

Санкт-Петербургский государственный университет
Экономический факультет
Кафедра информационных систем в экономике

КУРСОВАЯ РАБОТА

По направлению 38.03.05 – «Бизнес-Информатика»

ВНЕДРЕНИЕ ИНТЕРАКТИВНОЙ ПЛАТФОРМЫ ВИЗУАЛЬНОГО ПРЕДСТАВЛЕНИЯ ДАННЫХ ДЛЯ ПОДДЕРЖКИ ПРИНЯТИЯ РЕШЕНИЙ

Выполнил:

Бакалавриант 3 курса, группы 20.Б12-Э

Моисеенко Татьяна Михайловна

_____/Подпись/

Научный руководитель

Доцент, канд. экон. наук

Иванова Виктория Валерьевна

_____/Подпись/

Санкт-Петербург
2023

Оглавление

ВВЕДЕНИЕ	3
Глава 1. Теоретические аспекты применения BI-приложения.....	5
1.1. Понятие, особенности и общая характеристика BI.....	5
1.2. Обзор на зарубежные BI-системы.....	8
1.3. Обзор на российские/отечественные BI-системы	12
1.4. Обоснование выбора BI-платформы.....	15
Глава 2. Проектирование BI-приложения для анализа движения косметических товаров.....	17
2.1. Формирование задач проектирования и целей разработки BI-приложения	17
2.2 Анализ информационных потребностей и требований пользователей.....	18
2.3. Предварительная подготовка данных.....	18
2.4. Построение модели данных	23
Глава 3. Разработка BI-приложения.	29
3.1. Реализация BI-приложения в Power BI.....	29
3.2. Анализ экономической эффективности.....	34
Заключение.....	36
Список использованных источников:.....	37

ВВЕДЕНИЕ

В современном мире с огромной скоростью увеличивается объем данных, который требуется незамедлительно анализировать и обрабатывать. Анализ данных в реальном времени становится необходимостью для многих компаний, и это, в свою очередь, приводит к развитию новых инструментов и технологий. Возрастает спрос на системы класса Business Intelligence (BI) как средство, которое позволяет эффективно обрабатывать данные и использовать их для принятия управленческих решений. BI позволяет структурировать данные, выявлять тенденции, проводить анализ и прогнозирование, а также быстро реагировать на изменения внешней и внутренней среды. Системы BI помогают бизнесу принимать обоснованные и обдуманные решения на основе анализа реальных данных.

Пользователи BI приложений фокусируются на анализе данных, а не на их сборе, поскольку данные уже были подготовлены для анализа. BI позволяет аналитикам проводить интерактивный анализ данных и получать ответы на вопросы руководства и сотрудников разных уровней в режиме реального времени, что значительно повышает эффективность работы.

Актуальность и практическая значимость выбранного направления исследования обусловлена следующими причинами:

Наглядная аналитика обладает большим потенциалом в плане анализа данных. Она позволяет просматривать и анализировать большие объемы информации в удобном и доступном формате, позволяет смотреть на проблему со всех сторон и формировать адекватное решение, которое учитывает интересы всех заинтересованных сторон. Благодаря этому, компания может снизить риски и достичь более стабильных и устойчивых результатов.

Проблема: в условиях цифровой экономики, принятие решений требует от менеджеров специальных знаний и навыков, которые не всегда являются очевидными. Кроме того, для эффективного управления необходима определенная технологическая инфраструктура, включающая оперативный доступ к данным, отражающим текущее состояние компании.

В рамках вышеуказанной проблемы сформирована цель настоящей курсовой работы.

Цель курсовой работы – разработка интерактивной платформы визуального представления данных (BI-приложения) для поддержки принятия решений компании.

Исходя из цели исследования, обозначим следующие задачи курсовой работы:

1. Определить терминологическую базу
2. Провести анализ BI-систем, определить критерии выбора программного продукта
3. Определить правила и этапы разработки информационных панелей
4. Разработать BI-приложение для анализа движения товаров со склада в филиалы торговой сети.
5. Дать экономическую оценку приложению.

Предметом исследования является склад компании ООО «Дрогери Ритейл»

Информационная база представляет собой работы современных исследователей в области интерактивных платформ визуального представления данных, а также отчеты крупных консалтинговых компаний и аналитических агентств.

Структура курсовой работы представляет собой введение, 3 главы, заключение и список использованной литературы.

Во введении обоснована актуальность курсовой работы, раскрыта цель, сформулированы задачи, указаны объект и предмет исследования, прописана структура работы.

Первый раздел курсовой работы посвящен теоретическому анализу использования ВІ приложений в бизнесе. В рамках данного раздела изучены основные концепции и термины, связанные с ВІ-технологиями, проанализированы их возможности в контексте бизнес-процессов. Были выявлены наиболее эффективные подходы к внедрению ВІ-приложений, а также проведен обзор наиболее популярных инструментов и систем ВІ-аналитики, используемых на сегодняшний день.

Во второй главе курсовой работы дана краткая характеристика компании-заказчика ВІ-приложения ООО «Дрогери ритейл».

В третьем разделе произведена разработка ВІ-приложения для анализа движения товаров со склада в филиалы сети.

Заключение представляет собой результаты проведенной работы.

Глава 1. Теоретические аспекты применения BI-приложения.

1.1. Понятие, особенности и общая характеристика BI

BI-системы предоставляют собой мощный инструмент для анализа информации, агрегированной из баз данных различных форматов. Они позволяют существенно улучшить процесс принятия решений, представляя информацию под разным углом, то есть срезами, что безусловно удобно. В свою очередь, это позволяет быстрее и эффективнее выявлять драйверы изменения показателей, факторы, влияющие на бизнес, и определять возможные стратегии развития компании.

Качественная информационная поддержка является одним из ключевых факторов успешного функционирования средних и крупных предприятий. В настоящее время большинство крупных компаний в России уже внедрили или находятся в процессе внедрения специализированных информационных систем, таких как ERP, CRM, MES и другие. Такие системы позволяют эффективно автоматизировать и управлять бизнес-процессами, оптимизировать ресурсы и повысить производительность труда. [3, стр.59]

ERP (Enterprise Resource Planning) - системы способны управлять всеми процессами внутри компании, начиная от производства и заканчивая управлением ресурсами, финансами и клиентскими отношениями.

CRM (Customer Relationship Management) - системы предоставляют инструменты для управления клиентской базой, маркетинга, продаж и обслуживания клиентов.

MES (Manufacturing Execution System) - системы управляют производственными процессами на уровне производственной линии, позволяя оптимизировать производственные операции, уменьшить затраты на производство и повысить качество продукции.

При проектировании BI-системы необходимо учитывать ряд особенностей, которые связаны с требованиями бизнеса и интеграцией новой системы в существующую ИТ-архитектуру предприятия.

BI-системы способны обрабатывать большие объемы данных и анализировать их в режиме реального времени. В связи с этим, требования к производительности и надежности системы являются критически важными. Кроме того, BI-системы должны быть гибкими и настраиваемыми под конкретные нужды бизнеса.

Внедрение BI-системы требует интеграции с другими системами, уже используемыми на предприятии. На данном этапе могут возникнуть трудности, так как существующие системы могут иметь разные структуры данных и форматы обмена информацией. Интеграция новой системы должна быть осуществлена без нарушения работы уже существующих систем и минимальными затратами на модернизацию инфраструктуры.

BI-системы могут быть реализованы как внутреннее решение, так и в виде облачного сервиса. Каждый вариант имеет свои особенности и требования к инфраструктуре и безопасности. Решение о выборе типа реализации BI-системы должно быть основано на требованиях к производительности, доступности и защите информации.

BI-система предоставляет обширный набор инструментов визуализации данных, обеспечивая пользователей возможностью создания многочисленных видов диаграмм,

информационных панелей и других пользовательских объектов для анализа данных. Система позволяет пользователям взаимодействовать с данными непосредственно из графических объектов, что дает возможность быстро реагировать на изменения и проводить мгновенный анализ.

Сформируем основные требования к BI- системам:

- Гибкость и адаптивность.

Это означает, что BI-системы должны легко масштабироваться, а также быстро адаптироваться к изменениям в бизнес-процессах и требованиям пользователей.

- Простота использования.

Важно, чтобы BI-системы были интуитивно понятными и простыми в использовании.

- Масштабируемость.

BI-системы должны быть масштабируемыми, чтобы обрабатывать большое количество данных и поддерживать рост бизнеса.

- Высокая скорость и производительность.

BI-системы должны обеспечивать высокую скорость и производительность при выполнении запросов и генерации отчетов. Быстрый доступ к данным является ключевым фактором для принятия правильных решений в бизнесе.

- Безопасность и защита данных.

Данные, которые хранятся в BI-системах, могут быть конфиденциальными, поэтому важно, чтобы BI-системы обеспечивали высокий уровень безопасности, предотвращали несанкционированный доступ и утечки информации.

- Интеграция с другими системами.

BI-системы должны интегрироваться с другими системами, чтобы обеспечивать единый доступ к данным и совместную работу с другими приложениями.

- Поддержка различных источников данных.

BI-системы должны поддерживать различные источники данных, чтобы обеспечивать полноту и точность анализа. Данные могут поступать из различных источников, таких как базы данных, системы управления контактами, системы управления ресурсами предприятия и т.д.

