1

Московский государственный технический университет им. Н.Э. Баумана

Факультет «Информатика и системы управления»

Кафедра «Автоматизированные системы обработки информации и управления»



**Отчет**

**Лабораторная работа № 1**

**«Основные объекты бизнес - аналитики BI QlickView»**

**ИСПОЛНИТЕЛЬ:**

Белоусов Евгений

Группа ИУ5-51

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

"\_\_"\_\_\_\_\_\_\_\_\_\_\_2019 г.

**ПРЕПОДАВАТЕЛЬ:**

Селиверстова А.В.

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

"\_\_"\_\_\_\_\_\_\_\_\_\_\_2019 г.

Москва 2019

**Содержание отчета по лабораторной работе**

1. **Цель лабораторной работы**

Установка BI QlickView версии Personal Edition Ознакомиться с интерфейсом OLE DB для подключения к внешним источникам данных и основными объектами платформы бизнес аналитики BI QlickView. Изучить следующие объекты платформы - Мастер начала работы - Команды меню - Панели инструментов и строка состояния

1. **Задачи, выполняемые в лабораторной работе**

**2.1 Теоретическая часть**

**2.1.1 Средства бизнес аналитики**

* + - 1. **Существующие платформы бизнес аналитики**
* Реализация платформы бизнес аналитики на базе компонент системных сервисов, ROLAP, мондриана, MDX запоров и стандартными средствами визуализации
* Реализация платформы бизнес аналитики на базе компонент OLTP, Kettle, Pentaho BI
  + - 1. **Сравнительный анализ платформы бизнес аналитики**

Pentaho BI – пакет ориентированный более на работу с OLAP, и в частности, с OLAP движком Mondrian. Первое – это то, что сам по себе этот пакет предоставляет крайне скудные возможности по генерированию различных форм отчетов. Представлено только получение таблиц, но выбор нужных колонок и столбцов затруднен для оператора. Получение диаграмм, графиков и таблиц с периодом невозможно вообще. Эти проблемы частично решаемы с помощью нескольких дополнений к Pentaho BI server, в частности, CDF и Saiku.

Модуль CDF работает на базе Pentaho являясь графической надстройкой на нем. Позволяет создавать несколько различных видов диаграмм. Возможно только создание готовых решений администратором и затем пользователь сможет выбрать только из этих готовых решений. Создание своих отчетов затруднено оператором из-за сложности интерфейса и обилия технических настроек.

Saiku – самая подходящая альтернатива собственной системе. Многие идеи были взяты из Saiku. На тот момент система подавала большие надежды, так как была лишена вышеописанных недостатков. Интерфейс понятен для пользователя, технические детали отсутствуют, но так как система не автономна и в ее работе участвуют несколько различных модулей (Mondrian, например), то добиться слаженной работы этих модулей оказалось затруднительно. В этом случае сложна оказалась работа администратора. Каждому новому администратору приходилось бы изучать большое количество документации как по Mondrian, так и по Saiku, чтобы вникнуть в детали работы. Документация только на английском языке, что добавляет сложностей в поддержке системы.

Jaspersoft в плане поддержки и работы с OLAP почти ничем не отличается от Pentaho и недостатки там точно такие же. У всех этих систем (Pentaho + CDF, Saiku, Pentaho) есть несколько больших и решающих недостатков или же резюмируя: системы не гибки, использование их с имеющейся базой данных затруднено, скорость работы крайне низка

* + 1. **Платформа QlikView**
       1. **Ассоциативная модель данных QlikView**

Ассоциативная модель данных представляет собой таблицы,связанные по ключевым полям. Ассоциативная модель создается динамически в момент загрузки данных в QlikView, при этом поля с одинаковыми названиями ассоциируются друг с другом. При выборе пользователем значения в одной таблице автоматически выбираются все ассоциированные значения в других таблицах. Например, для модели представленной ниже, при выборе клиента, автоматически выбираются все заказы, сделанные этим клиентом, так как таблица заказов связана (ассоциирована) с таблицей клиентов. Но так как таблица заказов связана (ассоциирована) также с таблицей товаров, то кроме всех заказов данного клиента автоматически выбираются и все товары, купленные этим клиентом. И наоборот — при выборе товара, автоматически выбираются все заказы, где фигурирует этот товар, и все клиенты, сделавшие эти заказы, а значит купившие этот товар. Для ассоциативной модели не требуются заранее прописанные пути анализа (drill-down). Она позволяет начинать анализ данных с любого места модели в любом направлении. Ассоциативная модель позволяет связывать сразу несколько таблиц по одному полю с одинаковым названием. Каждые две таблицы могут быть связаны только одним полем, во избежание неоднозначности ассоциативной связи. При необходимости связи двух или более таблиц по нескольким полям сразу, необходимо формирание дополнительного поля в качестве ключа, путем комбинированния полей.

