

УДК 330.47:004

**ОСОБЕННОСТИ АНАЛИЗА ДАННЫХ В BI-СИСТЕМАХ****Дьяков А.М.,***к.э.н. доцент,**ФГБОУ ВО «Алтайский государственный аграрный университет»,**Барнаул, Россия***Аннотация**

Современное распространение информационных технологий меняет парадигму анализа данных, в том числе экономического, это отражается на его схеме, технологии, особенностях. Для того, чтобы разобраться в этих изменениях, в статье рассматривается, такое направление информационных технологий, как BI-системы, используемые для анализа данных.

Основной задачей была поставлена необходимость разобраться с технологией работы BI-систем, на примере Microsoft Power BI, в том числе оценить, чем она отличается от традиционных схем обработки и анализа данных, принятых в экономике и статистике.

В процессе исследования было выявлено что, технология работы с BI-системами имеет определенные особенности, в сравнении, с привычными схемами аналитической обработки данных, и включает загрузку данных, создание информационной модели, создание облачных панелей мониторинга для вывода в Web, на смартфоны и др. устройства.

Соответственно, сделаны выводы, что привычная схема анализа данных «Собрать данные» – «Сформировать таблицы» – «Обработать и проанализировать» – «Сделать выводы и принять решения» кардинально меняется. В современном варианте, с использованием ИТ, ее можно представить в таком виде: «Подключиться к данным» – «Структурировать информацию и создать БД» – «Создать информационную модель обработки и анализа данных, проанализировать» – «Сделать выводы и принять решения».

**Ключевые слова:** программное обеспечение как сервис, компьютерный анализ данных, интернет-сервисы для анализа данных, организация анализа данных, «Business intelligence», BI-системы.

### ***FEATURES OF DATA ANALYSIS IN BI-SYSTEMS***

***Dyakov A.M.,***

*candidate of economic sciences, associate professor,*

*Altai state agricultural university,*

*Barnaul, Russia*

#### **Annotation:**

The modern spread of information technologies changes the paradigm of data analysis, including economic, this is reflected in its scheme, technology, features. In order to understand these changes, the article discusses such a direction of information technology as BI-systems used for data analysis.

The main task was the need to understand the technology of BI systems, using the example of Microsoft Power BI, including assessing how it differs from traditional data processing and analysis schemes adopted in economics and statistics.

In the course of the study, it was revealed that the technology of working with BI systems has certain features, in comparison with the usual schemes for analytical data processing, and includes downloading data, creating an information model, creating cloud dashboards for displaying on the Web, on smartphones, etc. devices.

Accordingly, it was concluded that the usual scheme of data analysis «Collect data» – «Make tables» - «Process and analyze» - «Draw conclusions and make decisions" is radically changing. In the modern version, using IT, it can be presented as follows: «Connect to data» – «Structure information and create a database» – «Create an information model for data processing and analysis, analyze» – «Draw conclusions and make decisions».

**Keywords:** software as a service, computer data analysis, Internet services for data analysis, organization of data analysis, «Business intelligence», BI-systems.

Современное распространение информационных технологий меняет парадигму анализа данных, в том числе экономического, это отражается на его схеме, технологии, особенностях. Для того, чтобы разобраться в этих изменениях, стоит рассмотреть такое направление информационных технологий, как BI-системы, предназначенные для анализа данных.

Исследование, касаемо уровня развития информационно-коммуникационных технологий [5], свидетельствует о высокой доле использования программного обеспечения специального назначения в регионах Сибирского федерального округа и в целом об уже достаточно высоком уровне проникновения информационных технологий в экономику.

Business intelligence (BI) – обозначение компьютерных методов и инструментов, обеспечивающих получение, структурирование, обработку информации, строить модели анализа данных, и использовать их для организации аналитической работы. Наиболее известными примерами BI-систем являются SAP BusinessObjects BI Suite, Oracle Analytics Cloud, Tableau Desktop, а также Microsoft Power BI, которая рассматривается в нашей работе.

