pdssr Documentation

Release 0.0.3

Pacific Data Services

Jan 10, 2020

CONTENTS:

CHAPTER

ONE

INTRODUCTION

Pepsi-Frito uses Custom Data solutions to record and report vending sales. I am extremely familiar with this operation as I taught Custom Data Solutions the technology for 25 years and functioned as their first Chief Information Officer

I redesigned the entire process in the two years I was an employee.

More in *Contributions*

CHAPTER

TWO

BENEFITS

- Audit the data and rebate calculations. *Audit*
- Get answers without asking questions *Answers to Questions you should have asked*
- Custom Reporting
- Online analytical processing
- Startup is negligible
  - Just need copies of the files the distributors already report *Record Layouts*
  - A product master file
  - No computing resources, no long term contracts

Once installed:

1. Reporting data is loaded into staging tables
2. Verification of compliance with requirements and self consistency of data is ensured
3. Data is posted to history tables
4. Customer addresses are validated
5. Exceptional conditions are identified and reported
6. Rebate Calculations are made and reported
7. Rebate amounts are computed and an ACH file for transmission to a bank for payment is created
8. Spreadsheets for sales analysis are created

CHAPTER

THREE

FUNCTIONING REAL WORLD USAGE

Any distributor currently writing files in the CDS reporting format can easily load files into a local database in order to analyze the data.

We have provided many analyses that report data worthy of further inspection.

This approach is our “answers to questions you should have asked”.

CHAPTER

FOUR

AUDIT

Pacific Data Services will audit the data and rebate calculations for your distributors.

4.1 Findings

4.1.1 Item Accuracy

We will compare the items assigned by Custdata to the actual items sold by the distributors.

This affect sales history and rebate calculations

4.1.2 Addresses

Report incorrect addresses and undeliverable addresses.

4.1.3 Rebate

Calculate the actual rebate for the top 50 rebate amounts issued versus what Custdata computed.

4.2 Report augmentation

4.2.1 Answers to Questions you should have asked

Rather than spend time on the Custdata website creating spreadsheets and trying to spot interesting and unusual data we will create *Alert Reports* and notify you that there is research warranted.

Examples

- Newly introduced items not being bought by distributor
- Top selling items by category not being bought by distributor
- Top selling items by category not being bought by a vending operator
- Product declining sales for a vending operator
- Dropped products for a vending operator
- Parasitic promotion effect (Items not on promotion having declining sales during a promotion

as slot contention has had a parasitic effect.

CHAPTER

FIVE

SALES

5.1 Required Data

Distributors already report their data to custdata.com.

They have to do nothing but send us the same data files, no additional work.

5.2 Benefits

5.2.1 Manufacturers

Benefits

- Far lower costs
- Faster and Cheaper Custom Reports Generation
- Improved Data Quality
- Far Lower Costs
- Data Identified for them

  Currently the manufactures have to click through the Custdata website and create a spreadsheet and see if there is interesting data.

  We can analyze the data and inform them of interesting conditions, for example

  - New Vending Operators
  - Vendors with Declining Sales
  - Vendors not carrying products

New Capabilities

- Address Correction
-

5.2.2 Distributors

- Ability to create an operator facing website

5.2.3 Vending Operators

- Find new products to be offered

CHAPTER

SIX

BACKGROUND

6.1 Sales Analysis

6.1.1 Gather Product Statistics

- Total sales prior 12 months
- Percent of total sales by product
- Operator maximum percentage
- First sale date
- Trend line slope
- Seasonality

6.1.2 Operator Product Statistics

- Pct of sales for product by operator
- Ratio of percent of sales for this operator to median percentage
- Total dollars sales by product previous twelve months

6.1.3 Opportunities

6.1.4 Operator Call Report

- Product mix to standard mix
- Suggested products to carry

6.2 Answers to Questions you should have asked

We do the work of finding data that is interesting and provide email notices that these reports exist. Anyone subscribing to that event can get an email summary notification and click directly to the report, only having to login. We can support gmail or facebook authentication so the user doesn't need another username and password.

It couldn't be easier.

Answers phone friendy format or downloadable in a variety of formats.