Сформулируем задачи решаемые BI-приложениями для осуществления процесса принятия решений:

- Сбор и анализ данных для выявления тенденций и паттернов в продажах, производстве и бизнес-процессах.
- Анализ финансовых результатов и определение ключевых факторов, влияющих на рентабельность бизнеса.

- Выявление проблем и узких мест в бизнес-процессах и разработка рекомендаций для их устранения (оценка эффективности инвестиций, маркетинговых кампаний, показатели качества продукции, службы логистики и другие) [4].
- Планирование и бюджетирование. BI-приложения позволяют прогнозировать будущие потребности, определять возможные риски, следить за изменениями на рынке, анализировать тенденции и определять потенциальные возможности, принимать решения на основе данных.

Для того чтобы максимизировать отдачу от внедрения BI-технологий, компании должны иметь системы измерения бизнес-процессов [2]. Без знания показателей, на которые стоит ориентироваться руководству при принятии решений, невозможно делать оценки, прогнозы и выработать эффективную стратегию развития компании.

В книге “Analytics at Work” Томаса Дэвенпорта [5] описаны следующие этапы формирования BI-архитектуры для анализа данных при использовании BI:

1. Выгрузка данных из источников, определение бизнес-требований и целей BI-системы.
2. Выбор инструментов для работы с данными.
3. Разработка отчетности и панелей визуализации данных(дашбордов).
4. Формирование управленческих решений на основе бизнес-анализа.

Бизнес-аналитика (BI) — это обобщающий термин, который объединяет архитектуры, инструменты, базы данных, аналитические инструменты, приложения и методологии. [6]. Основная цель BI - обеспечить интерактивный доступ (иногда в режиме реального времени) к данным, обеспечить возможность манипулирования данными и предоставить бизнесу, менеджерам и аналитикам возможность проводить соответствующий анализ. Процесс BI основан на преобразовании данных в информацию, затем в решения и, наконец, в действия. Можно сделать вывод о том, что BI не является синонимом к понятию «бизнес-анализ», а входит в него [8].

В работе "Руководство по бизнес-аналитике: от интеграции данных к аналитике" Рика Шермана [7] BI представлен в виде совокупности понятий, которые можно разбить следующие группы:

1. Компоненты BI
 - Системы хранения данных (Data Warehouses) [9]
 - Инструменты для извлечения, трансформации и загрузки данных (ETL) [10]
 - Инструменты аналитики данных (OLAP) [11]
 - Визуализация данных (BI-отчеты, дашборды, диаграммы)
 - Интеллектуальный анализ данных (Data Mining)
 - Расширенная аналитика (Advanced Analytics)
2. Технологии BI:
 - Базы данных и системы управления базами данных (DBMS) [12]
 - Языки запросов (SQL, MDX)
 - Инструменты ETL (Informatica, Talend, Pentaho)

- Инструменты аналитики данных (SAP BusinessObjects, IBM Cognos, Microsoft Power BI)
- Инструменты визуализации данных (Tableau, QlikView, Microsoft Power BI)
- 3. Методологии BI:
 - Методология построения хранилищ данных (Kimball, Inmon)
 - Методология разработки BI-решений (Agile BI, Waterfall)
 - Методология работы с данными (Data Governance, Data Quality Management)
- 4. Типы данных используемых в BI:
 - Транзакционные данные
 - Данные из социальных сетей
 - Данные из IoT (Internet of Things)
 - Географические данные
 - Real-time данные
 - Структурированные данные
 - Неструктурированные данные
 - Полуструктурированные данные

В работе Стива Уильямса «Стратегия бизнес-аналитики и аналитика больших данных» [9] автор выделяет несколько специализированных BI-решений, которые могут быть использованы для мониторинга и контроля следующих бизнес-процессов:

- Эффективность управленческой деятельности
- Финансовая деятельность предприятия
- Мониторинга логистики и управления цепями поставок
- Управление рисками
- Управление персоналом

Таким образом, можно сделать вывод, что технология BI представляет собой мощный инструмент для обработки, анализа и визуализации данных. Комбинация компонентов, технологий и методологий BI позволяет бизнес-аналитикам и руководителям принимать обоснованные решения и повышать эффективность деятельности компании.

1.2. Обзор на зарубежные BI-системы

Конечной целью любой компании является успешное управление бизнесом и достижение поставленных целей. Для этого необходимо иметь доступ к качественной информации о бизнес-процессах и о том, как эти процессы влияют на результаты компании. В современном мире, где объемы данных растут с каждым днем, для успешного управления бизнесом необходимы инструменты, позволяющие быстро и точно анализировать большие объемы данных. Именно здесь важную роль играют BI-приложения.

При выборе BI-приложения необходимо учитывать множество факторов, таких как функциональность, скорость работы, стоимость, простота использования и многое другое. При выборе BI-приложения компании необходимо учитывать свои потребности, доступные бюджетные ресурсы, а также уровень сложности и доступность инструмента для сотрудников компании.

В данной части будут рассмотрены несколько популярных BI-приложений, будет проведен анализ их преимуществ и недостатков. По итогу анализа, мы сможем выбрать лучшее BI-приложение, соответствующее вашим потребностям и бюджету.

Tableau — это американская BI-платформа, которая предоставляет инструменты для визуализации данных, создания отчетов и дашбордов. Tableau является одним из наиболее популярных инструментов в своей категории, благодаря своей простоте использования и богатому функционалу.

Несмотря на то, что Tableau имеет ряд преимуществ, он также имеет и недостатки, которые необходимо учитывать при выборе инструмента для анализа данных.

Преимущества Tableau:

1. Простота в использовании.

Tableau обладает простым и интуитивно понятным интерфейсом, что позволяет быстро и эффективно анализировать данные без дополнительного обучения.

2. Богатый функционал.

Tableau предоставляет множество инструментов для визуализации данных, включая таблицы, диаграммы, карты и многое другое.

3. Возможность работы с различными источниками данных.

Tableau поддерживает работу с большим количеством источников данных, включая базы данных, файлы Excel и другие.

4. Возможность работы с данными в реальном времени.

Tableau позволяет работать с данными в реальном времени, что позволяет быстро реагировать на изменения в бизнес-процессах.

Недостатки Tableau:

1. Высокая стоимость.

Tableau является одним из самых дорогих BI-приложений на рынке, что может оказаться недоступным для небольших компаний.

2. Ограниченная возможность настройки.

В Tableau имеются некоторые ограничения в настройке отчетов и дашбордов, что может ограничивать возможности адаптации приложения под конкретные нужды компании.

3. Требования к аппаратным ресурсам.

Для работы Tableau необходимо достаточно мощное оборудование, что может потребовать дополнительных затрат на приобретение и обслуживание серверов.

4. Сложность интеграции с другими системами.

В случае, если компания использует несколько различных приложений и систем, интеграция Tableau с ними может потребовать дополнительных усилий и ресурсов.

Qlik — это бизнес-интеллект-платформа, которая позволяет пользователям извлекать, обрабатывать и анализировать данные из разных источников.

У компании 2 продукта:

- QlikView
- QlikSense

QlikView – корпоративная платформа, а QlikSense – платформа для персональной работы[15]. Qlik Sense имеет более простой и интуитивно понятный интерфейс и развитый подход Drag&Drop, который упрощает создание отчетов и дашбордов. Однако, каждый отчет в QlikSense имеет закрепленный дизайн, что ограничивает возможность создания более сложных проектов.

Поэтому ниже будут приведены основные преимущества и недостатки QlikView для бизнеса.

Преимущества QlikView:

1. Простота использования.

QlikView имеет интуитивно понятный интерфейс и позволяет пользователям создавать свои собственные отчеты и дашборды без необходимости знать SQL или другие языки программирования.

2. Визуализация данных.

QlikView позволяет визуализировать данные в режиме реального времени, что упрощает анализ и понимание данных.

3. Быстродействие.

QlikView имеет быстродействие и может обрабатывать большие объемы данных, что позволяет пользователю быстро находить ответы на свои вопросы.

4. Гибкость.

QlikView позволяет пользователям создавать свои собственные аналитические приложения и настраивать их под свои потребности.

Недостатки QlikView:

1. Высокая стоимость.

QlikView является достаточно дорогой платформой, что может быть проблемой для небольших компаний с ограниченным бюджетом.

2. Сложность настройки.

QlikView требует определенных знаний и навыков для настройки и использования его функций и возможностей.

3. Ограничения по интеграции.

QlikView имеет ограничения в интеграции с некоторыми сторонними системами, что может ограничивать его применение в некоторых сферах бизнеса.

Power BI — это платформа для бизнес-анализа, разработанная компанией Microsoft. Она предоставляет пользователям возможность создавать отчеты и дашборды на основе больших объемов данных из разных источников, таких как базы данных, Excel-файлы, облачные хранилища данных и другие источники данных. Ниже приведены основные преимущества и недостатки Power BI для бизнеса.

Преимущества Power BI:

1. Интеграция с другими продуктами Microsoft.

Power BI легко интегрируется с другими продуктами Microsoft, такими как Excel, SharePoint и другие, что упрощает работу с данными и повышает производительность.

2. Быстродействие.

Power BI позволяет обрабатывать большие объемы данных и создавать отчеты и дашборды в режиме реального времени, что ускоряет принятие решений.

3. Надежность и безопасность.

Power BI имеет высокий уровень надежности и безопасности, что позволяет сохранять и обрабатывать конфиденциальные данные без риска их потери или утечки.

