

Научная статья
<https://doi.org/10.24412/2220-2404-2024-9-12>
УДК 316.1



ПЕРСПЕКТИВЫ ПРИМЕНЕНИЯ ИНСТРУМЕНТОВ BI В СОЦИОЛОГИИ

Куракова Т.В.

ФГБОУ ВО «Кубанский государственный университет»

Аннотация. Данная статья посвящена перспективам применения технологий Business Intelligence для обработки и представления данных в процессе проведения социологических исследований. В рамках данной работы был проведен сравнительный анализ и оценена эффективность нескольких BI продуктов с целью выбора наиболее подходящего для применения в социологии, а также проведено социологическое исследование на тему «Организация работы общественного транспорта в г. Краснодар» методом опроса, данные которого были обработаны, а результаты - визуализированы с использованием Microsoft Power BI. По результатам проведенного исследования были подготовлены рекомендации, направленные на улучшение качества услуг общественного транспорта в г. Краснодар, а также составлен отчет (дашборд), который может применяться для обработки и анализа данных любого социологического опроса.

Ключевые слова: business intelligence, Microsoft Power BI, Tableau, социология, социологический опрос, обработка данных, анализ данных, визуализация, общественный транспорт.

THE PROSPECTS OF BI USAGE IN SOCIOLOGY

Tatiana V. Kurakova

Kuban State Technological University

Abstract. This article is devoted to the prospects of Business Intelligence technologies usage for processing and presenting data in the process of conducting sociological research. As part of this work, a comparative analysis was conducted and the effectiveness of several BI products was assessed in order to select the most suitable for use in sociology. The sociological study on the topic of "Organization of public transport in Krasnodar" was conducted using a questionnaire, the data of which were processed, and the results were visualized using Microsoft Power BI. Based on the results of the study, recommendations were prepared aimed at improving the quality of public transport services in Krasnodar, and a report (dashboard) was compiled that can be used to process and analyze data from any sociological survey.

Keywords: business intelligence, Microsoft Power BI, Tableau, sociology, sociological survey, data processing, data analysis, visualization, public transport.

Введение.

Роль социологии как науки и как прикладного вида деятельности в эпоху информационного общества значительно возрастает. Социология позволяет получить четкое представление о структуре общества и проблемах его функционирования и развития, как глобального и всеобъемлющего, так и локального характера, присущего отдельным сообществам и социальным группам; формирует научно подкрепленную базу для прогнозирования общественных процессов и активно взаимодействует с экономикой, политологией, психологией и рядом естественнонаучных дисциплин [1].

Современные информационные технологии вносят позитивные изменения в процесс

сбора и обработки данных в рамках проведения социологических опросов, а именно:

- автоматизированные формы позволяют отправить анкеты без личных встреч с респондентами, а также не использовать бумагу, что позволяет экономить природные ресурсы и соответствует ценностям устойчивого развития;

- программы для ЭВМ позволяют значительно сократить время обработки полученных результатов;

- совместное применение автоматизированных форм для проведения исследования и программ для ЭВМ для обработки результатов способствуют тому, что в социологический опрос может быть вовлечено неограниченное количество респондентов, что позволяет получить как можно более точные результаты;

- применение графических инструментов позволяет создать презентабельные и понятные широкому кругу пользователей визуализации.

Таким образом, выбор инструмента обработки данных, полученных в рамках социологического исследования, является значимым вопросом и определяет, как время, потраченное на сбор и обработку данных, так и точность полученных результатов и выводов.

Обсуждение.

Процесс сбора и обработки данных в рамках проведения социологического опроса можно разделить на несколько этапов:

- непосредственную фиксацию ответов респондента;
- подсчет количества заполненных анкет и ответов на каждый вопрос;
- подсчет структуры полученных ответов в процентном соотношении;
- анализ динамики полученных результатов (если необходимо);
- представление полученных результатов [2].

Автоматизировать вышеупомянутые этапы (кроме непосредственной фиксации ответов респондента) позволяют инструменты business intelligence (BI). BI-системы – это программные продукты, которые собирают информацию из разных источников, обрабатывают её в соответствии с запросами и потребностями пользователя и представляют в виде удобных отчетов. BI-системы многократно упрощают анализ данных, а с их помощью конечные пользователи могут принимать решения быстрее и эффективнее.

