



Акционерное общество «Флекс Софтваре Системс»
127055 Россия, г. Москва, ул. Новолесная 2, офис 3
Тел.: +7 (495) 788 - 03 - 25
e-mail: info@flexsoft.com
www.flexsoft.com

Установка базовых компонентов

BI Pentaho

Москва — 2020

Оглавление

1. Цели документа и область применения.....	3
2. Глоссарий.....	4
3. Настройка операционной системы.....	5
4. Pentaho BI Suite.....	7
5. Установка Pentaho BI Suite Community Edition	8
6. Запуск Pentaho BI-сервер.....	9
7. Начало работы с Pentaho BI-сервер.....	10
8. Выполнение анализа в Pentaho BI-Suite.....	18

1. Цели документа и область применения

Данный документ описывает процедуру установки базовых системных компонентов, необходимых для развёртывания программного обеспечения.

В документе приводится описание установки комплекса на примере ОС Oracle Linux 7 Update 6 (x86-64), процедура установки под другой редакцией и версией ОС может отличаться по части пакетной базы и процедур запуска сервисов. Документ предназначен для использования персоналом, ответственным за обслуживание и администрирование комплекса.

2. Глоссарий

Термин или сокращение	Определение
Oracle	компания Oracle и продукты, относящиеся к ней
FXL	«Core System FXL» - семейство продуктов, разрабатываемых и поддерживаемых компанией ФлексСофт
ОС	операционная система
ПО	программное обеспечение
СУБД	программное обеспечение для управления базами данных
APP	краткое наименование сервера приложений
Патч	исправление/доработка функционала FXL
Hot-Fix	внеплановое исправление ошибок, выявленных в работе FXL, требующие как можно быстрого исправления, поставляются не в рамках графика общей поставки патчей
Служба технической поддержки	служба технической поддержки клиентов АО «ФлексСофт»

3. Настройка операционной системы

- 1) должен быть корректно сконфигурирован файл `/etc/hosts`, в случае если в файле описан только `loopback`-интерфейс, необходимо добавить описание сетевого интерфейса и `hostname`;
- 2) На время установки рекомендуется деактивировать межсетевой экран, сделать это можно следующей командой, выполненной от имени супер-пользователя `root`:

```
#systemctl stop firewalld
```

- 3) Должен быть установлен следующий перечень пакетов:

- `oracle-rdbms-server-12cR1-preinstall.x86_64`
- `binutils-2.23.52.0.1`
- `compat-libcap1-1.10`
- `compat-libstdc++-33-3.2.3.x86_64`
- `compat-libstdc++-33-3.2.3.i686`
- `gcc-4.8.2`
- `gcc-c++-4.8.2`
- `glibc-2.17.x86_64`
- `glibc-2.17.i686`
- `glibc-devel-2.17.x86_64`
- `libaio-0.3.109.x86_64`
- `libaio-devel-0.3.109.x86_64`
- `libgcc-4.8.2.x86_64`
- `libgcc-4.8.2.x86_64`
- `libgcc-4.8.2.i686`
- `libstdc++-4.8.2.x86_64`
- `libstdc++-4.8.2.i686`
- `libstdc++-devel-4.8.2.x86_64`
- `ksh`
- `make-3.82`
- `sysstat-10.1.5`
- `numactl-2.0.9.x86_64`
- `numactl-devel-2.0.9.x86_64`
- `motif-2.3.4-7.x86_64`
- `motif-devel-2.3.4-7.x86_64`
- `zip / unzip`

4. Pentaho BI Suite

Продукт Pentaho состоит из отдельных модулей-компонентов, которые вместе формируют Pentaho BI Platform:

- Pentaho Reporting - модуль создания и представления отчетов;
- Pentaho Analysis - модуль для выполнения OLAP-анализа;
- Pentaho Dashboard - модуль для управления и прогнозирования;
- Pentaho Data Integration - модуль для интеграции данных из разных источников;
- Pentaho Data Mining - модуль для поиска скрытых зависимостей в данных.

Как и в случае с предыдущим продуктом, Pentaho BI Suite представлен в двух версиях: бесплатном Community Edition и коммерческом Enterprise Edition. Эти версии отличаются уровнем поддержки и наличием продвинутой, хотя и не очень востребованной небольшими компаниями функциональности: кластеризацией, мониторингом производительности и т.д. Оставшаяся часть этой статьи будет посвящена знакомству с бесплатной версией Pentaho BI Suite.

Стоит отметить, что кроме BI-продуктов компания Pentaho также предлагает продукты для интеграции данных. Этот DI-продукт (data integration) позволяет переносить данные из одной СУБД в другую, облегчая задачу сопровождения, если компания использует несколько СУБД в разных филиалах или выполняет переход с одной версии СУБД на другую. Кроме того, в ходе переноса данные могут быть преобразованы, если это необходимо, а сами интеграционные сеансы могут выполняться без вмешательства пользователя и по установленному расписанию, когда число активных пользователей СУБД не велико.

5. Установка Pentaho BI Suite Community Edition

Продукт Pentaho BI Suite разработан на языке программирования Java, поэтому обладает всеми преимуществами и недостатками, свойственными Java-приложениям. К преимуществам относится мультиплатформенность, благодаря которой Pentaho BI Suite может работать на разных ОС, при этом однозначно поддерживаются: Windows XP SP2, современные дистрибутивы Linux, Solaris 10 и MacOS X 10.4. Также для работы с этим BI-продуктом потребуется установленная версия JDK (комплекта разработчика Java) версии не менее 1.5 (5.0), так как в проекте используются возможности, появившиеся в Java только с этой версии. При установке Pentaho на Linux-систему может возникнуть проблема, связанная с компилятором GCJ (GNU compiler for Java), который используется по умолчанию на многих системах. Поэтому лучше отключить этот компилятор каким-либо способом и перейти на стандартный компилятор `javac`, входящий в JDK.

