Requirements specification for Teaching Students business process

1. General description of business process

a. A general description of the business process and a description of the performance metrics generated by this process, possible current analytical problems.

The academic year includes a comprehensive educational curriculum, with assessments such as tests, homework assignments, and examinations. These assessments are graded on a scale of 1 to 6, with 6 denoting exemplary performance.

A dedicated subject teacher leads each academic class, providing specialized instruction tailored to the subject matter. Different instructors may oversee classes within the same educational level and subject, each with their own unique pedagogical approaches. At the end of each academic year, students' academic average grade is comprehensively analysed. The analysis serves two purposes: evaluating the effectiveness of educators and reporting on progress of students.

Metrics:

- **Test scores**: from each test students achieve grades 1-6 which present their academic performance
- **Number of absences**: during each lesson teacher marks attendance of students and at the end we are able to calculate total number of absences

Achieve the performance of students -> at least 3.75 grade point average of student

Decrease in absenteeism at least by 5% compared to last month.

b. Typical questions

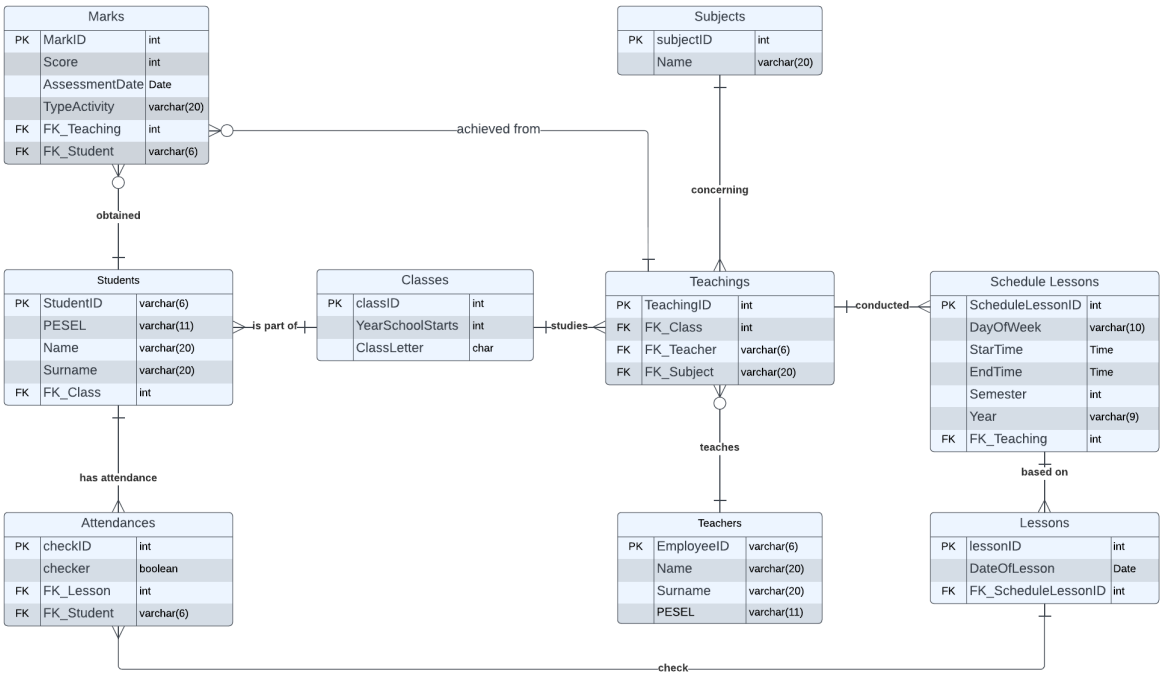
1. How many students graduated with an honours certificate in the previous year? (grade point average above 4.5)
2. At which level of learning (in which grade 4-8) children receive the highest grades?
3. Does the number of absences affect a student's final grade point average?
4. During which semester (winter/summer) there were more absences (last academics year)?
5. What is the level of education for boys and girls?
6. Whether the place of living (countryside/city) influences the students' academic average grade.
7. Are there any differences between genders when we take into account absences?
8. Is there any difference in number of absences when we take into account children place of living?
9. Does distance from school affect attendance?
10. Whether communication with parents through the teacher has a positive effect on the student's academic average grade and development?

c. Data

All information on progress and learning outcomes is recorded in the school system “EduTracker”, which is an electronic gradebook. This system stores information on all students, their grades in the subjects concerned and the teachers who issued them and carried out the teaching within the class.

In addition, data about employee and students are stored in the Director EXCEL sheet.

