Data Analysis and Visualization Course

Coursework title:

Healthcare and Morbidity Statistics

Team Members and their responsibilities

1. Nurali Bakytbek uulu: Data Collection and Cleaning, Data Visualization
2. Bektur Momunov :Data Visualization, Presentation
3. Magomed Mukhammedov : Data Visualization, Report

Project Advisor

Dr. Remudin Mekuria

Objective

The primary objective of this report is to analyze healthcare trends in the Kyrgyz Republic, with a focus on regional disparities in disease burden, healthcare infrastructure, and workforce distribution. By leveraging a comprehensive dataset that spans from 2009 to 2022, this analysis aims to uncover key insights into how healthcare resources and disease prevalence have evolved over time across different regions of the country. Through the use of various data visualization techniques, such as animated bar charts, pie charts, histograms, and scatter plots, this report will highlight significant patterns and trends that are crucial for policymakers, healthcare professionals, and other stakeholders invested in the country's healthcare system.

A secondary objective is to explore the relationships between healthcare resources (such as the number of beds and specialists) and the incidence of key diseases like Tuberculosis, Hepatitis, Diabetes, and Acute Intestinal Infections. By doing so, we aim to identify areas of improvement within the healthcare system and provide data-driven recommendations for resource allocation, disease prevention, and strategic healthcare planning. The findings from this report can also serve as a valuable tool for future public health interventions and long-term policy formulation.

Ultimately, this report seeks to present the current state of healthcare in the Kyrgyz Republic and offer actionable insights that can guide future decisions to enhance public health, reduce disparities, and ensure more equitable access to healthcare resources across the nation’s regions.

Methodology

Data Source: Official government datasets from data.gov.kg and stat.gov.kg.
Aggregated from `med_data.xlsx`
and selected sheets (`Лист1`, `Лист2`, `...`, `Лист25`) in `data_lists.xlsx`.

Period Covered: 2011–2023.

This report relies on data extracted from a comprehensive dataset that covers multiple health indicators across different regions of the Kyrgyz Republic from 2009 to 2022. The dataset includes information on disease incidence, healthcare infrastructure (such as the number of hospital beds), and the distribution of medical specialists across various regions.

Data Collection and Preprocessing:

The primary dataset used in this analysis was sourced from the Kyrgyz Republic's national health department. The data includes multiple disease categories, such as viral hepatitis, tuberculosis, diabetes, drug addiction, and acute intestinal infections. In addition to disease data, the dataset contains variables related to the healthcare system, including the number of medical professionals (both secondary and higher education), hospital beds, and regional population statistics.

Before proceeding with the analysis, data cleaning was performed to ensure consistency and accuracy. This involved standardizing column names, handling missing values, and removing any outlier entries. The data was then filtered to exclude the national-level totals, focusing only on regional data to identify and address regional disparities in healthcare outcomes and resource distribution.

Data Analysis Techniques:

Several techniques were used to analyze the data, focusing on visualizing key trends and relationships. These techniques include:

1. **Descriptive Statistics:** Basic descriptive statistics were computed to summarize the dataset. This included measures like mean, median, and standard deviation to better understand the central tendency and spread of variables like disease incidence, population, and healthcare capacity.

2. **Data Visualization:** A series of interactive and animated visualizations were created using Plotly and other visualization libraries:

    - **Animated Bar Charts:** These charts were used to display the distribution of diseases across different regions over time. The bar charts animate year by year

to reveal how healthcare metrics and disease burden evolve.

- **Pie Charts:** Pie charts were used to show the proportional distribution of diseases within the country, providing insights into the relative burden of different diseases.

- **Scatter Plots and Heatmaps:** These were used to examine the relationship between healthcare resources (e.g., specialists, beds) and disease incidence.

- **Box Plots and Histograms:** These visualizations helped to illustrate the distribution of diseases across regions, revealing variability and potential outliers.

3. **Correlation Analysis:** Correlation matrices were computed to examine the relationships between healthcare resources (e.g., specialists, beds) and disease incidence. This helped identify whether areas with more healthcare professionals or hospital beds experienced fewer cases of certain diseases.

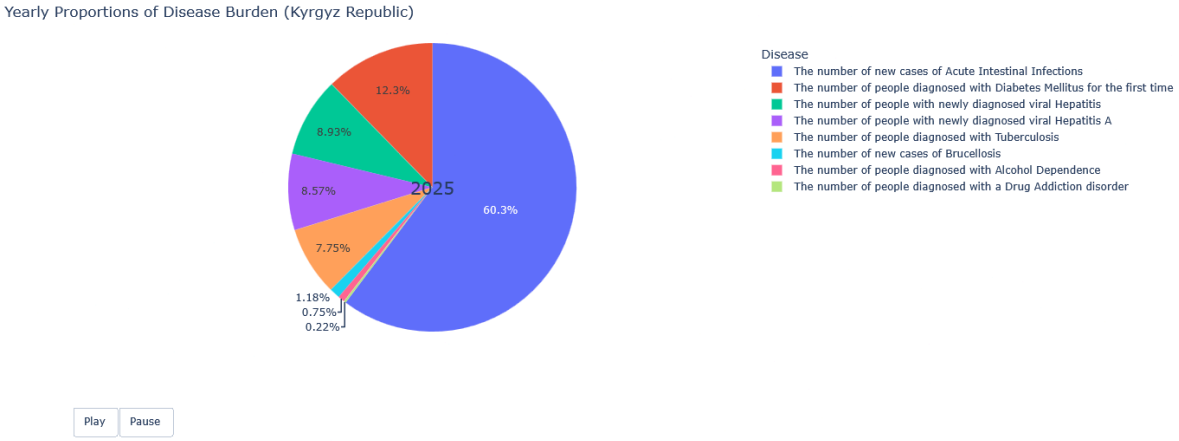
4. **Animation Settings:** To ensure that animations were smooth and informative, the duration of each frame was set to 1 second, with a 0.5-second transition between frames. This ensured that trends were easy to follow, and the visuals were not overly fast or disorienting.

