OVERVIEW OF DATA ANALYSIS FINDINGS

Analysis of Healthcare Trends

Insights and visualizations from our recent data analysis

Introduction to Data Analysis Report

Objective of the Analysis:

- Investigate healthcare trends in Kyrgyzstan from 2009 to 2022.
- Analyze regional disparities in disease burden and healthcare resources.

Primary Goal:

Provide data-driven recommendations for policy and resource allocation.

Data Overview

- Data mostly covers 2009 - 2023 across all regions of Kyrgyzstan.
- Key variables include disease incidence, healthcare infrastructure, and medical workforce.
- Regions: Excludes Kyrgyz Republic totals; focuses on regional data.
- Key Diseases: Hepatitis, Tuberculosis, Diabetes, Acute Intestinal Infections, etc.

Data Sources:

NATIONAL STATISTICAL COMMITEE
KYRGYZ REPUBLIC

open data
KYRGYZSTAN

Methodology Overview

```
df = df.drop(columns=["Number of marriages", "The number of divorces", "
```

```
df['Region'] = df['Region'].fillna(method='ffill')

df['Population'] = df['Population'].interpolate(method='linear').fillna(method='
for col in df.columns:
    if col == 'Region' and 'Year' and 'Population':
        continue

    df[col] = pd.to_numeric(df[col], errors='coerce')

    if df[col].isna().sum() > 0:
        if df[col].dtype in ['float64', 'int64']:
            df[col] = (
                df[col]
                .interpolate(method='linear')
                .fillna(method='bfill')
                .fillna(method='ffill')
                .fillna(df[col].mean())
            )
        else:
            df[col] = df[col].fillna(df[col].mode()[0])

numeric_cols = df.select_dtypes(include=['number']).columns
df[numeric_cols] = df[numeric_cols].round(0).astype('Int64')


df_filtered = df[df["Year"].between(2009, 2025)]
df_filtered
```

Data Cleaning & Preprocessing:

- Standardized column names and handled missing data.

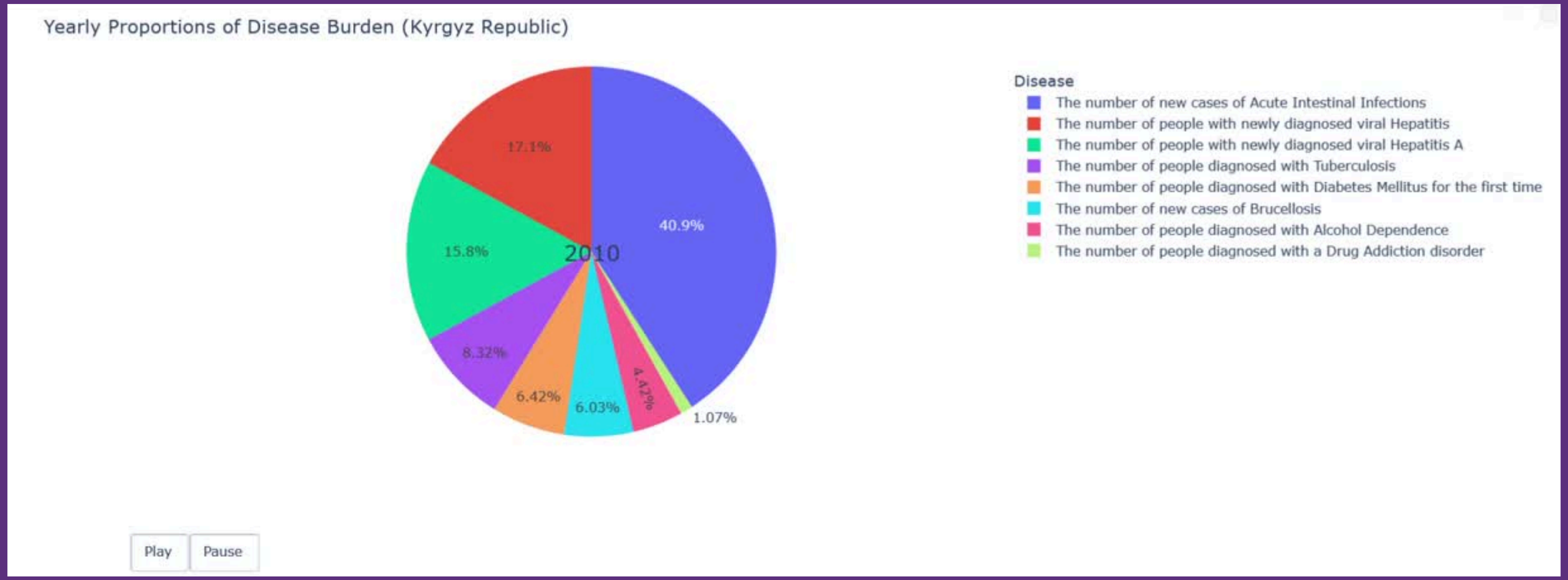
Visualization Tools:

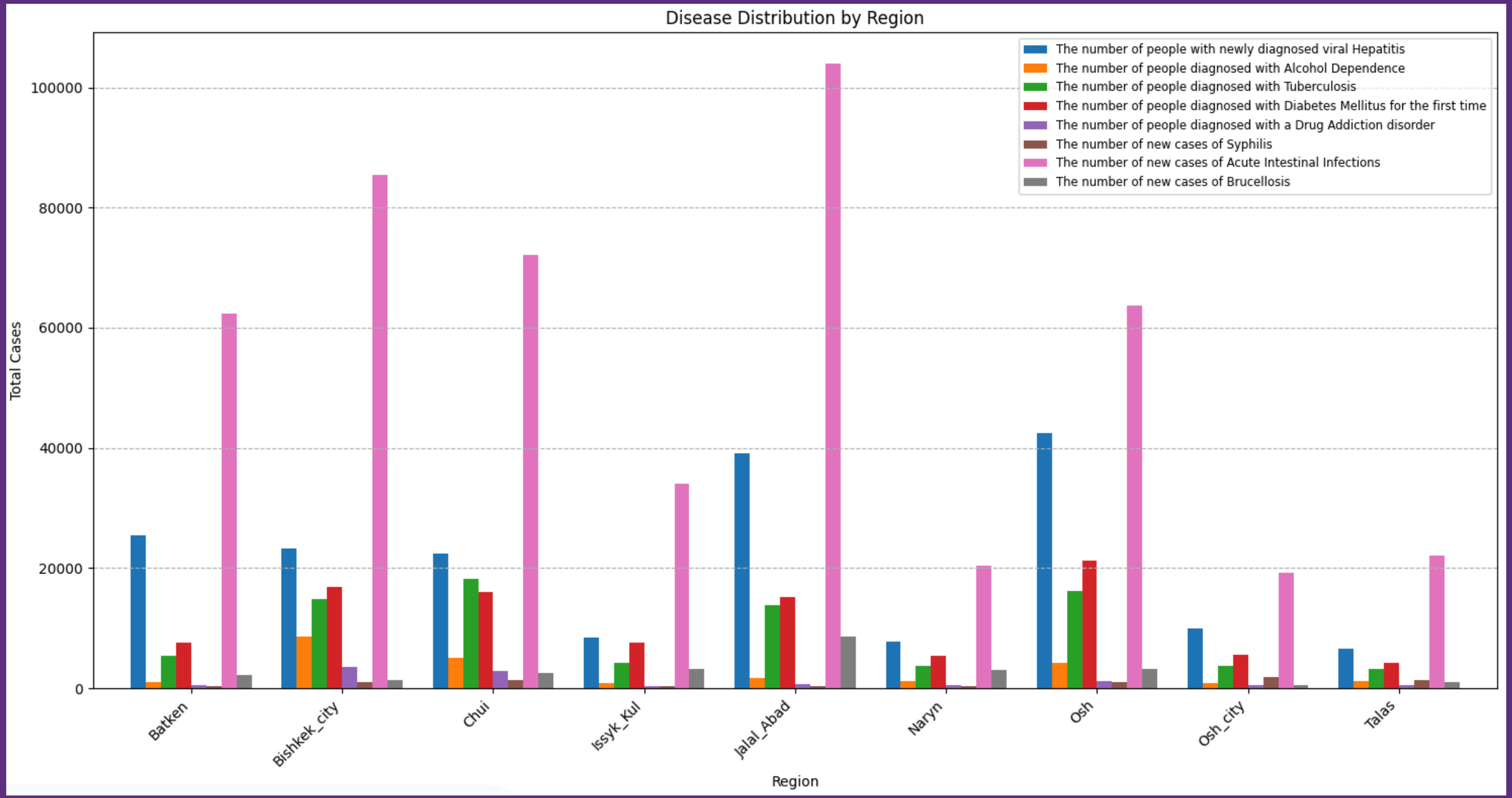
- Used Plotly for interactive and animated visualizations.

Analysis Techniques:

- Descriptive statistics to summarize data.
- Correlation analysis to explore relationships between resources and disease outcomes.

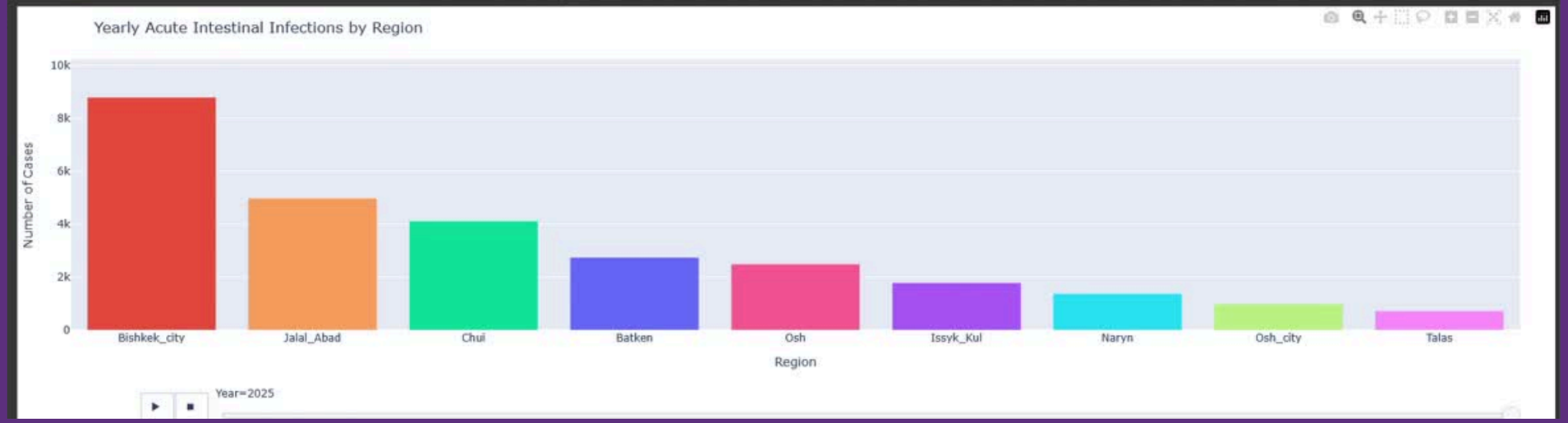
Let’s start with interactive plot




Disease Distribution by Region
The number of people with newly diagnosed viral Hepatitis
The number of people diagnosed with Alcohol Dependence
The number of people diagnosed with Tuberculosis
The number of people diagnosed with Diabetes Mellitus for the first time
The number of people diagnosed with a Drug Addiction disorder
The number of new cases of Syphilis
The number of new cases of Acute Intestinal Infections
The number of new cases of Brucellosis
Total Cases
0
20000
40000
60000
80000
100000
Batken
Bishkek_city
Chui
Issyk_Kul
Jalal_Abad
Naryn
Osh
Osh_city
Talas
Region

The number of new cases of Acute Intestinal Infections

```
)

fig = px.bar(
    agg,
    x='Region',
    y=value_column,
    color='Region',
    animation_frame='Year',
    range_y=[0, agg[value_column].max()*1.1],
    title=f'Yearly Acute Intestinal Infections by Region',
    labels={value_column: 'Number of Cases'}
)

fig.update_layout(
    xaxis={'categoryorder':'total descending'},
    showlegend=False,
    margin=dict(l=40, r=40, t=60, b=40)
)
fig.show()
```



Why so many?