Microsoft Power BI – это профессиональный инструмент для анализа и визуализации данных, программа может подключиться к различным источникам данных, связать их в едином массиве, пригодном для обработки и анализа информации, подготовить отчет в составе табличных и графических форм с интерактивными возможностями, а также разместить эти данные в облаке.

Важной особенностью программы является ее доступность для большого круга пользователей, это и бесплатная однопользовательская версия Power BI Desktop, и отсутствие необходимости наличия серьезных навыков программирования, и соответственно, привлечения IT-специалистов для решения необходимой задачи.

Power BI позволяет:

- консолидировать данные, которые могут находиться в разных источниках, в том числе разных базах данных и учетных системах;
- структурировать в формате БД и анализировать любые объемы данных;
- создать информационную модель анализа данных в виде различных форматов отчеты, использовать ее для обработки данных с помощью облачных вычислений.

Схема работы Microsoft Power BI представлена на рис. 1.

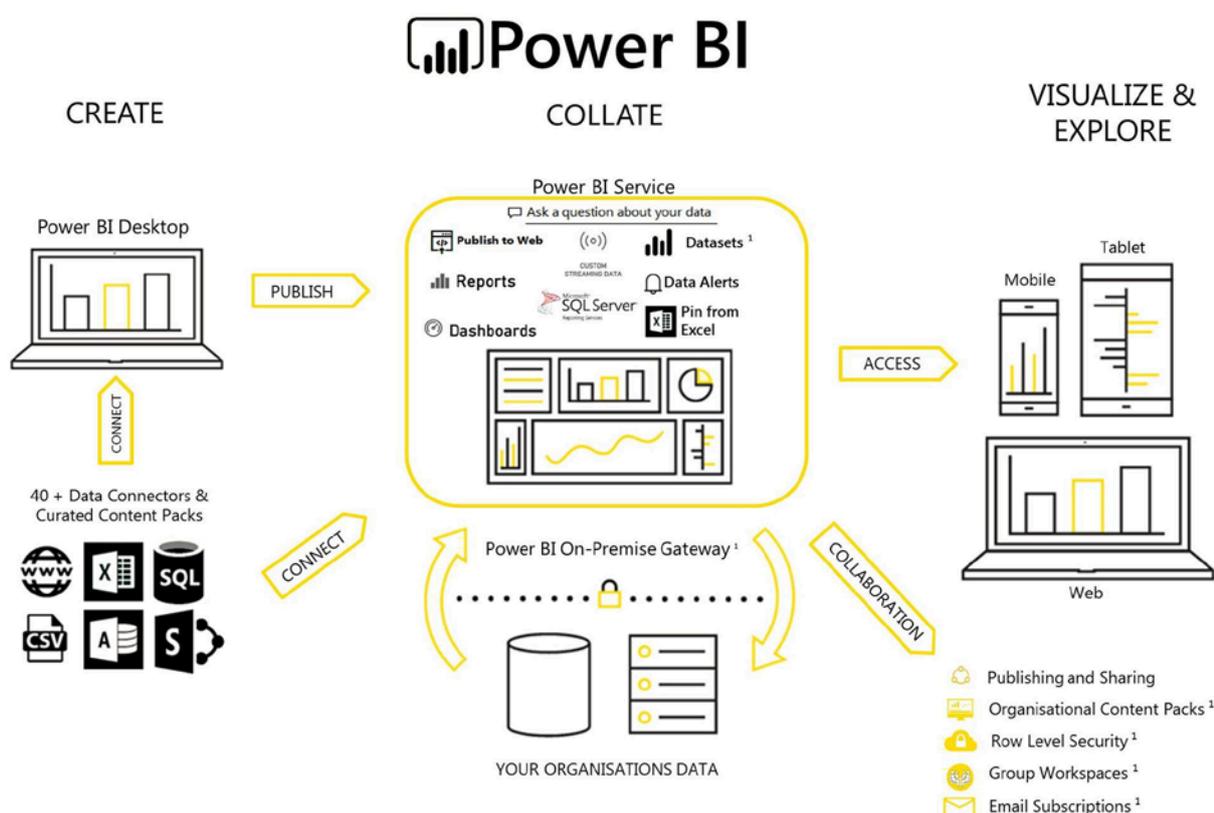


Рисунок 1 – Схема работы Microsoft Power BI [2]

