

# Requirements specification for Sale business process

## 1. General description of business process

### Students' frequency analysis – Frequency check-up

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

The attendance tracking process for a high school is as follows. At the beginning of each semester, the attendance officer for the school is assigned to handle the tracking process. Attendance data is collected no less frequently than once a week. As part of this process, a list of attendance records is prepared for each class in the school. Each attendance record is assigned a unique identifier in addition to the general record number. These attendance records are then updated with the students' attendance information, including the date, period, subjects and teachers conducting classes during which the student was absent and reason for any absence. Any missing attendance information is also recorded. Once the attendance information is collected, it is validated against the class schedule and attendance policies and stored for future reference. Any discrepancies or issues with the attendance information are also documented.

**The increase in the monthly average attendance per student so its level is no less than 85% monthly compared to the previous month.**

**The decrease in the number of unexcused absences per semester at a level not less than 10% compared to the previous semester.**

- b) Typical questions

What is the overall attendance rate for the school this month?

Which classes have the highest and lowest attendance rates this week?

What students have missed more than three classes this semester?

How does the attendance rate for this month compare to the attendance rate for the same month last year?

Which students have the highest and lowest attendance rates for this semester?

What are the most common reasons for student absences this month?

On which subject the attendance for each class is the lowest?

Which teacher conducts the class with the highest attendance?

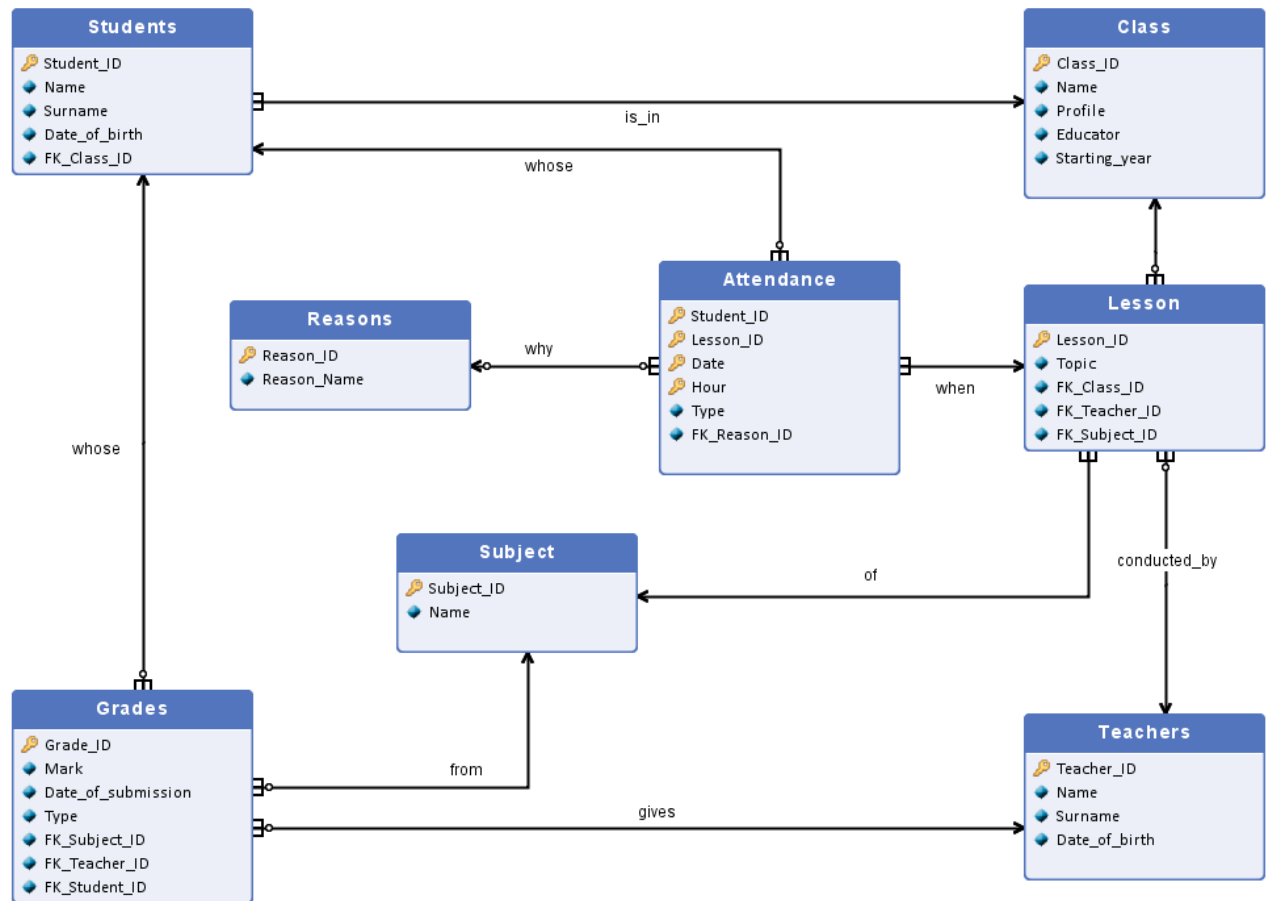
How does the attendance rate for this school compare to other schools in the district?

