Primary school “jedynka” - Data Warehouse design

Business process

The data warehouse is designed for Assess student performance process. This process is described in RequirementsProcessSpecification.

Relational Database Schema

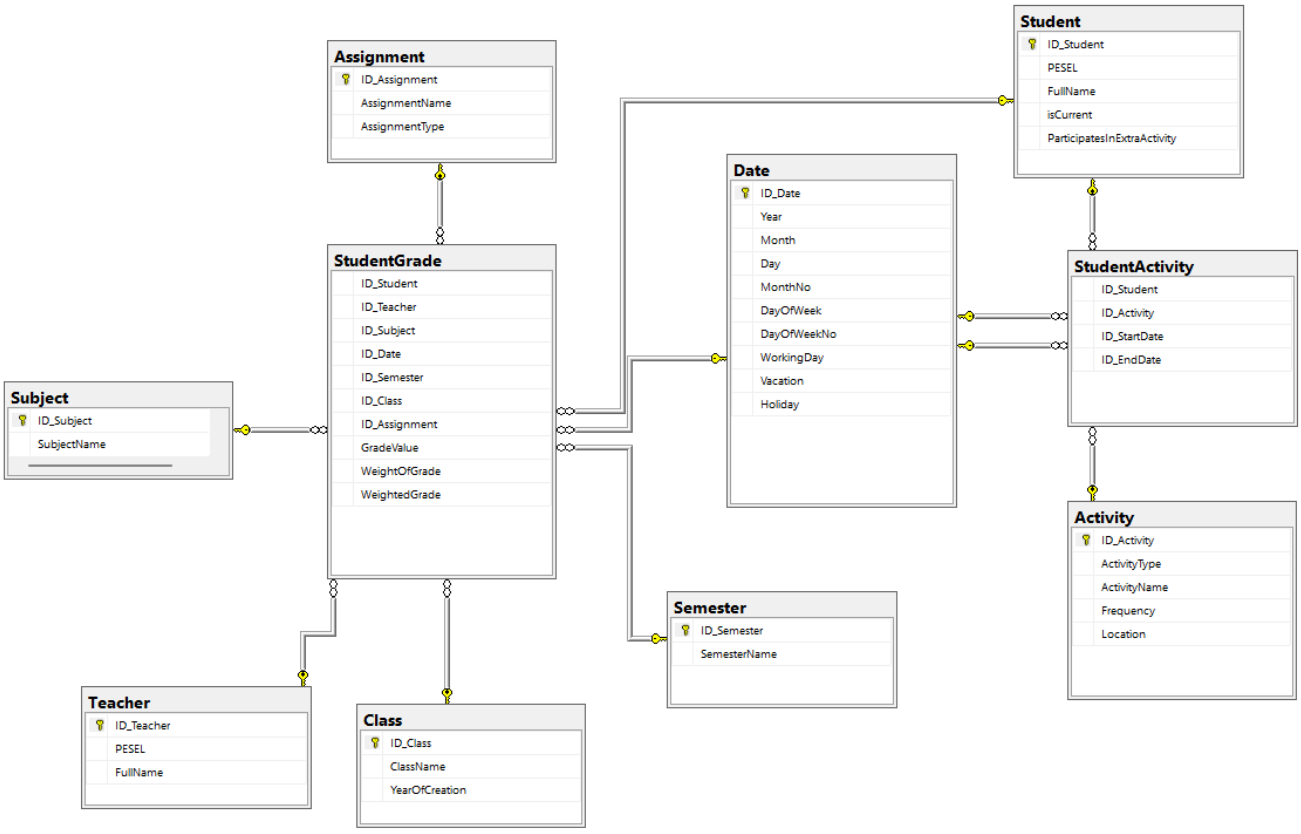

TABLE NAME	ATTRIBUTE	ATTRIBUTE TYPE	DESCRIPTION

StudentGrade (Fact table)	One tuple describes one fact of a student's grade		
	ID_Student	Numeric	FK to Student Student who received the grade
	ID_Teacher	Numeric	FK to Teacher The teacher who gave the grade
	ID_Subject	Numeric	FK to Subject The subject in which student was graded
	ID_Date	Numeric	FK to Date The date on which the grade was recorded
	ID_Semester	Numeric	FK to Semester The semester this grade belongs to
	ID_Class	Numeric	FK to Class The class student belongs to at the time of receiving a grade
	ID_Assignment	Numeric	FK to Assignment The assignment for which the grade was received
	GradeValue	Decimal	The numeric grade (scale 1-6, with increments by 0.25)
	WeightOfGrade	Integer	The weight of a grade (scale 0-6)
	WeightedGrade	Decimal	WeightedGrade equals to GradeValue * WeightOfGrade
StudentActivity (Fact table)	One tuple describes one fact of a student participating in a specific extracurricular activity		
	ID_Student	Numeric	FK to Student The student who participates in activity
	ID_Activity	Numeric	FK to Activity The extracurricular activity
	ID_StartDate	Numeric	FK to Date The date when student started the activity
	ID_EndDate	Numeric	FK to Date The date on which student ended the activity
Student (Dimension table)	One tuple describes one student		