Бесплатная версия Pentaho находится на Web-сайте портала SourceForge - известного репозитория проектов с открытым кодом; там же, кстати, находятся и дистрибутивы BI-решений от JasperSoft. Страница с компонентами BI-решения от Pentaho находится по этому адресу. Обязательно нужно загрузить последнюю стабильную версию Business Intelligence Server, так как это ядро продукта. В BIS-сервер также входят встроенные версии сервлет-контейнера Tomcat и локальной СУБД Hupersonic, так что им можно пользоваться сразу после установки. Дополнительно можно загрузить и клиентские приложения для работы с сервером:

- Report Designer
- Data Integration
- Design Studio
- Pentaho Metadata

Загруженные файлы необходимо развернуть в папке, выбранной для установки ПО.

6. Запуск Pentaho BI-сервер

Для начала работы с Pentaho BIS необходимо запустить его основные компоненты: сам BI-сервер и его административную консоль. Для этого необходимо выполнить следующие сценарии.

Для Linux:

```
%PENTAHO_BI_SERVER_INSTALLATION%\start-pentaho.sh
%PENTAHO_BI_SERVER_INSTALLATION%\administration-console\start.sh Для Windows:
%PENTAHO_BI_SERVER_INSTALLATION%\start-pentaho.bat
%PENTAHO_BI_SERVER_INSTALLATION%\administration-console\startup.bat
```

Примечание. Возможно, придется внести изменения в сценарий запуска административной консоли, явно указав месторасположение интерпретатора java.

Для остановки сервера можно использовать соответствующие сценарии с префиксом stop, находящиеся в этих же каталогах. После того как оба сценария запущены, можно войти в административную консоль с помощью любимого Web-браузера; ее URL-адрес по умолчанию: <http://localhost:8099/admin>. По умолчанию для администрирования создана учетная запись admin/password (имя пользователя/пароль). В административной консоли необходимо убедиться, что сервер успешно запустился: иконка Server Status (статус сервера) справа от логотипа в верхнем ряду.

7. Начало работы с Pentaho BI-сервер

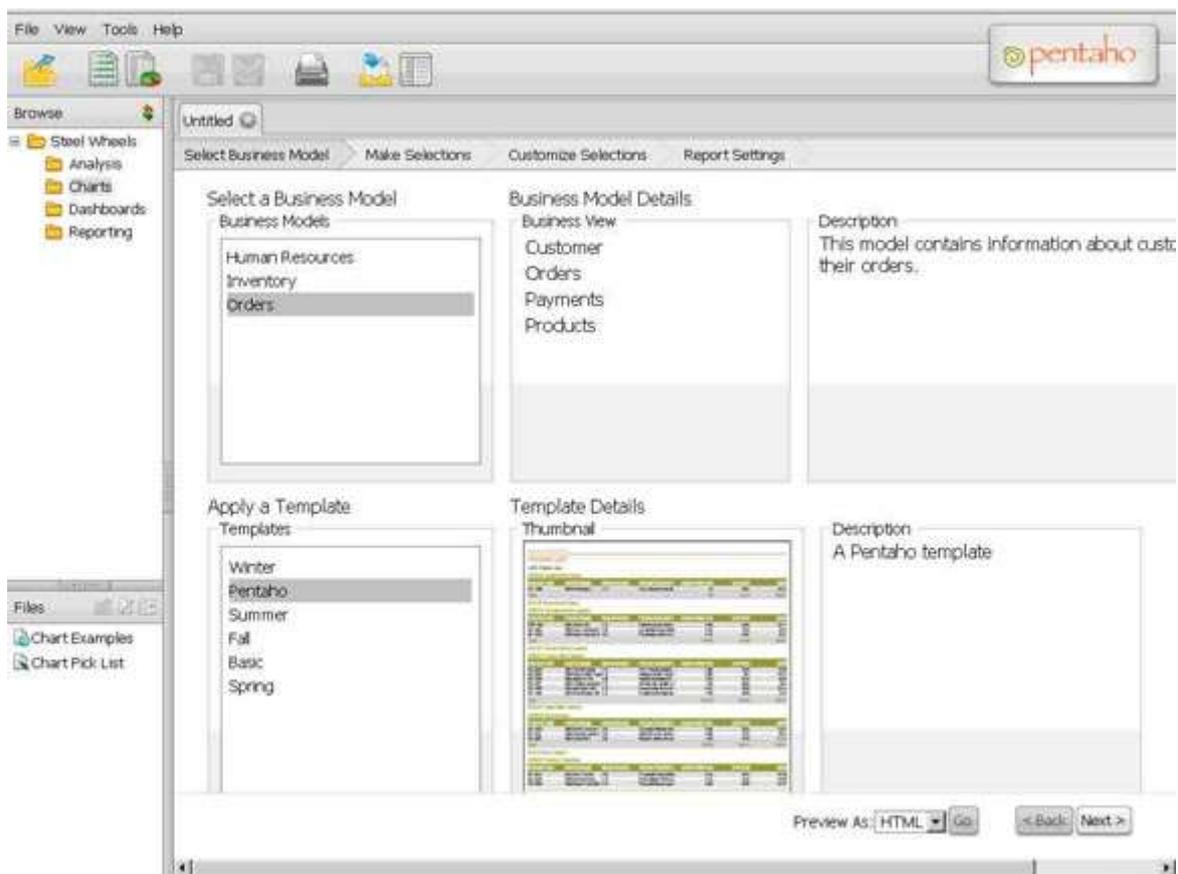
После запуска и проверки статуса сервера можно начинать использовать его на уровне обычного пользователя для решения BI-задач. Для этого нужно открыть в Web-браузере консоль пользователя по URL-адресу: <http://localhost:8080/>. Для входа в выпадающем списке пользователей необходимо нажать кнопку Login, выбрать пользователя Джо из списка и нажать кнопку Login. На рисунке 1 приведен пример GUI-интерфейса Pentaho. Возможности, предоставляемые этим интерфейсом, будут различаться в зависимости от уровня доступа пользователя (пользователь Джо является администратором и имеет доступ ко всей функциональности).

