**Project Documentation**

Team: Yessengaziyev Darkhan

Torebek Yerassyl

**Middle, secondary schools in Kazakhstan regions**

**Introduction**:

The study of the school system in the regions of Kazakhstan is a thorough analysis of key aspects of the educational process in the country. Education plays a vital role in shaping society, and its quality and effectiveness largely determine the future of the country. This project is aimed at studying the number of schools in the regions of Kazakhstan in order to identify weaknesses, as well as provide recommendations for its improvement.

In 1999, Kazakhstan became the first nation in Central Asia to implement free and obligatory preschool education (UNESCO, 2006). Following the global agenda, which outlines a roadmap for a better and sustainable future for all, Kazakhstan, like many other nations, is dedicated to accomplishing SDG4 "Quality of Education" (UN, 2015b) and hopes to enroll all children between the ages of three and six at 100% by 2030. There is a lack of research supporting current government policies and their implementation in schools, especially ECE settings, despite the state prioritizing and investing in free and mandatory education.

**Goal**:

The purpose of this study is to show the role played by the construction of schools and their main distribution across the country. Since, in the regions of our country, you can see the actual problems with schools. Moreover, given the rapid growth of the birth rate, one can clearly see the coming problem.

**Tasks**:

Study and analysis of the structure of the school system in different regions of Kazakhstan.

Assessment of the quality of education and the level of student achievement in various regions.

Identification of factors affecting the effectiveness of the school system in the regions of Kazakhstan.

Try to visualize the data and make some opinion

**Methodology**:

Data collection: Collect information on the structure, curricula, resources and performance indicators of schools in various regions of Kazakhstan using official reports.

Analysis: Analyze the collected data to assess the quality of education, student achievement, and factors affecting the effectiveness of the school system.

Comparative Study: Compare the results across different regions to identify patterns and differences

Recommendations: Develop recommendations based on the results of the study to address deficiencies and improve the overall quality and effectiveness of the school system in each region.

An important step is a comparative study of specific examples to develop several options in order to gain an in-depth understanding of the strengths and challenges of the education system.

Expected results:

Comprehensive understanding of the strengths and weaknesses of the school system in various regions of Kazakhstan.

Identification of factors contributing to differences in learning outcomes in different regions.

Development of targeted recommendations to improve the quality and effectiveness of the school system in each region.

Contribution to the development of sound policy and strategic planning for the development of education at both the regional and national levels.

**Project progress**:

1. **Data Selection**

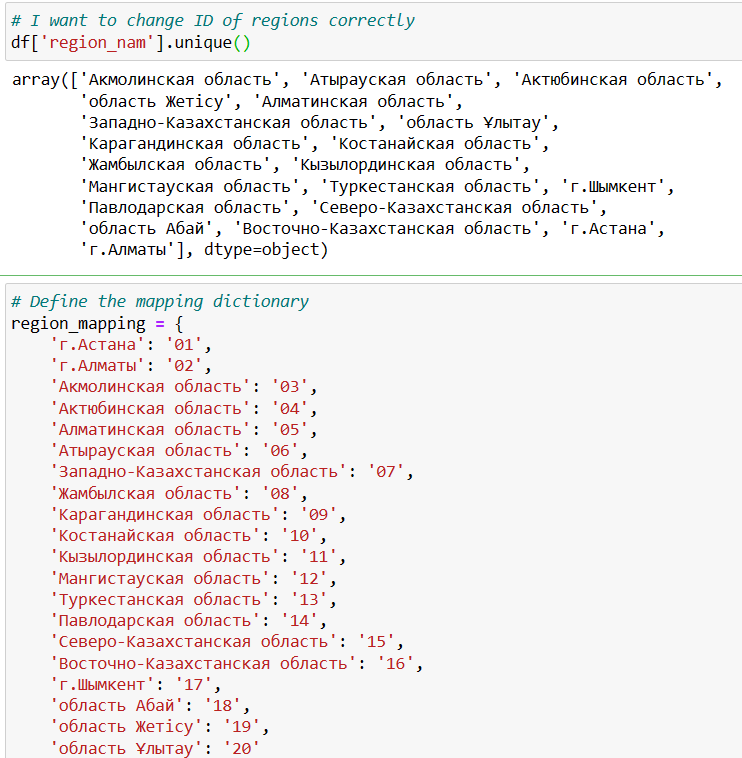
To begin with, we carried out extensive work on searching for a truthful database. We took the database from the official website <https://data.egov.kz/> where we can see information about 20 regions of Kazakhstan.

