

IBM Industry Customer Intelligence Solutions
Version 1.0.0

Lift Analytics for Retail Solution Guide

IBM

Note

Before using this information and the product it supports, read the information in "Notices" on page 59.

Product Information

This document applies to IBM Industry Customer Intelligence Solutions 1.0.0 and may also apply to subsequent releases.

Licensed Materials - Property of IBM

© **Copyright IBM Corporation 2015.**

US Government Users Restricted Rights – Use, duplication or disclosure restricted by GSA ADP Schedule Contract with IBM Corp.

Contents

Introduction	v
Chapter 1. Overview	1
Chapter 2. Installation and configuration.	3
Prerequisites	3
Contents of the download package	3
Overview of the installation and configuration process	3
Downloading the solution	4
Planning the installation and configuration of the solution	4
Modifying the sudoers file for the user who runs the installation	5
Computer firewall settings required for the installation	5
Disabling firewall settings during installation (optional)	5
Configuring your firewall settings on the installation node computer by using the Firewall script	6
Manually configuring firewall settings on the installation node	6
Configuring firewall settings on other computers that will contain content from the solution	7
The solution installer	7
Running the industry solution installer	8
Installing the InfoSphere Information Server software.	9
Creating the Retail Lift database.	9
Installing Netezza client software	9
Configuring the Netezza data source on your Information Server node	10
Installing the InfoSphere Information Server server-side content.	10
Creating a DataStage project.	11
Importing the InfoSphere Information Server package	11
Deploying the IBM Cognos content	11
Preparing the IBM Cognos content	12
Adding the Netezza library to your IBM Cognos BI installation	12
Creating a data source connection to the Lift Analytics database.	13
Deploying the IBM Cognos reports	13
Deploying Cognos images	14
Setting up the web page	14
Chapter 3. Lift Analytics for Retail DataStage jobs	15
Working with DataStage jobs	15
DataStage jobs	15
TLOG ETL process	19
Aggregation fact table ETL process	20
Chapter 4. Using the Lift Analytics for Retail solution	27
Choosing a Lift Analytics for Retail Report	27
Product Performance Trend report.	28
Customer Segment Performance report	29
Product Performance Summary report	29
Affinity Performance report	29
Affinity Relationship report	30
Appendix A. Accessibility features.	31
Appendix B. Troubleshooting	33
Troubleshooting a problem	33
Troubleshooting resources	33
Performance issues because Netezza database needs rebalancing	35
Performance is impacted after large amounts of data are deleted	35

Performance is impacted after large amounts of data have been inserted	35
After running the solution installer, Chef processes are still running after executing the cleanup.sh script	36
Appendix C. Importing data specifications for IBM Lift Analytics for Retail	37
Appendix D. Data dictionary	57
Notices	59
Index	63

Introduction

The IBM Lift Analytics for Retail solution gives you the information and insight that you need to provide proactive service to your customers.

Audience

This guide is intended for administrators and users of the Lift Analytics for Retail solution. It provides information on installation, configuration, and use of the Lift Analytics for Retail solution.

Finding information

To find product documentation on the web, including all translated documentation, access IBM® Knowledge Center (<http://www.ibm.com/support/knowledgecenter>).

Accessibility features

Accessibility features help users who have a physical disability, such as restricted mobility or limited vision, to use information technology products. Some of the components included in the Lift Analytics for Retail have accessibility features. For more information, see Appendix A, “Accessibility features,” on page 31.

Lift Analytics for Retail HTML documentation has accessibility features. PDF documents are supplemental and, as such, include no added accessibility features.

Forward-looking statements

This documentation describes the current functionality of the product. References to items that are not currently available may be included. No implication of any future availability should be inferred. Any such references are not a commitment, promise, or legal obligation to deliver any material, code, or functionality. The development, release, and timing of features or functionality remain at the sole discretion of IBM.

Samples disclaimer

Sample files may contain fictional data manually or machine generated, factual data compiled from academic or public sources, or data used with permission of the copyright holder, for use as sample data to develop sample applications. Product names referenced may be the trademarks of their respective owners. Unauthorized duplication is prohibited.

Chapter 1. Overview

The IBM Lift Analytics for Retail solution is a prepackaged solution that works with IBM Predictive Customer Insight.

IBM Lift Analytics for Retail V1.0 provides retail merchants and marketers with analytic insights into the business dynamics of selling patterns. It helps them to make quicker and better decisions about merchandising, assortment, and promotions. These decisions help to increase incremental sales and profits known as "lift".

The IBM Lift Analytics for Retail solution provides retailers with insights to make informed decisions about what products to carry, promote, and feature. It provides prescriptive and predictive merchandise management dashboards that put advanced analytics directly in the hands of merchants, planners, and marketers. IBM Lift Analytics for Retail provides insight into the potential impact of merchandising decisions without the need for data scientists or other analytics experts.

The Lift Analytics for Retail solution extends descriptive business analytics by adding affinity analysis, predictive analysis, and segmentation. It enables merchants and marketers to increase their understanding of selling affinities and patterns across multiple dimensions (for example, time, customer segment, channel, and location) to inform assortment, placement, pricing, and promotional decisions that can drive incremental sales and profits.

What need does the solution address?

When retailers make changes to the items that they carry or promote, it is difficult for them to predict the effect of their decisions on sales. For example, if a grocery store decides to sell videos, then this might result in an increase in popcorn sales. Or if they have a sale on a particular soft drink product, it might result in an increase in potato chip sales. The effect on sales of one product on other products is complex and difficult to predict.

Retailers have a great deal of data related to transactions, but the analysis of the data requires time-consuming and expensive analysis by experts.

Capabilities of the solution

The solution enables merchants, planners, and marketers to easily quantify the incremental cross shopping impact of products. Easy to understand decision-focused dashboards are at merchants' fingertips to guide day to day decisions. Flexible structure for inquiry enables insights for individual or groups of products across categories, channels, and customer segments. Companies can avoid making costly mistakes if they have an accurate prediction of the potential impact of changes in their merchandising programs.

A summary of the benefits

The Lift Analytics for Retail solution provides the following potential benefits:

- Increased incremental sales

- Increased profitability
- Increased average customer sales
- Improved inventory productivity
- Improved effectiveness of product promotions and assortment
- Descriptive and predictive analytics
- Easily accessible business intelligence (BI) trend reporting
- On-demand affinity analytics both point-in-time and trend over time
- Access to analytics without the need to involve data scientists or other expert resources
- Line-of-business user-friendly interpretation of analytic results delivered within the context of a business user's job at time of decision analysis
- Ability to access insights at the time of need allowing more prescriptive and predictive decision making
- Ability to make informed decisions using predictive analytics to understand the potential impact of merchandise and marketing program changes
- Avoidance of costly mistakes by enabling retailers to understand likely impacts before execution
- Ability to find under-served intersections of product and customer segments and address the need to increase sales to that segment
- Identification of opportunities to enhance lift (sales and profitability) by exposing variances in customer behavior

Chapter 2. Installation and configuration

The IBM Lift Analytics for Retail solution is for use with IBM Predictive Customer Intelligence.

The installation and configuration of the solution involves going to Passport Advantage, downloading the solution, and then configuring the various parts of it within your environment.

Prerequisites

Before you install the Lift Analytics for Retail solution there are prerequisites that must be met.

The following conditions must be met before the application is installed:

- IBM Predictive Customer Intelligence is installed and configured.
- You have a Linux system that you can install to. The solution installer works only on Linux.
- The following ports are available: 8080, 443, 9683, and 22
- All node computers to be used by the solution are accessible from the Solution Installer computer.
- The web interface must be accessed from a computer that can access the installation node computer.
- PureData System for Analytics (formerly know as a Netezza appliance)
- InfoSphere DataStage server
- You have administration rights and the ability to copy files between computers.

Contents of the download package

The Lift Analytics for Retail package, which you download from Passport Advantage, contains several parts.

The following parts are included in the package:

- Solution installer
- Netezza SQL for creating the database
- InfoSphere Information Server package, for performing regular loads of data
- Cognos BI content
 - a repository export with reports and package
 - a zip file with the Framework Manager model
 - report images
- Web content - a zip file containing the home page for the solution

Overview of the installation and configuration process

This is an overview of the general tasks that you must perform to install and configure the Lift Analytics for Retail solution.

You will need to perform the following tasks:

1. On the server from which you want to install the solution, download the tar.gz file from IBM Passport Advantage.
2. Plan where you want the various parts of the solution to reside
3. Open the ports that are required for the installation. You can do this by running the firewall.sh script, or you can choose to manually open the ports on the installation node. You will also have to open ports on the other node computers.
4. If the DataStage server and client is not installed, you must install it.
5. On the installation computer, run the setup.sh file to install the solution installer.
6. Run the solution installer to deploy files to the correct locations.
7. Create the Netezza SQL database on the data node computer.
8. Create the Netezza data source in Cognos BI.
9. Import the DataStage package.
10. Import the Cognos deployment package.
11. Install report images.
12. Copy the FM model to a client location.
13. Unzip the web page.

Downloading the solution

You must download the Lift Analytics for Retail solution from IBM Passport Advantage.

Before you begin

You must have a Passport Advantage account and be entitled to the package.

About this task

Passport Advantage contains package repositories. Instead of downloading the package from Passport Advantage and creating a repository locally, you access the repository on Passport Advantage.

Procedure

1. Set the Passport Advantage preference to connect to Passport Advantage.
2. Download the tar.gz file.
3. Move the tar file to the Linux node from which you want to run the installation.
4. Uncompress the tar file.

Planning the installation and configuration of the solution

Before you install the solution, you must decide which computers you want the various parts of the solution to reside on.

You must ensure that no application is listening on the ports that are used during the installation of the solution.

Modifying the sudoers file for the user who runs the installation

To run the Lift Analytics for Retail installation, you must be either **root** user or have **sudo** permission on each node computer, and you must disable requiretty during the installation.

To install with **sudo** user permissions, you must also add the user to the sudoers file.

Procedure

1. Log in as **root** user.
2. Enter the following command to open the sudoers file for editing:
`visudo -f /etc/sudoers`
3. Locate the following line:
Defaults requiretty
4. Press the I key to insert text.
5. Type a number sign (#) in front of Defaults requiretty to comment out the line. For example, the line will appear as
#Defaults requiretty
6. If you run the installer as a user with **sudo** user permissions, go to the end of the file, and add the following line for your user:
`username ALL=(ALL) NOPASSWD: ALL`
7. Press the Esc key, and enter :wq to save and close the file.
8. Repeat these steps on each computer on which you install a Lift Analytics for Retail node.

Computer firewall settings required for the installation

The solutions installer requires bidirectional communication with each node when you install a server component or server artifact component, and your firewall can interfere with this communication.

To ensure this bidirectional communication, you must configure your firewall settings for the installation. Alternatively, you can temporarily disable your firewall during the installation.

A script is provided to open all necessary ports on the installation node. You can either use the firewall.sh script, open the nodes manually, or disable the firewall settings.

In addition to opening ports on the installation node computer, you must open some ports on each computer to which content from the solution will be copied.

Disabling firewall settings during installation (optional)

If you choose to temporarily disable your computer's firewall during the installation of the IBM Lift Analytics for Retail application, ensure that you enable the firewall on each node after the installation is complete

About this task

Important: If you disable your computer's firewall during the installation, you must still ensure that the required ports are opened on each node computer when you restart your firewall.

Procedure

1. Log on to the computer as the root user or as a user with sudo permissions.
2. Type the following command:
`/etc/init.d/iptables stop`
3. Repeat these steps on each computer on which you install a node and on the server from which you are running the installation.

What to do next

After the installation is complete, on each node, type `/etc/init.d/iptables start` to restart your firewall.

Configuring your firewall settings on the installation node computer by using the Firewall script

You can run the `firewall.sh` script to open the ports on the installation node.

