HUMAN CAPITAL DATAWAREHOUSE

Dr.BH.Nagarajasri,
In-Charge Director, SVU Computer Centre

Abstract :
   This paper is about the development and operational analysis report for business user using various components of SAS (statistical Package)serving the growing needs of human capital Data Ware House in Information Technology industry. This paper is a guideline of Human Resource Management System Data Ware House.

   In this paper we propose a developments of Global HR data Mart – POC(Proof of concept)that appropriately source its data from the host of existing data sources its data from the Host of existing data sources and make the data available globally for viewing and analysis. It would provide single a windows for organizational HR data that cut across the multitudes of systems and repositories thus ensuing the consistency in management reports. It also support tp run adhoc quries in an efficient manner and generate analysis and operational reports in prevent report formats like PDF Excel and HTML and SAS EG and SAS web report Studio.

1. INTRODUCTION

1.1 Purpose of Project citation

   This paper contains the overall definition of the Project and is used to provide a common understanding of the Project to all business users and team members in the organization. This includes the development of Operational Analysis reports for business users using various components of SAS, SPSS, SAP serving the growing needs of human capital.

   The Project character contains the following information.

> Why the Project is undertaken
> The Scope of the Project

   The Document describes how the Project is executed –i.e. the Project Organization high level Architecture and end user deliverables. This document is a guideline of HRMS DW Human Resource Management System Data Ware House Project.

1.2 The Objective

   Development of global HR Data Mart –POC that appropriately source its data from the host of existing data sources and make data available globally for viewing and analysis. It may provide a single window for organizational HR data that cuts across the multitude of the System and repositories thus ensuring data consistency in management reports.

> Run adhoc queries in and efficient manner and generate analysis and Operational Reports in prevalent report formats like PDF Excel and HTML using SAS EG SPSS, SAP. and SAS web port Studio

1.3. SCOPE

- Integration of HR Source Systems to support HR Data Mart
- Create a Platform for HR operational reporting e.g Resource Management
- Build HR Data Mart
- Create a Platform for HR Analytical reporting e.g. Employees attrition Report w.r.t Training, department and reason for resign

EXECUTIVE SUMMARY

a. Business Challenges

   IT INDUSTRY HR data is scattered in various applications such as People Soft Payroll system and other satellite systems from various HR Support departments like CRP, Training and Corporate HR.

   This Poses the following Problems

   Data across these systems is not consistent and there is no way of producing a single view of data

   Systems are sometimes maintained by different entities, leading to uncontrolled versioning of data.

Due to the limitation in nature of such systems , it is cumbersome and sometimes not possible to adequately analyze the data and create reports that meet the requirement of the organization's Senior Management.

b. Implementation steps in Data Ware Housing

1. **Requirement analysis and capacity planning**

To overcome the business challenges mentioned above , IT industry has envisaged the need for a system that will provide a Single Window for Organizational HR data and cuts across the multitude of systems and repositories , thus ensuring data consistency in management reports. Offer a powerful analysis engine that would help in performance management resource management and employee development management through Turnover Analysis Cost Analysis Workforce Analysis and Training and Development Analysis. The first step in data warehousing involves defining enterprise needs defining architecture ,carrying out capacity planning and selecting the hardware and software tools. This step involves consulting the Senior Management as well as with the various Stake Holders.

2.Hardware Integration : Once the hardware and software have been selected they need to be put together by integrating the servers, the storage devices and the client software tools.

3.Physical Modeling : Modeling is a major step that involves designing the warehouse schema and views. For the data warehouse to perform efficiently physical modeling required. This involves designing the physical data ware house organization data placement data partitioning deciding on access methods and indexing.

4.Sources: The data from the source systems will need to go through an ETL Process. This step of designing and implementing the ETL process may involve identifying a suitable ETL tool vendor and purchasing an implementing the tool.

5,ETL : The data from the source systems will need to go through ETL process

The appropriate ETL tool must be identified to suit the needs of the enterprise.

6..Populate the data warehouse Once ETL tools have been agreed upon testing the tools will be required perhaps using the staging area. Once everything is working satisfactorily the ETL tools may be used in populating the data ware house given the schema and view definitions.