2. Data sources structures



"EduTracker”

Table name	Attribute	Attribute type	Description
Students	The student in the school network is identified by Personal Identification Number - studentID.		
	studentID	String – 6 characters	PK
	Name	String – 20 characters	First name
	Surname	String – 20 characters	Surname
	FK_Class	Numeric	FK auto generated
	PESEL	String – 11 characters	Polish Identification Number
Teachers	Teacher in the School network is identified by Personal Identification Number - employeeID. For teachers each employeeID ends with first (or second for “music”) letter of subject they teach.		
	employeeID	String – 6 characters	PK
	Name	String – 20 characters	First name
	Surname	String – 20 characters	Surname
	PESEL	String – 11 characters	Polish Identification Number
Attendances	The attendance grade applies to the pupil concerned during the lesson. It is updated after each lesson in the subject. Identified by artificial key.		
	CheckID	Numeric	PK, auto generated starting from 1 with increment 1
	checker	Boolean	For each subject, each student has an absence checker which changes its value to false after each specific lesson in which the student has been absent
	FK_Student	String – 6 characters	A FK from the student to whom the attendance information relates is provided
	FK_Lesson	Numeric	A FK from the lesson to which the attendance information of students relates
Marks	Each grade has its own unique key		
	markID	Numeric	PK, auto generated starting from 1 with increment 1

	Score	Numeric	Evaluation score. Ranges from 1 to 6
	AssessmentDate	Date	Date of assessment/entry of the grade into the system. This does not have to coincide with the time and date of the teacher's lesson.
	TypeActivity	String - 20 characters	Activity in class. Allowed values: exam, test, homework
	FK_Teaching	Numeric	FK Teaching in relation to which the assessment was given
	FK_Student	String – 6 characters	FK, Students ID number
Classes	Each class has its own number, but they repeat each year - that's why for classes we use artificial key.		
	classID	Numeric	PK, auto generated starting from 1 with increment 1
	ClassLetter	Character	For each year we have 3 groups. Thats why allowed letters are: A,B,C
	YearSchoolStarts	Numeric	The year in which students of this class started school
Subjects	Each subject has its own unique name that’s why it stands for the PK		
	subcjectID	Numeric	PK, auto generated starting from 1 with increment 1
	Name	String – 20 characters	Name of the subject. Allowed: Mathematics, Polish, English, Science, WOS, History, Music, Art, Religion, PE, IT, German.
Teachings	The teaching entity contains quite a few variables therefore we are not able to find a suitable compressed Composit Key. For this reason, an artificial key is used.		
	teachingID	Numeric	PK,

			auto generated starting from 1 with increment 1
	FK_Class	Numeric	FK, auto generated
	FK_Teacher	String – 6 characters	FK ID number of the teacher who conducts the subject
	FK_Subject	String – 20 characters	FK Type of the subject which is conducted
Lessons	Each scheduled lesson that was performed is called a lesson and has its own artificial key.		
	lessonID	Numeric	PK, auto generated starting from 1 with increment 1
	DateOfLesson	Date	Date when the lesson was conducted
	FK_ScheduleLesson	Numeric	FK Lesson is a part of lessons shedule
Schedule Lessons	Each scheduled lesson has its own artificial key because in one specific moment of time there are multiple different lessons scheduled.		
	ScheduleLessonID	Numeric	PK, auto generated starting from 1 with increment 1
	DayOfWeek	String – 10 characters	Day of the week when the subject for particular class takes place
	StartTime	Time	Starting time of the lesson
	EndTime	Time	End of lesson time
	Semester	Numeric	Number of the semester in which particular classes are taught. Winter semester – 1, Summer semester - 2
	Year	String – 9 characters	Years in which the plan applies. For example. For example, the

			learning year “2022/2023”
	FK_Teaching	Numeric	FK

RDB

Students(studentID, Name, Surname, PESEL, FK_Class REF Classes)

Teachers(employeeID, Name, Surname, PESEL)

Attendances(checkID, checker, FK_Student REF Students, FK_Lesson REF Lessons)

Marks(markID, Score, AssessmentDate, TypeActivity, FK_Teaching REF Teachings, FK_Student REF Students)

Classes(classID, ClassLetter, YearSchoolStarts)

Subjects(subjectID, Name)

Teachings(teachingID, FK_Class REF Classes, FK_Teacher REF Teachers, FK_Subject REF Subjects)

Lessons(lessonID, DateOfLesson, FK_ScheduleLesson REF ScheduleLessons)

ScheduleLessons(schLessonID, DayOfWeek, StartTime, EndTime, Semester, Year, FK_Teaching REF Teachings)

Relationships:

Obtained (Marks 0..n - 1 Students)
Students obtain marks. One student can obtain one, many or none marks, but each mark is obtained by only one student.
Achieved from (Marks 0..n - 1 Teachings)
Marks are achieved from teaching. From each teaching there can be achieved one, many or none marks, but each mark is achieved from one teaching.
Concerning (Subjects 1 – n Teachings)
Teachings concerns subjects. One teaching concerns only one subject, but one subject is being concern by many teachings.
Is part of (Students n – 1 Classes)
Students are a part of class. One student is a part of only one class, but in one class there can be many students.
Has attendance (Attendances n – 1 Students)
Students have attendance. One student has many attendances, but one attendance is for one student only.
Studies (Classes 1 – n Teachings)
Classes study teachings. One class studies many teachings, but one teaching is studied by only one class.