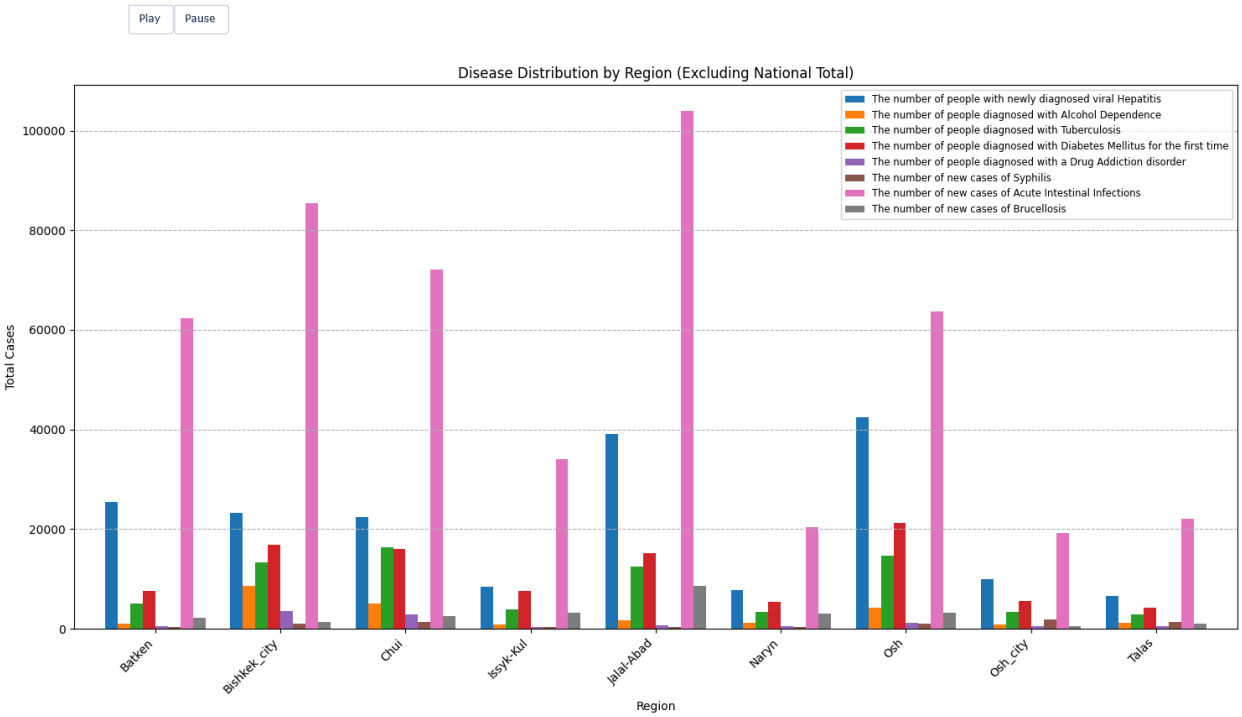
Limitations of Methodology:

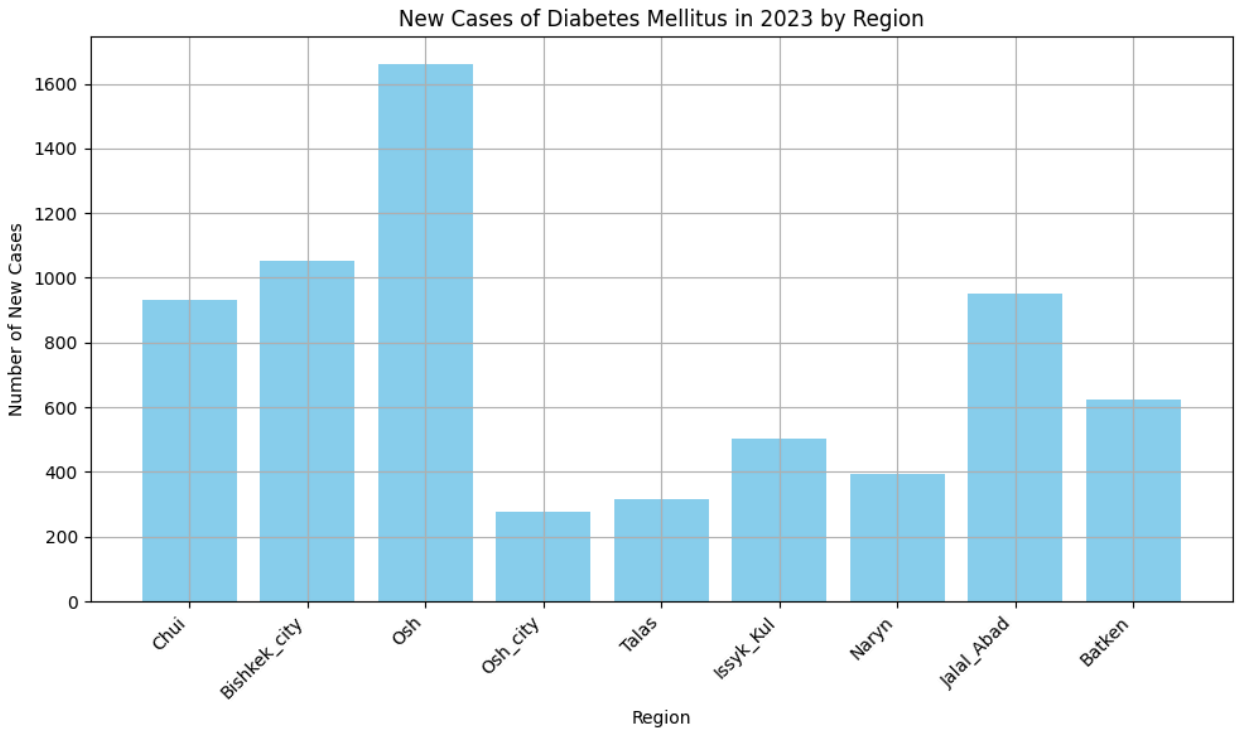
While this methodology provides an in-depth view of the healthcare landscape in the Kyrgyz Republic, several limitations exist. First, the dataset may not capture all factors influencing disease prevalence and healthcare access, such as socioeconomic status, environmental factors, or healthcare policy changes. Additionally, regional variation in data reporting practices could affect the accuracy of the results.