7.User Applications For the data ware house to be there must an end-user applications. This step involves designing and implementing applications required by the end users.

8..Roll out the warehouse and applications : Once the data warehouse has been populated and end-user applications tested. The warehouse system and the applications may be rolled out for the user community to use.

2.Proposed Data Ware House Solution

2.1 Business Requirements

To meet the Business Challenges IT Industry has identified the need to put in place and integrated data warehouse that the robust and scalable and provide an Enterprise wide reporting framework . The framework will address the reporting

CORPORATE HR	Standard /Parameterized Reports New Jionees details Employee's Details (Name Department Grade etc) Employees on deputation with location and allowance details Parameter Driven report providing details of employee(s)on basis of parameter value Ad-hoc Reports Any Reports that may be required on an ad-hoc basis
Unit HR	Standard /Parameterized Reports Secondment and Transfer Details List of People who have valid work permit and visa Leave Balance Details Loan Balance Details Seating arrangements details Ad-hoc Reports Any reports that may be required on an ad-hoc basis
CRP	Standard Parameterized Reports Resource Summary-Utilization Report Vertical wise resource List Bench analysis report

	Resource Allocation Sheet(RAS)
	RRF Tracker
	Ad-hoc Reports
	|Reports similar to those mentioned above but of an ad-hoc
Training and Development	Standard/Parameterized Reports
	Report on number of people trained by vertical technology
	Training Details by people, by Practice, by number of Hours
	Details of joinees not inducted within a month
	Individual's appearance in QCP exam, number of attempts result of QCP exam
	With alert generation (179 days from DOJ for Laterals 364 for Campus recruits) Quality certification program)
	Number of Hours of Training by individual, by Technology by time period
	Details of vendors providing Training with timelines and skill
	Sets
	Flags for QCP (lateral and campus recruits
	Pass and/or fail trend in QCP by time period (for ex. In the last 6 months) Automatic Calculation of the Date of Confirmation QCP deadline by inputting Date of joining
	SLA Matrix calculations on quarterly basis to conform to QMS
	Ad-hoc Reports
	Reports on training Conducted
	Ad-hoc request for department specific report requirement

In addition to providing for standard parameterized and ad-hoc reporting the framework will also enable analysis such as Turnover analysis Cost Analysis Workforce Analysis and T& D analysis
Data Ware Housing provides architectures and tools for business executives to systematically organize understand and user their data to make Strategic decisions.

According to W.H. Inmon "A Data Ware House is subject oriented integrated time variant no violent and nonvolatile collection of data in support of management's Decision Making Process."
The Primary aims in building a Data Ware House are to provide a single version of truth about the enterprise information and good performance for ad-hoc management queries required for Enterprise Analysis to manipulate animate and synthesize enterprise information. Modern Decision making requires strategic use information.

2.2 Implementation Guidelines
We present a number of guidelines for successful implementing a Data Ware House There are general guidelines not all of them applicable to every data warehouse project.
1. **Build incrementally**: Data warehouses build incrementally .In general it is recommended that the data mart may first build with one particular project in mind and once it is implemented a number of other sections of the enterprise may also wish to implement similar systems.
2. **Need a champion**: A data warehouse Project must have a champion who is willing to carry out considerable research into expected costs and benefits of the Project. Data ware Housing Project require inputs from many units in an enterprise and therefore need to be driven by someone who is capable of interacting with people in the enterprise and can actively persuade colleagues. Studies have shown that having a champion can help adoption and success of data warehousing projects.
3. **Senior Management Support** : A data warehouse Project must be fully supported by the Senior management. This can be sometimes be difficult since it may be hard to quantify the benefits of data warehouse technology and the managers may consider it as a cost without any explicit return on investment. Data ware housing project studies show that the management support is essential for the success of data warehousing project.
4. **Ensure Quality** :Improved data quality, when recognized by the senior Managers and Stake holders is likely lead to improved support for the Data Ware House Project.
5. **Corporate Strategy**: A data Ware House project must fit with Corporate Strategy
and business objectives.
6. Business Plan: The financial costs expected benefits and a project plan for a data ware house project must be clearly outlined and understood by Stake Holders.