4. Расширяемость.

Power BI имеет возможность расширения функционала путем добавления дополнительных плагинов и инструментов.

5. Наличие бесплатной версии и невысокая стоимость.

Недостатки Power BI:

1. Ограничения по количеству данных.

Power BI имеет ограничения по количеству данных, которые можно обработать, что может ограничивать его применение в некоторых сферах бизнеса. Для пользователей бесплатной версии Power BI, размер данных, которые можно загрузить в сервис, ограничен до 1 ГБ. Также существуют ограничения на количество запросов в день, количество данных, которые могут быть обработаны в секунду, и другие параметры. Для пользователей, имеющих платную лицензию Power BI Pro, ограничение на размер данных увеличивается до 100 ГБ. Кроме того, пользователи Power BI Premium могут загружать до 400 ГБ данных в сервис.

2. Неустойчивость приложения к большим объемам данных.

Power BI довольно часто зависает при импорте и обработке больших данных.

3. Ограничения по настройке

Power BI требует определенных знаний и навыков для настройки и использования его функций и возможностей.

4. Ограничения по интеграции.

Power BI имеет ограничения в интеграции с некоторыми сторонними системами, что может ограничивать его применение в некоторых сферах бизнеса.

Таблица, представленная ниже, содержит сравнительный анализ рассмотренных зарубежных BI-платформ, которые часто используются в бизнесе. В данном сравнении рассмотрены такие характеристики, как цена, возможности по работе с данными, простота использования, функционал, интеграция с другими системами, импорт данных из различных источников и наличие бесплатной демоверсии. Выбирая BI-платформу, необходимо учитывать свои потребности и возможности, чтобы получить максимальную пользу от ее использования.

Табл. 2. Сравнительная таблица зарубежных BI-платформ

Критерии	Tableau	Qlick	Power BI
Доступная цена	X	X	✓
Возможности по работе с Bigdata	✓	✓	✓/X
Простота использования	X	X	✓
Богатый функционал	✓	✓	✓
Интеграция с другими системами	✓/X	✓/X	✓
Импорт данных из различных источников	✓/X	✓/X	✓
Бесплатная демоверсия	14 дней	30 дней	60 дней

Составлено по: <https://mkt.tableau.com/> (дата обращения 01.04.2023),
<https://www.qlik.com/us/products/qlikview> (дата обращения 01.04.2023),
<https://powerbi.microsoft.com/ru-ru/> (дата обращения 01.04.2023)

1.3. Обзор на российские/отечественные BI-системы

Национальная программа "Цифровая экономика" является одной из ключевых стратегических инициатив России на ближайшие годы и направлена на модернизацию экономики страны с помощью применения новых технологий и цифровых инноваций. Одним из главных аспектов этой программы является проведение цифровой трансформации на базе отечественных решений.

Такой подход к цифровой трансформации обусловлен несколькими факторами. Во-первых, он способствует развитию отечественного IT-сектора и созданию новых рабочих мест, что является важным фактором в период экономических трудностей. Во-вторых, использование отечественных решений позволяет улучшить качество и безопасность технологий и продуктов, что в свою очередь повышает конкурентоспособность российской экономики на мировой арене.

Цифровая трансформация является неотъемлемой частью стратегии национального развития России до 2030 года, определенной в подпункте "д" пункта 1 и подпункте "д" пункта 2 Указа Президента РФ от 21 июля 2020 года № 474. [1]

Большинство российских BI-платформ предлагают широкий набор функциональных возможностей, включающих в себя аналитику данных, визуализацию, создание отчетов и дашбордов, а также инструменты машинного обучения.

Далее будет рассмотрено несколько российских BI-платформ, которые имеют хорошую репутацию на рынке и предлагают передовые технологии для анализа и визуализации данных.

Visiology — это российское BI-приложение, которое используется для анализа данных и создания дашбордов. Visiology является победителем конкурса «Лучшие цифровые решения» от Аналитического центра при Правительстве Российской Федерации.

Одним из основных преимуществ Visiology является простота использования и интуитивно-понятный интерфейс, что позволяет даже неопытным пользователям легко создавать и настраивать дашборды и отчеты. В приложении реализованы различные типы графиков и диаграмм, что позволяет визуализировать данные в разных форматах и уровнях детализации.

Также Visiology предлагает возможности для интеграции с другими BI-системами и базами данных, что упрощает процесс обмена и обработки данных. Важным преимуществом является возможность использования приложения в локальной сети без подключения к Интернету, что повышает безопасность и защиту данных.

К недостаткам можно отнести относительно высокую стоимость лицензии. Visiology стоит дороже, чем некоторые другие российские BI-платформы. К недостаткам также можно отнести ограниченную возможность кастомизации, приложение не позволяет пользователям настроить все функциональные возможности и внешний вид элементов отчета так, как это может быть доступно в других BI-платформах. Доступ к демоверсии приложения ограничен сутками, введенные данные будут утеряны.

Отсутствие большого сообщества пользователей и разработчиков: Visiology относительно недавно появилась на рынке и не имеет такого большого сообщества пользователей и разработчиков. Это может означать меньше общественных обсуждений и менее активную поддержку разработчиков.

Аналитическая платформа Форсайт — это российская BI-система, предназначенная для анализа и визуализации данных в различных сферах деятельности. Она разрабатывается компанией "Тензор" с 2004 года.

Рассмотрим преимущества программного продукта. Форсайт может извлекать данные из различных источников, включая реляционные базы данных, файлы Excel, текстовые файлы, OLAP-серверы и другие. Форсайт может масштабироваться как по объему данных, так и по числу пользователей, что позволяет использовать его в крупных корпоративных системах. Программа предоставляет возможность работать с файлами через настольную веб и мобильную версии. Для ознакомления с рабочей средой доступна бесплатная демоверсия на 30 дней. Форсайт обеспечивает согласованность данных в организации благодаря созданию единого хранилища данных и систем нормативно-справочной информации (НСИ).

К недостаткам платформы можно отнести то, что Форсайт не всегда может быть легко интегрирован с другими программными продуктами, что может вызывать проблемы при обмене данными и автоматизации бизнес-процессов. Несмотря на то, что Форсайт является достаточно популярной BI-платформой в России, сообщество ее пользователей не так широко, как у мировых аналогов, что может сказаться на доступности дополнительной информации и ресурсов.

Visary BI — это российская бизнес-аналитическая платформа, предназначенная для обработки больших объемов данных и создания отчетов и дашбордов. Платформа имеет интуитивно понятный интерфейс и множество функциональных возможностей, включая интеграцию с различными источниками данных, автоматическую генерацию отчетов и дашбордов, возможность проведения множественных анализов и создания прогнозов.

Одним из главных преимуществ Visary BI является высокая скорость работы благодаря оптимизированным запросам к базам данных и эффективному кэшированию. Также платформа обладает удобным интерфейсом для создания и настройки отчетов, в том числе с использованием сложных формул и алгоритмов. Visary BI позволяет оперативное формирование аналитических отчетов посредством интерактивных панелей визуального представления данных. Платформа позволяет поддерживать неограниченное число разнородных источников данных (ETL). Visary BI является кроссплатформенной программой, доступна мобильная версия. С целью улучшения эффективности рутинных задач и обеспечения доступа к актуальным результатам прогнозирования, в платформу внедрены различные механизмы искусственного интеллекта, такие как предиктивная аналитика, машинное обучение, Data Science и Data Mining.

Среди недостатков Visary BI можно отметить отсутствие многих продвинутых функций, которые есть в некоторых других BI-системах. Visary BI может быть сложно интегрировать с другими BI-системами или существующими приложениями, что может вызывать сложности в обмене данными между системами. Платформа предоставляет базовый набор функций и возможностей, но не всегда может удовлетворять требованиям пользователей. Некоторые пользователи отмечают ограничения в возможностях настройки пользовательского интерфейса Visary BI, что может снижать удобство использования системы. Visary BI достаточно дорогой, особенно для небольших и средних компаний, что может стать препятствием для их внедрения.

Yandex DataLens — это облачный сервис визуализации данных, который был разработан компанией Яндекс для бизнес-аналитики и инсайтов. Сервис предоставляет возможность создавать дашборды, отчеты и графики, а также производить анализ данных с помощью встроенных алгоритмов и возможностей машинного обучения.

Рассмотрим преимущества программного продукта. Yandex DataLens имеет интуитивно понятный интерфейс, который позволяет быстро и удобно создавать, и настраивать отчеты. Сервис также поддерживает несколько источников данных, включая Google Analytics, Яндекс.Метрику, PostgreSQL, MySQL и другие. Платформа использует AI-алгоритмы для автоматической обработки данных и построения предсказательных моделей.

Недостатком Yandex DataLens является то, что платформа не подходит для решения сложных аналитических задач, которые требуют более продвинутых алгоритмов и

специализированных инструментов. Yandex DataLens имеет ограниченный функционал для создания кастомных отчетов и дашбордов и ограничен интеграцией с другими BI платформами.

Таблица (табл. 2.), представленная ниже, содержит сравнительный анализ наиболее популярных отечественных BI-платформ. В данном сравнении рассмотрены такие характеристики, как цена, возможности по работе с данными, простота использования, функционал, интеграция с другими системами, импорт данных из различных источников и наличие бесплатной демоверсии.

