REPORT

TOPIC : PROJECT AND BILLING TIMESHEET TRACKING

Elaborated by :

Amira DOGHRI

Eya AKRIMI

Ines ZHIOUA

Raoua LAHOUAGUE

Maroua ABBOUR

Nour TRABELSI

4 ERP-BI5

2020-2021

THE HOUSE OF DIGITAL EXPERTS

List of contents

List of figures

List of tables

General Introduction

Business intelligence is a subject in full Evolution, addressing the general direction as the trades. it is the decision support tool that enables a whole different business activities and its view .This view environment requires knowledge of the different Trades of the company and implies some organizational specifications . The implementation of projects BI can not be done without having defined a comprehensive BI strategy. So we can say that the Business Intelligence is the process of distill

So we can say that the Business Intelligence is the process of distilling information and knowledge from data. With this, companies try to find business insights that can help in running or improving the company. Today, the BI market offers relatively comprehensive solutions through analysis concerning aspects of reporting and data consolidation. The key to successful analytics is gaining timely insights from data that leads to better decision making.

g.It resolve their business challenges by turning high volumes of data into actionable insights that can enable them . to advance their business objectives, achieve their revenue goals, and sustain a competitive advantage. The business insights can be disclosed via different data presentations like reports, dashboards and visualizations, or via analysis tools like OLAP or advanced statistical learning tools.

Usually there are three different ways to incorporate the business insights in the business:

- Managed reports that are periodically refreshed.
- Self-service analytics.
- Input for operational systems .

Business intelligence also includes all activities to gather, prepare or decide on the necessary data and data definitions. Examples of these activities are:

- Data warehouse modeling.
- Data integration and ETL (extraction, transformation, load).
- Data governance for data definitions and data ownership.
- Validating data quality.

Chapter 1: Project's Context

1. Introduction

In this chapter, we are going to present the context of our project, starting with study of existing, then specifying the objectives of our project and finally presenting the different solutions.

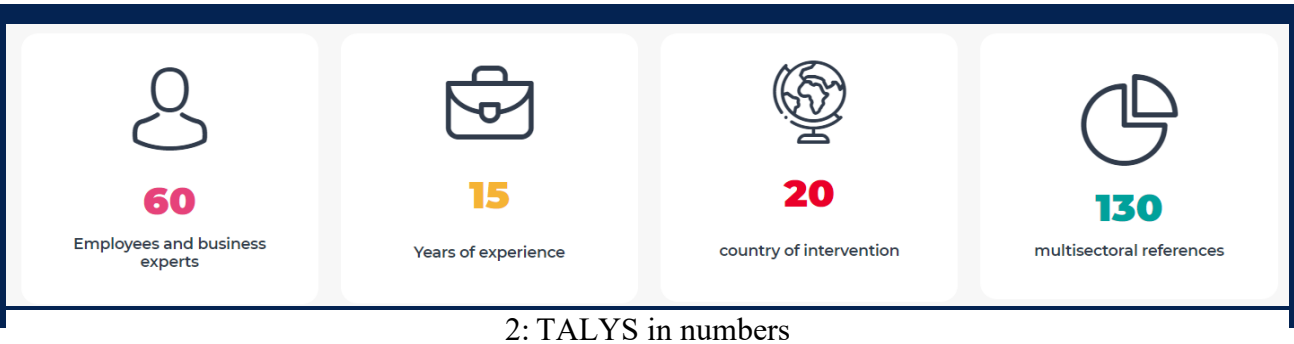
2. The host organization

TALYS is a group of technological companies specializing in Organization, **Information Systems** and **Digital Transformation**.

Since 2006, its two founders **Hatem MSADAA** and **Elyssa AOUNALLAH** have had the desire to **create added value for customers** by supporting them in **the realization and success of their** innovative **projects** and the transformation of their business.

Today, TALYS is positioned as a major player in the financial **sector** in Tunisia and has succeeded in extending its **scope of action on 3 continents (Africa, Europe and the Middle East).** This is thanks to a team of **functional and technical experts** who combine several business expertise: **Banking**, insurance, **microfinance** and **leasing** .

1 : position of TALYS



2: TALYS in numbers

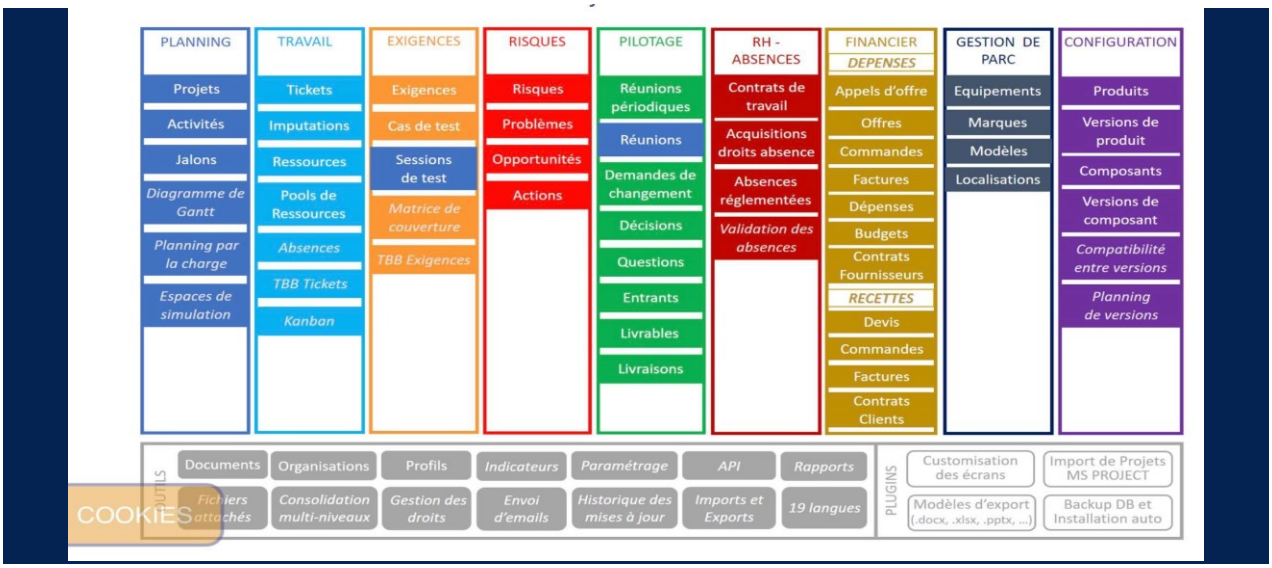
3. The used application

In a decidedly digital and connected world, TALYS manages their projects through the project management tool ProjeQtOr.