7. Training; For a data ware house Project to be successful , the users must be trained to use warehoutse and to understand its capabilities. Training of users and Professionals development of the Project team also be required since data warehousing is a complex task and skills of the project team are crucial to the success of the Project.

8. Adaptability: The Project should build in adaptability so that challenges may be made to the Data Ware House if and when required. Like any system a data warehouse will need to change as an enterprise's needs change.

9. Joint Management: The Project must managed both IT and Business professionals in the Enterprise. To ensure good communication with the stake holders and that the project is focused on assisting the enterprise's business professionals must be involved in the project along with the technical professionals.

2.3 SAS Data warehouse and Data Mining Solution
Above mentioned requirements are best scattered by deploying:
SAS ® Business Intelligence Server , and
SAS® Enterprise Guide

The first step in the implementation will involve designing the Logical data Model based on the Reporting requirements. The Logical data model will next be converted into a Physical Data Model in Base SAS which will be managed with the SAS SPDE(Scalable Performance Data Base Engine).The Physical Data Model implementation will provide the basic foundation layer for creating a reporting framework consisting of standard parameterized and ad-hoc reports which form part of the solution. The reporting framework will be configured on the SAS Business Intelligence Server

ARCHITECTURE

Enterprise Guide   Management Console |Information Map studio
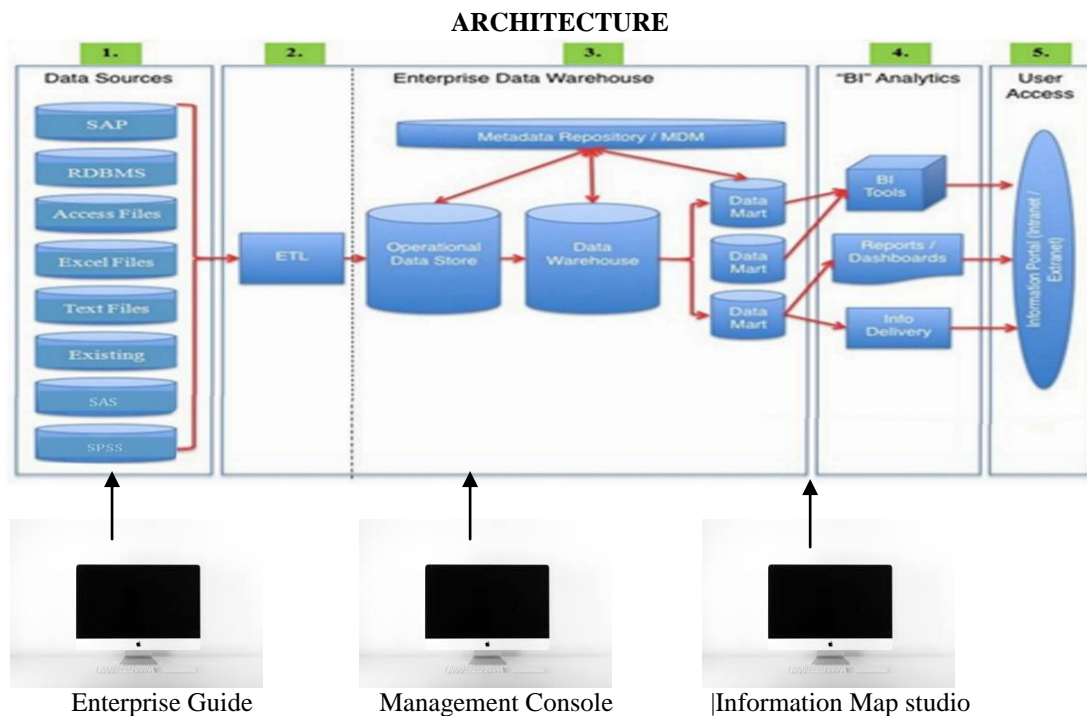
An ODS (operational Data Store) has been defined by Inmon and Imhoff as subject-oriented, integrated,volatile, current valued datastore,contianing only corporate detailed data. For the implementation an ODS just like implementing any database system a data model should be developed.