Соответственно, технология работы с программой отличается от привычных схем аналитической обработки данных, и выглядит так: в Power BI Desktop загружаются данные, создается макет анализа, с необходимыми таблицами, графиками и т.п., далее этот макет загружается Power BI Service, где создаются панели мониторинга для вывода в Web, на смартфоны и т.д., настраивается обновление данных с помощью шлюзов, подключение Вектор экономики | [www.vectoreconomy.ru](http://www.vectoreconomy.ru) | СМИ ЭЛ № ФС 77-66790, ISSN 2500-3666

пользователей, после чего данные (и соответствующие формы) становятся доступны пользователям, и сторонним приложениям для последующей работы.

Для того, чтобы, изучить возможности и технологию работы с программой, нами использован большой массив данных о коронавирусе (за 2020-2021г. по всем странам мира), теме, которая последнее время достаточно актуальна. В сети интернет можно найти данные, в частности на веб-сервисе GitHub [1] по ссылке <https://covid.ourworldindata.org/data/owid-covid-data.xlsx>, которые подойдут для того, чтобы разобраться в схеме аналитической обработки данных данных.

Нашей задачей будет рассмотреть, как выглядит технология работы с BI-системой, на примере Microsoft Power BI, и как она отличается от традиционных схем обработки и анализа данных в экономике и статистике, в частности которые можно найти у ряда авторов [4].

Во-первых, поскольку, данные часто распределены по нескольким источникам, это могут быть базы данных, веб-сайты, файлы Excel, Access, CSV и др., используя Power BI Desktop, необходимо объединить данные из нескольких источников (рис. 2) и создать взаимосвязанный массив данных, позволяющий эти данные эффективно обрабатывать и делать различные вычисления.