В BI-системах происходит процесс под названием ETL: extract, transform, load – извлечение, преобразование и загрузка. Вот как можно описать этапы работы:

- сбор данных: BI-системы способны извлекать их из множества разных источников – CRM-систем, баз данных, ERP-систем, Excel, Google Sheets Google Forms и другие источники данных как автоматически, так и запросу пользователя;
- обработка данных: BI системы позволяют загружать в отчет разные типы данных (логические, числовые, текстовые, время и дата и т. п.) и проводить с ними различные манипуляции – изменять формат, структурировать, анализировать и проводить расчеты;
- визуализация данных: данный этап является завершающим результатом работы BI системы, в результате манипуляций с данными со-

здаются дашборды (интерактивные информационные панели, на которые выводятся важные для пользователей системы показатели), которые содержат результаты проведенного анализа в удобной для восприятия форме [3].

Также, BI системы позволяют подгружать новые данные в уже созданный дашборд – в котором они будут обрабатываться автоматически в соответствии с ранее созданным алгоритмом, а выведенные графические результаты панели также будут меняться.

В рамках данной статьи будут рассмотрены следующие вопросы:

- основные характеристики ведущих BI систем и выбор наиболее удобного продукта для обработки результатов социологических опросов;
 - обработка и анализ данных реального социологического опроса на тему «Качество услуг общественного транспорта в г. Краснодар».
- В настоящее время лидерами в сфере BI являются 2 программных продукта: Microsoft Power BI и Tableau. Рассмотрим достоинства и недостатки каждой BI системы.

Каждый из вышеупомянутых BI продуктов предоставляет следующие необходимые для работы пользователя функции высокого качества:

- автоматическое обновление дашбордов и анализ данных в реальном времени;
 - возможность настраивать роли для пользователей с разным уровнем доступа к дашбордам;
 - наличие коннекторов для подключения внешних данных;
 - широкие возможности визуализации.
- Однако есть и существенные отличия между системами.

Microsoft Power BI – это сервис, разработанный компанией Microsoft, поэтому все, кто знаком с данной операционной системой, смогут разобраться в работе с Power BI быстрее, чем с Tableau; и напротив, для пользователей, работающих на MacOS, не представляется возможности работать в Power BI [4].

Power BI предлагает пользователям бесплатную desktop версию с широким функционалом, в то время как бесплатное использование Tableau возможно только в интернет-версии.

Power BI за счет надстройки Power Query предоставляет больше возможностей для очистки и подготовки данных, в то время как в Tableau лучше загружать уже подготовленные данные.

Относительно визуализации, Tableau обладает более широким спектром возможностей,

чем Power BI: от стандартных круговых и столбчатых графиков до древовидных и пузырьковых диаграмм, размещения данных на карте и возможности наложения графиков друг на друга. Все это дает более высокий уровень понимания данных и их контекста.

Основным отличием Tableau от других систем является возможность смешивания данных – для отрисовки одного графика можно скомбинировать данные из разных баз данных и источников. Еще одним преимуществом системы Tableau является и то, что несколько пользователей одновременно могут работать над отчетами в режиме реального времени, абсолютно не мешая друг другу, что не предусмотрено в функционале Power BI. Также Tableau может обрабатывать большое количество строк, чем Power BI.

Таким образом несмотря на то, что обе системы предоставляют одинаковый базовый функционал, у них есть и существенные отличия, которые могут повлиять на выбор пользователя.

Power BI лучше выбирать пользователям, которые:

- используют в работе операционную систему Windows,
- используют для анализа 1-2 источника данных;
- хотели бы использовать бесплатную версию программы;
- в работе над отчетом задействовано небольшое количество пользователей.

В тоже время Tableau лучше выбрать пользователям, которые:

- работают на MacOS;
- в анализе используют значительно большее количество строк;
- в работе над отчетом задействовано большое количество пользователей, которые обрабатывают данные в реальном режиме, и несколько источников данных;
- готовы получить платную лицензию.

В рамках написания данной статьи была выбрана система Power BI.

Далее, перейдем к описанию процесса построения дашборда и анализу полученных результатов социологического исследования. Данная работа подразделяется на следующие этапы:

- непосредственное проведение опроса с использованием Microsoft Forms,
- обработку данных и построение дашборда (загрузка данных из источников, обработка посредством редактора запросов Power Query, создание мер и визуализация);
- построение выводов на основе полученных данных.

Результаты.

В мае-июне 2024 г. с целью выявления проблем организации работы общественного транспорта в г. Краснодаре было проведено пилотажное социологическое исследование методом опроса, в котором приняли участие 150 респондентов. Респонденты оставляли ответы посредством заполнения Microsoft Forms.

