BI TERM PROJECT

Yayuan Zhang / Selina

MAY 8, 2022

ISM 6255 – KNOWLEDGE MANAGEMENT & BUSINESS INTELLIGENCE

CONTENT

Business justification

Background Information

Original Access database: Employee information management system.accdb

The employee management system is an indispensable part of enterprises. The benefit is the most important part and it depends on the work efficiency of employees. Facing with a large amount of employee information and constantly changes, manual process is inefficient and inaccurate. This database system provides users with managing employee information that is simplified with fast query methods to help enterprise improve management efficiency.

The system records the employee personal information, salary, department, employee skills, job level, attendance, personnel transfer and resignation information.

Even though this database includes some report generation functions, it still lacks of real-time analytical reports with dimensions. Companies also need rapidly generated reports/data to aid decision making, such as annual/quarterly/monthly report on the number of hired employees, report on number of soft skill holders and their annual salary, report on number of courses taken by professional level employees, and etc.

Development purpose

In order to provide the company's competitiveness, the company pays attention to the employee's personal ability and employee value (including potential value, such as some soft skills). Managers expect to develop high-potential employees and career advancement with the goal of realizing the employee's value proposition.

So, we need a report/table where we can track the real-time salary/bonus/training courses taken of employee's and realize all-round talent management.

The purpose of the project is to provide all employees with relevant information by accessing data by three dimensions. Naturally, all employees should not have access to all data, as this would pose a threat to personal privacy. For instance, a regular employee should not have seen the private information from another employees, such as home address or marital status, etc. Therefore, permission to access data is required. By linking data sets, I will convert large amounts of data into simple reports/PivotTables, it allows them to capture and analyze collected data through multiple dimensions.

It has a better overall understanding of the operation of enterprises to help clearly understand the current situation of employees and tap their potentials to provide valuable advice on employee promotion, salary adjustment, career development and training courses taken.

Goals

Identify relevant data: find out what information can be useful.

Build a multidimensional database and a data cube

Use BI tool (Microsoft Excel) and connect Data Mart

Delivery offline usage

Positive outcomes & adding value

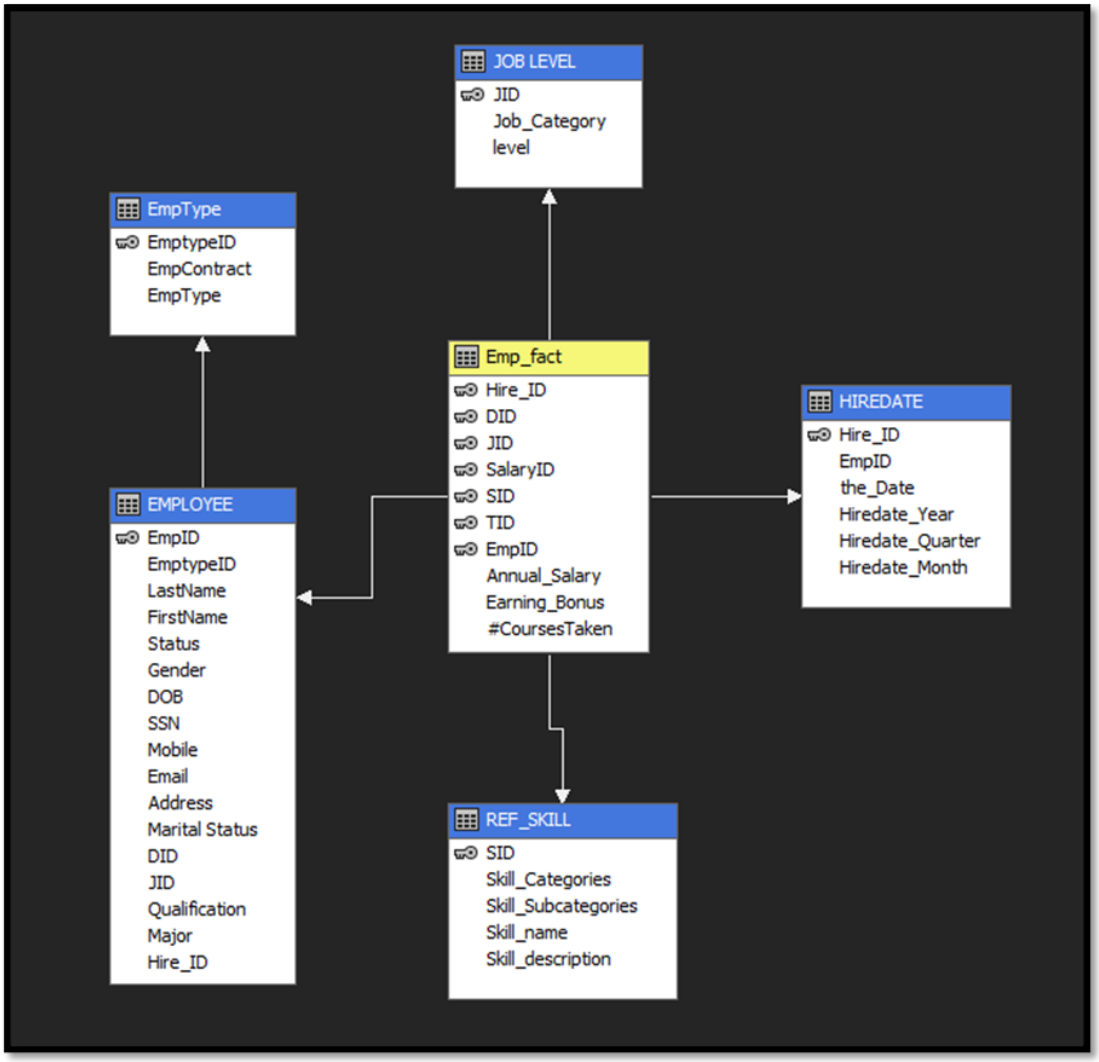
Improve work efficiency & Economic benefits