Populating an ODS involves an acquisition process of extracting transforming and loading data from Data Source systems. This Process called ETL. In practice this process more complex and tedious and may significant resources to implement. The ETL process consist of data extraction from source systems data transformation which includes data cleaning , and loading the data warehouse.
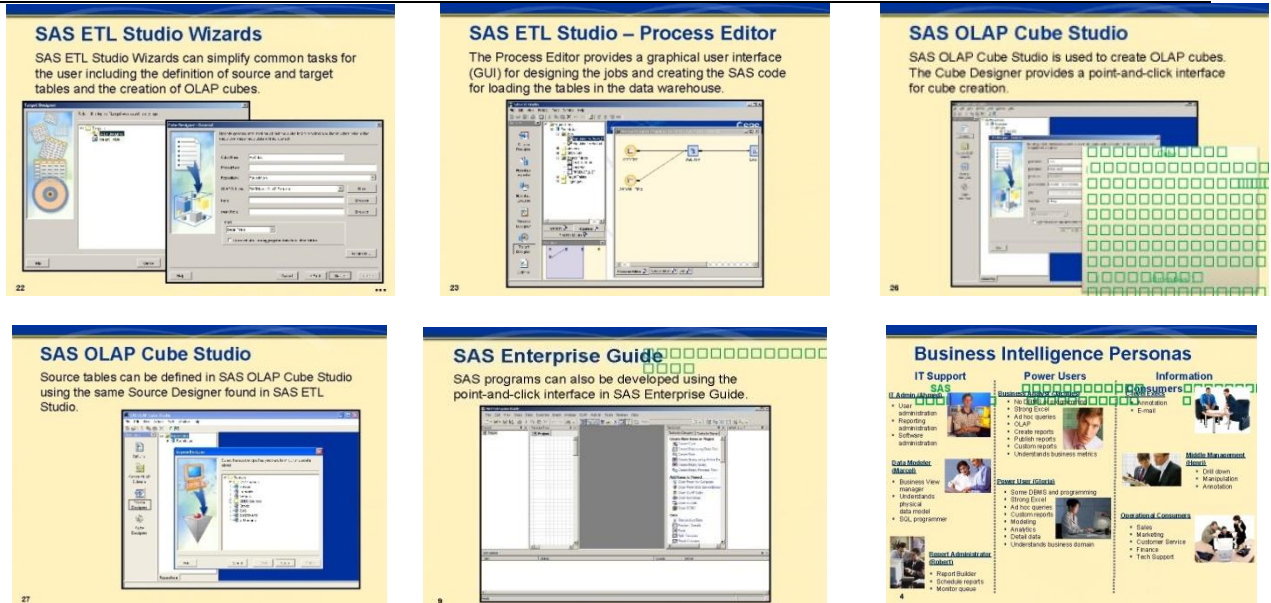
SAS ETL Studio enables you to perform the following tasks:

- the **E**xtraction of data from operational data stores
- the **T**ransformation of this data
- the **L**oading of the extracted data into your data warehouse or data mart.

Below given are some important SAS applications:

Industry	Usage
Pharmaceutical	Statistical Analysis, Reporting
Telecom	ETL, Reporting, Data Mining, Forecasting
Financials	ETL, Reporting, Data Mining, Financial research
Predictive modeling	DBMarketing, Activity-Based Management
Healthcare	ETL, reporting, Data Mining

SAS Vs. R

SAS	R
SAS is commercial software, so it requires a financial investment.	R is open source software. Hence anyone can use it.
SAS is an easiest analytical tool to learn. Even people with limited knowledge of SQL can learn it quickly.	R requires you to write complicated and lengthy codes.
SAS is a highly preferred choice by big companies and is quite technically advanced & user-friendly.	R is fast developing software; however, you need to keep upgrading it.

SAS has good graphical support but does not offer any customization.	Graphical support of R tool is very poor.

