

Home

Dashboard

Instruction to get API KEY

Overview

Data Report

NTViz's functions

NTViz Tasks

Sections

Provider Instruction

Gemini API key:

.....

Successfully connected to Gemini!

Tasks:

Functions:

Summarize & Goal

NTViz Tasks

 Filter Instruction  Requirements


Instruction: ▾



Select Model:


gemini-1.5-flash ▾

Upload a data file in .csv format:

 Drag and drop file here

Limit 200MB per file • CSV

Browse files

 crimea.csv 0.5KB

×

Successfully uploaded a CSV file with 24 rows of data.

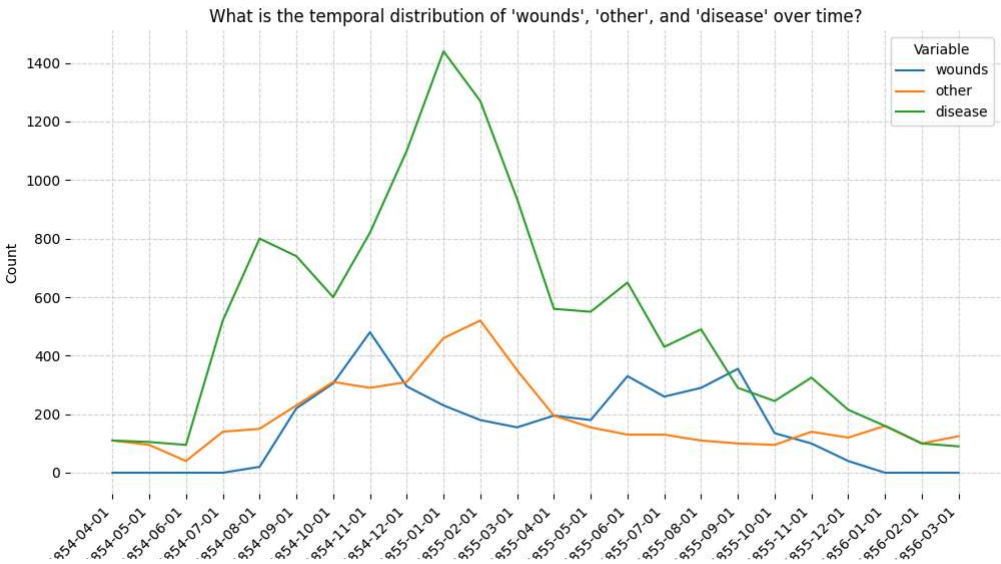
	date	wounds	other	disease
0	1854-04-01	0	110	110
1	1854-05-01	0	95	105
2	1854-06-01	0	40	95
3	1854-07-01	0	140	520
4	1854-08-01	20	150	800

No missing or duplicate values found in the data.


Generate Charts

★ Insight 1:

<pre>main() Goal Goal(question="What is the temporal distribution of 'wounds', 'other', and 'disease' over time?", visualization="Line chart of 'wounds', 'other', and 'disease' over 'date'", rationale="1. **Data Type:** 'date' is datetime, 'wounds', 'other', and 'disease' are numerical. 2. **Visualization Justifica...</pre>	
A visualization goal	
index int	0
question str	"What is the temporal distribution of 'wounds', 'other', and 'disease' over time?"
rationale str	"1. **Data Type:** 'date' is datetime, 'wounds', 'other', and 'disease' are numerical. 2. **Visualization Justification:** A line chart effectively visualizes trends over time for multiple variables. This allows for easy comparison of the fluctuations in the number of wounds, other cases, and disea..."
visualization str	"Line chart of 'wounds', 'other', and 'disease' over 'date'"



[°*↶•••?↷ Download Chart °*](#)

