|  |
| --- |
|  |
| Analysis of students’ performance at the Vilnius College of Higher Education (VIKO) from 2001 to 2009 |
|  |

|  |
| --- |
| Julia Giszczak, Dominik Karaś |

Obraz zawierający tekst, Czcionka, krąg, Grafika

Opis wygenerowany automatycznie

**Analysis of students at the Vilnius College of Higher Education (VIKO)**

Content:

[1. Introduction 2](#_Toc137487214)

[2. Development of BI system 2](#_Toc137487215)

[2.1. Original data and ETL process 2](#_Toc137487216)

[2.2. DWH - Star Scheme 4](#_Toc137487217)

[2.3. KPIs 5](#_Toc137487218)

[2.4. OLAP analysis and Interpretation 7](#_Toc137487219)

[3. Conclusions 13](#_Toc137487220)

[4. References 14](#_Toc137487221)

1. **Introduction**

The subject of the project is the analysis of students’ result in the Vilnius College of Higher Education (VIKO). The goal of this analysis is to find associations between students results with students’ personal details and background and find the answer to what influences students’ grades the most and why. The study covers the years 2001 – 2009 and includes 475 students.

Vilnius College of Higher Education (VIKO) is the largest accredited higher professional education institution in Lithuania. VIKO has over 10000 students, who study there under 42 study programmes in 8 faculties. The Faculty of Electronics and informatics is the biggest department.

1. **Development of BI system**

Obraz zawierający tekst, zrzut ekranu, oprogramowanie, Oprogramowanie multimedialne

Opis wygenerowany automatycznie

* 1. **Original data and ETL process**

**Table 1. Original data**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Subject | Studies | School | Each study programm | | |
| **Background** | | **Grade** |
| Study programme | Department | County | StudentId | StudentId | |
| Subjects group | Study Programme | City | Name | CourseId | |
| Numer of credits | National code | School type | Surname | Grade | |
| Subject Name | Specialization | School | School |  | |
| Course ID |  |  | Place of birth |  | |
|  |  |  | Date of birth |  | |
|  |  |  | Gender |  | |
|  |  |  | Grading year |  | |
|  |  |  | Entry mark |  | |
|  |  |  | Specialization |  | |

The data consisted of five directories and three separate excel files. Directories included:

* ET – Electronic engineering
* IS – Information systems
* KT – Computer engineering
* PK – Computer programming
* TK - Telecommunications

All of these directories included background and grades about each student from each study programme. The files contained in these directories required renaming the final project grading depending on the name of the study programme. The grades also required adjustments in terms of the value and as a result they were multiplied by 2. Another change that had to be made concerned the dates of birth, because some of them contained incorrect values such as 31st of November or 29th of February, which caused errors with calculating the age of the students. These kinds of values have been changed to the day before.

The other 3 excel files contained information about schools, studies and subjects of the students, but they did not require any adjustments.

The second step was uploading data to Power BI. This included separating the tables with students’ grades and background from the same files. Then, uploaded data required transforming which included deleting empty rows, errors, renaming the columns to the values from the first rows and changing the data type for some of the columns.

Another step was appending queries in order to concatenate rows from different tables into new tables containing information about all the students background and their grades.

The next step was adding new columns necessary to answer the questions later in the project. These included:

* Year of birth extracted from the date of birth column
* Entry year, based on the Source.Name column
* Age of the student, calculated with the formula: Age = grading year – year of birth
* Number of years, which described how many years it took the student to finish the study program. It was calculated with the following formula: NoYears = Grading year – entry year
* Entry mark interval, which grouped the students according to the following intervals:
  + 10-12
  + 12-14
  + 14-16
  + 16-18
  + 18-20
* Final mark, calculated as a arithmetic average of the grades for each student
* Final grade, calculated as a weighted average using number of credits for each subject as weights. This column was calculated using following formula:

Final grade = SUMX(RELATEDTABLE(Grades),

    Grades[Grade] \* RELATED(Subjects[Number of credits])

) / SUMX(RELATEDTABLE(Grades), RELATED(Subjects[Number of credits]))

* Final grade interval according to the following intervals:
  + 10-12 – Satisfactory
  + 12-16 – Good
  + 16-18 – Very good
  + 18-20 - Excellent
  1. **DWH - Star Scheme**

Fact table: Grades

Measures: Grade

Dimensions: School, Students, Studies, Subjects

**DWH:**

**Grades** (StudentId, Course ID, Source.Name, Grade)

**Schools** (City, County, School, School type)

**Student** (StudentId, Name, Surname, School, place of birth, gender, Grading Year, Entry Mark, Specialization, Entry Year, Year, Age, Entry Mark interval, Final Grade, Final grade interval)