Quality based Open-Source Project Organizer.

Complete, Collaborative, Quality based Open-Source Project Organizer.

ProjeQtOr (formerly Project'Or RIA) is a collaborative project management software. It is a tool designed to be a Project Organizer as a Rich Internet Application. Web based, it is very easy to use and targets to include every feature needed to the management of your projects.



3: Features of PROJEQTOR

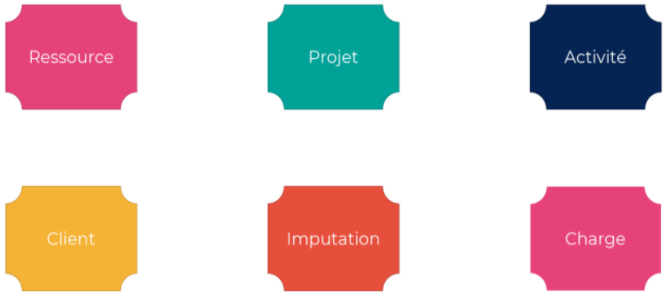
4. Problematic

The application limits

This project management tool is limited in the reporting and visualization area:

TALYS employees manage their **consulting** projects through ProjeQtor, in order to give a better-quality service, they need more visualizations on:

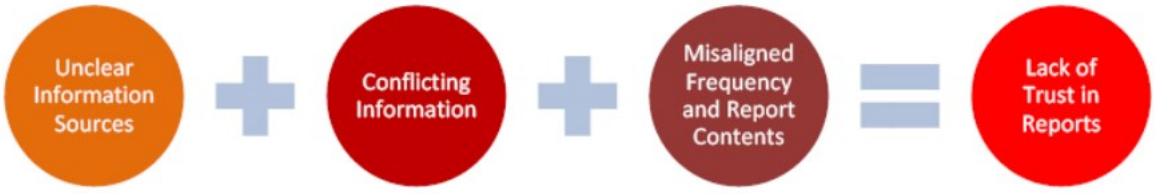
- Projects' deadlines: phases, delays, and timing
- Projects' and tasks assignment
- Teams/Consultants assigning to projects.
- Time/charge tracking and management.
- Billing progress
- Projects' activity sector



4: project keyword

Challenges

There could be scenarios where the information collection frequency and the quality of the collected data are not aligned with the Business Intelligence project requirements. This generally results in a lack of trust of the decision makers with the information obtained from the reports.



5: Challenges

In order to ensure the project's reliability, there are some crucial steps to do:

- Sources of information are identified and verified.
- Reporting mechanism is transparent, including the clear process through which the reports are created.
- Source used is accurate and relevant to the information required.
- Information is sufficiently specific and updated.

5. Conclusion

In this first chapter, we presented the context of our project in order to make things clear, and to help you more understand the main idea.

Chapter 2: Business Objectives & Functional requirements

1. Introduction

This chapter's main purpose is to present the specific steps or actions that must be taken to reach the company's goal

2. Business Objectives

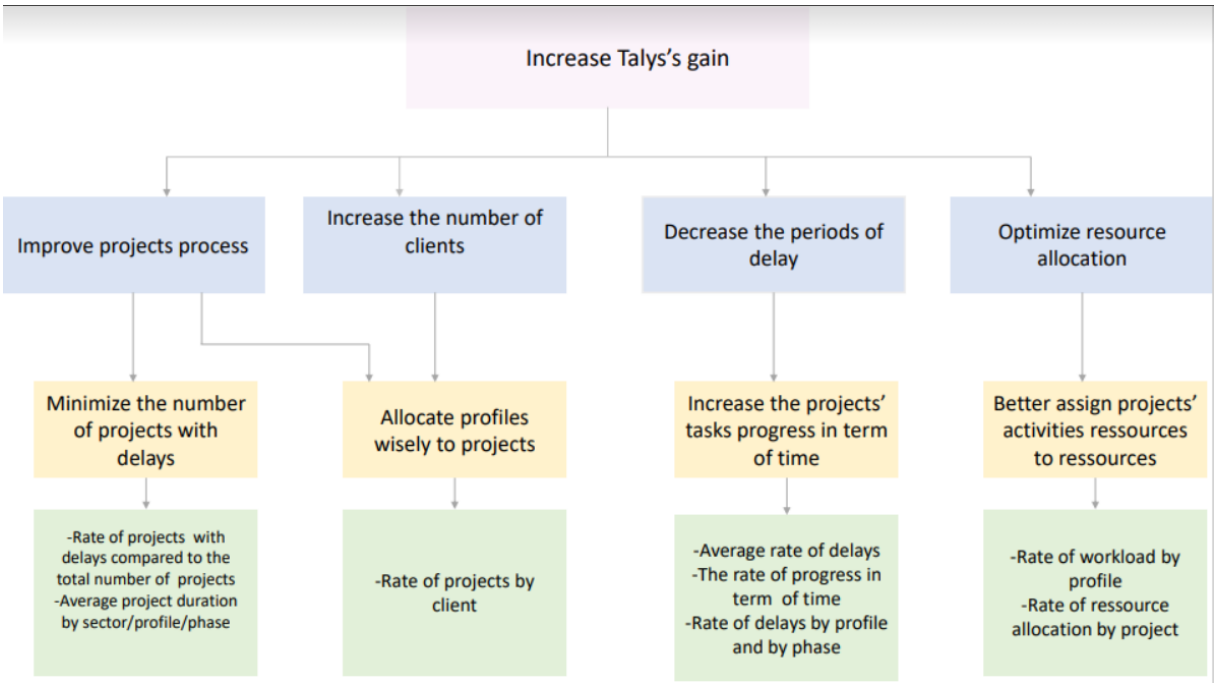
Business goals and objectives are part of the planning process. They describe what a company expects to accomplish throughout the year.

These goals and objectives might pertain to the company, departments, employees, customers and even marketing efforts.