Advantages of SAS

* SAS has an easy syntax which can be learned without any type of programming knowledge
* Ability to handle a large database with ease
* SAS is a very comprehensible language which can easily be debugged
* Its "log" window clearly states the error which helps you to debug your code easily
* SAS helps you thoroughly test and analyze the algorithm
* SAS is completely secured so you can't extract without a license in the office
* Makes statistical computing easier for the non-programming users
* Handles large size database effectively.

2.4 Human Data Ware House Methodology & Approach

The Phases :

Requirement analysis Phase: collects the Business requirements of the user for data reporting and analysis. Provide structured documentation of requirements for easy validation for Business user.

Design Phase: Physical data Model is frozen using the logical data model and data volumetric analysis .Authentication and access requirements for the business users are also incorporated in the design,

Construction and Testing Phase: The Physical data model realized through creation of the database and schema objects. The mapping , sessions and reports are developed and tested. All components of the Data Warehouse are put together amd tested for integration aspects.

Deployment Phase :once when all the data warehouse components (specified in the DW –Phase Matrix are build tested and accepted by the Business Users they are moved to the production environment for operation

The Layers

Presentation Layer: It is the Layer through which business users accesses HR DW. Business Objects is the OLAP tool used for accessing the Data ware house and ODS. Analytical reports and MI reports are generated from DW and ODS respectively.

Transform Layer Addresses the method and rules to be applied for data extraction cleansing reorganization and loading of the transformed data to the data structure within Data Ware House.

Structured layer: Basically represents the structure of data organization in the data warehouse. The structure will be aligned to cater the queries and reporting needs of the business users.
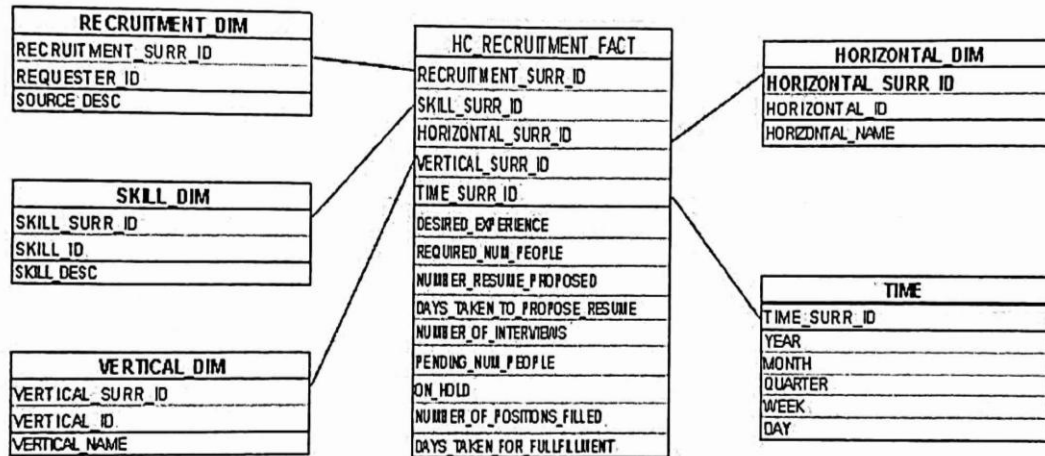
Operations Layer It specifies the post implementation rules for administration and ongoing operations of the Data Ware House