GUI-интерфейс Pentaho



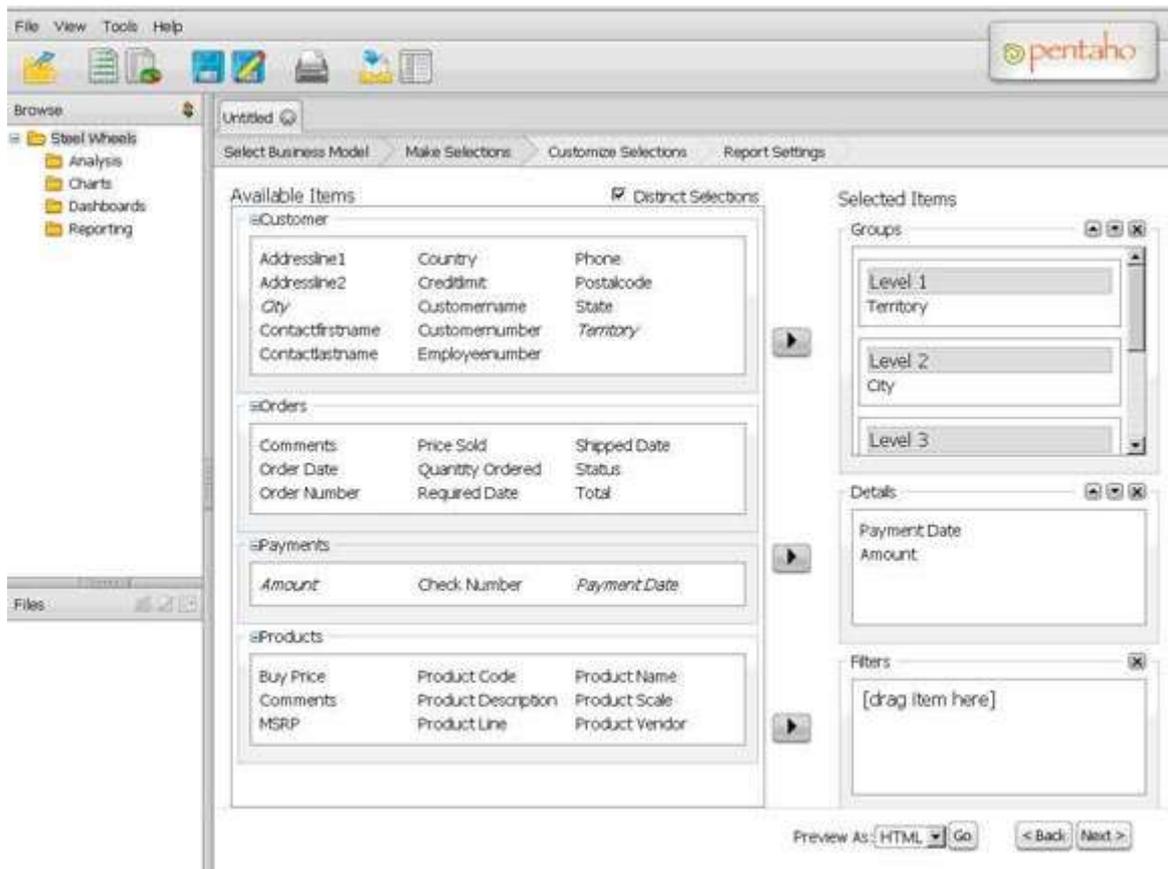
Воспользуемся основной возможностью, предоставляемой всеми BI-решениями, - созданием отчета по уже имеющимся в системе данным (Pentaho BI-сервер поставляется с предустановленным набором данных, которые можно использовать для ознакомления с продуктом). Для этого необходимо нажать кнопку **New Report** (новый отчет) в главном окне Pentaho, после чего запустится мастер создания отчетов. Первая форма для создания отчета показана на рисунке 2.