For each business objective, functional requirements are derived in order to be deployed in the project.

The functional requirements come from the specifications of the project. These are the required needs by the end user. They are the features and actions that the dashboard must obligatorily carry out. The final dashboard must meet the following functional requirements:

The project mainly aims to help the company **improve its performance and lead a better strategy.**



6:Objective analysis tree

7:the project objectives

3. functional requirements

1:Business Objectives

Business Objective	*Functional requirements*
Reduce project delays	View projects with delays (by phase / sector) View projects without delays View delay rate by phase View projects by sector View projects by duration and phase View phases that caused delays
Optimize resource allocation	View the delay rate by profile. View project with delays by profile
Ameliorate the billing progress	View billed projects. View billing project rate by sector View billing project rate by profile

4. Conclusion

After understanding the data and setting the objectives, we will present, in the next chapter, the methodology.

Chapter 3: Methodology of work

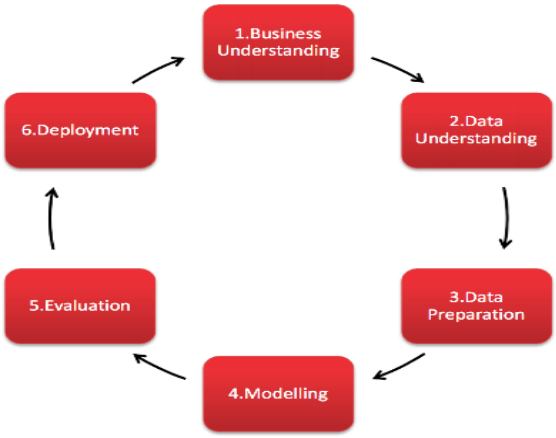
1. Introduction

The main objective of this chapter is to present the methodology that we will work on, in order to manage our project well

2. Methodology

What is CRISP-DM?

The **CR**oss **I**ndustry **S**tandard **P**rocess for **D**ata **M**ining (CRISP-DM) is a process model with six phases:



8 :Methodology

1) **Business understanding** – What does the business need?
2) **Data understandin**g – What data do we have / need? Is it clean?
3) **Data preparation** – How do we organize the data for modeling?
4) **Modeling** – What modeling techniques should we apply?
5) **Evaluation** – Which model best meets the business objectives?
6) **Deployment** – How do stakeholders access the results?

Why CRISP-DM?

➔ **Flexibility**: This methodology makes it possible for models and processes to be imperfect at the very beginning.

➔ **Long-term Strategy**: CRISP-DM methodology allows to create a long-term strategy based on short iterations at the beginning of project development. During first iterations, a team can create a basic and simple model cycle that can easily be improved in further iterations.

➔ **Functional Templates**: Using a CRISP-DM approach makes it possible to develop functional templates for project management processes.

Steps

Business Understanding: Give context to the goals and to the data so that the developer/engineer gets a notion of the relevance of data in a particular business model.

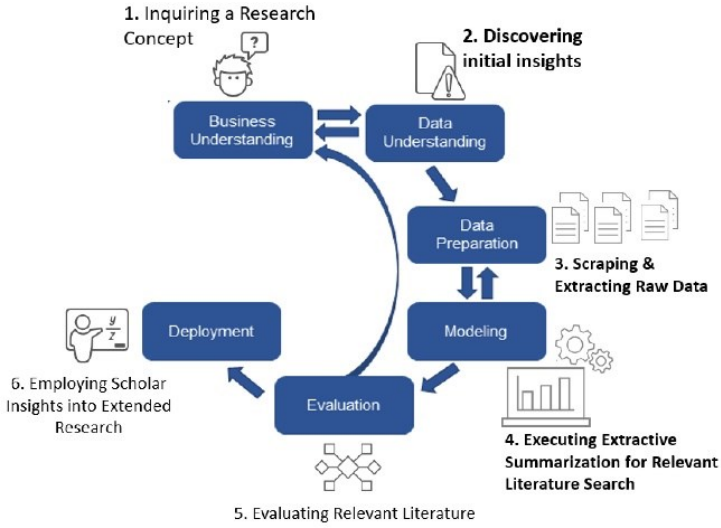
Data Understanding: Understanding what can be expected and achieved from the data. This process goes through checking the quality of the data, in several terms, such as data completeness, values distributions and data governance compliance. This is a crucial part of the project because it defines how viable and trustworthy can be the final results.

Data Preparation: Involves the ETL process that turns the pieces of data into something useful using algorithms and process.

Modeling: This is the core of any machine learning project. This step is responsible for the results that should satisfy the project goals, some algorithms such as k-means, hierarchical clustering, time series, linear regression, k-nearest neighbors, and several others, are the core code lines of this step in the methodology.

Evaluation: Verify that the results are valid and correct. We can expect two results - Wrong results: In such a case, we have to go back to the first step, in order to understand why the results are mistaken. - Accurate results: If we reach a satisfying test accuracy results, we can move on to the next step.

Deployment: Presenting the results in a useful and understandable manner, and by achieving this, the project should achieve its goals. It is the only step not belonging to a cycle.



10:Project cycle

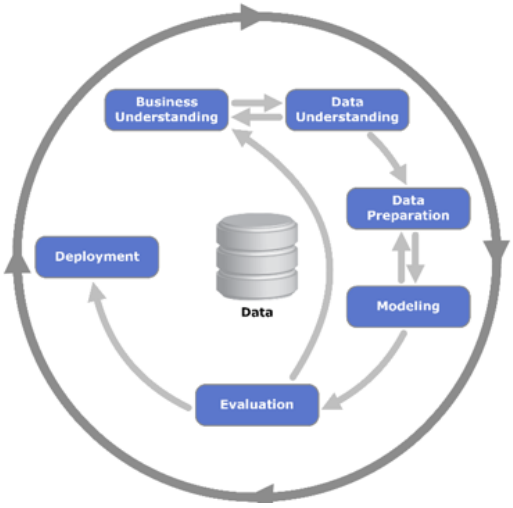
CRISP-DM VS Scrum

One key difference between the two approaches is what is being delivered. The deliverable for a Crisp-DM (Analytics) project is a piece of insight embedded into a decision space. The deliverable for an Agile (Software) project is a working piece of software or a product.