**Studies** (Department, Study Programme, national code, Specialization)

**Subjects** (Course ID, Number of credits, Study programme, Subject Name, Subject group)

**Relationships:**

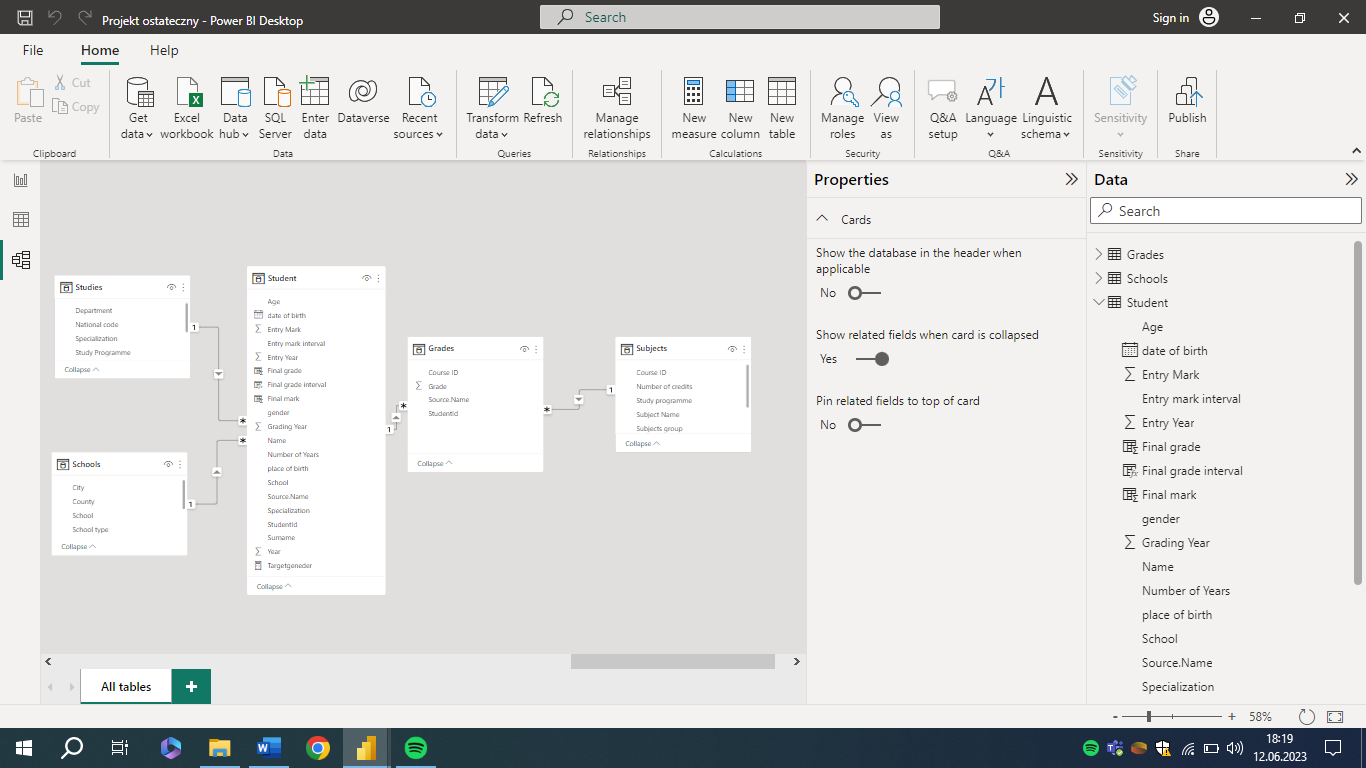
Grades (Course ID) – Subject (Course ID)

Grades (StudentId) – Student(StudentId)

Student(School) – Schools(School)

Student(Specialization) – Studies (Specialization)

**Figure 1. Star scheme**



## **KPIs**

* **Rate of the students which finished the program study in 3 years (86,53% - target 90%)**
* **Rate of the students which finished the program study in 4 years (6,74%)**
* **Rate of the students which finished the program study in 5 years (1,26%)**

The rate of the students who finished the study programme in 3 years is an important KPI, because it shows the teaching quality of the university. The higher this value is, the better the teaching quality. Therefore, the universities should strive to increase this value. Number of students who finish college in 2, 3, 4 or 5 years is shown on the figure below.