Мастер создания отчетов: выбор данных и шаблона отчета



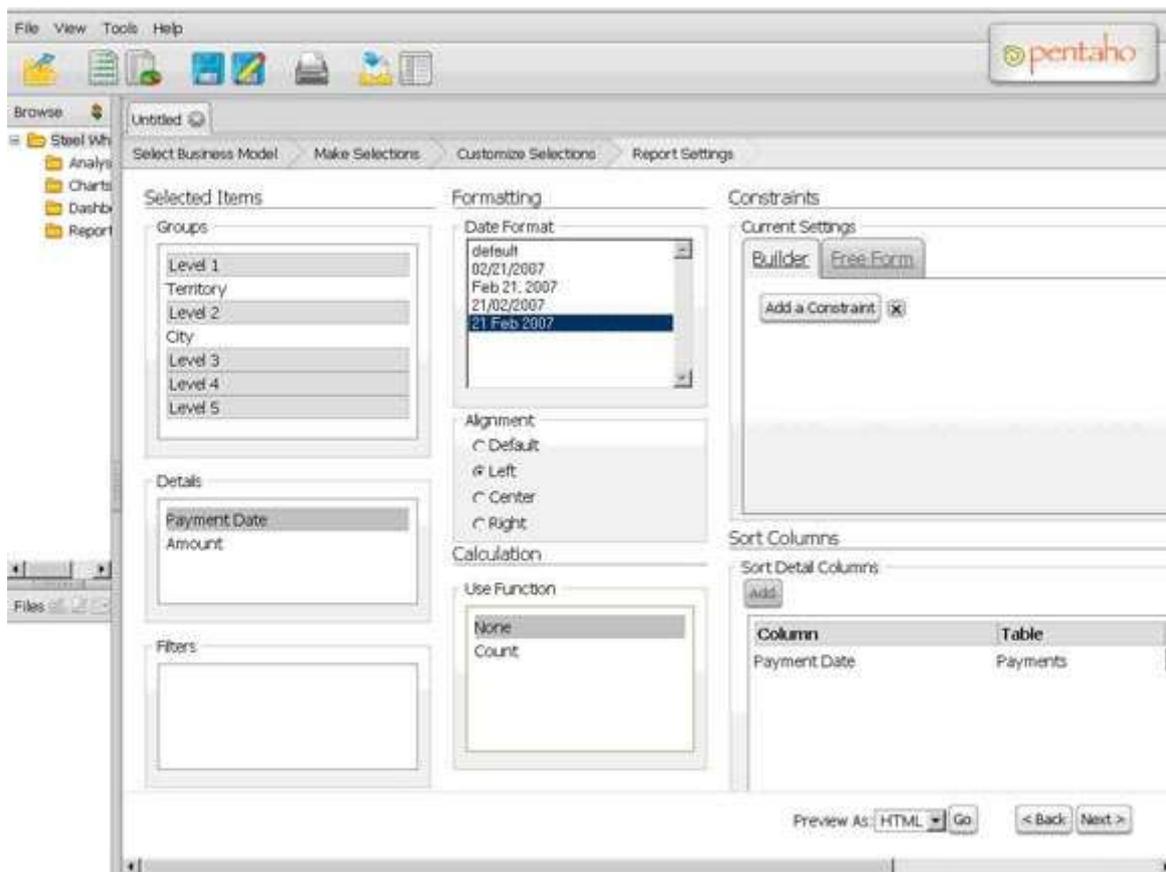
На Select Business Model (выбрать бизнес-модель) в списке Business Models (бизнес-модели) необходимо выбрать набор данных, для которого будет строиться отчет, и указать стиль шаблона для форматирования этого отчета. В данном случае был выбран набор данных Orders (заказы) и фирменный стиль шаблона Pentaho. После этого необходимо нажать кнопку Next, чтобы перейти к следующей форме мастера (рисунок 3).

Мастер создания отчетов: выбор столбцов для отчета и группировки



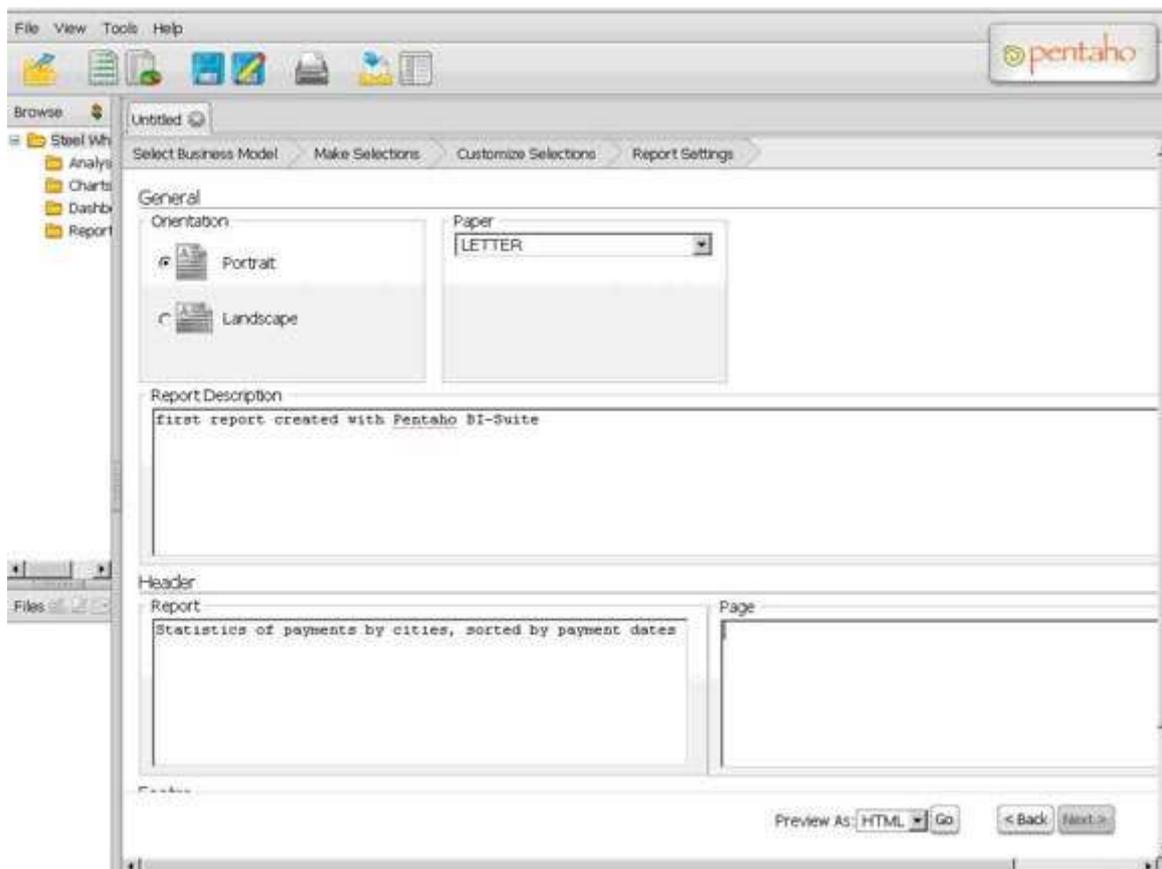
В форме Make Selections (выбрать данные) необходимо указать, какие столбцы из таблиц, входящих в набор данных, будут отображаться в отчете. При этом необходимо указать столбцы для группировки и собственно столбцы, содержащие данные. В данном случае используются два уровня группировки - сначала данные группируются по столбцу Territory (территория), а затем по столбцу City (город). Чтобы указать эти столбцы для группировки, необходимо просто перенести их мышью в соответствующие позиции в области Groups (группы) - Territory становится группировкой первого уровня (Level 1), а City - второго (Level 2).

