпоральных изменениях демонстрируют по-разному.

В разделе «Рейтинги» большее внимание графическим моделям, представляющим темпоральные изменения в социально-экономической и политической сфере. В условиях невозможности публикации сопроводительного текста в материалах рубрики, графики и диаграммы становятся более информативными. Этот параметр позволяет повысить интерактивность графиков в каждой из публикаций. При наведении на каждую из точек появляется полная сводка информации, позволяющая составить мнение о процентном соотношении каждого из выбранных респондентами варианта ответа в определенном временном промежутке.

В рубрике «Аналитические доклады» и «Аналитические обзоры» ведущую роль в публикациях играет текст. Часто материалы выходят вообще без дополнительных диаграмм и графиков, а данные преподносят в форме перечисления, что неудобно для неподготовленного читателя.

В разделе «Инфографика» к графикам и таблицам добавляют дополнительные визуальные элементы: фотографии, иконки. Каждую часть информации сопровождает краткий текст. Публикации в этом разделе выглядят более ярко и подходят для массового читателя.

Рубрика презентации по наполнению схожа с предыдущей, однако всю найденную по теме информацию авторы представляют в виде нескольких слайдов, на которых также размещают визуализацию данных разного типа: таблицы, столбчатые диаграммы, диаграммы-кривые и пр.

Среди материалов «ВЦИОМ новости» более тщательного рассмотрения заслуживают рубрики «Рейтинги», «Презентации» и «Инфографика», так как публикации в них более интересно оформлены графиками и диаграммами, чем в разделах «Аналитические доклады» и «Аналитические обзоры», где иллюстрации к текстам почти не используют.

В рубрике «Рейтинги» для читателя доступны 19 рейтингов по социально-экономическим показателям и политическим рейтингам. 3 исследования архивированы, однако остальные доступны для читателей и постоянно пополняются данными. Их авторы публикаций получают по итогам ежедневного всероссийского телефонного опроса «ВЦИОМ-СПУТНИК».

Для иллюстрации данных, полученных за 4 и больше лет (многие опросы были начаты в 2020 г. и ранее), редакторы рубрики используют диаграммы-кривые, к каждой из которых применен метод LOWESS, чтобы даже резкие изменения в общественных мнениях было проще воспринимать визуально. Каждое из исследований проиллюстрировано несколькими системами графиков, дополненных таблицами с данными. Все точки на кривых интерактивны, и при наведении курсора мышки на определенный участок читатель может ознакомиться с данными, собранными в выбранный месяц, и сравнить их (*публикация «Индекс социальных ожиданий»*).

Рубрика «Инфографика» существует вдвое дольше, чем «Рейтинги», первая доступная публикация датируется 2010 г. В процессе анализа публикаций, вышедших в этом разделе в период с 2020 г. по 2023 г., было замечено, что большинство представленных в них данных можно изобразить в формате круговой диаграммы (*публикация «И швец, и жнец»*).

Однако авторы практически не используют этот тип диаграмм, отдавая предпочтение простому списку с дополнительными рисованными иконками. Возможно, это сделано для лучшей визуализации с точки зрения изображений, однако при большом количестве чисел, как в приведённом примере, может усложнить восприятие.

Наиболее популярный тип графика в рубрике «Инфографика» – линейчатая гистограмма. С её помощью дизайнеры и авторы публикаций показывают разрыв между количеством проголосовавших за отдельные варианты ответа, в случае если их было немного.

Так или иначе, в исследованной рубрике диаграммы и графики как изобразительные элементы отходят на второй план, уступая место рисункам и прочим художественным объектам. Однако в этой ситуации всё так же можно заметить, насколько визуализация данных в виде графика упрощает восприятие информации.

Раздел «Презентации» содержит публикации нестандартного для СМИ формата – все необходимые для освещения тезисы размещены на слайдах. Однако именно такой способ оформления позволяет задействовать наибольшее количество изобразительных элементов, в том числе, и графиков и диаграмм.

Так, наиболее частой диаграммой всё так же является гистограмма, которую используют для отражения результатов опросов с ограниченным количеством вариантов ответа. Однако в отличие от инфографики, на одном слайде в презентациях ВЦИОМ может быть размещено больше одной гистограммы, а иногда они могут сочетаться с другими видами графиков, для более полного преподнесения информации (*публикация «Роль средней школы в формировании исторической памяти молодого поколения»*, *диаграмма «Отношение к предложенным школьным новациям»*).

Как видно по предыдущей иллюстрации, в рубрике «Презентации» более часто, чем в каком-либо разделе, используют круговые диаграммы. Их немного видоизменяют для лучшего соответствия стилю презентации, однако суть и функционал остаются теми же (*публикация Искусственный интеллект в России: оценки и ожидания отрасли, диаграмма «Взгляд отрасли на развитие ИИ»*).