About this task

The `firewall.sh` script opens the following ports on the solution installer node:

- 8080 incoming
- 443 incoming
- 9683 incoming
- 22 outgoing

Procedure

1. Log on to the computer that contains the solution installer node as the root user or as a user with sudo permissions.
2. Back up your existing firewall settings by typing the following command:
`/etc/init.d/iptables save.`
3. Run the `firewall.sh` script.

Manually configuring firewall settings on the installation node

If you prefer not to use the `firewall.sh` script, you can manually configure your firewall settings on the installation node.

Procedure

1. As the root user or as a user with sudo permissions, log on to the computer that contains the solution installer node .
2. Back up your existing firewall settings by typing the following command: .
`/etc/init.d/iptables save`
3. On the computer that contains the solution installer node, add rules for the ports that are required by the installer by typing the following commands:

```
iptables -I INPUT -p tcp --dport 8080 -j ACCEPT
iptables -I OUTPUT -p tcp --sport 8080 -m state --state RELATED,ESTABLISHED -j ACCEPT

iptables -I INPUT -p tcp --dport 443 -j ACCEPT
iptables -I OUTPUT -p tcp --sport 443 -m state --state RELATED,ESTABLISHED -j ACCEPT

iptables -I INPUT -p tcp --dport 9683 -j ACCEPT
iptables -I OUTPUT -p tcp --sport 9683 -m state --state RELATED,ESTABLISHED -j ACCEPT
```

```
iptables -I OUTPUT -p tcp --dport 22 -j ACCEPT
iptables -I INPUT -p tcp --sport 22 -m state --state
RELATED,ESTABLISHED -j ACCEPT
```

Configuring firewall settings on other computers that will contain content from the solution

On all node computers to which content from the application will be copied, the ports 443, 9683, and 8080 must be opened for the installation process.

Procedure

1. On each of the computers to which content for the solution will be copied, add rules for the ports that are required by the installer by typing the following commands:

```
iptables -I OUTPUT -p tcp --dport 443 -j ACCEPT
iptables -I INPUT -p tcp --sport 443 -m state --state RELATED,ESTABLISHED -j ACCEPT
iptables -I OUTPUT -p tcp --dport 9683 -j ACCEPT
iptables -I INPUT -p tcp --sport 9683 -m state --state RELATED,ESTABLISHED -j ACCEPT
iptables -I OUTPUT -p tcp --dport 8080 -j ACCEPT
iptables -I INPUT -p tcp --sport 8080 -m state --state RELATED,ESTABLISHED -j ACCEPT
```
2. Type the following command to save your firewall settings:

```
/etc/init.d/iptables save
```
3. Type the following command to restart your firewall:

```
/etc/init.d/iptables restart
```
4. Verify that your changes are applied by using the following command:

```
iptables -L
```

The added port appears in the list.
5. Repeat these steps on each computer on which you will install an IBM Lift Analytics for Retail node.

The solution installer

The solution installer provides automation that will help you to install the solution.

The Industry Solution Installer runs on the latest versions of Chrome and Firefox. It is not tested on Safari and does not run on Internet Explorer.

The Industry Solution provides the following functionality:

- Deploys the required files to defined nodes.
- You can create a new configuration or start from a configuration that you previously saved.
- Automated validation.
- Electronic license acceptance.
- Automated clean-up of installed files if the license is declined.

The setup.sh script installs nodejs, chef client, and chef server. It extracts and uploads the cookbooks and sets up everything so that chef can be used to install the software.

The setup.sh script takes the following five arguments (the fifth argument is a password, and if omitted, it will be prompted for):

```
sh setup.sh <username> <first name> <last name> <e-mail> <password>
```

Where

username: The user name to create in chef server (can be used to log on to the chef server) Must be all lowercase characters.

first name: The first name of the user that is created in chef (for display purposes only)

last name: The last name of the user that is created in chef (for display purposes only)

e-mail: A valid email address for notifications from chef server

password: A password for the user that is created in chef (this is optional: a password will be prompted for if it is not included on the command line).

Running the industry solution installer

You can now run the Industry Solution Installer to deploy files to the correct nodes. Nodes are the target computers on which you deploy the software.

Before you begin

Ensure that you know the following information for each of the nodes that are used in your solution and that you open the necessary ports.

- Host name or IP address
- Password of the root user ID or a user ID and password that has sudo permissions

About this task

In this task, you create an installation configuration for your solution by dragging nodes onto a canvas and then dragging components into the nodes.

Important: The Industry Solution Installer runs on the latest versions of Chrome and Firefox. It does not run on Internet Explorer and it is untested on Safari.

Procedure

1. Run the setup shell script `setup.sh`.
2. After the solution installer finishes, a browser window opens. If a browser is not installed, a URL is displayed from which you can connect to the UI. The default URL is `https://solution_installer_node/UI/Index.html`.
3. Click **New Configuration**
If you have a configuration that was previously saved, you can start from that saved configuration.
4. From the Mandatory Content List pane, select **Node**.
A visual representation of an empty node is displayed on the canvas.
5. In the Property Editor, type a name for the node, and in the **Host Name** box, specify the server on which the node will be placed.
For example, you might create a node that is named BI Node.
6. Enter a user name and password for a user that has sudo access.

If you click **Validate** before all the mandatory content is added, a message indicates that the node is not valid. For example, if you validate the BI Node before adding Cognos BI Content, then a message indicates that the node is not yet valid.

7. From the **Optional Content** list, drag any content that you want to add into the appropriate node. The reason that this content is optional is because you might already have some of the necessary components installed in your environment.
8. For every node that you want to add content to, define the node and then drag content objects onto it.
9. Click **Validate**.
10. After all nodes are defined, the contents are specified, and each node is validated, click **Run**.

Installing the InfoSphere Information Server software

If the server and client software for IBM InfoSphere Information Server is not installed and configured, you must install it from the location that you specified when you ran the solution installer.

Procedure

For instructions on installing and configuring IBM InfoSphere Server, refer to the documentation for IBM InfoSphere Information Server Version 9.1.2 (ibm.com/support/knowledgecenter/SSZJPZ_9.1.0/com.ibm.swg.im.iis.productization.iisinfsv.home.doc/topics/ic_homepage_IS.html?lang=en)

Creating the Retail Lift database

A script is provided for creating the database on the PureData System for Analytics.

Procedure

1. Login as Netezza user `nz`.
2. Change the directory to the folder where the installer deployed the database scripts, for example, `/opt/IBM/IS_Retail_LiftAnalytics_1.0`.
3. Use the following command to create the database `RTLIFT`:

```
nzsql -f Retail_LiftAnalytics_DB.sql
```

Installing Netezza client software

The Lift Analytics for Retail solution uses a Netezza database. Therefore, the Netezza client software must be installed on the nodes for Cognos BI and InfoSphere Information Server.

Procedure

1. Download the IBM Netezza Client Components V7.2 for Linux from the IBM Netezza Developer Network (https://www-01.ibm.com/marketing/iwm/iwm/web/reg/pick.do?source=swg-im-ibmndn&lang=en_US).
2. Extract the `nz-linuxclient-v7.2.0.0.tar.gz` installation package to a directory where everyone has read and execute access, for example, `tar -zxvpf nz-linuxclient-v7.2.0.0.tar.gz`

3. Change the directory to linux, and run the `./unpack` command and accept the defaults. Don't install the linux64 part because it only contains the 64-bit Netezza client software.
4. Return to the parent directory, and extract the `datadirect.package.tar.gz`
5. Go to the `linux/x86/32bit` directory and move the `linux32_x86_dd_dm.tar` file to the `/opt` folder, for example, `mv linux32_x86_dd_dm.tar /opt`
6. Change the directory to the `/opt` directory, and extract the `linux32_x86_dd_dm.tar` file. i.e., `tar -xvf linux32_x86_dd_dm.tar`. This will create the `/opt/odbc32v51` folder.
7. Go to the `odbc32v51` directory and edit the `odbc.ini` file using `vi` and add sections for your databases.
8. Change the directory to `/usr/local/nz/bin`, and run the `./nzds -host nzhost -u user -pw password` command, to test the installation.

Configuring the Netezza data source on your Information Server node

After installing the Netezza client software you must configure the data source.

Procedure

1. Log on as `dsadm`.
2. Change the directory to `/opt/IBM/InformationServer/Server/DSEngine`.
3. Copy the contents of the `/usr/local/nz/lib64/odbcinst.ini.sample` file to the `.odbcinst.ini` file.
4. Rename the existing `.odbc.ini` file to `.odbc.ini.bak`.
5. Copy the contents of the `/usr/local/nz/lib64/odbc.ini.sample` file to the `.odbc.ini` file.
6. Edit the `.odbc.ini` file to configure the data source to access your IBM PureData System for Analytics (Netezza appliance).
7. Save and close the file.
8. Create a symbolic link, `ln -s .odbc.ini odbc.ini`.
9. Add `/usr/local/nz/lib64` to the `LD_LIBRARY_PATH` variable in `dsenv`.
10. Add the variable `NZ_ODBC_INI_PATH=directory containing the odbc.ini file` in `dsenv`, for example, `NZ_ODBC_INI_PATH=/opt/IBM/InformationServer/Server/DSEngine`.
11. Restart the server engine and the ASB Agent. An ASB agent is a Java based background process that runs on all machines that have an IBM Information Server engine installed. When there is a request for service that needs processing by the IBM Information Server component, an ASB agent receives and conveys the request. For more information, see *Administrating IBM InfoSphere Information Server and IBM WebSphere Application Server services*.
12. Specify the data source name in the `uvodbc.config` file.
For example:
`<DSN_name>`
`DBMSTYPE = ODBC`
13. Save and close the file.

Installing the InfoSphere Information Server server-side content

You must install content that will be used by InfoSphere Information Server.

Procedure

1. Decompress the IBM InfoSphere Information Server Manager package, Retail_LiftAnalytics_IISvrContent.zip, on the Information Server node. A Retail_LiftAnalytics directory is created.
2. The Retail_LiftAnalytics directory and its content should be owned by dsadm user and dstage group. You can change the ownership of a file or directory with the following command:

```
chown -R dsadm:dstage Retail_LiftAnalytics
```
3. The DataStage job reads, writes, and moves files in these directories, so the dstage group should have read, write, and execute permissions.

```
find /opt/IBM/Retail_LiftAnalytics -type d -exec chmod 774 {} +
```

Creating a DataStage project

You must create a DataStage project.

Procedure

1. Invoke the IBM InfoSphere DataStage and QualityStage Administrator.
2. Open the **Projects** tab, and click **Add**.
3. Enter a project name, for example, LIFTANALYTICS, and click **OK**.
4. Select the new project and click **Properties**.
5. Open the **Permissions** tab.
6. Add users and groups. Ensure users and groups have sufficient roles, for example, DataStage and QualityStage Developer, and click **OK**.
7. Click **Close**.

Importing the InfoSphere Information Server package

You must move and deploy the DataStage package.

Procedure

1. Move the package, Retail_LiftAnalytics_IISClContent.pkg, to the DataStage client computer.
2. Invoke IBM InfoSphere Information Server Manager.
3. Add a new domain connection. Click **Add Domain** and enter the services tier and port in the **Domain** field, <hostname>:9080
4. Right-click the domain, and select **Deploy Package From > File**.
5. Navigate to the Retail_LiftAnalytics_IISClContent.pkg file, and click **Open**.
6. Click **Deploy**, and select an empty project to deploy into, for example, LIFTANALYTICS, and click **OK**.
7. Click **OK**.
8. Click **File > Exit**.
9. To view the deployed content, invoke IBM InfoSphere DataStage and Quality Stage Designer.

Deploying the IBM Cognos content

For IBM Cognos Business Intelligence, after you install the Netezza client, you'll define a 32-bit ODBC data source, create a data source connection in Cognos, and then deploy the content files for the the IBM Lift Analytics for Retail solution.

Preparing the IBM Cognos content

You must move the IBM Cognos® Framework Manager content for the IBM Lift Analytics for Retail solution to a Windows client computer.