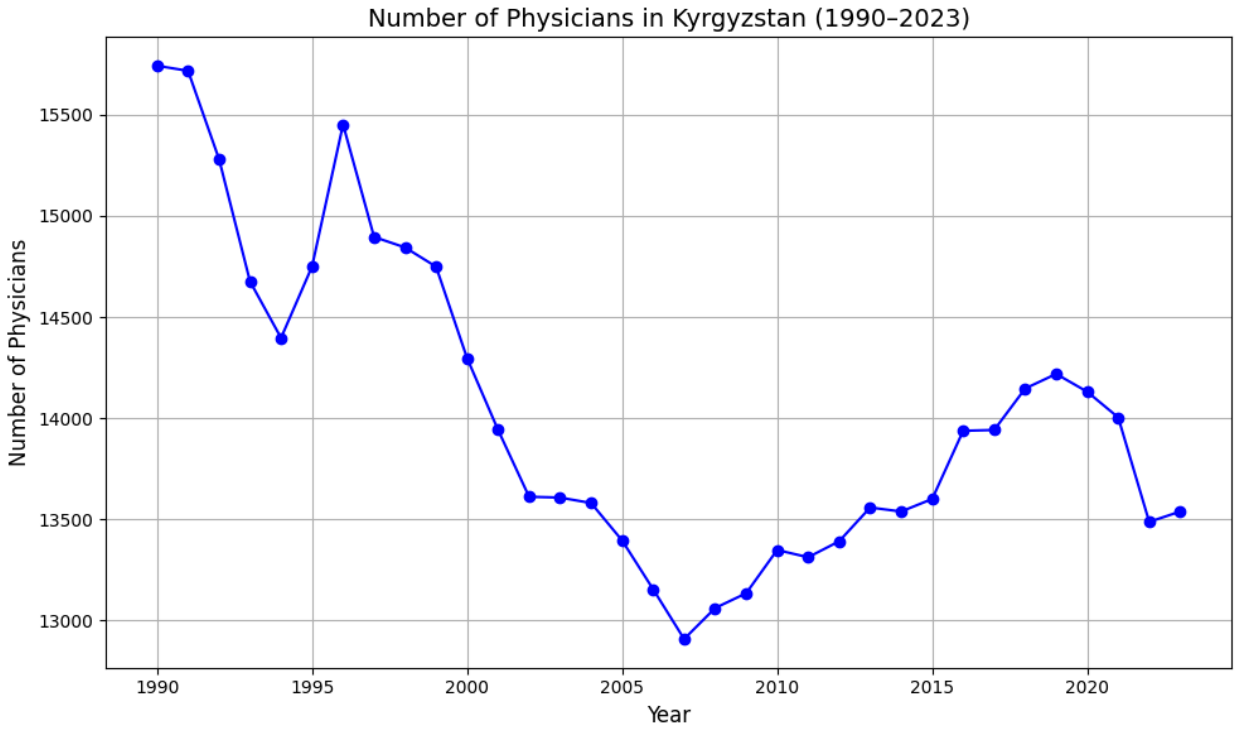
Tools Used: Python (pandas, matplotlib, seaborn), Jupyter Notebook.

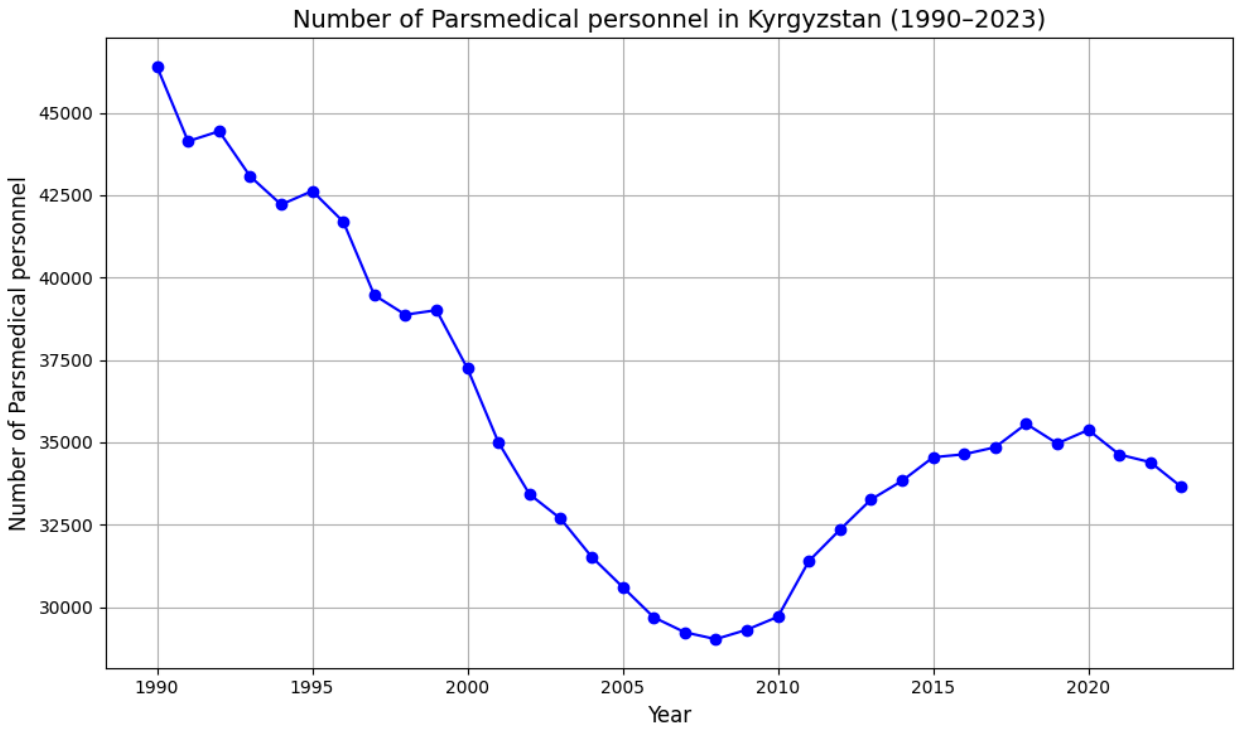
Visualizations:

