

ПОДДЕРЖКА БИЗНЕС-АНАЛИЗА С ИСПОЛЬЗОВАНИЕМ ХРАНИЛИЩА ДАННЫХ БАНКОВСКОЙ СИСТЕМЫ

Студент Медведева Л. Ю.

Научный руководитель — канд. физ.-мат. наук, доцент Головатый А.И.

Белорусский государственный университет,
факультет радиофизики и компьютерных технологий,
кафедра интеллектуальных систем

Минск, Беларусь

Хранилища данных (DWH) играют важную роль в цифровой трансформации, особенно в банковской сфере, где критически важны оперативность и точность принимаемых решений. Процесс их внедрения включает сбор, структурирование и анализ данных, а также тесное взаимодействие бизнес-аналитиков и технических специалистов. Бизнес-анализ обеспечивает сбор требований, проектирование моделей данных и проверку решений, что определяет соответствие хранилища бизнес-целям и успешность проекта.

Таким образом, важную роль в данном процессе играет четкое определение цели и задач разработки.

Целью работы было исследование применения методов бизнес-анализа для разработки и оптимизации различных шаблонов хранилищ данных с целью повышения эффективности обработки и использования информации в организационной среде [1].

Для поставленной цели необходимо было решить следующие задачи:

- исследовать концепцию и архитектуру хранилищ данных (DWH),
- оценить влияние методов бизнес-анализа на эффективность при разработке хранилища данных,
- продемонстрировать применение методов бизнес-анализа для создания хранилища данных банковской системы.

Важным этапом является выбор и обоснование инструментов и технологий,

применяемых в процессе исследования.

Были применены следующие инструменты:

- SQL: язык запросов, используемый для управления данными в хранилище, выполнения операций извлечения, обработки и анализа информации.

- DBeaver: инструмент для работы с базами данных, применяемый для разработки, тестирования и оптимизации SQL-запросов, а также администрирования хранилища [2].

- Apache Superset: платформа для визуализации данных, предназначенная для создания интерактивных дашбордов и аналитических отчетов, обеспечивающих наглядное представление данных.

- ERD-диаграммы: инструмент моделирования, используемый для проектирования структуры базы данных и визуализации связей между ее сущностями.

Первым шагом в процессе создания прототипа являлся сбор требований. На этом этапе бизнес-анализ обеспечивает взаимодействие с заказчиками для понимания их целей и задач. Это включало проведение интервью, опросов или рабочих сессий, направленных на выявление текущих проблем, потребностей и ожиданий. В банковском секторе важно обеспечить доступ к данным по клиентам, транзакциям и кредитным заявкам для анализа в реальном времени. Результаты сбора требований документируются в виде технических спецификаций, User Stories, которые служат основой для работы команды разработчиков.

Далее процесс моделирования. Этот процесс включает разработку концептуальных и логических моделей данных, позволяющих визуализировать структуру данных, их взаимосвязи и пути обработки. Для банковского DWH это модель, описывающая связи между клиентами, их счетами и транзакциями. В рамках моделирования учитываются бизнес-правила, чтобы гарантировать, что данные корректно отражают реальные процессы.

Заключительным этапом явилось тестирование. На это этапе разрабатывались сценарии тестирования для проверки соответствия созданного решения заявленным требованиям. Это включает оценку качества данных, корректности их обработки и

интеграции с другими системами. Например, проверялось, правильно ли агрегируются данные по клиентским счетам за разные периоды времени. Результаты тестирования фиксировались, при необходимости инициировались действия по устранению выявленных проблем, обеспечивая соответствие конечного решения потребностям бизнеса.

Разработка «Хранилища банковской системы» со стороны бизнес-анализа включала несколько ключевых этапов. На первом этапе была проведена разработка ERD-диаграмм, которые визуализировали структуру данных, ключевые сущности, их атрибуты и взаимосвязи. Диаграммы демонстрировали источники данных для заполнения хранилища, что обеспечило понимание процессов интеграции и обработки данных.

На следующем этапе определялась структура, объем и содержание данных для аналитических и операционных нужд. Были сформулированы задачи по включению данных, трансформациям, фильтрации и выборке. Также разрабатывались дашборды с ключевыми метриками и логикой визуализации, задавались правила проверки данных на целостность, актуальность и качество. Кроме того, были определены требования к интеграции данных из разных источников для их унификации и согласования.