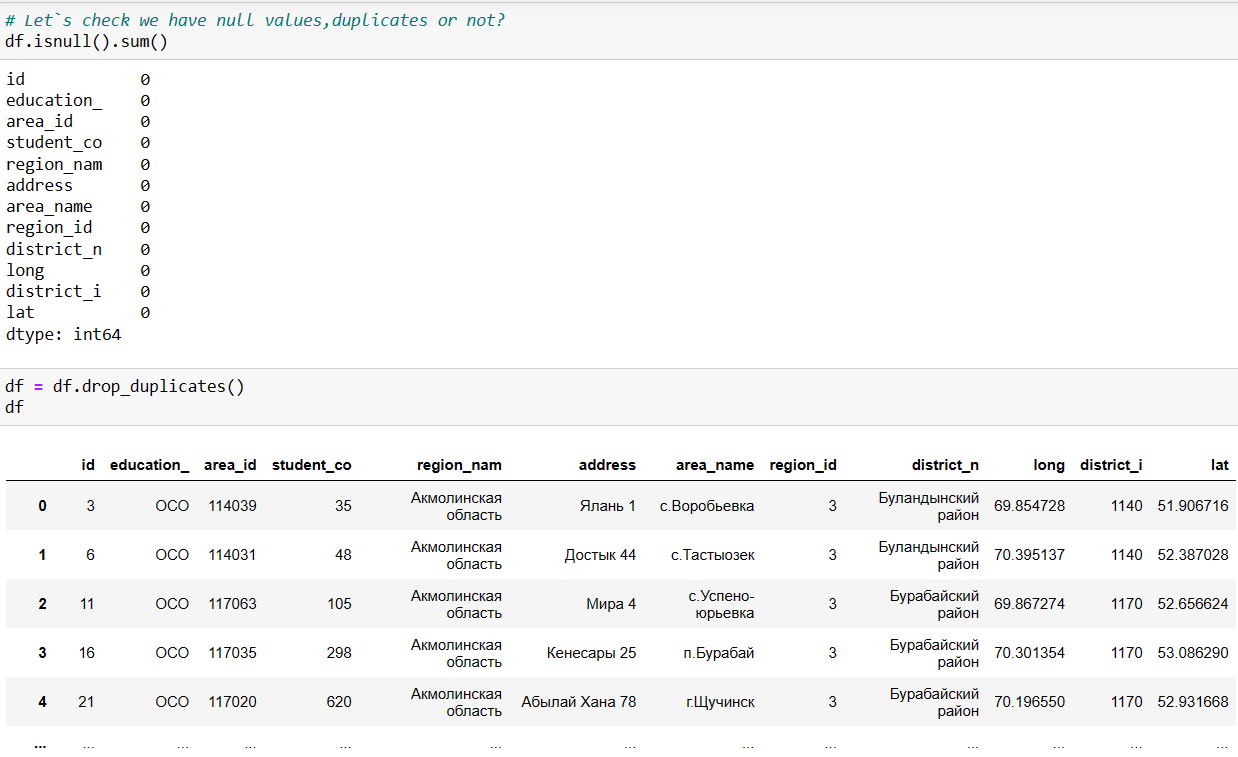
The structure of the table is simple, we have 7094 rows and 15 columns before cleaning the data.

**Parsing process.**

* Firstly, we initiated a web scraping process by sending a GET request to the website https://data.egov.kz/ to retrieve its HTML content. Using API Key, we parsed the HTML and identified specific elements containing data about regions in Kazakhstan. These data were then extracted and stored in a structured format. Subsequently, we determined that there are 20 regions in Kazakhstan. Upon completion, we successfully obtained data for 7094 rows and 15 columns.
* Focusing on the core data acquisition phase, we collected various information regarding each region, including statistical indicators such as population, area, and demographic characteristics. Despite the extensive amount of data available on the website, we deliberately limited our extraction to the most relevant information pertaining to our analysis. Post-scraping, our dataset comprised 7094 rows and 12 columns, providing a comprehensive overview of the regions of Kazakhstan.
* Here are the lines of code demonstrating our approach to importing and cleaning the dataset for further analysis.







After gathering all the necessary data from the website https://data.egov.kz/, the next step in our project is data cleaning, a crucial phase in our analysis pipeline. This process involves several steps to ensure the quality and consistency of our dataset.

To achieve these objectives, we utilized a mapping dictionary to convert regional names into standardized codes. Each region name was mapped to a corresponding code, ensuring consistency and accuracy in our dataset.

Furthermore, we performed data type conversions and handled null values. By inspecting unique values in each column, we identified and corrected data types to accurately represent the information. For instance, we converted string values to integer or float where appropriate, such as converting IDs and numerical data. Additionally, we addressed null values and ensured that our dataset was appropriately formatted and ready for further analysis.

Overall, the data cleaning process plays a crucial role in preparing our dataset for robust analysis, ensuring that our findings are accurate, reliable, and aligned with the objectives of our project.