Табл. 2. Сравнительная таблица российских BI-платформ

Критерии	Visiology	Форсайт	Visary BI	Yandex DataLens
Доступная цена	✗	✗	✗	бесплатно
Возможности работы с Bigdata	✓	✓	✓	✓
Простота использования	✓	✓	✓	✓
Богатый функционал	✓/✗	✓	✓	✗
Интеграция с другими системами	✓	✓/✗	✓	✓
Импорт данных из различных источников	✓	✓	✓	✓
Бесплатная демоверсия	1 сутки	30 дней	✗	✓

Составлено по: <https://datalens.yandex.ru/> (дата обращения 01.04.2023) , <https://ru.visiology.su/> (дата обращения 01.04.2023), <https://npc.ba/development/bi> (дата обращения 01.04.2023), <https://npc.ba/development/bi> (дата обращения 01.04.2023) , <https://www.fsight.ru/platform/> (дата обращения 01.04.2023)

1.4. Обоснование выбора BI-платформы

В условиях быстрого развития технологий и увеличения объемов данных, которые необходимо анализировать, BI-системы становятся неотъемлемой частью бизнес-процессов компаний [13].

Однако, выбор подходящей BI-системы может стать сложной задачей, требующей серьезного анализа и обоснования. На сегодняшний день на рынке представлено множество BI-систем, открытых и закрытых, с различным функционалом и стоимостью. В данной части будет рассмотрен выбор BI-системы Power BI и обоснованы причины, почему именно она была выбрана.

Среди причин выбора можно выделить следующие:

1. Цена

Power BI имеет довольно низкую цену по сравнению с аналогами на рынке BI-систем. Она доступна как отдельная BI-система, так и в составе пакета Microsoft Office 365.

2. Интеграция

Power BI легко интегрируется с другими продуктами Microsoft. Это облегчает работу с данными и делает использование Power BI более удобным и эффективным.

3. Визуализация данных

Power BI обладает мощными инструментами визуализации данных, позволяющими отображать данные в различных форматах и виде. Это делает работу с данными более понятной и интуитивно понятной.

4. Обработка больших объемов данных

Power BI позволяет обрабатывать большие объемы данных, что делает ее идеальной для использования в крупных компаниях.

5. Автоматизация

Power BI обладает мощными инструментами для автоматизации бизнес-процессов. Существует возможность настроить автоматическую загрузку данных из различных источников, обновление отчетов и дашбордов в реальном времени, а также отправку уведомлений в случае нарушения определенных пороговых значений. Power BI также поддерживает автоматический процесс создания отчетов и дашбордов на основе заданных шаблонов, что позволяет значительно сократить время на подготовку отчетности.

6. Бесплатная демоверсия

Бесплатная версия называется Power BI Desktop. С ее помощью пользователи могут создавать отчеты, дашборды и модели данных на своем компьютере. Однако для публикации этих отчетов и дашбордов в облаке Microsoft и доступа к ним из любого устройства необходимо приобрести подписку на один из тарифных планов Power BI Pro.

Бесплатная версия Power BI Desktop имеет множество функций, которые позволяют создавать профессиональные отчеты и дашборды. Некоторые функции могут быть ограничены, такие как доступность определенных типов источников данных, объем данных или количество пользователей, которые могут использовать один отчет или дашборд. Также отсутствует возможность обновления данных в автоматическом режиме, что доступно только в тарифных планах Power BI Pro и Premium.

Тем не менее, Power BI Desktop является мощным и бесплатным инструментом для создания отчетов и дашбордов, который может быть использован как для личных, так и для коммерческих целей.

Глава 2. Проектирование BI-приложения для анализа движения косметических товаров.

2.1. Формирование задач проектирования и целей разработки BI-приложения

Разработка BI-приложения для ООО «Дрогери-ритейл» является практической частью курсовой работы. Это предполагает практическую реализацию проекта на основе теоретических знаний, полученных в процессе обучения.

Компания ООО "Дрогери-ритейл" успешно работает под брендом "Улыбка радуги".

"Улыбка радуги" — это крупная федеральная сеть магазинов, которая специализируется на продаже товаров для ухода за собой и за домом. Компания предлагает широкий ассортимент товаров, включая хозяйственные товары, косметику, парфюмерию, товары для дома, сада и огорода, а также продукты питания. Компания предлагает товары от ведущих производителей, а также собственные бренды, которые отличаются высоким качеством и доступными ценами.

Одной из основных задач деятельности сети косметических магазинов, является контроль за движением товаров со склада в филиалы сети (магазины/торговые точки). Представление данных в визуально удобной форме позволяет выявлять успешные продукты и территории, управлять запасами и заказами, а также выявить проблемные зоны в работе склада или компании в целом, принять меры для устранения причин неэффективности.

Использование дашбордов на складе косметической сети магазинов позволяет оптимизировать следующие направления:

1. Управление запасами: дашборды могут помочь визуализировать уровень запасов на складе, включая количество и типы товаров, а также оценку оставшегося времени до истечения срока годности. Это позволяет менеджерам склада управлять запасами более эффективно, избегать избыточных запасов и нехватки товаров.
2. Оптимизация процессов: дашборды могут показать ключевые показатели производительности склада, такие как время обработки заказов, скорость отгрузки товаров, количество ошибок при комплектации заказов и другие. Эти данные помогают менеджерам оптимизировать процессы на складе, сокращать время выполнения заказов и улучшать качество обслуживания.
3. Мониторинг эффективности: дашборды могут отображать общую эффективность склада и его отдельных компонентов, таких как упаковка, погрузка и т.д. Это помогает менеджерам склада мониторить производительность и идентифицировать проблемные области, которые нуждаются в улучшении.
4. Прогнозирование спроса: дашборды могут отображать данные о продажах и спросе на товары в магазинах, что помогает менеджерам склада прогнозировать будущий спрос и планировать поставки товаров. Это позволяет избежать нехватки товаров и своевременно удовлетворять потребности клиентов.

В рамках процесса разработки необходимо решить следующие задачи:

1. Анализ предметной области в сфере логистики поставок косметических товаров и средств для быта компании ООО «Дрогери-ритейл».
2. Анализ информационных потребностей пользователей и формирование требований к информационной системе.
3. Комплексный анализ доступных источников данных, определение возможностей и ограничений по их использованию, а также выбор оптимальных методов обработки и реализации данных для BI-приложения.
4. Проектирование информационной системы в Power BI.
5. Анализ экономической эффективности использования разработанного приложения.

2.2 Анализ информационных потребностей и требований пользователей

Пользователями приложения являются руководители и специалисты отдела складской логистики. В ходе анализа были определены следующие информационные потребности пользователей:

1. Возможность анализировать информацию о движении товаров со склада в магазины в разрезе районов Санкт-Петербурга;
2. Возможность проводить анализ выполнения рейсов доставки заказов в магазины по продолжительности, месяцам и времени суток;
3. Возможность сравнивать показатели по районам Санкт-Петербурга и номерам магазинов;
4. Анализ структуры заказываемых магазинами товаров по категориям;
5. Возможность выявить магазины с максимальным и минимальным числом заказываемых товаров, а также самые заказываемые товары;

В ходе анализа информационных потребностей пользователей и специфики выбранной предметной области выявлены следующие требования пользователей к BI-приложению:

1. Классифицировать и представить данные по ключевым показателям выполнения рейсов;
2. Классифицировать и представить данные по ключевым показателям движения и структуры товаров;
3. Разработать комфортный интерфейс для пользователя, состоящий из визуализации информации в различных видах;

2.3. Предварительная подготовка данных

Курсовая работа по внедрению BI приложения предполагает не только разработку программного продукта, но и проверку качества данных, которые будут использоваться в этом приложении. Как и в случае с моделями машинного обучения, качество данных является важным фактором, влияющим на точность и полноту информации, получаемой в BI системе. Поэтому, перед разработкой BI приложения необходимо проверить данные на наличие потенциальных проблем, таких как дублирование, пропущенные значения, ошибки формата и

т.д. При подготовке к исследованию был проведен анализ источников данных и выполнены работы по подготовке данных. Качество учетных данных является важным фактором для эффективной информационной поддержки управления региональными продажами. Чтобы данные были точными, чистыми, полными и своевременными, необходимо объединить их в единую систему. В компании ООО «Дрогери-ритейл» для построения BI-приложений используется среда Power BI.

Описание данных, полученных из предоставленных источников представлено в табл. 2.3.1.

Таблица 2.3.1.

Данные	Разрешение
Данные по остаткам на мезонине	xlsx
Данные об отгрузке заказов	xlsx
Данные об информации о заказе	xlsx
Данные о коробках с группировкой по рейсам	xlsx
Данные о составе коробок для отгрузки	xlsx
Данные о формировании перемещении товаров из ячеек на мезонине в коробки для отгрузки	xlsx
Данные о времени выполнения рейса	xlsx
Данные о расположении магазинов	xlsx

Для проверки качества данных в данной работе будет использоваться язык программирования Python.

При анализе данных были обнаружены следующие проблемы:

- Дублирование

В ходе анализа данных были тщательно проверены все таблицы на наличие дубликатов. Дубликаты данных могут существенно исказить результаты анализа и привести к неверным выводам. Они могут быть вызваны ошибками ввода данных, техническими сбоями или другими причинами. В общем и целом, чем меньше дубликатов в данных, тем лучше.