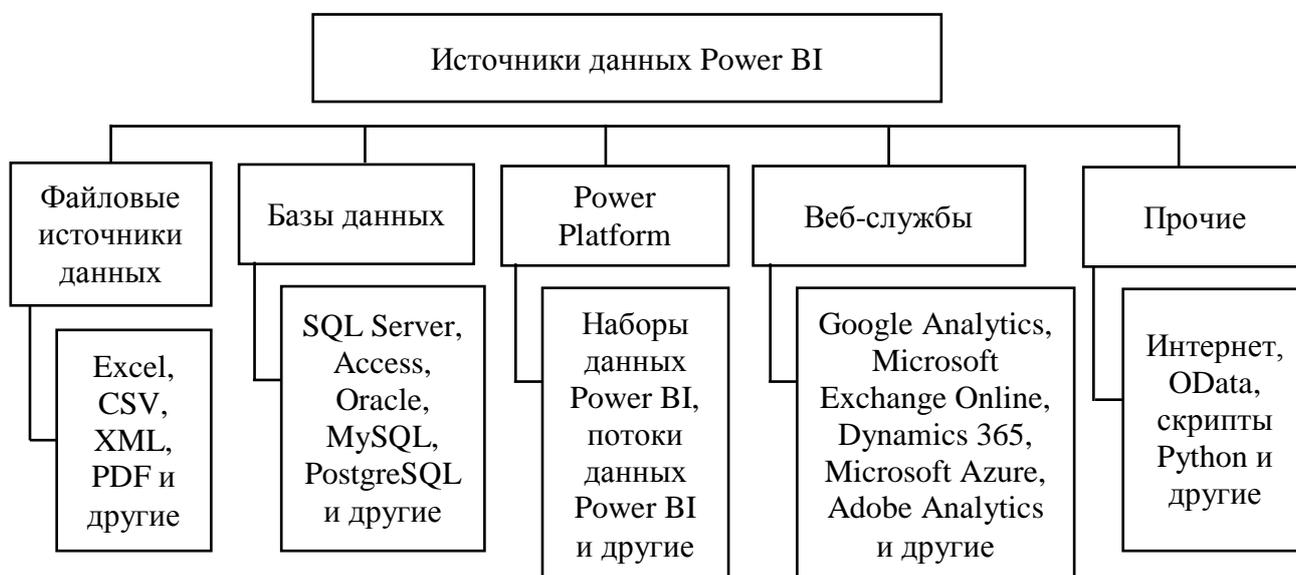
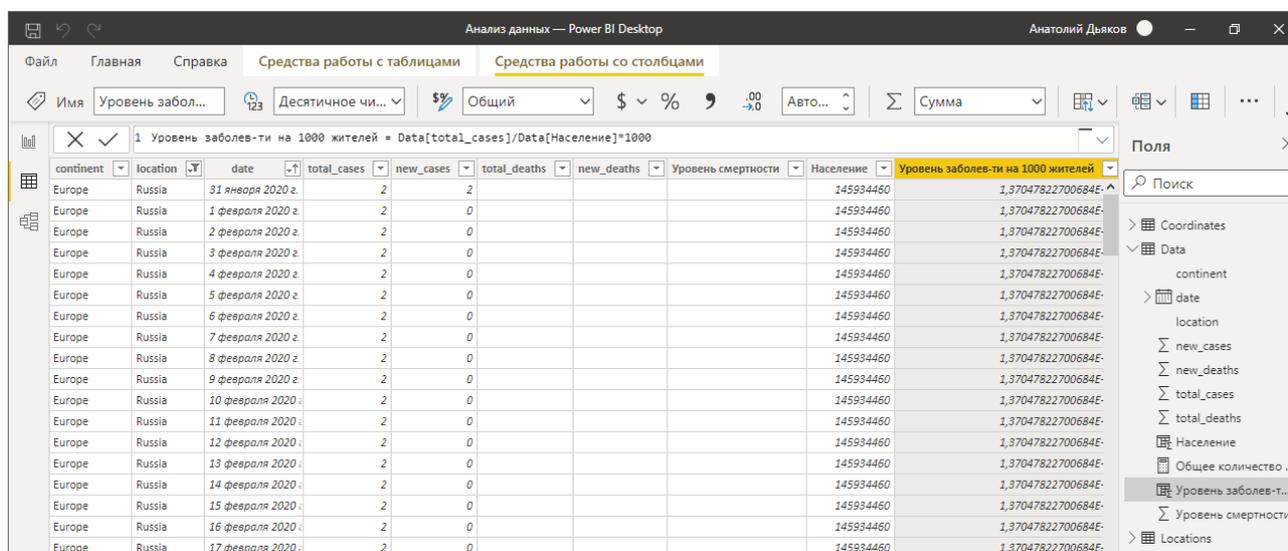


Рисунок 2 – Источники данных для Power BI [составлено автором]

Во-вторых, для того чтобы анализировать данные, часто требуется выполнить ряд подготовительных действий, в частности отфильтровать, отсортировать, удалить ненужные, разбить или соединить таблицы (рис. 3). Для этого используется язык формул Power Query M.

