**UNIVERSITY OF ECONOMICS AND LAW**

**FACULTY OF INFORMATION SYSTEMS**

-- 🙞 🕮 🙜 --

Logo

Description automatically generated

**Subject: Business Intelligence and Decision Support System**

**PROGRESS REPORT**

**BUILDING AND ANALYZING DATA HUMAN**

**RESOURCE PROCESSES IN BUSINESS ACTIVITIES**

Lecturer: **Nguyen Van Ho**

Teaching Assistant: **Nguyen Van Tuyen, Nguyen Anh Nhat**

Class code: **212MI3304**

Implementation: **Group 5 – Class K19406**

|  |  |
| --- | --- |
| **Nguyen Thi Thu** | **K194060820** |
| Trinh Nhu Y | K194060838 |
| Do Thi Thanh Phuong | K194060806 |
| Luu Thi Thuy | K194060822 |
| Duong Ngoc Minh Anh | K194060771 |

**LIST OF GROUP 13 MEMBERS AND WORK ASSIGNMENT**

|  |  |  |
| --- | --- | --- |
| **Student ID Code** | **Name** | **Level of contribution** |
| K194060820 | Nguyen Thi Thu | 100% |
| K194060838 | Trinh Nhu Y | 100% |
| K194060806 | Do Thi Thanh Phuong | 100% |
| K194060822 | Luu Thi Thuy | 100% |
| K194060771 | Duong Ngoc Minh Anh | 100% |

**TABLE OF CONTENTS**

[CHAPTER 1: INTRODUCTION 9](#_Toc104331862)

[1.1 Business case for the project 9](#_Toc104331863)

[1.2 Objectives of the project 9](#_Toc104331864)

[1.2.1 General Objectives 9](#_Toc104331865)

[1.2.2 Specific Objectives 9](#_Toc104331866)

[1.3 Research Objects 10](#_Toc104331867)

[1.4 Scope of the project 10](#_Toc104331868)

[1.5 Value and desired outcome of the project 10](#_Toc104331869)

[1.6 Structure of report: 10](#_Toc104331870)

[CHAPTER 2: THEORETICAL BASIS 12](#_Toc104331871)

[2.1 Overview about BI 12](#_Toc104331872)

[2.1.1 What is BI? 12](#_Toc104331873)

[2.1.2 BI Architecture 12](#_Toc104331874)

[2.1.3 Advantages of BI in enterprises 13](#_Toc104331875)

[2.1.4 BI Strategy for Business 14](#_Toc104331876)

[2.2 ETL process 14](#_Toc104331877)

[2.2.1 What is ETL? 14](#_Toc104331878)

[2.2.2 Why do you need ETL 15](#_Toc104331879)

[2.2.3 ETL process 15](#_Toc104331880)

[2.3 Data warehouse and data mart 17](#_Toc104331881)

[2.3.1 What are data warehouse and data mart? 17](#_Toc104331882)

[2.3.2 Snowflake and Star schemas 18](#_Toc104331883)

[2.3.2 Who needs Data warehouse and Datamart? 20](#_Toc104331884)

[2.3.4 Advantages and disadvantages of data warehouse 20](#_Toc104331885)

[2.4 KPIs 21](#_Toc104331886)

[2.4.1 KPIs Definition 21](#_Toc104331887)

[2.4.2 The advantages and disadvantages of KPIs 21](#_Toc104331888)

[2.4.3 Categories of KPIs 22](#_Toc104331889)

[Employee satisfaction levels 22](#_Toc104331890)

[2.5. MDX language for analyzing multi dimension data and OLAP 25](#_Toc104331891)

[2.5.1. What is MDX language? 25](#_Toc104331892)

[2.5.2. OLAP technique 25](#_Toc104331893)

[2.5.3. MDX method and structure 26](#_Toc104331894)

[2.6. Azure Analysis Service 26](#_Toc104331895)

[CHAPTER 3: REQUIREMENTS ANALYTICS AND INTRODUCTION TO BI SOLUTION 28](#_Toc104331896)

[3.1. Business process 28](#_Toc104331897)

[3.1.1. Human resource department 28](#_Toc104331898)

[3.1.2. The purpose of Human resource 28](#_Toc104331899)

[3.1.3. Human resource process 29](#_Toc104331900)

[3.2. Data Source 33](#_Toc104331901)

[3.2.1. Department 33](#_Toc104331902)

[3.2.2. Employee 33](#_Toc104331903)

[3.2.3. EmployeePayHistory 35](#_Toc104331904)

[3.2.4. EmployeeDepartmentHistory 35](#_Toc104331905)

[3.2.5. JobCandidate 36](#_Toc104331906)

[3.2.6. Shift 37](#_Toc104331907)

[3.3. Human resource Requirement Analysis 38](#_Toc104331908)

[3.3.1. Demography 38](#_Toc104331909)

[3.3.2. Performance 38](#_Toc104331910)

[3.3.3. Recruitment 38](#_Toc104331911)

[3.3.4. Salary, bonus, and benefits 38](#_Toc104331912)

[3.3.5. Job title 38](#_Toc104331913)

[3.4. IT requirements Analysis 38](#_Toc104331914)

[3.4.1. Dashboarding and Data Visualization 38](#_Toc104331915)

[3.4.2. Data Management 39](#_Toc104331916)

[3.4.3. Data Querying 39](#_Toc104331917)

[3.4.4. HR Reporting 39](#_Toc104331918)

[3.5.1. Surveying and evaluation 40](#_Toc104331919)

[Tableau 40](#_Toc104331920)

[Power BI 40](#_Toc104331921)

[Google Data Studio 40](#_Toc104331922)

[3.5.2. Proposing BI solution for the project 42](#_Toc104331923)

[CHAPTER 4: BUIDING DATA WAREHOUSE AND INTEGRATING DATA 44](#_Toc104331924)

[4.1. Designing Data Warehouse 44](#_Toc104331925)

[4.1.1. Bus Matrix 44](#_Toc104331926)

[4.1.2. Master Data 45](#_Toc104331927)

[4.1.3 Transaction Data 46](#_Toc104331928)

[4.1.4 Fact and dimension tables 46](#_Toc104331929)

[4.1.5. Snowflake model 53](#_Toc104331930)

[4.2. ETL process 54](#_Toc104331931)

[4.2.1. Dimension Table’s ETL Process 54](#_Toc104331932)

[4.2.2. Fact Table’s ETL Process 56](#_Toc104331933)

[CHAPTER 5: DATA ANALYTICS 58](#_Toc104331934)

[5.1. Data analytics with SSAS technology 58](#_Toc104331935)

[5.1.1. Building the cube 58](#_Toc104331936)

[5.1.2. Analysis with SSAS 58](#_Toc104331937)

[5.1.3. Building KPIs system 67](#_Toc104331938)

[CHAPTER 6: VISUALIZATION AND FORCASTING OR PREDICTIVE MODEL 72](#_Toc104331939)

[6.1. Report and dashboard systems (structure) 72](#_Toc104331940)

[6.2. Data analysis with Power BI 73](#_Toc104331941)

[CHAPTER 7: CONCLUSION AND FUTURE WORKS 76](#_Toc104331942)

[7.1. Results 77](#_Toc104331943)

[7.2. Limitations 77](#_Toc104331944)

[7.3. Future works 77](#_Toc104331945)

**LIST OF FIGURES**

Figure 2.1 Business Intelligence Archirecture………………………………………12

Figure 2.2 ETL Process………………………………………………………… …..16

Figure 2.3 Data warehouse……………………………………………………... …..18

Figure 2.4 Star Schema………………………………………………………… …..19

Figure 2.5 Snow schema……………………………………………………….. …..20

Figure 3.1 Recruitment Process…………………………………………………......34

Figure 4.1 SnowFlake model…………………………………………………… …..58

Figure 4.2 DimEmployees’s ETL process……………………………………..........59

Figure 4.3 DimAddress’s ETL process………………………………………… …..59

Figure 4.4 DimJobTitle’s ETL process……………………………………………...61

Figure 4.5 DimGender’s ETL process………………………………………………61

Figure 4.6 Fact Perfomance’s ETL process…………………………………………61

Figure 4.7 Fact Salary ETL process……………………………………………........61

Figure 4.8 Fact Recruitment’s ETL process……………………………………. …..62

Figure 4.9 Fact Demography’s ETL process……………………………………….62

Figure 5.1 Cube Structive…………………………………………………………..63

Figure 5.2 Numberr and propotion of employees for each age group in the

Entire company……………………………………………………………………..64

Figure 5.3 The number off employees by Gender………………………………….65

Figure 5.4 The number off employees by Country…………………………………65

Figure 5.5 The number of employees by Marital Status…………………………….66

Figure 5.6 The number of employees hired by organizational level……………….67

Figure 5.7 The number of full/part time employees hired by years…......................68

Figure 5.8 The number of full/part time employees hired by years……………….69

Figure 5.9 Sale revenue of sales staff………………………………………….......70

Figure 5.10 Number of days off each department…………………………………71

Figure 5.11 The number of years of work of all employees………………………72

Figure 5.12 AbsenceRate KPI……………………………………………………..74

Figure 5.13 Sales KPI………………………………………………………….......76

Figure 6.1 The dashboard of demography …………………………………………77

Figure 6.2 Recruitment…………………………………………………………….78

Figure 6.3 Perfomance……………………………………………………………..79

Figure 6.4 Top 10 the Employees have the most years of week…………………...80

**LIST OF TABLES**

Table 2.1 Diference between data warehouse and data mart…………………... 19

Table 2.2 Diference between star schema and snow schema…………………... 21

Table 3.1 Description of Department table…………………………………….. 37

Table 3.2 Description of Employees table……………………………………... 38

Table 3.3 Description of Employees PayHistory………………………………. 39

Table 3.4 Description of Employees DepartmentHistory……………………… 40

Table 3.5 Description of JobTitle table………………………………………… 41

Table 3.6 Description of Shift table……………………………………………. 41

Table 3.7 Comparative Analysis of BI and Data Visualization tools………….. 45

Table 4.1 Human Resource Bus Matrix………………………………………... 49

Table 4.2 Description of Dim\_Employees……………………………………... 52

Table 4.3 Description of Dim\_Address………………………………………… 53

Table 4.4 Description of Dim\_Gender…………………………………………. 53

Table 4.5 Description of Dim\_JobTitle………………………………………… 54

Table 4.6 Description of Dim\_Time…………………………………………… 55

Table 4.7 Description of FactDemography…………………………………….. 56

Table 4.8 Description of FactRecruitment……………………………………... 57

Table 4.9 Description of FactPerfomance……………………………………… 57

Table 4.10 Description of FactSalary…………………………………………... 57

**LIST OF ACRONYMS**

|  |  |  |
| --- | --- | --- |
| **STT** | **ACRONYMS** | **EXPLAIN** |
| 1 | ATS | Applicant Tracking System |
| 2 | BI | Business Intelligence |
| 3 | B.O.D | Board of Directors |
| 4 | CRM | Customer Relationship Management |
| 5 | DDL | Data Definition Language |
| 6 | DML | Data Manipulation Language |
| 7 | DW | Data Warehouse |
| 8 | ETL | Extract Transform Load |
| 9 | ERP | Enterprise Resource Planning |
| 10 | ERD | Entity Relationship Diagram |
| 11 | HR | Human Resource |
| 12 | KPIs | Key Performance Indicators |
| 13 | MDX | MultiDimensional eXpression |
| 14 | OLAP | Online Analytical Processing |
| 15 | SQL | Structured Query Language |
| 16 | UTC | Universal Time Coordinated |
| 17 | SSAS | SQL Server Analysis Services |

# **CHAPTER 1: INTRODUCTION**

## 1.1 Business case for the project

Adventure Works is a fictional bicycle wholesaler. The company has 97 different brands of bikes that grouped into three categories: mountain bikes, road bikes, and touring bikes. Moreover, Adventure Works also manufacture some of its own components. Several components, accessories and clothing are purchased from outside from vendors.