Another difference is that for Software projects developed using Agile methods focus on the needs of the end user. Gathering requirements using user-stories is a very powerful tool and works really well for software projects. Focusing on the end user may not always work well in an analytics context. Usually, focusing on the business objective works much better, because in most cases, the end user will never know of the existence of the analytics output.

Actually, although Scrum is the most used methodology but it's used for developing software projects that's why it's not a perfect match for analytics project.

Moreover, Crisp-Dm is better methodology for modern analytics project.



VS

3. Conclusion

After setting the objectives and understanding the methodology, in the next chapter we will present the data and explain each field in the database.

Chapter 4: Data source identification and description

1. Introduction

The ability to conduct effective research and analysis depends on data collection and a good understanding of a database.

2. Data tables identification

2:bill

Database	Tables	Columns	Description
Projeqtor	This table contains informations about bills	id	Bill id
		Name	Bill's name
		idProject	Id of project
		idCilent	Id of client
		Done	Payment status 1 done 0 not done
		paymentDate	Real date of payment
		PaymentDueDate	Date of payment

3:Project

Database	Tables	Columns	Description
Projeqtor	This table contains informations about projects	id	Id of project
		Name	Name of project
		idClient	Id of client
		projectCode	Code of project
		Done	Project done or under progress 1 done 0 not yet
		DoneDate	End project date
		paymentDelay	Number of days after payment date
		idStatus	Id of project status
		ClientCode	Code of client
		creationDate	Date of creation
		idResource	Id of employees

4:Client

Database	Tables	Columns	Description
Projeqtor	This table contains informations about clients of Talys	Id	Id of client
		name	Name of client
		clientCode	Code of client
		PaymentDelay	Delay of payment

5:Resource

Database	Table	Columns	Description
Projeqtor	This table contains informations about employees of Talys	Id	Id of employee
		Name	Name of employee
		idProfile	Id of profile
		isResource	0 free 1 not free
		idTeam	Id of team
		idClient	Id of client

6:Profile

Database	Table	Columns	Description
Projeqtor	This table contains description of profile	Id	Id of profile
		Name	Name of profile
		ProfileCode	Code of profile

7:Role

Database	Table	Columns	Description
Projeqtor	This table contains description of role	Id	Id of role
		Name	Name of profile

8:Status

Database	Table	Columns	Description
Projeqtor	This table contains description of status of project	Id	Id of status
		Name	Name of status
		SetDoneStatus	1 done 0 in progress

9:Team

Database	Table	Columns	Description
Projeqtor	This table contains description of team	Id	Id of team
		Name	Name of team

10:Activity

Database	Table	Columns	Description
Projeqtor	This table contains description of activity	Id	Id of activity
		IdProjet	Id of project
		Name	Name of activity
		CreationDate	Date of creation
		IdStatus	Id of status
		IdRessource	Id of resource
		DoneDate	End date of activity

11 :Assignment

Database	Table	Columns	Description
Projeqtor	This table contains description of assignment	Id	Id of assignment
		IdRessource	Id of resource
		IdProject	Id of project
		realStartDate	Real start date of assignment
		realEndDate	Real end date of assignment
		plannedStartDate	Planned start date of assignment
		plannedEndDate	Planned end date of assignment
		idRole	Id of role

3. Conclusion

After identifying the source of the data and clarifying it, now we are ready to set up the architecture of our solution that we will adopt throughout the project.

Chapter 5: Modeling and Loading of data warehouse

1. Introduction

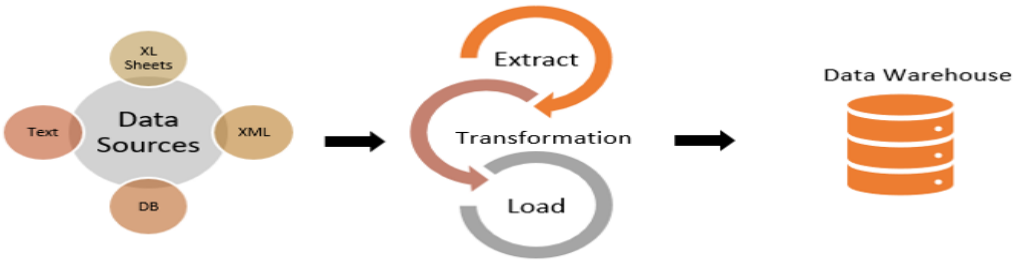
The main objective of this chapter is the Modeling and loading of a Data Warehouse

2. the Inmon approach

Bill Inmon, the father of data warehousing, came up with the concept to develop a data warehouse that starts with designing the corporate data warehouse data model, which identifies the main subject areas and entities the enterprise works with, such as customer, product, vendor, and so on.

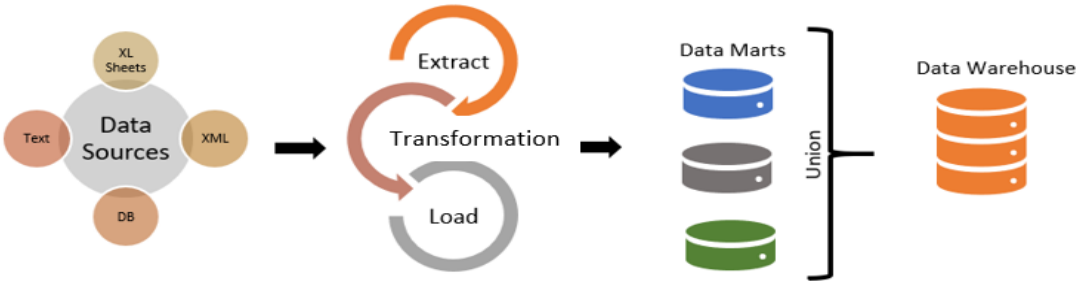
Bill Inmon's definition of a data warehouse is that it is a "subject-oriented, nonvolatile, integrated, time-variant collection of data in support of management's decisions".

3. Inmon vs Kimball



11:the Inmon approach

The Inmon approach is referred to as the top-down or data-driven approach, whereby we start from the data warehouse and break it down into data marts, specialized as needed to meet the needs of different departments within the organization, such as finance, accounting, HR.



12:the Kimball approach

The Kimball approach is referred to as bottom-up or user-driven, because we start from the user-specific data marts, and these form the very building blocks of our conceptual data warehouse. It's important to know from the outset which model best suits your needs so that it can be built into the data warehouse schema.