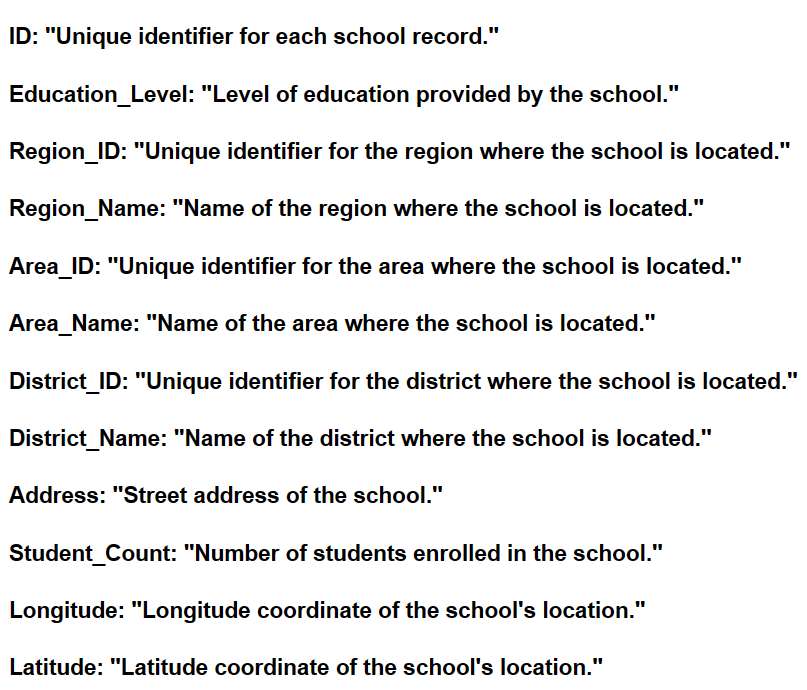
**Data Wrangling**

The first important step in our data wrangling process was to enhance the dataset by adding a new column named "Education\_Level", which categorizes the educational level of each entry.

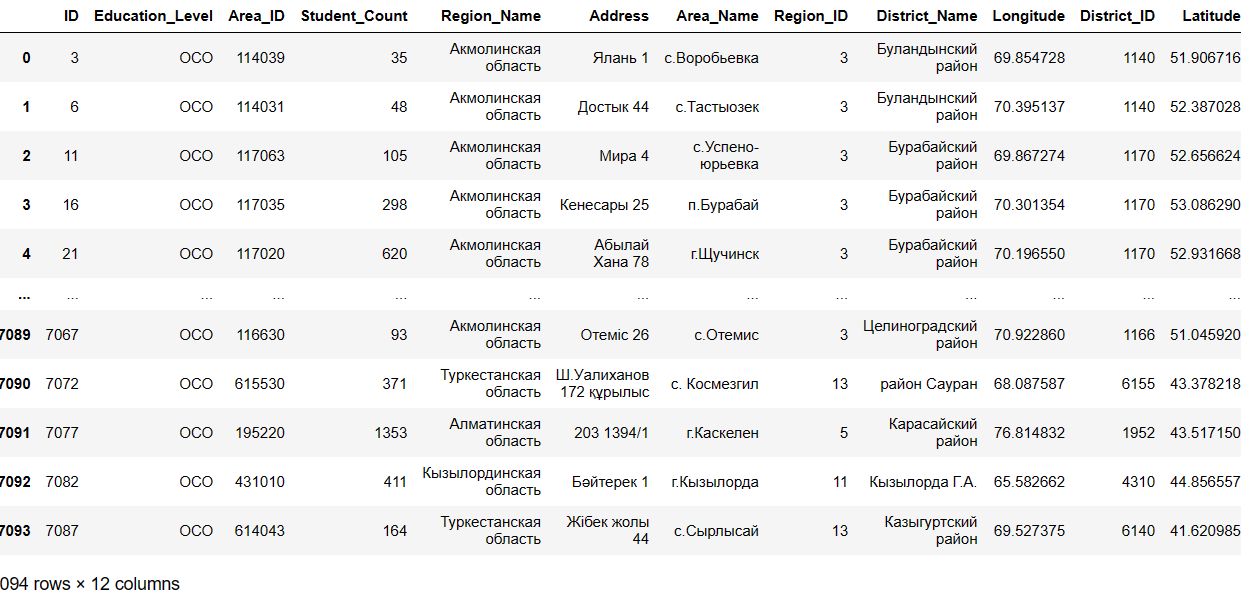
Next, we performed calculations to derive average values for various metrics related to our dataset, such as student count and geographical information. Specifically, we computed the average values for each combination of these metrics. Then, for each group, we calculated the mean value of the respective metric. These calculated average values were then assigned to new columns in the DataFrame. This process enables us to obtain average values for specific locations and other relevant metrics, facilitating further analysis and interpretation.