Adventure Works is not only selling bicycles, but it also provides accessories, clothing, and components. Many of those things are made by vendors, so Adventure Works stand as a reseller. Adventure Works serve the customer globally, including Australia, Canada, France, and Germany, United Kingdom, and United States. There are 2 business models in Adventure Works which are retail stores that sell bikes, and internet sales that serve individual customers. Usually Adventure Works sells in bulk to retail stores, which acts as resellers for its products.

To run the business activities, Adventure Works has a total of 290 employees that included in some functions such as sales, production, purchasing, engineering, finance, information services, marketing, shipping and receiving, and R&D.

Therefore, with the large scale of the company, each part of the business needs specific KPIs, so setting up a BI strategy with KPIs that match the goals of the business areas. Is very important. To manage performance effectively, businesses need to expand their Business Intelligence capabilities and reporting platforms to ensure it is the source of information for users in all departments: manufacturing, sales, and accounting, etc. This means that organizations need to extend their company-wide application of BI.

## 1.2 Objectives of the project

### **1.2.1 General Objectives**

To have more than 90% of employees happy with their work, pay, and bonus in the business, and to assist the organization's activities in running smoothly and efficiently. Recruit and pick appropriate applicants for the company in a timely manner.

### **Specific Objectives**

* Report on employee's ability to complete KPI.
* Determine the seniority of the employee to the determine the seniority allowance for the employee.
* Reasonable resource allocation at branches brings convenience to employees.
* Follow-up company recruitment activities.
* Monitor the working process of employees in the company so that the company considers salary adjustments for employees.

## Research Objects

* Recruitment procedure.
* Benefits and incentive compensation policy.
* Human resource report.
* SQL Server Business Intelligence.
* AdventureWorks database.

## Scope of the project

* + Analysis of the requirements.
  + Create KPIs consistent with the goals of your business.
  + Building a data warehouse that synthesizes information from all sources in the business areas of the business to ensure sufficient and fast provision for businesses.
* Enhance the reporting platform through the use of suitable BI tools and develop a team of data analysis.
* Making reports, predictive models from data to support the management of corporate personnel.

## Value and desired outcome of the project

* Theoretically:
* Understand the business processes of human resources.
* Understand the BI solution.
* Understand KPI assessment methodology.
* Understand the data analysis model and business intelligence system.
* In practice: Create documentation with business intelligence models and data analysis data with financial and business reports for HR module to help administrators control information effectively as well as make the right decision.

## Structure of report:

Report includes 7 chapters:

1. Introduction
2. Theoretical basis
3. Requirement analytics and introduction to BI solution
4. Building data warehouse and integrating data
5. Data analytics
6. Visualization and forecasting
7. Report

# **CHAPTER 2: THEORETICAL BASIS**

## Overview about BI

### **What is BI?**

Business intelligence (BI) leverages software and services to transform data into actionable insights that inform an organization’s strategic and tactical business decisions.

To achieve that goal, BI incorporates a combination of analytics, data management and reporting tools, plus various methodologies for managing and analyzing data.

### **BI Architecture**

Business intelligence architecture refers to the infrastructure that organizations use to define their data collection streams, information administration, and all the technology that supports their business intelligence.

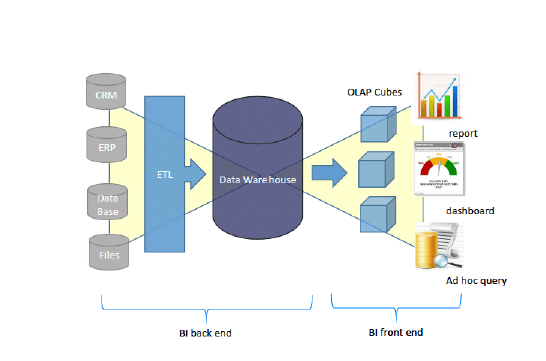
It includes the IT systems and software tools that are used to collect, integrate, store and analyze BI data and then present information on business operations 

Figure 2.1 Business Intelligence Architecture

Components of Business Intelligence Architecture:

* **Source systems:** gathering data from various data sources such as CRM, ERP, databases, files or APIs, depending on the requirements and resources of a company. They can also include secondary sources, such as market data and customer databases from outside information providers.
* **Data integration and cleansing tools:** To effectively analyze the data collected, an organization must integrate and consolidate different data sets to create unified views of them by using smart ETL engines in the background. This step ensures data is clean and prepared for the final stage: loading into a data warehouse and data mart.
* **Analytics data stores:** data warehouse structured data in a relational, columnar or multidimensional database and makes it available for querying and analysis. An enterprise data warehouse can also be tied to smaller data marts set up for individual departments and business units with data that's specific to their BI needs. A well-planned architecture should specify which of the different data stores is best suited for particular BI uses.
* **BI and data visualization tools:** This component performs data retrieval, data analysis, and data-mining tasks using the data in the data store. This component is used by the data analyst to create the queries that access the database. This component is generally represented in the form of an OLAP tool.
* **Dashboards, portals and reports:** in charge of presenting the data to the end-user in a variety of ways. This component is used by the data analyst to organize and present the data. This tool helps the end-user select the most appropriate presentation format, such as a summary report, map, pie or bar graph, or mixed graphs.

### **Advantages of BI in enterprises**

BI solutions help businesses gain the following competitive advantage:

* Cut labour costs: Save time and effort by generating standard reports for an organization.
* Reduce information congestion: Faster analysis, intuitive dashboards.
* Make data executable: Data collection and filtering become clean data.
* Better decision-making: Having accurate data and faster reporting capability provides for better business decisions.
* Decisions are taken more quickly: businesses can keep up with changes in the industry and anticipate customer needs.
* Orient the organization towards business goals.BI provides leaders with the ability to access data and gain a holistic view of their operations and provide data solutions to departments.

### **BI Strategy for Business**

A BI strategy will allow you to address all your data problems and needs, develop a cohesive system, and keep it maintained.

BI strategy has three main elements:

* Vision. Why are you building the BI practice in your company and what do you want to achieve?
* People and processes. Who will define and run the BI strategy? And how?
* Tools and architecture. Which dashboards and solutions do we want to build? For which areas? And how will they impact those areas?

**BI Strategy for Business includes six-step:**

Step 1. Assess your current BI ecosystem

Step 2. Create the vision

Step 3. Establish BI governance processes

Step 4. Build a BI roadmap

Step 5. Document a BI strategy

Step 6. Review your BI strategy every year

## 2.2 ETL process

### **2.2.1 What is ETL?**

ETL, which stands for extract, transform and load, is a data integration process that combines data from multiple data sources into a single, consistent data store that is loaded into a data warehouse or other target system.

* Extraction- Data are extracted from the available internal and external sources.
* Transformation- The goal of the cleaning and transformation phase is to improve the quality of data extracted from different sources.
* Loading- After extraction and transformation, data are loaded into the tables of the data warehouse.

### **2.2.2 Why do you need ETL**

* **Time-Efficiency:**An ETL tool allows you to collect, transform, and consolidate data in an automated way. As a result, you can save plenty of time and effort otherwise spent on importing data manually
* **Handle Complex Data Easily:**An ETL tool streamlines the tedious data cleansing tasks
* **Reduced Error Probability:**ETL tools automate several parts of a data process, reducing manual intervention and lowering error probability.
* **Improved Business Intelligence And ROI:**ETL tools helps companies to analyze their business data for taking critical business decisions.

### **2.2.3 ETL process**

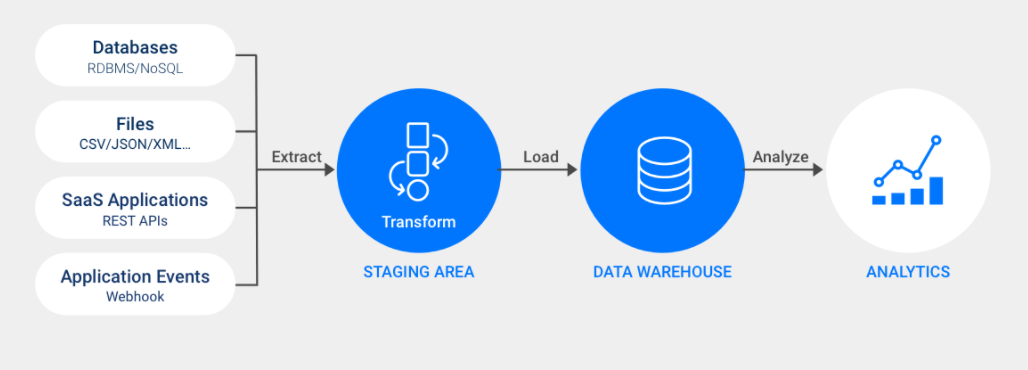
****

Figure 2.2 ETL Process

The ETL process is comprised of 3 steps that enable data integration from source to destination: data extraction, data transformation, and data loading.

**Step 1: Extraction**

To create a data warehouse, extraction typically involves combining data from these various sources into a single data set and then validating the data with invalid data flagged or removed. Extracted data may be several formats, such as relational databases, XML, JSON, and others. It is important to extract the data from various source systems and store it into the staging area first and not directly into the data warehouse because the extracted data is in various formats and can be corrupted also. Hence loading it directly into the data warehouse may damage it and rollback will be much more difficult.