 VizOps ▾

 Analyze Chart 1

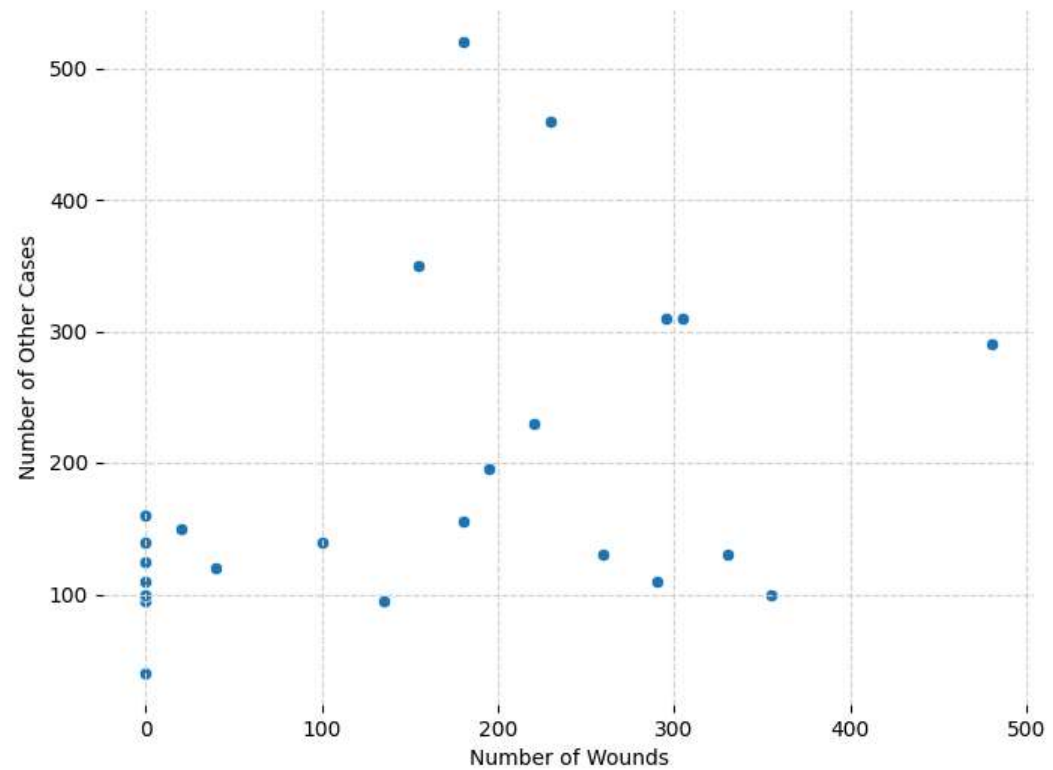
✱ Insight 2:

```
main() Goal Goal(question="What is the correlation between the number of 'wounds' and the number of 'other' cases?", visualization="Scatter plot of 'wounds' vs. 'other'", rationale="1. **Data Type:** 'wounds' and 'other' are both numerical. 2. **Visualization Justification:** A scatter plot is ideal for explori...
```

A visualization goal

index	int	1
question	str	"What is the correlation between the number of 'wounds' and the number of 'other' cases?"
rationale	str	"1. **Data Type:** 'wounds' and 'other' are both numerical. 2. **Visualization Justification:** A scatter plot is ideal for exploring the relationship between two numerical variables. It reveals the strength and direction of the correlation. 3. **Justification:** This visualization helps determine ...
visualization	str	"Scatter plot of 'wounds' vs. 'other'"

What is the correlation between the number of 'wounds' and the number of 'other' cases?