**2.1.2.2 Обработка данных в оперативной памяти (технология**

**in-memory)**

*Фундаментальным отличием QlikView является свой собственный патентованный механизм обработки данных, который обрабатывает все данные в оперативной памяти. Благодаря этому скорость выполнения аналитических расчетов в десятки и сотни раз превосходит традиционные реляционные СУБД, так как отсутствуют операции дискового ввода-вывода, традиционно являющиеся узким местом в производительности аналитическим систем. При загрузке данных в оперативную память они сжимаются в несколько раз, поэтому данные в памяти занимают меньше места, чем в системах-источниках. При больших объемах данных целесообразно использовать 64х-битную платформу. Ассоциативная модель данных физически представляет собой таблицы, загружаемые из системисточников целиком в оперативную память с помощью специального SQLподобного загрузочного скрипта QlikView, создаваемого пользователем. Таблицы могут быть загружены из разных источников – например, одна таблица из реляционной СУБД, вторая из текстового файла, третья из Excel и т.д. Когда таблицы загружены в оперативную память и связаны, для QlikView не имеет значения источник данных. Загруженные данные могут быть сохранены на диске вместе с приложением (отчетом) QlikView. При открытии пользователем этого приложения (отчета) повторно данные загружаются в оперативную память без повторного извлечения из систем-источников. При необходимости, повторное извлечение (обновление) данных из систем-источников (перезапуск загрузочного скрипта) может быть сделано вручную пользователем, либо автоматически по расписанию.*

**2.1.2.3 Версия QlikView Personal Edition**

Благодаря QlikView Personal Edition компания QlikTech представляет бесплатную версию программного обеспечения QlikView для личного использования. Эта версия предназначена для частных лиц, студентов или небольших новых компаний. QlikView Personal Edition — это полноценный продукт QlikView Desktop, который использует тот же установочный пакет. Единственное отличие заключается в том, что QlikView Personal Edition работает без лицензионного ключа.

**2.1.2.4 Мастер начала работы**

Мастер Приступая к Работе служит для создания документа QlikView путем загрузки данных и создания диаграммы.

**2.1.2.5 Команды меню**

Описанные в этой главе меню находятся в строке меню в верхней части экрана. Для большинства команд также можно настроить кнопки на панелях управления. Для получения дополнительных сведений см. раздел Настроить (Панели инструментов)...

**2.1.2.6 Меню «Правка»**

Раскрывающееся меню Правка располагается в верхней части экрана и включает следующие команды: Отменить изменение макета, Восстановить изменение макета, Вырезать, Копировать, Вставить, Средства форматирования, Удалить, Поиск.

**2.1.2.7 Меню «Выборка»**

Меню Выборки представляет собой раскрывающееся меню в верхней части экрана и содержит следующие команды: Назад, Вперед, Разблокировать, Заблокировать и тд...

**2.1.2.8 Панели инструментов и строка состояния**

QlikView имеет пять стандартных панелей инструментов и одну строку меню. Стандартные панели инструментов, как правило, включают кнопки задач, выполняемых с документами QlikView, а панель инструментов «Конструктор» обычно включает кнопки задач, выполняемых при создании или изменении макета документа. Панель инструментов навигации содержит наиболее часто используемые команды логических операторов в документе. Панель инструментов листа предлагает альтернативный метод поиска различных листов, а панель инструментов закладок — альтернативный способ доступа к закладкам. Каждая панель инструментов включается и выключается отдельно. Все панели инструментов являются настраиваемыми и могут включать любые кнопки доступных команд. Все панели инструментов можно переместить, щелкнув пунктирную линию, расположенную слева. Удерживая левую кнопку мыши нажатой, перетащите панель инструментов в любое положение. Панель инструментов можно закреплять с любой стороны окна приложения QlikView.

* + - 1. **Параметры пользователя**

Параметры пользователя — это параметры, которые относятся к способу работы пользователя и хранятся на компьютере, а не в файле документа.