12:table Inmon vs Kimball

	Kimball	Inmon
Introduced by	Introduced by Ralph Kimball.	Introduced by Bill Inmon.
Approach	It has Bottom-Up Approach for implementation.	It has Top-Down Approach for implementation.
Data Integration	It focuses Individual business areas.	It focuses Enterprise-wide areas.
Building Time	It is efficient and takes less time.	It is complex and consumes a lot of time.
Cost	It has iterative steps and is cost effective.	Initial cost is huge and development cost is low.
Skills Required	It does not need such skills but a generic team will do job.	It needs specialized skills to make work.
Maintenance	Maintenance is difficult.	Here maintenance is easy.
Data Model	It prefers data to be in De-normalized model.	It prefers data to be in normalized model.
Data Store Systems	In this, source systems are highly stable.	In this, source systems have high rate of change.

4. Model star schema

The star data model owes its name to its shape. This design model favours the user approach, the business orientation.
The reference table contains the facts. The facts or measures are the figures (such as results by sector). The satellite tables correspond to the dimensions. These are the user analysis axes.

Star schema is the fundamental schema among the data mart schema and it is simplest. This schema is widely used to develop or build a data warehouse and dimensional data marts. It includes one or more fact tables indexing any number of dimensional tables.

Why the model star schema in data warehouse:

- **Simpler Queries:**
  Join logic of star schema is quite cinch in compare to other join logic which are needed to fetch data from a transactional schema that is highly normalized.
- **Simplified Business Reporting Logic:**
  In compared to a transactional schema that is highly normalized, the star schema makes simpler common business reporting logic, such as as-of reporting and period-over-period.
- **Loading Cubes:**
  Star schema is widely used by all OLAP systems to design OLAP cubes efficiently. In fact, major OLAP systems deliver a ROLAP mode of operation which can use a star schema as a source without designing a cube structure.

5. Model explication

In our case we have one fact table that we judge necessary, which is joined to specific dimensions.

DimDate: a dimension that handles the time spans related to the fact table. It has many columns that are able to precisely describe the notion of time, with for example the name of the day, the name of the quarter etc…

DimSector: a dimension that contains different domains and category exploied in the project . its columns are Category, Name, Status, and done.

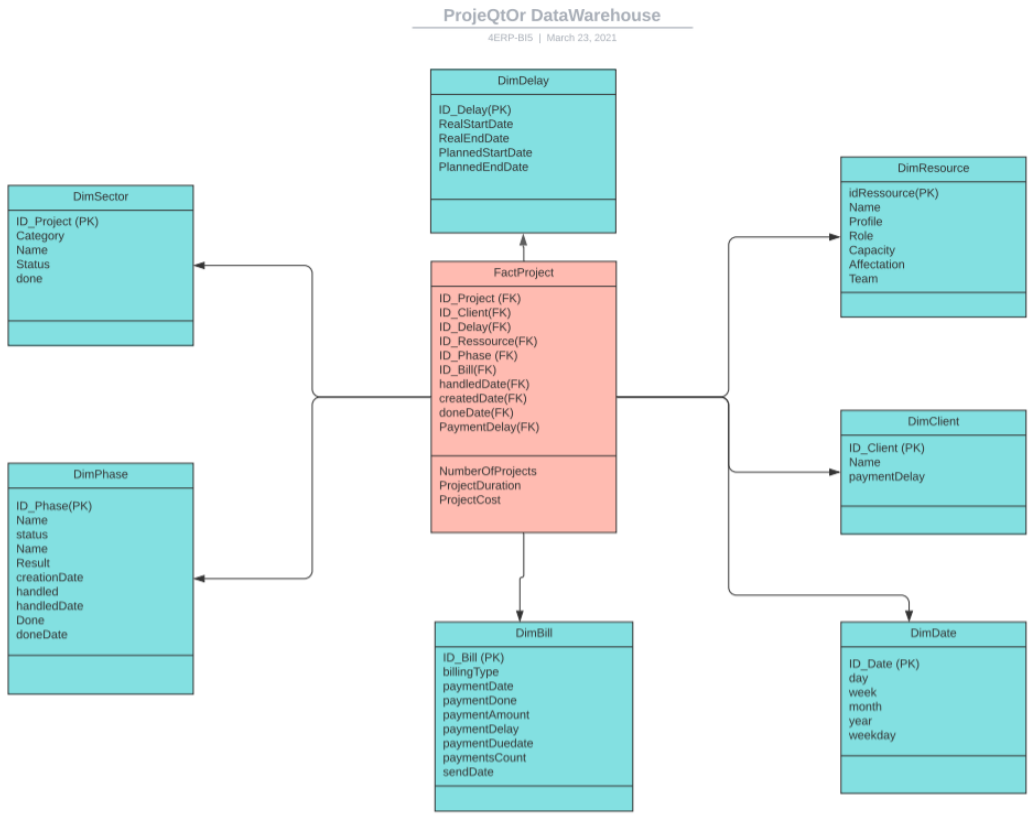
DimPhase: a dimension that indicates a fixed length events of one month or less to create consistency. it has many columns that describe this period of time such as the name, status,result, creationDate, handlded, handledDate, Done, DoneDate.

DimDelay: a dimension that specialize in the time by which something is late or postponed, its colmuns are RealStartDate, RealEndDate, PlannedStartDate, PlannedEndDate which are related to timing.

DimResource: a dimension that contains the different human resources who works for the company and it has many colmuns that define these resources such as the name, profile; role, capacity, team etc ...

DimClient: a dimension that describes the company's clients with just the columns Name and paymentdelay.