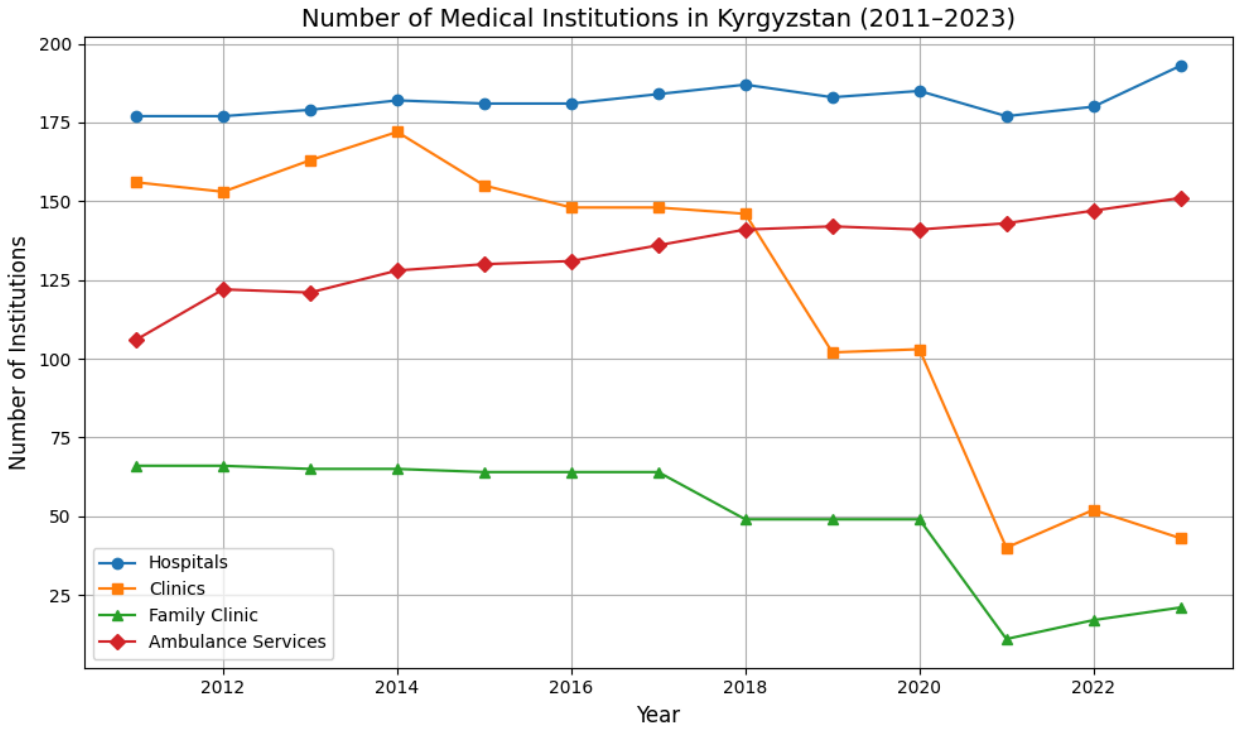


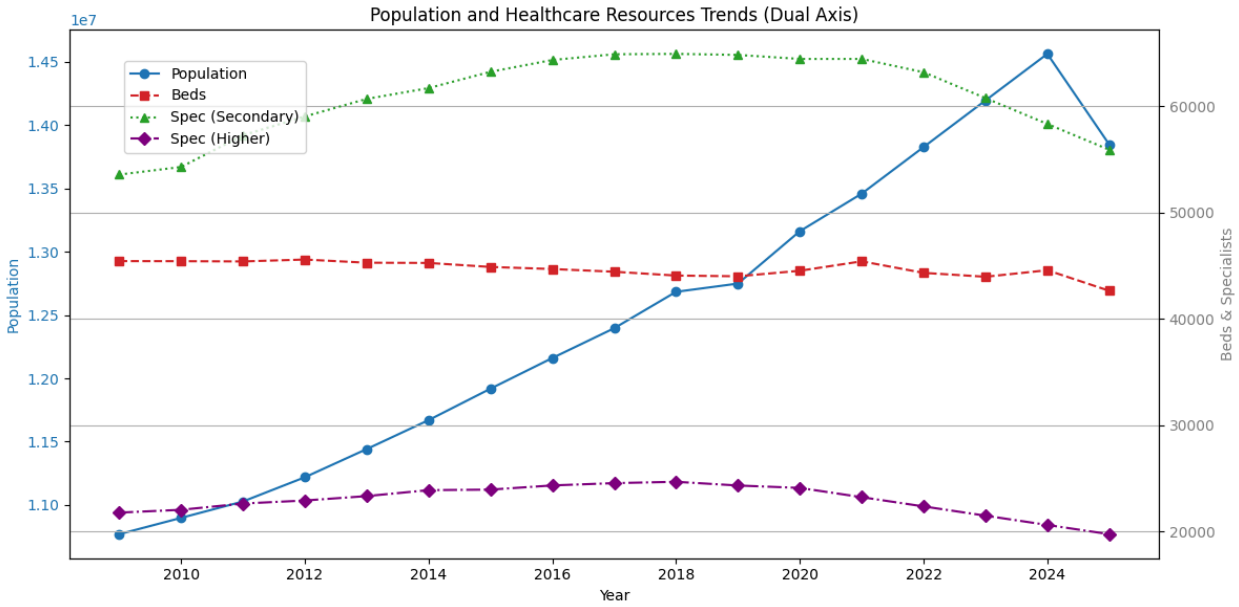


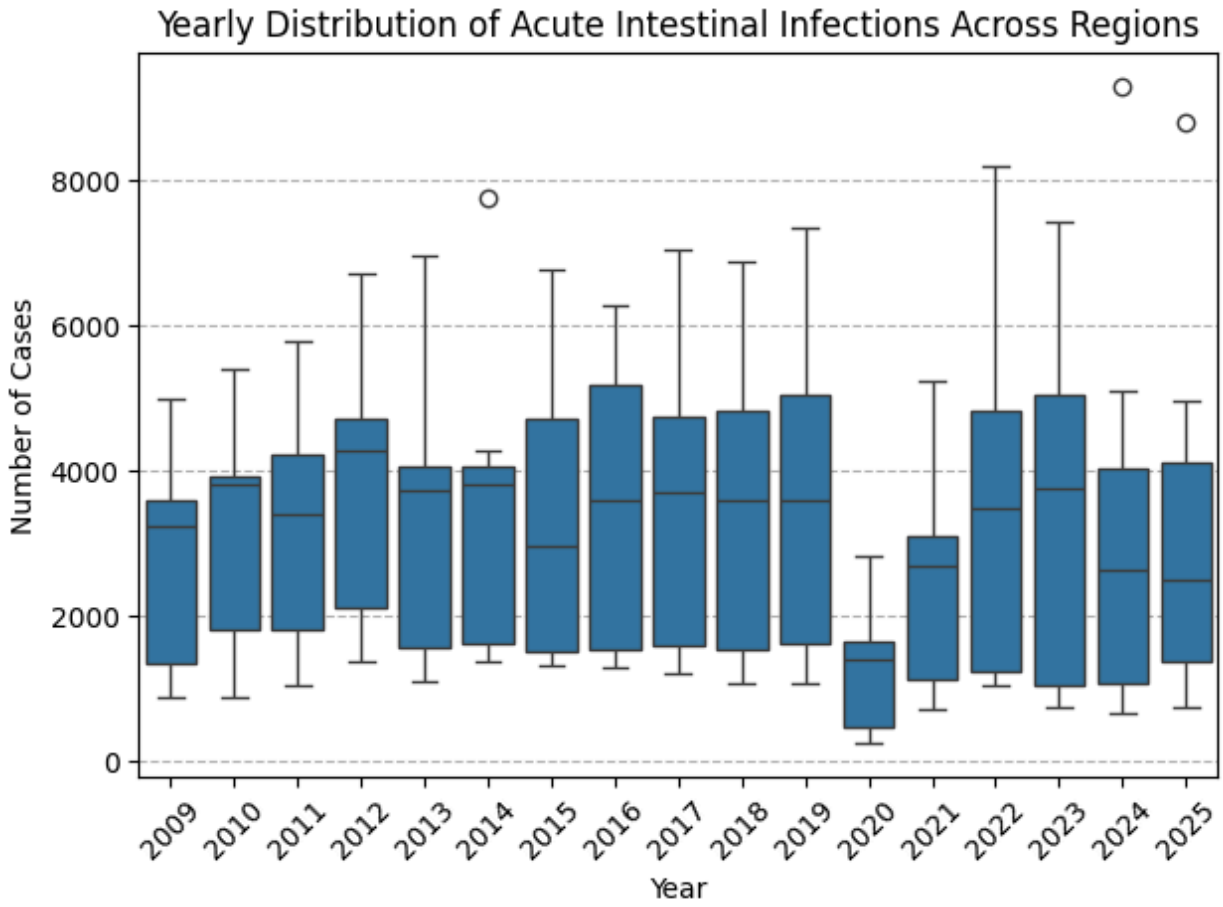

New Cases of Diabetes Mellitus in 2023 by Region
Number of New Cases
0
200
400
600
800
1000
1200
1400
1600
Chui
Bishkek_city