* + 1. **Форматы данных XML**
       1. **Основы XML**

Язык Extensible Markup Language (XML) можно назвать и языком разметки, и форматом хранения текстовых данных. Это подмножество языка Standard Generalized Markup Language (SGML); он предоставляет текстовые средства для описания древовидных структур и их применения к информации. XML служит основой для целого ряда языков и форматов, таких как Really Simple Syndication (RSS), Mozilla XML User Interface Language (XUL), Macromedia Maximum eXperience Markup Language (MXML), Microsoft eXtensible Application Markup Language (XAML) и open source-язык Java XML UI Markup Language (XAMJ).

**2.1.3.2 Структура XML**

Базовым блоком данных в XML является элемент. Элементы выделяются начальным тегом, таким как <Student>, и конечным тегом, таким как </Student>. Каждому начальному тегу должен соответствовать конечный тег. Если для какого-то начального тега отсутствует конечный тег, XML-документ оформлен неправильно, и синтаксический анализатор (парсер) не сможет проанализировать его надлежащим образом. Названия тегов обычно отражают тип элемента.Элемент Student содержит <LastName- фамилия>, < FirstName -имя>,< MiddleName -отчество>, < BirthDate - день рождения>,

< Gender -пол>, < University -названия ВУЗ-а>, и т.д. рис. Текст, содержащийся между тегами, включая пробелы, называется символьными данными.

**2.1.3.3 Достоинства и недостатки XML**

***Недостатки XML***

XML многословен и избыточен, что порождает документы большого объема, занимающие много дискового пространства и сетевых ресурсов. Предполагается, что он должен быть читабелен для людей, но трудно представить себе человека, пытающегося прочесть файл XML с 7 млн. узлов. Простейшие синтаксические анализаторы функционально не способны поддерживать широкий набор типов данных; по этой причине редкие или необычные данные, каких бывает много, становятся серьезным источником затруднений.

***Достоинства XML***

Одно из достоинств XML состоит в его относительной простоте. XML-документ можно составить в простом текстовом редакторе или текстовом процессоре, не прибегая к специальным инструментам или ПО. Базовый синтаксис XML состоит из вложенных элементов, некоторые из которых имеют атрибуты и содержание. Обычно элемент начинается открывающим тегом <тег> и заканчивается соответствующим закрывающим тегом </тег >. XML чувствителен к регистру и не игнорирует пробелы и табуляции. Он очень похож на HTML, но, в отличие от HTML, позволяет присваивать тегам имена для лучшего описания свох данных. К числу преимуществ XML относится самодокументирование, читабельный для людей и компьютеров формат, поддержка Unicode, что позволяет создавать документы на разных языках, и простые требования к синтаксису и синтаксическому анализу. К

сожалению, в РНР5 поддержка UTF-8 сопряжена с проблемами; это один из тех недостатков, которые привели к разработке РНР6.

* 1. **Практическая часть**

**2.2.1 Результаты, полученные в лабораторной работе:**

**2.2.1.1 Создание XML файла «Студент»**

<Student>

<StudentID>4</StudentID>

<University>МГТУ им. Н.Э. Баумана</University>

<Faculty>ИУ</Faculty>

<SubFaculty>ИУ-5</SubFaculty>

<GroupName>ИУ-51</GroupName>

<LastName>Белоусов</LastName>

<FirstName>Евгений</FirstName>

<MiddleName>Александрович</MiddleName>

<BirthDate>16.10.98</BirthDate>

<Gender>male</Gender>

<State>Russian Federation</State>

<StartDate>01.09.2017</StartDate>

<Enrollment>Общий конкурс</Enrollment>

<Grant>Стипендия</Grant>

<Dormitory>Общежитие</Dormitory>

<Address>Красная площадь, д.1</Address>

<City>Москва</City>

<Citizenship>Гражданин РФ</Citizenship>

<Country>РФ</Country>

<Phone>+7 800 555 35 35</Phone>

</Student>

**2.2.1.2 Создание XML файла «Студенты»С**

Файл студенты состоит из большого числа записей, идентичных записи в файле Студент, объединенных в общий тег.

1. **Выводы.**

При выполнении лабораторной работы №1, я установил BI QlickView версии Personal Edition, ознакомился с интерфейсом OLE DB для подключения к внешним источникам данных и основными объектами платформы бизнес аналитики BI QlickView. Изучил следующие объекты платформы - Мастер начала работы - Команды меню - Панели инструментов и строка состояния.