Для начала формирования отчета выгрузим совокупность полученных ответов в формате Excel из Microsoft forms, где каждая строчка соответствует одному респонденту. Далее загружаем полученные данные через редактор запросов Power Query в модель данных.

Для построения отчета и дальнейшей визуализации необходимо прописать меры. Для нашего отчета будет создано 2 меры: «Все ответы», описывающая общее количество ответов и «% от общего», описывающая процент от общего количества ответов.

Мера «Все ответы» описывается функцией `countrows` (со ссылкой на таблицу).

Мера «% от общего» описывается следующим образом (с заданием переменных):

```
var responses = [Все ответы]
var allresponses = CALCULATE([Все ответы], ALL(Responses), Responses[Вопросы] = SELECTEDVALUE(Responses[Вопросы]))
RETURN DIVIDE(responses, allResponses).
```

Далее, переходим к визуализации отчета - используем матрицы, круговые и линейные диаграммы для удобного предоставления данных (методом drag and drop). Финальный дашборд будет представлен на двух листах визуализации.

Первая часть дашборда представлена на рисунке 1.

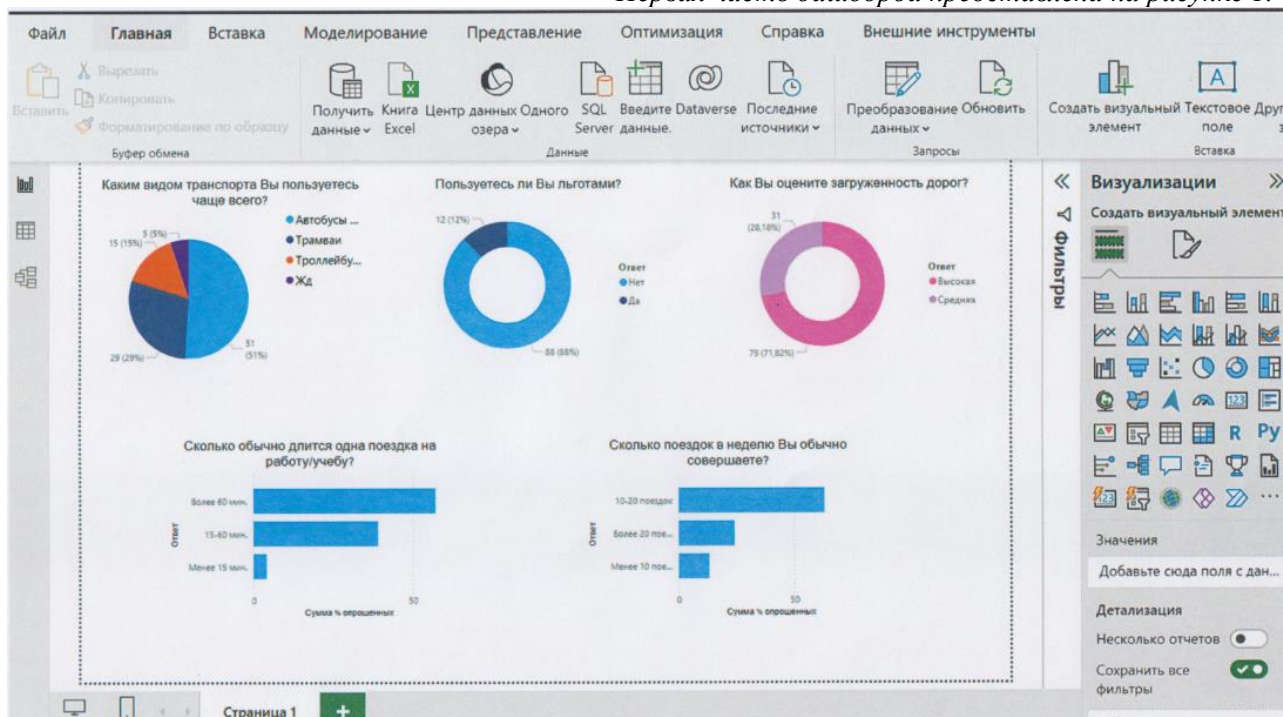


Рисунок 1 – Первая часть финального дашборда.
Вторая часть дашборда представлена на рисунке 2.