Кроме того, Power BI Desktop имеет специальный редактор запросов Power Query, он позволяет подключаться к одному или нескольким источникам данных, преобразовывать данные в соответствии с потребностями. В редакторе Power Query используется язык запросов DAX (Data Analysis eXpressions) – формульный функциональный язык запросов, для построения выражений и извлечения данных, которые расположены в табулярной модели – модели представления данных, схожей с многомерной OLAP-моделью.



continent	location	date	total_cases	new_cases	total_deaths	new_deaths	Уровень смертности	Население	Уровень заболеваемости на 1000 жителей
Europe	Russia	31 января 2020 г.	2	2				145934460	1,37047822700684E-
Europe	Russia	1 февраля 2020 г.	2	0				145934460	1,37047822700684E-
Europe	Russia	2 февраля 2020 г.	2	0				145934460	1,37047822700684E-
Europe	Russia	3 февраля 2020 г.	2	0				145934460	1,37047822700684E-
Europe	Russia	4 февраля 2020 г.	2	0				145934460	1,37047822700684E-
Europe	Russia	5 февраля 2020 г.	2	0				145934460	1,37047822700684E-
Europe	Russia	6 февраля 2020 г.	2	0				145934460	1,37047822700684E-
Europe	Russia	7 февраля 2020 г.	2	0				145934460	1,37047822700684E-
Europe	Russia	8 февраля 2020 г.	2	0				145934460	1,37047822700684E-
Europe	Russia	9 февраля 2020 г.	2	0				145934460	1,37047822700684E-
Europe	Russia	10 февраля 2020 г.	2	0				145934460	1,37047822700684E-
Europe	Russia	11 февраля 2020 г.	2	0				145934460	1,37047822700684E-
Europe	Russia	12 февраля 2020 г.	2	0				145934460	1,37047822700684E-
Europe	Russia	13 февраля 2020 г.	2	0				145934460	1,37047822700684E-
Europe	Russia	14 февраля 2020 г.	2	0				145934460	1,37047822700684E-
Europe	Russia	15 февраля 2020 г.	2	0				145934460	1,37047822700684E-
Europe	Russia	16 февраля 2020 г.	2	0				145934460	1,37047822700684E-
Europe	Russia	17 февраля 2020 г.	2	0				145934460	1,37047822700684E-

Рисунок 3 – Обработанные данные в программе [составлено автором]

При импорте нескольких таблиц, для того чтобы, провести анализ, необходимы связи между ними (рис. 4). После соединения таблиц можно работать с данными в обеих таблицах так, будто они находятся в одной.

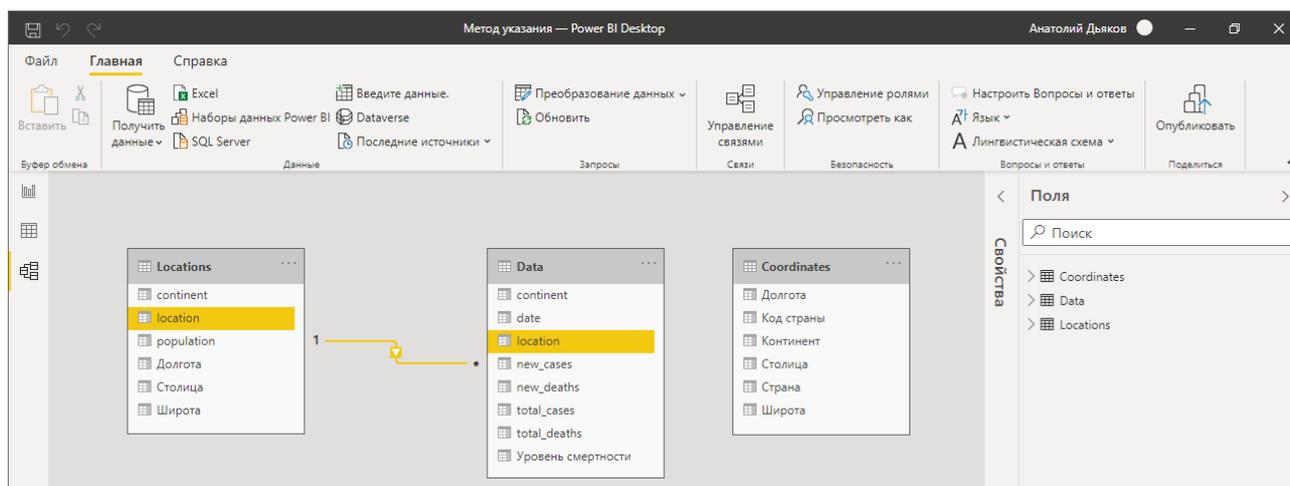


Рисунок 4 – Модель данных в Power BI Desktop [составлено автором]

Power BI Desktop значительно упрощает создание таких связей, во многих случаях, функция автообнаружения сделает все автоматически. Однако в некоторых случаях может потребоваться создать связи вручную или внести некоторые изменения в полученные связи.