	ID_Student	Numeric	PK (Surrogate key)
	PESEL	11 digits	Business key
	FullName	Varchar(60)	Student's full name (first and last name)
	isCurrent	Boolean	1 if it is current record of student, 0 if historical (SCD2 implementation)
	ParticipatesInExtra Activity	Boolean	1 if a student currently participates in some extracurricular activities, 0 is he is not (SCD2 implementation)
Teacher (Dimension table)	One tuple describes one teacher who can assign grades		
	ID_Teacher	Numeric	PK (Surrogate key)
	PESEL	11 digits	Business key
	FullName	Varchar(60)	Teacher's full name (first and last name)
Subject (Dimension table)	One tuple describes one subject taught in the school		
	ID_Subject	Numeric	PK (Surrogate key)
	SubjectName	Varchar(18)	Represents name of the subject. Allowed values: math, science, history, english, art, literature, music, physical education, computer science, geography, biology
Class (Dimension table)	One tuple describes one class (e.g. "1B," "3C")		
	ID_Class	Numeric	PK (Surrogate key)
	ClassName	Varchar(2)	Represents name of the class. Allowed values: 1A, 1B, 1C, 1D, 1E, 2A, 2B, 2C, 2D, 2E, 3A, 3B, 3C, 3D, 3E, 4A, 4B, 4C, 4D, 4E, 5A, 5B, 5C, 5D, 5E, 6A, 6B, 6C, 6D, 6E, 7A, 7B, 7C, 7D, 7E, 8A, 8B, 8C, 8D, 8E,
	YearOfCreation	Integer	The year when class was created

Semester (Dimension table)	One tuple describes one academic semester		
	ID_Semester	Numeric	PK (Surrogate key)
	SemesterName	Varchar(10)	Name of the semester. Allowed values: Summer####, Winter#### (#### indicates a year – e.g. 2021, 2024, 2025, etc.)
Date (Dimension table)	One tuple describes one calendar day		
	ID_Date	Numeric	PK (Surrogate key)
	Year	4 digits	Year
	Month	Varchar(10)	Month. Allowed values: January, February, March, April, May, June, July, August, September, October, November, December
	MonthNo	Numeric	Month numeric value. Range: 1…12
	Day	Numeric	Day numeric value. Range: 1…31
	DayOfWeek	Varchar(10)	Day of week. Allowed values: Monday, Tuesday, Wednesday, Thursday, Friday, Saturday, Sunday
	DayOfWeekNo	Numeric	A number from range 1-7, indicating a number of a day in a week (1 is Monday, 7 is Sunday)
	WorkingDay	Varchar(15)	Tells info if it is a working day or not. Allowed values: Working day, day off
	Vacation	Varchar(20)	Vacation time characteristic. Allowed values: Winter holiday, summer holiday, none
	Holiday	Varchar(50)	Type of holiday. Allowed values: C New Year's Day, Epiphany, Easter Sunday, Easter Monday, Labour Day, Constitution Day, Pentecost, Corpus Christi, Assumption of Mary, All

			Saints' Day, Independence Day, Christmas Day, Second Day of Christmas.
Activity (Dimension table)	One tuple describes one extracurricular activity		
	ID_Activity	Numeric	PK (Surrogate key)
	ActivityType	Varchar(12)	Type of extracurricular activity. Allowed values: sports, arts, volunteering, music
	ActivityName	Varchar(18)	Name of the activity. Allowed values: basketball, football, volleyball, running, choir, painting classes, community cleanup, school band, chess club
	Frequency	Varchar(20)	Tells info about how often students participates in certain activity. Allowed values: One time activity, 1-2 times a week, 3-5 times a week. 6-7 times a week, each 2 weeks, each month
	Location	Varchar(20)	Place where the activity is conducted. Allowed values: Playground, Library, Auditorium, Classroom, Gymnasium, Computer Lab, Science Lab, School Garden, Art Room, Music Room, Cafeteria, Hallway, Courtyard, Rooftop, Assembly Hall, Sports Field, Conference Room
Assignment (Dimension table)	One tuple represents one assignment that can be graded		
	ID_Assignment	Numeric	PK (Surrogate key)
	AssignmentName	Varchar(12)	Name of the assignment. The possible values are described by the following formula: AssignmentType#Number