°*↶•••?↷ Download Chart °*

⚙ VizOps ▾

🔍 Analyze Chart 2

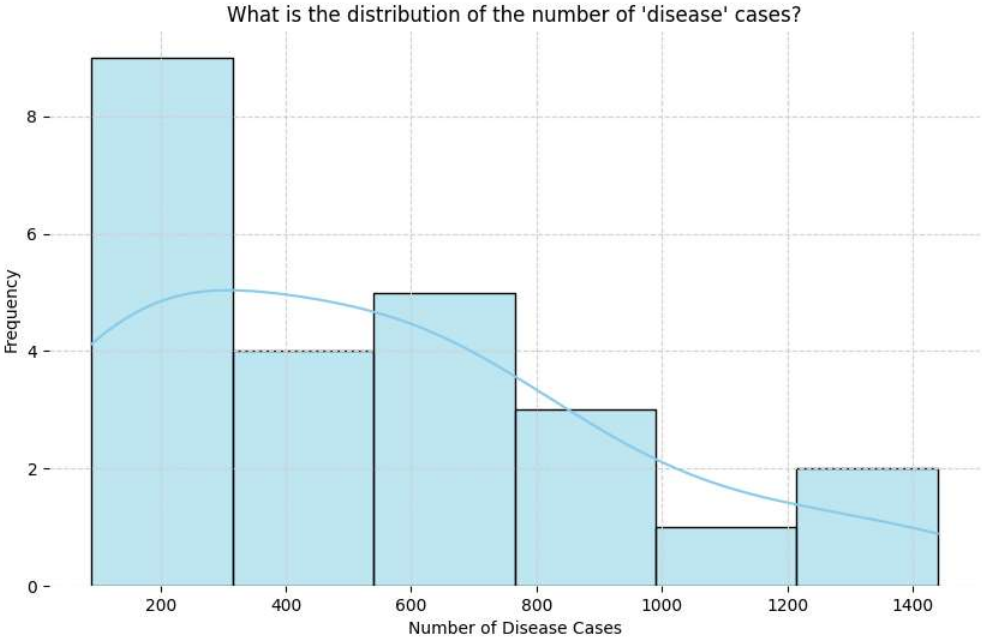
★ Insight 3:

```
main() Goal Goal(question="What is the distribution of the number of 'disease' cases?",  
visualization="Histogram of 'disease'", rationale="1. **Data Type:** 'disease' is
```

numerical. 2. **Visualization Justification:** A histogram effectively shows the frequency distribution of a numerical variable, revealing its...

A visualization goal

index	int	2
question	str	"What is the distribution of the number of 'disease' cases?"
rationale	str	"1. Data Type: 'disease' is numerical. 2. Visualization Justification: A histogram effectively shows the frequency distribution of a numerical variable, revealing its central tendency, spread, and skewness. 3. Justification: This visualization helps understand the typical number of disea..."
visualization	str	"Histogram of 'disease'"



[Download Chart](#)

VizOps

Analyze Chart 3

provider: gemini



Chart Analysis

Analysis Report

Analysis:

1. Chart Description:

- The chart is a histogram showing the distribution of the number of "disease" cases.
- The x-axis represents the number of disease cases.
- The y-axis represents the frequency, or how many times a particular number of disease cases occurred.
- The histogram uses bars to show the frequency of different ranges of disease cases. A kernel density estimate (KDE) line is overlaid, smoothing the distribution to highlight the overall shape. The title clearly states the purpose of the visualization.

2. Noticeable Patterns:

- The distribution of disease cases appears to be right-skewed (positively skewed). This means there are more cases with a lower number of diseases and a long tail towards higher numbers of diseases.
- There's a concentration of disease cases in the lower to mid-range (approximately 200-700 cases), suggesting a modal value in this region.
- The KDE line helps visualize the central tendency and spread of the data more clearly than the histogram alone. It confirms the right skew.
- There is at least one potential outlier at the high end of the distribution (around 1440 cases). This warrants further investigation.

3. Contextual Insights:

- The summary statistics show a mean of 526.67 disease cases and a standard deviation of 386.71. This high standard deviation, combined with the right skew, indicates considerable variability in the number of disease cases.
- The median (505 cases) is lower than the mean, further supporting the right skew. This suggests that a few instances of very high disease counts are pulling the average upward.
- The maximum value of 1440 disease cases is notably higher than the other values and is a potential outlier. This extreme value could be due to a specific event or anomaly not reflected in the other data points.

Investigating the date associated with this outlier (potentially 1854-12-01 or another date) is crucial to understanding its cause.

- Combining the chart with the dataset's information, we can see that the 'disease' variable is likely related to the 'wounds' and 'other' variables, although the exact relationship isn't clear without further analysis. The high variability in 'disease' cases might be explained by variations in 'wounds' and 'other' cases.