6.2.1 Examples

- Which operators are not selling recently introduced products
- Which operators have a sub-optimal mix of the top selling items by category
- Product declining sales for a vending operator
- Dropped products for a vending operator
- Parasitic promotion effect (Items not on promotion having declining sales during a promotion

as slot contention has had a parasitic effect.

6.3 Sales Analysis

List customers and the products they are should be selling but are not.

- Rank Products
-

As we are dealing with CDS reporting format and the scope of creation of customers in posting is beyond the scope of this section I will create a customer with a simple database view.

code:

```
create or replace view
customer_vw as
select distinct ship_to_cust_id
from etl_sale;
```

In order for these queries to work without recreating the database we assume an *effective date* which is the date of the last reported sale.

code:: create or replace view effective_date as select max(invoice_dt) report_date from etl_sale;

Total $ amount of sales for the preceeding 12 months based on last invoice date.

code:

```
create or replace view tot_sales_12 as
select sum(extended_net_amt) tot_extended_net_amt
from etl_sale,
     effective_date
where invoice_dt  > effective_date.report_date - interval '1 year';
```

code:

```
create or replace view customer_product_last_12_months as
select  etl_sale.ship_to_cust_id,
        case_gtin,
        sum(extended_net_amt) sum_extended_net_amt
from    etl_sale,
        effective_date
where   invoice_dt  > effective_date.report_date - interval '1 year'
group by ship_to_cust_id,
        case_gtin
order by sum(extended_net_amt) desc;
```

There are many ways to rank product.

Units, gross revenue, profit, turns...

This is based simply on gross revenue

code:

```
create or replace view product_rank_12_vw as
select etl_sale.case_gtin, sum(etl_sale.extended_net_amt) sum_ext,
       sum(etl_sale.extended_net_amt) * 100 / tot_sales_12.tot_extended_net_amt
from etl_sale,
     tot_sales_12
     effective_date
where invoice_dt  > effective_date.report_date - interval '1 year'
                group by case_gtin
                order by sum(extended_net_amt) desc
        ;
```

6.4 Product Not Sold

What products are customers not selling?

code:

```
create or replace view product_undersold_by_customer_vw as
select customers.ship_to_cust_id,
       top_products.case_gtin
from customers,
     top_products
where not exists
  ( select 'x' from etl_sale, top_products
    where
        etl_sale.ship_to_cust_id = customers.ship_to_cust_id  and
        etl_sale.case_gtin = top_products.case_gtin
  )
order by customers.ship_to_cust_id,
         case_gtin;
```

Now let us consider the amount that could be sold if these customers were to sell your standard ratio of product.

We do that by multiplying the ratio to the dollar total for the customer.

In the best case scenario, the vending operator will add a slot for the product, replacing with a product you don't distribute to him. There is a distinct possibility that the sale will be slightly parasitic and a lower volume product you do distribute will be replaced.

6.5 Audit

Pacific Data Services will audit the data and rebate calculations for your distributors.

6.5.1 Findings

Item Accuracy

We will compare the items assigned by Custdata to the actual items sold by the distributors.

This affect sales history and rebate calculations

Addresses

Report incorrect addresses and undeliverable addresses.

Rebate

Calculate the actual rebate for the top 50 rebate amounts issued versus what Custdata computed.

6.6 Benefits

6.6.1 Frito-Lay Pepsi

Audit

Identify product miscategorization. Product GTINS may be "guessed" from the product description, this is error prone.

ACH rebate payments

Rather than mail out checks, ACH transactions can be made, Pepsico-Frito can review and forward to its bank and rebate payments can be sent out electronically for about $0.01 per transaction.

This is faster, less expensive and less error prone than mailing out checks and more convenient for the distributors.

Anomolies

Report outliers on data reporting.

This is performed by the *javautil.org Condition Identification* package, written by Jim Schmidt for use by Trinity Technical Services customers to identify potential problems with data loading into the data warehouse.

Jim Schmidt demonstrated this while building an entire rebate program for a Custom Data Solutions client and it was used to replace the paper reporting system, saving a huge amount of time and increasing the identification of suspect data using metrics derived from historical data. This was completely designed and written by Jim Schmidt prior to joining CDS.