			- AssignmentType describes the type of the assignment (see below) - Number range 1 .. 100
	AssignmentType	Varchar(8)	Type of the assignment. Allowed values: exam, homework, project, quiz

Dimensional model

Fact dimensions

Fact 1 Receiving a Grade fact: Grade received by a student, received on a specific day, from specific subject and given by the specific teacher and for a specific assignment.

Fact table: StudentGrade

Granularity:

- A specified student
- A specified subject
- A specified teacher
- A specified class
- A specified date
- A specified assignment

Measures and aggregate functions:

- Number of grade facts – COUNT(1)
- Number of distinct students – DISTINCT COUNT(ID_Student)
- Average grade – SUM(WeightedGrade) / SUM(WeightOfGrade)

- Correlation coefficient between Average Grade and Extracurricular participation - CORR(ParticipatesInExtraActivity, SUM(WeightedGrade) / SUM(WeightOfGrade))

Fact 2 Student Activity fact: Participation of a student in an extracurricular activity, starting on a specified date and possibly ending on another specified date.

Fact table: StudentActivity

Granularity:

- A specified student
- A specified activity
- A specified starting date
- A specified ending date

Measures and aggregate functions:

- Number of activity facts – COUNT(1)

Dimensions definitions

Dimensions for Fact 1 Receiving a grade:

Dimension/ Dim. Attribute	Table/Column	Type
STUDENT	Student	Dimension
STUDENT PESEL	Student.PESEL	Dimension attribute
STUDENT FULL NAME	Student.FullName	Dimension attribute
IS CURRENT	Student.isCurrent	Dimension attribute
PARTICIPATES IN EXTRA ACTIVITY	Student.ParticipatesInExtraActivity	Dimension attribute
TEACHER	Teacher	Dimension
TEACHER PESEL	Teacher.PESEL	Dimension attribute
TEACHER FULL NAME	Teacher.FullName	Dimension attribute
SUBJECT	Subject	Dimension

SUBJECT NAME	Subject.SubjectName	Dimension attribute
CLASS	Class	Dimension
CLASS NAME	Class.ClassName	Dimension attribute
YEAR OF CREATION	Class.YearOfCreation	Dimension attribute
SEMESTER	Semester	Dimension
SEMESTER NAME	Semester.SemesterName	Dimension attribute
DATE	Date	Dimension
YEAR	Date.Year	Dimension attribute
MONTH	Date.Month	Dimension attribute
DAY	Date.Day	Dimension attribute
MONTH NUMBER	Date.MonthNo	Dimension attribute
DAY OF WEEK NUMBER	Date.DayOfWeekNo	Dimension attribute
DAY OF WEEK	Date.DayOfWeek	Dimension attribute
WORKING DAY	Date.WorkingDay	Dimension attribute
VACATION	Date.Vacation	Dimension attribute
HOLIDAY	Date.Holiday	Dimension attribute
ASSIGNMENT	Assignment	Dimension
ASSIGNMENT NAME	Assignment.AssignmentName	Dimension attribute
ASSIGNMENT TYPE	Assignment.AssignmentType	Dimension attribute
CLASS HIERARCHY	• Class.YearOfCreation •• Class.ClassName	Hierarchical dimension
DATE HIERARCHY	• Date.Year •• Date.Month •••Date.Day	Hierarchical dimension

