Assignment B6

Title
BAI tool for considering aspect of growth of the business.

Problem Statement
A Mall has number of items for sale. Build a required Database to develop BAI tool for considering one aspect of growth to the business such as organization of products based on demand and patterns use R Programming or other equivalent latest tools used in Industry or Use Hadoop, HDFS, HIVE, PIG, mongoBD Connectors for Hadoop and/OR other latest technology tools in the Hadoop Ecosystem for unstructured data analytics to effectively use advanced SQL functions and Greenplum extensions for in-database analytics. Use MADlib bigdata tools to solve analytics problems in-database used for this assignment.

Learning Objectives
1) To understand and use BAI tool.
2) To use R Programming or other equivalent latest tools for unstructured data analytics

Learning Outcome
1) Successfully use R Programming for unstructured data analytics

Software and Hardware Requirements
1) 64 bit machine with 64 bit operating system i.e. Fedora20
2) BI Publisher
3) R (R is available as a binary for many versions of Linux)

Theory
Business Intelligence:
Business intelligence (BI) is an umbrella term that combines architectures, tools, databases, analytical tools, applications, and methodologies. BI's major objective is to enable interactive access to data, to enable manipulation of data , and to give business managers and analysts the ability to conduct appropriate analysis. By analyzing historical and current data, situations, and performances, decision makers get valuable insights that enable them to make more informed and better decisions. The process of BI is based on the transformation of data to information, then to decisions, and finally to actions.

Business Analytics and Intelligence tool

Business intelligence tools are a type of application software designed to retrieve, analyze, transform and report data for business intelligence. The tools generally read data that have been previously stored, often, though not necessarily, in a data warehouse or data mart.

BI Publisher :

BI Publisher separates the creation of data from the process of formatting it for different uses. The engine can format any well-formed XML data, allowing integration with any system that can generate XML, including Web Services or any data source available through JDBC. BI Publisher can merge multiple data sources into a single output document.

Oracle BI Publisher is the reporting solution to author, manage, and deliver all your reports and documents easier and faster than traditional reporting tools. Use your web browser or familiar desktop tools to create everything from pixel-perfect customer facing documents to interactive management reports against practically any data source. View reports online or schedule them and deliver tens of thousands of documents per hour with minimal impact to transactional systems.

Template design:

BI Publisher report templates can be designed using the Microsoft Word, Adobe Acrobat, Microsoft Excel (standalone only) and Adobe Flash (standalone only). Templates created using these tools contain embedded fields with properties that determine how the XML data will be merged into the template, using Extensible Stylesheet Language syntax to precisely match the server's engine.

Template Builder for Word

Template Builder is an extension to Microsoft Word that simplifies the development of Rich Text Format templates. Templates created using Template Builder are transformed into XSL Stylesheets that can be used to generate PDF, RTF, Microsoft Excel and HTML outputs..

Layout Editor:

Starting with the 11g release, BI Publisher also offers a pure web based layout editor that allows users to create management reports and simple production reports in a WYSIWYG layout editor. The layout editor is written in pure DHTML. As with Rich Text Format templates Reports created in the web based layout editor are transformed into XSL stylesheets and can be viewed in the same output formats. In addition, the layout editor templates (.xpt) can also be viewed in an interactive viewer which allows re-sorting and interactive filtering of existing reports.

Adobe Acrobat:

XML Publisher templates can be designed in Adobe Acrobat 5.00 and above, using the native form field capabilities.

Adobe Flash:

The 10.1.3.3 release of Oracle BI Publisher offers support for Adobe Corporation's new document format for building interactive forms and reports, called Flex. You can build Flex templates, test them on your desktop, and deploy them to the BI Publisher server to generate

Flash output. Users are then able to run the reports from the BI Publisher user interface or schedule them for delivery to report consumers.

XSL Stylesheet:

In addition, to using the tools mentioned above users can also upload existing XSL stylesheets to run with BI Publisher.

R:

R is a language and environment for statistical computing and graphics. It is a GNU project which is similar to the S language and environment which was developed at Bell Laboratories (formerly AT&T, now Lucent Technologies) by John Chambers and colleagues. R can be considered as a different implementation of S. There are some important differences, but much code written for S runs unaltered under R.

R provides a wide variety of statistical (linear and nonlinear modelling, classical statistical tests, time-series analysis, classification, clustering, …) and graphical techniques, and is highly extensible. The S language is often the vehicle of choice for research in statistical methodology, and R provides an Open Source route to participation in that activity.

One of R's strengths is the ease with which well-designed publication-quality plots can be produced, including mathematical symbols and formulae where needed. Great care has been taken over the defaults for the minor design choices in graphics, but the user retains full control.

R is available as Free Software under the terms of the Free Software Foundation's GNU General Public License in source code form. It compiles and runs on a wide variety of UNIX platforms and similar systems (including FreeBSD and Linux), Windows and MacOS.

Mathematical Model

Let S be the system in consideration.

where:

S={s, e, X ,Y, DD, NDD, Fs, sc, fc}

s = start state of the system
{Ø}

e = end state of the system
result of analysis

X= set of inputs
= {x1}
x1= data entered in database
where x1 is a list of groceries

Y= set of outputs
{y1}

y1= result of analysis

DD= Deterministic data
{X}

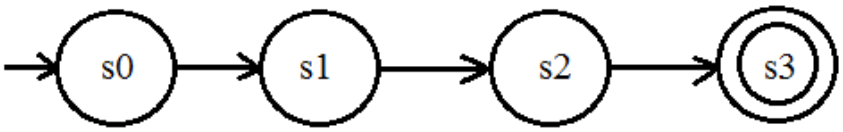
NDD= Non deterministic data
{Y}

Fs = set of functions
{f1,f2,f3,f4}
f1= insert data in database
f2= function to analyze database
f3= display result

sc = success case
result of analysis

fc= failure case
unable to read data from database.

State Diagram



s0= start state
s1= read from exel file
s2= calculate item frequency
s3= display result

Conclusion

In this way, we have studied how to use BI tools for considering various aspects of an organization and enhancing them.