Product Assignment

6.6.2 Distributors

Address Standardization

Correct and standardize all mailing addresses and report undeliverable USPS addresses for vending operators.

Geo Coding

Return latitude and longitude for all valid mailing addresses.

Product Identification

Phone App

Salesmen can generate reports from their phone and forward to vending operators.

Customer Facing Reports

6.6.3 Apps

Phone app for

6.6.4 Vending Operators

6.7 Product Identification

6.8 Contributions

Jim Schmidt and his company Trinity Technical Services introduced and mentored Custom Data Solutions on every technology they used since 1990.

In 1990 Chuck Schmidt, the owner and president of Custom Data Solutions was developing software in assembly language on alphamicro computers.

6.8.1 Unix/Linux

I introduced Chuck to Unix, starting with Interactive Unix, a Kodak supported version of AT&T Unix System 5 in 1991.

Thereafter Trinity became a reseller of MIPS computers and then Digital Equipment.

6.8.2 Oracle

Jim introduced Chuck to Oracle and provided training and technical assistance from 1992 to 2007. Trinity staff trained Custom Data Staff on Oracle Forms, Oracle Reports and the Database.

6.8.3 C

6.8.4 Java

6.8.5 Became Employee

In 2007 I made an agreement with Custom Data Solutions to become the CIO with the understanding that he could work remotely.

So I sold everything I owned, two houses, two boats, two jet-skis, two ATVs and the rest of my possessions.

When my youngest son went off to college, I drove him to Purdue and went to Michigan to fix a few problems for a couple of months, that turned out to be over a year.

During that time period

- Migrated 10 vending databases to a single database
- Upgraded from Oracle 7 to Oracle 11
- Installed Oracle Enterprise manager and taught how to use it
- Created a development and a test database. Formerly all testing was done in production.

6.8.6 Technical Standards

- Estabished the use of build tools, Ant and Maven
- Instituted source version control, CVS and Subversion
- Unit Testing
- Integration Testing
- Designed and build a document management service
- Migrated off Digital Unix to Linux
- Introduced virtualization, vmware, openvz and virtualbox
-

6.8.7 Products

Check 21

Document Management

Workbook Parser

Dexterous

6.8.8 Redesign

CDS had a prospective client that needed rebate process

Data Load

I replaced the java program that loaded the data into a stored procedure that loaded into staging tables.

Exception Processing

I employed the javautil Exception Processing system to identify load problems.

Unknown to me at the time, Custom Data Solutions identified problems with reports. The report modification was cumbersome, execution was slow, printing was slow and wasteful.

By generating metrics of minimum and maximum and standard deviations from the mean the identification of outlier input data was greatly improved.

Maintenance System

The maintenance and control system was all written in Apex, whereas CDS was using Oracle Forms. My staff trained CDS on developing in Apex.

Posting

Workbooks

6.8.9 Others

- Servlets
- Javascripts
- Excel workbooks
- Detachable tablespaces
- Materialized Views
- Virtual Private Databases

6.8.10 Javautil

javautil.org was Jim Schmidt's domain javautil.com still is

6.9 History of Jim Schmidt and Custom Data Solutions

6.9.1 Creation

Custom Data Solutions was initially called CDS for Chuck and Dee Schmidt.

Chuck is my older brother and Dee is his wife.

Chuck left his job at Michigan National Bank to offer data processing services, initially using a TRS-80 to print mailing labels.

He subsequently bought an Alpha Micro http://www.s100computers.com/Hardware%20Folder/Alpha%20Micro/History/History.htm and wrote some business applications.

On a visit to Detroit I explained to him the benefits of Unix and got him started, installing Interactive Unix from Kodak on a PC from 133 floppy disks in 1990.

Between 1983 and 1990 I worked as Vice President of International Banking Systems for RepublicBank Dallas, as a principal at a Distribution Requirements Planning Company and then at Computer Associates where I single handedly designed and wrote the ACH settlement software system in six weeks.

In 1990 I started Trinity Technical Services with two partners and ended up buying them both out.

My first contract was the selection of an ERP system for a distributor in Texas. The chosen vendor was CIM-JIT and it was my first exposure to Oracle. I soon became more proficient in the CIM-JIT software than the vendor and flew around country to modify to delivered software to comply with commitments to customers.