Datawarehouse Design (Star Schema) for the Global Human Capital Datawarehouse
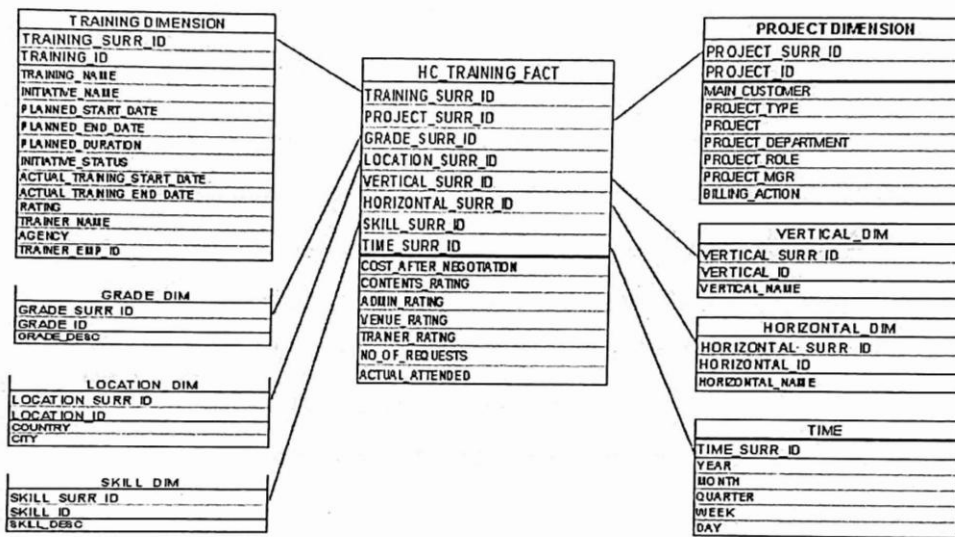
Employee Star Schema
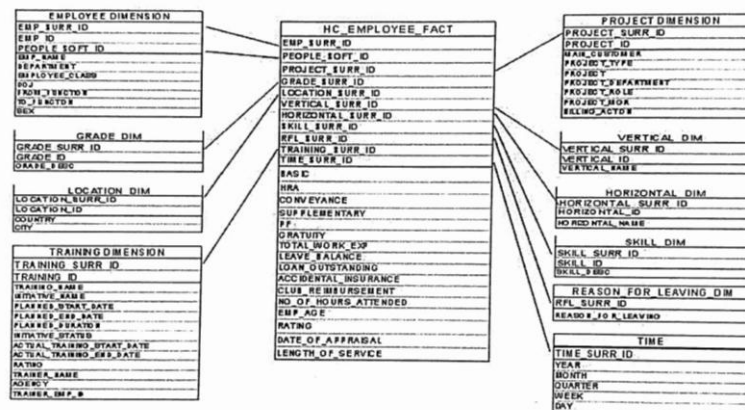
Recruitment Stay Schema

Training Star Schema

2.5 SAS Human Capital Management

- Measure the strategic value of your human Capital investments
- Predict human capital needs through advanced analytic techniques for data mining modeling and forecasting
- Learn from best practices by leveraging benchmark data and automatic calculation of Metrics –all though a web-accessible measurement system
- Surface Metrics into human capital scorecard for identifying cause-and effect linkages of measures associated with strategic objectives
- Build Models that help minimize operational risk and maximize ROI
- Align functional groups with overall organizational strategy through access the key intelligence.

Because this solution is from SAS you get exclusive features –all from the vendor . The web-browser interface lets you combine predefined and user-defined analytics and tap both workforce –specific and other types of information from across your organization

No other solution is more customizable offers more robust warehousing capabilities incorporates more award-winning analytics designed by HR Professionals –not statisticians –or spans more data sources and platforms while providing the widest range of reporting and analysis.

Learn the best Practices

A crucial element of measuring success is you ability to compare your organization to best practices.

Allowing you to set a Strategy that will positively affect future performance. Analytics can help you focus on how

Human capital impacts your business performance –something that traditional tools can't measure You can evaluate the cause-and-effect relationships between business objectives –using internal and external measures- and evaluate the effectiveness of your actions on organizational strategy. As you evaluate your key performance metrics against optimal measures. You can quickly and easily evaluate and adjust HR strategic objectives required to meet the organizational goals.

2.6 DATA MINING using DATA WARE HOUSING

DATA warehouse (DW) databases are popular sources of data Mining Applications because they contain a wealth of internal data from across boundaries which has gathered consolidated validated and cleansed in the extract /transform/load (ETL) process. DW databases may also contain external data such as regulations, demographic, or geographic data which when combined with internal organizational data offers a firm foundation for data mining.

A Data Ware House may be built as a centralized data warehouse a data ware house with data marts or a distributed data warehouse depending on the needs of an enterprise. In the data ware house with data marts architecture it is best to have the data marts extracting information from the central data warehouse to maintain integrity of data   A distributed data warehouse on the other hand is suitable for a large international organization with offices in several countries.