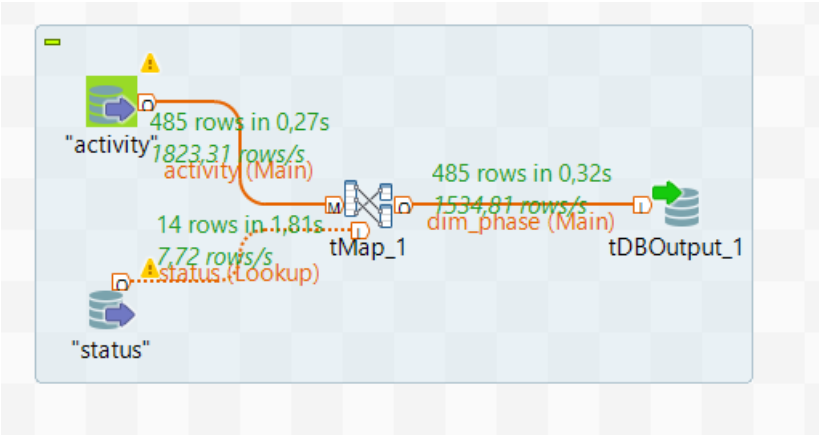
DimBill: a dimension that deals with the billing system and information about the payment date and type…. the columns of this dimension are billingType, paymentDate, paymentDone, paymentDelay …



13:Modeling of data Warehouse

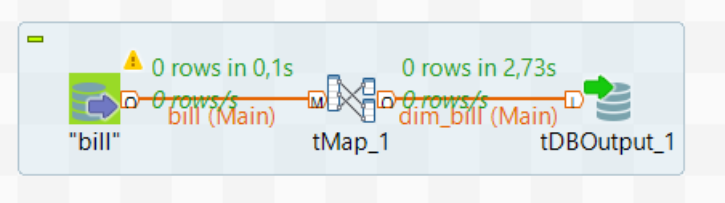
6. Data integration with Talend

Loading the Phase dimension from the source data table to the Datawarehouse



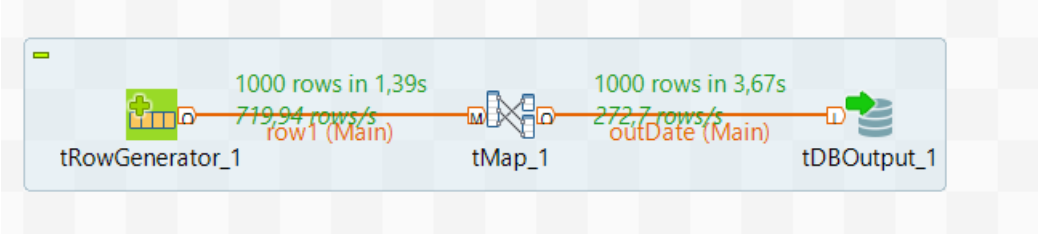
14:dim_phase

Loading the Bill dimension from the source data table to the Datawarehouse



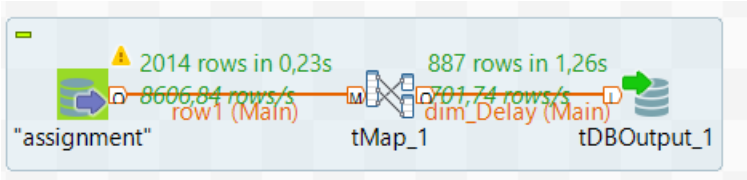
15:dim_bill

Loading the Date dimension generated through the tRowGenerator



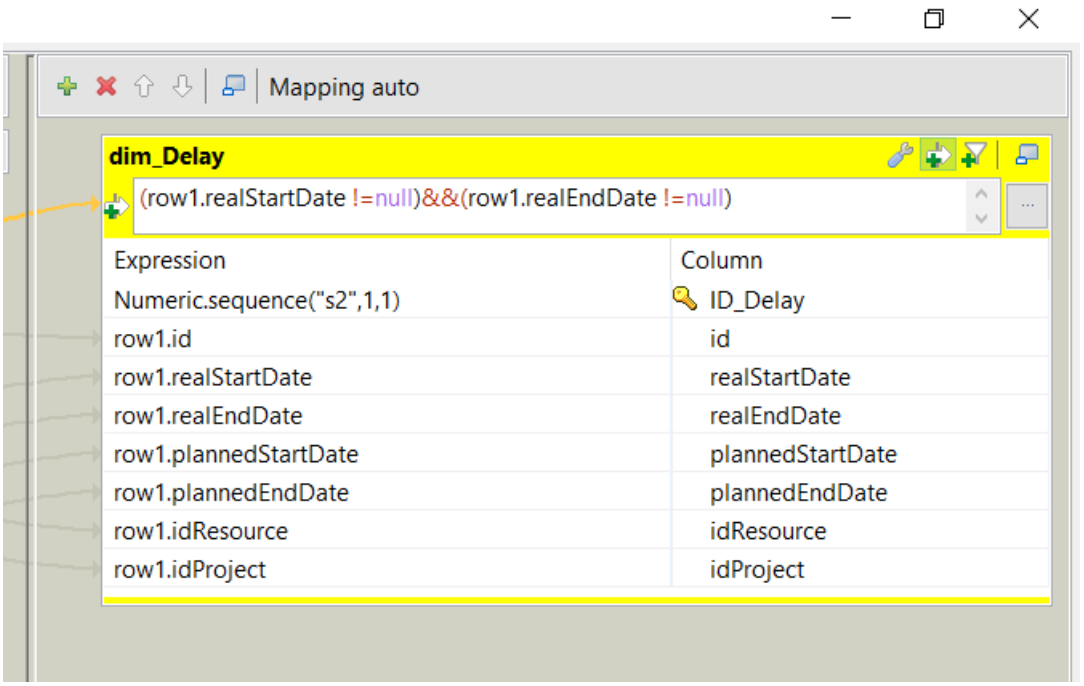
16:dim_date

Loading the Delay dimension from the source data table to the Datawarehouse



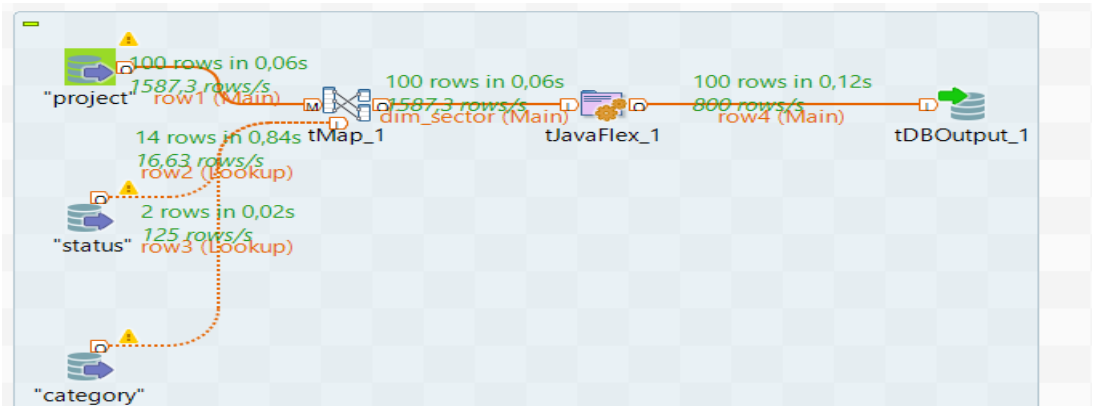
17:dim_delay

Filtered the delay rows , both the start date and the end date must contain data in order to calculate the project duration correctly