The Solution installer only supports copying content to Linux nodes. If you are modifying the Framework Manager model, it must be copied to a Windows computer. Instructions for doing this by using the browser's download capabilities are included in the SolutionInstaller > NodeRoot > Downloads > IS_<Solution_Name>_1.0 folder, in a file called IS_<SolutionName>_1.0_ClientDeployment Files.txt.

Procedure

1. Copy the IBM Cognos Framework Manager model file from the computer where you deployed the application to the computer where you installed IBM Cognos Framework Manager, and decompress the file.
The Framework Manager file is named `\BI\Retail_LiftAnalytics_FMPProjects.zip`.
2. Create a 32-bit ODBC System data source using the `odbcad32.exe` file found in `C:\Windows\SysWOW64` folder.

Adding the Netezza library to your IBM Cognos BI installation

To complete the installation of the server components, you must configure properties in IBM WebSphere Application Server for IBM Cognos Business Intelligence.

Procedure

1. In a web browser, type the address for your WebSphere Admin Console.
The Admin console address is: `http://BI-Node:9060/ibm/console`, where BI-Node is the name or IP address for your Cognos BI server computer.
2. Leave the **User ID** box empty, and click **Log in**.
3. Expand **Servers > Server Types > WebSphere application servers**, and click `server1`.
4. Under **Server Infrastructure**, expand **Java and Process Management**, and click **Process Definition**.
5. Under Additional Properties, click **Environment Entries**.
6. Click `LD_LIBRARY_PATH`.
7. In the **Value** box, add the following value to the path:
`/opt/odbc32v51/lib:/usr/local/nz/lib/`
This path points to the Netezza 32-bit library file.
For example, the **Value** box, should include the following path:
`/opt/ibm/cognos/c10_64/bin64:/opt/ibm/db2/V10.1/lib32:/opt/odbc32v51/lib:/usr/local/nz/lib/`
8. Click **Apply**.
9. In **Environment Entries**, click **New**.
 - a. In the **Name** box, enter `ODBCINI`.
 - b. In the **Value** box, enter `/opt/odbc32v51/odbc.ini`.
 - c. Click **OK**.
10. In **Environment Entries**, click **New**.
 - a. In the **Name** box, enter `NZ_ODBC_INI_PATH`.

- b. In the **Value** box, enter /opt/odbc32v51.
 - c. Click **OK**.
11. Click **Save**.

Creating a data source connection to the Lift Analytics database

You must create a data source connection to the Lift Analytics for Retail database.

Procedure

1. Open a web browser.
2. Go to the IBM Cognos BI portal URL.
For example, go to `http://bi_node_name/ibmcognos/`.
3. On the **Welcome** page, click **Administer IBM Cognos Content**.
4. Click the **Configuration** tab, and click **Data Source Connections**.
5. Click the **New Data Source** button .
6. In the **Name** box, type `RETAIL_LIFT_NETEZZA`, and then click **Next**.
7. From the **Type** list, select **Netezza (ODBC)**, and click **Next**.
8. Enter the database name, `RTLIFT` and Netezza user admin, and click **Next**
9. Test the connection, and if the test is successful, click **Finish**.

Deploying the IBM Cognos reports

You must deploy the Lift Analytics for Retail reports using IBM Cognos Administration.

Procedure

1. Open a web browser.
2. Go to the IBM Cognos BI portal URL.
3. On the **Welcome** page, click **Administer IBM Cognos Content**.
4. On the **Configuration** tab, click **Content Administration**.
5. On the toolbar, click the **New Import** button.
6. In the **Deployment Archive** pane, select the `Retail_LiftAnalytics_BIContent.zip` file, and click **Next**.
7. In the **Specify a name and description** pane, accept the default or enter a new name, and click **Next**.
8. In the **Select the public folders and directory content** pane, select all of the packages in the table, leave the **Options** as default, and click **Next**.
Ensure that the **Disable after import** option is cleared.
9. In the **Specify the general options** pane, accept the defaults, and click **Next**.
10. On the **Review the summary** page, click **Next**.
11. On the **Select an action** page, select **Save and run once**, and click **Finish**.
12. On the **Run with options** page, accept the defaults, click **Run**, and then click **OK**.
13. Select **View the details of this import after closing this dialog** and click **OK**.
14. In IBM Cognos Administration, click the **Home** button.
The content is available in **Public Folders > Lift Analytics for Retail 1.0**.

Deploying Cognos images

You must deploy the Cognos images.

Procedure

1. Copy the IBM Cognos report image file to the Cognos_Install_location/webcontent/ folder on the Business Intelligence node computer.
The report images file is named Retail_LiftAnalytics_Images.zip
2. Decompress the file using the following command:
unzip Retail_LiftAnalytics_Images.zip.
You should have a Cognos_Install_location/webcontent/IS_Images/Retail_LiftAnalytics folder that contains the report image files.

Setting up the web page

The solution includes a web page that is used as the home page for the solution.

Procedure

1. Move the Retail_LiftAnalytics_Web.zip file to the BI server in the Cognos webcontent folder.
2. Unzip the web content file .
3. Set up a link to the reports with your BI server names. Go to *BI Install folder/webcontent/RetailLift/config*.
4. Edit the report_url_config.js file to replace localhost with your BI server name.
5. Launch the web page by going to <http://servername/ibmcognos/RetailLift/> in a web browser.

Chapter 3. Lift Analytics for Retail DataStage jobs

The IBM Lift Analytics for Retail solution uses IBM InfoSphere DataStage to extract, transform, and load retail transactions for reporting.

Working with DataStage jobs

Before running the DataStage jobs, you must configure your DataStage jobs for your specific environment.

Updating the Parameter Sets

You must update the Parameter Sets to match your environment by completing the following steps:

1. In InfoSphere DataStage Designer, navigate to **Jobs\Common**, and double-click the psDBParams parameter set.
2. Open the **Parameters** tab, and update the values to match your target database connection.
3. Click **OK**.
4. Double-click the psFolders parameter set.
5. Open the **Parameters** tab, and update the values to match your DataStage Server directory structure. Root_Path is the parent directory. Pathnames must end with a directory delimiter.
6. Click **OK**.
7. Navigate to Jobs\TLOG\common, and double click the psTLog parameter set.
8. Open the **Parameters** tab, and update the values to match your environment.
9. Click **OK**.

Compiling jobs

To retain the updated default values, you must recompile all the DataStage jobs by doing the following steps:

1. From the InfoSphere DataStage Designer menu bar, click **Tools > Multiple Job Compile**.
2. In the Selection Criteria, click **Next**
3. In Selection Override, Add the Jobs folder to the selected items, and click **Next**.
4. In the Compiler Options, click **Next**.
5. Click **Start Compile**.
6. When the compilation finishes, click **Finish**.

DataStage jobs

This section describes the DataStage jobs that you will use to load the Lift Analytics for Retail reporting data.

Calendar

Job Calendars

Input None

Output

CI_DAY_D

CI_WEEK_D

Runtime parameters

StartDate: Start Date of the CI_DAY_D and CI_WEEK_D tables

EndDate: End Date of the CI_DAY_D and CI_WEEK_D tables

Job Dependencies

None

Frequency

Once, until the End Date.

CI_DAY_D and CI_WEEK_D are truncated before being loaded.

Segment Type

Job Segment_Type

Input segment_type.csv

Output

CI_SEGMENT_TYPE

Runtime Parameters

None

Job Dependencies

None

Frequency

As often as required.

New segment types are appended, and existing segment types are updated.

Segment

Job Segment

Input segment.csv

Output

CI_SEGMENT_D

Runtime Parameters

None

Job Dependencies

Segment_Type

Frequency

As often as required.

New segments are appended, and existing segments are updated.

Location

Job Location

Input location.csv

Output
FND_LOCATION

Runtime Parameters
None

Job Dependencies
None

Frequency
As often as required.
New locations are appended, and existing locations are updated.

Product

Job Product

Input product.csv

Output
CI_PRODUCT_LVLVD

Runtime Parameters
None

Job Dependencies
None

Frequency
As often as required.
New products are appended, and existing products are updated.

Customer Segment Type

Job Customer_Segment_Type

Input customer_segment_type.csv

Output
CUSTOMER_SEGMENT_MAP

Runtime Parameters
None

Job Dependencies
Calendars, Segment

Frequency
At least once per week.
New customer segment maps are appended, and existing customer segment maps are updated.

TLOG

Job TLog

Input [customer name]-tlog-YYYYMMDD.csv

Output

BI_F (used as input for the aggregation tables)

BI_R_F (stores returns but is not used)

Runtime parameters

Trailing TLog Window: Maximum number of weeks to accept trailing transactions before loading into the BI_F and BI_R_F tables.

TLog_Path: Directory that contains the input TLog files

Duplicate_Txn_Path: Directory to contain a log of rejected transactions when TXN_ID and UPC occur more than once.

Job Dependencies

Calendars

Location

Product

Frequency

Daily recommended.

Multiple tlog files can be processed in a single load.

Aggregation tables

Job Product_Affinity_Market_Basket

Input BI_F

Output

CI_MB_PROD_LVLA_WK_AGG

CI_MB_PROD_LVLB_WK_AGG

CI_MB_PROD_LVLC_WK_AGG

CI_MB_PROD_LVL_D_WK_AGG

CI_AFF_LVLA_WK_AGG

CI_AFF_LVL_B_WK_AGG

CI_AFF_LVLC_WK_AGG

CI_AFF_LVL_D_WK_AGG

CI_AFF_LVLA_YR_AGG

CI_AFF_LVL_B_YR_AGG

CI_AFF_LVLC_YR_AGG

CI_AFF_LVL_D_YR_AGG

Runtime parameters

StartDate, EndDate: Range of the transaction dates to extract from BI_F

Job Dependencies

Calendars

Customer_Segment_Type
Product
TLog
Frequency
Weekly recommended

TLOG ETL process

This section describes the process by which the TLog DataStage job loads the tlog transactions into the BI_F table.

The TLog DataStage job processes tlog transactions and stores them in the BI_F and BI_R_F tables. The BI_F table is used to load the aggregation fact tables.

The main sequencer job, TLog, can process one or more tlog files in the TLog_Path directory. The directory must only contain tlog files that will be processed. Multiple tlog files are processed one at a time in ascending order of the filename. The tlog file naming convention must be [customer_name]-tlog-YYYYMMDD.txt.

The TLog sequencer job calls the InboundTLogMainSEQ sequencer job, which reads and processes the tlog file and stores valid sales transactions in the STG_TLOG_RAW table and valid returns in the STG_R_TLOG_RAW table in the Netezza database.

The InboundTLogMainSEQ sequencer job formats the incoming tlog transaction record. Any records that do not conform to the expected format are logged to a file located in the FormatError_Path directory. Well formatted transactions flow through a series of data validations, such as valid Product and Location, and transformed for consumption by the Aggregation Fact ETL process. Invalid transactions are logged to a file located in the Reject_Path directory. Invalid transaction should be investigated and reprocessed if necessary.

Data in the STG_TLOG_RAW and STG_R_TLOG_RAW tables is moved into the STG_BI_F and STG_BI_R_F tables respectively after all the tlog files are processed through the InboundTLogMainSEQ sequencer job. New records are appended, and existing records, based on TXN_ID and UPC value combination, are updated.

Duplicate TXN_ID and UPC transactions are logged to a file in the Duplicate_Txn_Path directory and the job is stopped. Duplicate transactions should be investigated and deleted before continuing.

Valid transactions are moved from the STG_BI_F and STG_BI_R_F tables to the BI_F and BI_R_F tables respectively when the transaction date is less than or equal to the trailing window cutoff date. The trailing window cutoff date is the file date less the TrailingTLogWindow. The default is 1 week.

An email can be sent to interested parties when an exception occurs by providing the mail server name and email addresses in the NOTN_SendMail notification sequence in the TLog sequencer job.

Aggregation fact table ETL process

The Retail Lift Framework Manager model contains a number of aggregation tables that hold the measures for the BI Reports. This document outlines the ETL process that is required to build these tables.