**Step 2: Transformation**

Data extracted from the source server is raw and not usable in its original form. Therefore it needs to be cleansed, mapped and transformed. The process of data transformation is comprised of several sub-processes:

* Cleansing — inconsistencies and missing values in the data are resolved.
* Cleansing: inconsistencies and missing values in the data are resolved.
* Standardization: formatting rules are applied to the dataset.
* Deduplication: redundant data is excluded or discarded.
* Verification: unusable data is removed and anomalies are flagged.
* Sorting: data is organized according to type.
* Other tasks: any additional/optional rules can be applied to improve data quality.

**Step 3: Loading**

This process is to load the newly transformed data into a new destination (data lake or data warehouse.)

In a typical Data warehouse, huge volume of data needs to be loaded in a relatively short period (nights). Hence, load process should be optimized for performance.

In case of load failure, recover mechanisms should be configured to restart from the point of failure without data integrity loss.

Three types of Loading:

* Initial Load: populating all the Data Warehouse tables
* Incremental Load: applying ongoing changes as when needed periodically.
* Full Refresh: erasing the contents of one or more tables and reloading with fresh data.

Often, the three ETL phases are run in parallel to save time. For example, while data is being extracted, a transformation process could be working on data already received and prepare it for loading, and a loading process can begin working on the prepared data, rather than waiting for the entire extraction process to complete.

## 2.3 Data warehouse and data mart

### **2.3.1 What are data warehouse and data mart?**

A **data warehouse** is a type of data management system that is designed to enable and support business intelligence (BI) activities, especially analytics. Data warehouses are solely intended to perform queries and analysis and often contain large amounts of historical data. The data within a data warehouse is usually derived from a wide range of sources such as application log files and transaction applications.

A **data mart** is a curated subset of data often generated for analytics and business intelligence users. Data marts are often created as a repository of pertinent information for a subgroup of workers or a particular use case.

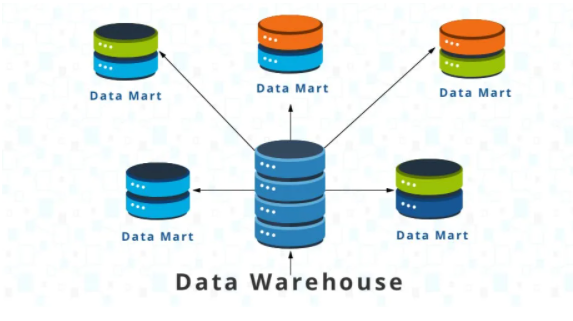


Figure 2.3 Data Warehouse

|  |  |  |
| --- | --- | --- |
| **Parameter** | **Data mart** | **Data Warehouse** |
| Usage | It helps to take tactical decisions for the business. | It helps to take a strategic decision. |
| Data Handling | Data warehousing includes a large area of the corporation, it takes a long time to process it. | Data marts are easy to use, design and implement as it can only handle small amounts of data. |
| Designing | Easy | Quite difficult. |
| Source | comes from very few sources. | comes from many sources. |
| Scope | Line-of-Business. | Enterprise-wide. |
| Size | Less than 100 GB. | Range from 100 GB to 1 TB+. |
| Implementation time | Data Mart is restricted to few months. | Data Warehouse can be extended from months to years. |

*Table 2.1 Diference between data warehouses and data mart*

### **2.3.2 Snowflake and Star schemas**

**Star schema** is the type of multidimensional model which is used for data warehouse. In star schema, The fact tables and the dimension tables are contained. In this schema fewer foreign-key join is used. This schema forms a star with fact table and dimension tables.

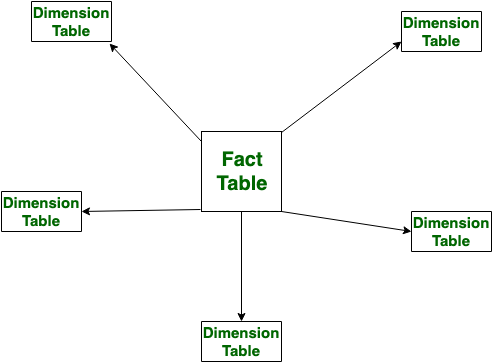
****

Figure 2.4 Star schemas

**The snowflake schema** represents a dimensional model which is also composed of a central fact table and a set of constituent dimension tables which are further normalized into sub-dimension table**s**

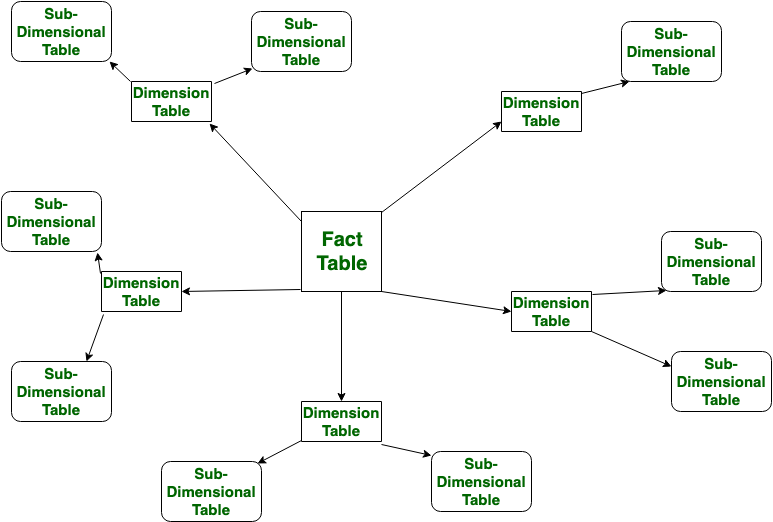


Figure 2.5 Snow Schema

|  |  |
| --- | --- |
| **Star Schema** | **Snowflake Schema** |
| Hierarchies for the dimensions are stored in the dimensional table. | Hierarchies are divided into separate tables. |
| Only single join creates the relationship between the fact table and any dimension tables. | A snowflake schema requires many joins to fetch the data. |
| Single Dimension table contains aggregated data. | Data Split into different Dimension Tables. |
| Star schema is a top-down model. | While it is a bottom-up model. |
| It takes less time for the execution of queries. | It takes more time than star schema for the execution of queries. |
| Normalization is not used. | Both normalization and denormalization are used. |
| It’s design is very simple. | It’s design is complex. |
| It has high data redundancy. | While it has low data redundancy. |

*Table 2.2 Diference between star schema and snowflake schema*

### **2.3.2 Who needs Data warehouse and Datamart?**

**Who needs Data Warehouse?**

* Decision-makers who rely on the mass amount of data
* Users who use customized, complex processes to obtain information from multiple data sources.
* It is also essential for those people who want a systematic approach for making decisions.
* If the user wants fast performance on a huge amount of data which is a necessity for reports, grids or charts, then Data warehouse proves useful.

**Who needs Data Mart?**

* Long sales cycles can use a data mart to help different departments align on the needs of the sales process.
* Who have many years of data can use a data mart to help departments focus on current data while still maintaining the historical information.
* Data marts enable them to manage large product sets better by focusing each product group on its individual mission.

### **2.3.4 Advantages and disadvantages of data warehouse**

**Advantages of data warehouse**

* The Data Warehouse helps to combine multiple data sources. It also saves time for users to access data from various sources.
* Restructuring and Integration make it easier for the user to use for reporting and analysis.
* The Data warehouse provides consistent information on various cross-functional activities. It is also supporting ad-hoc reporting and queries.
* A large amount of historical information is stored in a data center. This helps users to compare different times and trends to construct possible predictions.

**Disadvantages of data warehouse**

* Not an ideal option for unstructured data.
* Data warehouses' maintenance generates usually significant costs.
* Difficult to make changes in data types and ranges, data source schema, indexes, and queries.
* Data Warehouse can be outdated relatively quickly
* Difficulty in Compatibility: the use of data warehouse technology could likely require a helping hand from an independent BI team. With the intricacies of operating systems, it can be difficult for a business owner.
* Organisations need to spend lots of their resources on training and implementation purposes.

## 2.4 KPIs

### **2.4.1 KPIs Definition**

A key performance indicator (KPI) is a measurable value that demonstrates how effectively a company is achieving key business objectives. Organizations use KPIs to evaluate success at reaching targets. High-level KPIs may focus on the overall performance of the business, while low-level KPIs may focus on processes across departments like sales, marketing, HR, or support.

### **2.4.2 The advantages and disadvantages of KPIs**

**2.4.2.1 The Advantages of KPIs:**

**Measurable Results**

As the sole purpose of KPI is to watch for rate of progress, it displays precise results in the form of statistics, numbers or metrics. An employee or organization can easily measure or track the progress toward their goal and understand which part of the task requires more focus. In addition, KPIs give results daily, weekly or periodically depending on requirement or goal type.

**Alignment**

Keeping track of each employee's progress can be difficult if it's a large organization. Then, KPI helps them stay aligned with their goals because it helps everyone involved in the project to track progress.

This keeps everyone motivated because no one wants to see their name or progress marked as incomplete. In addition, it ensures everyone works consistently, systematically.

**Future Strategies**

Based on previous results, KPIs can help managers redesign or revise their strategy. So it helps them to plan or set future goals.

**Rewards**

Employee motivation is to be rewarded or increase salary after a job well done. That is why KPIs support each employee to have the opportunity to prove themselves as well as help managers see progress and deserved rewards. Not only that, it helps employees improve themselves and track their performance.

**2.4.2.2 The Disadvantages of KPIs:**

**Decrease in Quality**

Due to too much focus on short-term goals, it is possible that employees will not pay attention to the quality of work. As a result, setting financial goals and metrics tends to increase the weight of the task rather than the authenticity of the task.

**Standardization**

If employees focus too much on results, they will lose their creativity. Using traditional methods helps them get the job done faster.

**Loyalty**

KPIs are difficult to track the quality of work but only show levels of progress. As a result, loyalty between customers and businesses can be affected. When they do, they can lose customers and weaken the bond between them.

In short, KPIs are very useful for short-term goals, but not as effective for long-term goals. Therefore, it is necessary to consider all factors before implementing KPI for the organization.

### **2.4.3 Categories of KPIs**

### **Employee satisfaction levels**

Use the work environment survey and review the results obtained to detect problems, thereby improving a positive working environment and creating non-working efficiency for employees. When the atmosphere in the workplace is ensured, employee comfort is enhanced, improving their sense of well-being. Therefore, the brand is firmly built, and it is easy to attract new talent. Have employees do this by asking them to rate the work environment from 1 to 10 and take the average from all the answers

**Employee turnover**

This represents the average time that an employee works at the company, which shows their stable work. In other words, it shows the ability of companies to retain talent. The company will incur costs in the recruitment process if the employee turnover rate is high. The cause of the rotation may be because the company's salary is below the market rate or there is a poor working environment.

