



Bharatiya Vidya Bhavan's  
**SARDAR PATEL INSTITUTE OF TECHNOLOGY**

## **Advanced Data Visualization**

### **Experiment no. 3**

#### **Submitted To**

Prof. Pranav Nerurkar

#### **Submitted By**

Name : Mohit Narwaiye

Class : COMPS B

Batch : A

UID : 2021300085



## 1. Dataset

Can view the dataset here

<https://www.kaggle.com/datasets/redwankarimsony/heart-disease-data>

## 2. Description

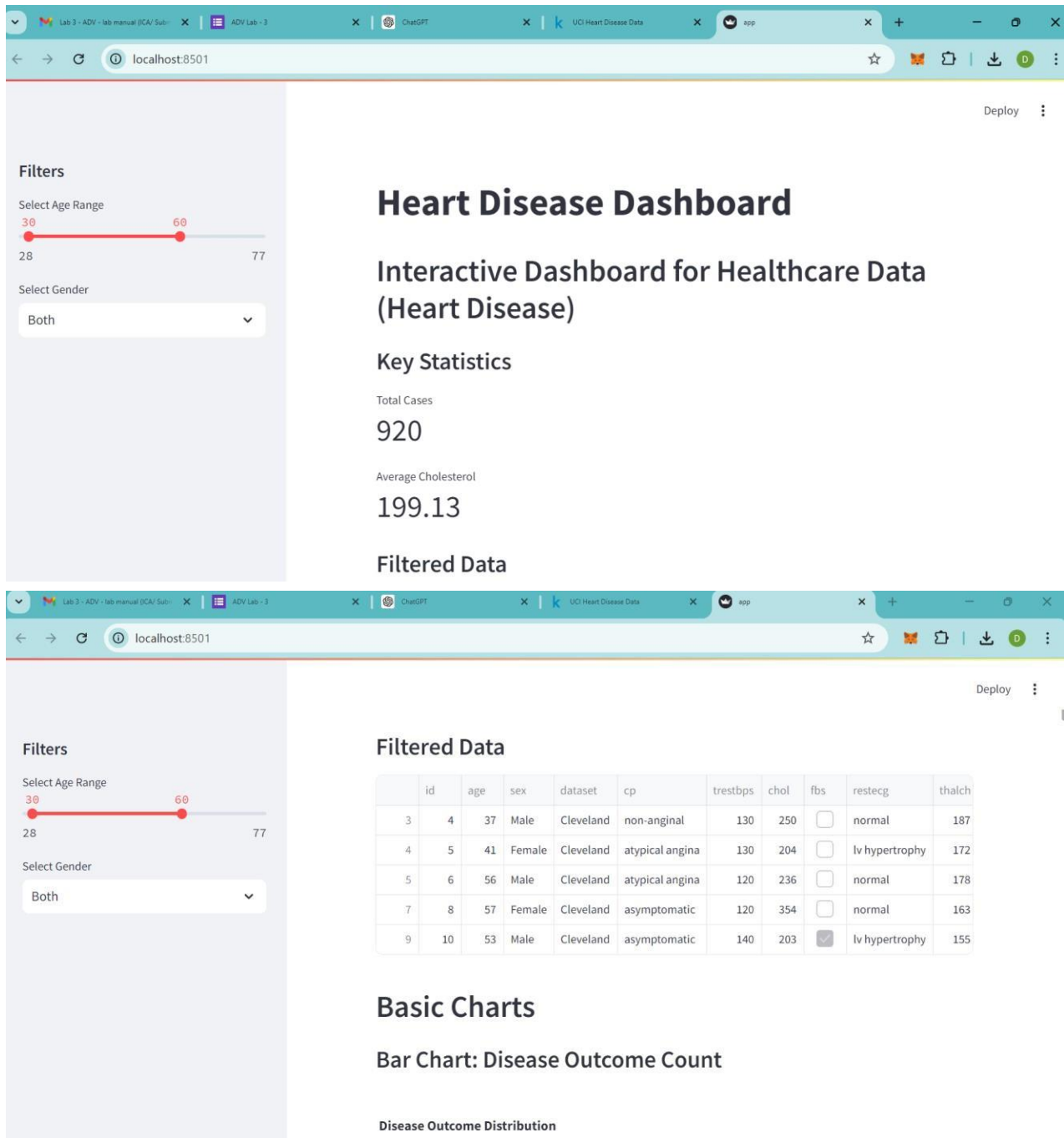
This is a multivariate type of dataset which means providing or involving a variety of separate mathematical or statistical variables, multivariate numerical data analysis. It is composed of 14 attributes which are age, sex, chest pain type, resting blood pressure, serum cholesterol, fasting blood sugar, resting electrocardiographic results, maximum heart rate achieved, exercise-induced angina, oldpeak — ST depression induced by exercise relative to rest, the slope of the peak exercise ST segment, number of major vessels and Thalassemia. This database includes 76 attributes, but all published studies relate to the use of a subset of 14 of them. The Cleveland database is the only one used by ML researchers to date. One of the major tasks on this dataset is to predict based on the given attributes of a patient that whether that particular person has heart disease or not and other is the experimental task to diagnose and find out various insights from this dataset which could help in understanding the problem more.