CI_MB_PROD_LVL_WK_AGG Fact Table

This set of tables is used in the ETL process and the BI Reports. They contain the Market Basket and Level measures. There is a table for each level in the Product hierarchy and in the default setup this is a total of four tables.

Table Name	Product Level
CI_MB_PROD_LVL_A_WK_AGG	Department: Top level of the Product Hierarchy
CI_MB_PROD_LVL_B_WK_AGG	Category: Second level of the Product Hierarchy
CI_MB_PROD_LVL_C_WK_AGG	Line: Third level of the Product Hierarchy
CI_MB_PROD_LVL_D_WK_AGG	Product: Lowest Level of the Product Hierarchy

The data for these tables is contained in four tables in the Retail Lift database. The four tables are:

Table Name	Source Data
BI_F	Transaction data provides the information about each retail transaction
CI_DAY_D	Day based calendar dimension provides the week identifier for the data being loaded during that week's update
CI_PRODUCT_LVL_D	Product Hierarchy data, provides the product level identifier for each of the four levels in the default Retail Lift reports
CUSTOMER_SEGMENT_MAP	Segment information for each customer provides the segment identifier for each customer in that week's transactions

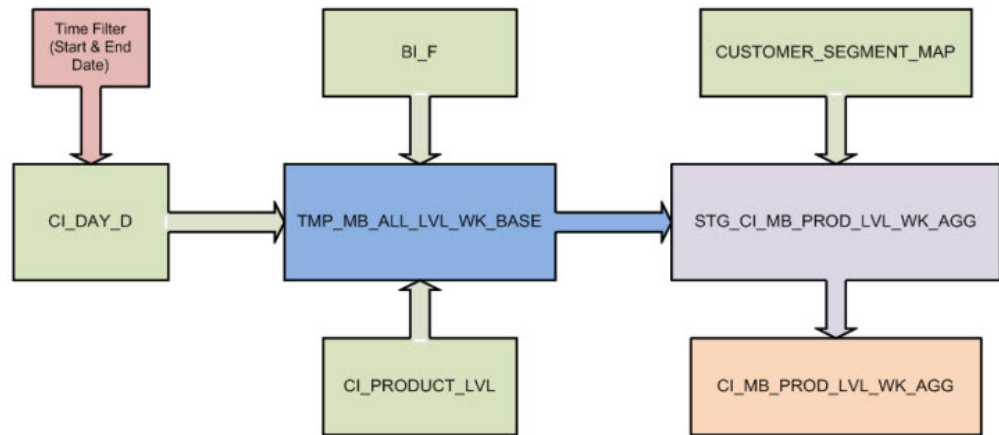
The DataStage® ETL runs through 3 steps for each of the **CI_MB_PROD_LVL_WK_AGG** tables.

The first step is that it loads the data from the **BI_F**, **CI_DAY_ID** and **CI_PRODUCT_LVL_D** table to get the transaction data for the time period defined by the Start and End date parameters. This data is loaded into the **TMP_MB_ALL_LVL_WK_BASE** table.

The second step combines the data from the **TMP_MB_ALL_LVL_WK_BASE** table with the data in the **CUSTOMER_SEGMENT_MAP** data. During this step the data is transformed and loaded into the **STG_CI_MB_PROD_LVL_WK_AGG** table. This process creates the measures listed in the following table:

Measure Name	Measure Definition	Tables
MBSALES	The sales amount aggregated at the Location and Segment levels. (Market Basket)	Level D Only
MBUNITS	The number of units sold aggregated at the Location and Segment levels. (Market Basket)	Level D Only
MBPROFITS	The profit amount aggregated at the Location and Segment levels. (Market Basket)	Level D Only
MBCNT	The count of transactions that contain the product aggregated at the Location and Segment levels. (Market Basket)	All Levels
LVL_SALES	The sales amount aggregated at the Product, Location and Segment levels. (Product Level)	All Levels
LVL_UNITS	The number of units sold aggregated at the Product, Location and Segment levels. (Product Level)	All Levels
LVL_PROFITS	The sales for the product aggregated at the Product, Location and Segment levels. (Product Level)	All Levels

The third step transfers the data from the staging table to the **CI_MB_PROD_LVL_WK_AGG** table for the appropriate Product Level. The ETL process flow for this set of tables is shown in the following figure.



CI_AFF_LVL_WK_AGG Tables

This set of tables is used in the ETL process, but not in the reports. These tables contain the From/To Market Basket measures. There is a table for each level in the Product hierarchy. In the default setup this represents a total of 4 tables.

Table Name	Product Level
CI_AFF_LVL_A_WK_AGG	Affinity Relationship for the Department level of the Product Hierarchy
CI_AFF_LVL_B_WK_AGG	Affinity Relationship for the Second level of the Product Hierarchy
CI_AFF_LVL_C_WK_AGG	Affinity Relationship for the Line level of the Product Hierarchy
CI_AFF_LVL_D_WK_AGG	Affinity Relationship for the Product level of the Product Hierarchy

The data for these tables is contained in four tables in the Retail Lift database. The four tables are:

Table Name	Source Data
BI_F	Transaction data provides the information about each retail transaction
CI_DAY_D	Day based calendar dimension provides the week identifier for the data being loaded during that week's update
CI_PRODUCT_LVL_D	Product Hierarchy data, provides the product level identifier for each of the four levels in the default Retail Lift reports
CUSTOMER_SEGMENT_MAP	Segment information for each customer provides the segment identifier for each customer in that week's transactions

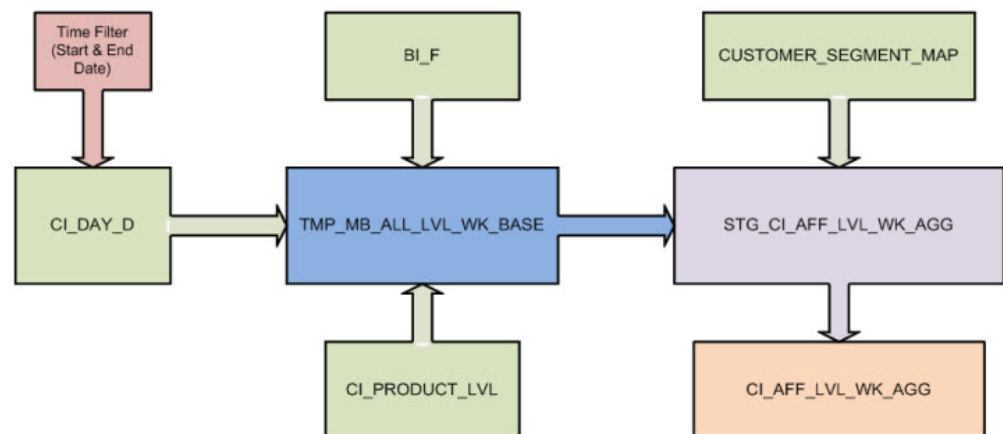
The DataStage ETL runs through three steps for each of the **CI_AFF_LVL_WK_AGG** tables.

The first step is it loads the data from the **BI_F**, **CI_DAY_ID** and **CI_PRODUCT_LVL_D** table to get the transaction data for the time period defined by the Start and End date parameters. This data is loaded into the **TMP_MB_ALL_LVL_WK_BASE** table. This table is used as the source of data for the second step of the ETL process.

The second step transforms the data from the **TMP_MB_ALL_LVL_WK_BASE** table to create the From/To relationship between all of the items in the same transaction. Each item in a transaction is first used on the From side of the affinity relationship. This item is then matched with the rest of the items in the same transaction on the To side. This is repeated until all of the items in the transaction appear on the From side once. The data is then loaded into the **STG_CI_AFF_LVL_WK_AGG** table for the appropriate Product Level. The measures that are created in this process are listed in the following table:

Measure Name	Measure Definition	Tables
FTMBSALES	The sales amount from the To side of the relationship when the From and To products are contained in the same transaction	All Levels
FTMBUNITS	The number of units sold from the To side of the relationship when the From and To products are contained in the same transaction	All Levels
FTMBPROFITS	The profit amount from the To side of the relationship when the From and To products are contained in the same transaction	All Levels
FTMBCNT	The distinct count of the transactions number from the To side of the relationship when the From and To products are contained in the same transaction	All Levels

The third step transfers the data from the staging table to the **CI_AFF_LVL_WK_AGG** table. The ETL process flow for this set of tables is shown in the following figure.



CI_AFF_LVL_YR_AGG Tables

This set of tables is used in the BI Reports. These tables contain the From/To Market Basket measures and the measures for the products on the From and To sides of the affinity relationship. There is a table for each level in the Product hierarchy. In the default product this represents a total of four tables.

Table Name	Product Level
CI_AFF_LVL_YR_AGG	Affinity Relationship for the Department level of the Product Hierarchy

Table Name	Product Level
CI_AFF_LVLB_YR_AGG	Affinity Relationship for the Second level of the Product Hierarchy
CI_AFF_LVL_C_YR_AGG	Affinity Relationship for the Line level of the Product Hierarchy
CI_AFF_LVL_D_YR_AGG	Affinity Relationship for the Product level of the Product Hierarchy

The data for these tables is contained in staging tables for the **CI_MB_PROD_LVL_WK_AGG** and **CI_AFF_LVL_WK_AGG** tables. The two staging tables are:

Table Name	Source Data
STG_CI_MB_PROD_LVL_WK_AGG	Level counts for the From and To side of the affinity relationship and the ALL_MBCNT for each Location and Segment
STG_CI_AFF_LVL_WK_AGG	The From/To Measures

The DataStage ETL runs through 3 steps for each of the **CI_AFF_LVL_YR_AGG** tables.

The first step takes the data from the **STG_CI_MB_PROD_LVL_WK_AGG** table and places it in the **STG_CI_AFF_LVL_YR_AGG_02** table.

The second step takes the data from the **STG_CI_AFF_LVL_YR_AGG_02** table and combines it with the data from the **STG_CI_AFF_LVL_WK_AGG** table. During this step the data from the **STG_CI_MB_PROD_LVL_YR_AGG_02** is used to create the measures that will represent the Sales, Number of Units, Profits and Number of Transactions for the From product and the To product in the affinity relationship. These measures are combined with the From/To measures in the **STG_CI_AFF_LVL_WK_AGG** table and placed in the **STG_CI_AFF_LVL_YR_AGG_03** table.

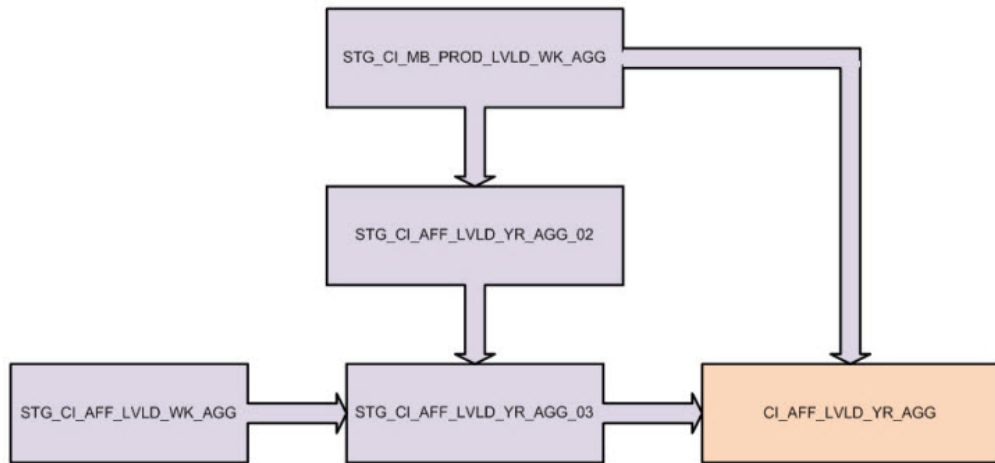
The third step takes the data from the **STG_CI_AFF_LVL_YR_AGG_03** table and combines it with the data from the **STG_CI_MB_PROD_LVL_WK_AGG**. The data from the **STG_CI_MB_PROD_LVL_WK_AGG** is aggregated at the Location and Segment level to create the ALL_MBCNT measure. This measure is added to the measures contained in the **STG_CI_MB_PROD_LVL_YR_AGG_03** table. This combined data is then loaded into the **CI_AFF_LVL_YR_AGG** table. The complete list of measures contained in the **CI_AFF_LVL_YR_AGG** table is shown in the following table.