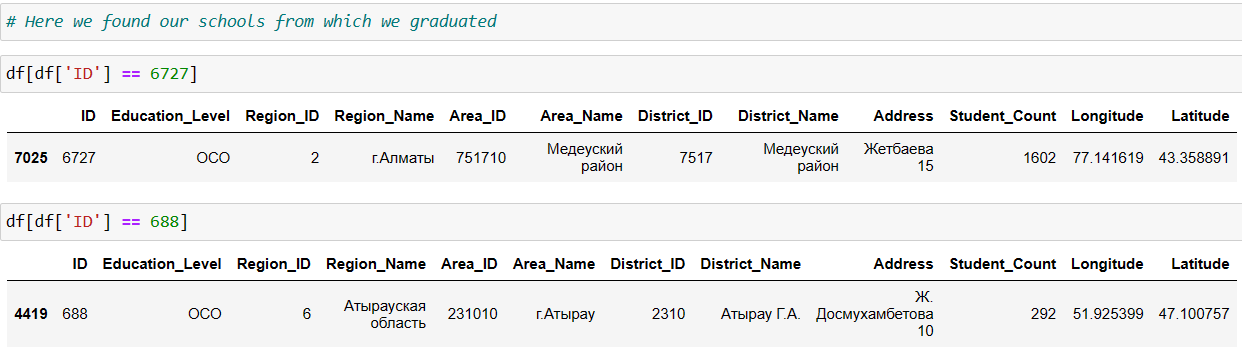
To streamline our dataset and focus on pertinent information, we initially dropped unnecessary columns. These columns, including 'sdu\_load\_i', 'geo\_coordi', and 'establish\_', were identified as not needed for our analysis and were consequently removed from the DataFrame.

Additionally, we renamed our columns to improve clarity and consistency within the dataset. Each column was assigned a more descriptive label to accurately represent its contents, such as 'ID,' 'Education\_Level,' 'Area\_ID,' 'Student\_Count,' 'Region\_Name,' 'Address,' 'Area\_Name,' 'Region\_ID,' 'District\_Name,' 'Longitude,' and 'Latitude.'



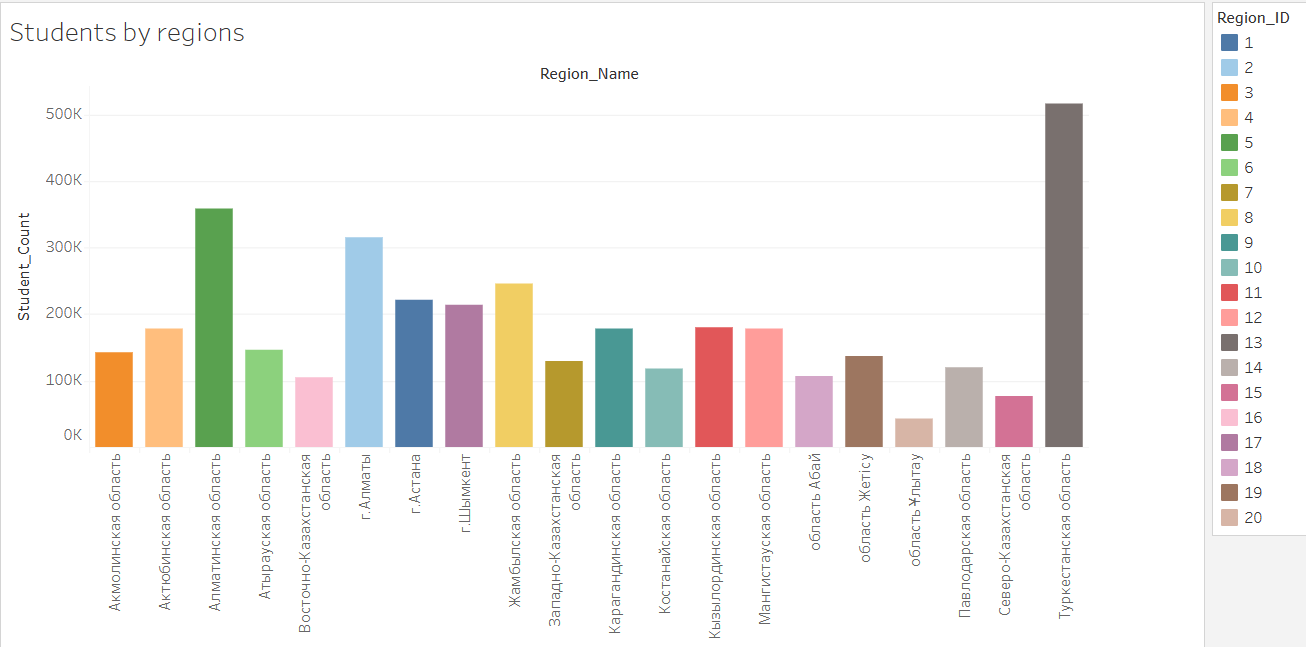
Our dataset, after these steps, contains insightful information ready for further analysis and exploration.

