

**TRAINING REPORT**

**OF**

**SIX MONTHS INDUSTRIAL TRAINING, UNDERTAKEN**

**AT**

**INFOSYS LIMITED**

**IN**

**CSE DEPARTMENT**

**ON**

**“Region Wise Sales Performance Analysis of Employees-A POC”**

**SUBMITTED IN PARTIAL FULFILLMENT OF THE DEGREE**

**OF**

**BTech (CSE)**

**Under the Guidance of:Submitted By:**

**Name:** Mr.AmrinderSingh Tarun Jindal (CUN130101379)

**Designation:** System Engineer

**Department**: ETA-BI

**ACKNOWLEDGEMENT**

It is my pleasure to be indebted to various people who directly or indirectly contributed in the development of those who influenced my thinking, behaviour and acts during the course of study.

I express my sincere gratitude to all for providing me an opportunity to undergo the Industry Oriented Hands on Training (IOHE) in Infosys Limited, Mysore.

I am thankful to “Mr. Amrinder Singh” for his support, cooperation, motivation and constant guidance provided during the training for constant inspiration, presence and blessings.

I also extend my sincere appreciation to my fellow teammates for their cooperation, our brainstorming sessions and sharing of knowledge in this project.

Lastly, I would like to thank my parents for their moral support.

REGION WISE SALES PERFORMANCE ANALYSIS OF EMPLOYEES-A POC

Undertaken at

Infosys Limited.

Mysore

**Submitted by:**

Tarun Jindal(CUN130101379)

**CERTIFICATE**

I hereby declare that the project work titled, **“Region Wise Sales Performance Analysis of Employees-A POC”** submitted as part of Bachelor’s degree in Computer Science, at Chitkara University, Rajpura, is an authentic record of our own work carried out under the supervision of Mr. Amrinder Singh.Gayatri Guddad

mRm

**Date: Verified by:** Mr.Amrinder SinghGayatri Guddad

mRm

Name(s):

Tarun Jindal

**Signature of Supervisor:**

**PREFACE**

To analyze region wise sales performance of employees.OSM Sports is a big and a famous sports equipment sellers. It was instituted in year 2010. They have their stories all over the world. People can also buy product online .They sell sports equipment of different sports and brands. A large number of people are working with this brand at different designations and across different locations. A large number of customers buy products of this brand. If the customer find the sports equipment faulty (or any other valid reason), he/she can return it. But for the last few years, their sale is going down and hence their profit is also getting reduced. So the higher management wants to know where they need to improve. For this, they need different type of reports/balanced scorecards/dashboards to analyze the sales made.

**Contents**

A.Project Undertaken 7

B. Introduction 8

C.Requirement Analysis 9-15

D. Modular description 16

E. Detailed Analysis 17-25

F. Design 26-38

G. Future Scope 39

H. Conclusion 40

I. References 41

**PROJECT UNDERTAKEN**

OBJECTIVE:

To analyze region wise sales performance of employees.

ABSTRACT:

OSM Sports is a big and a famous sportsequipment sellers. It was instituted in year 2010.They have their stories all over the world. People can also buy product online .They sell sports equipment of different sports and brands.A large number of people are working with this brand at different designations and across different locations. A large number of customers buy products of this brand. If the customer find the sports equipment faulty (or any other valid reason), he/she can return it. But for the last few years, their sale is going down and hence their profit is also getting reduced. So the higher management wants to know where they need to improve. For this, they need different type of reports/balanced scorecards/dashboards to analyze the sales made.

**INTRODUCTION**

Today, Companies all over the globe are moving their business to online market as it widens the scope of their business in a very short period of time. Sales isn’t just about the members. It’s about understanding the numbers-how to make them grow. That means every salesperson can understand customer’s need better and faster, identifying, promising leads faster and forecast result accurately, reflecting sales performance. A sales performance analysis is a way to measure sales progress over a period of time. The analysis allow the sales team to identify weakness in the sales strategy and make changes so that we can improve the results over the next period. This analysis could be used to set realistic sales performance goals for the business.

**REQUIREMENT ANALYSIS**

**Requirements in Scope:**

1. Tables should be created.

2. Cleanse the data using Informatica.

3. Data Warehouse should be created.

4. Generate reports using cognos.

5. Attractive reports should be generated using appropriate concepts, styles, patterns, etc.

**Software Requirements:**

1. DBMS: Microsoft Sql Server

2. ETL Tool : Informatica 9.5.1

3. Report Tool /Visualization : IBM Cognos 10.1 / Table

***Technologies used:***

* 1. Firstly data had to be generated. For this we have used SQL Server and MS-EXCEL
     1. **SQL SERVER:**

SQL Server 2008 (formerly codenamed "Katmai") was released on August 6, 2008, announced to the SQL Server Special Interest Group at the ESRI 2008 User's Conference on August 6, 2008 by Ed Katibah (Spatial Program Manager at Microsoft), and aims to make data management self-tuning, self-organizing, and self-maintaining with the development of SQL Server Always On technologies, to provide near-zero downtime. SQL Server 2008 also includes support for structured and semi-structured data, including digital media formats for pictures, audio, video and other multimedia data. In current versions, such multimedia data can be stored as BLOBs (binary large objects), but they are generic bit streams. Intrinsic awareness of multimedia data will allow specialized functions to be performed on them. According to Paul Flessner, senior Vice President of Server Applications at Microsoft, SQL Server 2008 can be a data storage backend for different varieties of data: XML, email, time/calendar, file, document, spatial, etc. as well as perform search, query, analysis, sharing, and synchronization across all data types.

Other new data types include specialized date and time types and a spatial data type for location-dependent data. Better support for unstructured and semi-structured data is provided using the new FILESTREAM data type, which can be used to reference any file stored on the file system. Structured data and metadata about the file is stored in SQL Server database, whereas the unstructured component is stored in the file system. Such files can be accessed both via Win32 file handling APIs as well as via SQL Server using T-SQL; doing the latter accesses the file data as a BLOB. Backing up and restoring the database backs up or restores the referenced files as well. SQL Server 2008 also natively supports hierarchical data, and includes T-SQL constructs to directly deal with them, without using recursive queries.

The full-text search functionality has been integrated with the database engine. According to a Microsoft technical article, this simplifies management and improves performance.