Use analytical tools effectively and integrate all data sources. Data will be updated in real-time mode and reduce manual operations, which greatly improves management and work efficiency. Through the automatic association of various management functions and simple page operations, they both can reduce labor input cost and time cost. It enhances the overall strength and competitiveness of an enterprise by avoiding a lot of repetitive work and improving work efficiency.

Enterprise management

Managers can grasp condition information at any time according to the report/PivotTables multidimensional. It allows them to retrieve the needed information to help the company make quick decisions.

Star schema



Why picked the facts/measures and dimensions?

Fact table: Emp_fact

I built the PK are Hire_ID, EmpID, DID, JID, SID, SalaryID, TID in this fact.

Within my fact table, I use measures Annual_Salary, earning bonus, #CoursesTaken, and emp fact count.

The most important question I must ask myself: what information I will need.

In my opinion, this is a good fact to have access to all the relevant data needed while protecting more personal information. This fact table covers the necessary information for the entire database without personal privacy information. This fact table has opportunities to observe the number of recruited employees, the number of employees with various skills in the company, their annual salaries and bonuses, and the number of training courses. But also, it allows clearly find the advice on employee promotion, and salary readjustment multidimensional.

Dimensions(4)

Hire Date: Year - Quarter - Month - Date

In my opinion, the entry date is a good dimension in this system. The date somebody joined the company is represented by the Hire_Date_ID in the fact and will lead to the date-dimension. This allows me for example to compare the number of resources and changes in career development that came IN and OUT over time.

Ideally, one dimension will be based on time. It allows viewing the annual salary, earning bonus and learning status of training courses of employees. It is more convenient to provide data analysis suggestions for employee promotion, salary adjustment, and training courses.

Type of employee: EmpType - EmpContract

Salaried employee: Annual salary, or stock optional

Hourly employee: hourly rate

Consultant: contact number, billing rate

This dimension is selected to more clearly present employees with different employment contracts. It allows uncovering the potential of different types of employees to rack the impact of annual salary and bonuses.

Job level: Job category - level

Job level dimension can be tracked by employee roles which are guaranteed to be assigned appropriately to employees in the organization. And whether to receive a salary or bonus that matches the job level of work.

Skill: Skill categories - skill subcategories - skill name

In this part, I divided data into more detailed dimensions, hard skill and soft skill. Hard skills are like the diploma, certificate, etc. Soft skill includes but not limited to time management, emotional intelligence, communication skill etc. This dimension is to check whether the ability is matched to their annual salary level and can check if some employees need training.

Post-mortem report for 3Ws

What went well?

Find the measures. Before I started this project, I asked myself one of the most important questions, what information will I need? These specs have easily told me

what measures I have to include.

The steps for ETL and Microsoft SQL Server, included build a cube, process the cube and brown the cube. Due to the equipment error, I practiced those steps many times when I was doing the assignment. So, this part went quite smoothly in my individual project.

Build the PivotTable, I did it when I was doing assignment. So, this part was relatively smooth.

What did not go well?

Original database(.accdb)

- There is no facts table in my original database, so I had to create a new table to complete the project.
- In the original database, there is no subdivision type of field in each table. So, I reorganized the database and focused on the table of hire data, job level, skill, and employee type. It allows my project able to search the corresponding employee information very well.
- Same ID should be set up with same data type. In fact, I spent a lot of time combing and supplement the source database. There are some missing parts in the original database, which made me unable to query cross these to go SQL Serve. For instance, when I tried to linking PK and FK in SQL Server, it is showing error due to some of my IDs are not written in the same data type. One of table set this data type as a number, and other table set a short text. I

must go back and reset the data type.

Some unknown system errors

Some unknown errors are encountered after completing the project led to redo this cube three times. After shutting down and restarting my laptop, the cube cannot run. I have checked many cases online, but still not find the corresponding answer for error. I guess it is because it has two logged-in users maybe. However, I did this cube all over again and decided not to shut down the laptop until committing.

What would you do different next time?

- Refine all raw data, attribute, tables as much detail as possible. When I build the database next time, I will improve the original data and attributes as much as possible. So that when I use the analysis tool later, I will do not need to add a lot of information and tables.
- Check data file(.accdb) has not changed before database up to SQL Server.
- Make sure the same ID from different table should be set to same data type, such as set as number or short text. It can cause successfully link the PK and FK in the star schema.
- Do not rush the next step until almost done.