Medical Data Visualizer - Project Report

Project Overview

This project is part of the **freeCodeCamp Data Analysis with Python Certification**.
It focuses on analyzing and visualizing medical examination data using **Pandas**, **Matplotlib**, and **Seaborn**.

Dataset

File name: `medical_examination.csv`
Source: freeCodeCamp / UCI Machine Learning Repository

Each row represents a patient and includes features like:

- Age (in days)
- Height (cm)
- Weight (kg)
- Blood pressure (`ap_hi`, `ap_lo`)
- Cholesterol and glucose levels
- Lifestyle (smoke, alco, active)
- Cardio diagnosis (target variable)

Analysis Performed

- **Added BMI & Overweight Column**
  Calculated BMI and labeled patients as overweight (`1`) if BMI > 25.
- **Normalized Cholesterol & Glucose**
  Re-labeled as `0` for normal, and `1` for above normal levels.
- **Generated Categorical Plot**
  Compared lifestyle features between people with and without cardiovascular disease.
  Variables used: `cholesterol`, `gluc`, `smoke`, `alco`, `active`, `overweight`.
- **Cleaned Data for Heatmap**
  Removed:
  - Invalid blood pressure values
  - Height and weight outliers (below 2.5% or above 97.5% quantiles)

- **Generated Heatmap**
  Visualized the correlation matrix between all numerical features.

📈 Output

📊 Categorical Plot

Shows the distribution of health/lifestyle factors grouped by the presence of heart disease (`cardio = 0 or 1`).

🔥 Heatmap

Displays the correlation between medical variables such as BMI, age, blood pressure, cholesterol, etc.

🛠️ Technologies Used

- Python 3
- Pandas
- Matplotlib
- Seaborn
- Spyder IDE

👩‍💻 Files in the Repository

File	Description
`medical_examination.csv`	Dataset
`medical_data_visualizer.py`	Main logic and analysis
`main.py`	Script to test the code
`README.md`	Project summary
`medical_visualizer_report.pdf` (optional)	PDF version of this report

✅ Certification

This project contributes to:
freeCodeCamp – Data Analysis with Python Certification

How to Run

To execute the project, run:

```
bash
CopyEdit
python main.py
```

Make sure the dataset and script files are in the same folder.

OUDHA BOCHRA