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

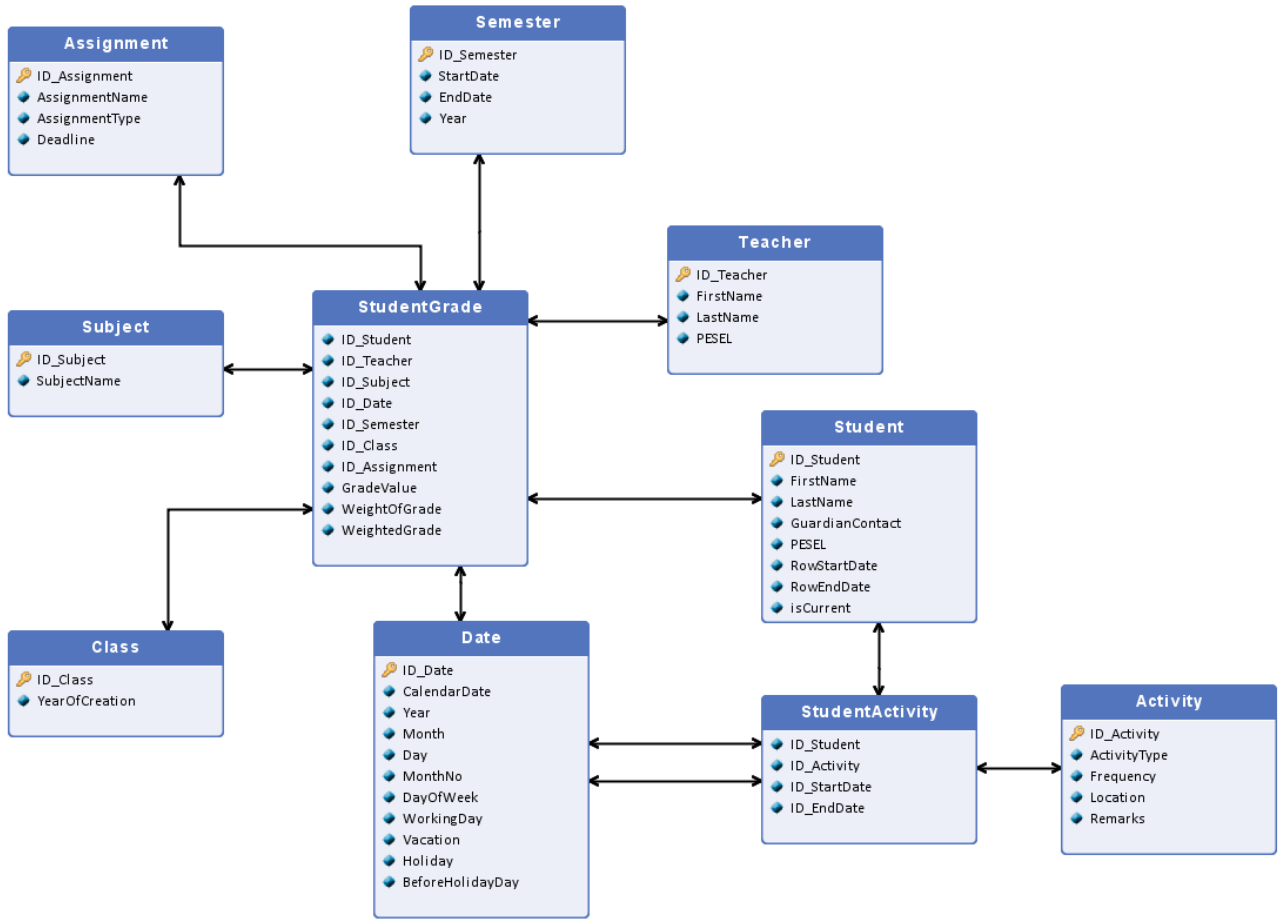

TABLE NAME	ATTRIBUTE	ATTRIBUTE TYPE	DESCRIPTION
StudentGrade (Fact table)	One tuple describes one fact of a student's grade		
	ID_Student	Numeric	FK to Student Student who received the grade
	ID_Teacher	Numeric	FK to Teacher The teacher who gave the grade
	ID_Subject	Numeric	FK to Subject The subject 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
	FirstName	Varchar(30)	Student's first name
	LastName	Varchar(30)	Student's last name
	GuardianContact	Varchar(40)	Email or phone number of a parent/guardian
	RowStartDate	Date	The date when this version of the student record became effective. If a new version is inserted later, that new version's RowStartDate will exceed this one
	RowEndDate	Date	The date when this version stopped being current. If it is still current then a date is set to 9999-12-31until a new version is created and then updated to the day before the new version's RowStartDate
	isCurrent	Boolean	1 if it is current record of student, 0 if historical (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
	FirstName	Varchar(30)	Teacher's first name
	LastName	Varchar(30)	Teacher's 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 (e.g. “Fall 2025,” “Spring 2026”)		
	ID_Semester	Numeric	PK (Surrogate key)
	StartDate	Date	Date the semester begins
	EndDate	Date	Date the semester ends
	Year	Integer	Year of the semester (e.g. 2025, 2026)
Date (Dimension table)	One tuple describes one calendar day		
	ID_Date	Numeric	PK (Surrogate key)
	CalendarDate	Date	Actual date
