

УДК 004.94

Ю.А. Сысоева

Харьковский национальный экономический университет имени Семена Кузнеця

КОМПЬЮТЕРНЫЕ ИНСТРУМЕНТЫ ВИЗУАЛИЗАЦИИ ДАННЫХ

В статье проанализированы имеющиеся классификации компьютерных инструментов визуализации данных. Предложена усовершенствованная классификация, которая учитывает назначение инструмента, вид программного продукта, аналитический и статистический функционал, наличие готовых шаблонов, интерактивность результатов визуализации и другие характеристики. В качестве примера разработанная классификация применена к инструменту визуализации данных Tableau Public.

Ключевые слова: визуализация данных, инструмент визуализации, классификация, Tableau Public.

Введение

Количество генерируемой в мире информации растет в геометрической прогрессии. Справиться с ее анализом помогает визуализация данных. За последние годы появилось множество компьютерных инструментов визуализации, позволяющих преобразовывать большие объемы абстрактных данных в форматы, эффективные для восприятия, познания и общения.

Но, как показал анализ, на сегодняшний день в посвященной этой теме литературе (например, [1, 2]) преобладают обзоры с перечнями инструментов и их кратким описанием, в то время как имеющиеся немногочисленные классификации являются неполными и в отдельных аспектах спорными.

Так, например, в работе [3] приведена такая классификация компьютерных инструментов визуализации:

- 1) функциональность инструмента:
 - пользователь (на кого рассчитан инструмент):*
 - профессионал;
 - новичок;
 - поддержка коллективной работы (сколько человек может одновременно разрабатывать модель):*
 - один пользователь;
 - много пользователей;
 - расширяемость (возможность внесения изменений):*
 - масштабируемость;
- 2) методы (характеристики используемых методов):
 - количество методов;*
 - тип метода;*
- 3) задача (характеристики решаемой задачи):
 - тип информации:*
 - количественный;
 - качественный;
 - предметная область:*
 - наука;

- бизнес;
- образование;
- 4) производитель (характеристики самого продукта):
 - фирма-производитель:*
 - отечественный;
 - иностранный;
 - тип продукта:*
 - коммерческий;
 - учебно-демонстрационный;
 - условно бесплатный.

Заметим, что в данной классификации функциональность инструмента ограничивается уровнем квалификации пользователя, поддержкой коллективной работы, расширяемостью и масштабируемостью.

В то время как наличие готовых шаблонов, интерактивность визуализации и другие важные характеристики вовсе не учтены. Деление же производителей инструментов на отечественных и зарубежных является, на наш взгляд, излишним – гораздо важнее многоязычность предлагаемых ими интерфейсов.

А в статье [4] предложена такая классификация инструментов визуализации и анализа данных:

- 1) категория:
 - анализ и построение диаграмм;*
 - приложение для ПК;*
 - геоинформационные системы (ГИС);*
 - библиотека;*
 - сетевой анализ;*
 - статистический анализ;*
 - анализ временных данных;*
 - извлечение данных;*
 - очистка данных;*
 - конвертирование данных;*
 - приложение/сервис для визуализации и анализа;*
 - облака слов;*
- 2) многоцелевая визуализация:
 - да;*
 - нет;*

3) отображение данных на карте:

да;

нет;

4) платформа:5) уровень квалификации пользователя:

1 – пользователи, умеющие работать с электронными таблицами;

2 – пользователи, продвинутые настолько, что не боятся потратить два часа на изучение нового приложения;

3 – опытные пользователи;

4 – пользователи с опытом программирования или специальными знаниями в области ГИС или сетевого анализа;

б) место хранения и обработки данных:

устройство пользователя;

внешний сервер;

публичный внешний сервер;

7) публикация результатов в веб:

да;

с плагином;

как изображение;

нет.

Трудно согласиться с утверждением, что такие разные явления как приложение для ПК, библиотека, сетевой анализ, статистический анализ и облако слов относятся к одному основанию, поскольку приложение для ПК и библиотека – это виды программного продукта, сетевой анализ и статистический анализ – назначение инструмента, а облако слов – метод визуализации данных.

Цель данной работы состоит в разработке новой классификации компьютерных инструментов визуализации, которая расширяет и улучшает предложенные ранее варианты.

Классификация инструментов

Предлагается усовершенствованная классификация инструментов визуализации данных:

1) назначение:

общего назначения;

специализированный;

вспомогательный.

К специализированным инструментам относятся, например, ГИС, инструменты бизнес-аналитики, сетевого анализа.