В таблице «Остатки на мезонине» было выявление дублирование данных. Ввиду большого объема исходных данных, было принято решение исключить дубликаты. Для удостоверения, в том, что таблица больше не содержит дубликатов, была проведена повторная проверка (рис. 2.3.1)

```
import pandas as pd

df = pd.read_excel('для 3.xlsx')

# Находим дубликаты и сохраняем их в новый DataFrame
duplicates = df[df.duplicated()]

# Выводим дубликаты на экран
print(duplicates)
```

	Код товара	Наименование товара	Кол-во штук	\
5661	3070912	S`cosmetic д/укл.лак Укл.и восст.300мл	12	
6587	3081118	INNAMORE колг.жен.40D BellaDaino p4	1	
8712	7507942	Bell тест д/лица BB cream 04	50	

	Штучная упаковка	Ширина шт	Глубина шт	Высота шт	Вес шт	\
5661	Шт-1ед-5.5x5.5x24	5.5	5.5	24.0	0.304	
6587	Шт-1ед-21.5x16x1.2	21.5	16.0	1.2	0.058	
8712	Шт-1ед-13x4x2.5	13.0	4.0	2.5	0.040	

	Коробочная упаковка	Ширина кор	Глубина кор	Высота кор	Вес кор	\
5661	Кор-12ед-22.5x17x26	22.5	17.0	26.0	0.10	
6587	Кор-100ед-51x36x24	51.0	36.0	24.0	0.43	
8712	Кор-270ед-43x27x36	43.0	27.0	36.0	0.52	

	Вес брутто	Штук в коробке	Ячейка на мезонине
5661	3.748	12.0	322-022-2
6587	6.230	100.0	409-024-3
8712	11.320	270.0	502-078-3


```
duplicates = df.duplicated()

if duplicates.sum() == 0:
    print('Дубликатов не обнаружено')
else:
    print('Найдены дубликаты:')
    print(df[duplicates])
```

Дубликатов не обнаружено

Рис. 2.3.1. Проверка на дубликаты

- Выбросы

Возник интерес, посмотреть какие товары преобладают на мезонине, а каких совсем немного. Для этого была проведена оценка распределения значений столбца «Код товара» и вычислена частота появления. Были обнаружили некоторые выбросы в данных, но было решено не удалять их, а оставить для ознакомления. Это может быть полезно для более глубокого понимания данных и возможного поиска причин, почему такие выбросы возникли. Однако, при использовании этих данных в будущих исследованиях или моделях, нужно учитывать эти выбросы и их влияние на результаты.

```
import pandas as pd
import matplotlib.pyplot as plt

%matplotlib inline

# чтение таблицы из файла Excel
df = pd.read_excel('для 3.xlsx')

# оценка распределения значений столбца "код товара" и их частоты
value_counts = df['Код товара'].value_counts()

# запись результатов в новый файл Excel
value_counts.to_excel('выбросы.xlsx', header=False)
```



```
import pandas as pd
import matplotlib.pyplot as plt

# создаем Data Frame
df = pd.read_excel('выбросы.xlsx')

# построение точечного графика
df.plot.scatter(x='Код товара', y='Количество')

# отображение графика
plt.show()
```

Рис. 2.3.2 Код проверки данных на выбросы

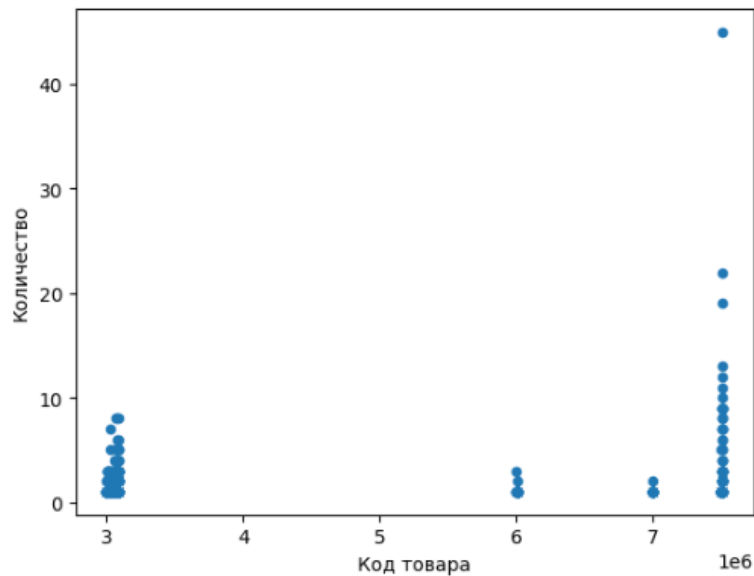


Рис. 2.3.3 График выбросов

В процессе изучения данных был проанализирован график распределения товаров на мезонине. Из полученных результатов было выявлено, что основное количество товаров на мезонине колеблется в пределах 10-15 единиц. Однако, был обнаружен товар «посылка интернет магазина маг.N580», который представлен на мезонине 45 раз. Несмотря на то, что такое значение может быть воспринято как выброс, в данной предметной области это не является аномалией и не будет подвержено удалению. Важно отметить, что проведение анализа на наличие выбросов позволяет выявлять не только ошибки в данных, но и важные особенности и закономерности в исследуемой предметной области.

- Пропущенные значения

Проведение анализа данных на пропущенные значения является критически важной задачей в области анализа данных. Это связано с тем, что отсутствие данных может привести к значительным искажениям в результатах анализа, а также уменьшить точность и надежность получаемых выводов. Пропущенные данные могут возникать из-за различных причин, таких как ошибки ввода данных, технические сбои или отсутствие информации в исходных источниках данных. Проверке на пропущенные значения были подвержены все таблицы. Пропущенные значения обнаружены только в таблице «Остатки на мезонине».

Ниже представлен код, по выявлению и последующему удалению пропущенных значений из таблицы «Остатки на мезонине»:

```
import pandas as pd

# чтение таблицы из файла Excel
df = pd.read_excel('для 3.xlsx')

# проверка на пропущенные значения
missing_values = df.isnull().sum()

if missing_values.sum() == 0:
    print('Пропущенных значений нет')
else:
    print('Найдены пропущенные значения:')
    print(missing_values)
```

```
Найдены пропущенные значения:
Код товара          0
Наименование товара 0
Кол-во штук         0
Штучная упаковка  0
Ширина шт           0
Глубина шт          0
Высота шт           0
Вес шт              0
Коробочная упаковка 237
Ширина кор          237
Глубина кор         237
Высота кор          237
Вес кор             237
Вес брутто          237
Штук в коробке      237
Ячейка на мезонине  0
dtype: int64
```

```
# проверка на пропущенные значения
missing_values = df.isnull().sum()

if missing_values.sum() == 0:
    print('Пропущенных значений нет')
else:
    print('Найдены пропущенные значения:')
    print(missing_values)
```

```
Пропущенных значений нет
```

Рис. 2.3.3 Код проверки данных на пропущенные значения

- Несоответствие данных формату

Анализ на несоответствие формату данных была подвержена таблица «Рейсы. Время рейса». Таблица содержит в себе 2 столбца с датами, которые до коррекции имели 2 разных формата даты. Формат даты может существенно влиять на корректность результатов анализа данных. Если в наборе данных присутствуют даты, записанные в различных форматах, то алгоритмы анализа могут некорректно интерпретировать эти данные, что может привести к ошибочным выводам. Поэтому важно убедиться, что все даты в наборе данных записаны в едином формате.

Ниже представлен код, приводящий все даты в таблице к единому формату.

```
import pandas as pd

# чтение таблицы из файла Excel
df = pd.read_excel('рейсы.xlsx', sheet_name=4)

# изменение формата ячеек столбца "Время начала"
df['Время начала'] = pd.to_datetime(df['Время начала'], format='%d.%m.%Y %H:%M')

# сохранение изменений в файл
df.to_excel('рейсы.xlsx', index=False)
```

Рис. 2.3.4 Код для приведения дат к единому формату

Цель подготовки и проверки качества данных заключается в том, чтобы гарантировать, что данные соответствуют требуемым стандартам и являются точными, полными и надежными. Этот этап является важным, так как на него напрямую влияет качество дальнейших аналитических выводов, которые могут быть сделаны на основе этих данных.

2.4. Построение модели данных

Следующим этапом разработки приложения является построение модели данных. Power BI использует реляционную модель данных. Реляционная модель данных означает, что данные организованы в таблицы, которые связаны между собой ключами: один-к-одному, один-ко-многим, многие-ко-многим. Каждая таблица представляет отдельный объект или сущность, а столбцы в таблицах представляют атрибуты этих объектов. Power BI использует язык запросов DAX (Data Analysis Expressions), который позволяет проводить мощный анализ данных в соответствии с реляционной моделью.

Когда данные, необходимые для работы системы, загружены, начинается процесс преобразования этих данных, который включает их очистку и подготовку для создания требуемой модели.