In the process I became quite proficient in fixing performance problems.

Convinced that Oracle was the finest relational database, I convinced Chuck to come down to my Dallas office for a quick introduction.

He immediately saw the benefits of SQL over procedural file manipulation and also became an Oracle Partner.

CDS somehow happened to land a contract to handle rebates for Leaf Gum.

6.9.2 Technologies

Unix

Oracle

Oracle Forms

Oracle Reports

PL/SQL

C

Java

JSP

Graphics

Javascript

Spreadsheets

Digital Unix

Linux

Build tools

Build scripts were shell scripts.

I introduced standardized build tools, first ant and then Maven

Unit Tests

I introduced junit for unit testing code and cobertura for code coverage analysis.

Integration Tests

6.9.3 Chuck Retires

Each client had a separate database version 7 of Oracle. I consolidated 10 databases into one, eliminated database links.

There was a daemon that ran that attempted to match customers on addresses.

I utilized javautils service to call the USPS API to actually standardize the addresses for exact matches and identify undeliverable addresses

Help

Make files C Pro*C Forms

6.10 Business Model

Custom Data Services was created by my brother Chuck. I consulted to them for over 25 years and taught him everything he knew from linux to relational databases, website creation, c and java.

In 2007 I accepted the position of CIO when my brother sold the company. I gave up my consulting practice on the promise that I could work remotely after my son went to college.

After I completely revised all of the infrastructure, which was horrible in more ways than can be credibly described including the fact that they were seven years out of date on Oracle and had 14 instances running on three different machines, I upgraded to Oracle 10 and migrated everything to one instance and finally into one schema.

After I sold my house and all my possessions and moved to Costa Rica, they fired me.

6.10.1 While CIO

- Document Management
- Migrated off of Digital Unix
- Changed backup procedures
- Introduced Software Version Control
- Created development and test databases
- Trained a cafeteria worker and a pizza delivery person to program
- Coded up the most complex spreadsheets for Frito
- Introduced virtualization vmware, openvz, virtualbox
- Upgraded to Oracle 10
- Introduced Oracle Enterprise Manager
- Materialized Views

- Detachable Tablespaces
- Virtual Private Databases
- Spring inversion of control, hibernate, aspects
- Unit Testing
- Check 21 solution
- Wikis for documentation procedures
- xml
- dexterous
- workbook parser

6.10.2 Terminated

They claimed ownership of javautil despite the facts * javautil.org was owned by Jim Schmidt * javautil.com was owned by Jim Schmidt

The code base dated back to 1999.

CDS sued Jim Schmidt and claimed ownership of the code, but I would have had to have their current employees testify that the code they were using was in production at my customer sites before I joined CDS.

A settlement was reached, I allowed them to continue to use the code as it was at them time while I retained all rights.

6.11 History of Jim Schmidt and Custom Data Solutions

6.11.1 Creation

Custom Data Solutions was initially called CDS for Chuck and Dee Schmidt.

Chuck is my older brother and Dee is his wife.

Chuck left his job at Michigan National Bank to offer data processing services, initially using a TRS-80 to print mailing labels.

He subsequently bought an Alpha Micro http://www.s100computers.com/Hardware%20Folder/Alpha%20Micro/History/History.htm and wrote some business applications.

On a visit to Detroit I explained to him the benefits of Unix and got him started, installing Interactive Unix from Kodak on a PC from 133 floppy disks in 1990.

Between 1983 and 1990 I worked as Vice President of International Banking Systems for RepublicBank Dallas, as a principal at a Distribution Requirements Planning Company and then at Computer Associates where I single handedly designed and wrote the ACH settlement software system in six weeks.

In 1990 I started Trinity Technical Services with two partners and ended up buying them both out.

My first contract was the selection of an ERP system for a distributor in Texas. The chosen vendor was CIM-JIT and it was my first exposure to Oracle. I soon became more proficient in the CIM-JIT software than the vendor and flew around country to modify to delivered software to comply with commitments to customers.

In the process I became quite proficient in fixing performance problems.

Convinced that Oracle was the finest relational database, I convinced Chuck to come down to my Dallas office for a quick introduction.

He immediately saw the benefits of SQL over procedural file manipulation and also became an Oracle Partner.