Dimensions for fact 2 Student Activity

DIM/DIM ATTRIBUTE	TABLE/COLUMN	TYPE
STUDENT	Student	Dimension

STUDENT PESEL	Student.PESEL	Dimension Attribute
STUDENT FULL NAME	Student.FullName	Dimension Attribute
STUDENT IS CURRENT	Student.isCurrent	Dimension Attribute
PARTICIPATES IN EXTRA ACTIVITY	Student.ParticipatesInExtraActivity	Dimension Attribute
ACTIVITY	Activity	Dimension
ACTIVITY NAME	Activity.ActivityName	Dimension Attribute
ACTIVITY TYPE	Activity.ActivityType	Dimension Attribute
ACTIVITY FREQUENCY	Activity.Frequency	Dimension Attribute
ACTIVITY LOCATION	Activity.Location	Dimension Attribute
DATE	Date	Dimension
YEAR	Date.Year	Dimension attribute
MONTH	Date.Month	Dimension attribute
DAY	Date.Day	Dimension attribute
MONTH NUMBER	Date.MonthNo	Dimension attribute
DAY OF WEEK	Date.DayOfWeek	Dimension attribute
DAY OF WEEK NUMBER	Date.DayOfWeekNo	Dimension attribute
WORKING DAY	Date.WorkingDay	Dimension attribute
VACATION	Date.Vacation	Dimension attribute
HOLIDAY	Date.Holiday	Dimension attribute
DATE HIERARCHY	• Date.Year •• Date.Month •••Date.Day	Hierarchical dimension

Checking the feasibility of queries based on the multidimensional model

1. What are the subjects for which the average grade improved the most in the last three months?

   Measure: Average grade
   Dimension: Subject (dimension attributes: Subject Name)
   Dimension: Date (dimension attributes: Day, Month)

2. What is the change in the average grades of all students from biology in comparison to the previous month?

   Measure: Average grade
   Dimension: Subject (dimension attributes: Subject Name)
   Dimension: Date (dimension attributes: Day, Month)

3. Did the average grade increase or decrease this month compared to the previous month?

   Measure: Average grade
   Dimension: Date (dimension attributes: Day, Month)

4. Did students who participated in extracurricular activities in the past month experience a positive change in the average grade from all subjects?

   Measure: Average grade
   Dimension: Student (dimension attributes: Participates in extra activity, Student PESEL)
   Dimension: Date (dimension attribute: Month)

5. Did the students with the highest number of assignments given in the past month experience a negative change in the average grade from all subjects past month?

   Measure: Average grade
   Measure: Number of grade facts
   Dimension: Student (dimension attributes: Student PESEL)
   Dimension: Date (dimension attributes: Month)

6. Is there a significant difference (significant difference is around half a grade) between the average grades of students who participate in extracurricular

activities compared to those who do not?

Measure: Average grade
Dimension: Student (dimension attributes: Participates in extra activity)

7. Has there been a student that participated in additional activities, which had an average grade below 2.0 from any subject he participated in?

   Measure: Average grade (students below 2.0)
   Dimension: Student (dimension attributes: Participates in extra activity, Student PESEL)
   Dimension: Subject (dimension attributes: Subject Name)

8. Which extracurricular activities are associated with the highest average grades rankings among students?

   Measure: Average grade
   Dimension: Extracurricular activity (dimension attributes: Activity Name)

9. How has the average grade of students involved in extracurricular activities changed over the past five years?

   Measure: Average grade
   Dimension: Date (dimension attributes: Year)
   Dimension: Student (dimension attributes: Participates in extra activity)

10. Is there a statistically significant correlation between the number of extracurricular activities a student participates in and their average grades?

    Measure: Average grade
    Measure: Correlation coefficient between Average Grade and Extracurricular participation
    Dimension: Student (dimension attributes: Participation in extracurricular activities)

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

TABLE NAME	COLUMN	SOURCE
StudentGrade	One tuple describes one fact of a student's grade	
	ID_Student	Student ID. Foreign key from dimension table. Based on StudentID column stored in Student Grading System database, table Student
	ID_Teacher	Teacher ID. Foreign key from dimension table. Based on TeacherID column stored in Student Grading System database, table Teacher
	ID_Subject	Subject ID. Foreign key from dimension table. Based on SubjectID column stored in Student Grading System database, table Subject
	ID_Date	Date ID. Foreign key from dimension table. Based on Date column stored in Student Grading System database, table Grade
	ID_Semester	Semester ID. Foreign key from dimension table. Based on SemesterID column stored in Student Grading System database, table Semester
	ID_Class	Class ID. Foreign key from dimension table. Based on ClassID column stored in Student Grading System database, table Class
	ID_Assignment	Assignment ID. Foreign key from dimension table. Based on AssignmentID column stored in Student Grading System database, table Assignment
	GradeValue	Value of the grade received (on a scale from 1 to 6 with increments by 0,25) - taken from GradeValue from Grade table in Student Grading System database

