

# Project Report — Demography meets Healthcare

## SIS\_DataMinds

**Author: Dilnaz Zeinolla & Zhumanova Zhanel**

### 1. Project Goal

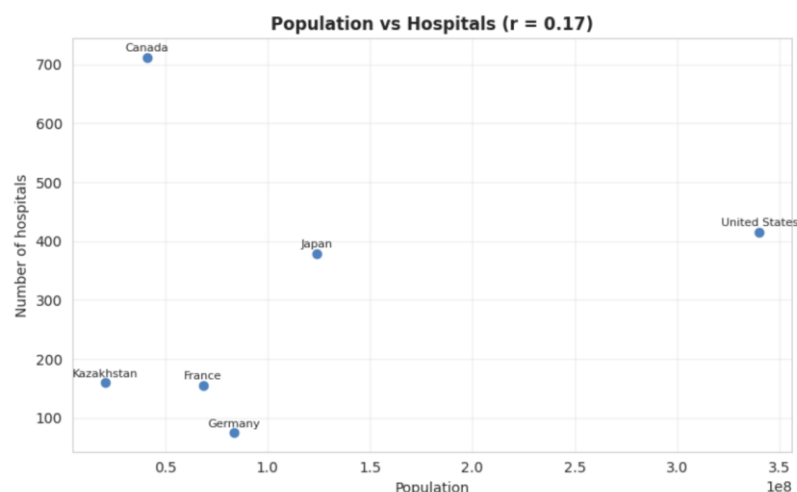
The goal of our project is to make data analysis using the knowledge learned in the lessons and as a topic we took the relationship between the population and the number of hospitals in different countries. We wanted to see if the number of hospitals is growing in proportion to the population or if other factors such as population density, age structure also play an important role. We investigated how hospital infrastructure varies by country and region, and how it relates to population structure - especially 65+ seniors.

### 2. Data Collection & Preparation

In our project, we used two data sources: World Bank API and Wikipedia. We first tapped into the World Bank API to get 2022 population and life expectancy data by country. The data came in JSON format, so we easily converted it to a table using the pandas library. Then, with the help of requests and BeautifulSoup, we made a web scraping of the Wikipedia page with information on the number of hospitals in different countries and compared this data with the names of countries from the API. After collecting the data, we cleaned them up: we removed duplicates and gaps, translated the population into millions and rounded to one decimal place, and normalized the names of countries so that they match in both tables. We then pooled the data from the Country column and added a calculated figure - the number of hospitals per million elderly people.

### 3. Exploratory Data Analysis & Visualization

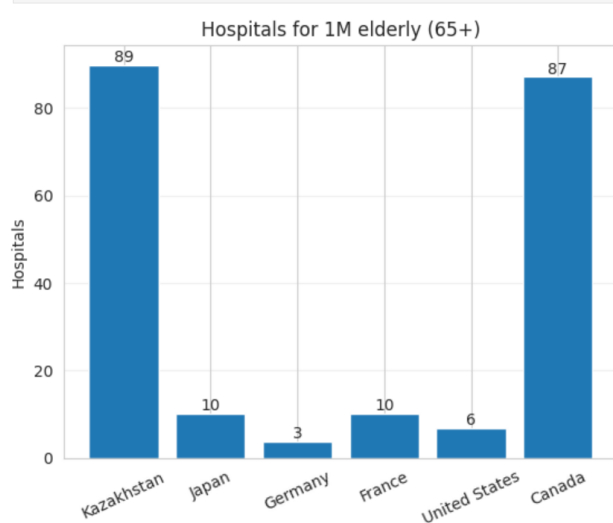
We have done basic EDA and built three main visualizations for our project.



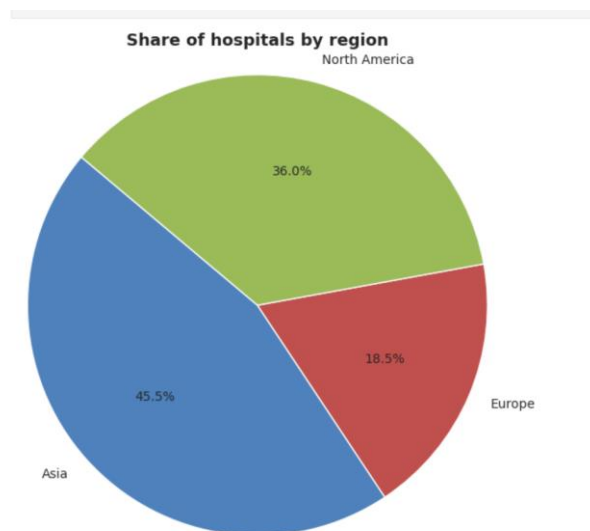
1. Scatter plot:

This graph shows how population size relates to the number of hospitals per country.

## 2.Bar chart: hospitals for 1 million elderly (65+).



This bar chart compares the number of hospitals per one million elderly people across different countries.



## 3. Pie chart:

This pie chart shows the percentage share of hospitals distributed across different regions, illustrating how hospital density varies by region.

## 4. Key Results & Insights

Our analysis showed that the number of hospitals is not proportionally equal to the population size. The number of hospitals does not depend on the population. The correlation coefficient was just 0.17, which shows that the number of hospitals mainly depends on the country's policies and healthcare system.

A key finding is that countries with a higher share of people aged 65 and over face greater pressure on hospitals, even if the total number of hospitals seems adequate. This highlights the importance of healthcare planning focused on the elderly for effective long-term strategies.

Regional differences show Asia has many small hospitals, Europe fewer large centers, and North America a balanced mix. Planning should focus on hospital density per population, not just total numbers, for better healthcare access.

## **5. Conclusion**

In this project, we went through a full cycle: collecting, cleaning, combining and analyzing data. Initially, we wanted to understand how the number of people is related to the number of hospitals and find out the availability of medical institutions. The results showed that the number of hospitals is not directly proportional to the population. Here the level of development of the health care system, the distribution of funding, the share of private clinics can affect. To come to this decision and make the analysis, we awarded all the necessary tools for collecting and preparing data, which we studied in the classroom and thereby show our level of understanding of the process of collecting, preparing and analyzing real data.