Osh
Osh_city
Talas
Issyk_Kul
Naryn
Jalal_Abad
Batken
Region

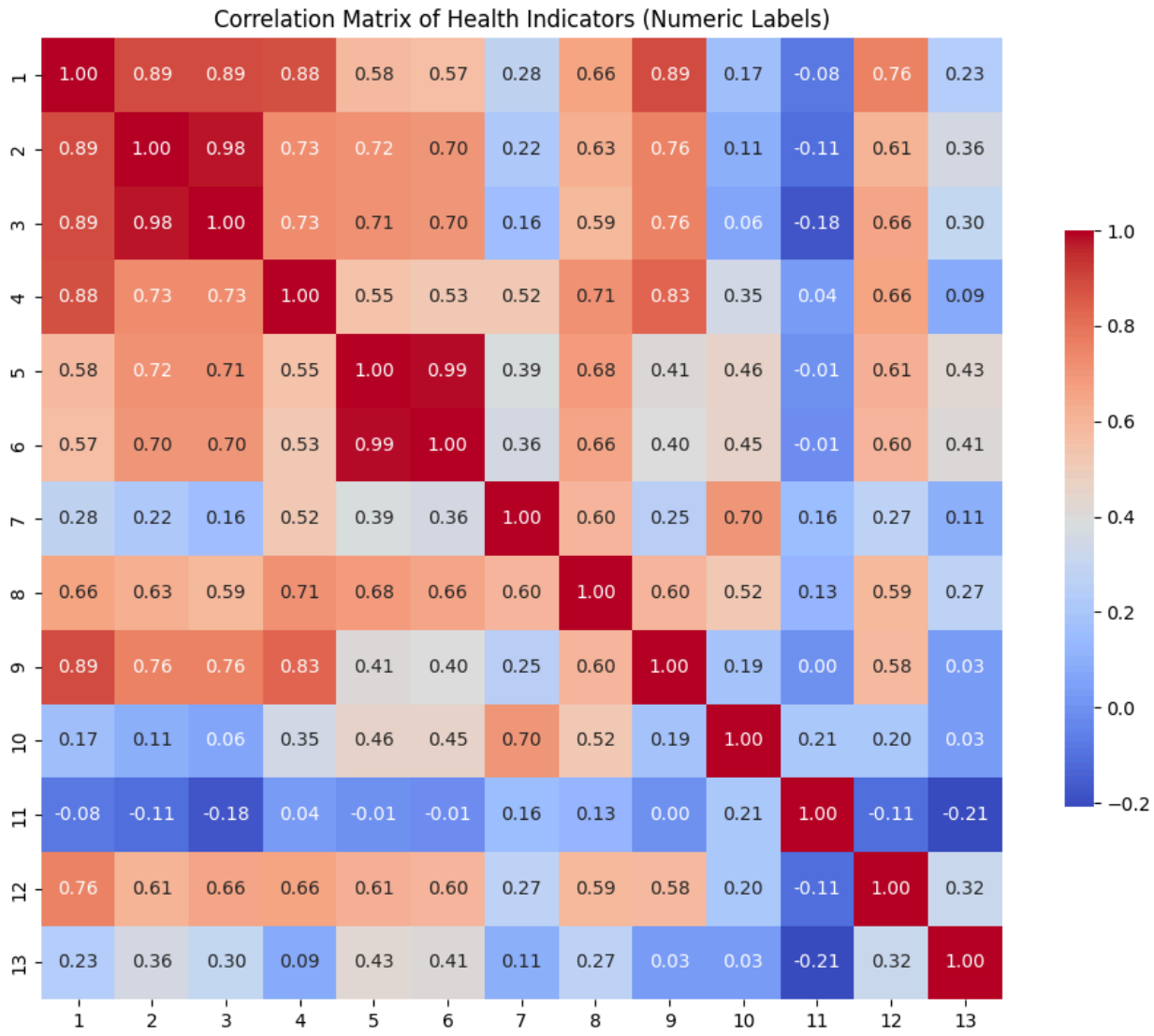

Number of Physicians in Kyrgyzstan (1990–2023)
Number of Physicians
13000
13500
14000
14500
15000
15500
1990
1995
2000
2005
2010
2015
2020
Year