Conducted (Teachings 1 – n Schedule lessons)
Teaching is conducted by schedule lessons. One schedule lesson conducts one teaching, but one teaching is conducted by many schedule lessons.
Based on (Schedule lessons 1 – n Lessons)
Lessons are based on schedule lessons. One schedule lesson is the base for many lessons, but one lesson is based on one schedule lessons.
Teaches (Teachings 0..n - 1 Teacher)
Teachers teach teachings. One teacher teaches one, many or zero teachings, but one teaching is taught by one teacher.
Check (Attendances n – 1 Lessons)
Attendance is checked during lessons. One lesson checks many attendances, but one attendance is checked during one lesson.

Director Excel

Sheet 1 (Information about employees who work in the school, each line describes one employee, line 1 is a header row):
Column A – Employee's PESEL (numeric, 0 decimal precision),
Column B – Employee's identification number (text),
Column C – Employee's name (text),
Column D – Employee's surname (text), in case of surname change the column is updated,
Column E – Gender of employee (man, woman) (text),
Column F – Position of employee (teacher, assistant, cook, cleaner, director), in school there is only one director , each row means employment of a given person on a given position; in case of position change the date of end of work on the current position is written and there is a new line with a new position and a new date of employment; (text),
Column G – Street and house number (text), in case of change the column is updated,
Column H – City (text), in case of change the column is updated,
Column I – Postal code (text), in case of change the column is updated,
Column J – Phone number (numeric, 0 decimal precision), in case of change the column is updated,
Column K – E-mail address (text), in case of change the column is updated,
Column L – Date of employment for position (Date in format year - month - day, e.g. 2013-12-09),
Column M – Date of dismissal of current position (Date in format year - month - day, e.g. 2013-12-09), it is not set if the employee currently works on a given position.

Sheet 2 (Information about students who attend the school, each line describes one student, line 1 is a header row):

Column A – Student's PESEL (numeric, 0 decimal precision),
Column B – Student's identification number (text),
Column C – Student's name (text),
Column D – Student's surname (text), in case of surname change the column is updated,
Column E – Date of birth (Date in format year - month - day, e.g. 2013-12-09),
Column F – Place of birth (text),
Column G – Gender of student (man, woman) (text),

Column H – Street and house number (text), in case of change the column is updated,
Column I – City (text), in case of change the column is updated,
Column J – Postal code (text), in case of change the column is updated,
Column K – Place of living (countryside, city) (text), in case of change the column is updated,
Column L – Parent phone number (numeric, 0 decimal precision), in case of change the column is updated,
Column M – Parent e-mail address (text), in case of change the column is updated.

3. Scenarios of analytical problems

Why can we observe an increase / decrease in grade point average?

1. Compare the number of people (on the same level for each class) who graduated with an honours certificate in the previous year and 2 years before. (Grade point average above 4.5).
2. At which level of learning (in which grade 4-8) children receive the highest grades? Based on previous year.
3. Compare the academic grade point average of students based on tutoring of teachers who work for this school less than 2 years or more (teacher who are currently employed). Based on previous year.
4. Whether the place of living (countryside/city) influences the students' grade point average.
5. Whether communication with parents gives the positive effect on the students' grade point average -> checked for the first semester of the previous academic year

What influences the high absence rate among students?

1. Does the number of absences affect a student's final grade point average? Based on the previous year.
2. Which class has the highest absence rate in the current semester?
3. Which subject has the highest number of absences?
4. During which semester (winter/summer) there were more absences (last academics year)?
5. Compare the number of absences among genders (boys and girls).
6. Compare the number of absences of children taking into account their age. For the previous year.
7. Does distance from school affect attendance? -> no data

4. Data needed for analytical problems

Analytical problem:

Why can we observe an increase / decrease in grade point average?