Далее нужно выбрать столбцы, содержащие данные, и точно так же перенести их в группу Details (подробная информация); для этого были выбраны столбцы Amount (сумма платежа) и Payment Date (дата платежа). Если какие-либо столбцы были добавлены по ошибке или в неправильном порядке, то их можно удалить/поменять местами с помощью соответствующих кнопок (вверх/вниз/удалить) в правом верхнем углу каждой области. Выполнив эти действия, необходимо нажать кнопку Next и перейти к следующей форме Customise selections (кастомизация выбранных данных), рисунок 4.

Мастер создания отчетов: настройка представления данных

На этом этапе создания отчета можно указать различные параметры отображения данных (выравнивание, стиль форматирования, тип сортировки). Изменим параметры отображения столбца Payment Date; для этого необходимо выделить столбец Payment Date, щелкнув по нему левой кнопкой мыши. В области Date Format (формат даты) необходимо выбрать стиль 21/02/2007, в области Alignment (выравнивание) необходимо выбрать опцию Left (по левому краю), а в области Sort Details Columns (сортировка столбцов с данными) нажать кнопку Add (добавить), чтобы отсортировать данные по возрастанию значений в столбце Payment Date. После этого можно нажать Next и перейти к завершающему этапу создания отчета - форме Report Settings (настройки отчета), рисунок 5.

Мастер создания отчетов: установка параметров отчета



В этой форме необходимо указать глобальные параметры отчета: размер (Paper) и ориентацию страниц отчета (Orientation), его описание (Report Description), заголовок (Header) и заключение (Footer).

Примечание. Параметры Page в областях Header и Footer будут применяться к отчету только в том случае, если он выводится в формате, где происходит разбиение на страницы (к таким форматам относится только PDF). Указав требуемые значения, можно выбрать желаемый формат отчета в выпадающем списке Preview As (предварительный просмотр в формате) и нажать кнопку Go (начать), чтобы увидеть пример выполнения отчета. На рисунке 6 показан пример созданного отчета в формате HTML.

Пример отчета, созданного Pentaho BI-suite

Statistics of payments by cities, sorted by payment dates
Territory: APAC
City: Auckland

Payment Date	Amount
05/07/2003	\$2 478,8
16/07/2003	\$37 754,33
29/11/2003	\$26 421,24
16/12/2003	\$23 294,8
24/06/2004	\$40 034,67
01/07/2004	\$41 418,55
24/12/2004	\$43 930,62
31/12/2004	\$59 693,29
23/05/2005	\$39 267,74
25/06/2005	\$34 332,29

City: Chatswood

Payment Date	Amount
03/08/2003	\$29 397,26
02/08/2004	\$41 297,14
03/01/2005	\$37 906,16