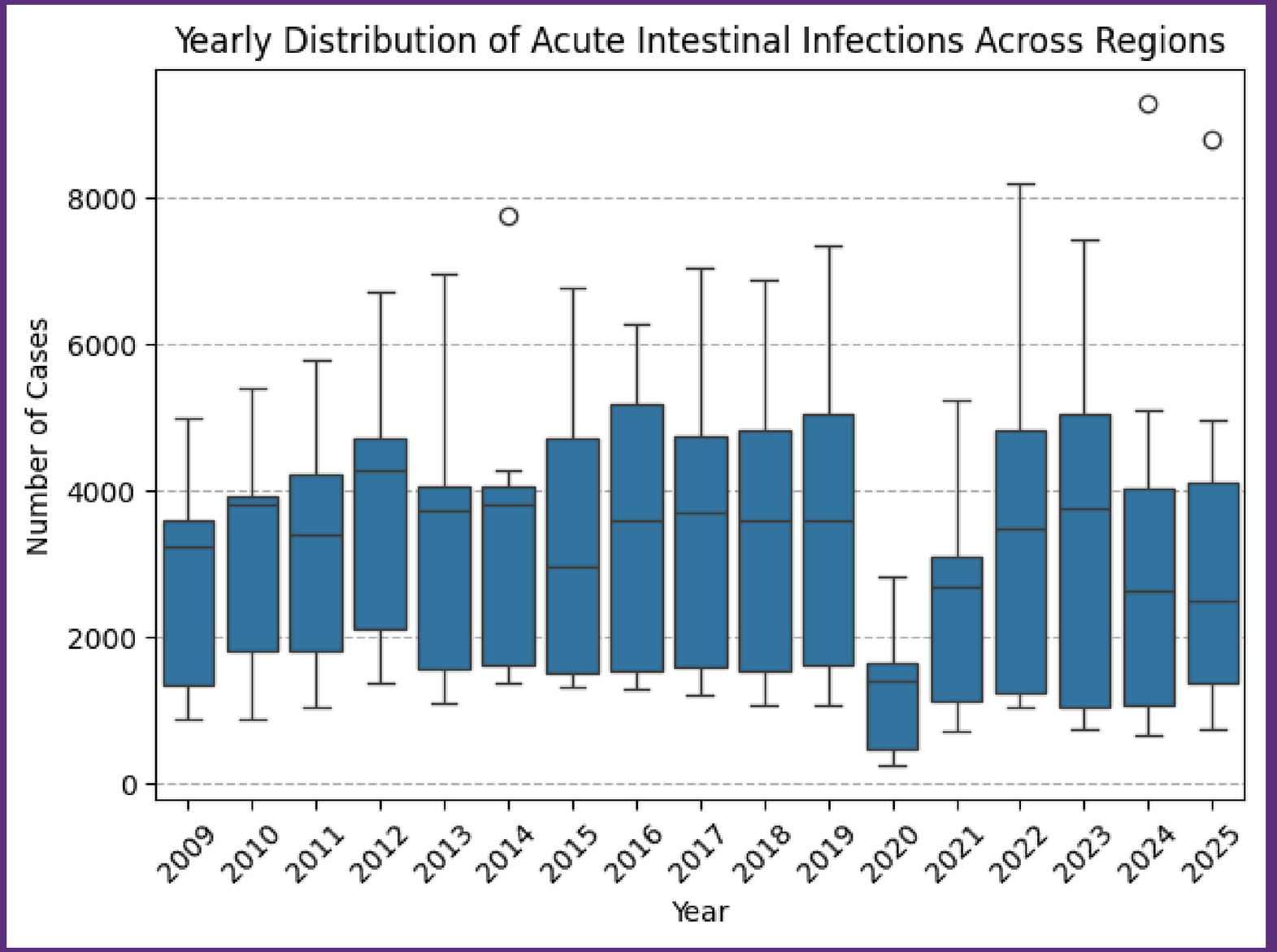
Unsafe Water Supply

Poor Sanitation Infrastructure

Seasonal Factors

Recommendations:

Strengthen Healthcare Reporting and Rapid Response


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

Proportion of Alcohol Dependence Cases by Region and Year

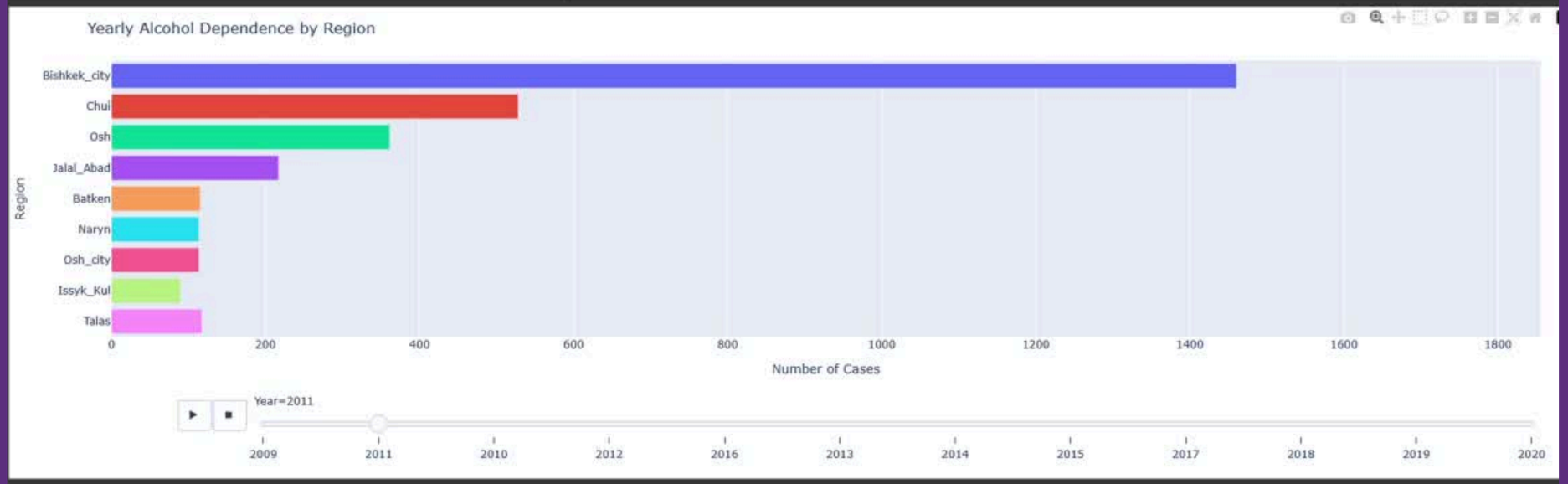
We used interactive plotly visualizations with many datasets

