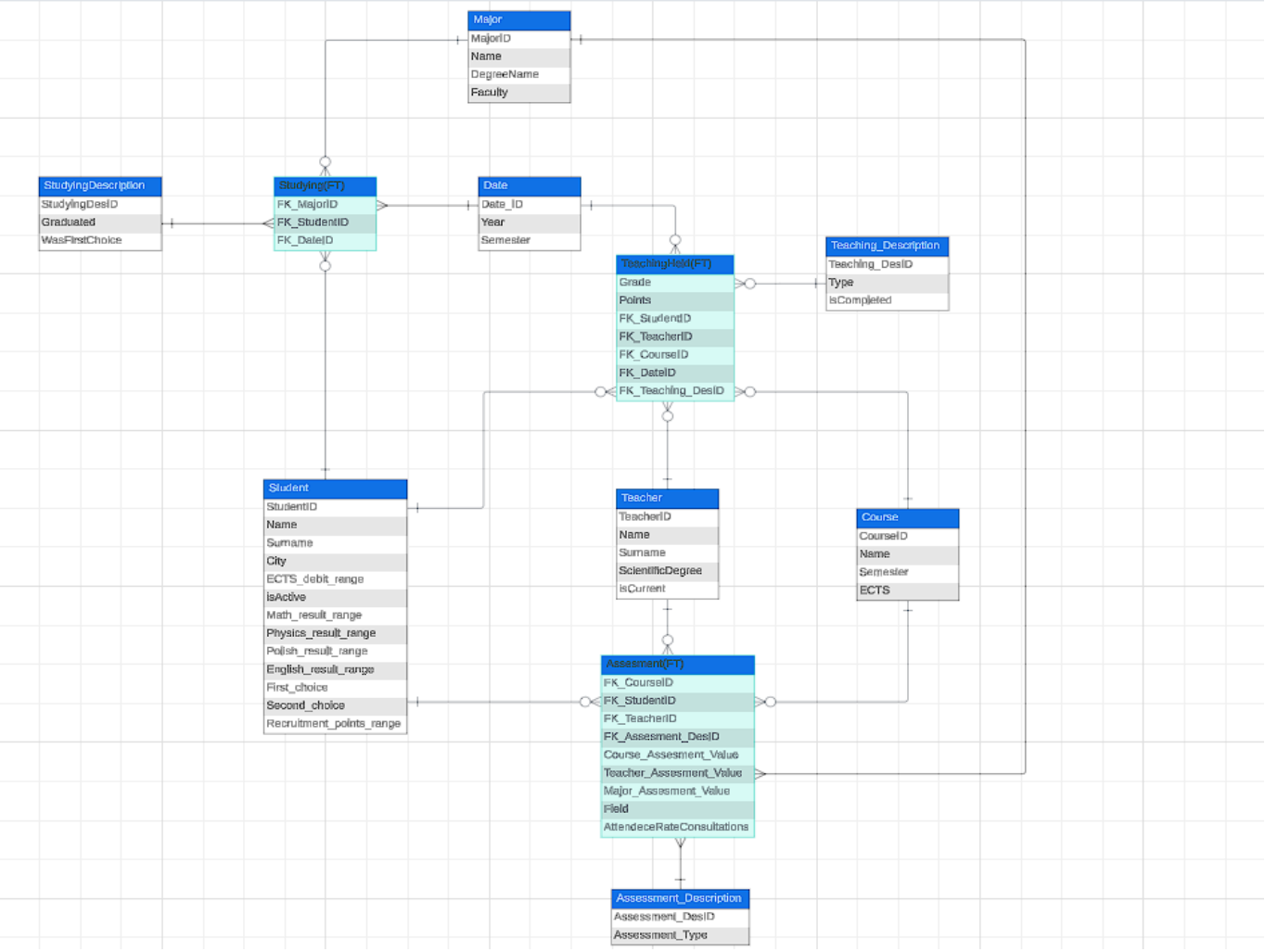
# Pomeranian University – Data warehouse design

## Business process

The Date warehouse is designed for Teaching business process. This process is described in the document Specification of business processes in Pomeranian University.