4. External Influences:

- Several external factors could influence the number of disease cases. These include:
 - **Seasonality:** Disease outbreaks often show seasonal patterns. Analyzing the 'date' variable could reveal seasonal trends.
 - **Environmental conditions:** Changes in weather, sanitation, or other environmental factors could impact disease spread.
 - **Public health interventions:** The presence or absence of public health measures (e.g., vaccination campaigns, quarantine) would significantly affect disease counts.
 - **Underlying population health:** The overall health and immunity of the population could influence susceptibility to disease.

5. Recommendations:

- **Investigate the outlier:** Determine the cause of the 1440 disease case outlier. Was there a specific event (e.g., epidemic, war) or data entry error?
- **Analyze seasonal trends:** Explore the relationship between the 'date' and 'disease' variables to identify any seasonal patterns in disease outbreaks.
- **Explore correlations:** Conduct a correlation analysis between 'disease', 'wounds', and 'other' variables to understand their interrelationships.
- **Consider additional variables:** If possible, include additional variables (e.g., population size, location, specific disease type) to improve the analysis and identify other contributing factors.

6. Conclusion:

The distribution of disease cases is highly variable and right-skewed, with a potential outlier suggesting a significant event or anomaly. Further investigation is needed to understand the causes of this variability, including exploring seasonal trends, correlations with other variables, and the influence of external factors. Addressing the outlier and conducting further analysis will provide a more complete understanding of disease patterns and inform effective public health strategies.

✳ Insight 4:

```
main() Goal Goal(question="How do the average values of 'wounds' and 'other' vary over time?", visualization="Line chart showing the rolling average of 'wounds' and 'other' over 'date'", rationale="1. **Data Type:** 'date' is datetime, 'wounds' and 'other' are numerical. 2. **Visualization Justification:** A li...
```

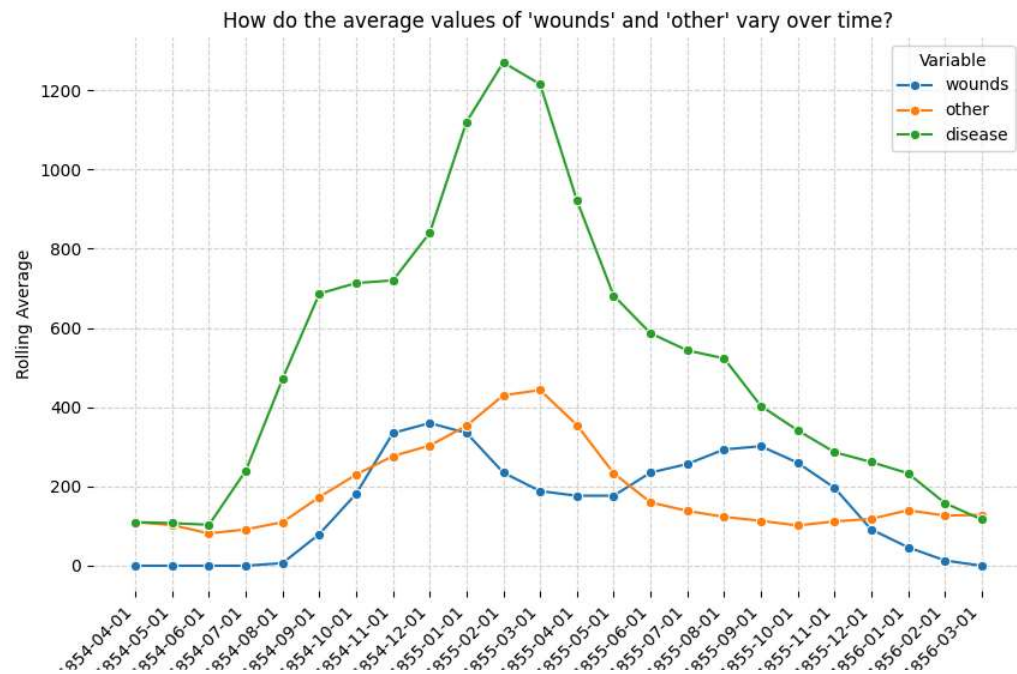
A visualization goal

index int 3


question str "How do the average values of 'wounds' and 'other' vary over time?"