	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
	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: Christmas, National Day of Independence, Grandmother's Day, none, ... (and other similar of this kind)
	BeforeHolidayDay	Varchar(65)	Before holiday day. Allowed values: tomorrow is Christmas, tomorrow is National Day of Independence, none, etc.
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 (e.g. once a month, once a week)

	Location	Varchar(20)	Place where the activity is conducted (e.g. gym, library, etc.)
	Remarks	Varchar(50)	Remarks about the activity, possibly some achievements as well
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 (e.g. quiz#3, homework#7, project#1, etc.)
	AssignmentType	Varchar(8)	Type of the assignment. Allowed values: exam, homework, project, quiz
	Deadline	Date	Date when assignment is due

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)
- WeightedGrade – GradeValue * WeightOfGrade
- Sum of WeightedGrade – SUM(WeightedGrade)

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)
- Duration in days – (EndDate – StartDate)

Dimensions definitions

Dimensions for Fact 1 Recieving a grade:

Dimension/ Dim. Attribute	Table/Column	Type
STUDENT	Student	Dimension
STUDENT ID	Student.ID_Student	Dimension attribute
STUDENT PESEL	Student.PESEL	Dimension attribute
STUDENT FIRST NAME	Student.FIrstName	Dimension attribute
STUDENT LAST NAME	Student.LastName	Dimension attribute
GUARDIAN CONTACT	Student.GuardianContact	Dimension attribute
TEACHER	Teacher	Dimension
TEACHER ID	Teacher.ID_Teacher	Dimension attribute