Measure Name	Measure Definition	Tables
F_LVL_SALES	The sales amount for the product on the From side of the affinity relationship. This measure is aggregated at the Location and Segment levels	All Levels

Measure Name	Measure Definition	Tables
F_LVL_UNITS	The number of units sold for the product on the From side of the affinity relationship. This measure is aggregated at the Location and Segment levels	All Levels
F_LVL_PROFITS	The profit amount for the product on the From side of the affinity relationship. This measure is aggregated at the Location and Segment levels	All Levels
F_MBCNT	The distinct count of the transaction number for the product on the From side of the affinity relationship	All Levels
T_LVL_SALES	The sales amount for the product on the To side of the affinity relationship. This measure is aggregated at the Location and Segment levels	All Levels
T_LVL_UNITS	The number of units sold for the product on the To side of the affinity relationship. This measure is aggregated at the Location and Segment levels	All Levels
T_LVL_PROFITS	The profit amount for the product on the To side of the affinity relationship. This measure is aggregated at the Location and Segment levels	All Levels
T_MBCNT	The distinct count of the transaction number for the product on the To side of the affinity relationship	All Levels
FTMBSALES	The sales amount from the To side of the relationship when the From and To products are contained in the same transaction	All Levels
FTMBUNITS	The number of units sold from the To side of the relationship when the From and To products are contained in the same transaction	All Levels
FTMBPROFITS	The profit amount from the To side of the relationship when the From and To products are contained in the same transaction	All Levels

Measure Name	Measure Definition	Tables
FTMBCNT	The distinct count of the transactions number from the To side of the relationship when the From and To products are contained in the same transaction	All Levels
ALL_MBCNT	The number of transactions at the Location	All Levels

The ETL process flow for this set of tables is shown in the following figure.

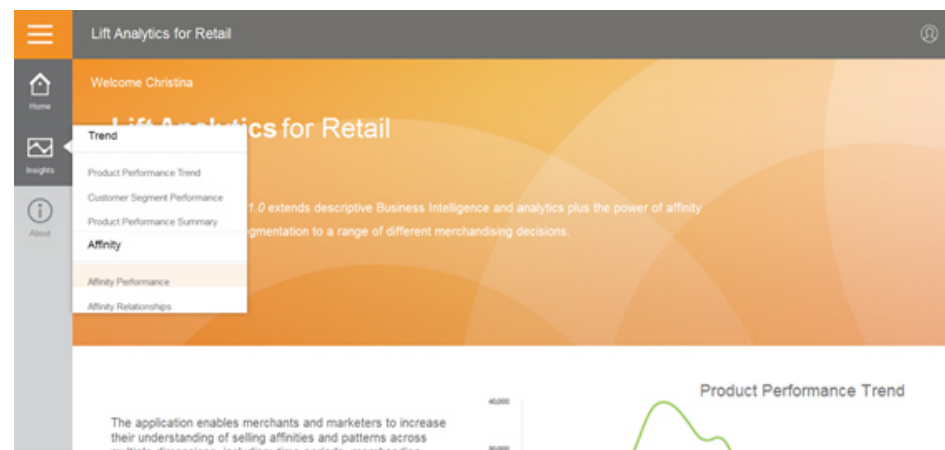


Chapter 4. Using the Lift Analytics for Retail solution

The IBM Lift Analytics for Retail solution provides quick access to reports that provide insight related to product performance trends and sales affinities.

The following analytics reports are available:

- Product Performance Trend
- Customer Segment Performance
- Product Performance Summary
- Affinity Performance
- Affinity Relationship

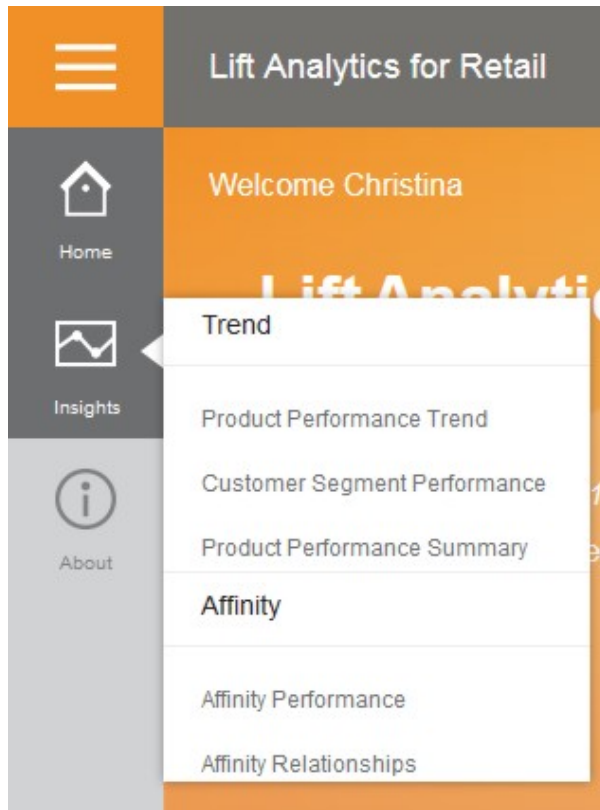


Choosing a Lift Analytics for Retail Report

From the Home page you can click the **Insights** button to view a trend or affinity report.

Procedure

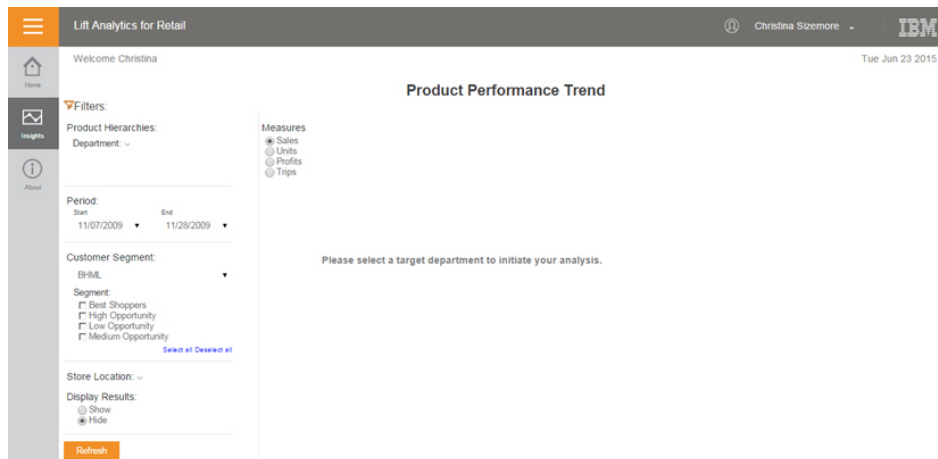
Click the Insights button to display a list of reports.



Product Performance Trend report

The Product Performance Trend report allows you to show sales performance over time for department, category, or product line.

It provides insight into trends in product performance. You can view trends in performance by sales, units, profits, and trips for selected departments, categories, products.



Filters allow you to display results for any level of the product hierarchy. You can select multiple levels. Click the refresh button each time you make a change in your selections.

You can select a time period and a customer segment. You can change the measure to display results for sales, units, profits, or trips. At any time, you have the option of switching between chart view and table view.

The Display Report section allows you to hide the results. This allows faster navigation of all the prompt hierarchies without instigating data calls.

Customer Segment Performance report

The Customer Segment Performance report allows you to analyze your data by RFM (Recency, Frequency, Monetary) segments.

It shows you how your sales break down in terms of your RFM high, medium, and low value customers.

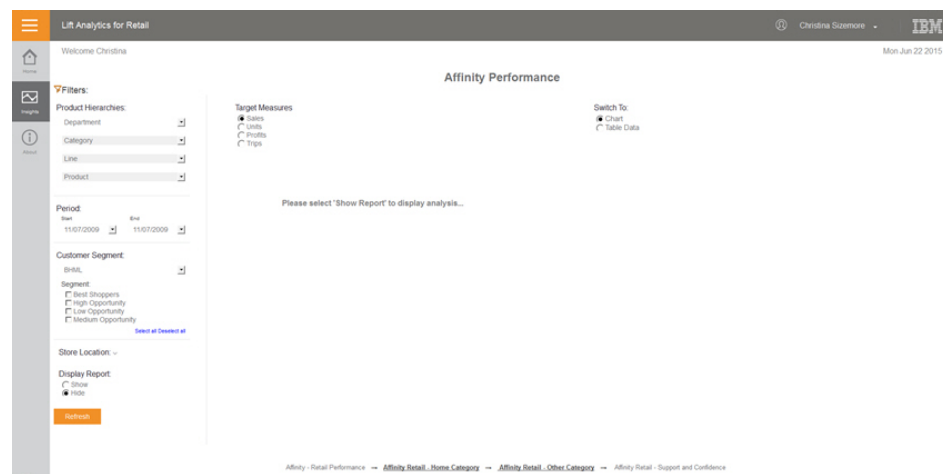
Product Performance Summary report

The Product Performance Summary report shows you how your sales are distributed amongst your departments, categories, and product lines.

You can display results in chart or table view.

Affinity Performance report

The Affinity Performance report helps you to understand the impact to your business resulting from such things as halo sales.



For example, an Affinity Performance report might show you that swim goggles account for a relatively small amount of revenue, but products, such as bathing suits, that have a strong affinity with swim goggles make up a large amount of revenue. The report may indicate that swim goggle sales are more important to the business than is indicated by their sales numbers .

Affinity Relationship report

The Affinity Relationship report indicates how closely sales of a category or product line are associated with other categories or product lines.

The report allows you to explore which consequent products are frequently purchased with the item selected from product hierarchy. It shows you the trend in performance across several possible metrics, including sales, units, profit, and number of trips for selected departments, categories, lines and products. The report also allows you to select the time periods and specific customer segments.

Appendix A. Accessibility features

Accessibility features help users who have a physical disability, such as restricted mobility or limited vision, to use information technology products.

For information about the commitment that IBM has to accessibility, see the IBM Accessibility Center (www.ibm.com/able).

IBM Cognos HTML documentation has accessibility features. PDF documents are supplemental and, as such, include no added accessibility features.

Report output

In IBM Cognos Administration, you can enable system-wide settings to create accessible report output. For more information, see the *IBM Cognos Business Intelligence Administration and Security Guide*. In IBM Cognos Report Studio, you can enable settings to create accessible output for individual reports. For more information, see the *IBM Cognos Report Studio User Guide*. You can access the previously mentioned documents at IBM Knowledge Center (<http://www.ibm.com/support/knowledgecenter>).

Appendix B. Troubleshooting

This section contains troubleshooting information for the Lift Analytics for Retail installation.

Troubleshooting a problem

Troubleshooting is a systematic approach to solving a problem. The goal of troubleshooting is to determine why something does not work as expected and how to resolve the problem.

Review the following table to help you or customer support resolve a problem.

Table 1. Troubleshooting actions and descriptions

Actions	Description
A product fix might be available to resolve your problem.	Apply all known fix packs, or service levels, or program temporary fixes (PTF).
Look up error messages by selecting the product from the IBM Support Portal, and then typing the error message code into the Search support box (http://www.ibm.com/support/entry/portal/).	Error messages give important information to help you identify the component that is causing the problem.
Reproduce the problem to ensure that it is not just a simple error.	If samples are available with the product, you might try to reproduce the problem by using the sample data.
Ensure that the installation successfully finished.	The installation location must contain the appropriate file structure and the file permissions. For example, if the product requires write access to log files, ensure that the directory has the correct permission.
Review all relevant documentation, including release notes, technotes, and proven practices documentation.	Search the IBM Knowledge Center to determine whether your problem is known, has a workaround, or if it is already resolved and documented.
Review recent changes in your computing environment.	Sometimes installing new software might cause compatibility issues.