- c) Data

All data are extracted from the electronic grade book system – “EGradelit”. The system stores information about each student's attendance record, including the date of each absence or tardy and the reasons for the absence or tardy are extracted from specially created EXCEL file.

## 2. Data sources structures

### *EGradelit*



Students			
Set of all students in high school			
Name	Primary Key	Type	Description
Student_ID	Yes	natural number	Student's ID
Name	No	String – 20 characters	Student's Name
Surname	No	String – 20 characters	Student's Surname
Date_of_birth	No	Date	Student's date of birth in format YYYY-MM-DD
FK_Class_ID	No	Natural number	Class's ID

Class			
Set of all classes in school			
Name	Primary Key	Type	Description
Class_ID	Yes	natural number	Class's ID
Name	No	String – 2 characters	Class's name
Profile	No	String – 20 characters	Class's profile
Educator	No	String – 5 characters	Class's educator
Starting_year	No	String – 4 characters	Year of creation of class

Teachers			
Set of all teachers in school			
Name	Primary Key	Type	Description
Teacher_ID	Yes	natural number	Teacher's ID
Name	No	String – 20 characters	Teacher's Name
Surname	No	String – 20 characters	Teacher's Surname
Date_of_birth	No	Date	Teacher's date of birth in format YYYY-MM-DD

Subject			
Set of all subjects in school			
Name	Primary Key	Type	Description
Subject_ID	Yes	natural number	Subject's ID
Name	No	String – 20 characters	Subject's Name

Grades			
Set of all grades of all students in school			
Name	Primary Key	Type	Description
Grade_ID	Yes	natural number	Grade's ID

Mark	No	natural number	Mark from interval {1,2,3,4,5,6}
Type	No	Text (semestral, final)	Semestral or Final
Date_of_submission	No	DateTime	Date of submission of grade in format YYYY-MM-DD HH:MM
FK_Student_ID	No	Natural number	Student's ID
FK_Teacher_ID	No	Natural number	Teacher's ID
FK_Subject_ID	No	Natural number	Subject's ID

Attendance			
Set of all attendances of all students in school			
Name	Primary Key	Type	Description
Student_ID	Yes	natural number	Student's ID
Lesson_ID	Yes	natural number	Lesson's ID
Date	Yes	Date	Date of Lesson
Hour	Yes	DateTime	Hour of Lesson
Type	No	text (present, absent, late)	Type of attendance eg. Present, absent, late
FK_Reason_ID	No	Natural number	Reason's ID

Lesson			
Set of all lessons of all classes in school			
Name	Primary Key	Type	Description
Lesson_ID	Yes	natural number	Lesson's ID
Topic	No	String – 40 characters	Topic of the lesson
FK_Class_ID	No	Natural number	Class's ID
FK_Teacher_ID	No	Natural number	Teacher's ID
FK_Subject_ID	No	Natural number	Subject's ID

Reasons			
Set of all possible reasons of absences			
Name	Primary Key	Type	Description
Reason_ID	Yes	natural number	Reason's ID
Reason_Name	No	String – 40 characters	Reason description

## Attendance EXCEL

**Sheet 1** (Information about school events on which students' absence was recorded, each line describes one event, line 1 is a header row):

Column A - Student ID (numeric, 0 decimal precision),

Column B - Class ID (numeric, 0 decimal precision),

Column C - Event ID (numeric, 0 decimal precision),

Column D - Event name (text),

Column E - Event date (Date in format year - month - day, e.g., 2022-11-20),

Column F - Event location (text),

Column G - Event type (text, e.g., field trip, assembly, etc.),

Column H - Reason ID (numeric, 0 decimal precision), in case of absence.

## 3. Scenarios of analytical problems

Why was there an increase / decrease in absences this month?

1. Compare the number of absences of every student in relation to the previous months.
2. What is the most frequently present reason of absence in current and previous months?
3. Compare the average number of absences of classes in relation to other high schools.
4. What is the impact of the hour of lesson start on the average number of absences?
5. What is the rate of unexcused absences in current month in relation to previous months?
6. Is there an increase in average number of absences of a student week after grade submission date at the end of semester in comparison to the attendance week before the submission?

What impact on the attendance have interior (teachers, subjects, events) and exterior (holidays, weather) factors?