## Relational Database schema




| TABLE NAME | ATTRIBUTE | ATTRIBUTE TYPE | DESCRIPTION |
|---|---|---|---|

| **TeachingHeld (Fact Table)** | **One tuple describes one fact of Teaching** | | |
|---|---|---|---|
| | StudentID | Numeric | FK_StudentID Student |
| | TeacherID | Numeric | FK_TeacherID Teacher |
| | CourseID | Numeric | FK_CourseID Course |
| | DateID | Numeric | FK_DateID Date |
| | Teaching_DesID | Numeric | FK_ Teaching_DesID Teaching_Description |
| | Grade | Numeric | Grade obtained for given teaching by student |
| | Points | Numeric | Points obtained for given teaching by student |
| **Studying (Fact table)** | **One tuple describes one fact of Teaching** | | |
| | MajorID | Numeric | FK_MajorID Major |
| | StudentID | Numeric | FK_StudentID Student |
| | DateID | Numeric | FK_DateID Date |
| **Assessment (Fact Table)** | **One tuple describes one fact of assessing a course and a teacher** | | |
| | CourseID | Numeric | FK_CourseID Course |
| | StudentID | Numeric | FK_StudentID Student |
| | TeacherID | Numeric | FK_TeacherID Teacher |
| | Assesment_DesID | Numeric | FK_Assesment_DesID Assesment_Description |
| | Course_Assesment_Value | Numeric | Assesment given from Student for Course that he/she participated in (from 1 to 5) |
| | Teacher_Assesment_Value | Numeric | Assesment given from Student for Teacher who conducted Teaching in given Course (from 1 to 5) |
| | AttendenceRate Consultaution | Numeric | The attendence rate of a student on cosultations from a specific course |
| **Teacher (Dimension Table)** | **One tuple describes one teacher** | | |
| | TeacherID | Numeric | PK |
| | Name | Varchar(20) | First name |
| | Surname | Varchar(20) | Surname |
| | ScientificDegree SCD | Varchar(30) | Scientific degree. Allowed values: Master, bachelor, doctor, engineer, associate |

| | | | |
|---|---|---|---|
| | isCurrent | Bool | Is information about teacher current ?<br>Allowed values: yes, no (SCD2 implementation) |
| **Student (Dimension Table)** | **One tuple describes one student** | | |
| | StudentID | Numeric | PK |
| | Name | Varchar(20) | First name |
| | Surname | Varchar(20) | Surname |
| | City | Varchar(20) | Hometown of a student |
| | ECTS_debit_range | Varchar(20) | Amount of ECTS points debit that the student has.<br>Allowed values:<br>*No – if ECTS_debit = 0<br>*Low – if ECTS_debit >0 && ECTS_debit <12<br>*Medium – if ECTS_debit >=12 && ECTS_debit < 18<br>*High – if ECTS_debit >=18 |
| | Math_result_range | Varchar(10) | Result from mathematics matura exam in percentages<br>Allowed values:<br>*Low – if Math_result >= 0 && Math_result < 40<br>*Mid – if Math_result >= 40 && Math_result < 80<br>*High – if Math_result >= 80 |
| | Physics_result_range | Varchar(10) | Result from physics matura exam in percentages<br>Allowed values:<br>*Low – if Physics_result >= 0 && Physics_result < 40<br>*Mid – if Physics_result >= 40 && Physics_result < 80<br>*High – if Physics_result >= 80 |
| | Polish_result_range | Varchar(10) | Result from polish language matura exam in percentages |