Документация включала описание структуры данных, критериев их фильтрации, контроля качества и алгоритмов обработки. Для дашбордов разработано техническое задание с перечнем метрик, схемами визуализации и требованиями к интерфейсу. Также были описаны шаги обработки данных и правила их объединения.

Хранилище данных в DBeaver организовано на трех уровнях (рис.1): Core, где содержатся сырые данные, Stage для их промежуточной обработки, и Datamart, где хранятся агрегированные данные, готовые для аналитики и отчетов.

Были построены различные чарты, такие как линейные графики, столбчатые диаграммы и тепловые карты, для анализа динамики и структуры данных. Например, разработаны графики для отображения ежеквартальной динамики остатков и структуры по банкам.

Созданные чарты объединялись в интерактивные дашборды с возможностью фильтрации данных и обновления визуализаций в реальном времени.

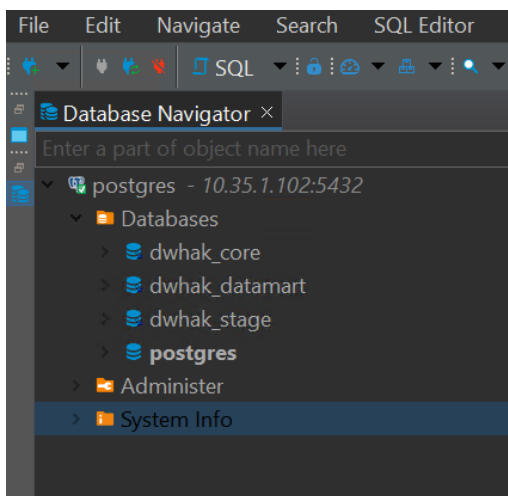


Рисунок 1 Схема базы данных

Визуализация данных осуществлялась в Superset (рис. 2).

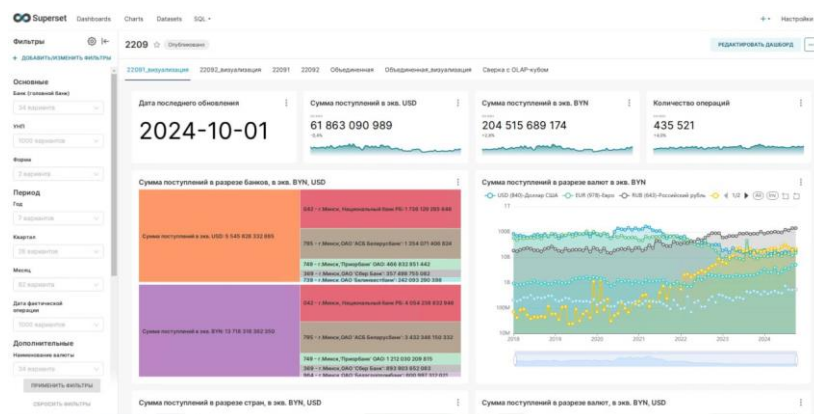


Рисунок 2 Пример дашборда

Для создания дашбордов использовала SQL-запросы, настраивала визуализации, фильтры и параметры аналитики, что обеспечило точность и удобство работы с данными.

Разработанное хранилище данных предназначено для эффективного управления и анализа данных в банковской системе. В процессе его создания использовались передовые методы проектирования хранилищ данных, включая многоуровневую архитектуру (stage, core, datamart), а также современные инструменты для обработки и визуализации данных. Основой разработки стали платформы, такие как DBeaver и Apache Superset, обеспечивающие высокую производительность, надежность и удобство работы с данными.

Литература

1. Что такое хранилище данных Data Warehouse и зачем оно бизнесу [Электронный ресурс]. – Режим доступа: <https://selectel.ru/blog/data-warehouse/> Дата доступа: 21.05.2024.
2. DBeaver [Electronic resource]. –. Режим доступа: [DBeaver Community | Free Universal Database Tool](#) Дата доступа: 21.05.2024.
3. Как бизнес-аналитику выстроить работу в проектах DWH (Data Warehouse). Структура работ Режим доступа: <https://habr.com/ru/articles/807913/>Дата доступа: 21.05.2024