Для начала, нужно выбрать схему модели данных. Основными схемами в Power BI являются:

- Звезда (Star schema)
- Снежинка (Snowflake schema)
- Плоская (Flat schema)

Рассмотрим каждую из схем подробнее:

- Звезда

Эта модель является наиболее распространенной и простой для понимания. Она состоит из одной основной таблицы фактов, которая содержит ключи для связи с несколькими таблицами измерений [14]. Таблицы фактов содержат данные о фактах, таких как продажи или заказы, а таблицы измерений содержат дополнительные данные о сущностях, таких как клиенты, продукты и т.д. Пример схемы типа «Звезда» представлен на рис. 2.4.1

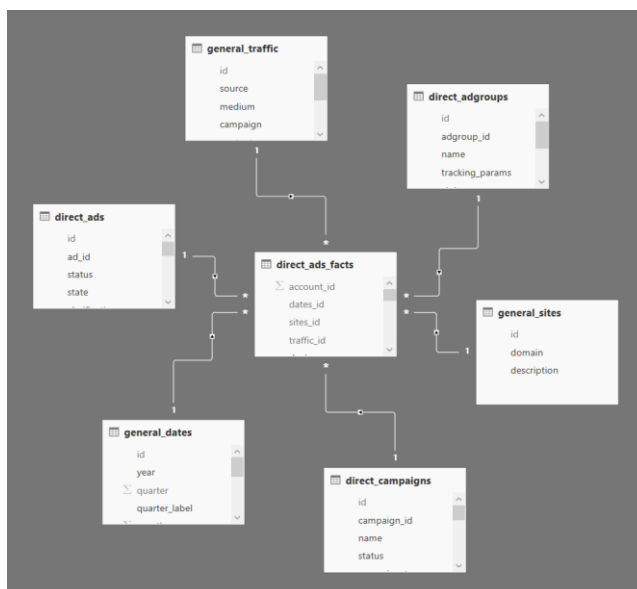


Рис 2.4.1. Пример схемы «Звезда»

Источник: URL: https://powerbirussia.ru/2018/03/27/power_bi_data_modeling/ (Дата обращения: 12.04.2023)

- Снежинка

Эта модель является расширением модели "Звезда" и используется в случае, когда таблицы измерений имеют дополнительные уровни связанных данных. Эта модель использует множество таблиц измерений, каждая из которых может содержать данные в виде таблиц, связанных с другими таблицами измерений. Пример схемы типа «Снежинка» представлен на рис. 2.4.2.

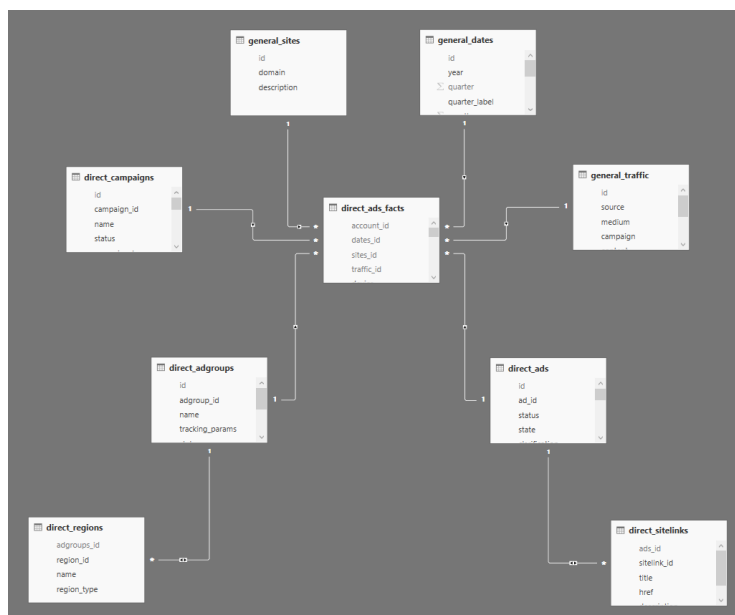


Рис 2.4.2. Пример схемы «Снежинка»

Источник: URL: https://powerbirussia.ru/2018/03/27/power_bi_data_modeling/ (Дата обращения: 12.04.2023)

- Плоская

Эта модель используется для простых приложений, когда не требуется большое количество таблиц. Она состоит из одной таблицы, которая содержит все данные о фактах и измерениях.

Для проектируемого приложения, на основе анализа данных, выбрана схема «Снежинка».

Далее, перейдём к загрузке данных в Power BI. Все данные предоставлены в формате книг Excel, поэтому для загрузки не требуется использование дополнительных коннектеров.

В Power BI были загружены следующие данные:

- Таблица «Остатки на мезонине».

Данная таблица нужна для предоставления информации об остатках на складе каждой из позиций продукции с указанием габаритов и веса индивидуальной упаковки и целой коробки, количестве штук в коробке. Для каждого товара указывается его размещение, то есть номер ячейки на мезонине. Таблица «Остатки на мезонине» представлена в табл. 2.4.1.

Поля	Описание
Код товара	Уникальный идентификатор товара
Наименование товара	
Количество штук	Количество штук на мезонине
Штучная упаковка	Габариты индивидуальной упаковки товара
Ширина	
Высота	
Вес	
Коробочная упаковка	Габариты коробки пачки товаров
Ширина коробки	
Глубина коробки	
Высота коробки	
Вес коробки	Вес пустой коробки
Вес брутто	Вес товара+ вес коробки
Количество штук в коробке	
Номер ячейки на мезонине	Место расположения товаров на мезонине

Таблица 2.4.1.

- Таблица «Отгрузки»

Данная таблица нужна для предоставления информации о том, как формируется заказ, во сколько этапов его отправляют, какие позиции содержатся в заказе, а именно код товара и его наименование. Таблица «Отгрузки» представлена в табл. 2.4.2.

Поля	Описание
Заказ	Уникальный код заказа
Отправка	Код отправления
Код товара	
Наименование товара	

Таблица 2.4.2.

- Таблица «Рейсы. Заказы»

Данная таблица нужна для предоставления информации о номере заказа, в какой магазин будет отправлен заказ, перечень товаров входящих в заказ и их количество. Один заказ = один магазин. В одном заказе несколько позиций товаров. Таблица «Рейсы. Заказы» представлена в табл. 2.4.3.

Поля	Описание
Заказ	Уникальный код заказа
Номер магазина	
Код товара	
Количество товара	

Таблица 2.4.3.

- Таблица «Рейсы. Коробки»

Данная таблица нужна для предоставления информации о том какая позиция товара находится в какой коробке, для какого заказа. Ниже представлена структура таблицы «Рейсы. Коробки» в табл. 2.4.4.

Поля	Описание
Код товара	
Номер коробки	
Заказ	

Таблица 2.4.4.

- Таблица «Рейсы. Рейсы-коробки»

Данная таблица нужна для предоставления информации о том, какие коробки едут одним рейсом. Один рейс может содержать в себе несколько коробок. Структура таблицы «Рейсы. Рейсы-коробки» представлена в табл. 2.4.5.

Поля	Описание
Номер рейса	
Номер коробки	

Таблица 2.4.5.

- Таблица «Рейсы. Ячейки»

Данная таблица нужна для предоставления информации о том, с какой ячейки на мезонине формируются товары для выполнения рейса. Структура таблицы «Рейсы. Ячейки» представлена в табл. 2.4.6.

Поля	Описание
Номер рейса	
Номер ячейки	

Таблица 2.4.6.

- Таблица «Рейсы. Время рейса»

Данная таблица нужна для предоставления информации о рейсе, во сколько машина вышла на рейс и когда его завершила. Структура таблицы «Рейсы. Время рейса» представлена в табл. 2.4.7.

Поля	Описание
Номер рейса	
Время начала рейса	
Время окончания рейса	
Район	

Таблица 2.4.7.

- Таблица «Рейсы. Магазины»

Данная таблица нужна для предоставления информации о рейсе, во сколько машина вышла на рейс и когда его завершила. Структура таблицы «Рейсы. Магазины» представлена в табл. 2.4.8.

Поля	Описание
Код магазина	
Район	

Таблица 2.4.8.

Для создания модели данных используются таблицы, которые были ранее загружены. В процессе построения данные из разных таблиц связываются между собой на основе наличия одинаковых названий полей. Полученная модель представлена на рисунке 2.4.3.

Так как данные были получены не из строго структурированной базы данных, а предоставлены в виде файлов excel, то при формировании модели возникла проблема с не полным дублированием данных, а именно было обнаружено дублирование кодов. Несмотря на то, что был проведен анализ качества данных средствами python, так как проверка проводилась на полное дублирование.

Ввиду достаточного большого объема имеющихся данных и незначительного числа дубликатов, решением неполного дублирования стало создание дубликатов проблемных таблиц, проведена группировка данных, исключение повторяющихся кодов и объединение с таблиц-дубляжей с исходными таблицами с помощью инструмента ETL, Microsoft Power Query.

Теперь, перейдем ко второму этапу разработки ВІ приложения, который связан с созданием пользовательского интерфейса и отчетов. Эти шаги будут представлены в третьей главе данной работы.