**Figure 2. Number of students by number of years to finish the study programme**

Obraz zawierający tekst, zrzut ekranu, oprogramowanie, Ikona komputerowa

Opis wygenerowany automatycznie

* **Rate of the students who enter the university with the entry mark between 16-20 (38,52%- target 50%)**

Another important measure is the rate of the students with the entry mark interval between 16-20. The higher this KPI value is the better students university attracts. The rate of the students for each entry mark interval is shown on the figure below.

**Figure 3. Number of students by entry mark interval**

Obraz zawierający tekst, zrzut ekranu, oprogramowanie, Ikona komputerowa

Opis wygenerowany automatycznie

* **Rate of students by gender**

In VIKO only 57 out of 475 students are women. The target value of this measure is for women to make up at least 40% of all students. To achive this goal university should try to attract women to choose technical subjects.

**Figure 4. Number of students by gender**

Obraz zawierający tekst, zrzut ekranu, oprogramowanie, Ikona komputerowa

Opis wygenerowany automatycznie

* 1. **OLAP analysis and Interpretation**

**Questions**

1. **Which subjects and their groups influence the final grading the most?**

Obraz zawierający tekst, oprogramowanie, Oprogramowanie multimedialne, Ikona komputerowa

Opis wygenerowany automatycznie

Based on the graph showing the number of credits for each group of subjects, we can see that the greatest impact on the final grade has the grade from final project, which has 20 credits. Second most important group of subjects for the final Maintenance of Electronic Equipment Specialization Subject and Telecommunication Equipment Specialization Subject both of them had 5 credits.

The greatest impact on the final grade from subject also had final project (20 credits). Subject with the second highest number of credits were automatics and controllers, English for Specific Purposes, Information administration basics and Mathematics.

For every study program the subject and subject group with the greatest impact was final project with the 20 credits. But the number of credits even in subjects with the same name may vary from department to department.

Subject group with highest amount of credits(despite the final project):

* For Computer Engineering
  + Subject group: Professional activities practices (3.33 credits)
  + subject: final practice (7 credits)
* For Computer Programming
  + Subject group: Professional activities practices (3.33 credits)
  + Subject: Final practice (7 credits)
* For Electronic Engineering
  + Subject group: Maintenance of Electronic Equipment Specialization Subjects (5 credits) Telecommunications Equipment Specialization Subject (5 credits)
  + Subject: Professional Practice (7 credits)
* For Information Systems
  + Subject group: Information Technology Administration Specialization Subject (4 credits)
  + Subject: information administration basics (6 credits), Information Technology (6 credits), Mathematics (6 credits)
* For Telecommunication
  + Subject group: Data transmission Technology Specialization Subjects (4 credits), Digital Network Technology Specialization Subjects (4 credits)
  + Subject: Mathematics (6 credits)

1. **What is the percentage of students finishing their studies in VIKO EIF the year they were supposed to do that?**

Obraz zawierający tekst, zrzut ekranu, oprogramowanie, Ikona komputerowa

Opis wygenerowany automatycznie

In the analyzed period 86,53 percent of students, which is equal to 411 students, completed their study program in three years. 32 students (6,74%) finished in 4 years and 6 students finished in 5 years.

Students who finished the study program in 2 years had the highest average of final grade equals 15,37. The analysis shows that longer studies last the lower average students had.

The graphs show that people who studied 4 years tended to have lower entry mark but difference between final grade average between students who studied 3 and 4 years was small.

1. **Does having academic failure influence the final mark and how?**

To answer this question we need e.g. information about number of exam each student failed or how many times he tried.

1. **Does having academic vacation influence the final mark and how?**

To answer this question we need data about academic vacation and e.g. national holidays.

1. **What is the correlation between the student’s entry mark and his studies, as well as his final project’s grading?**

Obraz zawierający tekst, zrzut ekranu, oprogramowanie, Ikona komputerowa

Opis wygenerowany automatycznie

The graphs show that students who had the entry mark in the range 18-20 had the highest final grade. Second group with where people who had entry mark in the range 14-16.

Students with the highest entry mark chose Software Department. The second choice was Electronic Department. The Study program they chose first was Computer programming), second one was Electronic Engineering.

Specializations with the students with the highest entry mark was Database Systems (15,74), Internet Technology (15,72), Programming and Computer Systems (15,69) and Computer Maintenance And Networks Administration (15,68)

The final project grade range was 13.70-16.96 and the average equals 15.15. 449 students graduated with a good results i.e. a grade in the range 14-16. Only 26 had grades between 16-18 and no one had the result above 18.