Spatial data will be stored in two types. A "Flat Earth" (GEOMETRY or planar) data type represents geospatial data which has been projected from its native, spherical, coordinate system into a plane. A "Round Earth" data type (GEOGRAPHY) uses an ellipsoidal model in which the Earth is defined as a single continuous entity which doesnot suffer from the singularities such as the international dateline, poles, or map projection zone "edges". Approximately 70 methods are available to represent spatial operations for the Open Geospatial Consortium Simple Features for SQL, Version 1.1.

SQL Server includes better compression features, which also helps in improving scalability. It enhanced the indexing algorithms and introduced the notion of filtered indexes. It also includes Resource Governor that allows reserving resources for certain users or workflows. It also includes capabilities for transparent encryption of data (TDE) as well as compression of backups. SQL Server 2008 supports the ADO.NET Entity Framework and the reporting tools, replication, and data definition will be built around the Entity Data Model. SQL Server Reporting Services will gain charting capabilities from the integration of the data visualization products from Dundas Data Visualization, Inc., which was acquired by Microsoft. On the management side, SQL Server 2008 includes the Declarative Management Framework which allows configuring policies and constraints, on the entire database or certain tables, declaratively. The version of SQL Server Management Studio included with SQL Server 2008 supports IntelliSense for SQL queries against a SQL Server 2008 Database Engine. SQL Server 2008 also makes the databases available via Windows PowerShell providers and management functionality available as Cmdlets, so that the server and all the running instances can be managed from Windows PowerShell.The final SQL Server 2008 service pack (10.00.6000, Service Pack 4) was released on September 30, 2014.

* 1. The next step is to fetch theNecessary details of the product ordered by the customer

Including the details like which employee delivered the product, on which date and the total bill amount. This is done based on some functional requirements. For this a separate target table is designed fetching all the required data using INFORMATICA.

3.2.1**INFORMATICA**:

Informatica is a widely used ETL tool (Extract, Transform and Load) for extracting the source data and looking it into the target after applying the required transformation. Serving as the foundation for all data integration projects, the Informatica Platform lets IT organizations initiate the ETL process from virtually any business system, in any format. As part of the Informatica Platform, Informatica PowerCenter delivers robust yet easy-to-use Enterprise ETL capabilities that simplify development and deployment of database.

Steps used to create a simple Integration sample:

* Open PowerCenter Designer and configuring Domains by providing Domain Name, Gateway host and Gateway port
* Select Repository in which user want to work and provide Login Credentials
* Then in source analyzer import source and in target designer design target table as per requirements and then create ODBC connections and new User DSN. Select required ODBC connections and generate target table.
* In Mapping Designer Drag and drop source and target. Then add transformations like sequence generator, union, joiner, rank ,filter etc. as per user requirements
* Open PowerCenter Workflow Manager and create session and provide source and target path of tables used in mappings .If target table is database then provide database server name.

Then create workflow and run workflow and check session logs in PowerCenter Workflow monitor.

It is a service oriented Architecture and has the capability to share service and resource across several machines. Informatica ETL product, known as Informatica Power Center consists of 3 main components.

1. Informatica Power Center Client Tools: These are the development tools installed at developer end. These tools enable a developer to
   * Define transformation process, known as mapping. (Designer)
   * Define run-time properties for a mapping, known as sessions ([Workflow Manager](https://en.wikipedia.org/wiki/Workflow_management))
   * Monitor execution of sessions (Workflow Monitor)
   * Manage repository, useful for administrators (Repository Manager)
   * Report Metadata (Metadata Reporter)
2. Informatica Power Center Repository: Repository is the [backend](https://en.wikipedia.org/wiki/Backend) of Informatica tools. Repository is a data inventory where all the data related to mappings, sources, targets etc. is kept. This is the place where all the metadata for an application is stored. All the client tools and Informatica Server fetch data from Repository.
3. Informatica Power Center Server: In the Server, all the executions take place. Server makes physical connections to sources/ targets, fetches data, applies the transformations mentioned in the mapping and loads the data in the target system.

### Types of transformations

* **Active**: One that can change the number of rows being output.
* **Passive**: One that cannot change the number of rows being output.
* **Connected** : One that has at least one link with other transformations or objects in a mapping
* **Unconnected** : One that does not have any link with any transformations or objects in a mapping

Also, Active transformation are recognized by any of the following three factors:

* the transformation changes the row type
* the transformation changes the row order
* transaction control happens in the transformation

A transformation which has these properties is said to be an "active transformation", otherwise it is said to be a "passive transformation".

Some of the transformations are:

* Source qualifier - converts source data types to Informatica data types; joins two source tables or files of the same type
* Expression - transforms data with functions without affecting cardinality or data order
* Lookups - enhances data elements with lookup values; may be connected or unconnected
* Filter - filters incoming data
* Router - routes data to selected output paths
* Sorter - sorts data and optionally removes duplicate rows
* Aggregator - [aggregates data](https://en.wikipedia.org/wiki/Data_aggregation)
* Joiner - [joins](https://en.wikipedia.org/wiki/Join_(SQL)) relational data midstream; can be used when source data is of multiple types
* Normalizer - used with flat file input containing arrays or repeated data to create [normal form](https://en.wikipedia.org/wiki/Normal_form_(databases)) relational data
* Update Strategy - determines whether data is to be updated, inserted or deleted in cases where there are multiple output targets
* Transaction Control - creates transactions and controls commits and [rollbacks](https://en.wikipedia.org/wiki/Rollback_(data_management))
* Sequence Generator - generates a sequence of unique values as per the specified parameters.
  1. Lastly we had to generate report using Report Studio Cognos.
     1. **COGNOS:**

IBM Cognos Business Intelligence is a web-based, integrated business intelligence suite by IBM. It provides a toolset for reporting, analysis, and monitoring of events and metrics. The software consists of several components to meet the different information requirements in a company.A successful reporting platform implementation in a business intelligence environment requires great attention to be paid from both the business end users and IT professionals.   
The fact is that the reporting layer is what business users might consider a data warehouse system and if they do not like it, they will not use it. Even though it might be a perfectly maintained data warehouse with high-quality data, stable and optimized ETL processes and faultless operation. It will be just useless for them, thus useless for the whole organization.

Steps used to create a simple report using:

* + Create a new and select Report Model Project from the Project Wizard.
  + Next we need to create a data source for the project, we just need to click on edit to specify the server name and database that will be used from that server. The connection string is automatically created.
  + In the IBM Cognos Welcome page, click Advanced reports to open Report Studio.
  + As per user requirements create a new report which could be list, crosstab, pie chart, bar chart etc.
  + Insert data to user report and apply formatting to reports as per requirements
  + We can also change properties and apply join and other operation to query

Then run reports and observe output.