If the items in the table did not guide you to a resolution, you might need to collect diagnostic data. This data is necessary for an IBM technical-support representative to effectively troubleshoot and assist you in resolving the problem. You can also collect diagnostic data and analyze it yourself.

Troubleshooting resources

Troubleshooting resources are sources of information that can help you resolve a problem that you are having with an IBM product.

Support Portal

The IBM Support Portal is a unified, centralized view of all technical support tools and information for all IBM systems, software, and services.

The IBM Support Portal lets you access all the IBM support resources from one place. You can tailor the pages to focus on the information and resources that you need for problem prevention and faster problem resolution. Familiarize yourself with the IBM Support Portal by viewing the demo videos (https://www.ibm.com/blogs/SPNA/entry/the_ibm_support_portal_videos).

Find the content that you need by selecting your products from the IBM Support Portal (<http://www.ibm.com/support/entry/portal>).

Before contacting IBM Support, you will need to collect diagnostic data (system information, symptoms, log files, traces, and so on) that is required to resolve a problem. Gathering this information will help to familiarize you with the troubleshooting process and save you time.

Service request

Service requests are also known as Problem Management Reports (PMRs). Several methods exist to submit diagnostic information to IBM Software Technical Support.

To open a PMR or to exchange information with technical support, view the IBM Software Support Exchanging information with Technical Support page (<http://www.ibm.com/software/support/exchangeinfo.html>).

Fix Central

Fix Central provides fixes and updates for your system's software, hardware, and operating system.

Use the pull-down menu to navigate to your product fixes on Fix Central (<http://www.ibm.com/systems/support/fixes/en/fixcentral/help/getstarted.html>). You may also want to view Fix Central help.

IBM developerWorks

IBM developerWorks[®] provides verified technical information in specific technology environments.

As a troubleshooting resource, developerWorks provides easy access to the most popular practices, in addition to videos and other information: developerWorks (<http://www.ibm.com/developerworks>).

IBM Redbooks

IBM Redbooks[®] are developed and published by the IBM International Technical Support Organization, the ITSO.

IBM Redbooks (<http://www.redbooks.ibm.com>) provide in-depth guidance about such topics as installation and configuration and solution implementation.

Software support and RSS feeds

IBM Software Support RSS feeds are a quick, easy, and lightweight format for monitoring new content added to websites.

After you download an RSS reader or browser plug-in, you can subscribe to IBM product feeds at IBM Software Support RSS feeds (<https://www.ibm.com/software/support/rss>).

Log files

Log files can help you troubleshoot problems by recording the activities that take place when you work with a product.

Error messages

The first indication of a problem is often an error message. Error messages contain information that can be helpful in determining the cause of a problem.

Performance issues because Netezza database needs rebalancing

To prevent or mitigate performance issues related to your Netezza database, you can perform a rebalance operation. A rebalance is automatically performed when you restart Netezza.

To check if rebalancing of the SPUs is necessary, perhaps due to a hardware failure, log in to the server using the nz account and run the command: `nzhw rebalance -check`.

If you receive a message indicating that a rebalance is needed, then ensure that no users are on the system, and execute the following two commands to shut down and restart Netezza:

```
nzstop
```

```
nzstart
```

Performance is impacted after large amounts of data are deleted

If large amounts of data have been deleted from your Netezza database, not truncated, you must purge the deleted data from the table in order to return the table to peak performance.

Login as the user nz and perform the following command:

```
nzreclaim -db < database name > -t < table name > -records
```

Performance is impacted after large amounts of data have been inserted

After large volumes of data is inserted to your Netezza database, it is advisable to groom the tables. The command can be executed as either of the users nz or admin.

Connect to the database through a SQL session and execute the following statements:

```
groom table < table name >;
```

```
generate statistics on < table name >;
```

After running the solution installer, Chef processes are still running after executing the cleanup.sh script

After running the cleanup.sh script on the solution installer node, all chef processes should be terminated. However, in some cases chef processes are still running.

If you try to run the solution installer setup.sh file again, you may receive a message indicating that required ports are in use, or creation of the chef user might fail.

The solution to this issue is to reboot the installation node after having run "cleanup.sh".

Note: This only applies to the solution installer node and not the client/target nodes.

Appendix C. Importing data specifications for IBM Lift Analytics for Retail

This section summarizes the overall requirements for delivering customer data to IBM Lift Analytics for Retail. It serves as a business-level reference for retail customers. Any deviations in the file formats will require customization of the data loading process. Data is expected to be delivered and processed weekly.

Acceptable data formats

IBM will only accept data in plain text formats (ASCII for standard single-byte data, or UTF-8 or multi-byte data). The format is character-delimited flat files, using the pipe character | as the delimiter.

Table 2. Field Type Definitions

Type	Size
Small Integer	2 Bytes/5 Digits
Integer	4 Bytes/10 Digits
Big Integer	8 Bytes/19 Digits
Date	YYYY-MM-DD hh:mm:ss

Product Identifier. It is important that you provide one, unambiguous code that identifies a product over time. The product identifier should align with what the consumer perceives is a product. This ensures that IBM's modeling process properly identifies a product's behavior. Not providing a sellable item identifier in this manner will result in inaccurate models.

Value Added Tax (VAT) / Global Services Tax (GST). Value Added Tax (also known as VAT, TVA, or GST) is a sales tax that is built into the shelf price for some non-U.S. customers (U.S. customers do not need to consider this file type). In order to accurately develop models and report financial benefits, it is necessary that VAT is accounted for properly in the various data feeds. At a high level, the requirements are as shown in the following table.

Table 3. Value Added Tax (VAT) / Global Services Tax (GST)

Data Set	Field	VAT/GST	Notes
TLog Data	UnitScanPrice	Include	A separate field is provided in the interface to hold VAT amounts. However, VAT should still be included in the UnitScanPrice, UnitRegularPrice and SalesRevenue field values.

Table 3. Value Added Tax (VAT) / Global Services Tax (GST) (continued)

Data Set	Field	VAT/GST	Notes
TLog Data	UnitRegularPrice	Include	A separate field is provided in the interface to hold VAT amounts. However, VAT should still be included in the UnitScanPrice, UnitRegularPrice and SalesRevenue field values.
Current Optimization Cost	UnitOptCost	Exclude	
Current Retail (Base) Price Data	UnitRegularPrice	Include	

Data requirements

The following sections detail what data sets are necessary for IBM to begin processing.

Data Sets and Schedule. The following table presents an overview of the data sets that are needed for the IBM applications.

Table 4. File Type History and Update Frequency

Data Set	Minimum History	Recommended History	Update Frequency
T-Log Data	1 year	2 years	Daily
Product Hierarchy	Current	Current	As mappings change
Store Data	All stores that were active in T-Log period, even if closed	All stores that were active in T-Log period, even if closed	As stores change
SI Segment Types	Current	Current	As mappings change
SI Segment Type to Segment Map	Current	Current	As mappings change
SI Customer to Segment Map	Current	Current	As mappings change

Product Hierarchy Data. Product Hierarchy refers to the hierarchical grouping or classification of products for the purpose of merchandising and pricing. Lift Analytics allows a four level hierarchy. It is flexible but as an example could include Department, Category, Line and SKU. It is important to include a sellable item identifier for all products, including currently discontinued items, appearing in the TLog data. The sellable item is defined as a product entity that can be purchased by a consumer at the point of sale. For example, a single bar of soap is a sellable item whereas a case of 100 soap bars is not a sellable item. TLog data needs to be tied to the sellable item identifier. Sellable items have to be uniquely identifiable by the Item Identifier.

Table 5. Product Hierarchy Data

Field Name	Required	Type	Notes
Product level A Code	Yes	Varchar(50)	This is the code for the highest roll up of products in the tree, for instance a department code
Product Level A Name	Yes	Varchar(50)	This is the code for the highest roll up of products in the tree, for instance a department code
Product Level B Code	Yes	Varchar(50)	This is the code for the 3rd level roll up of products in the tree, for instance a category code
Product Level B Name	Yes	Varchar(50)	This is the code for the 3rd level roll up of products in the tree, for instance a category code
Product Level C Code	Yes	Varchar(50)	This is the code for the 2nd level roll up of products in the tree, for instance a line code
Product Level C Name	Yes	Varchar(50)	This is the code for the 2nd level roll up of products in the tree, for instance a line code
Product Level D Code	Yes	Varchar(50)	This is the unique product identifier at the lowest level, such as a UPC
Product Level D Name	Yes	Varchar(50)	This is the unique product name that corresponds with the code

Important: The unique product level D code:

- is required
- is the product key that is used to model and optimize
- should be the key that is used in all other data sets that require a product key
- should uniquely and unambiguously define an item over time
- is usually a regional product identifier (UPC, EAN, JAN), but does not have to be

Location Data. Recommended file name is `location.csv`. This data should include all stores that are listed in the historical and current TLog data, as well as all currently open stores.

Table 6. Location Data

Field Name	Required	Field Type	Default	Notes
LocationKey	Yes	Varchar(50)		Customer Location Key which uniquely identifies the location across the company
AdZone	No	Varchar(50)	(null)	Administration zone
LocationStatus	Yes	Int		Status of Location: Open (1), Closed (0)
LocationName	Yes	Varchar(50)		Location name or description
OpenDate	No	Date		Date format: YYYY-MM-DD
CloseDate	No	Date		Date format: YYYY-MM-DD
DistributionCenter	No	Varchar(50)		
AddressLine1	No	Varchar(255)	(null)	Address line 1
AddressLine2	No	Varchar(255)	(null)	Address line 2
City	No	Varchar(50)	(null)	City
County	No	Varchar(50)	(null)	County
StateProvince	No	Varchar(50)	(null)	State or province
PostalCode	No	Varchar(50)	(null)	Postal code (ZIP code in U.S.)
Country	No	Varchar(50)		
StoreFormat	No	Varchar(50)	(null)	
SqFt	No	Varchar(50)	(null)	Square footage of the Location
MedianAge	No	dec(9.4)	(null)	Median age of customers at Location
MedianIncome	No	dec(9.4)	(null)	Median income of customers at Location
PercentEthnic	No	dec(9.5)	(null)	Percent of ethnic customers
DateLastRemodel	No	Date	(null)	Date Location was last remodeled. Prefer 4-digit year in date. Date format: YYYY-MM-DD
LocationLevel1Code	No	Varchar(25)	(null)	Code for first level of location hierarchy. These fields are helpful in creating store groupings.