В-третьих, непосредственно для анализа данных используются визуализации (также называемые визуальными элементами), которые отображают, в определенном формате (диаграммы, таблицы, карты и др., рис. 5), данные из модели Power BI Desktop.



Рисунок 5 – Визуализации данных в Power BI [составлено автором]

В Power BI Desktop есть редактор, который предназначен для создания отчетов, создание любого визуального элемента производится путем компоновки данных из таблиц базы данных.

В-четвертых, нужно созданную информационную модель загрузить в службу Power BI Service [3], это облачный интернет-сервис, на настоящий момент, это очень популярная форма предоставления программного обеспечения пользователям [6].

Данная служба позволяет закрепить созданные визуальные элементы на панели мониторинга, находящегося в облаке, нами была создана следующая панель мониторинга – рис 6.

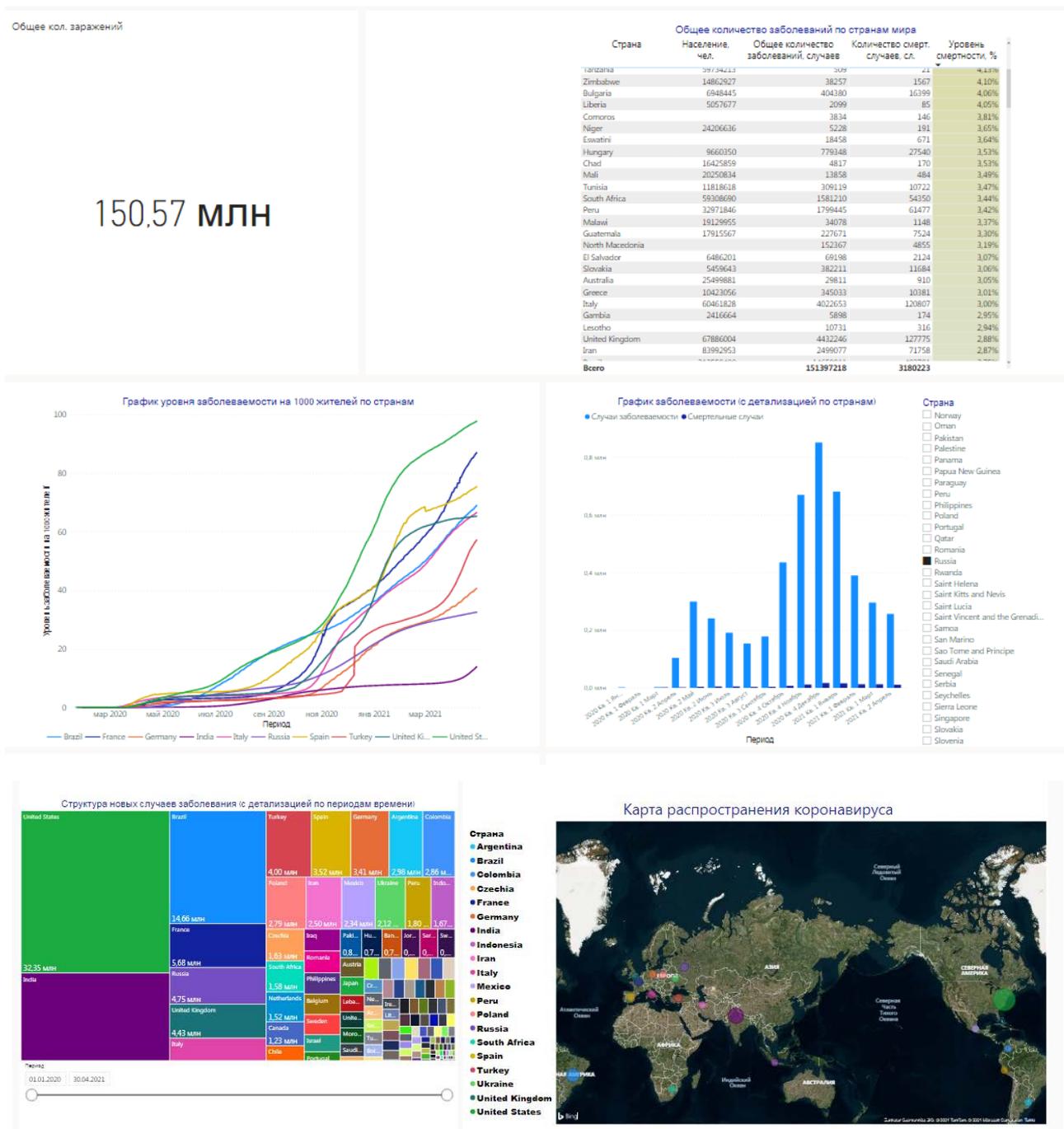


Рисунок 6 – Панель мониторинга в Power BI Service [составлено автором]

Созданную панель мониторинга можно передать в мобильное приложение, для этого необходима настройка, прежде всего в плане форматирования. Перед использованием, приложение «Power BI для Android», необходимо установить и войти в него под своей учетной записью.

В качестве вывода, можно отметить, что привычная схема анализа данных «Собрать данные» – «Сформировать таблицы» – «Обработать и проанализировать» – «Сделать выводы и принять решения» кардинально меняется. Сейчас, это уже выглядит, примерно, таким образом: «Подключиться к данным» – «Структурировать информацию и создать БД» – «Создать информационную модель обработки и анализа данных, проанализировать» – «Сделать выводы и принять решения». Соответственно, важным является, понять эти новые принципы, научиться ими пользоваться, применяя BI-системы.