**Final result looks like this:**  


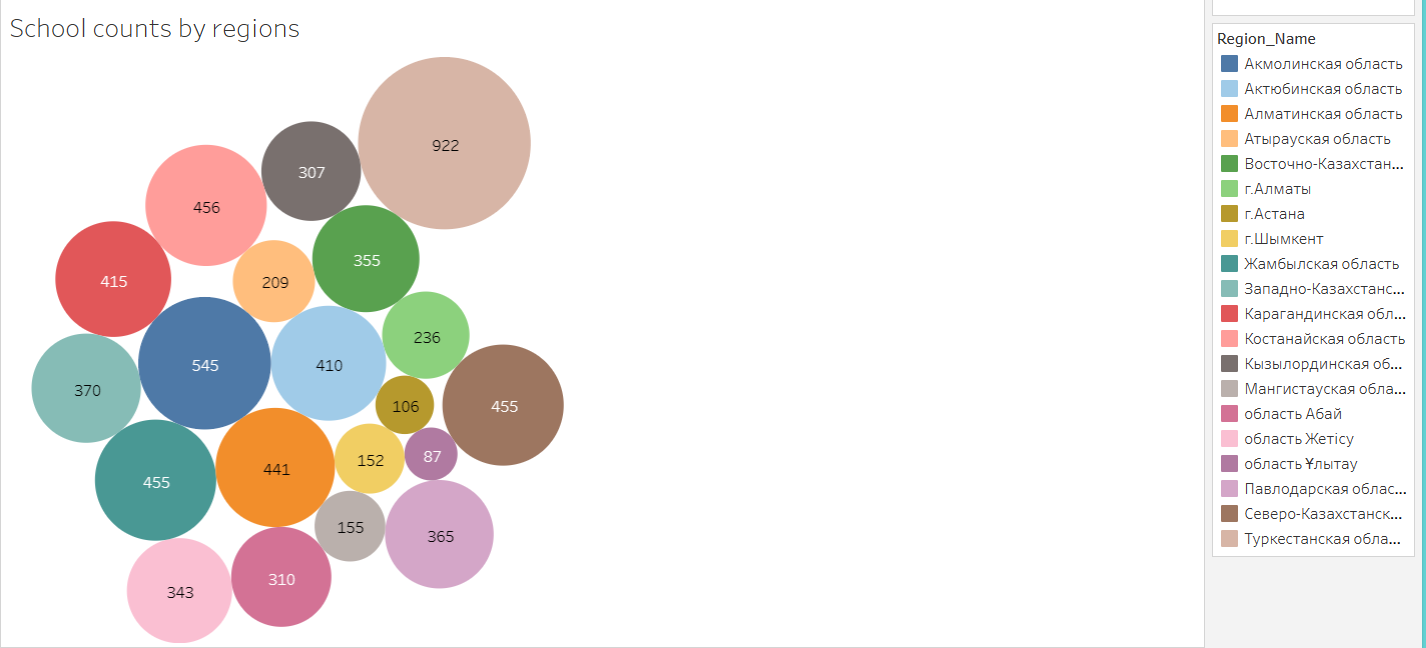
****

**Visualization of DataFrame**

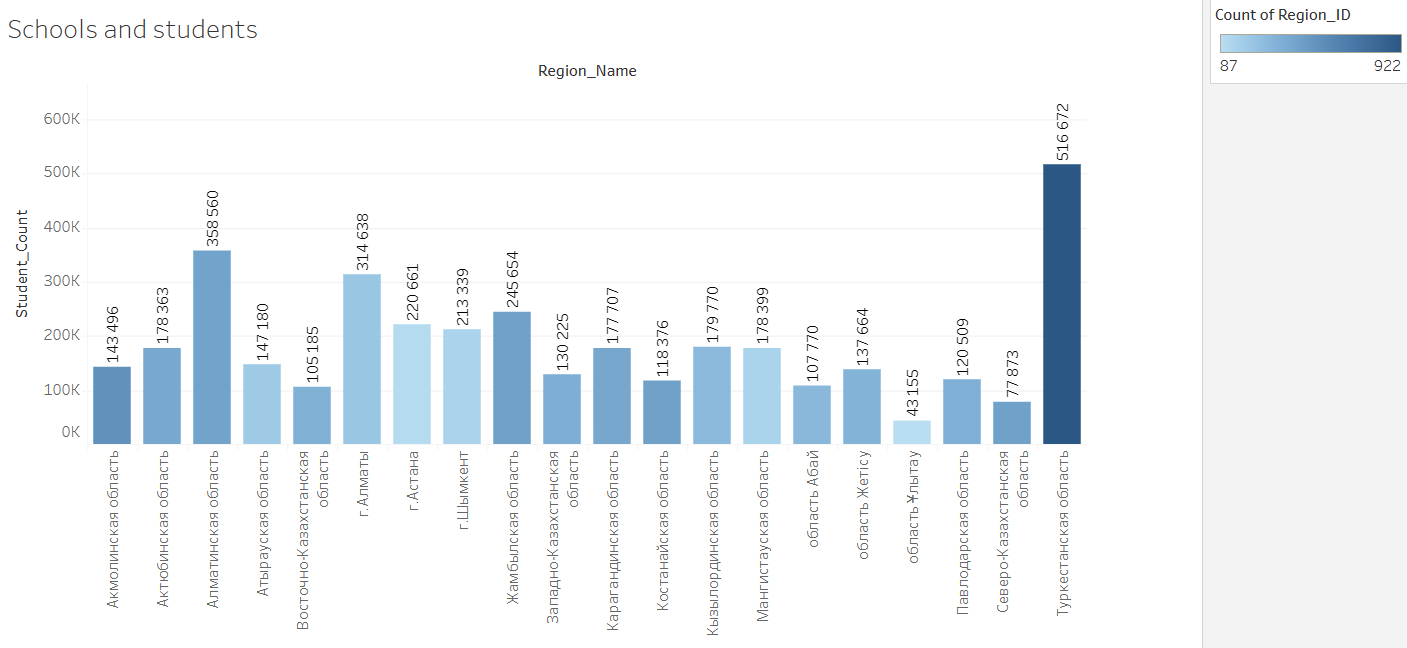
In order to visually show our visualization, the Tableau program was used.



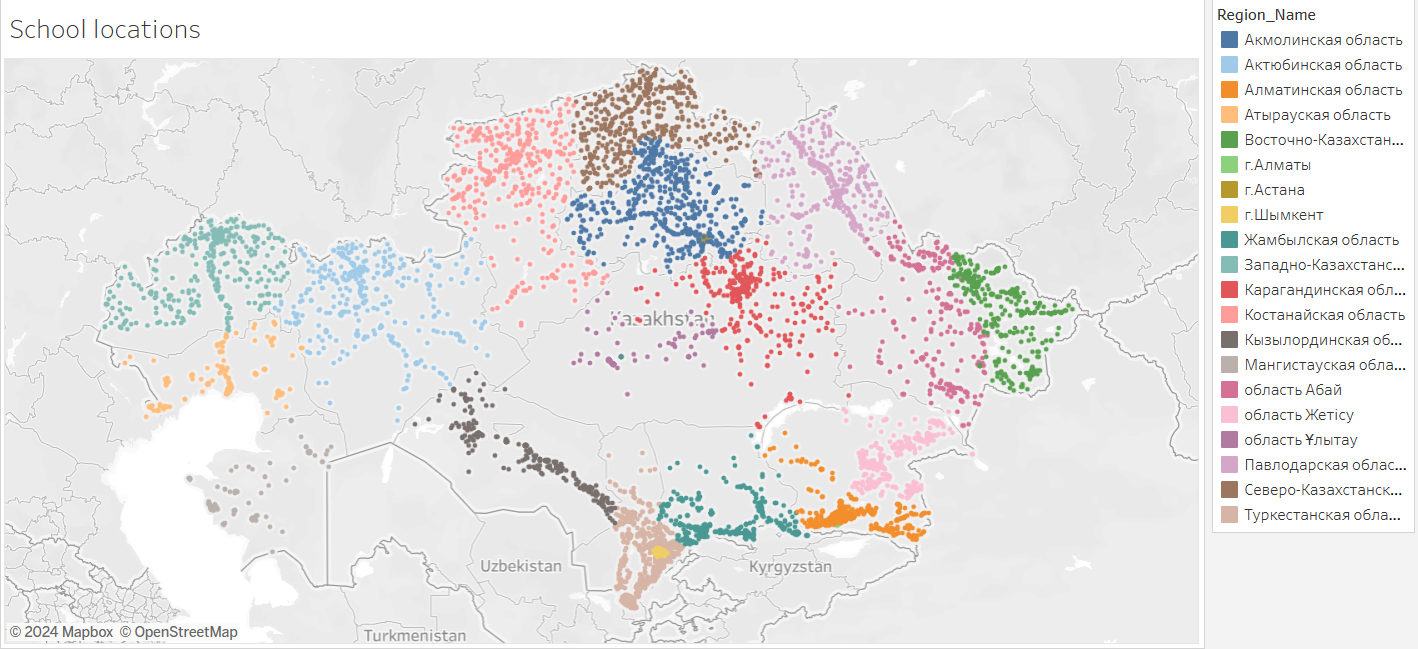
Here we can see the number of students in each region of Kazakhstan. As you can see, most of the students are in the Turkestan region, and least of all in the Ulytau region. This will help us pay more attention to the spread of students by region



The number of schools in each region is also indicated here. Given the number of students in the Turkestan region, there are many schools, respectively. There are also not many students in the East Kazakhstan region, but there are not a few schools.

