

Data Analysis Report

Dataset Summary:

Number of Rows: 1000

Number of Columns: 14

Insights:

Detailed Insights

The provided dataset contains 1000 rows and 14 columns, representing various patient characteristics and medical measurements. The analysis of the summary statistics and correlation matrix reveals several key trends, anomalies, and correlations.

Summary Statistics Insights

- **Age****: The average age of the patients is 49.24 years, with a standard deviation of 17.86 years. The age range is between 20 and 80 years.
- **Gender****: The majority of the patients are male (76.5%), with a mean value of 0.765.
- **Chest Pain****: The average chest pain level is 0.98, with a standard deviation of 0.95. The chest pain levels range from 0 to 3.
- **Resting Blood Pressure****: The average resting blood pressure is 151.75 mmHg, with a standard deviation of 29.97 mmHg. The resting blood pressure range is between 94 and 200 mmHg.
- **Serum Cholesterol****: The average serum cholesterol level is 311.45 mg/dL, with a standard deviation of 132.44 mg/dL. The serum cholesterol range is between 0 and 602 mg/dL.

Correlation Matrix Insights

- **Strong Positive Correlations****:
 - * Chest pain and resting blood pressure (0.222)
 - * Chest pain and serum cholesterol (0.175)
 - * Resting blood pressure and serum cholesterol (0.127)
 - * Slope and number of major vessels (0.527)
 - * Slope and target (0.797)
- **Strong Negative Correlations****:
 - * Gender and serum cholesterol (-0.121)
 - * Exercise angina and chest pain (-0.027)
- **Moderate Positive Correlations****:
 - * Age and fasting blood sugar (0.050)
 - * Resting blood pressure and fasting blood sugar (0.192)
 - * Serum cholesterol and fasting blood sugar (0.278)
- **Moderate Negative Correlations****:
 - * Age and max heart rate (-0.040)

* Gender and resting blood pressure (-0.070)

****Anomalies and Outliers****

1. ****Patient ID****: The patient ID column has a large range (103368 to 9990855), which may indicate that the patient IDs are not consecutive or that there are missing values.
2. ****Serum Cholesterol****: The serum cholesterol column has a large range (0 to 602 mg/dL), which may indicate that there are outliers or errors in the data.
3. ****Fasting Blood Sugar****: The fasting blood sugar column has a large number of zeros (75th percentile is 0), which may indicate that many patients have normal fasting blood sugar levels.

****Data-Driven Actions****

1. ****Targeted Interventions****: Patients with high chest pain levels, high resting blood pressure, and high serum cholesterol levels may benefit from targeted interventions to reduce their risk of cardiovascular disease.
2. ****Screening and Monitoring****: Patients with high fasting blood sugar levels, high resting blood pressure, and high serum cholesterol levels may require regular screening and monitoring to prevent the development of cardiovascular disease.
3. ****Lifestyle Modifications****: Patients with high risk factors for cardiovascular disease may benefit from lifestyle modifications, such as diet and exercise changes, to reduce their risk.

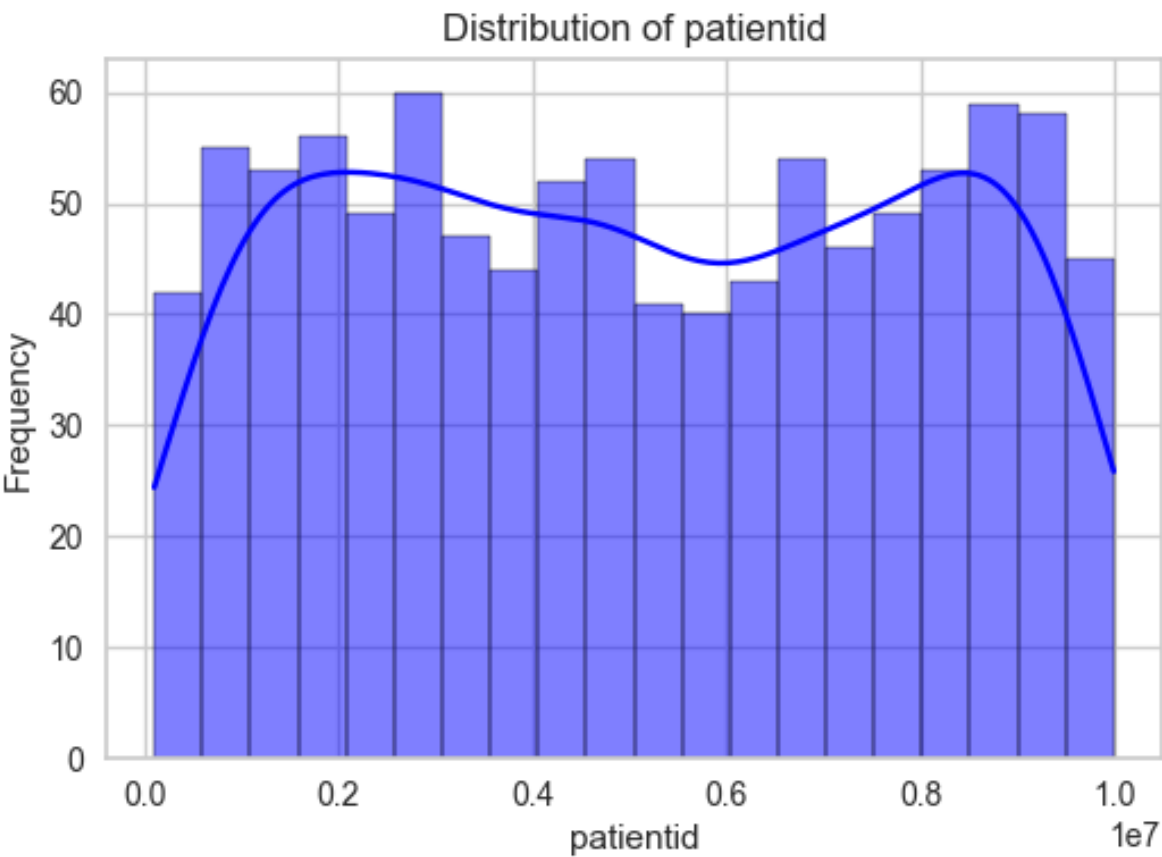
****PDF Report****

A PDF report can be generated to summarize the key findings and insights from the analysis. The report can include:

1. ****Introduction****: A brief introduction to the dataset and the analysis objectives.
2. ****Summary Statistics****: A summary of the key statistics for each column, including means, standard deviations, and ranges.
3. ****Correlation Matrix****: A summary of the correlation matrix, highlighting strong positive and negative correlations.
4. ****Anomalies and Outliers****: A discussion of the anomalies and outliers identified in the data.
5. ****Data-Driven Actions****: A summary of the data-driven actions recommended based on the analysis.
6. ****Conclusion****: A conclusion summarizing the key findings and insights from the analysis.

The PDF report can be used to communicate the results of the analysis to stakeholders and to inform decision-making related to patient care and cardiovascular disease prevention.

Statistical Graphs:



Feature Correlation Heatmap