not only Alcohol + Drug Addiction

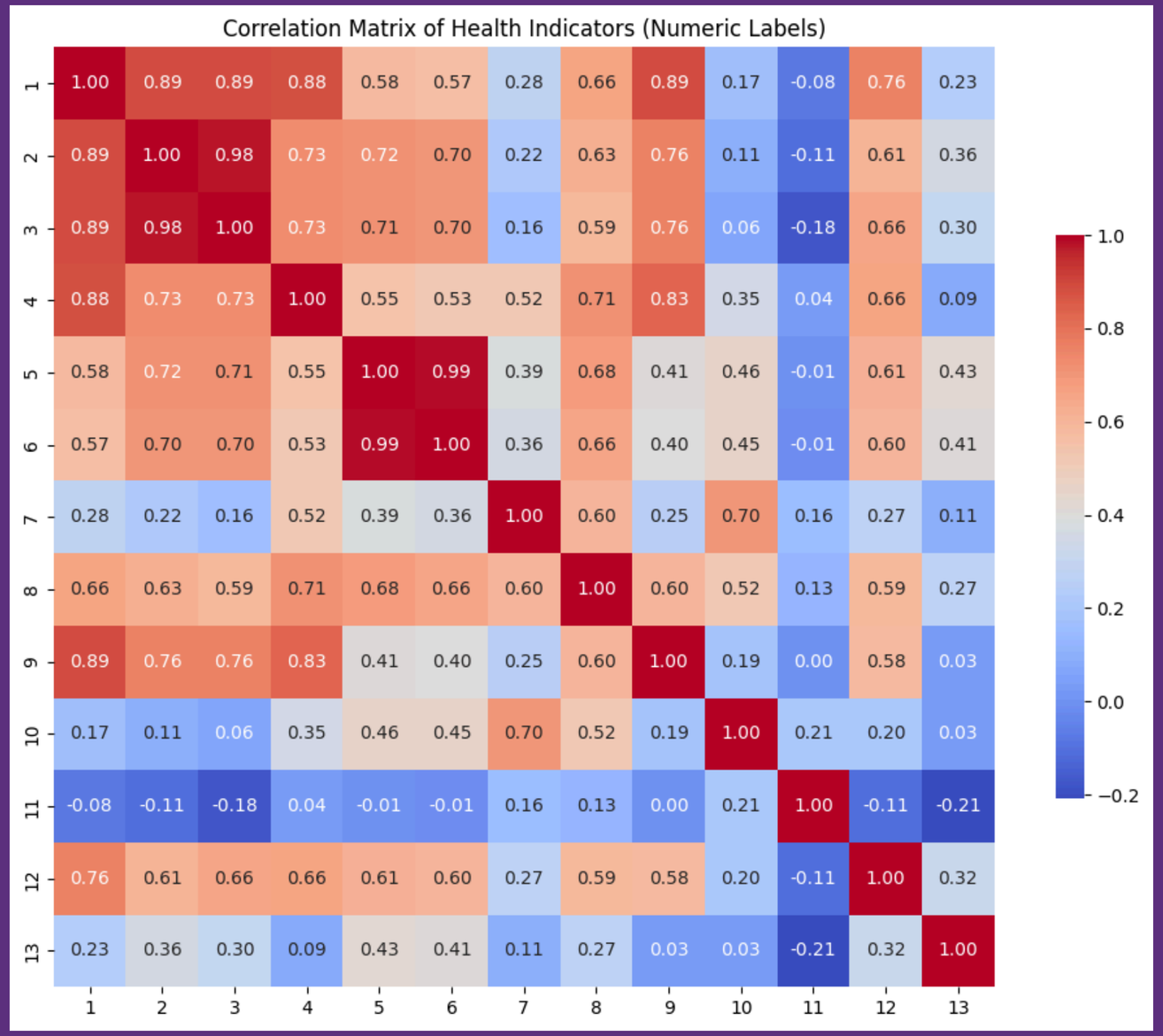
Recommendations:

Strengthen Healthcare Reporting and Rapid Response

```
    orientation='h'
)

fig.update_layout(
    xaxis={'categoryorder': 'total descending'},
    showlegend=False,
    margin=dict(l=40, r=40, t=60, b=40)

)

fig.show()
```




```
# Dual-Axis Line Chart: Population & Healthcare Resources Trends
import matplotlib.pyplot as plt
data = pd.read_csv('filtered.csv')
line_data = data.groupby('Year').agg({
```