| | | | |
|---|---|---|---|
| | | | Allowed values:<br>*Low – if<br>Polish_result >= 0 &&<br>Polish_result < 40<br>*Mid – if<br>Polish_result >= 40 &&<br>Polish_result < 80<br>*High – if<br>Polish_result >= 80 |
| | English_result_range | Varchar(10) | Result from english language matura exam in percentages<br>Allowed values:<br>*Low – if<br>English_result >= 0 &&<br>English_result < 40<br>*Mid – if<br>English_result >= 40 &&<br>English_result < 80<br>*High – if<br>English_result >= 80 |
| | First_choice | Varchar(20) | The name of the course that was the first choice of the student |
| | Second_choice | Varchar(20) | The name of the course that was the second choice of the student |
| | Recruitment_points_range | Varchar(10) | Amount of recruitment points calculated from the matura exam results<br>Allowed values:<br>*Low – if<br>Recruitment_points >= 0 &&<br>Recruitment_points < 40<br>*Mid – if<br>Recruitment_points >= 40 &&<br>Recruitment_points < 80<br>*High – if<br>Recruitment_points >= 80 |
| | isActive | Boolean | Is information about student current? (SCD2 implementation) |
| **Major (Dimensioin Table)** | **One tuple describes one major** | | |
| | MajorID | Numeric | PK |
| | Name | Varchar(20) | Name of the major |
| | DegreeName | Varchar(20) | Scientific degree that the student obtains after graduating from this major |

| | | | |
|---|---|---|---|
| | Faculty | Varchar(50) | Faculty that the major belongs to |
| **Course (Dimension Table)** | **One tuple describes one course** | | |
| | CourseID | Numeric | PK |
| | Name | Varchar(40) | Name of the course |
| | Semester | Numeric | The semester on which this course is conducted |
| | ECTS | Numeric | Amount of ECTS points this course is worth |
| **Date (Dimension Table)** | **One tuple describes the date of the conducted teaching** | | |
| | Date_ID | Numeric | PK |
| | Year | 4 digits | The year that the teaching/studying was conducted in |
| | Term | Numeric | The semester that the teaching/studying was conducted in |
| **StudyingDescription (Dimension Table)** | **One tuple describes the studying details** | | |
| | StudyingDesID | Numeric | PK |
| | Graduated | Varchar(3) | Did student graduate? Possible values: yes/no |
| | WasFirstChoice | Varchar(3) | Was the major that a student is pursuing his/her first choice in recruitment stage? |
| **Assessment_Description (Dimension Table)** | **One tuple describes assesments' type** | | |
| | Assesment_DesID | Numeric | PK |
| | Type | Varchar(13) | If Course_assesment_Value or Teacher_assesment_Value equals to:<br>1 – very bad<br>2 – not that bad<br>3 – mediocre<br>4 – good<br>5 - extraordinary |
| **Teaching_Description (Dimension Table)** | **One tuple describes teaching details** | | |
| | Teaching_DesID | Numeric | PK |
| | Type | Varchar(10) | Type of the conducted teaching<br>Allowed values: Stationary, online, hybrid |
| | isCompleted | Boolean | Is the teaching completed? Allowed values: yes, no (SCD2 implementation) |

# Dimensional model

Fact definitions

**Fact 1 Teaching fact:** Teaching a specified curse in specified semester on specified major with specified time of conduction and with specified type. Taught by a specified teacher with specified academic title and specified number of publications. Teaching conducted with a specified student of specified major and with specified number of ECTS missing.

Fact table: Teaching

Granularity:

- a specified teaching type
- a specified start date of course
- a specified student with specified number of ECTS missing
- a specified teacher with specified academic title
- a specified course from specified semester with specified number of ECTS

Measures and aggregation functions:
Number of teaching facts - COUNT(1)
Number of points collected - SUM(Points)
Number of completed teachings - COUNT(1), isCompleted is yes
Average grade - SUM(Grade)/Number of teaching facts

**Fact 2 Assessment fact:** Assessment of a specified course and specified teacher given by specified student on specified major in specified semester.

Fact table: Assessment

Granularity:

- a specified student with specified number of ECTS missing
- a specified teacher with specified academic title
- a specified course from specified semester with specified number of ECTS
- a specified major

Measures and aggregation functions:

Number of assessments facts for given major - COUNT(1)
Number of good/poor assessments – COUNT(1), Assessment_type is good/poor
Major assessment(average value of courses assessments that belongs to given major) – SUM(Course_Assessment)/COUNT(Course_Assessment)

**Fact 3 Studying fact:** Studying on a specified major with specified starting date done by specified student.

Fact resulting from “many to many” relationship.
Fact table: Studying

Granularity:

- a specified student with specified number of ECTS missing
- a specified major with specified start year, major and faculty name, and a degree after graduation

Measures and aggregation functions:
Number of students - COUNT(1)
Number of graduations - COUNT(1), Graduated is yes
Numbers of major being first choice - COUNT(1), WasFirstChoice is yes

# Dimension tables

Dimensions for **Fact 1 Teaching fact:**

| **DIMENSION/DIMENSION ATTRIBUTE** | **TABLE/COLUMN** | **TYPE** |
|---|---|---|
| COURSE | Course | Dimension |
| Course name | Course.Name | Dimension attribute |
| Semester | Course.Semester | Dimension attribute |
| ECTS | Course.ECTS | Dimension attribute |
| TEACHER | Teacher | Dimension |
| Teacher Name | Teacher.Name | Dimension attribute |
| Teacher Surname | Teacher.Surname | Dimension attribute |
| Scientific degree SCD | Teacher.ScientificDegree | Dimension attribute |
| STUDENT | Student | Dimension |
| Student Name | Student.Name | Dimension attribute |
| Student Surname | Student.Surname | Dimension attribute |
| City | Student.City | Dimension attribute |
| ECTS debit | Student.ECTS_debit_range | Dimension attribute |
| DATE HIERARCHY | *Date.Year<br>**Date.Term | Hierarchical dimension |
| DATE | Date | Dimension |
| Teaching Year | Date.Year | Dimension attribute |
| Teaching Term | Date.Term | Dimension attribute |
| TEACHING_DESCRIPTION | TeachingDescription | Dimension |
| Teaching Type | TeachingDescription.Type | Dimension attribute |
| Completeness of teaching | TeachingDescription.IsCompleted | Dimension attribute |

| MAJOR | Major | Dimension |
|---|---|---|
| Major name | Major.Name | Dimension attribute |
| Degree name | Major.DegreeName | Dimension attribute |
| Faculty | Major.Facluty | Dimension attribute |

Dimension table for **Fact 2 Assessment fact:**

| **DIMENSION/DIMENSION ATTRIBUTE** | **TABLE/COLUMN** | **TYPE** |
|---|---|---|
| COURSE | Course | Dimension |
| Name | Course.Name | Dimension attribute |
| Semester | Course.Semester | Dimension attribute |
| ECTS | Course.ECTS | Dimension attribute |
| TEACHER | Teacher | Dimension |
| Name | Teacher.Name | Dimension attribute |
| Surname | Teacher.Surname | Dimension attribute |
| Scientific degree | Teacher.ScientificDegree | Dimension attribute |
| STUDENT | Student | Dimension |
| ECTS debit | Student.ECTS_debit_range | Dimension attribute |
| MAJOR | Major | Dimension |
| Major name | Major.Name | Dimension attribute |
| ASSESSMENT DESCRIPTION | Assesment Description | Dimension |
| Course Assessment | Assesment_Description.Assesment_Type | Dimension attribute |
| Teacher Assessment | Assesment_Description.Assesment_type | Dimension attribute |

Dimension table for **Fact 3 Studying fact**

| **DIMENSION/DIMENSION ATTRIBUTE** | **TABLE/COLUMN** | **TYPE** |
|---|---|---|
| STUDENT | Student | Dimension |
| Student Name | Student.Name | Dimension attribute |
| Student Surname | Student.Surname | Dimension attribute |
| ECTS_debit | Student.ECTS_debit_range | Dimension attribute |
| isActive | Student.isActive | Dimension attribute |
| DATE HIERARCHY | *Date.Year<br>**Date.Term | Hierarchical dimension |
| DATE | Date | Dimension |
| Studying year | Date.Year | Dimension attribute |
| Studying term | Date.Term | Dimension attribute |
| MAJOR | Major | Dimension attribute |
| Name | Major.Name | Dimension attribute |
| STUDYING DESCRIPTION | StudyingDescription | Dimension |
| Graduation of studying | StudyingDescription.Graduated | Dimension attribute |
| Major chosen | StudyingDescription.WasFirstChoice | Dimension attribute |