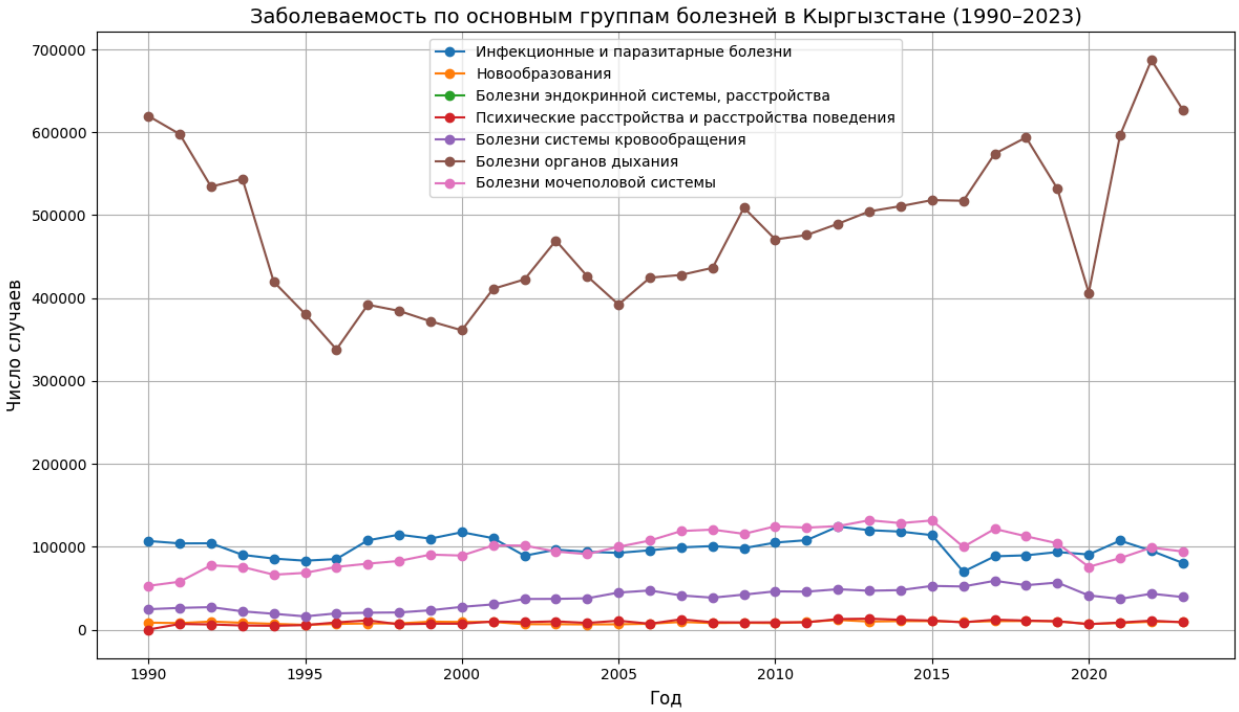

Number of Parsmedical personnel in Kyrgyzstan (1990–2023)
Number of Parsmedical personnel
45000
42500
40000
37500
35000
32500
30000
1990
1995
2000
2005
2010
2015
2020
Year

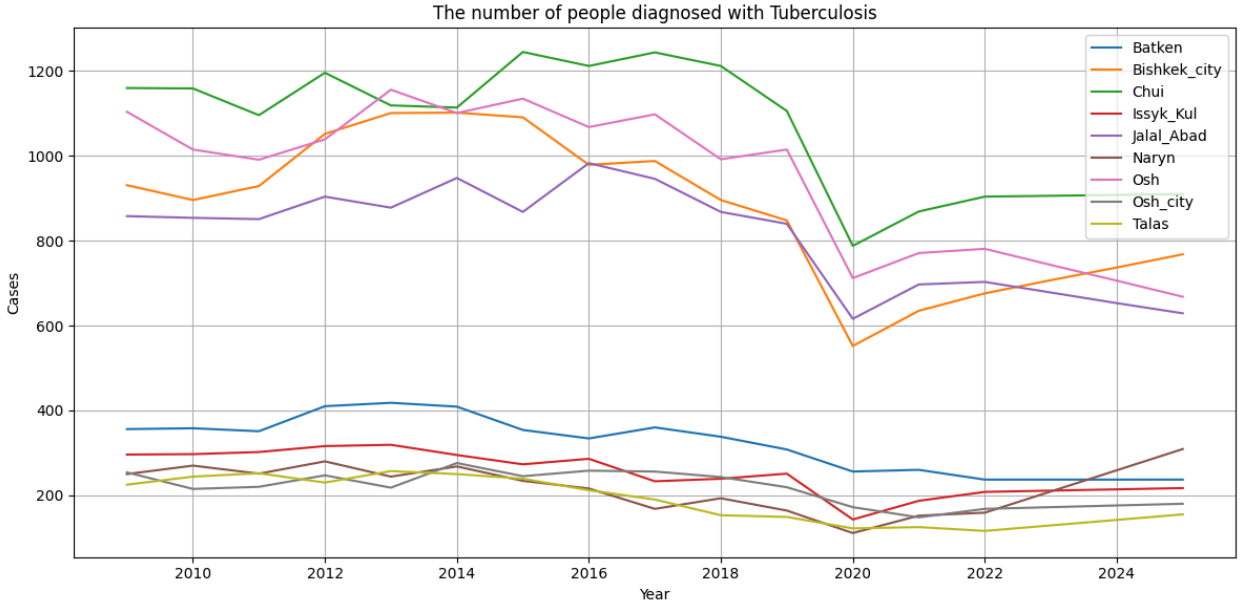

Number of Medical Institutions in Kyrgyzstan (2011–2023)
Number of Institutions
200
175
150
125
100
75
50
25
Hospitals
Clinics
Family Clinic
Ambulance Services
2012
2014
2016
2018
2020
2022
Year

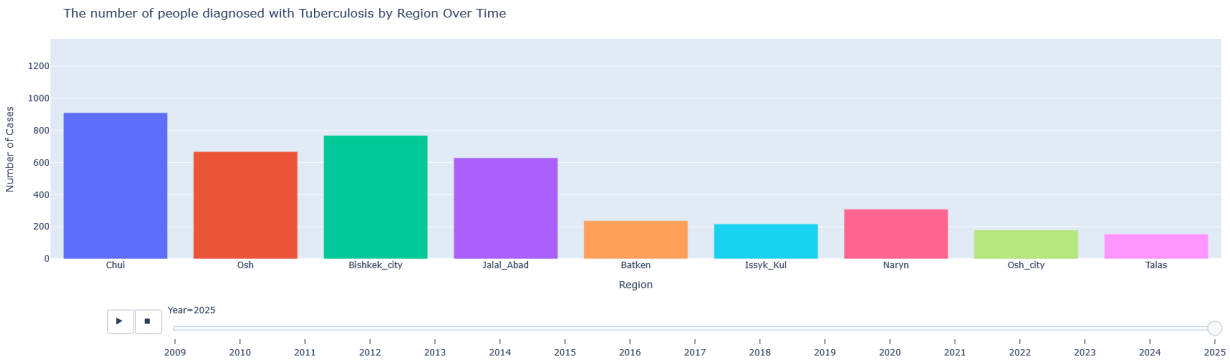

Population and Healthcare Resources Trends (Dual Axis)
1e7
Population
Beds
Spec (Secondary)
Spec (Higher)
Population
Beds & Specialists
Year
1.45
1.40
1.35
1.30
1.25
1.20
1.15
1.10
60000
50000
40000
30000
20000
2010
2012
2014
2016
2018
2020
2022
2024