CDS somehow happened to land a contract to handle rebates for Leaf Gum.

6.11.2 Technologies

Unix

Oracle

Oracle Forms

Oracle Reports

PL/SQL

C

Java

JSP

Graphics

Javascript

Spreadsheets

Digital Unix

Linux

Build tools

Unit Tests

Integration Tests

6.11.3 Chuck Retires

Help

Make files C Pro*C Forms

6.12 Business Model

6.12.1 Custom Data Solutions

http://custdata.com

Relationship

Custom Data Services was created by my brother Chuck. I consulted to them for over 25 years and taught him everything he knew from linux to relational databases, website creation, c and java.

In 2007 I accepted the position of CIO when my brother sold the company. I gave up my consulting practice on the promise that I could work remotely after my son went to college.

After I completely revised all of the infrastructure, which was horrible in more ways than can be credibly described including the fact that they were seven years out of date on Oracle and had 14 instances running on three different machines, I upgraded to Oracle 10 and migrated everything to one instance and finally into one schema.

After I sold my house and all my possessions and moved to Costa Rica, they fired me.

They stole my javautil code, there was a lawsuit, but I would have had to have their current employees testify that the code they were using was in production at my customer sites before I joined CDS.

A settlement was reached, I allowed them to continue to use the code while retaining all rights.

6.12.2 Processing Steps

Data is uploaded to the data processing service.

The file loads are analyzed at CDS using an old version of my javautil conditionidentification package.

The data is loading into prepost tables.

The data is posted

A snapshot of the data is made once a day for online web reporting in the form of spreadsheets.

Rebate programs are designed and the sales are tracked and the rebates calculated.

Rebate checks are mailed out.

6.13 Product Offering

So I can offer

- Data Reporting
- Web Site for Data
- Sales Analytics*
- Address Validation and Correction
- ACH processing for rebates
- Availability to Vending operators
- A version that runs on Frito/Pepsico laptops so they can work offline while traveling
- High level management financial analysis

6.13.1 Benefits

- Availability of site to operators for tracking rebate progress
- Internationalization works outside the U.S.

- Lower cost
  - Using open source database allows for many database servers at minimal incremental cost
  - Using Asian resources is far less expensive

6.13.2 How to get started

I simply need

- the data that is currently reported by the distributors
- product master list

No cost, no additional work, just send me the data and let me do my magic.

6.14 Project Description

This project showcases javautil.org functionality in a simple real world example.

This is a simple application the provides real functionality for sales reporting and rebate processing.

The objective is to dive into a wide variety of technologies rather than go deep into any given technology.

With a working app exploring additional features of this technology is easy.

6.14.1 Software Description

Distributors of various products report their sales as specified by the data processor, custdata.com. This software reads the data currently being reported and

- Loads the data into ETL (Extract Transform and Load) tables
- Runs a variety of tests to ensure the self consistency of the data and compliance with reporting rules.
- Posts the data
- Extracts data into various reporting formats
- Analyses Data
- Creates spreadsheets for users
- Validates customer mailing addresses and standardizes them
- Computes rebates
- Provides Vending operator spreadsheets
- Creates ACH files for submitting to a bank to issue rebates to customers

6.14.2 Tech Steps

Install the required software Configure for build Build

6.14.3 History

Processing Steps

6.15 Technologies

6.15.1 Application

- **Java**
  - JDBC
  - Hibernate
- **Build Tools**
  - Maven
- **Postgres**
  - SQL
- **Testing**
  - Unit Testing - SureFire
  - Integration Testing
- Spring * JSON RPC * Dependency Injection * Spring Boot
- YAML
- Sphinx
- javautil

6.15.2 Deployment

Amazon Web Services with high availability and security, unlimited scalability and robust administration.

Service Contracts monthly.

