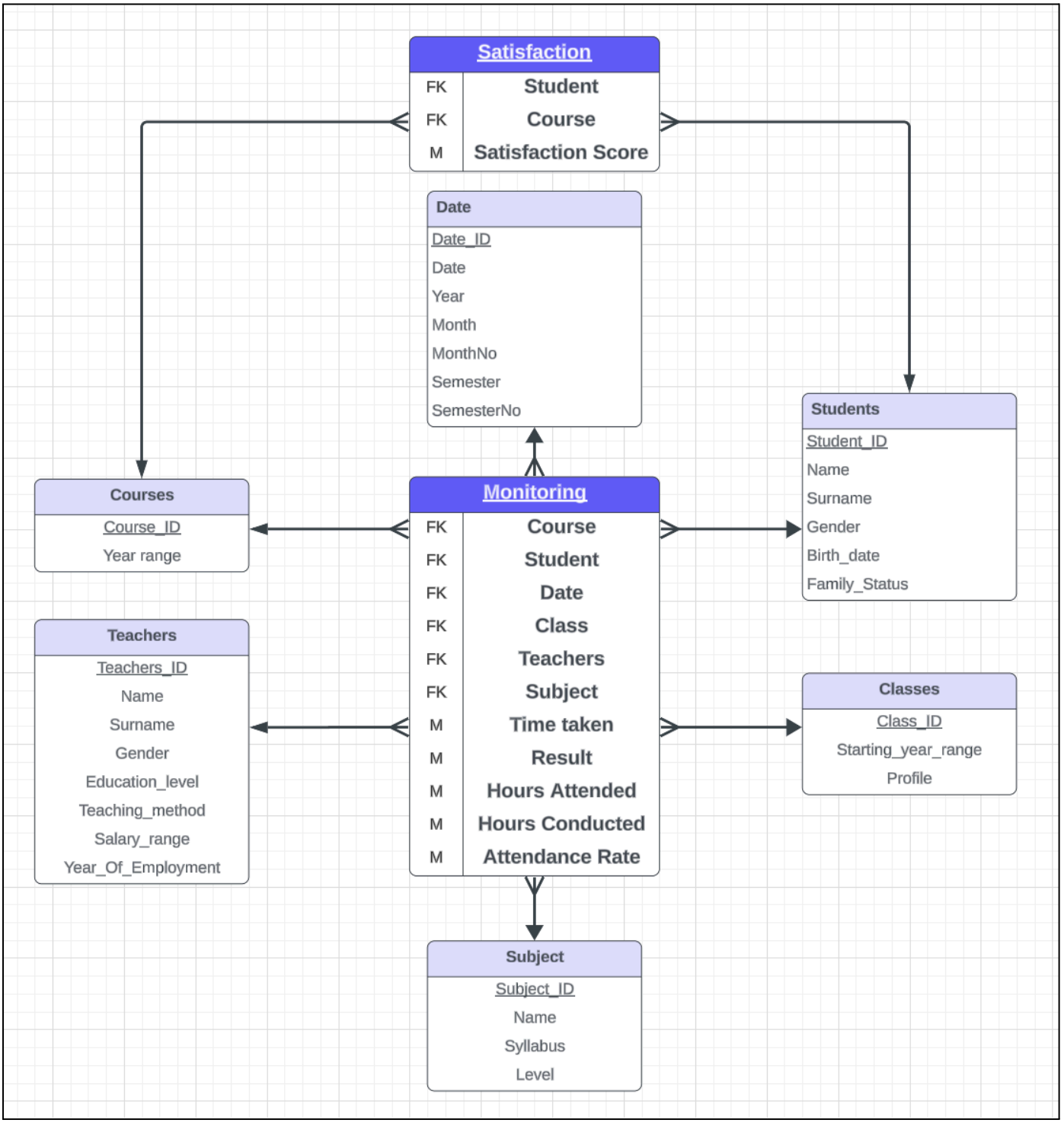
Business process

The Data warehouse is designed for monitoring students' processes. This process is described in the document “Requirements specification for monitoring students process “in “Gall Anonim" High school network