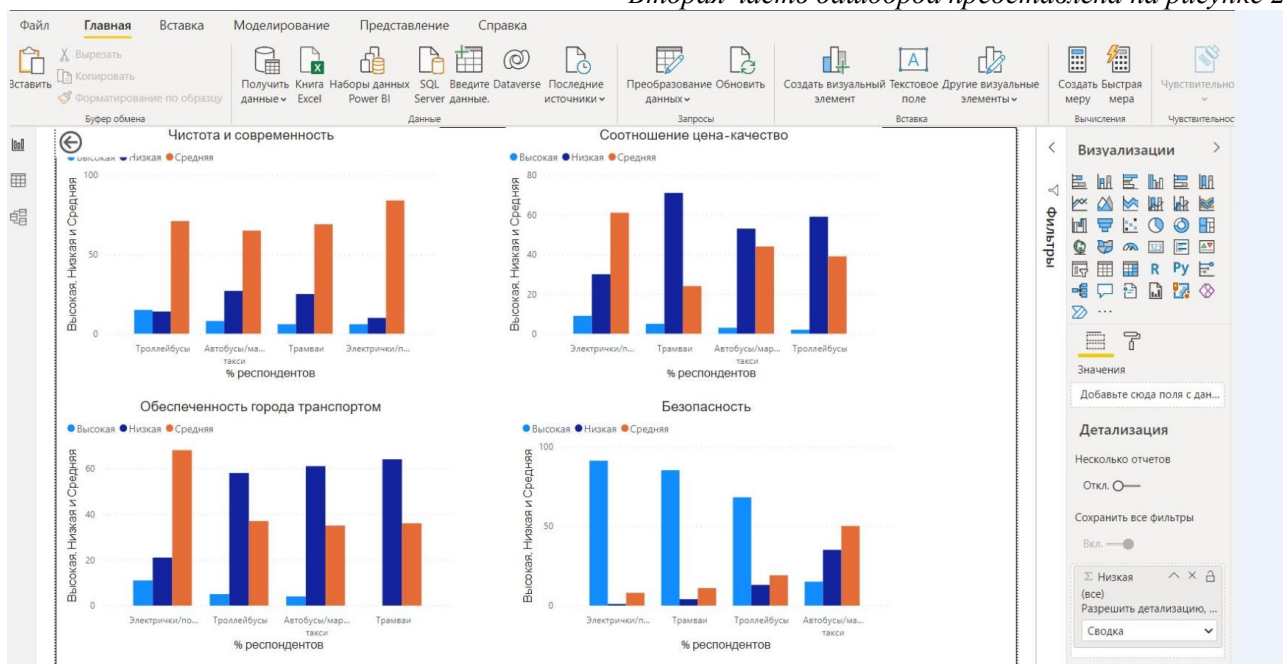


Рисунок 2 – Вторая часть финального дашборда.

Исходя из дашборда, проведенное социологическое исследование имеет следующие результаты.

51% респондентов на вопрос: «Каким видом общественного транспорта Вы пользуетесь чаще всего?» ответили, что автобусами и маршрутными такси (как городскими, так и междугородными), 29% опрошенных выбрали трамвай,

15% опрошенных наиболее часто пользуются троллейбусами, и только 5% выбрали поездка и электрички.

Ответы на вопрос: «Сколько времени Вы в среднем тратите на дорогу на учебу/работу при использовании общественного транспорта» и

«Сколько раз в неделю Вы пользуетесь общественным транспортом?» распределились следующим образом:

– 57% респондентов тратят на дорогу до места учебы/работы более 60 минут, 39% – от 15 до 60 минут и только 4% респондентов тратят на поездку менее 15 минут;

– 63% опрошенных совершают от 10 до 20 поездок на общественном транспорте в неделю, 24% – более 20 поездок и 13% – менее 10 поездок.

88% респондентов не пользуются льготами при использовании общественного транспорта (12% пользуются, соответственно). 79% опрошенных оценили загруженность дорог в дневное время (с 08:00 до 20:00) как «высокую» и 21% – как умеренную, вариант «низкая» не выбрал ни один респондент.

Также, респондентам было предложено оценить качество оказываемых транспортных услуг в разрезе следующих критериев: степень удовлетворенности чистотой и современностью транспортного парка, соотношением цена/качество по оказываемым транспортным услугам, наполненностью пассажирами во время поездки, обеспеченностью города общественным транспортом, безопасностью во время поездки.