6.16 Record Layouts

This details the reporting format as last posted by custdata.com on their website. This is available at

```
http://web.archive.org/web/*/custdata.com
```

6.16.1 Inventory

field_name	type	offset	length	format
DISTRIBUTOR_ID	DIGITS	0	10	{:0>10s}
MFR_ID	DIGITS	10	10	{:0>10s}
MFR_PRODUCT_ID	DIGITS	20	8	{:0>8s}
COMMENTS	TEXT	28	96	{:<96s}
CASES	INTEGER	124	6	{0:06d}
BOXES	INTEGER	130	6	{0:06d}
UNITS	INTEGER	136	6	{0:06d}
CASE_GTIN	DIGITS	142	14	{:0<14s}
FILLER	LITERAL	156	12	{:<12s}
RECORD_TYPE	LITERAL	168	2	{:<2s}

6.16.2 Customer

field_name	type	offset	length	format
FILLER_00_05	LITERAL	0	6	{:>6}
CLASS_OF_TRADE	TEXT	6	4	{:<4s}
SHIP_TO_CUST_ID	TEXT	10	10	{:0>10s}
CUST_NM	TEXT	20	30	{:<30s}
ADDR_1	TEXT	50	30	{:<30s}
ADDR_2	TEXT	80	30	{:<30s}
CITY	TEXT	110	25	{:<25s}
STATE	TEXT	135	2	{:<2s}
POSTAL_CD	DIGITS	137	9	{:0>9s}
TEL_NBR	DIGITS	146	10	{:0>10s}
NATIONAL_ACCT_ID	TEXT	156	10	{:0>10s}
SPECIAL_FLG	TEXT	166	1	{:>1s}
FILLER_1	LITERAL	167	1	{:>1s}
RECORD_TYPE	LITERAL	168	2	{:>2s}

6.16.3 Inventory Total

field_name	type	offset	length	format
HEADER	LITERAL	0	10	{:>1s}
FILLER36	LITERAL	10	36	{:>36}
INVENTORY_DT	DATE	46	8	%Y%m%d
FILE_CREATION_DT	DATE	54	8	%Y%m%d
RECORD_CNT_REPORTED	INTEGER	62	9	{0:09d}
FILLER97	LITERAL	71	97	{:>97}
RECORD_TYPE	LITERAL	168	2	{:>2}

6.16.4 Customer Total

field_name	type	offset	length	format
HEADER	LITERAL	0	10	{:>10s}
FILLER_127	LITERAL	10	127	{:>127s}
CUSTOMER_COUNT	INTEGER	137	9	{0:09d}
FILLER_22	LITERAL	146	22	{:>22s}
RECORD_TYPE	LITERAL	168	2	{:>2s}

6.16.5 Sales

field_name	type	offset	length	format
DISTRIB_ID	TEXT	0	10	{:0>10s}
MFR_ID	TEXT	10	10	{:0>10s}
MFR_PRODUCT_ID	TEXT	20	8	{:0>8s}
SHIP_TO_CUST_ID	TEXT	28	10	{:0>10s}
INVOICE_CD	TEXT	38	10	{:0>10s}
INVOICE_DT	DATE	48	8	%Y%m%d
SHIP_DT	DATE	56	8	%Y%m%d
FILLER	LITERAL	64	9	{:<9s}
EXTENDED_NET_AMT	INTEGER	73	9	{0:09d}
DISTRIB_PRODUCT_REF	TEXT	82	12	{:<12s}
PRODUCT_DESCR	TEXT	94	30	{:<30s}
CASES_SHIPPED	INTEGER	124	6	{0:06d}
BOXES_SHIPPED	INTEGER	130	6	{0:06d}
UNITS_SHIPPED	INTEGER	136	6	{0:06d}
CASE_GTIN	TEXT	142	14	{:0>14s}
FILLER_12	LITERAL	156	12	{:>12s}
RECORD_TYPE	LITERAL	168	2	{:<2s}

6.16.6 Sales Total

field_name	type	offset	length	format
HEADER	LITERAL	0	10	{:>10s}
FILLER_28	LITERAL	10	28	{:>28s}
SALES_START_DT	DATE	38	8	%Y%m%d
SALES_END_DT	DATE	46	8	%Y%m%d
FILE_CREATE_DT	DATE	54	8	%Y%m%d
SALES_REC_CNT	INTEGER	62	9	{:0>9d}
SUM_EXT_NET_AMT	INTEGER	71	11	{:0>11d}
FILLER_86	LITERAL	82	86	{:>86s}
RECORD_TYPE	LITERAL	168	2	{:<2s}