2.7 SAS Enterprise Miner
**SAS ENTERPRISE MINER FOR REAL-TIME CUSTOMER RELATIONSHIP MANAGEMENT**
Benefits
1. **Immediate insight into customer action and business transactions**
2. **Personalized customer Interaction based on real-time information**
3. **Seamless integration of operational data storage enterprise applications and data mining**
4. **Accurate Predictions based on comprehensive up-to-the second information De**
5. **More Efficient business operations**

2.8 DELIVERING END-TO-END BUSINESS INTELIGENCE AND INDUSTRIAL LENGTH DATA MINING

Role of Analytical Applications in ZLE

Analytical applications are a critical driver of real-time Customer Relationship Management (CRM) and zero latency e-business operations. The Compaq Zero Latency Frame work incorporates core analytical capabilities so that events can be analyzed as they happen. The resulting knowledge is made available enterprise-wide to trigger instantaneous actions leading to :

1. Ultra-personalized customer service and highly targeted web marketing campaigns
2. More effective global supply chain and business process management
3. Real-time Fraud Protection and churn management

Analytical Applications
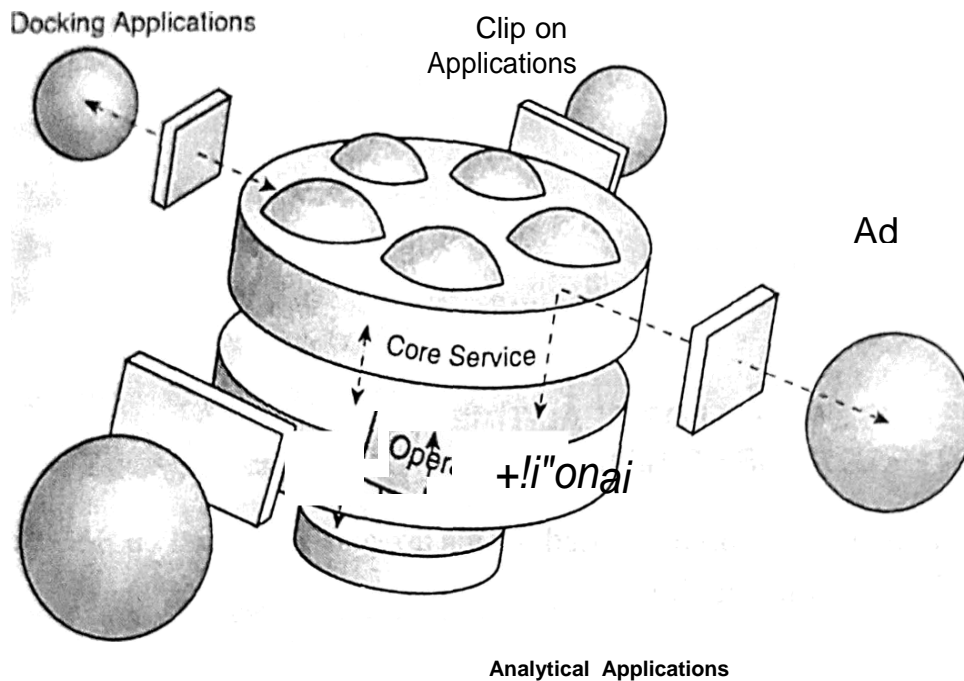
ZLE is an extensible architectural approach for real time integration of enterprise services and data – Synchronize and route information across various enterprise applications –Integrate data into a Storage Cache for the enterprise –Provide data to Business Intelligence applications for analysis –Support transaction access applications needing transactional access to up to the minute data from multiple systems

SAS as e-Intelligence Enterprise Miner

Compaq ZLE partner SAS Institute provides best-of –breed analytic, Data Mining and Real Time CRM Technologies integral to ZLE Solutions. SAS e-Intelligence and Enterprise Miner applications enable users to better understand and predict on-line Customer Behavior. They can integrate the information from these applications with data from data from other channels and customer Touch points for end-to-end business Intelligence and immediate competitive advantage.