City: Glen Waverly

Payment Date	Amount
10/12/2003	\$37 878,55
28/07/2004	\$12 334,82

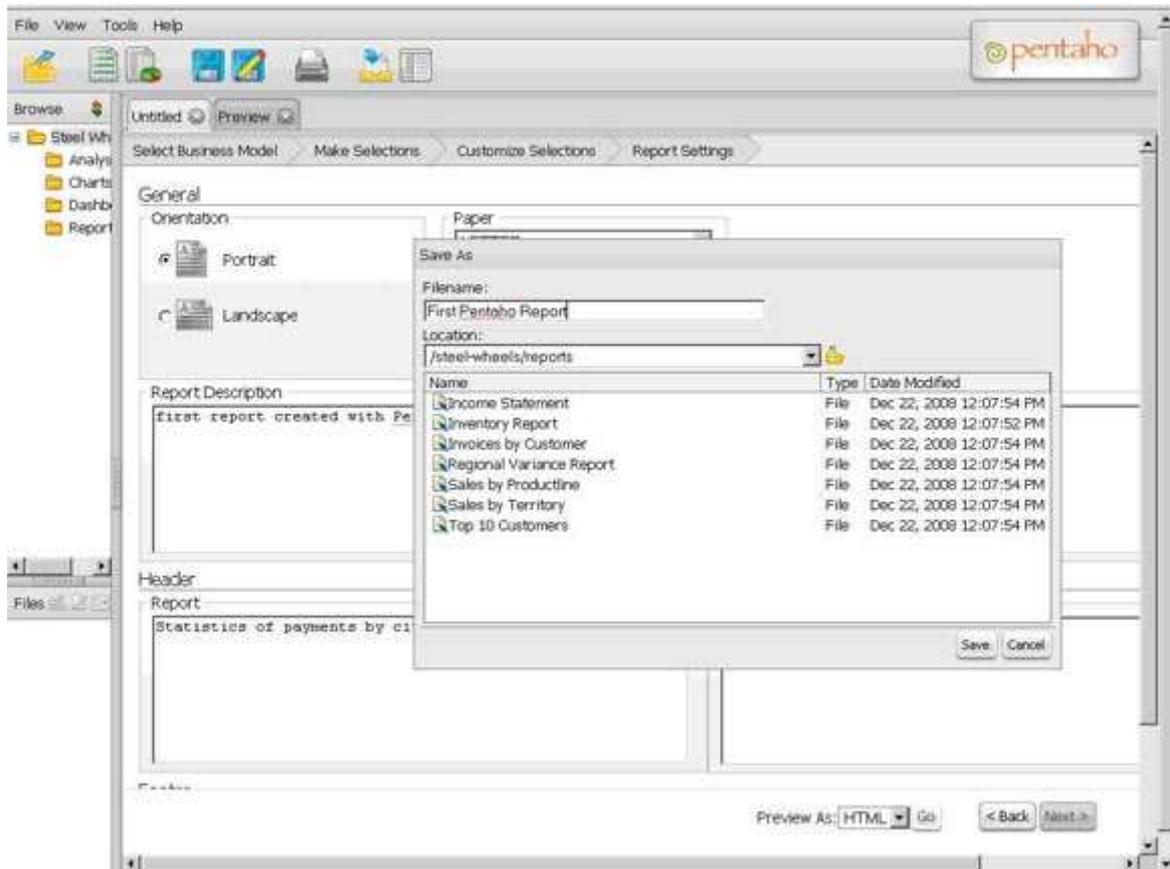
City: Melbourne

Payment Date	Amount
20/05/2003	\$50 397,66
31/05/2003	\$9 730,18
10/03/2004	\$49 637,57
15/12/2004	\$85 691,32

City: North Sydney

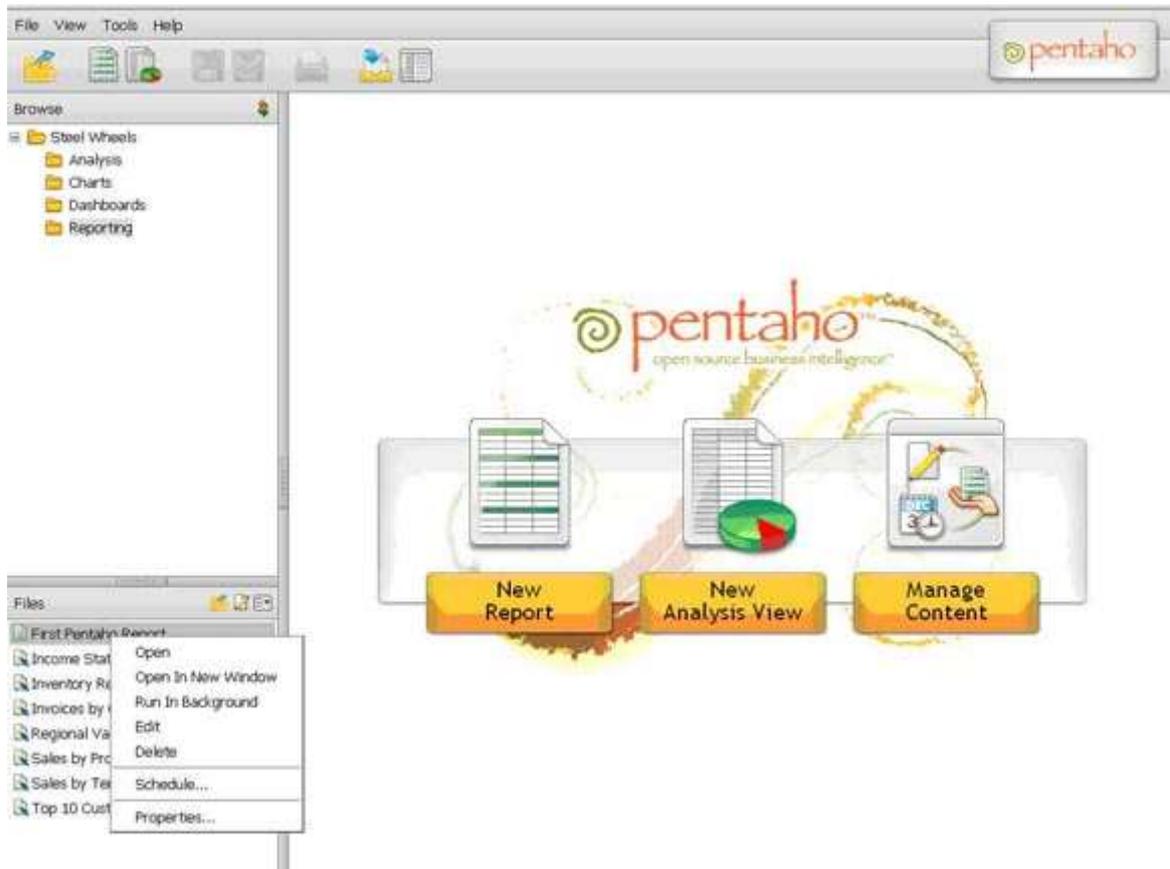
После того как отчет был создан, его стоит сохранить для дальнейшего использования; для этого необходимо выбрать вкладку с отчетом, чтобы снова переключиться на него. Эта вкладка называется Untitled (у отчета пока нет имени) и находится под панелью инструментов GUI-интерфейса. Затем нужно нажать кнопку Save As (сохранить как - пиктограмма с изображением дискеты с карандашом), выбрать папку для сохранения (они указываются относительно сервера, а не локальной файловой системы), ввести название файла отчета и нажать кнопку Save. Пример сохранения отчета показан на рисунке 7.