This is calculated by: the number of employees leaving the company in a year divided by the number of employees of the company in that year. Then multiply by 100 to get the annual sales rate. ([L / Avg] x 100).

There are many reasons why employees leave their jobs: a need for a higher salary, dissatisfaction with their current job, or a change in family circumstances.

**Rate of employee absenteeism**

Sick or injured leave in the UK was estimated to have been around 141.1 million days in 2018, according to a report released by the Office for National Statistics. This equates to 4.4 days per worker.

This KPI can help detect problems with the way the company is performing, thereby predicting employee absences in the organization.

Employees are allowed access to the system to track their absence history. They can then reduce absenteeism. Visibility of a coworker's reasonable absence also increases team spirit and strengthens co-worker relationships when working on group projects.

**Average recruitment time**

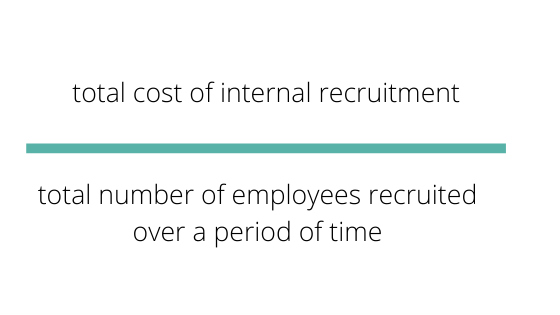
To improve and save company time and resources, access to metrics like these is essential. The recruitment period is calculated from the time the job is advertised until the start of new recruitment at the company. The lower this number, the better.

**Recruitment costs**

Another human resources KPI to consider is recruitment costs. In other words, how much money is spent to acquire the resources needed to hire a new employee? When calculating this, there are various factors to consider:

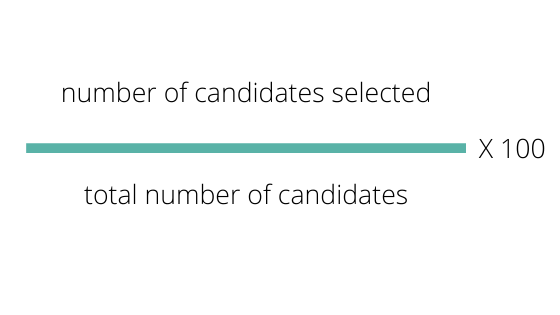
* Recruitment costs: advertising, subscription to employment portals, cost of the recruiter’s time, ATS and other software.
* Training costs: education, materials.

The formula is: (Total costs of internal recruitment + total costs of external recruitment) / (total number of employees recruited over a period of time).



**Recruitment conversion rate**

The recruiting conversion rate assists us in determining the most effective ways or channels for generating the best candidate for the job. This information allows us to compare and improve any recruitment process in order to cut costs. How is this determined? It's quite straightforward: (Number of selected candidates / total number of candidates) x 100



**Workplace accidents**

Workplace accidents are another significant HR indicator since they allow you to see if your occupational health and safety policies are being executed appropriately and effectively. In some industries, this information is more sensitive, thus it should be included in the department's scorecard. The total number of accidents resulting in harm per million hours worked minus the number of workers exposed to hazards equals this data. (Number of accidents multiplied by 1,000,000) / (total number of hours worked - people exposed to risks).



## 2.5. MDX language for analyzing multi dimension data and OLAP

### **2.5.1. What is MDX language?**

The MDX language is the query language for multidimensional databases, it is similar to SQL for relational databases, but it is the language of computation, it is therefore similar in syntax to the spreadsheet formula.

The MDX language has strong support for multi-dimensional data blocks, it supports data access in an intuitive and simple way. User can extract data from cubes with any number of dimensions. MDX was designed for major data analysis systems to easily process queries. MDX can handle data in a flexible way , MDX's are also very flexible. Complex queries in SQL as pivots are easy in MDX.

MDX does not support data definition languages, but the MDX manipulation language supports the implementation of a number of data definition operations.

### **2.5.2. OLAP technique**

OLAP stands for Online Analytical Processing, and it is a form of software that allows users to simultaneously analyze data from several databases. It's a program that lets analysts extract and evaluate market data from numerous angles.

For analysts, grouping, aggregating, and joining data is a typical operation. In relational databases, these processes consume a lot of resources. OLAP allows data to be pre-calculated and pre-aggregated, which speeds up analysis.

OLAP databases are partitioned using one or more cubes. The cubes are designed in such a way that creating and displaying reports is simple. The term "online analytical processing" (OLAP) refers to the process of analyzing data over the internet.

OLAP has four types of analytical operations:Roll-up:

* Drill-down.
* Slice and dice.
* Pivot (rotate).

### **2.5.3. MDX method and structure**

MDX has a structure comparable to SQL, but it is more expensive when dealing with multidimensional databases. The MDX query is written in the following format:

* The SELECT clause is used to specify the result set's dimensions.
* The FROM clause specifies the data source (cube) from which the data is retrieved.
* To filter the output data, the WHERE clause is utilized to determine the data truncation dimension.

The MDX query syntax is identical to that of SQL, with the following exceptions:

* Curly braces are used around sets or members in MDX syntax to identify them.
* While an MDX query can have up to 128 dimensions, only the first 5 have aliases (subname). In an MDX statement, an axis can be addressed by its order or, if existent, its sub-name. The FROM clause in a SQL query can refer to several data tables. An MDX query, on the other hand, can only have one FROM clause. The LookupCube function can be used to look up information from other blocks by value.
* The data cut direction is specified by the WHERE clause. If a dimension isn't specified in the specification, it's assumed that it doesn't exist
* The data cut direction is described using the WHERE clause. SQL Analysis Services still evaluates a dimension if it is not stated in the WHERE clause, but it is filtered by its default amount of members. The WHERE clause can help refine the data by changing the filtering for a specific dimension.

### **2.6. Azure Analysis Service**

Azure Analysis Services is a fully managed platform as a service (PaaS) that offers cloud-based enterprise data models. To aggregate data from many data sources, create metrics, and secure your data in a single, trustworthy tabular semantic data model, use advanced mashup and modeling tools. Users can execute ad hoc data analysis using tools like Power BI and Excel with ease and speed thanks to the data model.

You may create a server on the Azure interface in minutes. You may also create servers using declarative templates using Azure Resource Manager templates and PowerShell.

Many Azure services interface with Azure Analysis Services, allowing you to create complex analytics solutions. Secure, role-based access to your essential data is provided by integrating with Azure Active Directory. Include an activity that imports data into the model to integrate with Azure Data Factory pipelines. Custom code can be used to lightweight orchestrate models using Azure Automation and Azure Functions.

# **CHAPTER 3: REQUIREMENTS ANALYTICS AND INTRODUCTION TO BI SOLUTION**

## 3.1. Business process

### **3.1.1. Human resource department**

An HR department is tasked with maximizing employee productivity and protecting the company from any issues that may arise within the workforce. HR responsibilities include compensation and benefits, recruitment, firing, and keeping up to date with any laws that may affect the company and its employees.

Job title of Human resource department:

* Human resource assistant
* Human resource generalist
* Human resource manager
* Human resource director
* Vice President of HR

### **3.1.2. The purpose of Human resource**

The purpose of human resources departments is to manage the numerous needs of company employees. From financial to legal matters, human resources departments help in maximizing the efficiency of an organization.

**Staffing Needs**:

Human resources departments are mainly known for their responsibility of recruiting employees. This includes creating position announcements which involve identifying job duties and determining the skill level and requirements that are needed for a position. Personnel screen applicants by completing reference checks, administering tests, and interviewing potential candidates.

**Compensation**

Fair payment is a huge concern for employees. Human resources department personnel evaluate and make changes to an organization’s pay structure by researching compensation trends with the aim of providing employees with industry-acceptable pay. They also assess policies and help an organization comply with the various federal, state, and local laws in regards to compensation

**Benefits**

Developing and effectively administrating an employee benefits program is a great recruitment and retention tool that organizations use. Common benefits of a job include health insurance, dental insurance, and a retirement plan. Human resources personnel must research and analyze information to obtain the best packages for these benefits which should provide adequate coverage for employees while minimizing the costs for the employer.

**Performance Appraisal**

Human resources departments work in conjunction with individual department managers to evaluate employee performance. This department is responsible for creating the entire performance appraisal process. This process frequently occurs annually for many organizations. Thus, it’s important that the process is simple and straight-forward, yet is effective.

**Law Compliance**

Companies are responsible for the safety and equitable treatment of their employees. Laws exist to protect employees in a variety of areas. Human resources departments keep companies up-to-date with regulations by staying aware of these laws, disseminating the proper information, and creating procedures to ensure adherence to the laws by everyone.

### **3.1.3. Human resource process**

#### **3.1.3.1. Recruitment process**

Recruitment process is the first step in creating a powerful resource base. The process undergoes a systematic procedure starting from sourcing the resources to arranging and conducting interviews and finally selecting the right candidates.

**Recruitment Planning**

Recruitment planning is the first step of the recruitment process, where the vacant positions are analyzed and described. It includes job specifications and its nature, experience, qualifications and skills required for the job, etc.

A structured recruitment plan is mandatory to attract potential candidates from a pool of candidates. The potential candidates should be qualified, and experienced with the capability to take on the responsibilities required to achieve the objectives of the organization.

This step includes Identifying Vacancies, Job Analysis, and Job Evaluation

**Recruitment Strategy**

Recruitment strategy is the second step of the recruitment process, where a strategy is prepared for hiring the resources. After completing the preparation of job descriptions and job specifications, the next step is to decide which strategy to adopt for recruiting the potential candidates for the organization.

**Searching the Right Candidates**

Searching is the process of recruitment where the resources are sourced depending upon the requirement of the job. After the recruitment strategy is done, the searching of candidates will be initialized. This process consists of two steps −

* Source activation − Once the line manager verifies and permits the existence of the vacancy, the search for candidates starts.
* Selling − Here, the organization selects the media through which the communication of vacancies reaches the prospective candidates.

**Screening / Shortlisting**

Screening starts after the completion of the process of sourcing the candidates. Screening is the process of filtering the applications of the candidates for the further selection process.

Screening is an integral part of the recruitment process that helps in removing unqualified or irrelevant candidates, which were received through sourcing. The screening process of recruitment consists of three steps

Reviewing of Resumes and Cover Letters: the resumes of the candidates are reviewed and checked for the candidates’ education, work experience, and overall background matching the requirement of the job

Conducting Telephonic or Video Interview: In this process, after the resumes are screened, the candidates are contacted through phone or video by the hiring manager.