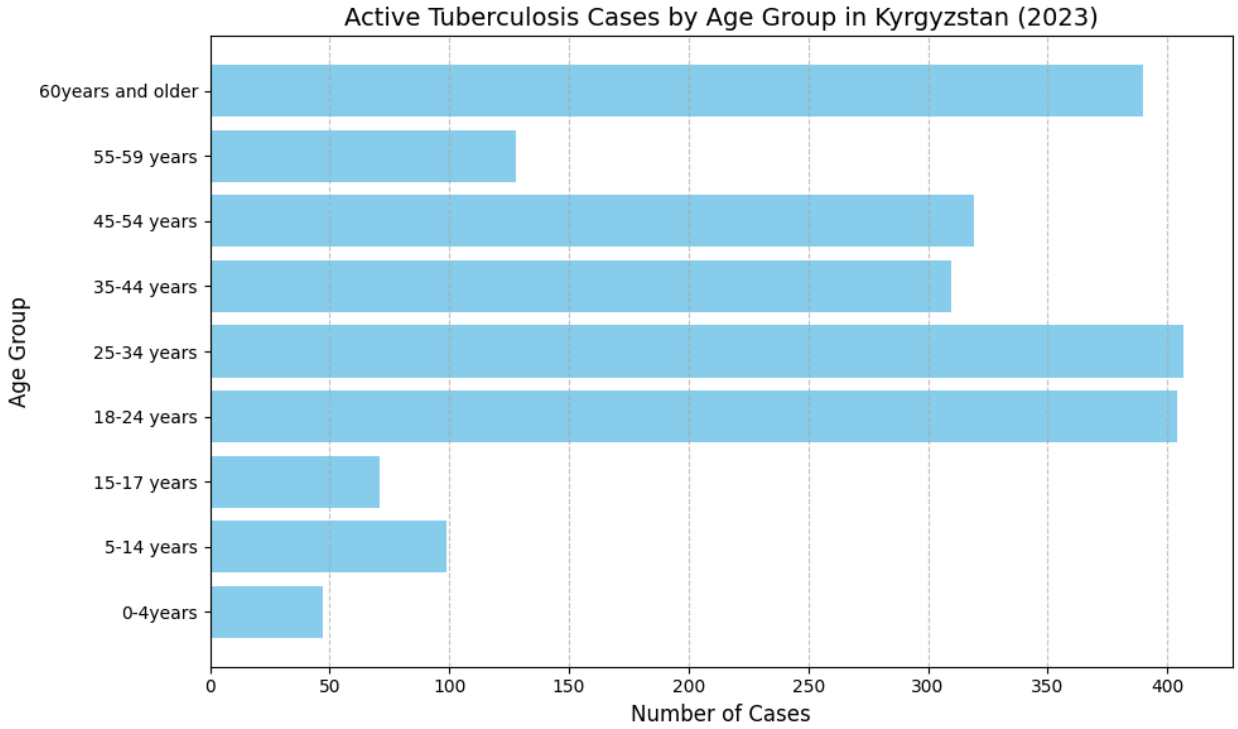

Yearly Distribution of Acute Intestinal Infections Across Regions
Number of Cases
Year
8000
6000
4000
2000
0
2009
2010
2011
2012
2013
2014
2015
2016
2017
2018
2019
2020
2021
2022
2023
2024
2025