Relational Database schema



ENTITY SETS	ATTRIBUTE	DESCRIPTION	ATTRIBUTE TYPE
Monitoring (Fact Table) one tuple describes the summary of a student at the end of each month in GA High School	**Student ID**	**FK Student** Student who wrote an exam.	**Numeric**
	Course ID	**FK Course** id of course	**Numeric**
	Subject ID	**FK Subject** Subject that this course is on	**Numeric**
	Teacher ID	**FK Teacher** teacher that teaches this course	**Numeric**
	Class ID	**FK Class** class to which student attends	**Numeric**
	Date ID	**FK Date** Exam date	**Numeric**
	Result	**The overall score [in %]**	**Decimal(1,2)** Range : [0.00, 1.00]
	Time_taken	**The time taken to write an exam**	**Numeric** max value: 180
	Hours_Attended	**The amount of hours that a given student attended**	**Numeric**
	Hours_Conducted	**The amount of conducted hours**	**Numeric**
	Attendance_rate	**the attendance rate [in %]** attended = hours attended /conducted lessons	**Decimal(1,2)** Range : [0.00, 1.00]
Level of satisfaction (Fact Table) one tuple represents students' satisfaction level at the end of academic year to every course	**Student_ID**	**FK Student** Student who rated the course	**Numeric**
	Course ID	**FK Course** Satisfaction on specific course	**Numeric**
	Satisfaction_Level	**Level of student's satisfaction** Between 0 and 10	**Numeric**
Students (Dimension Table) one tuple describes one teenagers attending to the GA High School	**Student_ID**	**PK Student**	**Numeric**
	Name	**Name of the student**	**varchar(30)**
	Surname	**Surname of the student**	**varchar(40)**
	Gender	**Gender of the student** Allowed values: Male, Female	**varchar(8)**
	Family status	**family status of student** Allowed values: full family, separated family, one parent, no parents	**varchar(20)**

ENTITY SETS	ATTRIBUTE	DESCRIPTION	ATTRIBUTE TYPE
Date (Dimension Table) one tuple describes one month in calendar	**Date ID**	**PK Date**	**Date**
	Date	**Date**	**Date**
	Year	**Year in calendar**	**varchar(4)**
	Month	**Month in calendar** Allowed values: January, February, March, April, May, June, July, August, September, October, November and December.	**varchar(10)**
	MonthNo	**Number of the month**	**numeric** Range: [1,12]
	Semester	**Semester of teaching** **Winter:** September - January **Summer:** February - June	**varchar(10)**
	SemesterNo	**Number of the semester:** **1:** Winter **2:** Summer	**numeric** Allowed values: 1 or 2
Teachers (Dimension Table) one tuple describes a teacher working in GA High School	**Teacher_ID**	**PK Teacher**	**Numeric**
	Name	**Name of the teacher**	**varchar(30)**
	Surname	**Surname of the teacher**	**varchar(40)**
	Gender	**Gender of the teacher** Allowed values: Male, Female	**varchar(8)**
	Education Level	**The education of the teacher**	**varchar(40)**
	Teaching_Method	**method of teaching** - Mass (everyones does together) - Individualized (by themself) - Group (in groups)	**varchar(15)**
	salary_range	**range of salary of a teacher** Allowed values: high, average,low	**varchar(7)**
	Year of employment	**year of employment in GA school**	**varchar(4)**
Courses (Dimension Table) one tuple represents one course from the particular subject teached by specific teacher for given class	**Course_ID**	**PK Course** Course of exact subject that is teached by one teacher and is for one class	**Numeric**
	Year_range	**The year range, when this course were conducted** Allowed values: between 2014 and 2015, between 2015 and 2016 , between 2016 and 2017 , between 2017 and 2018, between 2018 and 2019 , between 2020 and 2021, between 2021 and 2022, between 2022 and 2023, between 2023 and 2024	**varchar(24)**

ENTITY SETS	ATTRIBUTE	DESCRIPTION	ATTRIBUTE TYPE
Subjects (Dimension Table) one tuple represents subject	**Subject_ID**	**PK Subject**	**Numeric**
	Name	**name of the subject**	**varchar(40)**
	Syllabus	**The syllabus of the subject**	**varchar(50)**
	Level	**level of difficulty** Allowed values: Expanded, Basic	**varchar(10)**
Classes (Dimension Table) one tuple represents one class	**Class ID**	**Class identifier**	**Numeric**
	Profile	**Profile of class**	**varchar(40)**
	Starting year range	**The starting year range, when this course is conducted** Allowed values: between 2014 and 2015, between 2015 and 2016 , between 2016 and 2017 , between 2017 and 2018, between 2018 and 2019 , between 2020 and 2021, between 2021 and 2022, between 2022 and 2023, between 2023 and 2024	**varchar(24)**

Dimensional model

Monitoring fact

monitoring students that include checking attendance in the specific month for specific students and examining students. conducted exams are to specific students that are signed in classes, and from a specific course that is held by a specific teacher and about a certain subject.