## Checking the feasibility of queries based on the multidimensional model

1. Compare the passing rate of students of first semester that studies their first choice major to those who studies their second choice.
   Measure: Number of completed teachings
   Dimension: StudyingDescription (dimension attributes: WasFirstChoice)
   Dimension: Date (dimension attributes: Semester)

2. What are the hardest courses? (passing rate)
   Measure: Number of completed teachings
   Dimension: Course (dimension attribute: Name, Semester)

3. Compare passing rate of first semester of each major depending on students' math recruitment points. (We assume that each major has advanced math subjects in first semester)
   Measure: Number of completed teachings
   Dimension: Student (dimension attribute: Math_result_range, StudentID)
   Dimension: Major (dimension attribute: Name)

4. What number of applicants for the top five most popular majors are from Gdańsk?
   Measure: Number of Students
   Dimension: Major (dimension attribute: Name)
   Dimension: Student (dimension attribute: City)

5. Compare passing rates in first semester of studies of students depending on their recruitments points.
   Measure: Number of completed teachings
   Dimension: Student (dimension attribute: Reqruitment_points_range)
   Dimension: Course (dimension attribute: Semester)

6. How many students who rated the classes poorly (1,2) attended the consultations?
   Measure: Number of Students
   Dimension: Assessment_description (dimension attribute: Assessment_Type)
   Dimension: Assessment (dimension attribute: AttendanceRateConsultation)

7. Which courses were graded the worst in each major by students?
   Measure: Number of good/poor assessments
   Dimension: Major (dimension attribute: Name)

8. Which teacher was graded the worst in each major by students?
   Measure: Number of good/poor assessments
   Dimension: Assessment (dimension attribute: TeacherID)

9. What are the top three best rated courses?

   Measure: Number of good/poor assessments
   Dimension: Course (dimension attribute: Name)

10. Does the teacher's assessment influence the consultations attendance rate?

Measure: Number of good/poor assessments
Dimension: Assessment (dimension attribute: AttendanceRateConsultation)
Dimension: Assessment (dimension attribute: TeacherID)

# Checking if there are data in the data sources needed to fill in the data warehouse

| TABLE NAME | COLUMN | SOURCE |
|---|---|---|
| **TeachingHeld** | **One tuple describes one fact of conducted teaching.** | |
| | StudentID | Student ID. Foreign key from dimension table. Based on StudentID stored in a table Student from SQL Database source |
| | TeacherID | Teacher ID. Foreign key from dimension table. Based on Teacher stored in a table Teacher from SQL Database source |

| | | |
|---|---|---|
| | CourseID | Course ID. Foreign key from dimension table. Based on DegreeCourseID stored in a table DegreeCourse from SQL Database source |
| | DateID | Date ID. Foreign key from the dimension table |
| | Teaching_DescriptionID | Teaching description ID. Foreign key from the dimension table |
| | Grade | Grade given by the teacher, taken from the Teachings table from SQL Database source |
| | Points | Points obtained by the student, taken from the Teachings table from SQL Database source |
| **Assessment** | **One tuple describes one fact of course and teacher assessment.** | |
| | CourseID | Course ID. Foreign key from dimension table. Based on DegreeCourseID stored in a table DegreeCourse from SQL Database source |
| | StudentID | Student ID. Foreign key from dimension table. Based on StudentID stored in a table Student from SQL Database source |
| | TeacherID | Teacher ID. Foreign key from dimension table. Based on Teacher stored in a table Teacher from SQL Database source |
| | Assesment_DesID | Assesment description ID. Foreign key from dimension table |
| | Course_Assesment_Value | Assesment value regarding the course given by a student. Taken from column E in Survey Excel source. Allowed values: 1-5 |
| | Teacher_Assesment_Value | Assesment value regarding the teachergiven by a student. Taken from column D in Survey Excel source. Allowed values: 1-5 |
| | AttendanceRateConsultations | Attendance rate throught the semester of a student in consultations from a given course. Based on Attendence rate column in Sheet 2 of Excel file source. |
| **Studying** | **One tuple describes one fact of student pursuing a major** | |
| | MajorID | Major ID. Foreign key from dimension table. Based on |