	WeightOfGrade	The weight of a grade (scale 0-6) - taken from WeightOfGrade from Grade table in Student Grading System database
	WeightedGrade	WeightedGrade equals to GradeValue * WeightOfGrade – sourced from GradeValue and WeightOfGrade columns from Grade table in Student Grading System database
StudentActivity	One tuple describes one fact of a student participating in a specific extracurricular activity	
	ID_Student	Student ID. Foreign key from dimension table. Based on StudentID column stored in Student Grading System database, table Student
	ID_Activity	Activity ID. Foreign key from dimension table. Based on ActivityType and ActivityName stored in Extracurricular activities Excel file, Sheet 2, columns B and C
	ID_StartDate	StartDate ID. Foreign key from dimension table. Based on StartDate stored in Extracurricular activities Excel file, Sheet 2, column E
	ID_EndDate	EndDate ID. Foreign key from dimension table. Based on StartDate stored in Extracurricular activities Excel file, Sheet 2, column F
Student	One tuple describes one student (Implementation of SCD 2)	
	ID_Student	Student ID. Surrogate key - generated by database
	PESEL	Business key taken from PESEL from Student table in Student Grading System database
	FullName	Student's full name. Taken from FirstName and LastName columns in Student table in Student Grading System database
	isCurrent	"1" if information is current, otherwise "0" (SCD2 implementation)

	ParticipatesInExtraActivity	"1" if a student currently participates in some extracurricular activities, "0" if he is not (SCD2 implementation)
Teacher	One tuple describes one teacher who can assign grades	
	ID_Teacher	Teacher ID. Surrogate key - generated by database
	PESEL	Business key taken from PESEL from Teacher table in Student Grading System database
	FullName	Teacher's full name. Taken from FirstName and LastName columns in Teacher table in Student Grading System database
Subject	One tuple describes one subject taught in the school	
	ID_Subject	Subject ID. Surrogate key – generated by a database
	SubjectName	Represents name of the subject. Allowed values: math, science, history, english, art, literature, music, physical education, computer science, geography, biology – taken from SubjectName column from Subject table in Student Grading System database
Class	One tuple describes one class (e.g. "1B," "3C")	
	ID_Class	Class ID. Surrogate key – generated by the database
	ClassName	Represents name of the class. Allowed values: 1A, 1B, 1C, 1D, 1E, 2A, 2B, 2C, 2D, 2E, 3A, 3B, 3C, 3D, 3E, 4A, 4B, 4C, 4D, 4E, 5A, 5B, 5C, 5D, 5E, 6A, 6B, 6C, 6D, 6E, 7A, 7B, 7C, 7D, 7E, 8A, 8B, 8C, 8D, 8E – taken from ClassID column from Class table in Student Grading System database
	YearOfCreation	The year when class was created - taken from YearOfCreation column from Class table in Student Grading System database
Semester	One tuple describes one academic semester	

	ID_Semester	Semester ID. Surrogate key – generated by database
	SemesterName	Name of the semester. Allowed values: Summer####, Winter#### (#### indicates a year – e.g. 2021, 2024, 2025, etc.) - taken from SemesterName column from Semester table in Student Grading System database
Date	One tuple describes one calendar day. All the data in this table are generated tuple by tuple based on any calendar, before ETL process.	
Activity	One tuple describes one extracurricular activity	
	ID_Activity	Activity ID. Surrogate key - generated by database
	ActivityType	Type of extracurricular activity. Allowed values: sports, arts, volunteering, music. Taken from Extracurricular activities Excel file, Sheet 2, Column B
	ActivityName	Name of the activity. Allowed values: basketball, football, volleyball, running, choir, painting classes, community cleanup, school band, chess club Taken from Extracurricular activities Excel file, Sheet 2, Column C
	Frequency	Tells info about how often students participates in certain activity. Allowed values: One time activity, 1-2 times a week, 3-5 times a week. 6-7 times a week, each 2 weeks, each month. Taken from Extracurricular activities Excel file, Sheet 2, column D
	Location	Place where the activity is conducted. Allowed values: Playground, Library, Auditorium, Classroom, Gymnasium, Computer Lab, Science Lab, School Garden, Art Room, Music Room, Cafeteria, Hallway, Courtyard, Rooftop,

		Assembly Hall, Sports Field, Conference Room. Taken from Extracurricular activities Excel file, Sheet 2, Column H
Assignment	One tuple represents one assignment that can be graded	
	ID_Assignment	Assignment ID. Surrogate key – generated by a database
	AssignmentName	Name of the assignment. The possible values are described by the following formula: AssignmentType#Number - AssignmentType describes the type of the assignment (see below) - Number range 1 .. 100. Taken from AssignmentName column from Assignment table in Student Grading System database
	AssignmentType	Type of the assignment. Allowed values: exam, homework, project, quiz. Taken from AssignmentType from Assignment table in Student Grading System database