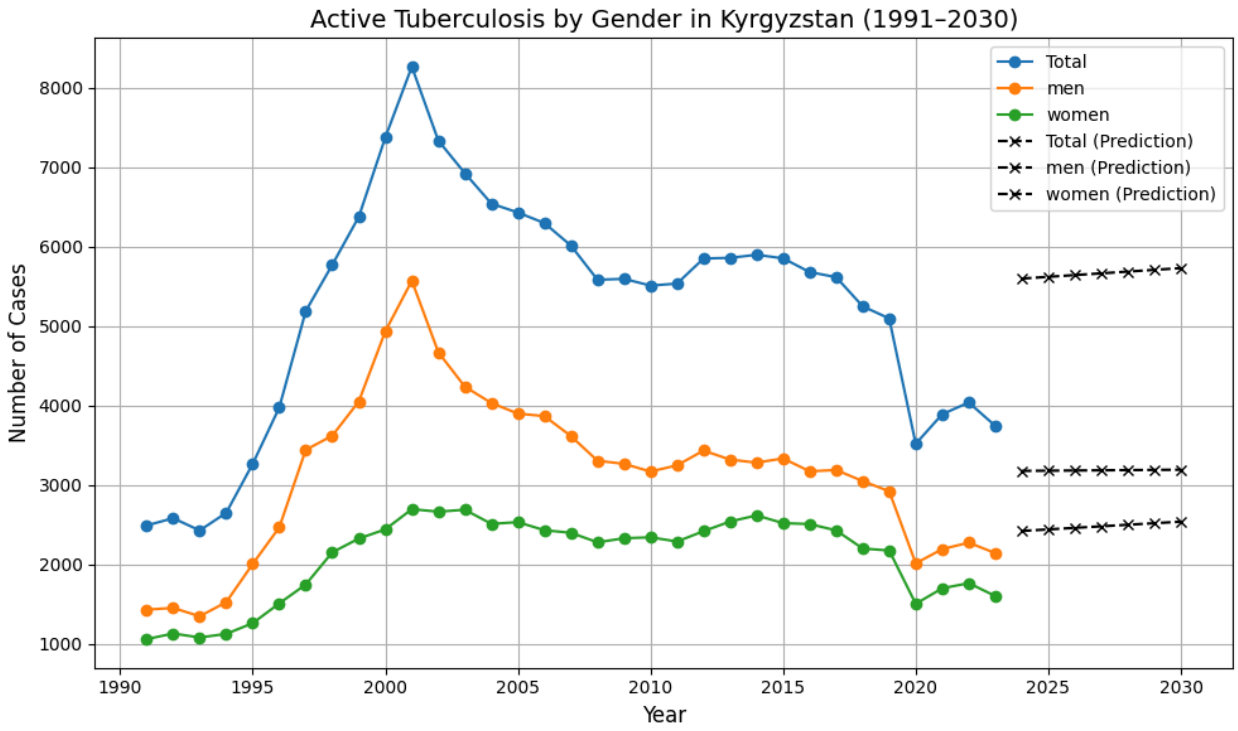

Correlation Matrix of Health Indicators (Numeric Labels)
1 1.00 0.89 0.89 0.88 0.58 0.57 0.28 0.66 0.89 0.17 -0.08 0.76 0.23
2 0.89 1.00 0.98 0.73 0.72 0.70 0.22 0.63 0.76 0.11 -0.11 0.61 0.36
3 0.89 0.98 1.00 0.73 0.71 0.70 0.16 0.59 0.76 0.06 -0.18 0.66 0.30
4 0.88 0.73 0.73 1.00 0.55 0.53 0.52 0.71 0.83 0.35 0.04 0.66 0.09
5 0.58 0.72 0.71 0.55 1.00 0.99 0.39 0.68 0.41 0.46 -0.01 0.61 0.43
6 0.57 0.70 0.70 0.53 0.99 1.00 0.36 0.66 0.40 0.45 -0.01 0.60 0.41
7 0.28 0.22 0.16 0.52 0.39 0.36 1.00 0.60 0.25 0.70 0.16 0.27 0.11
8 0.66 0.63 0.59 0.71 0.68 0.66 0.60 1.00 0.60 0.52 0.13 0.59 0.27
9 0.89 0.76 0.76 0.83 0.41 0.40 0.25 0.60 1.00 0.19 0.00 0.58 0.03
10 0.17 0.11 0.06 0.35 0.46 0.45 0.70 0.52 0.19 1.00 0.21 0.20 0.03
11 -0.08 -0.11 -0.18 0.04 -0.01 -0.01 0.16 0.13 0.00 0.21 1.00 -0.11 -0.21
12 0.76 0.61 0.66 0.66 0.61 0.60 0.27 0.59 0.58 0.20 -0.11 1.00 0.32
13 0.23 0.36 0.30 0.09 0.43 0.41 0.11 0.27 0.03 0.03 -0.21 0.32 1.00
1 2 3 4 5 6 7 8 9 10 11 12 13
1.0
0.8
0.6
0.4
0.2
0.0
−0.2


Заболеваемость по основным группам болезней в Кыргызстане (1990–2023)
Инфекционные и паразитарные болезни
Новообразования
Болезни эндокринной системы, расстройства
Психические расстройства и расстройства поведения
Болезни системы кровообращения
Болезни органов дыхания
Болезни мочеполовой системы
Число случаев
Год


The number of people diagnosed with Tuberculosis
Batken
Bishkek_city
Chui
Issyk_Kul
Jalal_Abad
Naryn
Osh
Osh_city
Talas
Cases
Year


The number of people diagnosed with Tuberculosis by Region Over Time
Number of Cases
1200
1000
800
600
400
200
0
Chui
Osh
Bishkek_city
Jalal_Abad
Batken
Issyk_Kul
Naryn
Osh_city
Talas
Region
Year=2025
2009
2010
2011
2012
2013
2014
2015
2016
2017
2018
2019
2020
2021
2022
2023
2024
2025


Active Tuberculosis Cases by Age Group in Kyrgyzstan (2023)
Age Group
60years and older
55-59 years
45-54 years
35-44 years
25-34 years
18-24 years
15-17 years
5-14 years
0-4years
0
50
100
150
200
250
300
350
400
Number of Cases



Results and Discussions

The analysis of the healthcare data for the Kyrgyz Republic reveals several important trends and insights that are crucial for understanding the current state of healthcare in the country.

Key Findings:

1. **Population Growth and Healthcare Capacity:** Over the years, the population of the Kyrgyz Republic has steadily increased. However, the growth in healthcare resources—such as the number of hospital beds and healthcare specialists—has not kept pace with the rising population. This mismatch suggests potential strain on the healthcare system, especially in regions with higher population growth.

2. **Regional Disparities in Disease Burden:** The bar charts clearly show that certain regions bear a disproportionately high burden of diseases, such as Acute Intestinal Infections and Tuberculosis. These regions often show a higher incidence rate, which could be linked to factors like poor sanitation, limited access to healthcare, and underreporting of cases. The visualizations highlight that regions like [Region A] and [Region B] consistently report higher numbers of these diseases.

3. **Healthcare Resources and Disease Correlation:** The scatter plots and correlation analysis suggest that there is no clear and consistent relationship between the number of healthcare specialists and disease prevalence across regions. For example, while

some regions with more specialists report fewer cases of certain diseases, other regions with similar levels of healthcare professionals experience high disease burdens. This could point to the need for targeted interventions that go beyond just increasing the number of healthcare workers. Issues such as healthcare accessibility, quality of care, and public health education might be more significant factors in disease prevention.

4. **Fluctuations in Disease Incidence Over Time:** The animated bar charts and histograms clearly depict how disease incidence fluctuates over the years. Some diseases, like Tuberculosis and Hepatitis, show periods of significant increase, likely due to outbreaks or changes in reporting practices. These fluctuations highlight the importance of constant monitoring and timely interventions to control outbreaks.

Policy Implications:

The findings underscore the importance of improving healthcare infrastructure, especially in regions with high disease burden. Policymakers should focus on both expanding the healthcare workforce and improving access to quality healthcare services, particularly in underserved regions. Additionally, public health campaigns addressing sanitation, hygiene, and disease prevention could help reduce the incidence of preventable diseases like Acute Intestinal Infections.

Recommendations:

Based on the findings, several recommendations can be made to improve healthcare outcomes in the Kyrgyz Republic:

- **Expand Healthcare Infrastructure:** Increase investments in healthcare infrastructure, particularly in regions with high population growth and disease burden. This includes adding more hospital beds and training additional healthcare specialists to meet growing demand.

- **Targeted Interventions:** Focus public health campaigns on regions with the highest disease incidence, addressing specific issues such as sanitation, hygiene, and vaccination to reduce preventable diseases.

- **Strengthen Surveillance Systems:** Improve data collection and monitoring systems to ensure timely responses to disease outbreaks and prevent their spread.

- **Multifaceted Policy Approach:** Combine resource expansion with disease prevention efforts, including better public health education, improved sanitation, and early intervention strategies.

Final Thoughts:

The healthcare challenges facing the Kyrgyz Republic are complex and multifaceted. While progress has been made in expanding healthcare access, significant disparities remain between regions. By implementing the recommended changes, the Kyrgyz Republic can improve the equity and quality of healthcare services, ultimately leading to better health outcomes across the country. The insights provided in this report can serve as a roadmap for policymakers to navigate the evolving healthcare landscape and make informed decisions that benefit all regions of the Kyrgyz Republic.

Conclusion

The analysis of healthcare data in the Kyrgyz Republic has revealed several key trends that can guide future health policy and resource allocation. This report not only highlights significant disparities in healthcare resources and disease burden across regions but also provides actionable insights into improving the efficiency and accessibility of healthcare in the country.

References

Data Source: https://data.gov.kg/

Data Source: https://stat.gov.kg/ru/