SAS Solution for e-intelligence With more information available to business than ever before , via the internet SAS e-intelligence instantly transforms raw data from websites trading hubs automated exchanges and more to actionable knowledge.

E-intelligence combines new knowledge with information from traditional internal and external sources to optimize the effectiveness of CRM, marketing, Supplier relationship management.

SAS Enterprise Miner Data Mining unlocks the value hidden in today's wealth of web transactions, Sales Histories point-of-sales data and more. The ZLE frame work feeds up-to-th-second click stream and operational data into SAS's best-of-breed data mining application which in turn used rich information to enable accurate real-time insights into purchasing pattern behavior and trends .

SAS is solution for CRM

SAS e-intelligence and SAS Enterprise Miner power the SAS solution for CRM, which is designed to maximize return on customer relationships. The real time insights and instant actions, they enable provide companies who implement the ZLE framework with vital e- business benefits such as:

1. Higher hit rates for web and other marketing campaigns.
2. More effective on line cross-selling and up-selling.
3. Lowe business risk and increased customer loyalty.
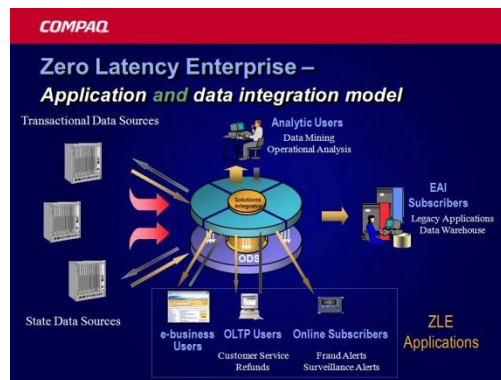
- S software meaning Statistical Analysis Software which is used for Data analytics
- R and Python are two widely used alternative tools of SAS.
- SAS was developed by Jim Goodnight and John Shall in1970 at N.C. University
- SAS allows you to access raw data files & data in an external database of any kind
- SAS architecture is divided mainly of three parts 1)Client Tier 2) Middle Tier 3) Back tier
- To use SAS software you need to follow four steps which are: 1)Access Data, 2)Management 3) Analyze Data, 4) Presentation
- SAS Program consists of three basic steps: Data Step, Proc Step, and Output Step
- SAS data analytical tool is widely used in sectors like Pharmaceutical, Telecom, Financials, Predictive modeling, and Healthcare
- SAS is commercial software while R is open source software
- One of the drawbacks of SAS model is that it is not an open source tool. So, algorithms used in SAS procedures are not available for common use.

3.REPORTS

A. Analysis Report
Resource Summary Report:-Grade wise distribution of resources across departments can be depicted in stacked pie.

Employee Attrition Report: Attrition of Employees by department .It can be further drill down to grade and employee id level to the lowest granularity w.r.t. Time Dimension (Year, qtr, month and day lowest granularity) Thus the reports illustrates number of Employees who resigned by attending the particular training grouped by department. It also gives analysis of assorted reasons of resigning

B. Table Report
Operational reports
1. **Total Training Period report:** This report illustrates sum of the Training hours of employees grouped by department. Further details like number of employees who attend training and their qcp /orientation dates can be drilled across.
2. **Report on channel Expenses :** The report shows recruitment expenses by various channels .Details of the number of positions filled in, dropped closed at a particular time can be further drilled across.
3. **Positions brought forward group by Department :** This report contains number of positions proposed group by Department. It gives details as to how many position have been filled through referral vendor and through various other resources.
4. Thus report describes the project application details Employee Project start date and his onsite Project details can be drilled across.

CONCLUSION
In this Project it is found that Information Technology play a vital role in Human Capital Management System Based on Analysis Reports and ?Technical Reports generated by SAS Package ,The High level(Senior) Management and Middle level Management can take decisions on how to use Human Capital in the It Industry. The Predictions given by SAS are considered and they may implement the best practices among the employees .This gives a solution to overcome Business Challenges with respect to human resources management in IT Industry.