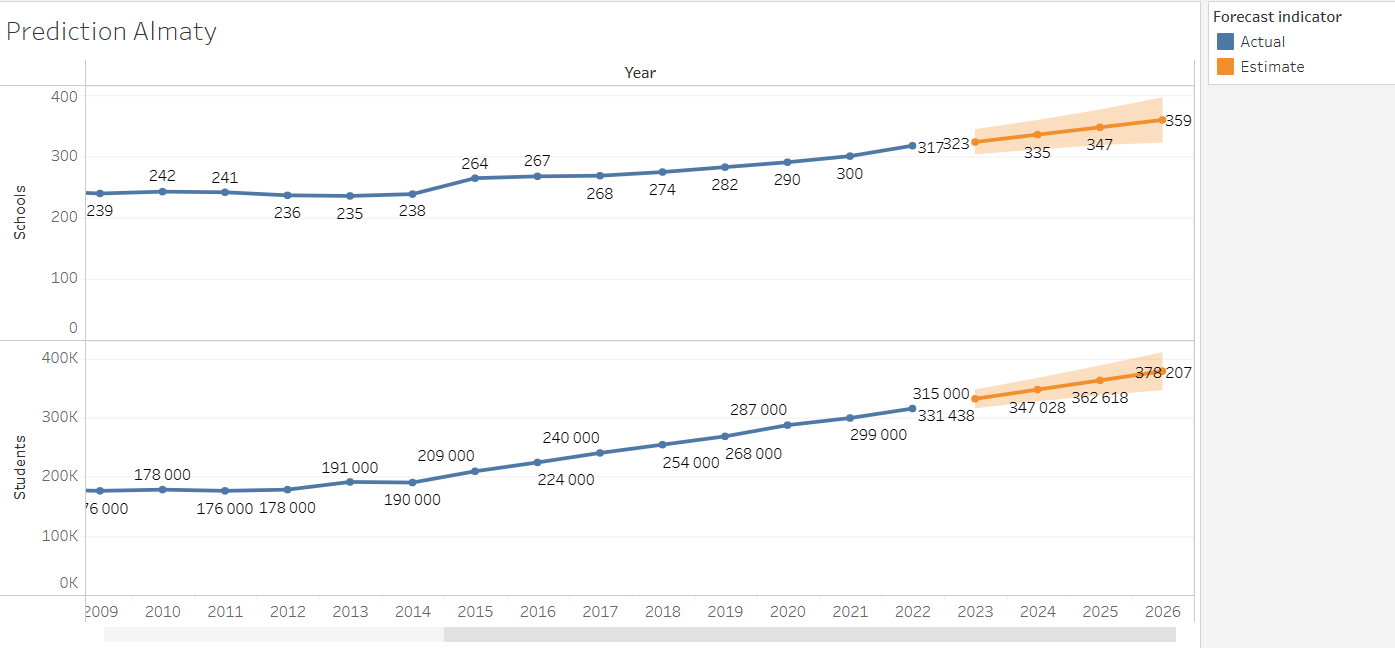


Here we can see the dependence of the number of students on the number of schools. The darker the pillar, the more schools there are, and here you can clearly see in which region there is a shortage of schools or an overabundance of them in relation to the number of students



Due to the fact that we have Latitude and Longitude data, we were able to display the location of each school by country. In the right part you can see the color of each region. On our map, we can see that there are fewer schools in the west of the country, and most in the south and north





Here we have made a small forecast of schools and students in the next two years in the city of Almaty. Why Almaty? This city was chosen according to several criteria. Such as the economy, population and priority

**Recommendations**:

What can we say in the end, schools are an integral part of the country's development. The school is the foundation of a prosperous state. But as you can see from the results of the study, there are not enough schools in our country, given the increase in the birth rate of children. This is most noticeable in the west of the country. Our country is also known for its achievements in various international Olympiads, which tells us that even under such conditions, children can find the very source of knowledge.

First, we would like to recommend that our country start building more schools. Using the example of Finland, Norway and Sweden, one can see that private schools are unpopular in these countries, where public schools are preferred more.

Secondly, allocate more budget for the development of education and science. In Germany, 94 trillion tenge was spent on education in 2023, and 1.85 trillion tenge was allocated in Kazakhstan.