1. *Compare the number of people (<u>on the same level</u> for each class) who graduated with an honours certificate in the previous year and 2 years before. (Grade point average above 4.5)*

   Number of the class (4-8) – "EduTracker", table Class, column YearSchoolStarts -> calculated based on current year

Number of people with honours certificate for previous year – count people with average from the grades from the whole year higher or equal to 4.5, "EduTracker", table Marks, column score -> sum of grades/number of grades >= 4.5

Number of people with honours certificate from 1 year earlier – similarly calculated

2. *At which level of learning (4-8) children receive the highest grades? Compare classes -> Based on previous year.*

   Class Number – "EduTracker", table Classes, column YearSchoolStarts -> calculated based on current year

   Grade point average – mean grade from each school level (4-8), "EduTracker", table Marks, column Score -> sum of all scores for each level/number of scores

3. *Compare the academic grade point average of students based on tutoring of teachers who work for this school less than 2 years or more (teacher who are currently employed). Based on previous year.*
   Years worked – this will be determined based on Director Excel, column M should be empty (it determines the date of dismissal), column L will help to calculate years of working for this school.

   Grade point average - mean grade from each group of teachers (currently working, less or more than 2 years), "EduTracker", table Marks, column Score -> sum of all scores for each level/number of scores.

4. *Whether the place of living (countryside/city) influences the students' grade point average.*

   StudentID – identification number of the student

   Place of living – this will be determined based on Directors EXCEL, sheet 2, column K

   Grade point average – average calculated from the grades from the whole year, "EduTracker", table Marks, column score -> sum of grades/number of grades

5. *Whether communication with parents gives the positive effect on the students' grade point average -> checked for the first semester of the previous academic year*

   StudentID – identification number of the student

   CommunicationConducted – no such information

   Grade point average before – average calculated from the grades from each student, calculated before the date of the meeting with parents -> also no information about the date of the meeting

   Grade point average after – average calculated from the grades from each student, calculated after the date of the meeting with parents -> one month after

What influences the high absence rate among students?

1. *Does the number of absences affect a student's final grade point average? Based on the previous year?*

StudentID – "EduTracker", table Student, column studentID

Number of absences – "EduTracker", table Attendance, column Check -> sum all attendance with "check" as False -> person was absent

Final grade point average – average calculated from the grades from the whole year, "EduTracker", table Marks, column score -> sum of grades/number of grades

2. *Which class has the highest absence rate in the current semester?*

   Class Number – "EduTracker", table Classes, column YearSchoolStarts -> calculated based on current year

   ClassLetter – "EduTracker", table Classes, column ClassLetter

   Number of absences – "EduTracker", table Attendance, column Check -> sum all attendance with "check" as False -> person was absent

3. *Which subject has the highest number of absences?*

   Subject name - "EduTracker", table Subject, column Name

   Number of absences – "EduTracker", table Attendance, column Check -> sum all attendance for particular subject with "check" as False -> person was absent

4. *During which semester (winter/summer) there were more absences (last academics year)?*

   Semester – "EduTracker", table Schedule Lessons, column Semester

   Number of absences – "EduTracker", table Attendance, column Check -> sum all attendance with "check" as False -> person was absent

5. *Compare the number of absences among genders (boys and girls).*

   Number of absences – "EduTracker", table Attendance, column Check -> sum all attendance with "check" as False -> person was absent

   Gender – this will be determined based on Directors EXCEL, sheet 2, columns G

6. *Compare the number of absences of children taking into account their age. For the previous year.*

   Age – this will be determined based on Directors EXCEL, sheet 2, columns E. We calculate age of student based on their date of birth and current date.

   Number of absences – "EduTracker", table Attendance, column Check -> sum all attendance with "check" as False -> person was absent

7. *Does distance from school affect attendance? -> no data*

   StudentID – identification number of the student

   Distance from school – there's no such information available in both data sources. The proposals for acquiring such information:

   - enhance Director Excel with distance to shopping malls
   - distance analysis with the use of tools like google.com/maps

Number of absences – "EduTracker", table Attendance, column Check -> sum all attendance with "check" as False -> person was absent

It is not possible to build a BI system to support Director in solving the problem of communication between parents and teachers. We need to introduce additional activities to overview the effectiveness of this business process. We suggest introducing a survey system conducted among students one month after the meeting with their parents. Such questionnaire should contain as minimum the following set of questions:

- Did your parent talk to you about learning?
- Did a parent want to help you with your studies?
- Did you find it beneficial to talk to your parents? / Did it make you learn more?
- Would you like a parent to help you with your studies more often?

Survey results are manually uploaded to the survey system and entered into the excel sheet. Sample structure of the survey sheet:

- Column A – studentID
- Column B – Talk conducted (whether the talk was conducted -> yes/no)
- Column C – Help with study (whether parents help with studying -> yes/no)
- Column D – Talk influence on learning (numeric from 1 to 5, where 1 – minimal influence, 5 – maximal influence)
- Column E – Help preferences (numeric from 1 to 5, where 1 – don't want/need help, 5 – want help from parents)