TEACHER PESEL	Teacher.PESEL	Dimension attribute
TEACHER FIRST NAME	Teacher.FirstName	Dimension attribute
TEACHER LAST NAME	Teacher.LastName	Dimension attribute
SUBJECT	Subject	Dimension
SUBJECT ID	Subject.ID_Subject	Dimension attribute
SUBJECT NAME	Subject.SubjectName	Dimension attribute
CLASS	Class	Dimension
CLASS ID	Class.ID_Class	Dimension attribute
CLASS NAME	Class.ClassName	Dimension attribute
YEAR OF CREATION	Class.YearOfCreation	Dimension attribute
SEMESTER	Semester	Dimension
SEMESTER ID	Semester.ID_Semester	Dimension attribute
SEMESTER START DATE	Semester.StartDate	Dimension attribute
SEMESTER END DATE	Semester.EndDate	Dimension attribute
SEMESTER YEAR	Semester.Year	Dimension attribute
DATE	Date	Dimension
DATE ID	Date.ID_Date	Dimension attribute
CALENDAR DATE	Date.CalendarDate	Dimension attribute
DATE YEAR	Date.Year	Dimension attribute
DATE MONTH	Date.Month	Dimension attribute
DATE MONTH NO	Date.MonthNo	Dimension attribute
DATE DAY	Date.Day	Dimension attribute
DATE DAY OF WEEK	Date.DayOfWeek	Dimension attribute
WORKING DAY	Date.WorkingDay	Dimension attribute
VACATION	Date.Vacation	Dimension attribute
HOLIDAY	Date.Holiday	Dimension attribute
BEFORE HOLIDAY DAY	Date.BeforeHolidayDay	Dimension attribute
ASSIGNMENT	Assignment	Dimension
ASSIGNMENT ID	Assignment.ID_Assignment	Dimension attribute
ASSIGNMENT NAME	Assignment.AssignmentName	Dimension attribute
ASSIGNMENT TYPE	Assignment.AssignmentType	Dimension attribute
ASSIGNMENT DEADLINE	Assignment.Deadline	Dimension attribute
CLASS HIERARCHY	• Class.YearOfCreation •• Class.ClassName	Hierarchical dimension
SEMESTER HIERARCHY	• Semester.Year •• Semester.StartDay •• Semester.EndDate	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 NAME	Student.Name	Dimension Attribute
STUDENT AGE	Student.Age	Dimension Attribute
STUDENT GENDER	Student.Gender	Dimension Attribute
STUDENT CLASS	Student.Class	Dimension Attribute
ACTIVITY	Activity	Dimension
ACTIVITY NAME	Activity.Name	Dimension Attribute
ACTIVITY TYPE	Activity.Type	Dimension Attribute
ACTIVITY LOCATION	Activity.Location	Dimension Attribute
START DATE	Date	Dimension
START YEAR	Date.Year	Dimension Attribute
START MONTH	Date.Month	Dimension Attribute
START DAY	Date.Day	Dimension Attribute
END DAY	Date	Dimension
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 weighted grade improved the most in the last three months?

   Measure: Weighted average grade
   Dimension: Subject (dimension attributes: Subject Name)
   Dimension: Date (dimension attributes: Grade Day, Grade Month)

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

   Measure: Percentage change in weighted average grade
   Dimension: Subject (dimension attributes: Subject Name = Biology)
   Dimension: Date (dimension attributes: Grade Month, Grade Day)

3. Did the variance in student grades increase or decrease this month compared to the previous month?

   Measure: Variance in student grades
   Dimension: Date (dimension attributes: Month)

4. Did students who joined and participated in extracurricular activities in the past month experience a positive change in the overall weighted average grade from

all subjects?

Measure: Change in weighted average grade
Dimension: Student (dimension attributes: Participation in extracurricular activities, Student name)
Dimension: Date (dimension attributes: Month)

5. Did the students with the highest amount of assignments that have a deadline in the same month as they were assigned experience a negative change in the overall weighted average grade from all subjects during this month?

   Measure: Change in weighted average grade
   Dimension: Student (dimension attributes: Number of assignments with same-month deadlines, Student name)
   Dimension: Date (dimension attributes: Month)

6. Is there a significant difference (significant difference is around half a grade) between the average weighted grades of students who participate in extracurricular activities compared to those who do not?

   Measure: Difference in weighted average grade
   Dimension: Student (dimension attributes: Participation in extracurricular activities, Student name)

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

   Measure: Weighted average grade (students below 2.0)
   Dimension: Student (dimension attributes: Participation in extracurricular activities)
   Dimension: Subject (dimension attributes: Subject Name)

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

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

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

   Measure: Weighted average grade
   Dimension: Date (dimension attributes: Year)
   Dimension: Student (dimension attributes: Participation in extracurricular activities)

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

    Measure: Correlation coefficient (Extracurricular participation vs. Weighted average grade)
    Dimension: Student (dimension attributes: Number of extracurricular activities)