Сохранение отчета для последующего использования



Теперь этот отчет можно найти через панель Browse (просмотр) в левой части GUI-интерфейса, если открыть папку (Steel Wheels/Reports), в которой он хранится. Для запуска этого отчета необходимо открыть его контекстное меню, щелкнув по нему правой кнопкой мыши, и выбрать опцию Open (открыть). Для редактирования необходимо выбрать в этом меню опцию Edit (редактировать), тогда отчет откроется в режиме редактора (рисунок 8).

Контекстное меню отчета



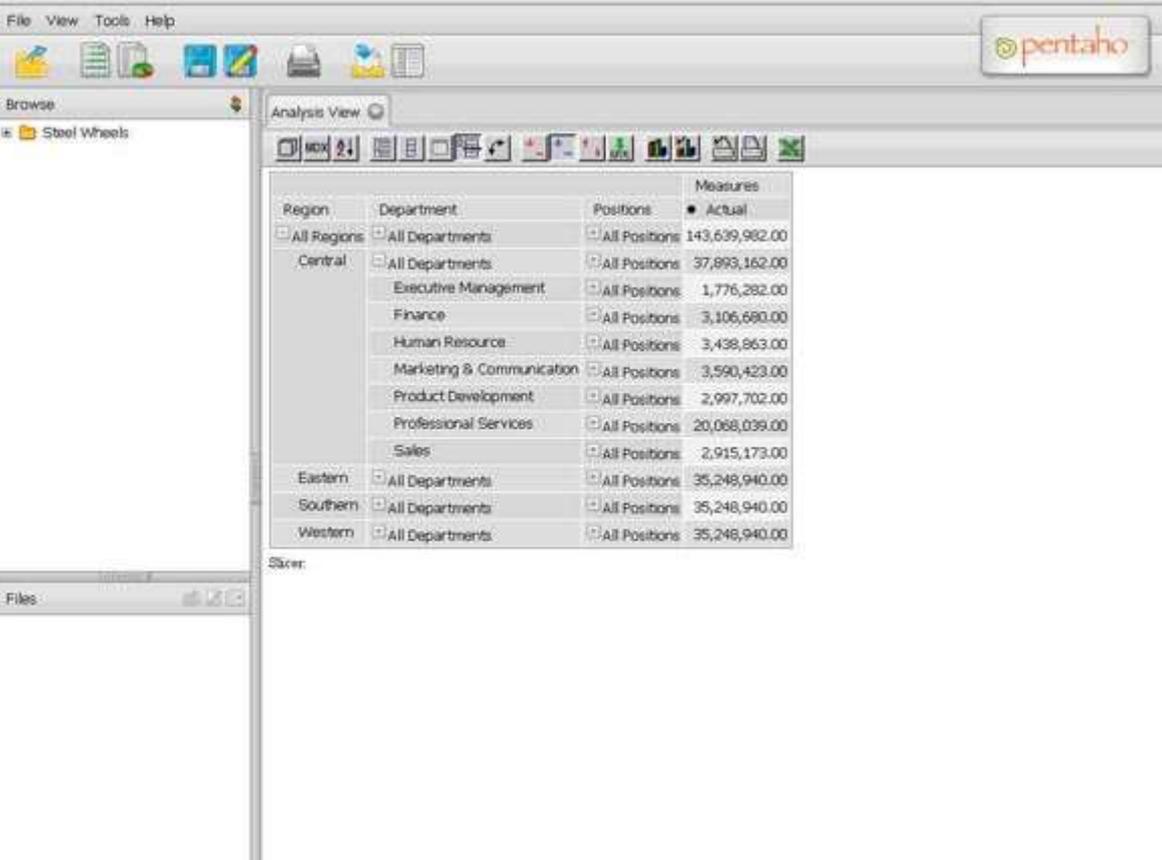
Опция Schedule в контекстном меню позволяет выполнять отчет по расписанию в указанные периоды времени, а опция Properties (свойства) позволяет устанавливать глобальные параметры отчета, например такие, как права доступа. Это позволяет рядовым пользователям использовать отчеты, которые были созданы для них другими, более опытными пользователями.

Для завершения работы с Pentaho BI-suite необходимо выбрать опцию Logout (выйти из системы) в меню File.

8. Выполнение анализа в Pentaho BI-Suite

Основной целью BI-продуктов является не создание статических по своей природе отчетов, а выполнение интерактивного анализа различной степени сложности. Попробуем провести анализ данных с помощью Pentaho BI Suite; для этого необходимо нажать кнопку New Analysis View (новое аналитическое представление) на главном экране GUI-интерфейса. После этого будет показано диалоговое окно, в котором в выпадающем списке Schema (схема данных) надо будет выбрать значение SampleData, а в выпадающем списке Cube (тип анализа) выбрать значение QuadrantAnalysis и нажать кнопку ОК. После этого будет показано окно, изображенное на рисунке 9.