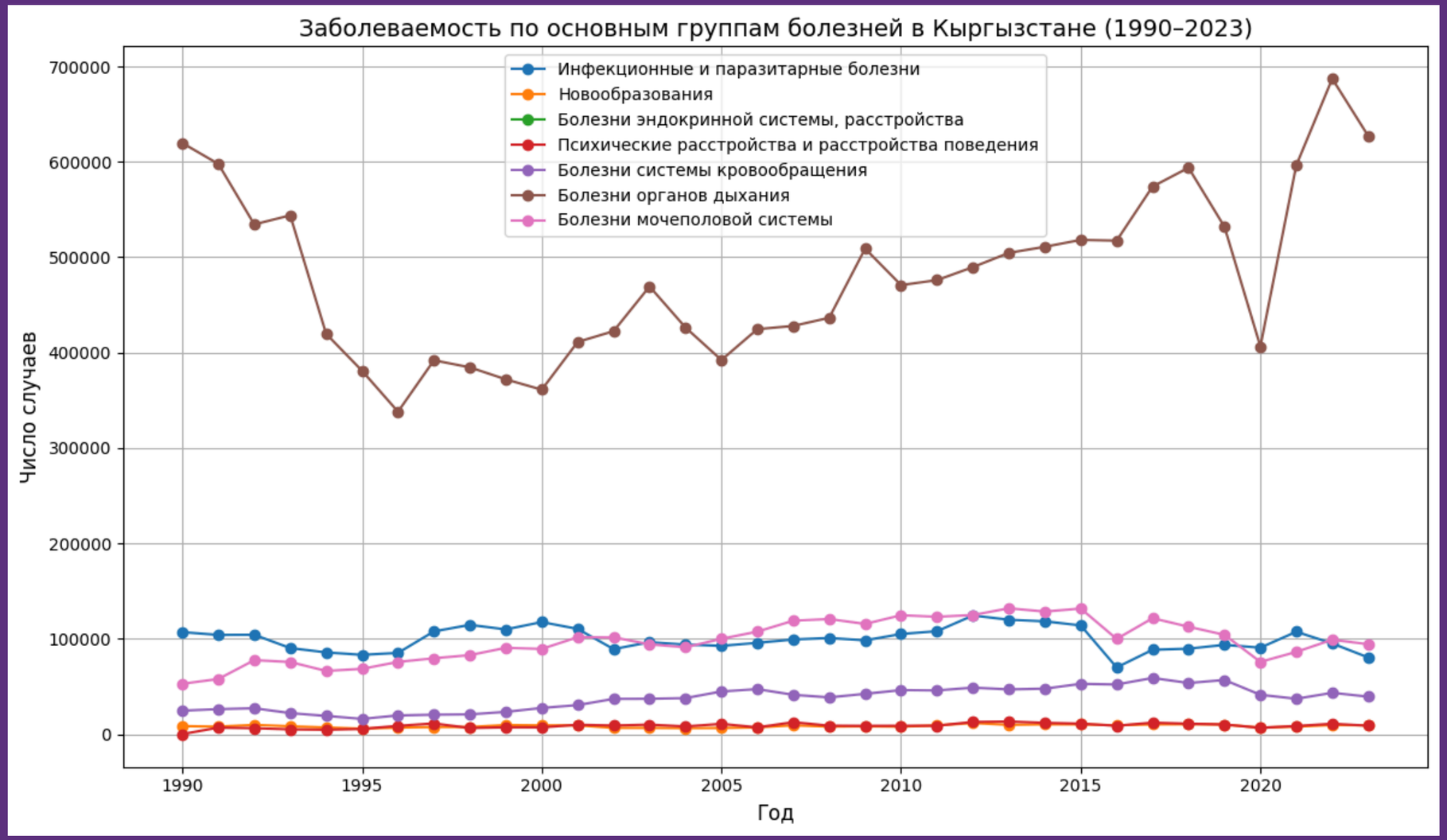


Correlation Matrix

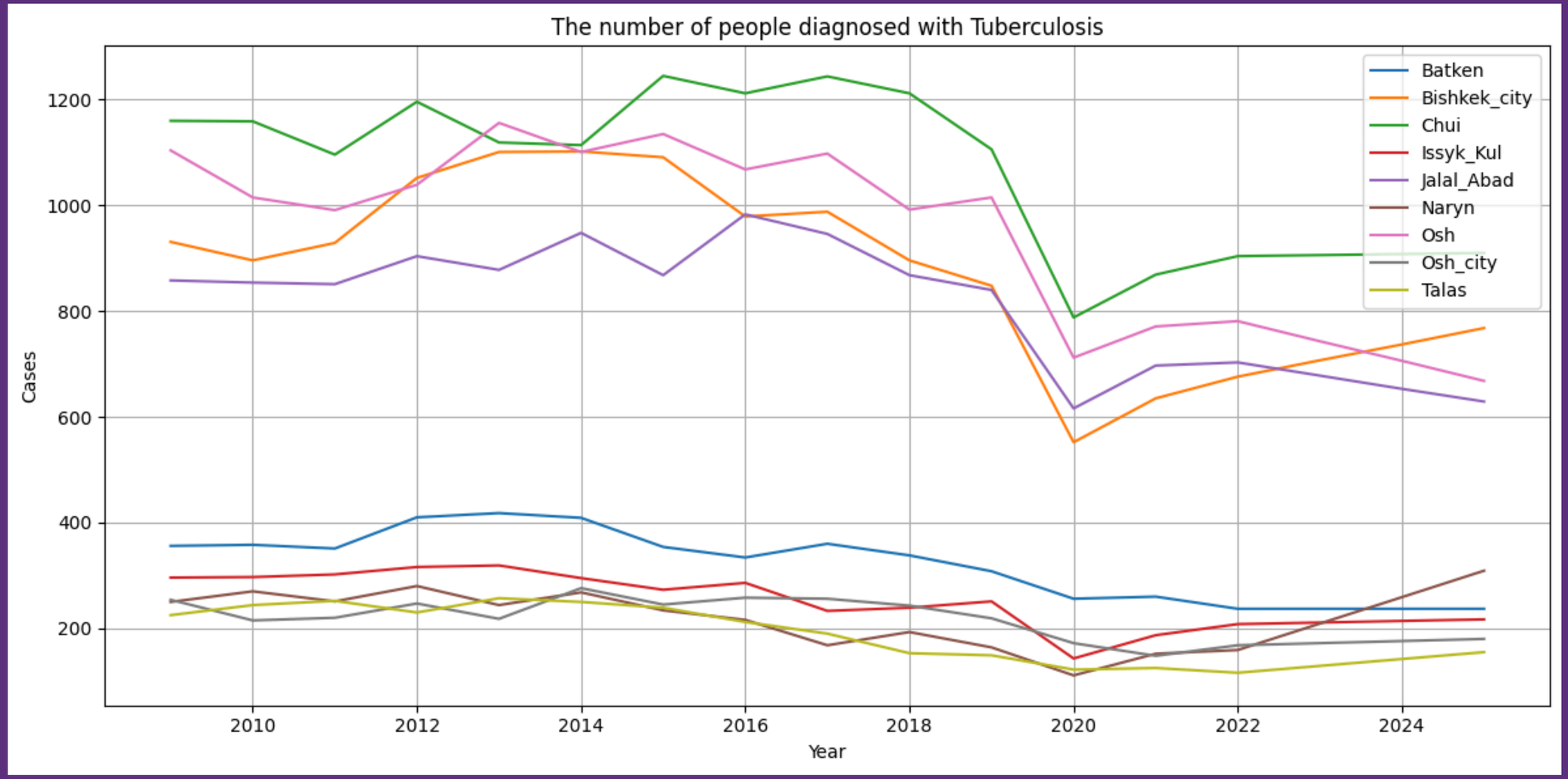
This correlation highlights the relationships between key health indicators and healthcare capacity across regions, offering insights into how different factors influence one another.

```
Variable mapping:
1 → Population
2 → Number of beds
3 → Number of specialists with secondary medical education
4 → Number of specialists with higher medical education
5 → The number of people with newly diagnosed viral Hepatitis
6 → The number of people with newly diagnosed viral Hepatitis A
7 → The number of people diagnosed with Alcohol Dependence
8 → The number of people diagnosed with Tuberculosis
9 → The number of people diagnosed with Diabetes Mellitus for the t
10 → The number of people diagnosed with a Drug Addiction disorder
11 → The number of new cases of Syphilis
12 → The number of new cases of Acute Intestinal Infections
13 → The number of new cases of Brucellosis
```