Fact table: Monitoring

Granularity:

- a specific month
- a specific course
- a specified subject that its on
- a specified teacher that holds it
- a specified student with the specified gender and family status
- a specified class
- a specified time taken to write an exam
- a specified results of parts
- a specified hours of attendance

Measures and aggregation functions:

COUNT(1) - count all exams for student
SUM (Result) - sum of points gained from the exams
SUM (Result)/COUNT(1) - average result of one student in one course
DISTINCT COUNT (Student) - amount students that has wrote exam
SUM (attendance_rate) - sum of attendance rates
SUM (attendance_rate)/COUNT(1) - average attendance rates

Satisfaction fact:

Level of satisfaction of specified students. Collected at the end of the academic year for each enrolled course

Fact table: Satisfaction

Granularity:

- a specific course that satisfaction was submitted to, and specific year of course
- a specified student, with the specified class, with the specified gender and family status
- specific rate of satisfaction

Measures and aggregation functions:

SUM(**Satisfaction_Level**)-sum of satisfaction of student
COUNT(1)-number of submitted satisfactions of student
SUM(**Satisfaction_Level**)/COUNT(1)-average of submitted satisfactions of student

Dimension definitions:

Dimensions for Fact 1 Monitoring students fact:

DIMENSION/DIMENSION ATTRIBUTE	TABLE/COLUMN	COLUMN TYPE
Student_ID	**Student**	**Dimension**
Name	**student.Name**	**Dimension attribute**
Surname	**student.Surname**	**Dimension attribute**
Gender	**student.Gender**	**Dimension attribute**
family_status	**student.family_status**	**Dimension attribute**
Course_ID	**Course**	**Dimension**
Year range	**Course.year_range**	**Dimension attribute**
Date ID	**Date**	**Dimension**
Year	**Date.Year**	**Dimension attribute**

Month	Date.Month	Dimension attribute
MONITORING DATE HIERARCHY	•Year ••Semester •••Month	Hierarchical dimension
Semester	Date.Semester	Dimension attribute
Teacher_ID	teacher	Dimension
Name	teacher.name	Dimension attribute
Surname	teacher.surname	Dimension attribute
Gender	teacher.gender	Dimension attribute
Education Level	teacher.education_level	Dimension attribute
Salary range	teacher.salary_range	Dimension attribute
Teaching_method	teacher.Teaching_method	Dimension attribute
year_of_employment	teacher.year_of_employment	Dimension attribute
Subject_ID	subject	Dimension
Name	subject.Name	Dimension attribute
Syllabus	subject.Syllabus	Dimension attribute
SUBJECT HIERARCHY	• Name ••Level •••Syllabus	hierarchical dimension
Level	subject.Level	Dimension attribute
Class_ID	class	Dimension
Starting year	class.Starting_year	Dimension attribute
Profile	class.Profile	Dimension attribute

Dimensions for Fact 2: submitting satisfaction fact

DIMENSION/DIMENSION ATTRIBUTE	TABLE/COLUMN	COLUMN TYPE
Student_ID	**Student**	**Dimension**
Name	**student.Name**	**Dimension attribute**
Surname	**student.Surname**	**Dimension attribute**
Gender	**student.Gender**	**Dimension attribute**
Course_ID	**Course**	**Dimension**
Year range	**Course.year_range**	**Dimension attribute**

Checking the feasibility of queries based on the multidimensional model

1. Does a year of studying have an impact on average level satisfaction?

Measure: Average Satisfaction Level
Dimension: Course (Year_range)

```
SELECT
  [Measures].[Satisfaction Level] ON COLUMNS,
  [Dim Courses].[Year Range].Members ON ROWS
FROM DW
```

2. Is there any correlation between students' satisfaction of the course and the level of education of a teacher that runs it?

Measure: Satisfaction Level
Dimensions: Teacher(Education Level)
Dimension: Course (Course_ID)

3. Compare the results in relation to the family status of students:

Measure: Average Exam Results
Dimension: Students(Family Status)

```
SELECT
    [Measures].[Average Exam Results] ON COLUMNS,
    [Dim Students].[Family Status] ON ROWS
```

4. What are the results of the overall best-students in the current and the previous month?

Measure: Average Exam Results
Dimension: Student(Student_ID, Name, Surname)
Dimension: Date(Year, Month)

5. Compare the exam results from each subject from this month to the previous months.

Measure: Average Exam Results
Dimension: Subject(Subject_ID, Subject)
Dimension: Date(Year, Month)

6. Analyze the correlation between attendance and exam results this month to the previous

Measures: Average Attendance Rate, Average Exam Results
Dimension: Date (Year, Month)

7. Does final exam grade have an impact on satisfaction?

Measures: Average Exam Result, Average Satisfaction Level

8. Which method gives the best result for given subjects?

Measure: Average Exam Results
Dimension: Teacher(Teaching Method)
Dimension: Subject (Subject_ID)

9. Identify the profiles that have the overall best and worst exam results.

Measure: Average Exam Results
Dimension: Class(Profile)