Несмотря на то что чистоту и современность транспортного парка по всем видам указанного транспорта респонденты оценили, как «среднюю», соотношением цены и качества оказываемых услуг большинство остается недовольным (больше всего выбрали вариант «низкая»).

Согласно ответам респондентов, можно предположить, что жители Краснодарского края испытывают нехватку в отношении всех видов транспорта, за исключением поездов и электричек.

В то же время, безопасностью вождения большинство респондентов остались удовлетворены, наиболее безопасными видами транспорта, согласно ответам, являются поезда/электрички и трамваи.

Также, респондентам было предложено в свободной форме ответить на вопрос: Чтобы вы

улучшили в работе общественного транспорта? Наиболее часто встречающимися ответами являются (ответы указаны в порядке убывания):

– сокращение времени ожидания транспорта;

– снижение количества людей в поездке;
– установка и бесперебойная работа кондиционера в летнее время;

– снижение цен на транспортные услуги.

Таким образом, на основе полученных ответов можно сделать следующие выводы о ситуации с общественным транспортом в г. Краснодар:

– наиболее популярными видами транспорта в городе являются автобусы и трамваи,

– общественный транспорт используется очень активно, и средняя продолжительность поездки составляет около часа,

– в целом горожане не в достаточной степени удовлетворены услугами общественного транспорта по причине низкой обеспеченностью транспортными средствами и загруженностью дорог.

Заключение.

Созданная модель данных обработки социологического опроса имеет следующие преимущества:

– скорость обработки данных (расчеты и построение дашборда заняли 20 минут времени);

– графическая визуализация является динамической;

– в созданный дашборд можно подгружать новые ответы респондентов (если есть такая необходимость) и наблюдать в реальном режиме автоматический пересчет результатов;

– созданную модель данных можно многократно использовать для других социологических опросов.

Таким образом, применение систем бизнес-аналитики в социологии позволяет повысить скорость и улучшить качество обработки анкет, а также создать красивую и удобную визуализацию.

Конфликт интересов

Не указан.

Рецензия

Все статьи проходят рецензирование в формате double-blind peer review (рецензенту неизвестны имя и должность автора, автору неизвестны имя и должность рецензента). Рецензия может быть предоставлена заинтересованным лицам по запросу.

Литература:

Conflict of Interest

None declared.

Review

All articles are reviewed in the double-blind peer review format (the reviewer does not know the name and position of the author, the author does not know the name and position of the reviewer). The review can be provided to interested persons upon request.

1. Вялых Н. А. Социология для общества, общество для социологии или социология для социологии? // Вестник Томского Государственного университета. Философия. Социология. Политология. 2022. № 66. С. 139-149.

2. Ахтырский А. А. Цифровые методы в социологических исследования // Российский научный журнал Телескоп: журнал социологических и маркетинговых исследований. 2022. № 3. С. 53-59

3. Чибикова Т. В., Крумина К. В. Алгоритм использования программного продукта Microsoft Power BI в анализе и визуализации данных // Прикладная математика и фундаментальная информатика. 2023. № 2. С. 45-50.

4. Иванова И. Ю., Соловьев Д. О. Типовые процессы формирования операционных и аналитических отчетов // Наукосфера. 2024. № 6-1. С. 255-258.

References:

1. Vyalykh N. A. *Sociology for society, society for sociology or sociology for sociology?* // *Bulletin of Tomsk State University. Philosophy. Sociology. Political Science.* 2022. No. 66. Pp. 139-149.

2. Akhtyrsky A. A. *Digital methods in sociological research* // *Russian scientific journal Telescope: journal of sociological and marketing research.* 2022. No. 3. Pp. 53-59

3. Chibikova T. V., Krumina K. V. *Algorithm for using Microsoft Power BI software for data analysis and visualization* // *Applied Mathematics and Fundamental Informatics.* 2023. No. 2. Pp. 45-50.

4. Ivanova I. Yu., Soloviev D. O. *Typical processes for managing operational and analytical reports* // *Naukosphere.* 2024. No. 6-1. P. 255-258.

Информация об авторе:

Куракова Татьяна Валерьевна, кандидат философских наук, доцент кафедры социологии, правопедения и работы с персоналом, Институт фундаментальных наук, ФГБОУ ВО «Кубанский государственный технологический университет», ladomir20@mail.ru

Tatyana V. Kurakova, PhD in Philosophy, Associate Professor of the Department of Sociology, Law and Human Resources, Place of work: Department of Sociology, Law and Human Resources, Institute of Fundamental Sciences, Kuban State Technological University.