Вспомогательные же инструменты предназначены для подготовки данных к визуализации. Они позволяют извлекать данные, фильтровать их, конвертировать в нужные форматы и т.п.

2) уровень квалификации пользователя:

начальный;

средний;

продвинутый;

пользователь с опытом программирования или специальными знаниями;

3) программный продукт:

язык программирования;

библиотека;

приложение/сервис.

К языкам программирования относятся, например, R и Python, а к библиотекам – D3, ChartJS, Highcharts JS, Raphaël и др.

4) платформа:5) распространение:

коммерческий;

условно-бесплатный;

бесплатный;

6) многоязычный интерфейс:

да;

нет;

7) необходимость регистрации:

да;

нет;

8) многопользовательский режим:

да;

нет;

9) типы визуализируемых данных:

качественные;

качественные и количественные;

10) поддерживаемые базы данных и форматы файлов:11) количество реализованных методов визуализации:

один;

несколько.

В некоторых инструментах визуализации реализован лишь один метод, например:

Timeline JS3 дает возможность визуализировать данные лишь в виде хронологической линии,

Wordle – в виде облака слов, в то время как другие инструменты дают возможность реализовать целый ряд методов визуализации.

12) группы, к которым относятся реализованные методы визуализации:

визуализация данных;

визуализация информации;

визуализация концепций;

визуализация метафор;

визуализация стратегий;

комплексная визуализация.

Данная классификация основана на результатах работы [5], в которой предложена классификация методов визуализации по природе объекта визуализации, характеру объекта, уровню грануляции и характеру мышления.

13) аналитический и статистический функционал:14) наличие готовых шаблонов:

да;

нет;

15) отображение данных на карте:

да;
ограниченно;
с плагином;
нет;

16) создание дашбордов (dashboard):

да;
нет.

Дашборд (приборная доска) – это средство визуализации ключевых бизнес-индикаторов, позволяющее быстро получить представление о состоянии какого-либо аспекта деятельности предприятия или процесса (например, текущем финансовом состоянии, тенденции продаж, эффективности производства) [6].

Пример дашборда приведен на рис. 1.

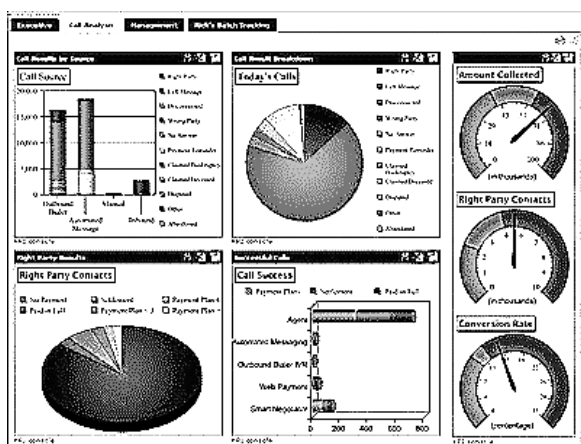


Рис. 1. Пример дашборда

17) интерактивная визуализация:

да;
нет;

18) место хранения и обработки данных:

устройство пользователя;
внешний сервер;
публичный внешний сервер;

19) форматы экспорта результатов:

20) публикация результатов в веб:

да;
с плагином;
как изображение;
нет.

В качестве примера применим предложенную классификацию к популярному компьютерному инструменту визуализации данных Tableau Public, разработанному американской компанией Tableau Software:

- 1) назначение: *специализированный (бизнес-аналитика);*
- 2) уровень квалификации пользователя: *продвинутый;*
- 3) программный продукт: *приложение/сервис;*
- 4) платформа: *Windows, OS X;*

5) распространение: *бесплатный;*

6) многоязычный интерфейс: *да;*

7) необходимость регистрации: *да;*

8) многопользовательский режим: *нет;*

9) типы визуализированных данных: *качественные и количественные;*

10) поддерживаемые базы данных и форматы файлов:

Excel Workbooks;

Text Files (.csv, .tab, .tsv, .txt);

Statistical Files (.sav, .sas7bdat, .rdata, .rda);

Odata;

Microsoft Windows Azure Marketplace Data Market;

11) количество реализованных методов визуализации: *несколько;*

12) группы, к которым относятся реализованные методы визуализации:

визуализация данных;

визуализация информации;

визуализация концепций;

13) аналитический и статистический функционал (рис. 2):