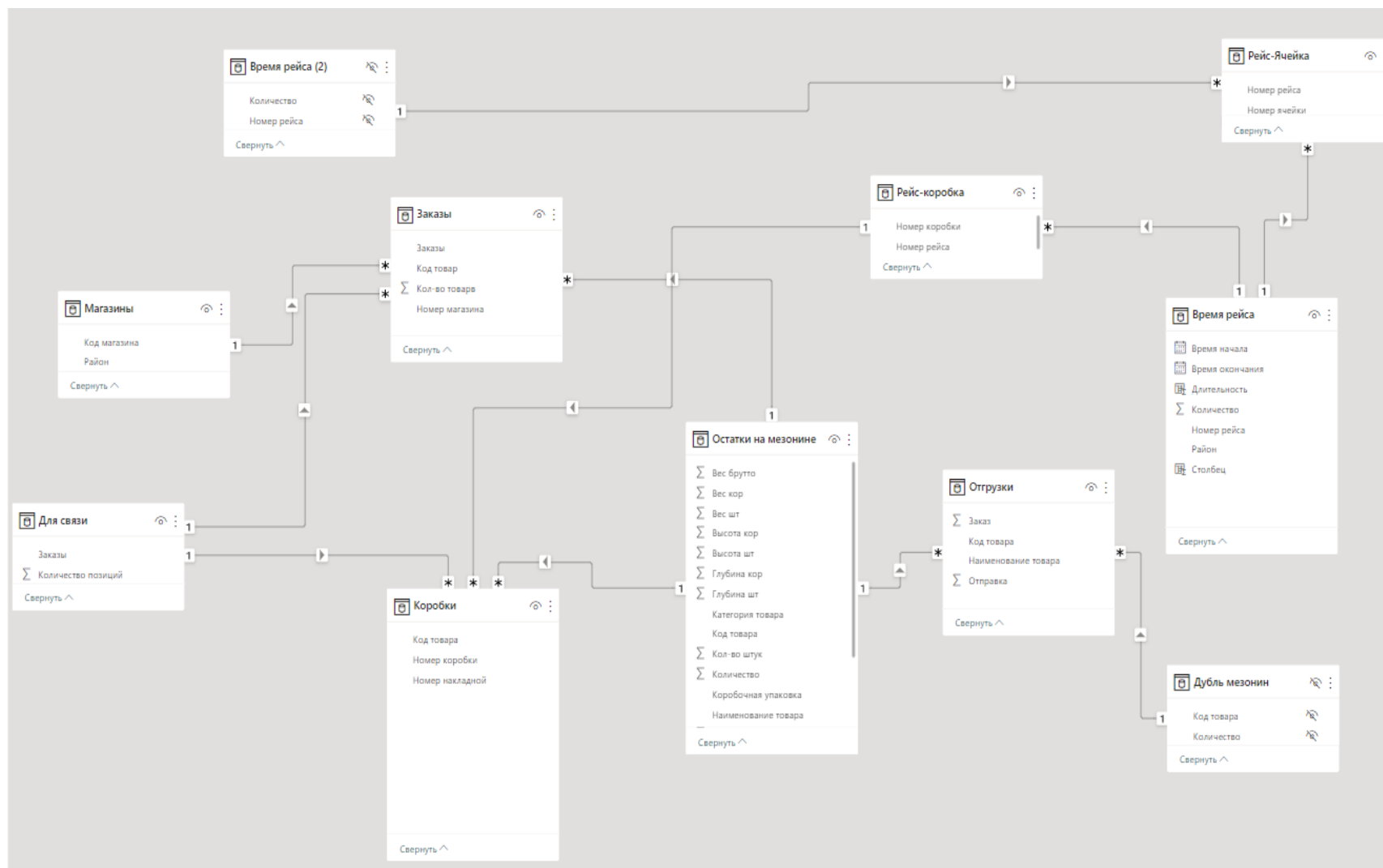


Рис 2.4.3. Модель данных

Глава 3. Разработка BI-приложения.

3.1. Реализация BI-приложения в Power BI

Для пользователя создается интерфейс, включающий следующие компоненты: листы, фильтры, ключевые объемные показатели, объекты в виде диаграмм, карт и списков, формулы, кнопки и макросы. Также производится настройка работы фильтров и разработка темы в корпоративном стиле.

Приложение состоит из одного листа и включает в себя два режима: «Рейсы», «Магазины». В режиме «Рейсы» содержится вывод информации по продолжительности рейсов, их распределению по времени суток и месяцам. С помощью инструмента «Кнопка», организовано объединение между режимами. Перемещение между режимами осуществляется с помощью клавиши ctrl. Кнопки расположены в правом верхнем углу листа с целью комфортной навигации (рис.).



Рис. 3.1.1 Кнопки переключения между режимами

С помощью фильтров «Дата», «Время суток» появляется возможность сравнения показателей по датам, времени выполнения рейсов. Данные фильтры построены на основе полей из таблицы «Рейсы. Время рейса». Срез реализованы по типу «Между», но можно посмотреть только одну дату или конкретное время (рис.).

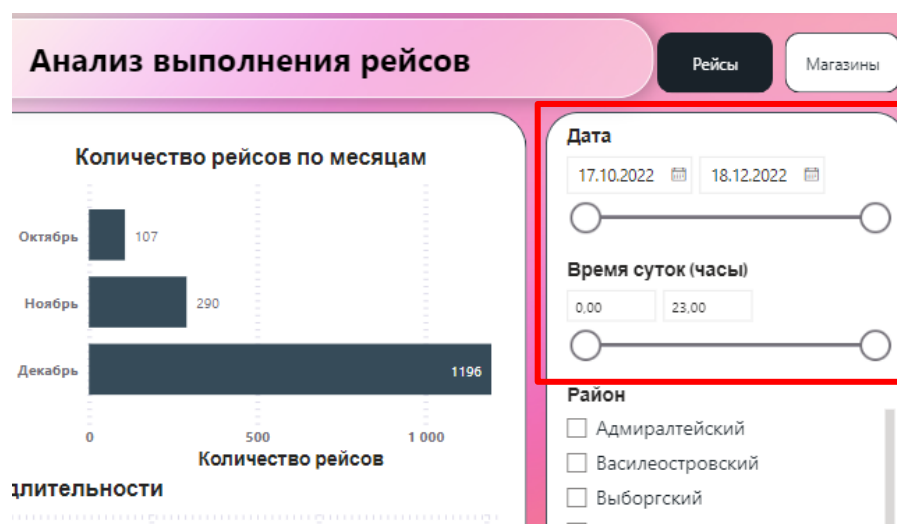


Рис. 3.1.2 Срезы «Дата» и «Время суток (часы)»

В реализацию добавлена возможность отследить изменение динамики рассматриваемых показателей по районам Санкт-Петербурга. Можно выбрать как один, так и несколько районов. Данный фильтр построен на основе полей из таблицы «Рейсы. Время рейса». Срез реализован по типу «Список» (рис.).



Рис. 3.1.3 Срез «Район»

Далее перейдем к описанию гистограммы с группировкой «Распределение рейсов по времени». Данная гистограмма сформирована на основе данных таблицы «Время рейса». Стоит отметить, что на гистограмме не отображен период с 7 до 18 часов, так как в это время отгрузки не производятся.



Рис. 3.1.4 Гистограмма «Распределение рейсов по времени»

Линейчатая диаграмма с накоплением под названием «Количество рейсов по месяцам» отображает насколько часто выполнялись рейсы в определенный месяц.

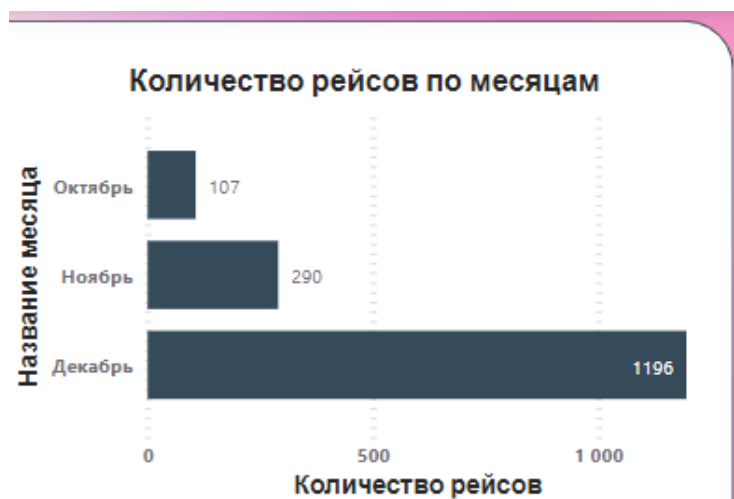


Рис. 3.1.5 Линейчатая диаграмма с накоплением «Количество рейсов по месяцам»

Следующими графическими элементами данной страницы являются график «Количество рейсов по длительности» и карточка «Средняя длительность». Данные элементы позволяют отследить среднее время рейса и количество рейсов, которые выполняются определенное время, а также какое самое распространенное и самое нераспространенное время выполнения рейса.



Рис. 3.1.6 График «Количество рейсов по длительности» и карточка «Средняя длительность»

Далее следует режим «Магазины». В данном режиме представлены количественные данные о заказах товаров по районам и магазинам. При помощи фильтров «Район» и «Номер магазина» можно получить информацию об интересующем нас районе и/или магазине для проведения дальнейшего анализа.