### Библиографический список:

1. GitHub, веб-сервис для хостинга IT-проектов и их совместной разработки [Электронный ресурс]. – Режим доступа – URL: <https://github.com/> (дата обращения: 24.05.2021).
2. Microsoft Docs, библиотека технической документации для конечных пользователей, разработчиков и IT-специалистов, работающих с продуктами Microsoft [Электронный ресурс]. – Режим доступа – URL: <https://docs.microsoft.com/en-us/power-bi/> (дата обращения: 24.05.2021).
3. Power BI Service, облачный сервис платформы Power BI от Microsoft [Электронный ресурс]. – Режим доступа – URL: <https://app.powerbi.com/> (дата обращения: 24.05.2021).
4. Белокурено Н.С. Использование методов статистики в экономическом анализе / Н.С. Белокурено // Приоритеты мировой науки: эксперимент и научная дискуссия: сборник материалов Международной научно-практической конференции (15 июня 2020 г.). – Кемерово: ЗапСибНЦ, 2020. – С. 38-41.
5. Галкин Д.Г. Оценка уровня развития информационно-коммуникационных технологий в регионах Сибирского федерального округа / Д.Г. Галкин // Международный журнал гуманитарных и естественных наук. – 2020. – №10-2 (49)

6. Дьяков А.М. Интернет-сервисы поддержки бизнеса, возможности и проблемы развития / А.М. Дьяков, В.А. Якоб // Вектор экономики. – 2019. – №6 [Электронный ресурс]. – Режим доступа – URL: [http://www.vectoreconomy.ru/images/publications/2019/6/innovationmanagement/Dyakov\\_Yakob.pdf](http://www.vectoreconomy.ru/images/publications/2019/6/innovationmanagement/Dyakov_Yakob.pdf) (дата обращения: 24.05.2021).

*Оригинальность 94%*