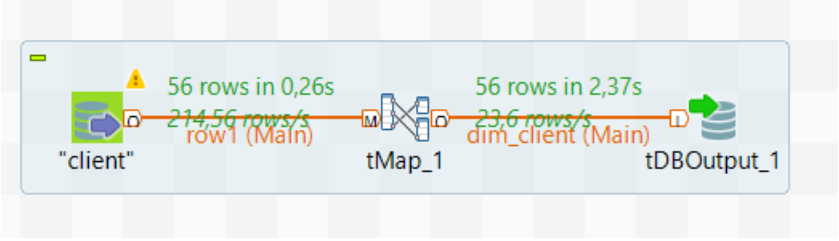
18:filter dim_delay:

Loading the Sector dimension from the source data table to the Datawarehouse



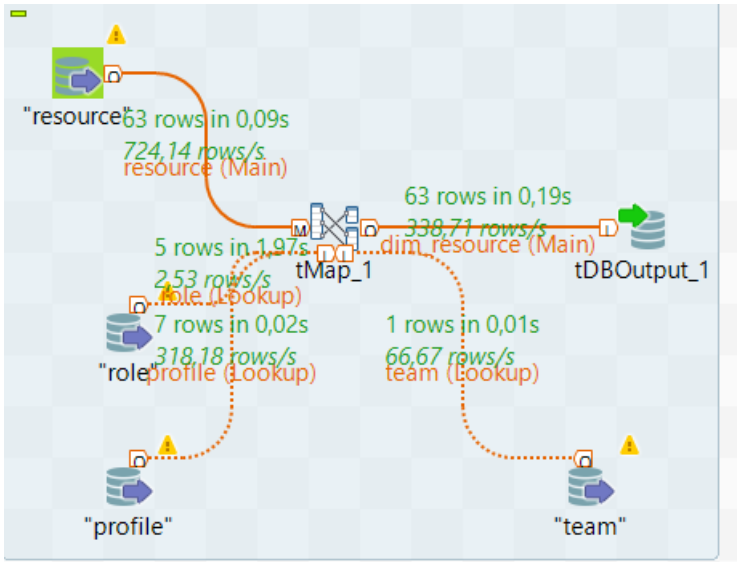
19:dim_sector

Loading the Client dimension from the source data table to the Datawarehouse



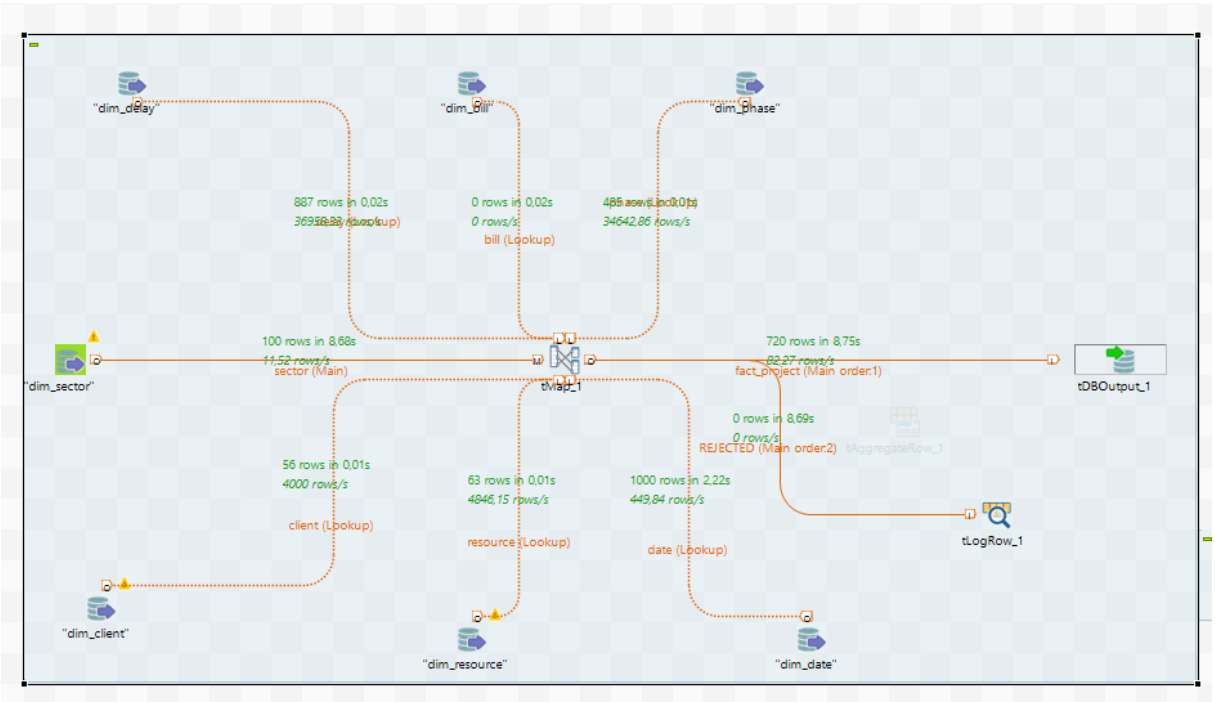
20:dim_client

Loading the Resource dimension from the source data table to the Datawarehouse