Identifying the top candidates: s. In this process, the cream/top layer of resumes is shortlisted, which makes it easy for the hiring manager to make a decision.

**Evaluation and Control**

Evaluation and control is the last stage in the process of recruitment. In this process, the effectiveness and the validity of the process and methods are assessed. Recruitment is a costly process, hence it is important that the performance of the recruitment process is thoroughly evaluated.



Figure 3.1 Recruitment Process

#### **3.1.3.2. Payroll process**

**Step 1**: Establish your employer identification number.

The first step in processing payroll is to establish your EIN and your state and local tax IDs. The government uses these identifications to track your business's payroll taxes and ensure you're meeting requirements.

**Step 2**: Collect relevant employee tax information.

Before you start processing payroll, your employees will have to fill out various tax forms so you can account for allowances and [other tax details](https://www.businessnewsdaily.com/2228-payroll-taxes-bndmp.html). These forms include the W-4 and [I-9](https://www.businessnewsdaily.com/15783-w9-1099-tax-forms.html) (if it is a new employee). There are various state and local forms you will have to provide, but these will depend on where your business is operating.

**Step 3**: Choose a payroll schedule.

Once you have the relevant tax and legal information to set up payroll, you can [choose a schedule that works best for your business](https://www.businessnewsdaily.com/11287-running-payroll-frequency.html). There are four main schedules: monthly, semimonthly, biweekly and weekly. It's important to understand each plan before deciding which is best for your business. Once you choose a schedule, set up a calendar with paydays, and make note of the days when you'll have to process payroll for your workers to get their money on that defined day.

Build in important quarterly tax dates, holidays, and annual tax filing dates. Keep in mind that you'll have to do this at the start of every year. You'll also want to establish the preferred delivery method for each employee. For example, many businesses allow employees to choose between paperchecks and direct deposit.

**Step 4**: Calculate gross pay.

Now that you've set a payroll schedule, you can start processing your first payroll. To do this, you must calculate each employee's gross pay, which is the total number of hours an employee works in a given pay period multiplied by their hourly rate.

Start by calculating the number of hours an employee has worked in a given pay period, and take note of [overtime hours](https://www.businessnewsdaily.com/15862-overtime-pay-and-how-to-calculate-it.html). The extra time has to be paid out at a higher rate consistent with [federal law](https://www.dol.gov/whd/overtime_pay.htm). If an hourly worker puts in more than 40 hours per week, you'll have to pay time and a half, or an employee's hourly wage plus half that wage.

**Step 5**: Determine each employee's deductions.

Gather information from your workers' W-4s, federal and state requirements, insurance requirements, and benefits requirements to determine each employee's deductions. This can get complicated; [each state collects different taxes](https://www.payroll-taxes.com/state-tax) from small businesses, so you'll have to research your state's policies before you complete this step.

**Step 6**: Calculate net pay, and pay your employees.

Subtract each employee's deductions from their gross pay. The amount left over is the employee's net pay, or take-home pay. This is the amount you'll pay each employee. You'll have to hold the deductions and pay them with your payroll taxes each month or quarter, depending on the schedule you establish.

**Step 7**: Keep payroll records, and make any necessary corrections.

As you process payroll, it's important to keep records of your transactions for tax and compliance purposes. If an employee disputes payment or the IRS needs some kind of documentation down the line, you need to have records at the ready. Especially in the case of an employee disputing a paycheck, it's important to maintain records, including year-to-date payment, so you can sort out any issues that arise.

**Step 8**: Be mindful of ongoing considerations.

Keep in mind that you have to file your business's taxes quarterly and annually. It's important to [consult an accountant](https://www.businessnewsdaily.com/8039-find-small-business-accountant.html) to ensure you understand how your payroll taxes fit into this aspect of your operations. You'll also have to report any new hires to the IRS. When you work with a payroll solution or an accountant, this usually isn't your responsibility.

## 3.2. Data Source

Human Resources module has 6 data tables:

* Department.
* Employee.
* EmployeePayHistory.
* EmployeeDepartmentHistory.
* JobCandidate.
* Shift

### **3.2.1. Department**

Department data table is used to store departments of the enterprise: It will store the department name and each department will have a unique ID as the primary key. Each department will belong to a different group. In addition, this table also stores the date of the last modification related to a department.

|  |  |  |
| --- | --- | --- |
| **Column Name** | **Signification** | **Data Type** |
| Department Id | Primary key for Department records. | Smallint |
| Name | Name of the department. | Name:nvarchar(50) |
| Group Name | Name of the group to which the department belongs. | Name:nvarchar(50) |
| Modified day | Date and time the record was last updated | Datetime |

*Table 3.1 Description of Department table*

### **3.2.2. Employee**

This data sheet stores all information about the company's employees such as corporate ID, marital status, ID number, position, date of birth and so on.

|  |  |  |
| --- | --- | --- |
| **Column Name** | **Signification** | **Data Type** |
| BusinessEntityID | Primary key for Employee records. Foreign key to BusinessEntity.BusinessEntityID. | Int |
| NationalIDNumber | Unique national identification number such as a social security number. | Nvarchar(15) |
| LoginID | Network login. | Nvarchar(256) |
| OrganizationNode | Where the employee is located in corporate hierarchy. | Hierarchyid |
| OrganizationLevel | The depth of the employee in the corporate hierarchy. | Smallint |
| JobTitle | Work title such as Buyer or Sales Representative. | Nvarchar(50) |
| BirthDate | Date of birth. | Date |
| MaritalStatus | M = Married, S = Single | Nchar(1) |
| Gender | M = Male, F = Female | Nchar(1) |
| HireDate | Employee hired on this date. | Date |
| SalariedFlag | Job classification. 0 = Hourly, not exempt from collective bargaining. 1 = Salaried, exempt from collective bargaining. | Flag:bit |
| VacationHours | Number of available vacation hours. | Smallint |
| SickLeaveHours | Number of available sick leave hours. | Smallint |
| CurrentFlag | 0 = Inactive, 1 = Active | Flag:bit |
| Rowguid | ROWGUIDCOL number uniquely identifying the record. Used to support a merge replication sample. | Uniqueidentifier |
| ModifiedDate | Date and time the record was last updated. | Datetime |

*Table 3.2 Description of Employee table*

### **3.2.3. EmployeePayHistory**

This is a data sheet used to store pay history for employees.

|  |  |  |
| --- | --- | --- |
| **Column Name** | **Signification** | **Data Type** |
| BusinessEntityiD | Employee identification number. Foreign key to Employee.BusinessEntityID. | Int |
| RatechangeDate | Date the change in pay is effective. | Datetime |
| Rate | Salary hourly rate. | Money |
| PayFrequency | 1 = Salary received monthly, 2 = Salary received biweekly | Tinyint |
| ModifiedDate | Date and time the record was last updated. | Datetime |

*Table 3.3 Description of EmployeePayHistory table*

### **3.2.4. EmployeeDepartmentHistory**

Employee Department History data sheet is a place to store employees' workplace. It will store where the employee ID and the department ID knows which department the employee works for. It also tells that the shift ID if we combine this table with the shift table will get the employee's shift name. It also records the start date of work in one department and the end date (if any) when the employee quit the job or moved to another department.

|  |  |  |
| --- | --- | --- |
| **Column Name** | **Signification** | **Data Type** |
| Business Entity ID | Employee identification number. Foreign key to Employee.BusinessEntityID. | Int |
| Department ID | Department in which the employee worked including currently. Foreign key to Department.DepartmentID. | Smallint |
| Shift ID | Identifies which 8-hour shift the employee works. Foreign key to Shift.Shift.ID. | Tinyint |
| Start Date | Date the employee started work in the department. | Date |
| End Date | Date the employee left the department. NULL = Current department. | Date |
| Modified day | Date and time the record was last updated. | Datetime |

*Table 3.4 Description of EmployeeDepartmentHistory table*

### **3.2.5. JobCandidate**

The job candidate datasheet is where candidate profiles are stored

|  |  |  |
| --- | --- | --- |
| **Column Name** | **Signification** | **Data Type** |
| JobCandidateID | Primary key for JobCandidate records. | Int |
| BusinessEntityID | Employee identification number if applicant was hired. Foreign key to Employee.BusinessEntityID. | Int |
| Resume | Résumé in XML format. | XML (CONTENT HumanResources.HRResumeSchema Collection) |
| Modifiedday | Date and time the record was last updated. | Datetime |

*Table 3.5 Description of JobTitle table*

### **3.2.6. Shift**

The Shift data sheet is a place to store the work shifts as well as the enterprise. Here each shift will have its own ID and this ID is the primary key. Each shift will have a start time and an end time.

|  |  |  |
| --- | --- | --- |
| **Column Name** | **Signification** | **Data Type** |
| ShiftId | Primary key for Shift records. | Tinyint |
| Name | Shift description. | Name:nvarchar(50) |
| Start time | Shift start time. | Time(7) |
| End Time | Shift end time. | time(7) |
| Modified day | Date and time the record was last updated. | Datetime |

*Table 3.6 Description of Shift table*

## 3.3. Human resource Requirement Analysis

### **3.3.1. Demography**

With demography analysis will help the company figure out where the majority of the company’s employees come from, The average age of employees in each department and the number, proportion of employees for each age group in the entire company, the number of employee by Gender.

### **3.3.2. Performance**

Performance analysis will help the company know-how company’s performance through some questions:

* How many days off do employees have in a month/quarter/year?
* Has the employee achieved KPIs?

### **3.3.3. Recruitment**

Recruitment analysis will help the company know How many employees are needed for the vacancy and know the number of employee hired by Organizational level, by years and department.

### **3.3.4. Salary, bonus, and benefits**

With salary analysis, the company will know How long was the last time an employee got a raise

### **3.3.5. Job title**

With Job title analysis, the company will know How long an employee in his / her current position has and Then calculate seniority and seniority allowance for employees

## 3.4. IT requirements Analysis

### **3.4.1. Dashboarding and Data Visualization**

Data dashboards are progress reports as a data visualization provides an objective view of performance metrics and serves as an effective foundation for further dialogue immediately understood. A dashboard is a business intelligence tool used to display data visualizations in a way that is immediately understood.

Data visualization is a way of representing data under visual images, charts, and tables. From there, convey information to viewers more vividly and understandably.

HR dashboard is a tool for businesses that allows Human Resource teams to help track, analyze, report on HR KPIs, and display data, usually to gain deeper insight into the overall well-being of the HR department. This dashboard incorporates an HR analytics platform that makes it easy to combine data from all systems and deep-explore this data directly in the dashboard. Thereby, the human resources department will optimize workplace management, improve employee performance and improve the quality of personnel in recruitment.