* IBM Cognos Business Intelligence is a web-based, integrated business intelligence suite by IBM. It provides a toolset for reporting, analysis, and monitoring of events and metrics. The software consists of several components to meet the different information requirements in a company.

A successful reporting platform implementation in a business intelligence environment requires great attention to be paid from both the business end users and IT professionals.   
The fact is that the reporting layer is what business users might consider a data warehouse system and if they do not like it, they will not use it. Even though it might be a perfectly maintained data warehouse with high-quality data, stable and optimized ETL processes and faultless operation. It will be just useless for them, thus useless for the whole organization.

**MODULAR DESCRIPTION**

**MS EXCEL**

5.1.1 Creation of data in unstructured format in different categories like employee data, sales data, returned item data, sales data, store data.

**MS SQL 2012**

5.2.1 Import the data files (.txt, .xml etc.) from various sources in Informatica and convert these flat files into database files.

**INFORMATICA**

5.3.1 After importing all the files in Informatica, remove various anomalies such as commas, semi colons, colons and convert the flat file to database files.

5.3.2 Modify the target table as per the requirements such as adding or removing attributes, assigning primary keys or foreign keys

5.3.3 In the mapping section map the source table to the target table using a wide array of tools such as router, filter, lookup, union, join.

5.3.4 Create a task in the workflow manager, making changes in the mapping and file source path, presql, serial and parallel processing. Run the workflow.

**IBM COGNOS FRAMEWORK MANAGER**

5.4.1 Design the fact table and dimension table in the framework manager. The fact table consists of the purchase table whereas the dimension table consists of productid, orderdate, year among others.

**IBM COGNOS REPORTING STUDIO**

5.5.1 Import the package containing fact and dimension table present in the presentation view created in the framework manager

5.5.2 Create effective reports depicting sales products category wise, payment method wise, year wise, total sales in a particular year, total sales.

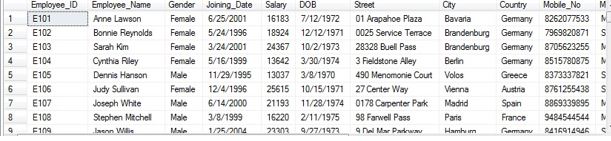
**DETAILED ANALYSIS**

**TABLES:**

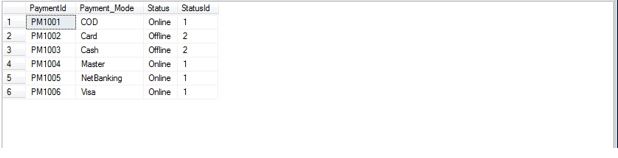
1) DimCustomer:

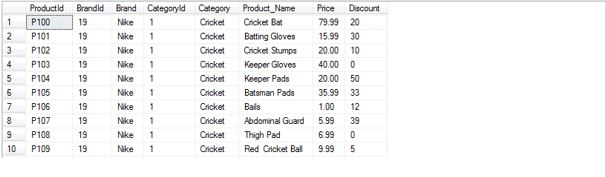


2) DimEmployee:

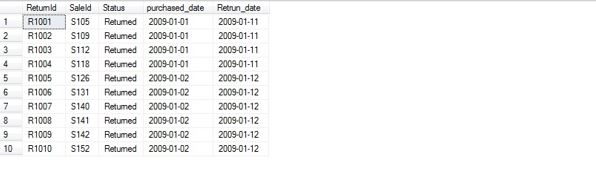


3) DimOrderDate:

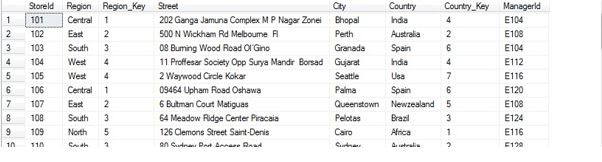
4) DimPayment:

5) DimProduct:

6) Dimreturn:

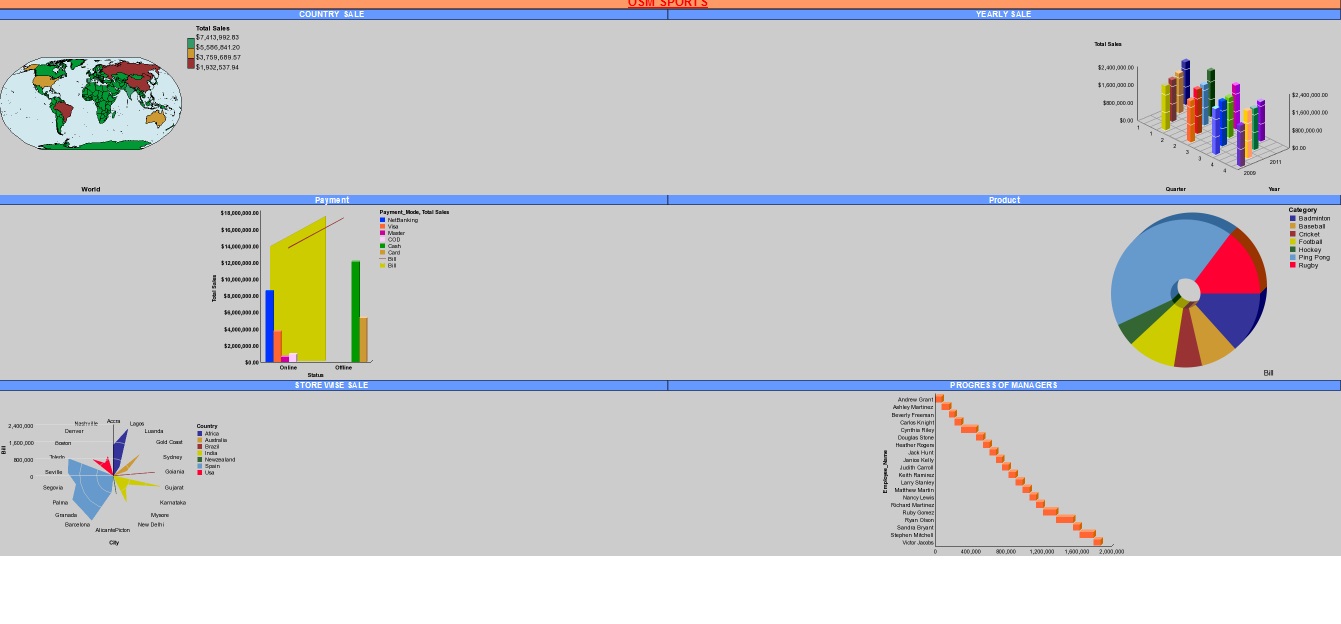


7) DimSales:

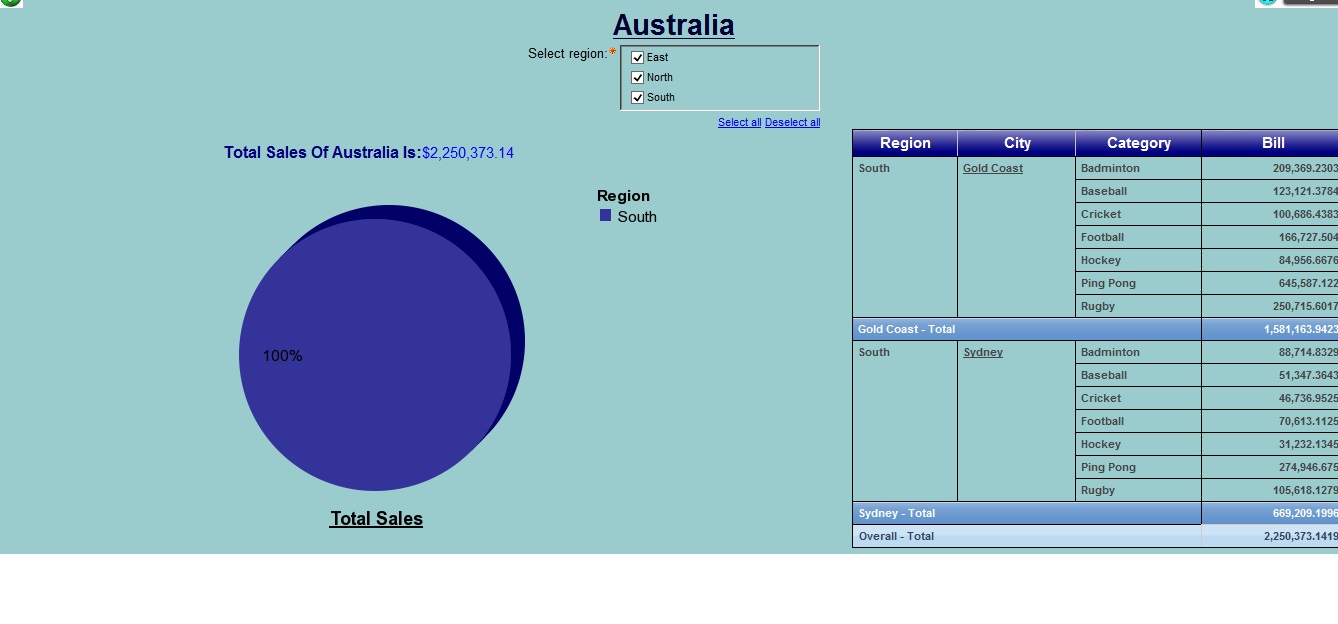
8)DimStore:

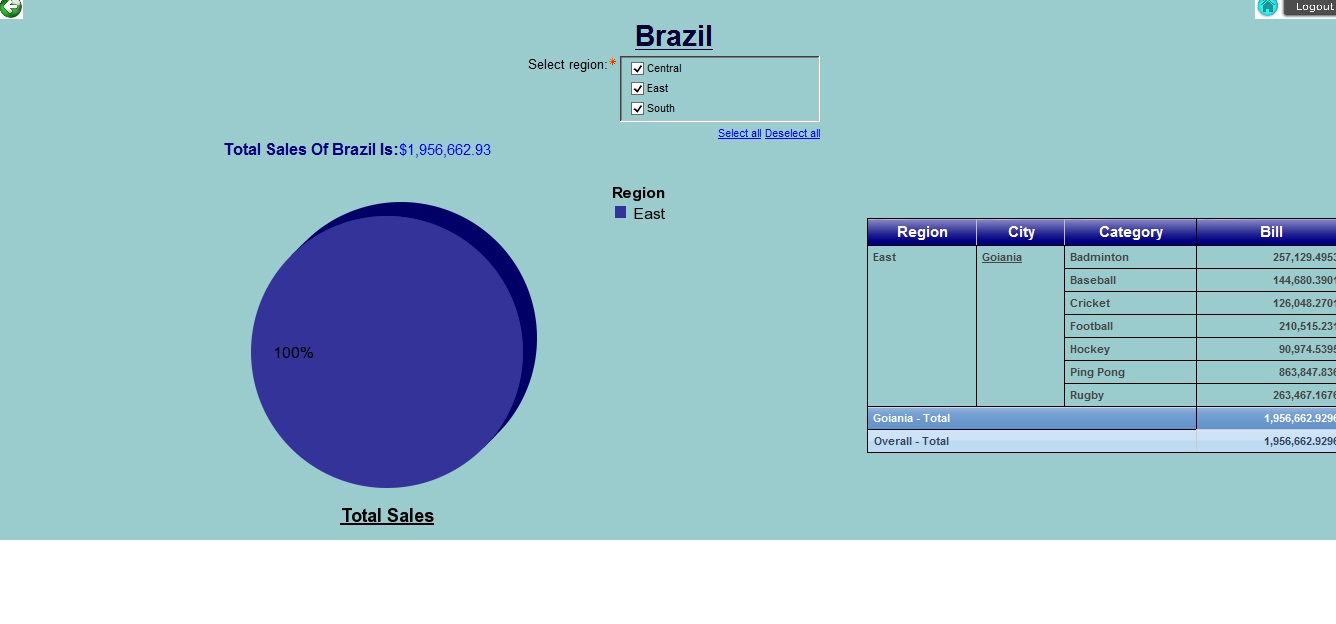
9)FactSales:

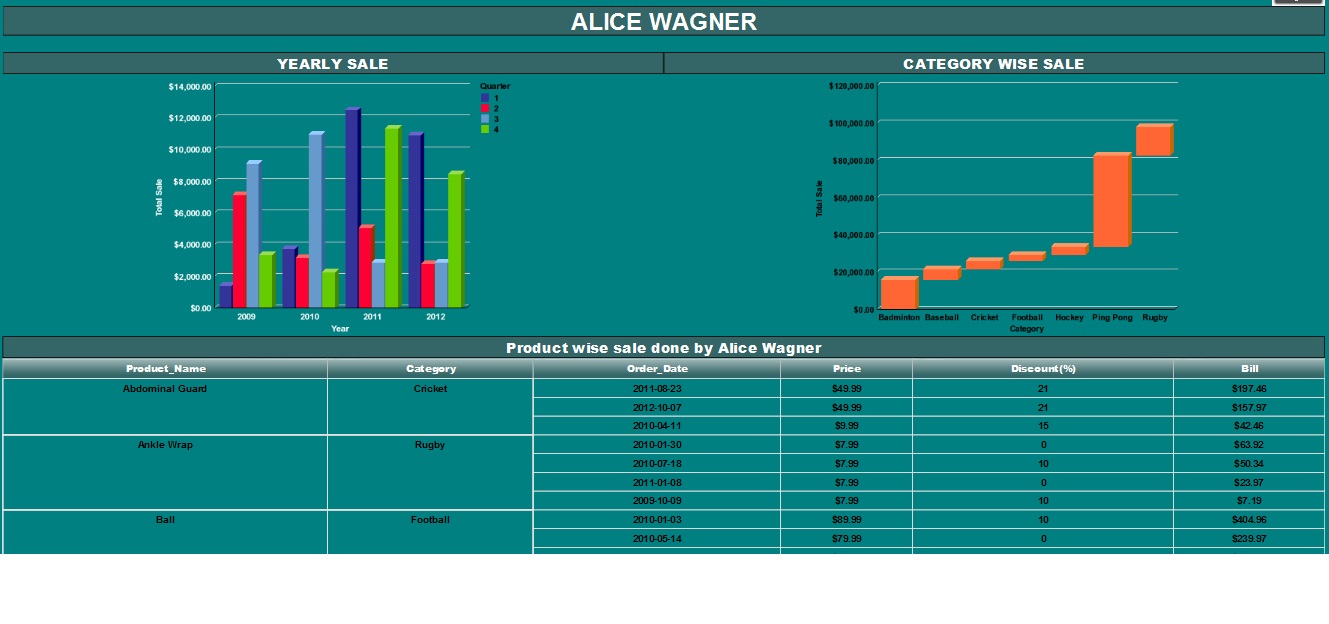
**DESIGN**

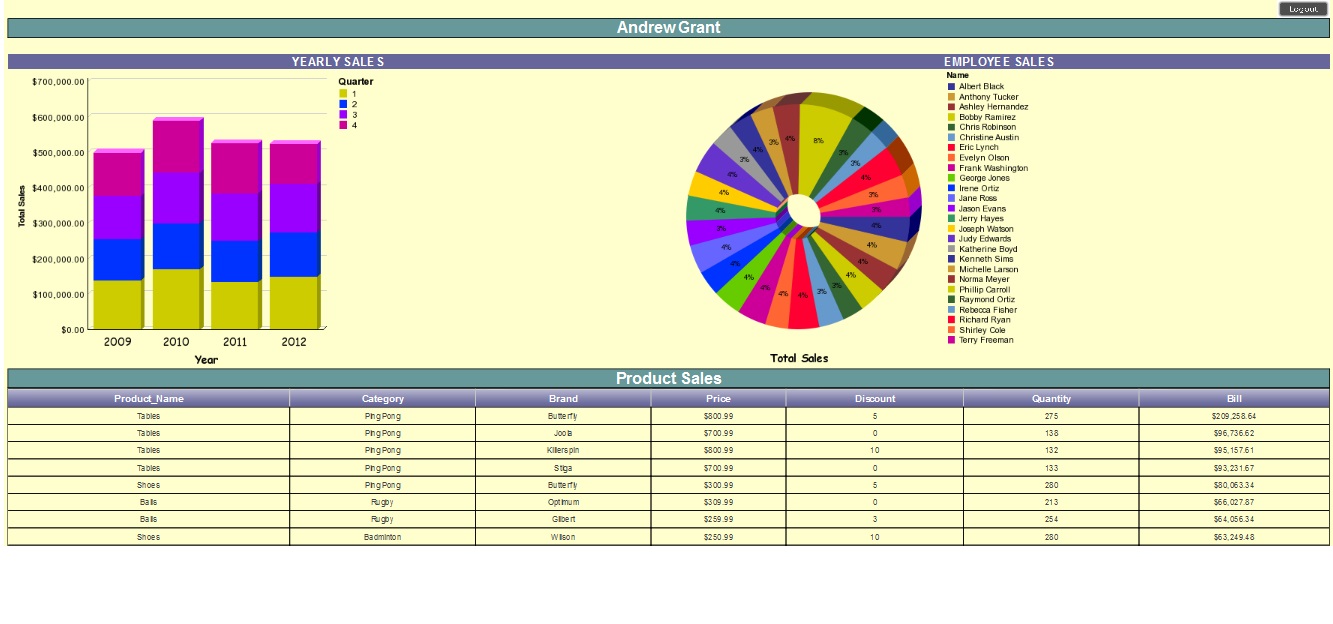


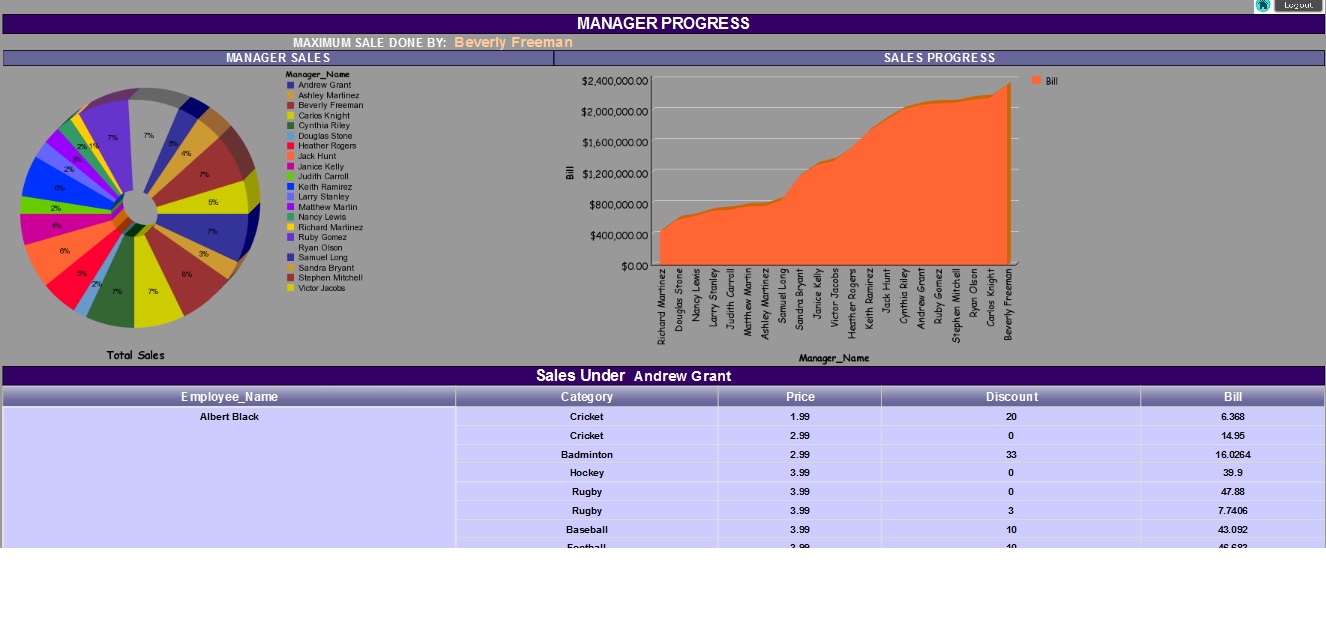
# **C:\Users\tarun02.trn\Desktop\africa.jpg**