rationale str "1. **Data Type:** 'date' is datetime, 'wounds' and 'other' are numerical. 2. **Visualization Justification:** A line chart effectively shows trends over time, and using a rolling average smooths out short-term fluctuations, revealing underlying trends. 3. **Justification:** This allows for a clear..."

visualization str "Line chart showing the rolling average of 'wounds' and 'other' over 'date'"



[°*↯••?↯ Download Chart °*](#)

 VizOps ▾

 Analyze Chart 4

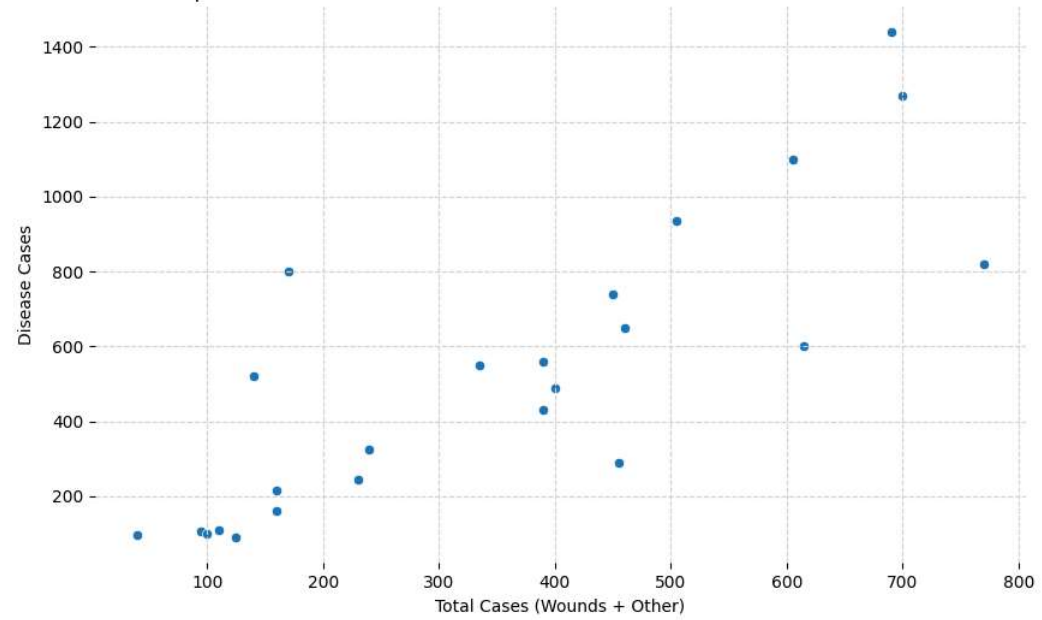
✱ Insight 5:

```
main() Goal Goal(question="Is there a relationship between the total number of 'wounds' and 'other' cases and the number of 'disease' cases?", visualization="Scatter plot of ('wounds' + 'other') vs. 'disease'", rationale="1. **Data Type:** 'wounds', 'other', and 'disease' are numerical. 2. **Visualization Justi...
```

A visualization goal

index	int	4
question	str	"Is there a relationship between the total number of 'wounds' and 'other' cases and the number of 'disease' cases?"
rationale	str	"1. **Data Type:** 'wounds', 'other', and 'disease' are numerical. 2. **Visualization Justification:** A scatter plot is used to visualize the relationship between two numerical variables. Here, we sum 'wounds' and 'other' to create a single variable representing the total number of cases. 3. **Jus...
visualization	str	"Scatter plot of ('wounds' + 'other') vs. 'disease'"

Is there a relationship between the total number of 'wounds' and 'other' cases and the number of 'disease' cases?



°*↵•••?↵ Download Chart °*

⚙ VizOps ▾

🔍 Analyze Chart 5