Интересным стал тот факт, что для визуализации данных в презентациях авторы использовали даже картодиаграммы, в тех случаях, когда это уместно. Этот непопулярный в публикациях ВЦИОМ способ построения диаграммы также нашёл свое место в публикациях издания (*публикация «Тяжелая ноша, или лишний вес и как с ним бороться»*, *диаграмма «ВЦИОМ 35 лет на рынке исследований»*).

Таким образом, в случае с изданием «ВЦИОМ новости», для демонстрации изменений темпоральных данных, авторы чаще всего выбирали столбчатую или линейчатую гистограмму, диаграмму-кривую или таблицу. Такая небольшая выборка типов графиков

соответствует стилю издания, а также уместна для визуализации в не интерактивном формате, что особенно важно для публикаций в рубриках «Инфографика» и «Презентации».

При сравнении тезисов, оформленных в формате диаграммы и в виде списка, можно заметить, что графическая форма представления данных упрощает их восприятие и делает публикацию более понятной для читателя, что, в том числе, сказывается и на популярности публикации.

Литература

1. Алексеев Д. Особенности исследования содержания динамических СМИ // URL: http://www.dalekseev.ru/sites/default/files/attachments/articles/issledovanie_temporalnyh_smi_2004.02.09.pdf (дата обращения: 20.02.2024).
2. Афанасьев, В. Н. Анализ временных рядов и прогнозирование: учебник / В. Н. Афанасьев; Ай Пи Ар Медиа – Саратов, Оренбургский гос. ун-т. – Оренбург: 2020 – 286 с.
3. Дутого Т. Радиальные диаграммы: значения, особенности построения, примеры применения // Международная научно-практическая конференция «Наука. Образование. Культура». - Комрат: 2021. - С. 104-108.
4. Кадочникова Е.И. Статистический анализ пространственных данных: учебное пособие / Е.И. Кадочникова, Ю.А. Варламова. – Казань: Издательство Казанского университета, 2023. – 140 с.
5. Конотопов П.Ю. Аналитика: методология, технология и организация информационно-аналитической работы / П.Ю. Конотопов, Ю.В. Курносов. - М.: РУСАКИ, 2004 г. - 512 с.
6. Круглова Е. Инфографика и визуализация данных. - Москва: Экономический факультет МГУ имени М.В.Ломоносова, 2018. - 63 с.
7. Мартынчук И. Г. Прогнозирование мультисезонных нагрузочных процессов в эластичных вычислительных системах // Изв. вузов. Приборостроение. 2023. Т. 66, No 11. С. 907—916. DOI: 10.17586/0021-3454-2023-66-11-907-916.
8. Муращенко С.В. Визуальная составляющая интернет-издания как фактор его позиционирования и продвижения / Муращенко С.В., Лабзина И.А., Трошина С.Г. - Современные исследования социальных проблем 2017: Том 8, No 3. - С. 111-124.
9. Романова И. К. Современные методы визуализации многомерных данных: анализ, классификация, реализация, приложения в технических системах // Наука и Образование. МГТУ им. Н.Э. Баумана. Электрон. журн. - 2016. - №No 03. - С. С. 133–167.
10. Сосновская Я. Типология визуализаций // Прикладной анализ данных в социальных науках URL: <https://education.yandex.ru/handbook/data-analysis/article/tipologiya-vizualizacij> (дата обращения: 18.02.2024).
11. Хименко В.И. Диаграмма рассеяния в анализе случайных потоков событий // Информационно-управляющие системы. - 2016. - №4. - С. 45-52.
12. Шевченко В.Э. Эффективность восприятия медиаинформации: визуализация контента // Информационная безопасность регионов. - 2013. - №2(13). - С. 45-52.
13. Яу Н. Искусство визуализации в бизнесе. Как представить сложную информацию простыми образами / Нейтан Яу; пер. с англ. Светланы Кировой. — М.: Манн, Иванов и Фербер, 2013. — 352 с.
14. Кливленд Уильям С. Робастная локально взвешенная регрессия и сглаживающие диаграммы рассеяния // Американская статистическая ассоциация. Том 74. № 368 (декабрь 1979 г.). С. 829-836.
15. Кливленд Уильям С. Надежное восприятие: теория, эксперименты и применение к разработке графических методов // Журнал Американской статистической ассоциации. - 1984, - Т. 79, № 387. - С. стр. 531-554.
16. Рупперт Дэвид. Статистика и анализ данных для финансового инжиниринга. Springer New York, 2011. С. 118.