In this project, we use related BI platforms such as Tableau, Power BI, and Excel to provide visualize the following cleaning and integration process with the following features:

* Dashboard.
* Interactive data visualization.
* Filtering
* Ability to drill-down and drill up.
* Geospatial visualization and Maps.

### **3.4.2. Data Management**

The data management process includes a combination of different functions that collectively aim to make sure that the data in corporate systems is accurate, available, and accessible to ensure greater visibility and more accurate results overall, combining data from multiple data sets and different file formats from disparate sources to create a single, data warehouse or dataset ready for processing or analysis:

* Data Modeling.
* Data Blending.
* OLAP and Multi-Dimensional Analysis

### **3.4.3. Data Querying**

Data querying is a request for a database’s data so we can retrieve or manipulate it, perform calculations, automate tasks, or dig deeper through data mining, which uncovers hidden trends and relationships between data points.

### **3.4.4. HR Reporting**

HR reporting should provide a comprehensive view of what's working and what's not, highlighting specific issues that affect people management. The HR data and metrics used to indicate workforce trends will depend on the people management goals of your organization.

* Human resource monitoring: Capture the situation of the departments by the criteria of employees.
* Information management: Helping leaders manage the growth of each team and department, making employee retention decisions when they perceive risks such as longer employee training time with employees leaving.

**3.5 Comparative Analysis of BI and Data Visualization Tools**

### **3.5.1. Surveying and evaluation**

|  |  |  |  |
| --- | --- | --- | --- |
|  | **Tableau** | **Power BI** | **Google Data Studio** |
| **Features** | Two main features make Tableau stand out -built-in data blending and real-time collaboration. The latter feature comes with several ways to share reports.  Another selling point of Tableau is the flexibility of its dashboards.  But Tableau does require you to carry out initial data preparation. | A major advantage of Power BI is its simplicity to share visualizations, including real-time interactive dashboards. Other features depend on the version.  The main downside is that, on all versions, users have reported performance issues like timeouts and freezes when streaming or importing large data sets. | A top feature of Google Data Studio is its capability to turn raw data into visualizations. Sharing is just like using Google Drive.  there are only 50 functions from the calculation — fewer than the other data visualization tools. Lastly, Google Data Studio can only work with clean data. |
| **Visualization** | Tableau offers unique ways to display analytics that are not present in any other data analysis visualization tools. | Power BI data visualization is highly customizable. Like Google Data Studio, it uses the drag-and-drop method for building charts. You can also improve your reports with an expansive library of built-in visuals. | Google data visualization is enhanced by a library of built-in visual types. You can only use and modify the built-in visuals; you cannot add your own. |
| **Usability** | Easy to use. | Instantly familiar to Windows users but complicated for those inexperienced with excel. | Good usability. |
| **Integration** | You can connect Tableau to more than 30 different data source types, including data systems organized in file formats, relational and non-relational data systems, and cloud systems. | There is limited support for Google services.  Power BI supports include Salesforce, Facebook, SQL databases, SAP, and, of course, Microsoft Analysis Services.  In addition, you can use stored files on your computer. | Most Google services are fully integrated. Google Data Studio also supports external connectors like SuperMetrics and Funnel but has a hefty price tag. |
| **Price** | - Tableau Desktop: Individual: $999/year  Enterprises: $1,999/ year.  - Tableau Online: $500 per year per user  - Tableau Server: $10,000 for every 10 users. | A free one and an enterprise one.  Enterprise version: $9.99 per user per month. | Free. |

*Table 3.7 Comparative Analysis of BI and Data Visualization Tools*

Data visualization tools like Tableau, Google Data Studio, and Power BI all enhance your capability to display and examine analytics.

The company chose Power BI based on the size of the company as well as the purpose of its use (Microsoft).

* Make a list of data sources and ensure that your BI tool can test them. Set up data analysis processes and ensure high-quality data. Consider a data warehouse's architecture.
* Build various types of reports and a dashboard.
* System test.
* User training manual system.

### **3.5.2. Proposing BI solution for the project**

#### **3.5.2.1 Data Source:**

The data source may be a database, a flat-file, live measurements from physical devices, scraped web data, or any of the myriad static and streaming data services that abound across the internet.

Data Resource is raw data taken from a daily database of business and transactions corporate data.

This data source is taken from many different departments such as sales, purchasing, human resources,...and from various sources such as ERP, SCM, and CRM.

#### **3.5.2.2 ETL (Extract, Transform and Load)**

Data extraction: Sources could include legacy applications like Mainframes, customized applications, Point of contact devices like ATM, Call switches, text files, spreadsheets, ERP, data from vendors, partners amongst others.

Data transformation: The ETL process adds value and changes data such that insightful BI reports can be generated.

Data loading: Loading data into the target data warehouse database is the last step of the ETL process At this stage, the transformed data is uploaded to the repository.

#### **3.5.2.3 Building DataWarehouse:**

This is a database that is designed in a different model from a regular OLTP database and is a data store that holds the organization's long-term data.

Data Warehouse can only be read and be updated/written by ETL tool to convert data from Data Source to Data Warehouse not used for writing or normal updating applications

#### **3.5.2.4 Data analytics:**

Tools: Tableau, Excel, Power BI, SSAS, MDX.

Works: People analytics, workforce analytics, or talent analytics, revolves around analyzing people's problems using data.

Analyze the data to answer some of the questions or requests we posed earlier.

# **CHAPTER 4: BUIDING DATA WAREHOUSE AND INTEGRATING DATA**

## 4.1. Designing Data Warehouse

### **4.1.1. Bus Matrix**

#### **4.1.1.1 What is Bus Matrix?**

Bus Matrix is the tool for designing and communicating the enterprise data warehouse bus architecture. Besides the technical design considerations, the bus matrix is used as input to prioritize DW/BI projects with business management as teams should implement one row of the matrix at a time.

**4.1.1.2 Structure:**

Consists of rows and columns where the rows of the matrix are business processes and columns are dimensions.

#### **4.1.1.3 Benefits:**

- There is a common overall architectural design framework.

- Bus Matrix allows us to easily identify which facts have the same dimension.

- It is possible based on Bus Matrix to evaluate which tables can be combined into one report.

#### **4.1.1.4 Design of Bus Matrix:**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | Dim\_Emplyee | Dim\_Address | Dim\_Gender | Dim\_JobTitle | Dim\_Time |
| Fact\_Demography | x | x | x |  |  |
| Fact\_Recruitment | x |  |  | x | x |
| Fact\_Performance | x |  |  | x |  |
| Fact\_Salary | x |  |  | x | x |
| Fact\_Seniority allowances | x |  |  |  | x |

*Table 4.1 Human Resource Bus Matrix*

The Columns in the BusMatrix represent common dimensions used in corporate employee attributes, including Address, Gender, JobTitle, Time, and Employee.

The Row in the BusMatrix represents events related to work progress, including Performance, Demography, Recruitment, Salary, and Seniority allowances.

After locating the core processes and identifying the dimensions in employee attributes, you can easily see the logical and interrelated relationships between them by looking at the dimensions of evidence.

The "X" label in the matrix boxes denotes the relationship between the dimensions and the truth.

### **4.1.2. Master Data**

#### **4.1.2.1 What is master data?**

Master data is the set of identifiers that provides context to business data. It is the core data that is essential for running operations within a business enterprise or unit.

#### **4.1.2.2 Analysis:**

To get started with HR analytics, a human resource manager needs to map and collect all the relevant data.

HR Analytics helps your organization become more strategic, data helps use tackle current issues and also plan better for future activities.

* Analyze data on employee accommodation from which to know that employees mainly live in areas, thereby allocating resources reasonably at branches to bring convenience to employees.
* Analyze data on job titles to know if there are vacancies and need to recruit employees, thereby making a timely recruitment plan for the company.
* Analyze the data revenue of each employee to evaluate whether the employee has achieved the KPI or not, thereby offering appropriate solutions to improve the achieved performance.
* The data salary increase last analysis is how long to implement the salary increase process for employees
* Analyze seniority data to determine the seniority level for employees.

After researching a number of HR requirements, we figured out how to attract talent and make our workplace productive by evaluating some standards above:

* Which job is the employee suitable for?
* Based on the above criteria, is the salary paid to excellent individuals suitable?
* In which departments are employees performing better than the rest?
* Employees in which city are the most crowded in the company?

To adapt to those requirements, we collected and identified our employees by

* Gender: In addition to gender equality, other genders must also be respected. In some departments, it is necessary to consider which gender will work better.
* Job Title: Depending on the nature of the job, the human resources department evaluates the appropriate salary. Consider following the same job position with similar companies.
* Address: Based on employee address analysis assign them to convenient branches for work.
* Employee: Defined all general data about an employee such as name, address, email, department, salaries-flag… We can see how many employees in each department, which age group has the most employees, and also calculate their salaries.
* Time: Keep our normal date and time dimension table small, all our facts link to the UTC date/time keys, then if we need to report/group by a different time zone we just have to join through the time zone bridge table and link the local date/time keys back to the date and time dimension tables.

### **4.1.3 Transaction Data**

*Transaction Data* is data describing an event and is usually described with verbs. Transaction data always has a time dimension a numerical value and refers to one more objects.

**Demography**: this is the connection between Employee, Gender, Address to easily analyze the demographics of each employee across gender, marital status.

**Performance**: the most significant task in HR analytics is evaluating how good an employee is by their attendance in the workplace. This can connect with Employee to calculate.

**Recruitment**: this is the connection between Employee and Job Title, we can know the time from recruiting someone to complete successfully.

**Salary**: we separate into this fact table in order not to duplicate the same data about salary (rate, rate date change) in Employee.