## 3. Metadata

1. id (Unique id for each patient)
2. age (Age of the patient in years)
3. origin (place of study)
4. sex (Male/Female)
5. cp chest pain type ([typical angina, atypical angina, non-anginal, asymptomatic])
6. trestbps resting blood pressure (resting blood pressure (in mm Hg on admission to the hospital))
7. chol (serum cholesterol in mg/dl)
8. fbs (if fasting blood sugar > 120 mg/dl)
9. restecg (resting electrocardiographic results)  
-- Values: [normal, stt abnormality, lv hypertrophy]
10. thalach: maximum heart rate achieved
11. exang: exercise-induced angina (True/ False)
12. oldpeak: ST depression induced by exercise relative to rest
13. slope: the slope of the peak exercise ST segment
14. ca: number of major vessels (0-3) colored by fluoroscopy
15. thal: [normal; fixed defect; reversible defect]
16. num: the predicted attribute

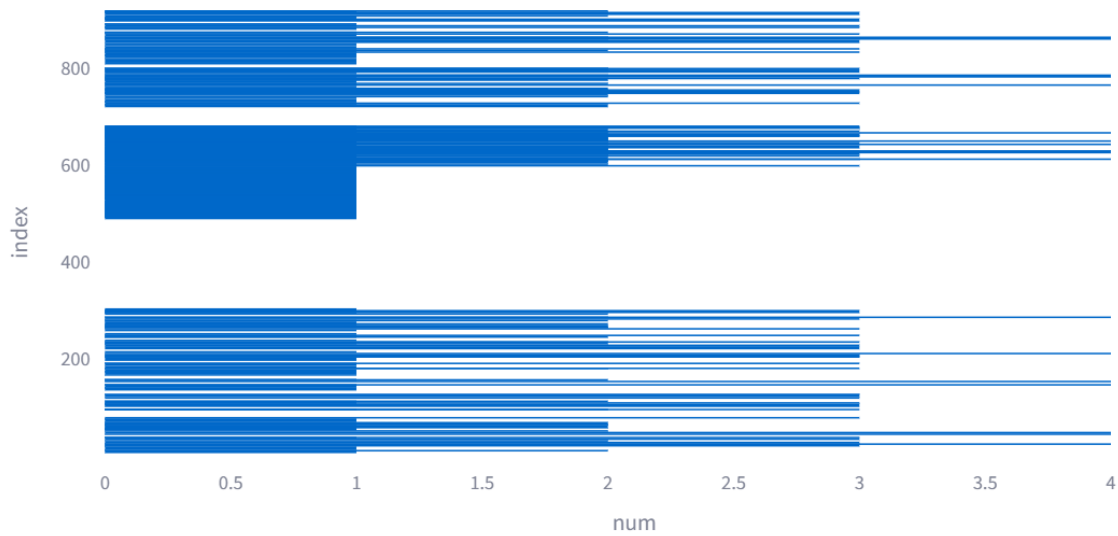


## 4. Visualizations and Observations





**Disease Outcome Distribution**

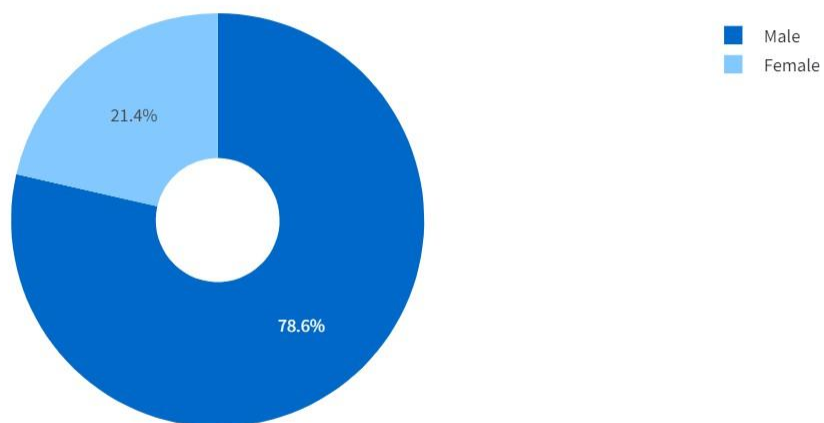


Observation: The bar chart shows the distribution of heart disease outcomes.