**Conclusion**:

The study of the school system in the regions of Kazakhstan is crucial to ensure equal access to quality education and promote socio-economic development. Analyzing the key aspects of the provision of education and its results, we saw what problems our country may face in the future. This study is also intended to provide valuable information and recommendations to improve the effectiveness of the school system across the country.

**References:**

1. [Teach for Qazaqstan: еліміздің аймақтарындағы балалардың әлеуетін ашыңыз](https://teachforqazaqstan.org/)
2. <https://www.scopus.com/record/display.uri?eid=2-s2.0-85175306237&origin=resultslist&sort=plf-f&src=s&sid=09dd5f26ce7ec49ae3979b0eab602643&sot=b&sdt=b&cluster=scosubjabbr%2C%22ENVI%22%2Ct%2C%22SOCI%22%2Ct%2C%22EART%22%2Ct%2C%22PHYS%22%2Ct%2C%22COMP%22%2Ct%2C%22MATE%22%2Ct%2C%22ARTS%22%2Ct%2C%22BIOC%22%2Ct%2C%22MULT%22%2Ct%2C%22DECI%22%2Ct&s=TITLE-ABS-KEY%28educational+AND+reform+kazakhstan%29&sl=32&sessionSearchId=09dd5f26ce7ec49ae3979b0eab602643&relpos=4>
3. <https://www.scopus.com/record/display.uri?eid=2-s2.0-85187132225&origin=resultslist&sort=plf-f&src=s&sid=09dd5f26ce7ec49ae3979b0eab602643&sot=b&sdt=b&cluster=scosubjabbr%2C%22ENVI%22%2Ct%2C%22SOCI%22%2Ct%2C%22EART%22%2Ct%2C%22PHYS%22%2Ct%2C%22COMP%22%2Ct%2C%22MATE%22%2Ct%2C%22ARTS%22%2Ct%2C%22BIOC%22%2Ct%2C%22MULT%22%2Ct%2C%22DECI%22%2Ct&s=TITLE-ABS-KEY%28educational+AND+kazakhstan+school%29&sl=32&sessionSearchId=09dd5f26ce7ec49ae3979b0eab602643&relpos=9>
4. <https://www.scopus.com/record/display.uri?eid=2-s2.0-85173963105&origin=resultslist&sort=plf-f&src=s&sid=09dd5f26ce7ec49ae3979b0eab602643&sot=b&sdt=b&cluster=scosubjabbr%2C%22ENVI%22%2Ct%2C%22SOCI%22%2Ct%2C%22EART%22%2Ct%2C%22PHYS%22%2Ct%2C%22COMP%22%2Ct%2C%22MATE%22%2Ct%2C%22ARTS%22%2Ct%2C%22BIOC%22%2Ct%2C%22MULT%22%2Ct%2C%22DECI%22%2Ct&s=TITLE-ABS-KEY%28educational+AND+kazakhstan+school%29&sl=32&sessionSearchId=09dd5f26ce7ec49ae3979b0eab602643&relpos=28>
5. <https://www.scopus.com/record/display.uri?eid=2-s2.0-85066448481&origin=resultslist&sort=plf-f&src=s&sid=09dd5f26ce7ec49ae3979b0eab602643&sot=b&sdt=b&cluster=scosubjabbr%2C%22ENVI%22%2Ct%2C%22SOCI%22%2Ct%2C%22EART%22%2Ct%2C%22PHYS%22%2Ct%2C%22COMP%22%2Ct%2C%22MATE%22%2Ct%2C%22ARTS%22%2Ct%2C%22BIOC%22%2Ct%2C%22MULT%22%2Ct%2C%22DECI%22%2Ct&s=TITLE-ABS-KEY%28educational+AND+kazakhstan+AND+middle+school%29&sl=32&sessionSearchId=09dd5f26ce7ec49ae3979b0eab602643&relpos=5>
6. <https://www.scopus.com/record/display.uri?eid=2-s2.0-85179061412&origin=resultslist&sort=plf-f&src=s&sid=09dd5f26ce7ec49ae3979b0eab602643&sot=b&sdt=b&cluster=scosubjabbr%2C%22ENVI%22%2Ct%2C%22SOCI%22%2Ct%2C%22EART%22%2Ct%2C%22PHYS%22%2Ct%2C%22COMP%22%2Ct%2C%22MATE%22%2Ct%2C%22ARTS%22%2Ct%2C%22BIOC%22%2Ct%2C%22MULT%22%2Ct%2C%22DECI%22%2Ct&s=TITLE-ABS-KEY%28education+AND+kazakhstan+AND+school%29&sl=32&sessionSearchId=09dd5f26ce7ec49ae3979b0eab602643&relpos=50>