### **4.1.4 Fact and dimension tables**

**Dim\_Employee**

Dim\_Employee table is used to store employee information.

|  |  |  |
| --- | --- | --- |
| **Column Name** | **Signification** | **Data Type** |
| EmployeeKey | Primary key for Dim\_Employee records. | Int |
| BusinessEntityID | Primary key for Employee records. Foreign key to BusinessEntity.BusinessEntityID. | Int |
| FirstName | First name of the person. | Nvarchar(50) |
| LastName | Last name of the person. | Nvarchar(50) |
| MaritalStatus | M = Married, S = Single | Nchar(1) |
| Age | Age of the person | Int |
| AgeGroup | Age groups from 20 to over 71 | Varchar(5) |
| DepartmentName | Name of the department. | Nvarchar(50) |
| DepartmentGroupName | Name of the group to which the department belongs. | Nvarchar(50) |
| ShiftName | Shift description. | Nvarchar(50) |
| SalariedFlag | Job classification. 0 = Hourly, not exempt from collective bargaining. 1 = Salaried, exempt from collective bargaining. | Varchar(7) |

*Table 4.2 Description of Dim\_Employee*

**Dim\_Address**

Dim\_Address table is used to store employee addresses with hierarchy from country region -> province -> city -> street.

|  |  |  |
| --- | --- | --- |
| **Column Name** | **Signification** | **Data Type** |
| AddressKey | Primary key for Dim\_Address records. | Int |
| AddressID | Primary key for Person.Address records. | Int |
| AddressLine1 | First street address line. | Nvarchar(60) |
| City | Name of the city. | Nvarchar(30) |
| Province | Name of the province. | Nvarchar(50) |
| CountryRegion | Name of the country region. | Nvarchar(50) |

*Table 4.3 Description of Dim\_Address*

**Dim\_Gender**

Dim\_Gender table is used to store employee gender.

|  |  |  |
| --- | --- | --- |
| **Column Name** | **Signification** | **Data Type** |
| GenderKey | Primary key for Dim\_Gender records | Int |
| Gender | M = Male, F = Female | Nvarchar(1) |
| GenderName | Name of the gender | Varchar(6) |

*Table 4.4 Description of Dim\_Gender*

**Dim\_JobTitle**

Dim\_JobTitle table is used to store employee job titles.

|  |  |  |
| --- | --- | --- |
| **Column Name** | **Signification** | **Data Type** |
| JobTitleKey | Primary key for Dim\_JobTitle records. | Int |
| JobTitle | Work title such as Buyer or Sales Representative | Nvarchar(50) |
| OrganizationLevel | The depth of the employee in the corporate hierarchy. | Smallint |

*Table 4.5 Description of Dim\_JobTitle*

**Dim\_Time**

Dim\_Time table is used to store time with hierarchy from year -> quarter -> month -> day.

|  |  |  |
| --- | --- | --- |
| **Column Name** | **Signification** | **Data Type** |
| TimeKey | Primary key for Dim\_Time records. | Nvarchar(255) |
| Date |  | Datetime |
| Year |  | Int |
| Month |  | Nvarchar(2) |
| Day |  | Nvarchar(2) |
| DateName |  | Nvarchar(30) |
| DayNumberOfWeek |  | Int |
| DayNumberOfMonth |  | Int |
| DayNumberOfYear |  | Int |
| WeekNumberOfYear |  | Int |
| Quarter |  | Int |

*Table 4.6 Description of Dim\_Time*

**FactDemography**

FactDemography table contains references to DimAddress, DimEmployee and DimGender tables.

|  |  |  |
| --- | --- | --- |
| **Column Name** | **Signification** | **Data Type** |
| DemographyKey | Primary key of FactDemography table | Int |
| EmployeeKey | Foreign key of FactDemography table, primary key of DimEmployee table | Int |
| GenderKey | Foreign key of FactDemography table, primary key of DimGender table | Int |
| AddressKey | Foreign key of FactDemography, primary key of DimAddress table | Int |
| MaritalStatus | M= Married, S = Single | Nchar(10) |

*Table 4.7 Description of FactDemography*

**FactRecruitment**

FactRecruitment table contains references to DimEmployee and DimJobTitle tables.

|  |  |  |
| --- | --- | --- |
| **Column Name** | **Signification** | **Data Type** |
| RecruitmentKey | Primary key of FactRecruitment table | Int |
| JobTitleKey | Foreign key of FactRecruitment table, primary key of DimJobTitle table | Int |
| EmployeeKey | Foreign key of FactRecruitment table, primary key of DimEmployee table | Int |
| StartDate | Employee started on this date | Nvarchar(255) |
| YearsOfWork | Number years of work of employee | Int |
| OrganizationLevel | The depth of the employee in the corporate hierarchy. | Int |

*Table 4.8 Description of FactRecruitment*

**FactPerformance**

FactPerformance table contains references to DimEmployee table.

|  |  |  |
| --- | --- | --- |
| **Column Name** | **Signification** | **Data Type** |
| PerformanceKey | Primary key of FactPerformance table | Int |
| EmployeeKey | Foreign key to FactPerformance, primary key for DimEmployee table | Int |
| VacationHours | Number of vacation hours. | Smallint |
| SickLeaveHours | Number of sick leave hours. | Smallint |
| Bonus | Number of bonus | Int |
| SalesYTD | Number of sales YTD | Int |
| SaleLastYear | Number of sale last year | Int |
| StartDate | The date which the employee starts working | Int |

*Table 4.9 Description of FactPerformance*

**FactSalary**

FactSalary table contains references to DimEmployee table.

|  |  |  |
| --- | --- | --- |
| **Column Name** | **Signification** | **Data Type** |
| SalaryKey | Primary key of FactSalary table | Int |
| EmployeeKey | Foreign key of FactSalary, primary key of DimEmployee table | Int |
| Rate | Salary hourly rate. | Money |
| RateChangeDate | Date the change in pay is effective | Nvarchar(255) |

*Table 4.10 Description of FactSalary*

### **4.1.5. Snowflake model**

Diagram

Description automatically generated

*Figure 4.1 Snowflake model*

## 4.2. ETL process

### **4.2.1. Dimension Table’s ETL Process**

**DimEmployee:**

A screenshot of a computer

Description automatically generated

*Figure 4.2 DimEmployee’s ETL process*

**DimAddress:**

A screenshot of a computer

Description automatically generated with medium confidence

*Figure 4.3 DimAddress’s ETL process*

**DimJobTitle:**

A screenshot of a computer

Description automatically generated

*Figure 4.4 DimJobTitle’s ETL process*

**DimGender:**

A screenshot of a computer

Description automatically generated

*Figure 4.5 DimGender’s ETL process*

### **4.2.2. Fact Table’s ETL Process**

**FactPerformance:**

A screenshot of a computer

Description automatically generated

*Figure 4.6 FactPerformance’s ETL process*

**FactSalary:**

A screenshot of a computer

Description automatically generated

*Figure 4.7 FactSalary’s ETL process*

**FactRecruitment:**

A screenshot of a computer

Description automatically generated with medium confidence

*Figure 4.8 FactRecruitment’s ETL process*

**Fact Demography:**

A screenshot of a computer

Description automatically generated with medium confidence

*Figure 4.9 Fact Demography’s ETL process*

# **CHAPTER 5: DATA ANALYTICS**

## 5.1. Data analytics with SSAS technology

### **5.1.1. Building the cube**

A screenshot of a computer

Description automatically generated with medium confidence

*Figure 5.1 Cube Structure*

### **5.1.2. Analysis with SSAS**

**SQL Server Analysis Services (SSAS)** is a multi-dimensional OLAP server as well as an analytics engine that allows you to slice and dice large volumes of data. It is part of Microsoft SQL Server and helps perform analysis using various dimensions.

* + - 1. **. Demography**

**Number and proportion of employees for each age group in the entire company:**

Graphical user interface, text, application, email

Description automatically generated

*Figure 5.2 Number and proportion of employees for each age group in the entire company*

The company has an ageing workforce (the age of employees is from 30-70 years old). The majority of employees are in the age group of 30 -45 years, accounting for 61%. The percentage of employees assigned to the retirement age bands is very high ( from 61-70 years old, accounting for 7.2%). An aging workforce has a lot of experience in the job, but it also causes many problems for the company such as the high cost of paying employees along with social benefits. HR department needs to come up with HR policies to recruit younger employees.

**The number of employees by Gender:**

Graphical user interface, text, application

Description automatically generated

*Figure 5.3 The number of employees by Gender*

Beside the age issue, the gender imbalance in Adventure company is quite high, with female employees accounting for 29% of the total number of employees.

**The number of employees by country region:**

Graphical user interface, text, application, email

Description automatically generated

*Figure 5.4 The number of employees by country region*

From the employee's country regionThe majority of the company's employees are concentrated in the United States 284 employees (accounting for nearly 98,3%).The number of employees in the external area is only six employees. The HR department can review allowances for remote employees or provide accommodation for employees.

**The number of employees by Marital Status:**

Graphical user interface, application, Word

Description automatically generated

*Figure 5.5 The number of employees by Marital Status*

Marital status of employees in the company: the number of employees who are married (M=146) and the number of employees who are single (S=144) are almost equal.

* + - 1. **. Recruitment**

**The number of employees hired by Organizational level:**

Graphical user interface, application, Word

Description automatically generated

*Figure 5.6 The number of employees hired by Organizational level*

The number of employees hired in the corporate hierarchy of the company such as the executive level, only one employee is hired. Here, the most hired fourth level is the staff in the production department which needs many employees to carry out the production process.

**The number of full/part time employees hired by years:**

Graphical user interface

Description automatically generated

*Figure 5.7 The number of full/part time employees hired by years*

The data shows the number of full-time employees or the number of part-time employees from year to year. True is exempt from collective bargaining. The company has a lower percentage of exempt employees than employees. Human resources should increase the percentage of this number of employees. The company recruited many part-time employees with the number of 63 and 130 respectively in 2008 nad 2009.

**The number of employees hired by department name and years:**

Graphical user interface, application

Description automatically generated

*Figure 5.8 The number of full/part time employees hired by years*

The table shows the number of employees for each department over the years. Based on the above information, the HR department can issue annual recruitment rounds. Total Manuafacturing staff hired in 2009 the highest with 107 employees.

* + - 1. **. Performance**

**Sales revenue of sales staff:**

Graphical user interface, text, application

Description automatically generated

*Figure 5.9 Sales revenue of sales staff*

The data shows the increase or decrease in sales of each Sales department employee between this year (Sales YTD) and last year (Sales last year). The employee with the highest KPI of the year is Stephen. The number of employee KPIs this year increased much higher than last year, and employee Linda had the highest revenue growth rate of three times compared to last year. Only three employees lost sales compared to last year (Pamela, Garret, and Lynn) but the number of reductions was not large.

**Number of days off of each department:**

Graphical user interface, application

Description automatically generated

*Figure 5.10 Number of days off of each department*

The table shows each relationship between Department, Sick leave hours, and Vacation hours. Help HR manage salary and total time off of each department to rank the departments. The Production department has the highest sick leave our (7929) and vacation hours (8787). From here, it is possible to evaluate the number of days off of each employee, closely monitor management, and improve the quality of human resources.

* + - 1. **. Job title**

**The number of years of work of all employee:**

Graphical user interface, application

Description automatically generated

*Figure 5.11 The number of years of work of all employee*

The table shows the employee's years of service in each job position. The average number of years working for employees in the company is 13 years, and employees in the company have high working seniority. Employee work engagement shows employee satisfaction with company policies, and employees help the company grow stronger (increased sales as above).