10. Does the length of employment of a teacher have an impact on the exam result over time?

Measure: Average Exam Results
Dimension: Teacher(Year_Of_Employment)
Dimension: Date(Year, Month)

C1-2014
C2-2014
C3-2014

Checking if there are Date in the Date sources needed to fill the Date warehouse

TABLE NAME	COLUMN	SOURCE
Monitoring (Fact Table) one tuple describes the summary of a student at the end of each month in GA High School	**date_id**	date id, foreign key from dimension table. based on date of attendance and exams stored in toGAther source
	course_id	course id, foreign key from dimension table. based on the course_id stored in toGAther source
	class_id	class id, foreign key from dimension table. based on the class_id stored in toGAther source
	teacher_id	teacher id, foreign key from dimension table. based on the teacher_id stored in toGAther source and data in excel sheet 2
	student_id	student id, foreign key from dimension table. based on the student_id stored in toGAther source and data in excel sheet 1
	subject_id	subject id, foreign key from dimension table. based on the subjectt_id stored in toGAther source
	time_taken	The time taken to write an exam. taken from time_taken from exams table from toGAther source
	attendance_rate	the attendance rate for each month. taken from lessons_attended column from attendance table and divided by hours_lessons_conducted column from lessons table*100%. both data are from toGAther source
	result	The overall score [in %] taken from result column from exams table
Level of satisfaction (Fact Table)	**Student_ID**	student id, foreign key from dimension table. based on the student_id stored in toGAther source and data in excel sheet 1

one tuple represents students' satisfaction level at the end of academic year to every course	**Course ID**	course id, foreign key from dimension table. based on the course_id stored in toGAther source
	Satisfaction_Level	Level of student's satisfaction. this is based on column Satisfaction_Level in table level of satisfaction stored in toGAther sourc
Date (Dimension Table) one tuple describes one month in calendar	**Date ID**	PK Date
	Year and month data in this table are generated tuple by tuple based on any calendar	
	Semester	**Semester of teaching, data in this table is based on months, months from** September - January represent semester 1 and February - June represent semester 2

ENTITY SETS	ATTRIBUTE	DESCRIPTION
Students (Dimension Table) one tuple describes one teenagers attending to the GA High School	**Student_ID**	Student Id. the key gained from Student id column from Student table toGAther source
	Name	Name of the student, this data is from Name column from students table from toGAther source
	Surname	Surname of the student, this data is from Surname column from students table from toGAther source
	Gender	Gender of the student, this data is from Gender column from students table from toGAther source
	Family status	family status of student, this data is from family status column from family status- Column I from excel sheet 1. this is a slowly changing dimension, change usually occurs not more often than once in a 3 years
Teachers (Dimension Table) one tuple describes a teacher working in GA High School	**Teacher_ID**	Teacher Id. the key gained from teacher column from teacher table from toGAther source
	Name	Name of the teacher, this data is from Name column from teachers table from toGAther source
	Surname	Surname of the teacher, this data is from Surname column from teachers table from toGAther source
	Gender	Gender of the teacher, this data is from Gender column from teachers table from toGAther source F-female, m-male
	Education Level	The education of the teacher, his data is from education_level column from teachers table from toGAther source. this is a slowly changing dimension, change usually occurs not more often than once in a 3 years
	Teaching_Method	method of teaching , this data is based on column teaching_menthod from teachers table in toGAther source: M- Mass, I-Individualized, G-Group

	salary_range	range of salary of a teacher Allowed values: high, average,low, this column is based on salary stored in column A in excel sheet 2. this is a slowly changing dimension, change usually occurs not more often than once in a year
	Year of employment	year of employment in GA school, stored in in column E in excel sheet 2
Courses (Dimension Table) one tuple represents one course from the particular subject teached by specific teacher for given class	**Course_ID**	Course ID - the key gained from course column from Course table from toGAther source
	Year range	The year range, when this course was conducted. This data is based on column Year_Range from Courses table in toGAther source

ENTITY SETS	ATTRIBUTE	DESCRIPTION
Subjects (Dimension Table) one tuple represents subject	**Subject_ID**	Subject ID - the key gained from Subject column from Subject table from toGAther source
	Name	name of the subject. This data is based on column Name from Subject table in toGAther source
	Syllabus	The syllabus of the subject. This data is based on column Syllabus from Subject table in toGAther source
	Level	level of difficulty(Expanded, Basic). This data is based on column Level from Subject table in toGAther source
Classes (Dimension Table) one tuple represents one class	**Class ID**	**Class identifier -** the key gained from class column from Classes table from toGAther source
	Profile	**Profile of class.** This data is based on column Profile from Classes table in toGAther source
	Starting_year_range	**The starting year range, when this course is conducted.** This data is based on column Starting_Year_Range from Classes table in toGAther source