Окно для выполнения интерактивного анализа

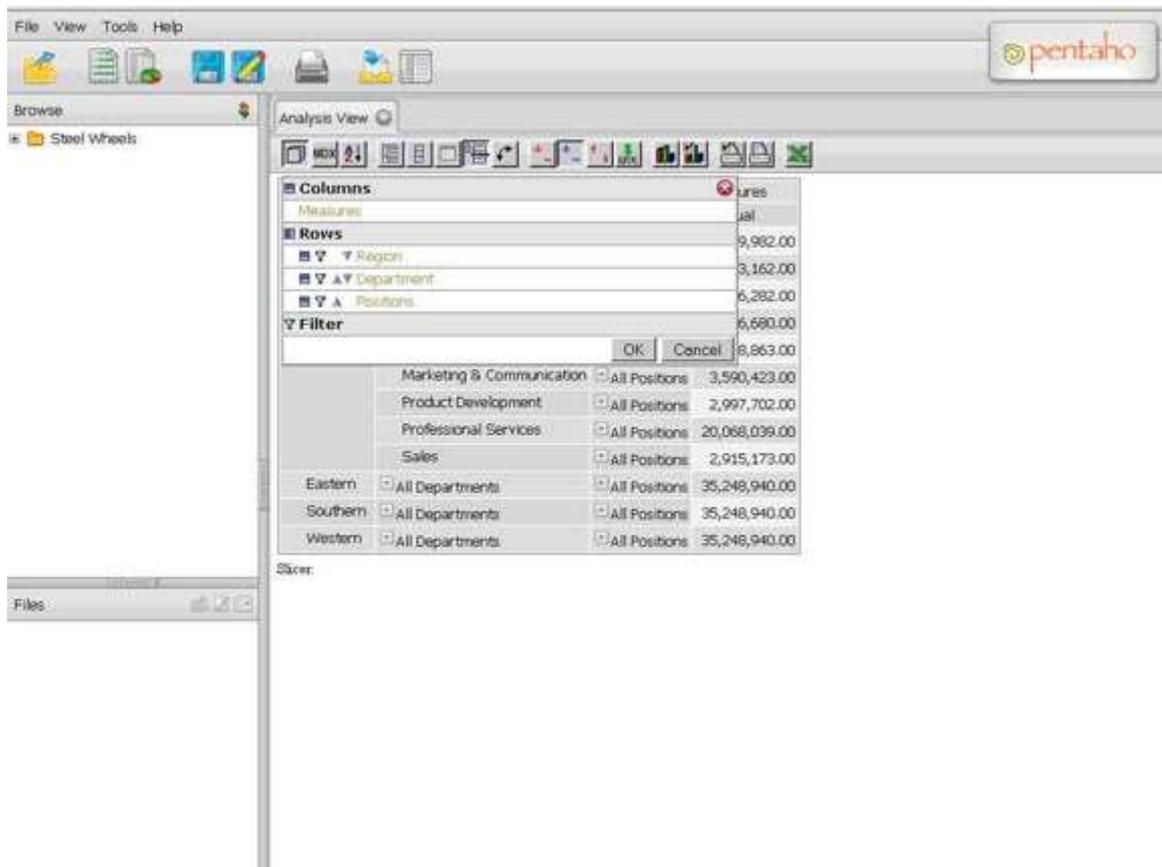


Region	Department	Positions	Measures
All Regions	All Departments	All Positions	143,639,982.00
Central	All Departments	All Positions	37,893,162.00
	Executive Management	All Positions	1,776,282.00
	Finance	All Positions	3,106,680.00
	Human Resource	All Positions	3,438,863.00
	Marketing & Communication	All Positions	3,590,423.00
	Product Development	All Positions	2,997,702.00
	Professional Services	All Positions	20,068,039.00
	Sales	All Positions	2,915,173.00
Eastern	All Departments	All Positions	35,248,940.00
Southern	All Departments	All Positions	35,248,940.00
Western	All Departments	All Positions	35,248,940.00

В этом окне при помощи символов "+" у соответствующих элементов можно сворачивать и разворачивать наборы данных, чтобы выводить только нужную информацию. По умолчанию выводится вся информация из набора данных без какой-либо фильтрации или изменения

порядка представления. Попробуем изменить порядок представления, выведя информацию о региональных отделах, которые превысили свой бюджет. Для этого необходимо перейти к мастеру OLAP-навигатор (рисунок 10), нажав на кнопку с изображением синего кубика (на крайней левой позиции панели инструментов вкладки Analysis View).

Мастер OLAP-навигатор



В окне этого мастера необходимо:

- перенести Region из области Rows в область Columns (для этого используется пиктограмма синего кубика слева от слова Region).
- перенести Measures из области Columns в область Filter (для этого используется пиктограмма воронки слева от слова Measures).
- сделать так, чтобы показывалось отклонение от намеченного бюджета (для этого необходимо щелкнуть левой кнопкой мыши по слову Measures, в появившемся диалоговом окне выбрать опцию Variance (отклонение) и нажать кнопку ОК).

После этого выпадающее окно мастера навигатора примет такой вид, как на рисунке 11 и, если нажать в нем кнопку ОК, то структура отчета будет изменена, как показано на рисунке 11.

Подготовка отчета для поиска региональных отделов с превышением бюджета

The screenshot shows the Pentaho Analysis View interface. The 'Columns' shelf contains 'Region' and the 'Filter' shelf contains 'Measures (Measure=Variance)'. The pivot table displays the following data:

	al	Eastern	Southern	Western		
Region						
Department	38.00	238,921.00	-445,079.00	-738,873.00		
Positions	51.00	350,130.00	-359,870.00	-541,664.00		
Measures (Measure=Variance)	71.00	-56,205.00	-38,205.00	-38,205.00		
	58.00	-16,518.00	-16,518.00	-28,518.00		
	29.00	-29,165.00	-29,165.00	-39,165.00		
Sales	All Positions	-195,361.00	-184,603.00	33,074.00	-16,926.00	-26,926.00
Product Development	All Positions	142,509.00	161,478.00	-18,323.00	-323.00	-323.00
Executive Management	All Positions	195,144.00	267,360.00	-24,072.00	15,928.00	-64,072.00

Из этого отчета можно сделать вывод, что бюджет отдела Executive management в регионах Eastern и Western был превышен на 24072 и 64072 соответственно. Этот интерактивный

аналитический отчет можно сохранить для последующего использования точно так же, как и простой отчет, как было показано ранее.