### **5.1.3. Building KPIs system**

* + - 1. **. AbsenceRate KPI**

Table

Description automatically generated

Chart

Description automatically generated with medium confidence

*Figure 5.12 AbsenceRate KPI*

1. **Value Expression:**

[Measures].[AbsenceRate] = ([Measures].[Sick Leave Hours] +

[Measures].[Vacation Hours]) / ( 2080 \* [Measures].[Fact Performance Count])

1. **Goal Expression:**

case

when [Dim Employee].[Department Group Name] is [Dim Employee].[Department

Group Name].&[Executive General and Administration] Then 0.06

when [Dim Employee].[Department Group Name].CurrentMember is [Dim

Employee].[Department Group Name].&[Inventory Management] Then 0.05

when [Dim Employee].[Department Group Name].CurrentMember is [Dim

Employee].[Department Group Name].&[Manufacturing] Then 0.045

when [Dim Employee].[Department Group Name].CurrentMember is [Dim

Employee].[Department Group Name].&[Quality Assurance] Then 0.06

when [Dim Employee].[Department Group Name].CurrentMember is [Dim

Employee].[Department Group Name].&[Research and Development] Then 0.03

when [Dim Employee].[Department Group Name].CurrentMember is [Dim

Employee].[Department Group Name].&[Sales and Marketing] Then 0.04

else 0.04

End

1. **Status Expression:**

case when

KPIVALUE("KPIAbsenceRate")-KPIGOAL("KPIAbsenceRate")<0 then 1

when KPIVALUE("KPIAbsenceRate")-KPIGOAL("KPIAbsenceRate")=0 then 0

else -1

end

1. **Trend Expression:**

case when ISEMPTY(PARALLELPERIOD([Dim Time].[Year].[Year],1,[Dim

Time].[Year])) then 0

when

[Measures].[AbsenceRate] > (PARALLELPERIOD([Dim Time].[Year].[Year] , 1 ,

[Dim Time].[Year]),[Measures].[AbsenceRate]) then 1

when

[Measures].[AbsenceRate]= (PARALLELPERIOD([Dim Time].[Year].[Year] , 1 ,

[Dim Time].[Year]),[Measures].[AbsenceRate]) then 0

else -1

end

* + - 1. **. Sales KPI**



Table

Description automatically generated









A picture containing table

Description automatically generated

*Figure 5.13 Sales KPI*

1. **Value Expression**

[Measures].[Sales Last Year]

1. **Goal Expression:**

case

when [Dim Job Title].[Job Title] is [Dim Job Title].[Job Title].&[Sales Representative] then 1600000

else 0

end

1. **Status Expression:**

case

when

KPIVALUE("KPISales")/KPIGOAL("KPISales") > 0.9 then 1

else -1

end

# **CHAPTER 6: VISUALIZATION AND FORCASTING OR PREDICTIVE MODEL**

## 6.1. Report and dashboard systems (structure)

Table

Description automatically generated with medium confidence

## 6.2. Data analysis with Power BI

**Demography**

Graphical user interface, application

Description automatically generated

*Figure 6.1 The dashboard of demography*

According to the chart of the total number of employees by age group, the age group 30 - 45 accounts for the highest, followed by the age group 46 - 60 and finally the age group 61 - 7. The age group 30 - 45 accounts for the highest. because at this age they have the best experience and ability to work. For the age group 46 - 60, they will have more experience to be able to develop for their company. The lowest surveyed age is because they can't work like in their 30s - 45s, at this age they will gradually lose their ability to work.

According to the graph of the proportion of employees by sex, more than half of men are 71.03% and women are only 28.97%. The proportion of male employees accounting for more than half is also easy to see that because men will be the breadwinners in the family, they will be the ones going out to earn money, while women will mostly work when they get married. do more housework and women also limit their work than men.

According to a survey of the number of employees in each region, the United States accounts for the highest of all regions. Countries such as Canada, Australia, France, Germany, and the UK account for a very small percentage. Because the United States is the 3rd most populous country in the world, it also has a larger proportion of employees in this country than in other countries. And that's because the United States is a very developed country that creates opportunities for employees to grow.

Looking at the chart of the employee's marital status, the marriage and marital status are quite equal.

**Recruitment**

Graphical user interface, chart, application

Description automatically generated

*Figure 6.2 Recruitment*

When looking at the chart, we can see that the 4th level accounts for a rather high percentage compared to other levels, accounting for more than half of 63.85%, the second level is level 3 with 23.99%, followed by the second level. 2 with 9.12%, level 1 with 2.7%.

According to the survey chart of the percentage of employees in the departments, the production department has the highest proportion of 179 departments, followed by sales at 18, purchasing 14, and the lowest percentage of employees in other departments. departments such as engineering, human resources, quality assurance.

According to the graph of the number of employees employed, in 2009 was the most employed and in 2006 accounted for the lowest percentage.

**Performance**

Graphical user interface

Description automatically generated

*Figure 6.3 Performance*

Looking at the chart, we see that these are the top 10 employees with the highest revenue last year. Occupying the highest is Varkey Chudukatil with 2.4M, second is Tsoflias with 2.3M, third place with 2.1M in revenue is named Ito.

The chart of the top 10 employees with the highest sales to date, with 4.3M in revenue was led by an employee named Milchell, followed by Pak employee with 4.1M in revenue, the third is Blythe with revenue of 3.8M.

According to the chart of the top 5 departments with the highest number of sick hours, the first place is the manufacturing department with 7.9k, the two Sales and Purchasing departments with 0.6k, and the last two with the number of hours off. 0.5k break is Information Service and Finance.

The chart of the top 5 employees with the most sick hours, the employee with the most sick leave is named Hill, second is Liu, followed by Brown, Johnson and finally Kane with 124 hours.

The graph shows that the top 5 employees with the most vacation hours are also the employees who take the most sick leave as shown above, still ranked first by Hill with 206 hours and the last person to change Johnson with the number of hours. 162 hours.

**Job Title**

**Chart, bar chart

Description automatically generated**

*Figure 6.4 Top 10 the Employees havve the most Years Of Work*

Here are the top 10 employees with the most years of service, the first person is Johnson with 42 years of service, Brown and Liu with 41 years of work only 1 year away from the first person, the fourth is Miller with 38 years. years of work and the 10th place holder is Walters with 27 years of work experience

**CHAPTER 7: CONCLUSION AND FUTURE WORKS**

**7.1. Results**

* After implementing the BI solution, we had a clearer view of the human resources situation of our company.
* BI solutions help companies better understand the human resources situation in terms.
* Determine the company's staffing needs: The company needs to recruit more young employees because the company has an ageing workforce. In addition, recruiting more female staff due to the high gender disparity. Need to recruit more official staff because the number of part-time employees of the company is high.
* From Bi solution can promote a culture of employee care, and identify the needs of employees. Based on the address, marital status, age, and working seniority to determine bonus, social benefits, and pension for employees, this promotes employee morale and aids your brand image externally.
* Identify employee performance work based on KPI.

**7.2. Limitations**

* The limitations of the HR department have not been exploited yet
* The data is not enough so some parts cannot be mined. For example, there is no salary to exploit the salary section, or the JobCandidate table section has very little data to exploit the recruitment KPI section.
* Since it was the first time, the team had little experience, had to do it many times and still didn't do it well

**7.3. Future works**

In the future, the team will clarify some issues about recruitment, salary and employee loyalty to know whether the recruitment is effective or not, and how the level of company's recognition and preference for the workforce is. Also finding the weak point of the department and find ways to improve it.

Currently, the BI solution is only applied in a part, Human Resource, with the current effectiveness, we will try to apply the BI solution to other parts of the company with the top priority being the department. sales department.

**REFERENCES**

[1] 5 Purposes of the Human Resources Department. (2021, May 18). Human Resources Degrees. <https://www.humanresourcesmba.net/lists/5-purposes-of-the-human-resources-department/>

[2] Advantages and Disadvantages of Performance Indicators. (2018). Unboxed Training & Technology. <https://www.unboxedtechnology.com/blog/advantages-and-disadvantages-of-performance-indicators/>

[3] A Comparative Analysis of Top 6 BI and Data Visualization Tools in 2018. (2018). ActiveWizards: Data Science and Engineering Lab. <https://activewizards.com/blog/a-comparative-analysis-of-top-6-bi-and-data-visualization-tools-in-2018/>

[4] D&Apos;Angelo, M. (2021, December 10). How to Process Payroll in 8 Easy Steps. Business News Daily. <https://www.businessnewsdaily.com/12008-how-to-process-payroll.html>

[5] Education, I. C. (2021, April 15). OLAP. Ibm. https://www.ibm.com/cloud/learn/olap

[6] Education, I. C. (2022, May 3). ETL (Extract, Transform, Load). ETL (Extract, Transform, Load). <https://www.ibm.com/cloud/learn/etl#:%7E:text=ETL%2C%20which%20stands%20for%20extract,warehouse%20or%20other%20target%20system>.

[7] Enterprise Data Warehouse Bus Matrix | Kimball Dimensional Modeling Techniques. (2018). Kimball Group. <https://www.kimballgroup.com/data-warehouse-business-intelligence-resources/kimball-techniques/dimensional-modeling-techniques/enterprise-data-warehouse-bus-matrix/>

[8] Human Resources (HR). (2020, July 29). Investopedia. <https://www.investopedia.com/terms/h/humanresources.asp>

[9] Kopecky, C. (2019). What is a database query? SQL and NoSQL queries explained. Educative: Interactive Courses for Software Developers. <https://www.educative.io/blog/what-is-database-query-sql-nosql>

[10] M. (2022a, February 16). What is Azure Analysis Services. Microsoft Docs. <https://docs.microsoft.com/en-us/azure/analysis-services/analysis-services-overview>

[11] M. (2022b, April 1). The Basic MDX Query (MDX). Microsoft Docs. <https://docs.microsoft.com/en-us/analysis-services/multidimensional-models/mdx/mdx-query-the-basic-query?view=sql-analysis-services-2022>

[12] Stedman, C., & Vaughan, J. (2019, October 30). What is data management and why is it important? SearchDataManagement. <https://www.techtarget.com/searchdatamanagement/definition/data-management>

[13] Taylor, D. (2022, April 16). ETL (Extract, Transform, and Load) Process in Data Warehouse. Guru99. <https://www.guru99.com/etl-extract-load-process.html>

[14] What is Master Data? (2018). TIBCO Software. <https://www.tibco.com/reference-center/what-is-master-data>

[15] 5 Purposes of the Human Resources Department. (2021, May 18). Human Resources Degrees. <https://www.humanresourcesmba.net/lists/5-purposes-of-the-human-resources-department/>

[16] Data Analytics in Business, VNU-HCM (2016). Ho Trung Thanh, Trieu Viet Cuong, Vu Thuy Hang, Le Hoang Su, Le Thi Kim Hien