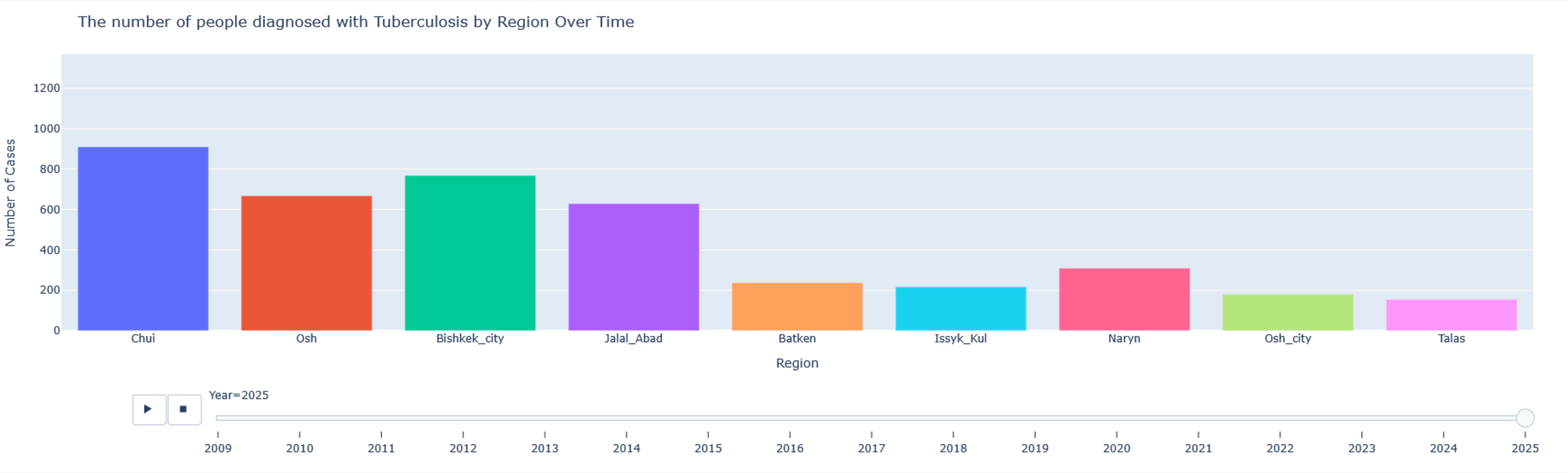
By type of the diseases



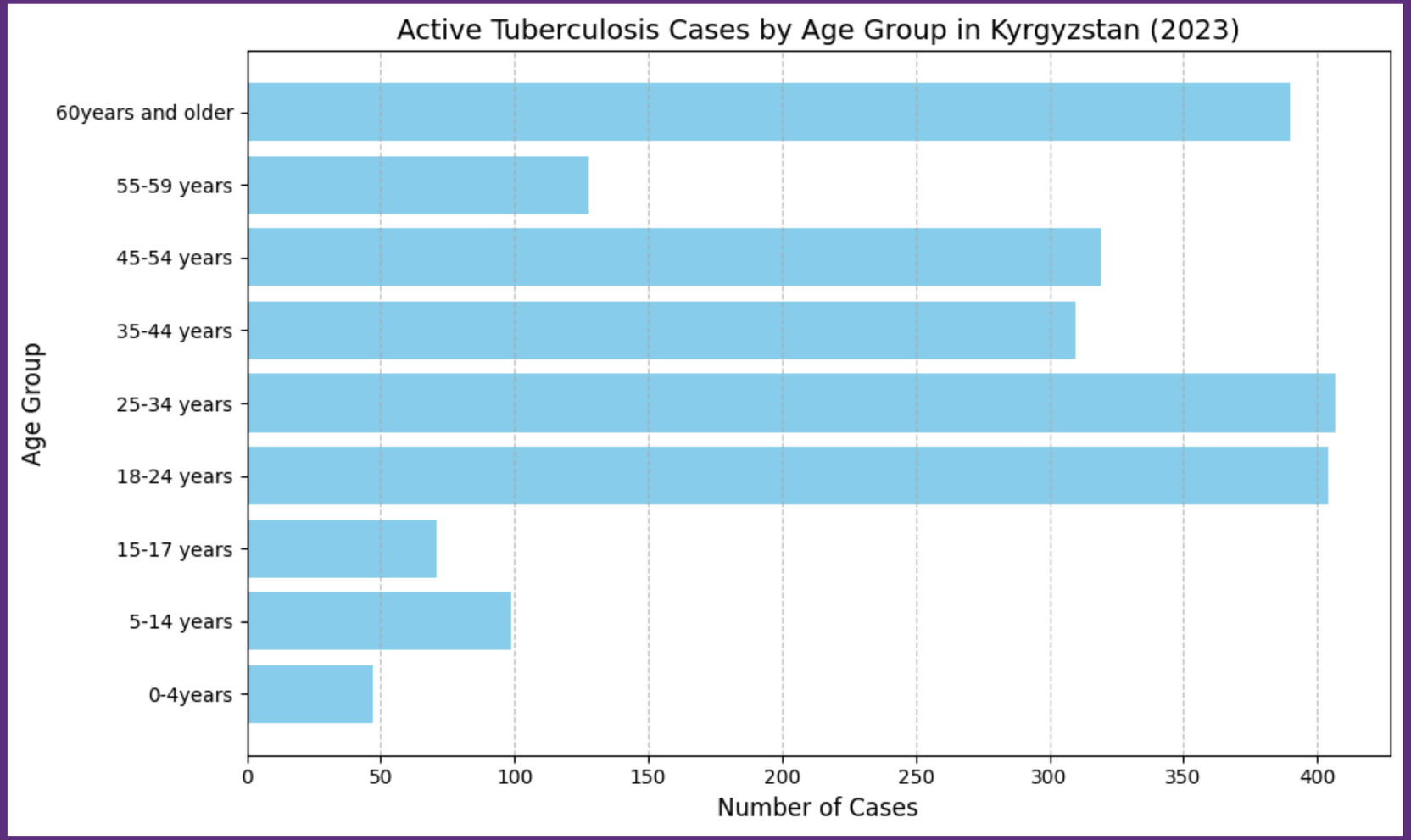
Let’s focus on Tuberculosis



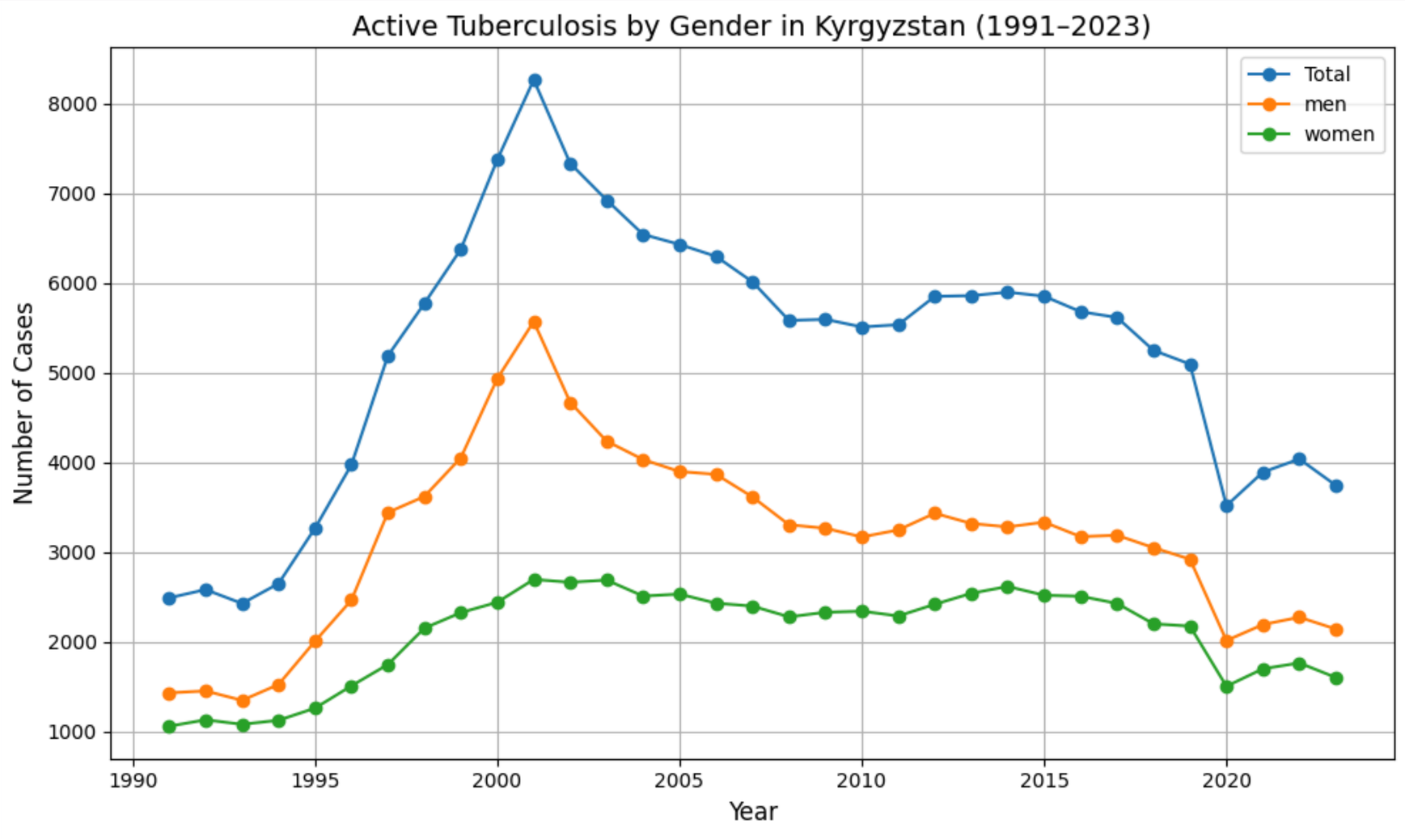
Interactive plot for visual understanding



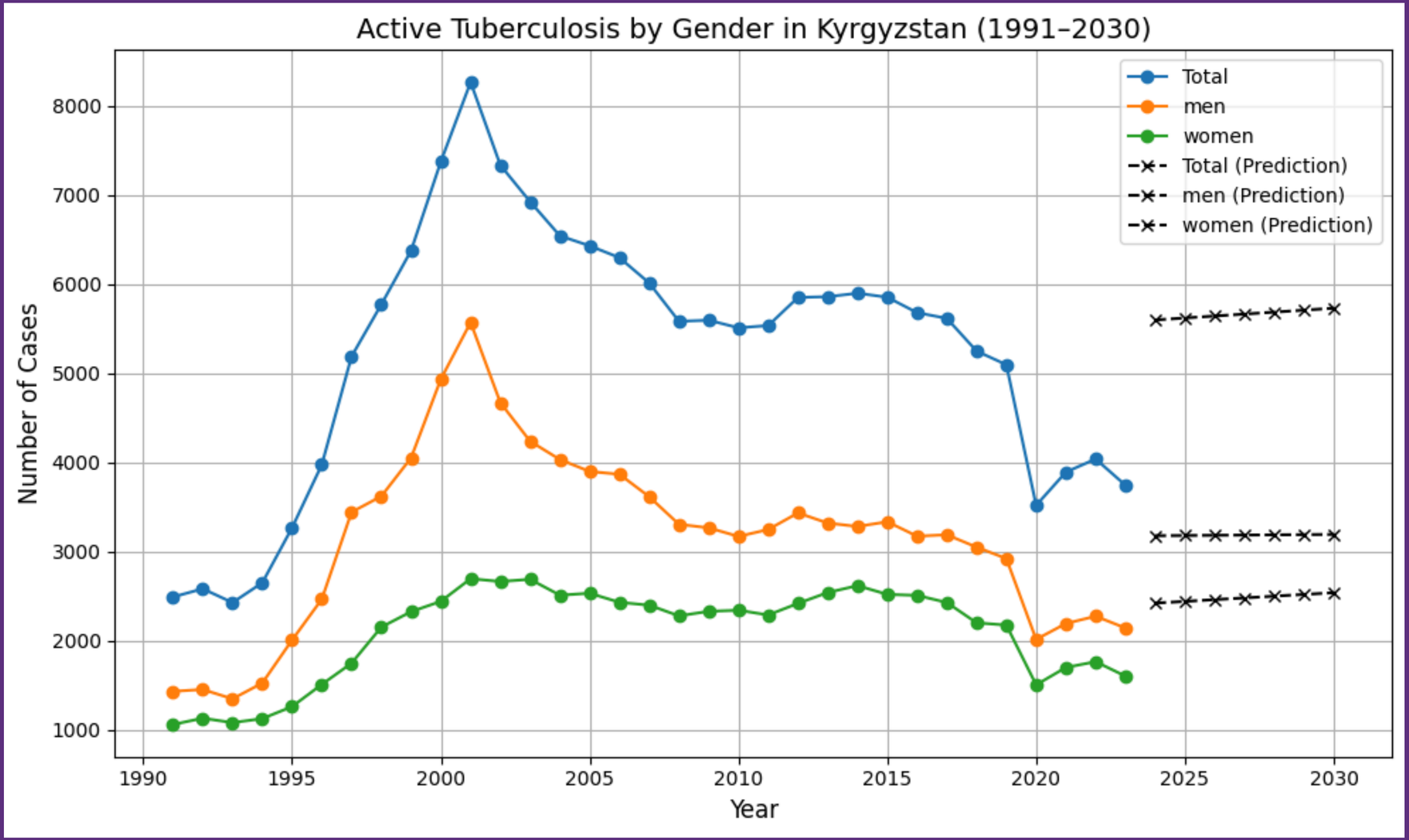
By Age


Active Tuberculosis Cases by Age Group in Kyrgyzstan (2023)
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
Age Group
0
50
100
150
200
250
300
350
400
Number of Cases

By gender


Active Tuberculosis by Gender in Kyrgyzstan (1991–2023)
Total
men
women
Number of Cases
8000
7000
6000
5000
4000
3000
2000
1000
1990
1995
2000
2005
2010
2015
2020
Year

Our Prediction:

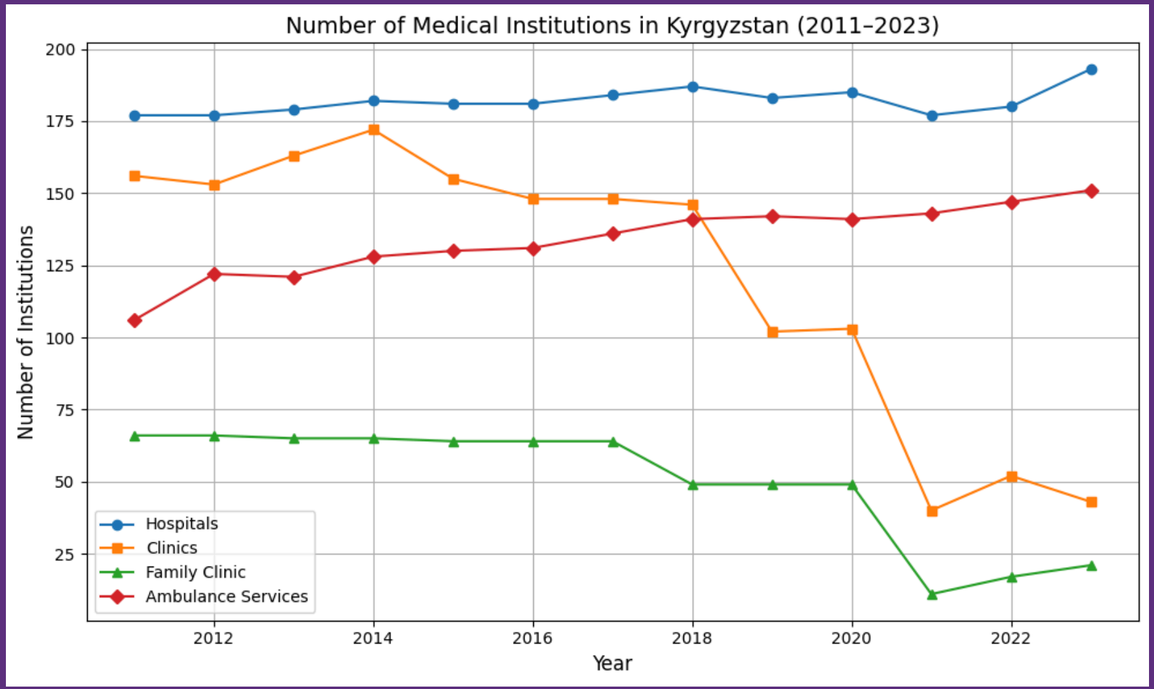


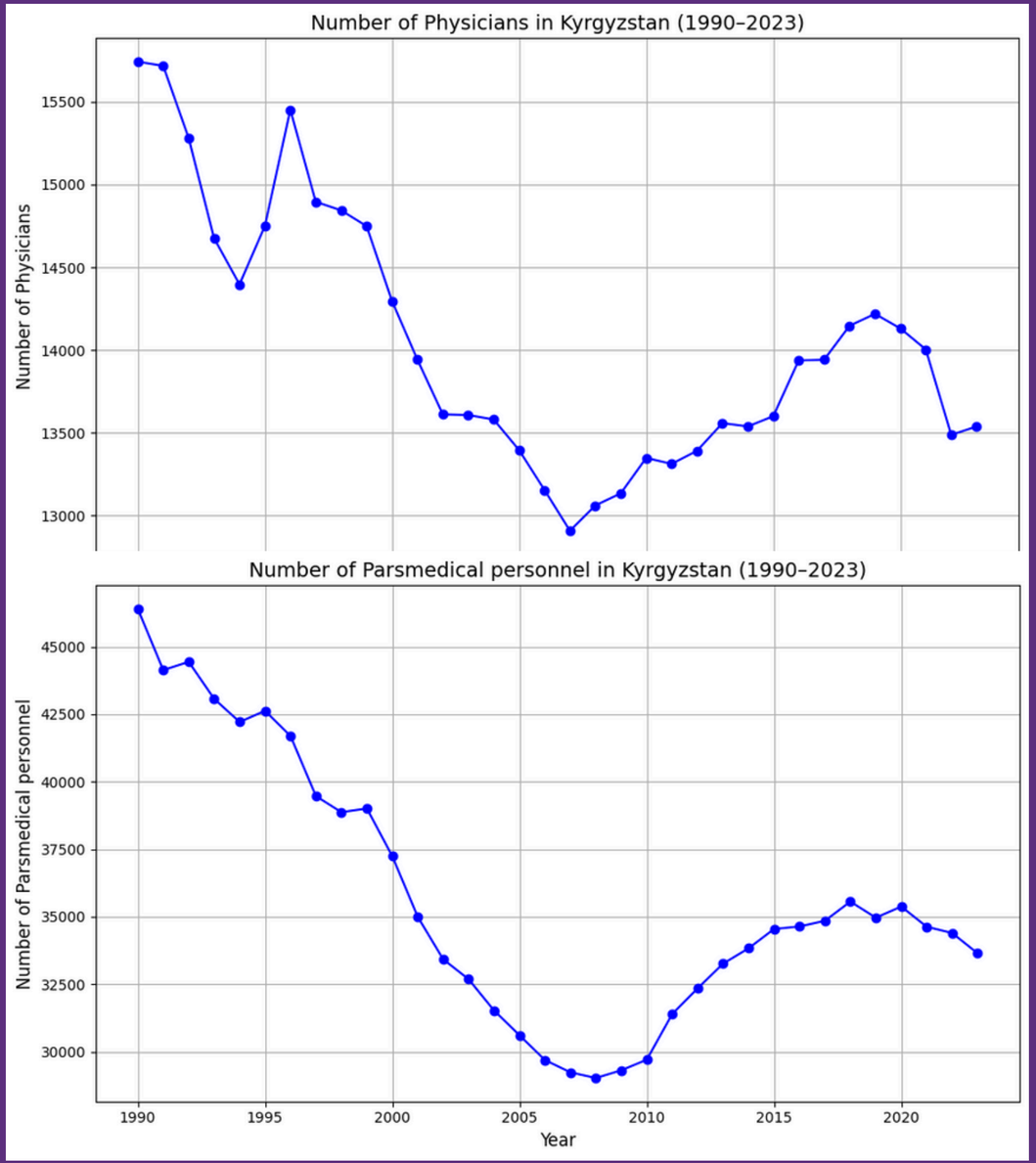
Now imagine:

The number of incidence increases

The number of population as well

But what about hospitals and doctors?





Conclusion

Recommendations:

- Expand Healthcare Infrastructure
- Targeted Interventions
- Strengthen Surveillance Systems:

The analysis of healthcare data in country 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.

Our 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

Members:

- **Nurali Bakytbek uulu**: Data Collection & Cleaning & Visualization
- **Bektur Momunov** :Data Cleaning & Visualization, Presentation
- **Magomed Mukhammedov** : Data Analysis, Report

Thank you for your time!

Do you have any questions?