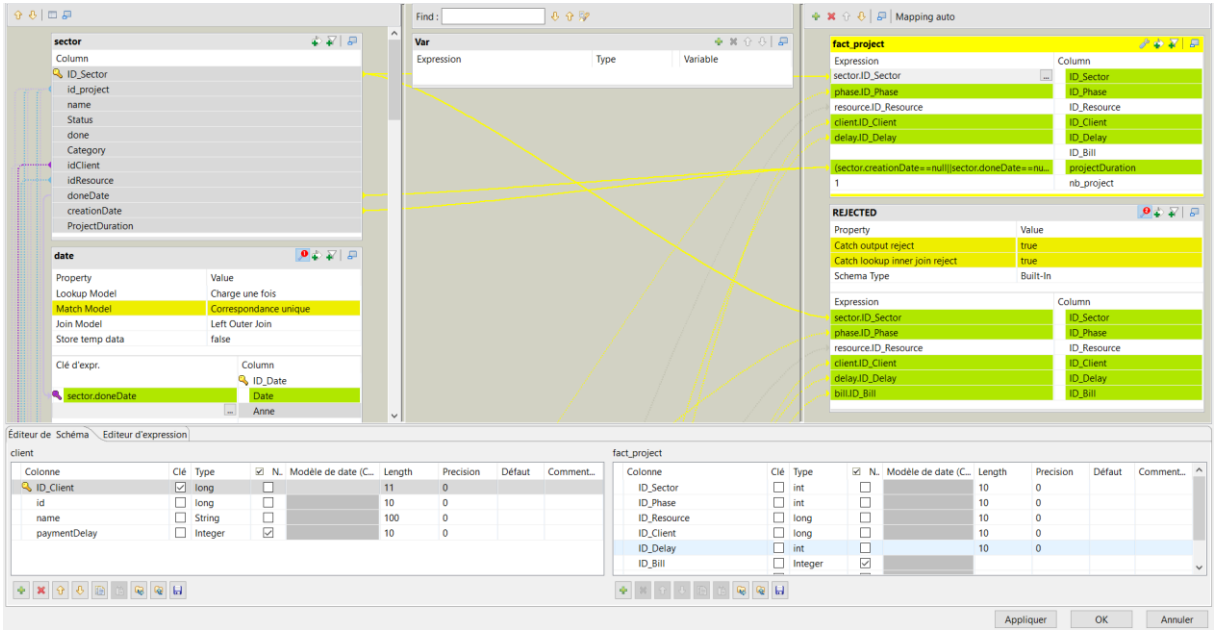
21:dim_resource

Loading the Fact table with dim Sector as the main data source and the others as lookup dimensions



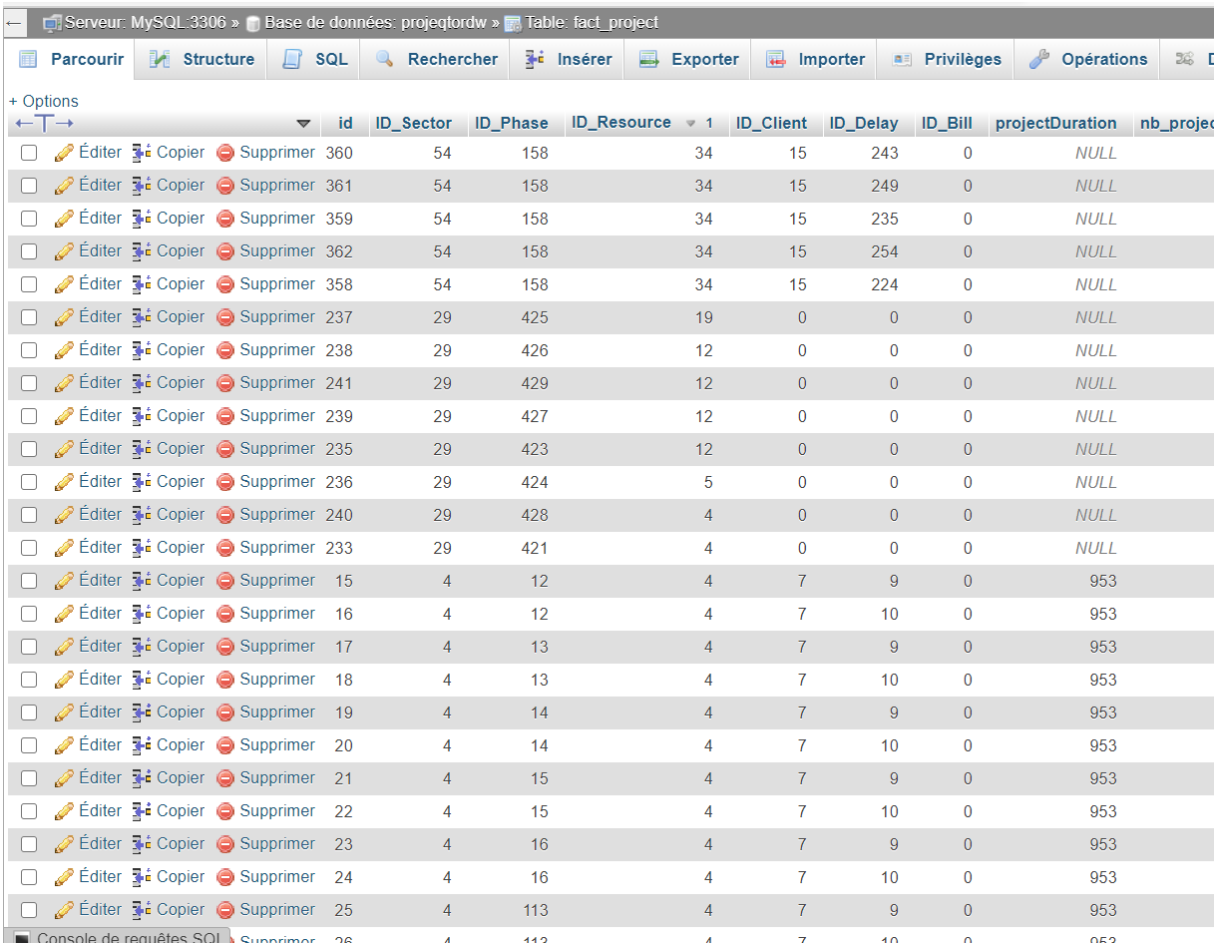
22:Fact_project

The tMap Component structure with the join between the Datawarehouse tables



23:tMap Fact_project

Résultat



24: table Fact_project