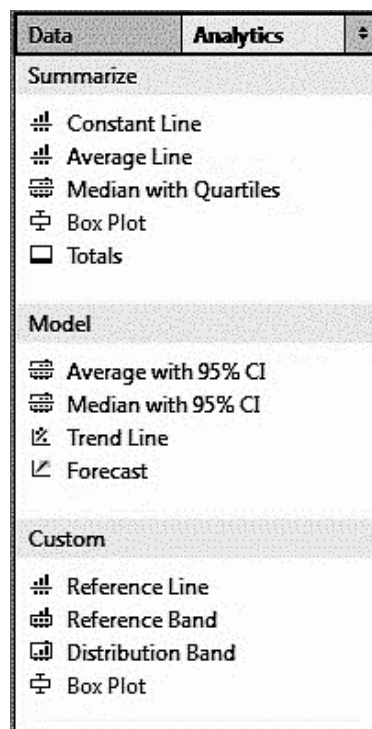


Рис. 2. Панель «Аналитика» инструмента визуализации Tableau Public

среднее значение;

медиана;

квартили;

«ящик с усами» (boxplot);

итоги;

доверительные интервалы;

линии тренда;

прогнозные оценки;

- 14) наличие готовых шаблонов: *нет*;
- 15) отображение данных на карте: *да*;
- 16) создание дашбордов: *да*;
- 17) интерактивная визуализация: *да*;
- 18) место хранения и обработки данных: *публичный внешний сервер*;
- 19) форматы экспорта результатов:
Tableau Workbook (.twbx);
Hyper Text Markup Language (.html);
Portable Document Format (.pdf);
Portable Network Graphics (.png);
- 20) публикация результатов в веб: *да*.

Выводы

В данной работе предложена классификация компьютерных инструментов визуализации данных, которая учитывает назначение инструмента, вид программного продукта, аналитический и статистический функционал, наличие готовых шаблонов, интерактивность результатов визуализации и другие характеристики. В качестве примера проведена подробная классификация инструмента Tableau Public. Приведенная классификация может быть использована при разработке системы поддержки принятия решений в сфере визуализации данных.

КОМП'ЮТЕРНІ ІНСТРУМЕНТИ ВІЗУАЛІЗАЦІЇ ДАНИХ

Ю.А. Сисоева

В статті проаналізовані існуючі класифікації комп'ютерних інструментів візуалізації даних. Запропонована вдосконалена класифікація, яка враховує призначення інструмента, вид програмного продукту, аналітичний та статистичний функціонал, наявність готових шаблонів, інтерактивність результатів візуалізації та інші характеристики. Як приклад розроблена класифікація застосована до інструменту візуалізації даних Tableau Public.

Ключові слова: візуалізація даних, інструмент візуалізації, класифікація, Tableau Public.

COMPUTER DATA VISUALIZATION TOOLS

Yu.A. Sysoyeva

In the article existing classifications of computer data visualization tools are analyzed. An improved classification is proposed. It takes into account the tool usage, the type of software, analytic and statistical functionality, the availability of ready-made templates, the interactivity of visualization results, and other characteristics. As an example, developed classification is used to classify data visualization tool Tableau Public.

Keywords: data visualization, visualization tool, classification, Tableau Public.

Список литературы

1. Крам Р. Инфографика. Визуальное представление данных / Р. Крам. – СПб. : Питер, 2015. – 384 с.
2. Яу Н. Искусство визуализации в бизнесе. Как представить сложную информацию простыми образами / Н. Яу. – М. : Манн, Иванов и Фербер, 2013. – 352 с.
3. Гаврилова Т.А. Визуальные методы работы со знаниями: попытка обзора / Т.А. Гаврилова, Н.А. Гулякина // Искусственный интеллект и принятие решений. – № 1. – 2008. – С. 15-21.
4. Machlis Sh. Chart and image gallery: 30+ free tools for data visualization and analysis [Electronic resource] / Sh. Machlis // Computerworld. – Access mode : <http://www.computerworld.com/article/2506820/business-intelligence/business-intelligence-chart-and-image-gallery-30-free-tools-for-data-visualization-and-analysis.html>.
5. Lengler R. Towards A Periodic Table of Visualization Methods for Management / R. Lengler, M. J. Eppler [Electronic resource]. – Access mode : http://www.visual-literacy.org/periodic_table/periodic_table.pdf.
6. Батьянов Д. Dashboards: как это делают профи. Часть 1 [Электронный ресурс] / Д. Батьянов. – Режим доступа : http://perfect-excel.ru/publ/excel/dashboard/dashboards_kak_ehto_delajut_profj_chast_1/11-1-0-36.

Поступила в редакцию 14.03.2016

Рецензент: д-р экон. наук, проф. А.И. Пушкарь, Харьковский национальный экономический университет имени Семена Кузнеця, Харьков.