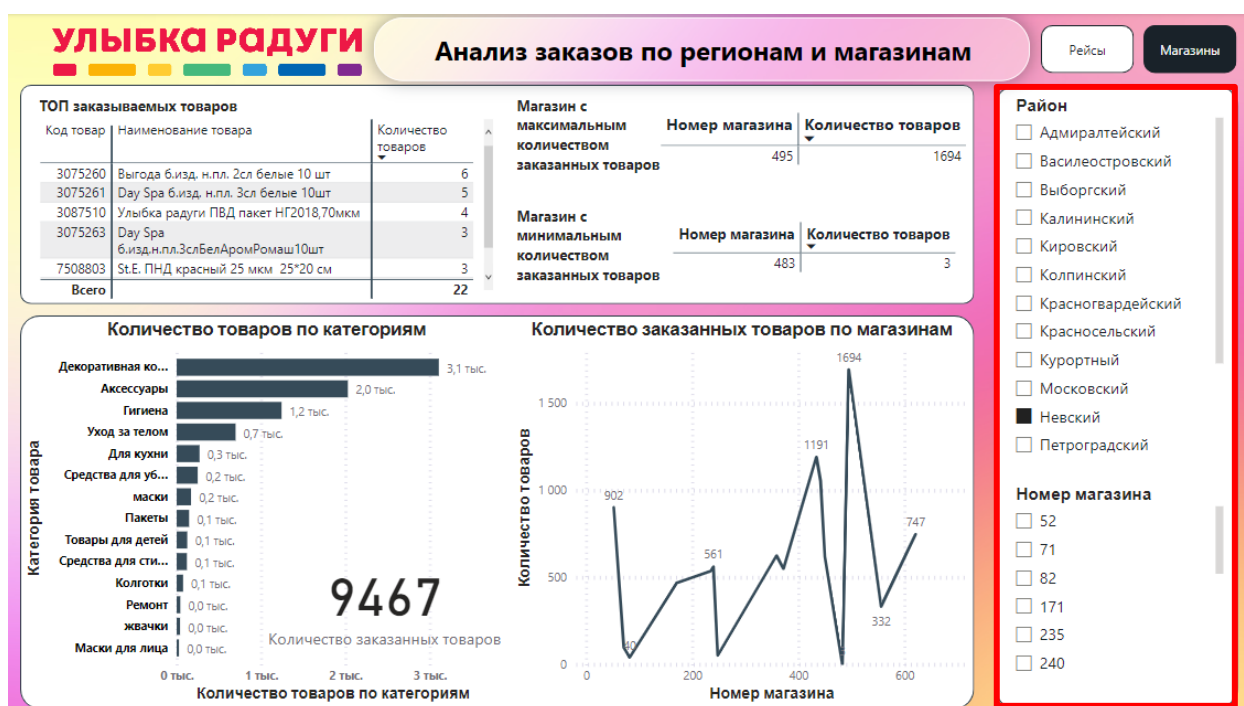


Рис. 3.1.7 Срезы «Район» и «Номер магазина»

В верхней части страницы располагаются таблицы «Топ заказываемых товаров», «Магазин с максимальным количеством заказанных товаров» и «Магазин с минимальным количеством заказанных товаров». Эти элементы позволяют отследить и выявить самые заказываемые и менее популярные товары, а также магазины с наибольшим и наименьшим числом заказов.

ТОП заказываемых товаров			Магазин с максимальным количеством заказанных товаров	
Код товар	Наименование товара	Количество товаров	Номер магазина	Количество товаров
3075260	Выгода б.изд. н.пл. 2сл белые 10 шт	6	495	1694
3075261	Day Spa б.изд. н.пл. 3сл белые 10шт	5		
3087510	Улыбка радуги ПВД пакет НГ2018,70мкм	4		
3075263	Day Spa б.изд.н.пл.3слБелАромРомаш10шт	3		
7508803	St.E. ПНД красный 25 мкм 25*20 см	3		
Всего		22		

Магазин с минимальным количеством заказанных товаров		
Номер магазина	Количество товаров	
483	3	

Рис. 3.1.8 Таблицы «Топ заказываемых товаров», «Магазин с максимальным количеством заказанных товаров» и «Магазин с минимальным количеством заказанных товаров»

Следующие элементы, расположенные ниже таблиц – линейчатая диаграмма с группировкой «Количество товаров по категориям» и карточка «Количество заказанных товаров».



Рис. 3.1.9 Линейчатая диаграмма с группировкой «Количество товаров по категориям» и карточка «Количество заказанных товаров»

Расположенный рядом график «Количество заказанных товаров по магазинам» позволяет сравнить количество заказанных товаров по магазинам.



Рис. 3.1.10 График «Количество заказанных товаров по магазинам»

Таким образом, было реализовано приложение в BI-системе Power BI с описанным выше интерфейсом.

3.2. Анализ экономической эффективности

При оценке экономической эффективности внедрения аналитической системы невозможно использовать стандартные методы оценки инвестиционных проектов. Это характеризуется тем, что отсутствует возможность оценить финансово доход от внедрения приложения, так как оно имеет стратегический характер и имеет задачу облегчить принятие управленческих решений благодаря предоставленному аналитическому инструментарию. Следовательно, нужно разработать метод оценки экономической эффективности, который будет учитывать как финансовые, так и показатели эффективности внедрения приложения, которые невозможно оценить финансово. Однако, можно выделить потенциальные выгоды для бизнеса, которые возможно получить при использовании разработанного BI-приложения:

- Облегченный процесс планирования и определения стратегии складской логистики и доставки товаров в магазины;
- Сокращение затрат на доставку товаров;
- Оптимизация времени выполнения рейсов;
- Анализ структуры спроса на определенный товар в определенном районе/магазине;
- Повышение скорости принятия управленческих решений благодаря оперативному предоставлению отчетности;

Также стоит отметить, что внедрение BI-приложения снижает издержки на поддержку BI-приложения. Компании нет необходимости содержать штат сотрудников, которые будут выполнять те BI-задачи, которые автоматизированы в приложении, так как

такие процессы, как выгрузка данных и их преобразование, а также визуализация уже входят в его функционал и достаточно иметь одного VI-специалиста, который будет заниматься непосредственно его поддержкой.

Таким образом, успешность инвестиций в разработку и внедрение VI-приложения определяется улучшением целевых показателей, на которые оно влияет. Однако, это затрудняется тем, что обычно довольно тяжело выявить аспекты бизнеса, на которые влияет приложение, то есть определить долю в изменении показателя, которая была достигнута благодаря внедрению VI-приложения.

Заключение

В результате проделанной работы были изучены теоретические основы ВІ как подхода к анализу информации в сфере бизнеса и ВІ-приложений как технической реализации ВІ в виде ПО и программных платформ. ВІ-технологии увеличивают скорость принятия решений, потому что автоматизация процесса обработки информации позволяет существенно сократить время сотрудника затраченное на выполнение повторяющихся действий технического характера, что позволяет повысить эффективность работы сотрудников в целом.

Для обоснованного выбора среды разработки ВІ-приложения был проведен обзор на лидирующие зарубежные и отечественные ВІ-системы. Выбор пал на Power ВІ, ввиду его богатого функционала и комфортабельности для пользователя.

Во второй части настоящей курсовой работы были проведены мероприятия по подготовке исходных данных к дальнейшему использованию. А именно, проверка на качество, трансформация данных и последующая очистка.

В заключительной главе была разработана и реализована интерактивная платформа визуального представления данных, в соответствии с анализом предметной области и информационных потребностей пользователей. Были разработаны инструменты для анализа информации и интуитивно понятный пользовательский интерфейс.

Разработанное ВІ-приложение позволяет существенно сократить время, затрачиваемое на формирование отчетности, позволяет проводить анализ динамики движения товаров со склада в магазины, оптимизировать мероприятия, связанные с доставкой товаров и проводить сравнительный анализ районов присутствия магазинов.

Список использованных источников:

1. Паспорт проекта Цифровая экономика // URL: <https://base.garant.ru/72296050/> . (дата обращения: 15.03.2023).
2. Белоглазов, Д.С. Управленческое решение: сущность и классификация // Прорывные научные исследования, закономерности, перспективы. – 2017. – С.143-150.
3. Ильяшенко О. Ю. Инновационное развитие ИТ-архитектуры предприятия посредством внедрения системы бизнес-аналитики / И.В.Ильин, А.А.Лепехин //Наука и бизнес: пути развития, №8, 2017, стр. 59-66.
4. Куклина, Е.А. Разработка управленческих решений// Е.А. Куклина//Экономика и управление. – 2017. – №8. – С. 13-16
5. Davenport, Thomas H., Jeanne G. Harris, Robert Morison. Analytics at work: smarter decisions, better results //McGraw-Hill. - 2009. С.288. // URL: <https://books.google.ru/books?id=Jy-8bBjjC38C&hl=ru> (дата обращения: 15.04.2023).
6. Ramesh Sharda, Dursun Delen, Efraim Turban. Business Intelligence and Analytics Systems for Decision Support // Pearson Education Limited. -2014, С.678
7. Rick Sherman. Business Intelligence Guidebook: From Data Integration to Analytics // Elsevier Inc. -2015, С.525
8. Steve Williams, Business Intelligence Strategy and Big Data Analytics// Elsevier Inc. - 2016, С.240.
9. URL: https://en.wikipedia.org/wiki/Data_warehouse (дата обращения: 25.04.2023).
10. URL: https://en.wikipedia.org/wiki/Extract,_transform,_load (дата обращения: 25.04.2023).
11. URL: <https://ru.wikipedia.org/wiki/OLAP> (дата обращения: 25.04.2023).
12. URL: <https://en.wikipedia.org/wiki/Database> (дата обращения: 25.04.2023).
13. URL: <https://learn.microsoft.com/ru-ru/power-bi/> (дата обращения: 25.04.2023).
14. URL: <https://radacad.com/power-bi-basics-of-modeling-star-schema-and-how-to-build-it> (дата обращения: 25.04.2023).
15. URL: <https://www.qlik.com/us/> (дата обращения: 25.04.2023).