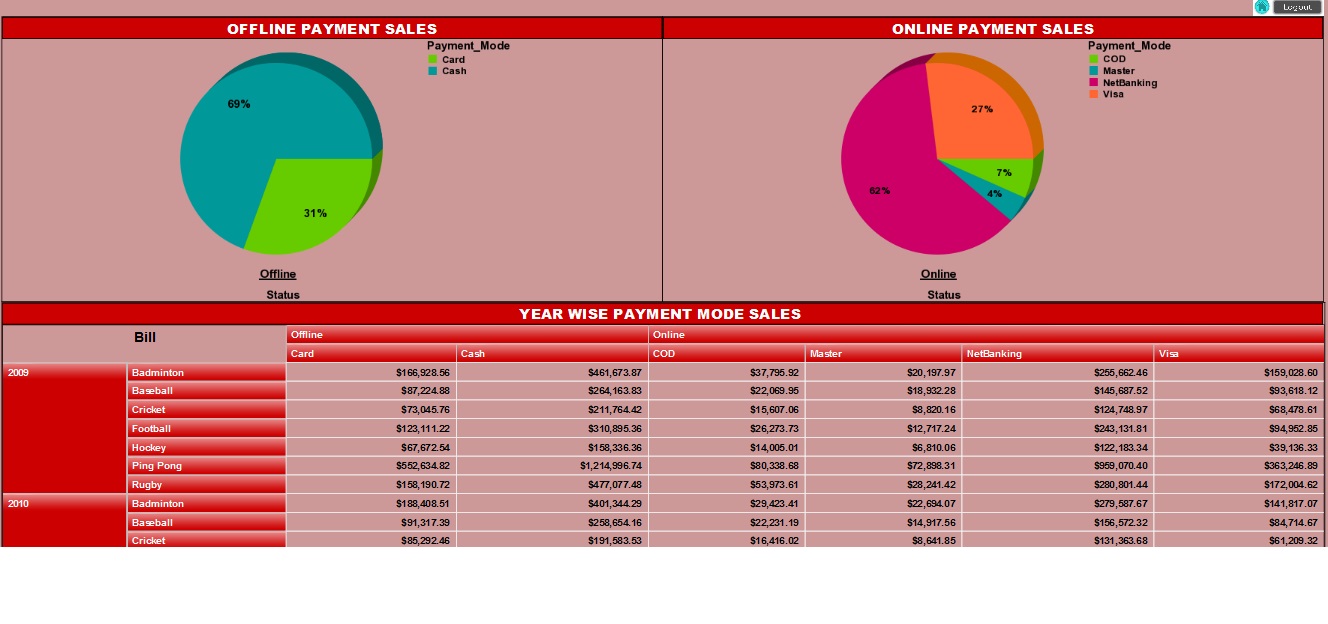


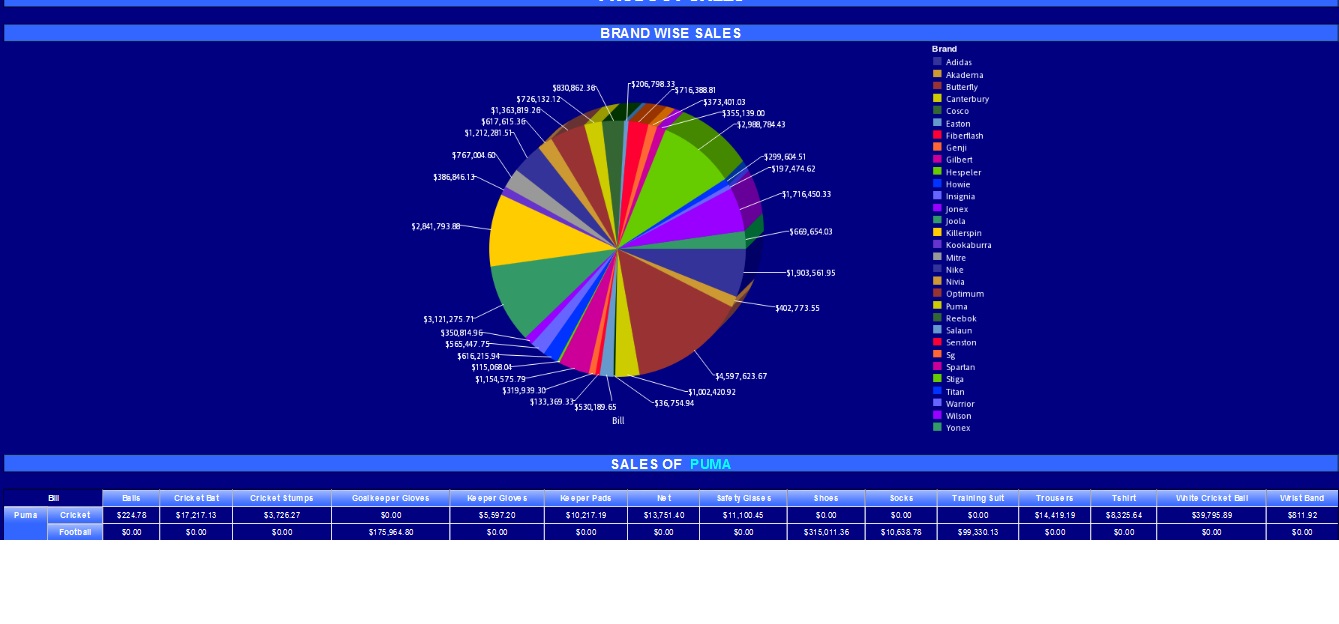


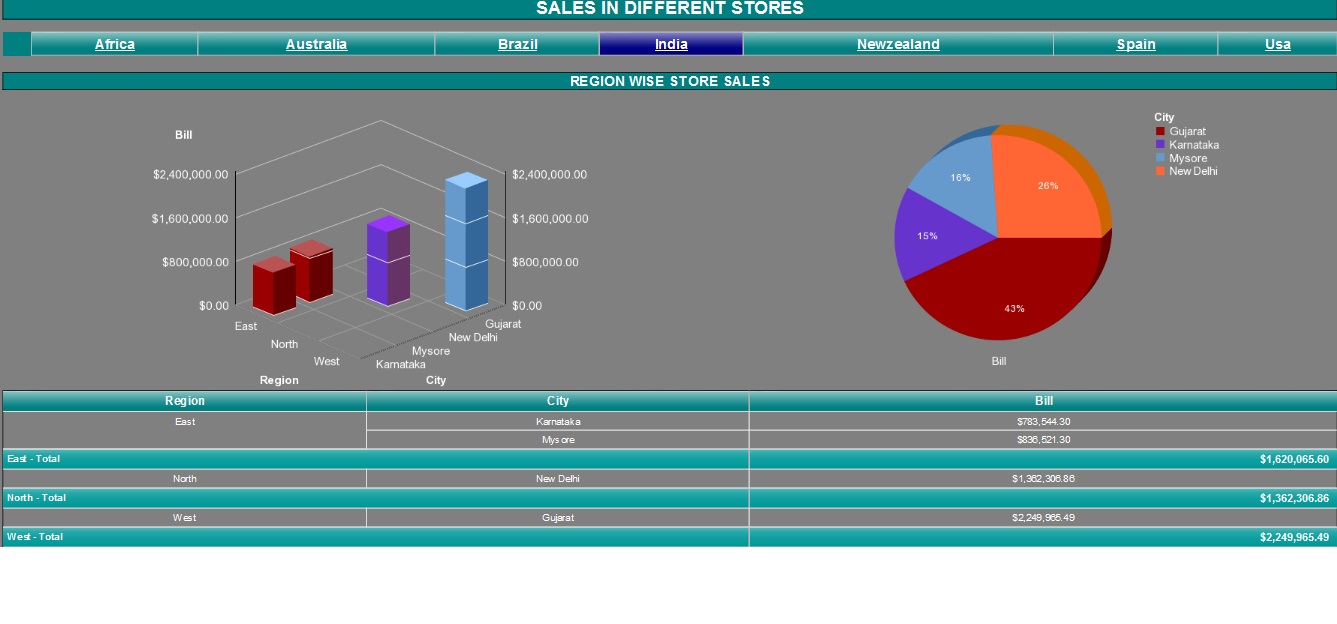


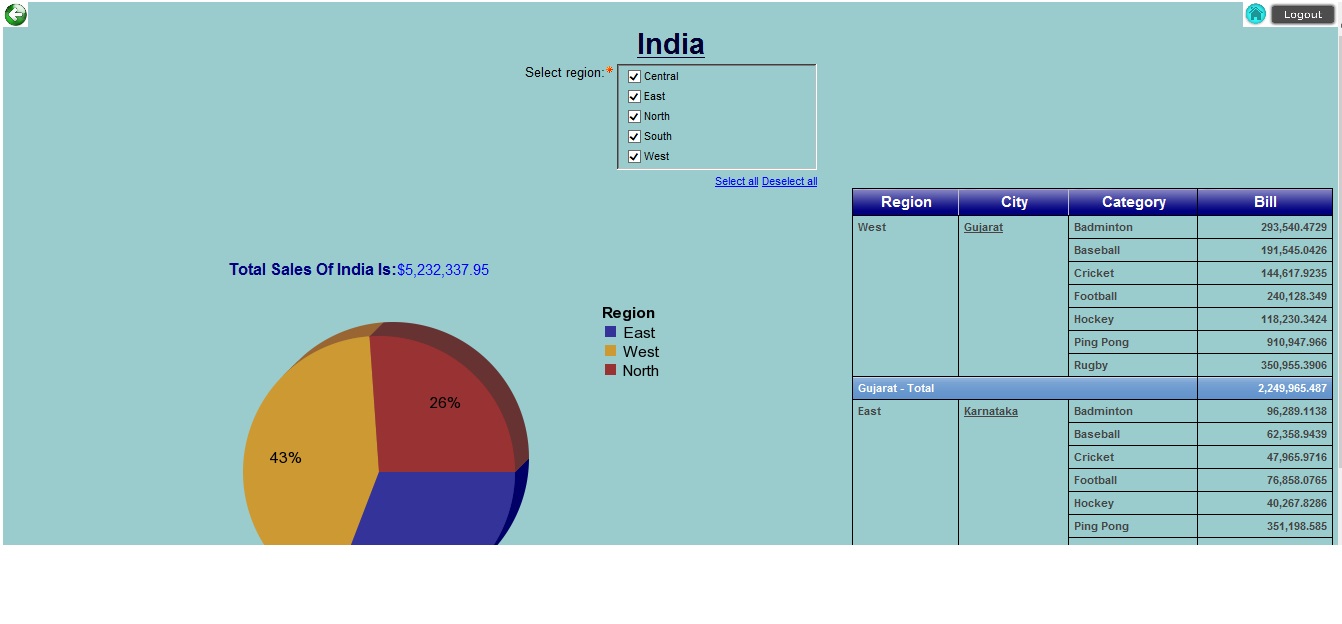


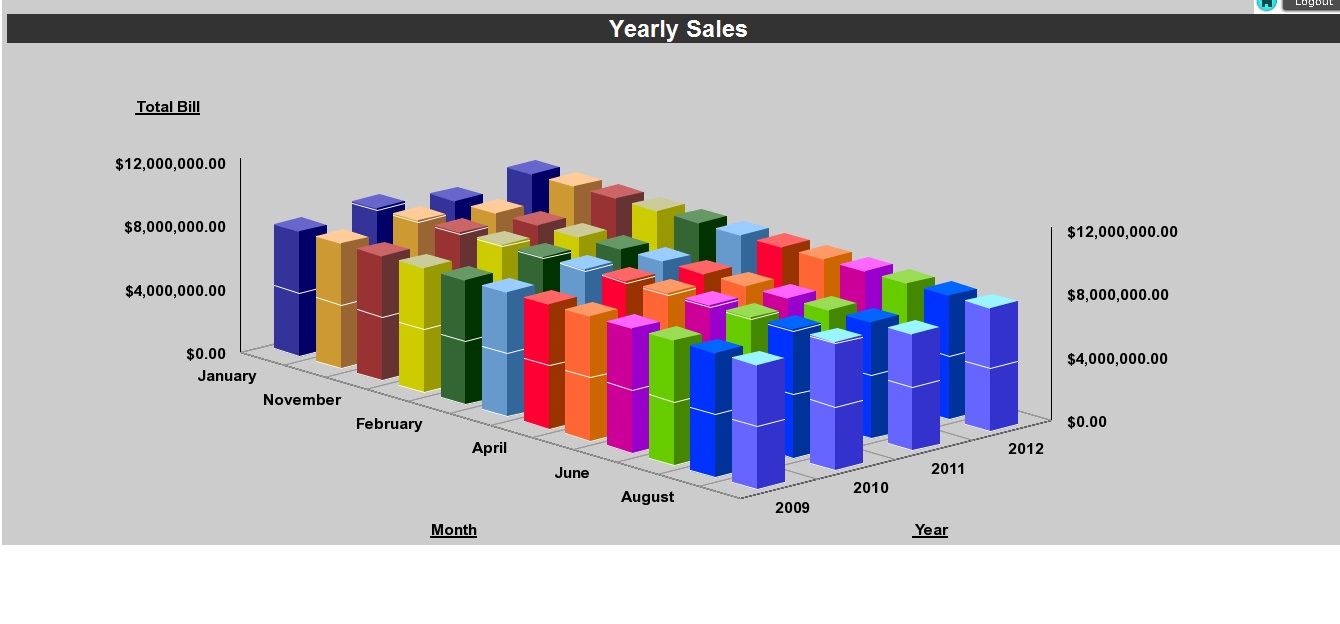






****

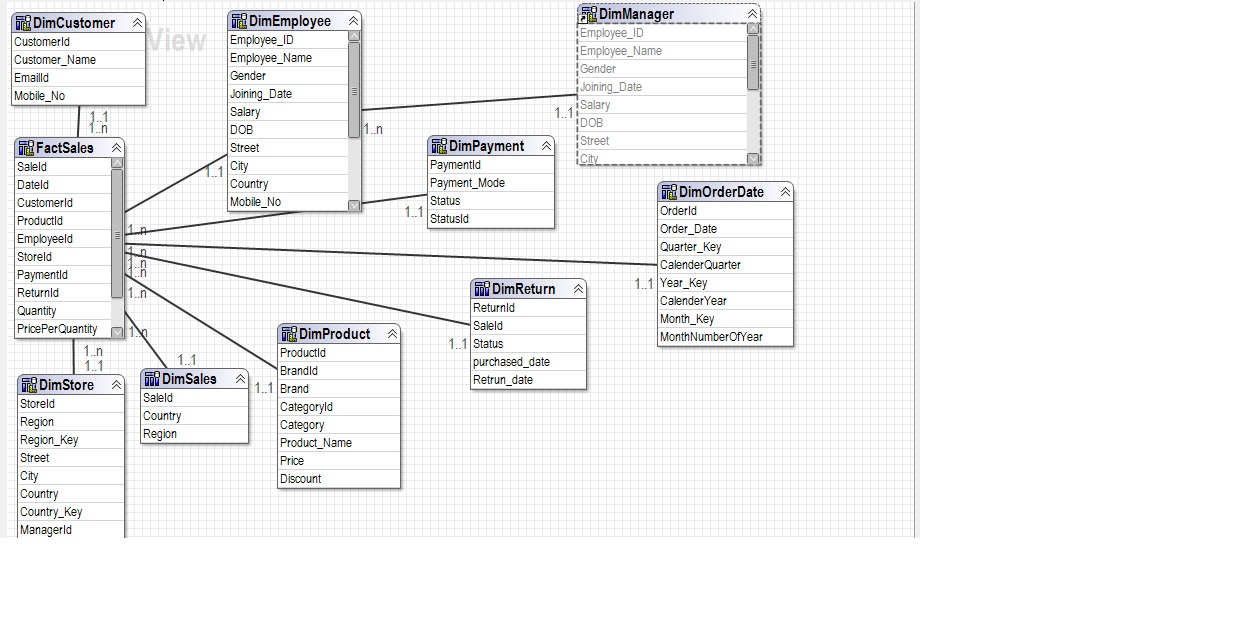






****

Framework Manager(Physical View):



***FUTURE ENHANCEMENTS***

* Improve quality of the products delivery of context and offers that increase the likelihood of purchase.
* To gain deeper understanding of customer’s needs, preferences and understanding.
* To analyze which store has made maximum profit in a particular year according to which sales of other stores can be improved.
* To check which product is in maximum demand, so that discounts can be offered on other products so as to maximize their sales too.

***CONCLUSION***

This project explains about the analysis of the data that was gathered through different sources. This analysis of data helps the company in removal of any backlogs. We come to know about the sales in different years, profit/loss for different time period, leading to sales improvement of their business. This results in letting us know which store made the maximum sales in a particular year, which employee had done the maximum sale etc.

# ***REFERENCES***

1. <http://etl-tools.info/informatica/tutorial.html>
2. <http://www.cognos-bi.info/tutorial.html>
3. <https://en.wikipedia.org/wiki/Business_intelligence>
4. <https://www.informatica.com/in/products/data-quality.html>
5. <http://www.folkstalk.com/2011/12/transformations-in-informatica-9.html>