1. How weather affects average number of absences?
2. What is the average level of absences in the week before summer holidays?
3. What impact on average number of absences have the lesson with specific teacher?
4. What impact on average number of absences have specific subject?
5. What type of events are the most frequently missed by students from each class and what are the most frequently present reasons of absence?
6. What type of mental illnesses make impact on absences of students?
7. How event locations influence number of absent students and what are the most frequently present reasons of absence?

## 4. Data needed for analytical problems

Analytical problem: Why was there an increase / decrease in absences this month?

1. Compare the number of absences of every student in relation to the previous months.
  - Number of attendances of type "absent": *EGradel*, table *Attendance*, column *Type*
  - Student number: *EGradel*, table *Attendance*, column *Student\_ID*
  - Lesson's month: *EGradel*, table *Attendance*, column *Date*
2. What is the most frequently present reason of absence in current and previous months?
  - Number of attendances of type "absent": *EGradel*, table *Attendance*, column *Type*
  - Reason name: *EGradel*, table *Reasons*, column *Reason\_name*
  - Lesson's month: *EGradel*, table *Attendance*, column *Date*
3. Compare the average number of absences of classes that started learning in the same year in relation to other high schools.
  - Number of attendances of type "absent": *EGradel*, table *Attendance*, column *Type*
  - Year of starting school: *EGradel*, table *Class*, column *Starting\_year*
  - Number of absences in other schools: there's no such information available in both data sources. Information can be acquired through:  
<https://www.gov.pl/web/edukacja-i-nauka>
4. What is the impact of the hour of lesson start on the average number of absences?
  - Number of attendances of type "absent": *EGradel*, table *Attendance*, column *Type*
  - Hour of the lesson: *EGradel*, table *Attendance*, column *Hour*
5. What is the rate of unexcused absences in current month in relation to previous months?
  - Number of attendances of type "absent": *EGradel*, table *Attendance*, column *Type*
  - Reason with name "unexcused": *EGradel*, table *Reasons*, column *Reason\_name*
  - Lesson's month: *EGradel*, table *Attendance*, column *Date*
6. Is there an increase in average absences week after grade submission date of each subject at the end of semester in comparison to the attendance week before the submission?
  - Number of attendances of type "absent": *EGradel*, table *Attendance*, column *Type*
  - Day of the lesson: *EGradel*, table *Attendance*, column *Date* with restriction about one week before and after submission date
  - Submission date: *EGradel*, table *Grade*, column *Date\_of\_submission*
  - Subject: *EGradel*, table *Subject*, column *Subject\_ID*

Analytical problem: What impact on the attendance have interior (teachers, subjects, events) and exterior (holidays, weather) factors?

1. How weather affects average number of absences?
  - Weather states – there's no such information available in both data sources. The proposals for acquiring such information:

- Cooperating with weather stations that provide forecast services and data (e.g., Pogoda&Radar)
- Number of attendances of type “absent”: *EGradel*t, table *Attendance*, column *Type*
- 2. What is the average level of absences in the week before summer holidays?
  - Number of attendances of type “absent”: *EGradel*t, table *Attendance*, column *Type*
  - Day of the lesson: *EGradel*t, table *Attendance*, column *Date* with restriction about one week before summer holidays
- 3. What impact on average number of absences have the lesson with specific teacher?
  - Number of attendances of type “absent”: *EGradel*t, table *Attendance*, column *Type*
  - Teacher ID: *EGradel*t, table *Teachers*, column *Teacher\_ID*
- 4. What impact on average number of absences have specific subject?
  - Subject: *EGradel*t, table *Subject*, column *Subject\_ID*
  - Number of attendances of type “absent”: *EGradel*t, table *Attendance*, column *Type*
- 5. What type of events are the most frequently missed by students from each class and what are the most frequently present reasons of absence?
  - Output will be calculated based on *Attendance EXCEL*, sheet 1, column *B*, *G* and *H*, calculating sum of occurrences of the same event type in column *G* for each class specified in column *B* connecting it by column *H* to *EGradel*t, table *Reasons*, column *Reason\_ID* to obtain information about reason name.
- 6. What type of mental illnesses make impact on absences of students?
  - Number of attendances of type “absent”: *EGradel*t, table *Attendance*, column *Type*
  - Type of illnesses: there’s no such information available in both data sources.
- 7. How event locations influence number of absent students and what are the most frequently present reasons of absence?
  - Output will be calculated based on *Attendance EXCEL*, sheet 1, column *A*, *F* and *H*, for each specific event location in column *F* calculate sum of occurrences of different students from column *A* connecting it by column *H* to *EGradel*t, table *Reasons*, column *Reason\_ID* to obtain information about reason name.