| | | |
|---|---|---|
| | | MajorID stored in a table Major from SQL Database source |
| | StudentID | Student ID. Foreign key from dimension table. Based on StudentID stored in a table Student from SQL Database source |
| | DateID | Date ID. Foreign key from dimension table. |
| **Teacher** | **One tuple describes one teacher with specified scientific degree (implementation of SCD2)** | |
| | TeacherID | Teacher ID, taken from Teacher table in SQL Database source |
| | Name | Teacher's name, taken from Teacher table in SQL Database source |
| | Surname | Teacher's surname taken from Teacher table in SQL Database source |
| | ScientificDegree | Teacher's scientific degree, taken from the SQL Database source |
| | isCurrent | If information (about scientific degree) is current - "1", otherwise - "0" |
| **Student** | **One tuple describes one student** | |
| | StudentID | Student ID, taken from the Student table in SQL Database source |
| | Name | Student's name taken from Student table in SQL Database source |
| | Surname | Student's surname taken from Student table in SQL Database source |
| | City | City that the student comes from, taken from column B (City), Sheet 2 in Excel source |
| | ECTS_debit_range | Amount of ECTS points debit that the student has, taken from table Student in SQL Database source |
| | isActive | If ECTS debit does not exceed 18 - "1", otherwise - "0" |
| | Math_result_range | Math matura exam results, taken from column C (Math result), Sheet 1, Excel source |
| | Physics_result_range | Physics matura exam results, taken from column D (Physics result), Sheet 1, Excel source |
| | Polish_result_range | Polish matura exam results, taken from column E (Polish result), Sheet 1, Excel source |

| | | |
|---|---|---|
| | English_result_range | English matura exam results, taken from column F (English result), Sheet 1, Excel source |
| | First_choice | First choice of course, taken from column G, Sheet 1, Excel source |
| | Second_choice | Second choice of course, taken from column H, Sheet 1, Excel source |
| | Recruitment_points_range | Recruitment point level, taken from column I, Sheet 1, Excel source |
| **Major** | **One tuple describes one major** | |
| | MajorID | Major ID, taken from the table DegreeCourse from SQL Database source |
| | Name | Major name, taken from the table DegreeCourse from SQL Database source |
| | DegreeName | Degree name obtained after graduating, taken from DegreeCourse table in SQL Database source |
| | Faculty | Faculty name, taken from table Faculty in SQL Database source |
| **Course** | **One tuple describes one course** | |
| | CourseID | CourseID, taken from the table Course from SQL Database source |
| | Name | Name, taken from the table Course from SQL Database source |
| | Semester | Semester, taken from the table Course from SQL Database source |
| | ECTS | ECTS, taken from the table Course from SQL Database source |
| **Date** | **One tuple describes the date of the conducted teaching** | |
| | DateID | Date ID. Surrogate key generated by the database |
| | Year | Starting year of the specific major, taken from DegreeCourse table in the SQL Database source |
| | Semester | Semester on which the teaching was conducted. Taken from the Course table in the SQL Database source |
| **Studying Description** | **One tuple describes the studying details** | |
| | StudyingDesID | Studying description ID. Surrogate key generated by the database |

| | | |
|---|---|---|
| | Graduated | Value is determinated by value of the ECTS_debit_range attribute from the Student table<br>Allowed values:<br>Yes – if ECTS_debit_range=No<br>No – if ECTS_debit_range !=No |
| | WasFirstChoice | Attribute indicates whether the major that student is pursuing was his/her first-choice major. Taken from column G, Sheet 1, Excel source |
| **Assessment Description** | **One tuple describes assessments' type** | |
| | Assessment_DesID | Assessment Description ID. Surrogate key generated by the database |
| | Assessment_Type | Assessment value in categories, taken from column D or E in Survey Excel source |
| **Teaching Description** | **One tuple describes teaching details** | |
| | Teaching_DesID | Teaching Description ID. Surrogate key generated by the database |
| | Type | Type of a teaching, taken from the Teachings table in SQL Database source |
| | isCompleted | Attribute indicates if the student completed the teaching, taken from the Teachings table in SQL Database source. |