Table 6. Location Data (continued)

Field Name	Required	Field Type	Default	Notes
LocationLevel1Descr	No	Varchar(50)	(null)	Description/name of this level of location hierarchy
LocationLevel2Code	No	Varchar(25)	(null)	Code for second level of location hierarchy
LocationLevel2Descr	No	Varchar(50)	(null)	Description/name of this level of location hierarchy
LocationLevel3Code	No	Varchar(25)	(null)	Code for third level of location hierarchy
LocationLevel3Descr	No	Varchar(50)	(null)	Description/name of this level of location hierarchy
LocationLevel4Code	No	Varchar(25)	(null)	Code for fourth level of location hierarchy
LocationLevel4Descr	No	Varchar(50)	(null)	Description/name of this level of location hierarchy
LocationLevel5Code	No	Varchar(25)	(null)	Code for fifth level of location hierarchy
LocationLevel5Descr	No	Varchar(50)	(null)	Description/name of this level of location hierarchy
LocationLevel6Code	No	Varchar(25)	(null)	Code for sixth level of location hierarchy
LocationLevel6Descr	No	Varchar(50)	(null)	Description/name of this level of location hierarchy
LocationLevel7Code	No	Varchar(25)	(null)	Code for seventh level of location hierarchy
LocationLevel7Descr	No	Varchar(50)	(null)	Description/name of this level of location hierarchy
LocationLevel8Code	No	Varchar(25)	(null)	Code for eighth level of location hierarchy
LocationLevel8Descr	No	Varchar(50)	(null)	Description/name of this level of location hierarchy
LocationLevel9Code	No	Varchar(25)	(null)	Code for ninth level of location hierarchy
LocationLevel9Descr	No	Varchar(50)	(null)	Description/name of this level of location hierarchy
LocationLevel10Code	No	Varchar(25)	(null)	Code for tenth level of location hierarchy

Table 6. Location Data (continued)

Field Name	Required	Field Type	Default	Notes
LocationLevel10Descr	No	Varchar(50)	(null)	Description/name of this level of location hierarchy
LocationLevel11Code	No	Varchar(25)	(null)	Code for eleventh level of location hierarchy
LocationLevel11Descr	No	Varchar(50)	(null)	Description/name of this level of location hierarchy
LocationLevel12Code	No	Varchar(25)	(null)	Code for twelfth level of location hierarchy
LocationLevel12Descr	No	Varchar(50)	(null)	Description/name of this level of location hierarchy
Division Code	No	Varchar(50)		Banner or division code
Division Name	No	Varchar(128)	(null)	Banner or division name
Channel	No	Varchar(50)	(null)	In Store or Online
LocationType	Yes	Integer	0	0=Store 1=Market
Timezone	No	Varchar(75)	(null)	Format = 'GMT +/- HH:00 GMT Example: GMT -08:00 GMT
Source Location Key	No	Varchar(50)	(null)	Store that will be used as a reference to assign a new store to the appropriate LGC.
Store Manager	No	Varchar(50)	(null)	Store Manager's name
IsFranchiseStore	No	SmallInt (2)	0	Possible values are 0 (false) or 1 (true); indicates a franchise store for use with franchise store application pricing functionality
Latitude	No	dec(9.7)	0	
Longitude	No	dec(10.7)	0	

T-Log Data. Transaction log, or T-log data serves as the basis for Market Basket Analysis. All transactions should be included. Recommended file name:[CUSTOMER NAME]-tlog-YYYYMMDD.txt. Historic and daily TLog files should contain records covering a single day's worth of transactions. The file name should contain the date covered by the file.

Table 7. T-Log Standard Format

Field Name	Required	Field Type	Default	Notes
SequenceNum	Yes	int		Uniquely identifies a line within a transaction. The sequence number starts from 1
Register ID	Yes	Varchar(30)		Uniquely identifies a register in a store. In the case of web store, the Register ID should be the web transaction id
Location Key	Yes	Varchar(30)		Uniquely identifies the store at which a given transaction occurred. The key should match a location stored in IBM
Transaction ID	Yes	Int(20)		Uniquely identifies the transaction within the company
Transaction Date	Yes	Date		Date that the transaction occurred. Use the following date format: YYYYMMDD
Transaction Time	Yes	Time		Time that the transaction occurred. Time format: HH:MI:SS

Table 7. T-Log Standard Format (continued)

Field Name	Required	Field Type	Default	Notes
Scan Type	Yes	Varchar(30)		<p>Scan Type distinguishes each line in the transaction depending on whether the scan recorded a regular sale, a return, a discount or an exchange.</p> <p>A four character scan code is used in this field. The following scan types are permitted:</p> <ul style="list-style-type: none"> • SALE: All items that scan as regular sales • DISC: All items that scan as discount sales • GCFR: Gift Card Free: Applies to gift cards that are given away for free with products • RETN: All Items that are returned • EXCH: All Items that are exchanged. There is no distinction between different types of exchanges • TEND: The form of tender used

Table 7. T-Log Standard Format (continued)

Field Name	Required	Field Type	Default	Notes
Item Key	Yes	Varchar(50)		<p>Values can vary depending on the Scan Type. The following values should be applied based on the Scan Type:</p> <p>For Scan Type SALE, RETN or EXCH, the ItemKey field will hold a unique product identifier (Sellable Item Identifier) such as a UPC or SKU that matches the unique product key in the Sellable Item file. All identical product keys should be rolled up into a single record within a given transaction</p> <p>For Scan Type DISC or GCFR, the ItemKey field will hold a Discount Code, an identifier that ties to a predefined discount type. All the available discount types and their unique codes should be defined in the customer supplied data dictionary.</p> <p>For Scan Type TEND, the ItemKey fields will hold a Tender Code, an identifier that ties the line to a predefined set of tender type values.</p> <p>All the available tender types and their unique codes should be defined in the customer supplied data dictionary</p>

Table 7. T-Log Standard Format (continued)

Field Name	Required	Field Type	Default	Notes
Quantity	Yes	dec (9.2)		Indicates quantity of products sold in the transaction. This value must always be positive
Scan UOM	Yes	Varchar(2)	EA	Indicates how the purchase was counted. EA = Each WT = Weight Note that the two-character abbreviation is capitalized. Also note that the Scan UOM value does not interfere with the UOM in the Sellable Item file. The two UOM codes are separate

Table 7. T-Log Standard Format (continued)

Field Name	Required	Field Type	Default	Notes
Total Scan Price	Yes	dec (9.2)		<p>For a SALE record: Total Price for a SALE record must be a positive value and should generally be calculated as follows:</p> <p>Unit Base Price of the Item * Quantity Purchased = Total Price</p> <p>Total Price for a SALE record should not include discounts. During file processing, we apply all associated discounts (captured by the DISC records) to the corresponding SALE record in order to arrive at the true price the customer paid for the item (the scan price). The item's scan price is key to our models. It is critical that discounts are not reflected both in the starting Total Price and through DISC records. If that is the case, we will double process discounts and arrive at a scan price which is artificially low</p>

Table 7. T-Log Standard Format (continued)

Field Name	Required	Field Type	Default	Notes
Total Scan Price Continued	Yes	dec (9.2)		<p>For a DISC record: Total Price must have a negative value and it should capture the total discount amount for the discount type:</p> <p>Unit Discount Amount for the Discount Type * Quantity = Discount Amount for the Discount Type</p> <p>For a RETN record: Total price must have a negative value</p> <p>For a TEND record: Total Price will capture the amount of the basket paid for by the specific TEND type. Baskets can have multiple TEND records. TEND records should exclude sales tax</p>
Unit Regular Price	No	dec (9.2)		<p>The unit base price of the product would only be sent with SALE, RETN, or EXCH scan types.</p>

Table 7. T-Log Standard Format (continued)

Field Name	Required	Field Type	Default	Notes
Reference ID	No	Varchar(30)		Ties line items within a transaction together in order to link records capturing discounts or packages to the base sales record for a product. Formatted as a comma separated list of sequence numbers, where each sequence number refers to the corresponding sequence number of the product that is associated with the discount or package
Unit List Cost	No	dec (9.4)		The cost per unit of the product specified would only be sent in the SALE , RETN , or EXCH scan types
Customer Number	No	Varchar(30)		Uniquely identifies the customer that purchased the basket that is represented in the transaction (club or loyalty card)
Original Transaction ID	No	Varchar(50)		Used for return items to link the return to the original transaction id the item was purchased
Package ID	No	Varchar(50)		Used to identify and link products which are being promoted together on a package-type promotion
Reserved 1	No	Varchar(50)		Reserved for future use

Table 7. T-Log Standard Format (continued)

Field Name	Required	Field Type	Default	Notes
VAT Amt	No	dec (9.2)		VAT amount for the SALE scan types for non-US customers Note that VAT quantities must be included/excluded from other fields in the TLog record as indicated by the table in section 2.2 VAT should be provided as a positive value across all Scan types
Reserved 2	No	Varchar(50)		Reserved for future use
UserField1	No	Varchar(30)		Populated as necessary for special or customer specific purposes
UserField2	No	Varchar(30)		Populated as necessary for special or customer specific purposes
UserField3	No	Varchar(30)		Populated as necessary for special or customer specific purposes

A data dictionary will all applicable codes used for line items with Scan Type = **DISC** is required. The data dictionary should contain the codes and a short description of the discounts.

Table 8. Discount codes for line items with Scan Type = DISC

Field Name	Required	Field Type	Default	Notes
Discount Code	Yes	Varchar(30)		Code used in the Item Key field in T-Log when the Scan Type = DISC
Description	Yes	Varchar(50)		Description of the discount code

A data dictionary with all applicable codes used for line items with Scan Type = **TEND** is required. The data dictionary should contain the codes and a short description of the discounts.

Table 9. Discount codes for line items with Scan Type = TEND

Field Name	Required	Field Type	Default	Notes
Tender Code	Yes	Varchar(30)		Code used in the Item Key field in T-Log when the Scan Type = TEND
Description	Yes	Varchar(50)		Description of the discount code

A data dictionary with all applicable codes used for line items with Scan Type = GCFR is required. The data dictionary should contain the codes and a short description of the discounts.

Table 10. Discount codes for line items with Scan Type = TEND

Field Name	Required	Field Type	Default	Notes
GCFR Code	Yes	Varchar(30)		Code used in the Item Key field in T-Log when the Scan Type = GCFR
Description	Yes	Varchar(50)		Description of the discount code

This example illustrates how this format will be used for a basket that contains regular sale items combined with discounts at the item level and transaction level. The basket is paid for using a gift card and a visa card.

Table 11. Regular purchases and item level and transaction level discounts

Seq No.	Register ID	Location Key	Transaction ID	Transaction Date	Scan Type	Item Key	Qty	Total Scan Price	Ref. ID	Unit Cost
1	9874	94070	135678	20141001	Sale	DVDPLAYr	1	\$140		\$30
2	9874	94070	135678	20141001	Sale	SRND_SPKRS	2	\$80		\$10
3	9874	94070	135678	20141001	Disc	COUPON	1	-\$10	2	
4	9874	94070	135678	20141001	Disc	EMPTY_DISC	1	-\$25	2	
5	9874	94070	135678	20141001	Sale	UNV_RMTE	1	\$120		\$50
6	9874	94070	135678	20141001	Disc	EMPTY_DISC	1	-\$25	0	
7	9874	94070	135678	20141001	Tend	GIFT_CRD	1	-\$50		
8	9874	94070	135678	20141001	TEND	VISA	1	-\$230		

In the specified format, products, item level discounts, transaction level discounts and tender types are all captured as separate lines. Every line in a transaction has unique sequence number. The sequence number starts at 1 for every transaction.

In this example, DVDPLAYR and SRND_SPKRS are regular sale items. The SRND_SPKRS come with a coupon that is redeemed at the register. Line 3 indicates that a discount was applied and the value of the discount is reflected in the TotalScanPrice field. The Ref Id field in Line 3 references the SRND_SPKRS item to indicate that the discount is tied to the that specific product. The person purchasing this basket is an employee so the employee discount is seen in Line 4

and Line 6 as EMPLY_DISC (employee discount). The retailer has a policy that employee discounts are not allowed on the latest release DVD Player so in the basket, only theSRND_SPKRS and the UNV_RMTE get the employee discount.

Tender types are also included as line items. In this example, a gift card valued at \$50 was used to pay for part of the purchases. This tender is indicated in line 7. The remaining amount was paid with a visa card as seen in line 8.

This example demonstrates how mix and match type discounts are encoded. In this example, the item that is free is not the same as the item that is promoted. There are two options to encoding these types of discounts. In both cases, it is important that the TLOG data accurately reflects which items are tied to the discount.

- Option 1. In the first option, the free item is treated as a discount and there is one line item with Scan Type = **DISC** that captures the total discount. The discount line must be tied to the items that were involved in that particular promotion. This link is established by the Reference ID column for the **DISC** line item. The Reference ID should contain the sequence numbers of all the line items involved. IBM will distribute the discount among all items involved in the promotion in a weighed manner.

