**Lab4 report**

**Data Sources**

The data for the analysis were taken from the official statistics of the regions, which include the following parameters:

* Number of students who completed grade 11.
* Number of students who completed grade 9.
* Percentage of students who continued their education after grade 9.
* Poverty levels across regions.

**Data Preprocessing**

In the course of data preprocessing, the following steps were performed:

1. **Handling Missing Values**: Missing values in the columns were replaced with the median value to eliminate data distortions.
2. **Data Normalization**: Data regarding the number of students were normalized to ensure comparability.
3. **Categorical Data Encoding**: The names of the regions were converted into a numerical format for analysis.

**Correlation Analysis**

A correlation analysis was conducted between the following variables:

* Number of students who completed grade 11.
* Number of students who completed grade 9.
* Percentage of students who continued their education after grade 9.
* Poverty level.

The results of the correlation analysis are presented on a heatmap, which visually assesses the degree of relationship between the variables.

**Data Visualization**

As part of the work, graphs and maps were created to visualize the following aspects:

* Distribution of students across regions.
* The relationship between the level of education and poverty.

**Conclusion**

The laboratory work revealed key dependencies between the level of education and the level of poverty in various regions.

*Ospan Ramazan it2-2310, 38045*