1. **What is the correlation between the study results and student’s background?**

Obraz zawierający tekst, oprogramowanie, Ikona komputerowa, Oprogramowanie multimedialne

Opis wygenerowany automatycznie

Student’s background can be represented by many measures, which include their place of birth, region they are from or the their entry mark. Taking into account entry mark, we can see that on average, the students with higher entry mark had higher final grades. However, difference regarding average final grade between students with the lowest and highest entry mark was only 0,14, which means very weak correlation between these two measures.

The highest average of final grade was noted among students that used to go to school in Kelme, but the difference with other cities is not significant. Therefore, we can assume that the city of the school that students used to go to does not have much of an effect on the later performance either.

As shown on the map, the best highest average final grade was noted among students from the county of Siauliu. County with the lowest average was Utenos with the average of 14,88. Based on the difference between these two regions we can say that students from Siauliu were better than the students from Utenos by 0,65 on average.

Taking place of birth as a measure of student’s background we can see there is one outstanding region named Kelme, from which students got the highest average final grade of 16,32. The worst in terms of average final grade of the students regions can be seen in the eastern and the border regions of the country.

1. **Does coming from different kind of complementary education influences studies in VIKO EIF?**

Obraz zawierający tekst, oprogramowanie, Ikona komputerowa, Oprogramowanie multimedialne

Opis wygenerowany automatycznie

Taking into account the average of final grade of the students, it should be noted that students that used to go to Lyceum before college had the best study results. What is interesting, the worst group in terms of this measure were students that went to private schools. However, they were only a small percentage of all students and their number was too low to draw any conclusions. In terms of number of years to finish the study programme, it may be seen that the students from secondary schools and gymnasiums were the worst, but because they were also by far the largest group, we should not make any conclusions based on this chart.

1. **Which departments, study programmes and specializations in them have better students and what is the reason for that?**

Obraz zawierający tekst, oprogramowanie, Ikona komputerowa, Oprogramowanie multimedialne

Opis wygenerowany automatycznie

Obraz zawierający tekst, zrzut ekranu, oprogramowanie, Ikona komputerowa

Opis wygenerowany automatycznie

The VIKO University consists of 4 Department, 5 Study Programs and 12 specialization.

Based on the average of final grade the Electronic Department had the best students, but Students from Software Department had the higher average of entry mark.

Study program with the best students was the Electronic Engineering, even though Computer programming students entry mark was higher by 0.4.

From specializations the Maintenance of electronic equipment had the best students based on the average of final grade and Computer Maintenance and Networks administration had the best students based on entry mark.

Both the specialization and the study program with the students with the highest average of the final grade were part of the electronics department.

1. **Does student’s personal data, such as age or gender has any influence to the final project grading?**

Obraz zawierający tekst, oprogramowanie, Ikona komputerowa, Oprogramowanie multimedialne

Opis wygenerowany automatycznie

By analyzing the influence of gender on the final project grading, it can be concluded that this factor is not statistically significant. The difference between the average of women and men was 0,18 in favor of men. However, the numbers of males and females differed significantly, so no conclusions should be drawn from this. Regarding age, we do not have any clear correlation either. The only significant trend that can be seen is among students aged 32, but it is probably caused by a very small number of students in this age, who happened to have a slightly worse study results than their younger peers

1. **Do students, who had any other higher education diploma prior to VIKO EIF study better?**

To answer this question we need information about other higher education diplomas.

1. **Do students who didn’t enter VIKO EIF immediately after finishing their school due to some reasons have additional problems in studies?**

To answer this question we need information about whether students had any break between finishing school and entering college.

1. **What is the correlation between the studies and the form of studies?**

To answer this question we need information about form of studies.

1. **Conclusions**

* The rate of the students who finished their study programme in 3 years was 86,53% which is equal to 411 students. 32 (6,74% of the students) finished in 4 years and only 6 of them in 5 years.
* The highest final grade had the students whose entry mark was in range 18-20. It’s worth to mention that second group of students in terms of the final grade were students with the entry mark between 14-16.
* The correlation between students’ background and their study results is not clear. The school that the student used to go to before college did not have significant effect on the final grade. The highest impact on the study results had the place of birth, because the students from the eastern and the border regions of the country had worse study results than students from such regions as Kelme.
* In terms of the impact of the complementary education on the studies in VIKO EIF we cannot make any conclusions. Although the students from Lyceum had slightly better results on average, the differences between number of students for each group was too big to compare them. an analysis of the effects of age and gender on school performance also did not lead to any specific conclusions
* It’s hard to clearly decide which department had the best students because students from the electronic department had the highest final grade and students from software departments had the highest entry mark.

1. **References**

* <https://learn.microsoft.com/en-us/dax/sumx-function-dax>
* <https://learn.microsoft.com/en-us/dax/>