Table 12. Option 1 - Quantity based discounts: Buy 2 Get 1 Free (Purchased Item is not the same as the Free Item)

Seq No.	Register ID	Location Key	Transaction ID	Transaction Date	Scan Type	Item Key	Qty	Total Scan Price	Ref. ID	Unit Cost
1	9874	94070	1234567	20141001	Sale	Canned Pineapple	1	\$3		\$1
2	9874	94070	1234567	20141001	Sale	Canned Mango	1	\$3		\$1
3	9874	94070	1234567	20141001	Sale	Canned Pear	1	\$3	2	\$1
4	9874	94070	1234567	20141001	Disc	Buy 2, 1 Free	1	-\$3	1, 2, 3	

- Option 2. An alternate way to send the same information is to create discount records for each item involved in the promotion.

Table 13. Option 2- Quantity based discounts: Buy 2 Get 1 Free (Purchased Item is not the same as the Free Item)

Seq No.	Register ID	Location Key	Transaction ID	Transaction Date	Scan Type	Item Key	Qty	Total Scan Price	Ref. ID	Unit Cost
1	9874	94070	1234567	20141001	Sale	Canned Pineapple	1	\$3		\$1
2	9874	94070	1234567	20141001	Disc	Buy 2, 1 Free	1	-\$1	1	
3	9874	94070	1234567	20141001	Sale	Canned Mango	1	\$3		\$1
4	9874	94070	1234567	20141001	Disc	Buy 2, 1 Free	1	-\$1	3	
5	9874	94070	1234567	20141001	Sale	Canned Pear	1	\$3		\$1
6	9874	94070	1234567	20141001	Disc	Buy 2, 1 Free	1	-\$1	5	

This example demonstrates how mix and match type discounts or quantity based discount are encoded. In this example, the purchased item(s) is the same as the

item that is free. There are two options to encoding these types of discounts. In both cases, it is important that the TLOG data accurately reflects which items are tied to the discount.

- Option 1. In the first option, there is one discount record that captures the complete discount for the purchase. The reference id links the discount to the purchased item. This is the preferred option.

Table 14. Option 1 - Quantity based discounts: Buy 2 Get 1 Free (Purchased Item is the same as the Free Item)

Seq No.	Register ID	Location Key	Transaction ID	Transaction Date	Scan Type	Item Key	Qty	Total Scan Price	Ref. ID	Unit Cost
1	9874	94070	1234567	20141001	Sale	Canned Pineapple	3	\$9		\$1
2	9874	94070	1234567	20141001	Disc	Buy 2, 1 Free	1	-\$3	1	

- Option 2. An alternate way to encode the same information is to create a discount line item for every purchased item in the transaction that is related to the specific discount. Each discount record contains a portion of the discount (value of the free item) that is attributed to the purchased item.

Table 15. Option 2 Quantity based discounts: Buy 2 Get 1 Free (Purchased Item is the same as the Free Item)

Seq No.	Register ID	Location Key	Transaction ID	Transaction Date	Scan Type	Item Key	Qty	Total Scan Price	Ref. ID	Unit Cost
1	9874	94070	1234567	20141001	Sale	Canned Pineapple	1	\$3		\$1
2	9874	94070	1234567	20141001	Disc	Buy 2, 1 Free	1	-\$1	1	
3	9874	94070	1234567	20141001	Sale	Canned Pineapple	1	\$3		\$1
4	9874	94070	1234567	20141001	Disc	Buy 2, 1 Free	1	-\$1	3	
5	9874	94070	1234567	20141001	Sale	Canned Pineapple	1	\$3		\$1
6	9874	94070	1234567	20141001	Disc	Buy 2, 1 Free	1	-\$1	5	

Potential data issues:

- A sequence number is embedded within the TransactionID, making every TransactionID value unique
- Discontinued items have been removed from historical T-Log data
- Closed stores have been removed from the historical T-Log data
- Products that were previously in the implementation categories have since moved to new categories and are not included in the historical T-Log extract
- T-Log data contains products that cannot be found in the Product file
- T-Log data contains stores that cannot be found in the Store file
- Selective product information (e.g., certain brands) not included in data set
- Certain weeks missing from the historical data
- No gap data provided to cover the missing weeks between the historical data feed and the first update

There are three segmentation interfaces:

1. Segment Type
2. Segment Type to Segment Map
3. Customer to Segment Map

The combination of these interfaces will provide a holistic picture of the shopping behavior of customers in various segments.

Recommended file name: SEGMENT_TYPE.CSV. The Segment Types interface allows customers to send information about their defined segment classes. The Segment Type ID is the unique key in the interface. Segment Type information can be updated through this interface.

Field Name	Required	Field Type	Default	Notes
Segment Type ID	Yes	Integer		Unique number identifying the Segment Type. Value > 10. 0 - 9 is reserved
Segment Type Name	Yes	Varchar(128)		Short name for the Segment Type
Segment Type Description	No	Varchar(256)		Description of the Segment Type
Primary Segment Type ID	No	Integer		Related segment type - Primary
Secondary Segment Type ID	No	Integer		Related segment type - Secondary

Recommended file name: SEGMENT.CSV. The Segment Type to Segment Map interface allows customers to use the different segments that are tied to each segment type. The data for this interface requires the Segment Type data to have already been defined in the application.

Field Name	Required	Field Type	Default	Notes
Segment Type ID	Yes	Integer		Unique number identifying the segment type
Segment Type Name	Yes	Varchar(128)		Short segment name
Segment Type Description	No	Varchar(256)		Description of the segment
Primary Segment Type ID	No	Integer		Segment id of the primary segment type
Secondary Segment Type ID	No	Integer		Segment id of the secondary segment type

Recommended file name: CUSTOMER_SEGMENT_MAP.CSV. A customer may belong once to each segment type. Anonymous transactions will be assigned to a default segment during loading.

Field Name	Required	Field Type	Default	Notes
Customer ID	Yes	Integer		Unique number identifying the customer over time
Segment ID	Yes	Integer		Unique number identifying the segment across all segment types
Date	Yes	Date	Current [®] Date	The date that the segment is effective, format YYYY-MM-DD

This list contains potential customer data issues.

- Field layout of files does not match data dictionary definition
- Status codes are not fully explained or data contains more status codes than are defined
- Weekly updates contain the same data as the previous week

Appendix D. Data dictionary

The data dictionary describes the Netezza database tables and attributes included with the Lift Analytics for Retail solution. In the first section, it identifies tables and attributes that are, or could be, used by the reports. Below that, are supporting tables used to import and transform transaction data into the datamarts used by the reports.

The data dictionary is available as a Microsoft Excel file (data_dictionary_v4.xlsx) from the following location:

Industry Solutions techdoc <http://www.ibm.com/support/docview.wss?uid=swg27046047>

Notices

This information was developed for products and services offered worldwide.

This material may be available from IBM in other languages. However, you may be required to own a copy of the product or product version in that language in order to access it.

IBM may not offer the products, services, or features discussed in this document in other countries. Consult your local IBM representative for information on the products and services currently available in your area. Any reference to an IBM product, program, or service is not intended to state or imply that only that IBM product, program, or service may be used. Any functionally equivalent product, program, or service that does not infringe any IBM intellectual property right may be used instead. However, it is the user's responsibility to evaluate and verify the operation of any non-IBM product, program, or service. This document may describe products, services, or features that are not included in the Program or license entitlement that you have purchased.

IBM may have patents or pending patent applications covering subject matter described in this document. The furnishing of this document does not grant you any license to these patents. You can send license inquiries, in writing, to:

IBM Director of Licensing
IBM Corporation
North Castle Drive
Armonk, NY 10504-1785
U.S.A.

For license inquiries regarding double-byte (DBCS) information, contact the IBM Intellectual Property Department in your country or send inquiries, in writing, to:

Intellectual Property Licensing
Legal and Intellectual Property Law
IBM Japan Ltd.
19-21, Nihonbashi-Hakozakicho, Chuo-ku
Tokyo 103-8510, Japan

The following paragraph does not apply to the United Kingdom or any other country where such provisions are inconsistent with local law: INTERNATIONAL BUSINESS MACHINES CORPORATION PROVIDES THIS PUBLICATION "AS IS" WITHOUT WARRANTY OF ANY KIND, EITHER EXPRESS OR IMPLIED, INCLUDING, BUT NOT LIMITED TO, THE IMPLIED WARRANTIES OF NON-INFRINGEMENT, MERCHANTABILITY OR FITNESS FOR A PARTICULAR PURPOSE. Some states do not allow disclaimer of express or implied warranties in certain transactions, therefore, this statement may not apply to you.

This information could include technical inaccuracies or typographical errors. Changes are periodically made to the information herein; these changes will be incorporated in new editions of the publication. IBM may make improvements and/or changes in the product(s) and/or the program(s) described in this publication at any time without notice.

Any references in this information to non-IBM Web sites are provided for convenience only and do not in any manner serve as an endorsement of those Web sites. The materials at those Web sites are not part of the materials for this IBM product and use of those Web sites is at your own risk.

IBM may use or distribute any of the information you supply in any way it believes appropriate without incurring any obligation to you.

Licensees of this program who wish to have information about it for the purpose of enabling: (i) the exchange of information between independently created programs and other programs (including this one) and (ii) the mutual use of the information which has been exchanged, should contact:

IBM Software Group
Attention: Licensing
3755 Riverside Dr.
Ottawa, ON
K1V 1B7
Canada

Such information may be available, subject to appropriate terms and conditions, including in some cases, payment of a fee.

The licensed program described in this document and all licensed material available for it are provided by IBM under terms of the IBM Customer Agreement, IBM International Program License Agreement or any equivalent agreement between us.

Any performance data contained herein was determined in a controlled environment. Therefore, the results obtained in other operating environments may vary significantly. Some measurements may have been made on development-level systems and there is no guarantee that these measurements will be the same on generally available systems. Furthermore, some measurements may have been estimated through extrapolation. Actual results may vary. Users of this document should verify the applicable data for their specific environment.

Information concerning non-IBM products was obtained from the suppliers of those products, their published announcements or other publicly available sources. IBM has not tested those products and cannot confirm the accuracy of performance, compatibility or any other claims related to non-IBM products. Questions on the capabilities of non-IBM products should be addressed to the suppliers of those products.

All statements regarding IBM's future direction or intent are subject to change or withdrawal without notice, and represent goals and objectives only.

This information contains examples of data and reports used in daily business operations. To illustrate them as completely as possible, the examples include the names of individuals, companies, brands, and products. All of these names are fictitious and any similarity to the names and addresses used by an actual business enterprise is entirely coincidental.

If you are viewing this information softcopy, the photographs and color illustrations may not appear.

This Software Offering does not use cookies or other technologies to collect personally identifiable information.

Trademarks

IBM, the IBM logo and ibm.com are trademarks or registered trademarks of International Business Machines Corp., registered in many jurisdictions worldwide. Other product and service names might be trademarks of IBM or other companies. A current list of IBM trademarks is available on the Web at “ Copyright and trademark information ” at www.ibm.com/legal/copytrade.shtml.

Index

A

accessibility 31

C

Cognos

- adding Netezza library 12
- creating data source connection 13
- deploying content 12
- deploying images 14
- deploying reports 13
- preparing content 12

configuration 3

- modifying sudoers file 5
- overview 3
- planning 4
- prerequisites 3
- solution installer 7

D

data dictionary 57

data specifications

- importing 37

database 9

DataStage

- aggregation fact table ETL process 20
- creating a project 11
- jobs 15

download package 3, 4

F

firewall settings 5

- configure with script 6
- disabling 5
- manually configuring 6
- on all computers 7

firewall.sh 6

H

home page 14

I

InfoSphere Information Server 9

- importing the package 11

- installing server-side content 11

installation 3

- modifying sudoers file 5

- overview 3

- planning 4

- prerequisites 3

- solution installer 7

N

Netezza

- configuring data source 10

- installing client software 9

R

report

- Affinity Performance 29

- Affinity Relationship 30

- choosing 28

- Customer Segment Performance 29

- Product Performance Summary 29

- Product Performance Trend 28

Retail Lift database 9

S

sudoers file 5

T

TLOG ETL process 19

troubleshooting 33

W

web page 14