## Pie Chart: Gender Distribution



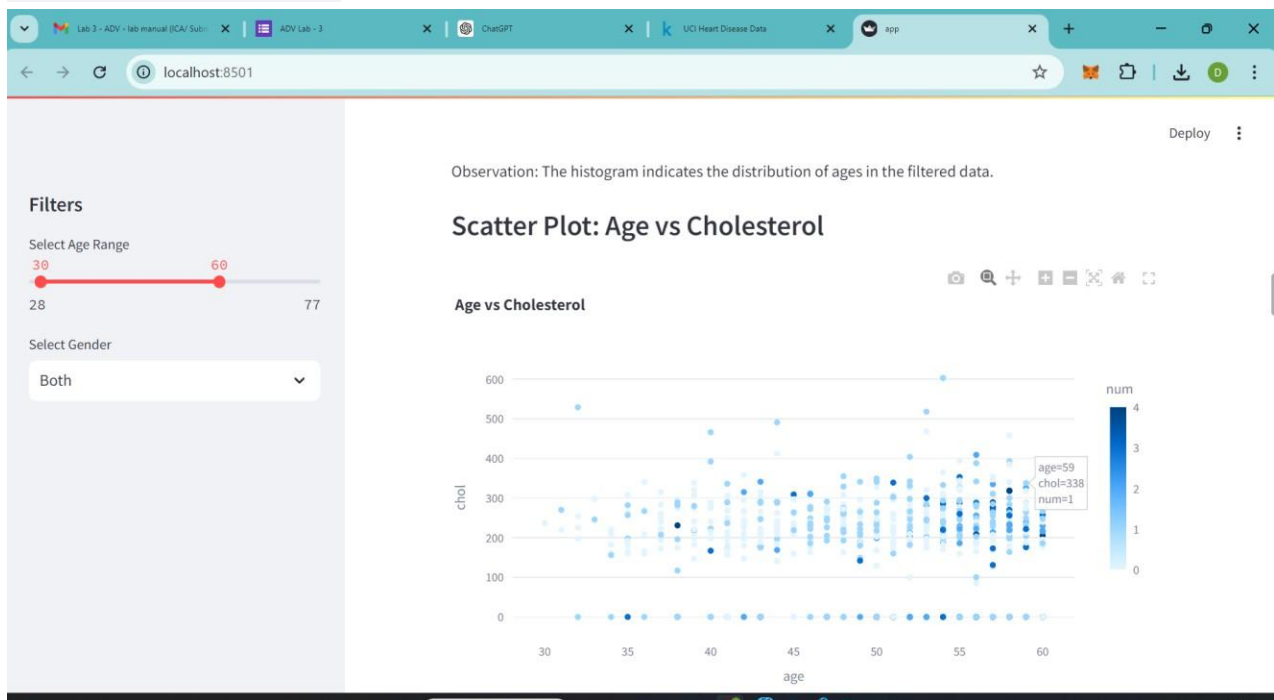
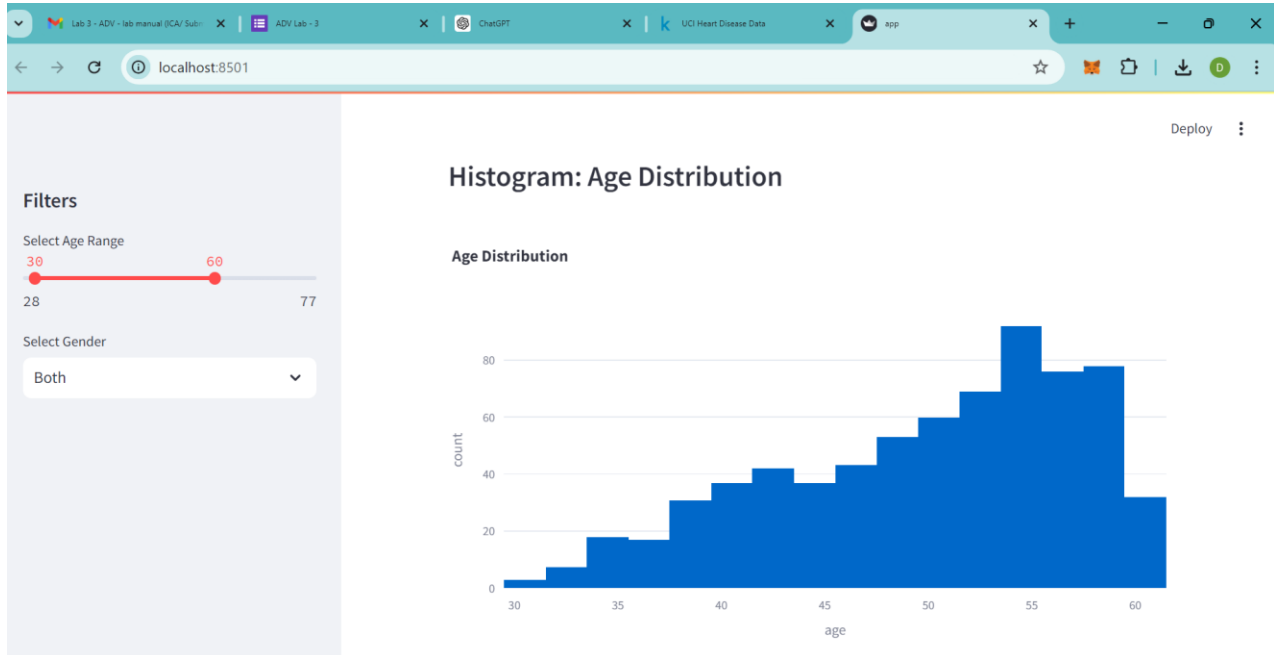
**Gender Distribution**



Observation: This pie chart shows the gender distribution of patients.

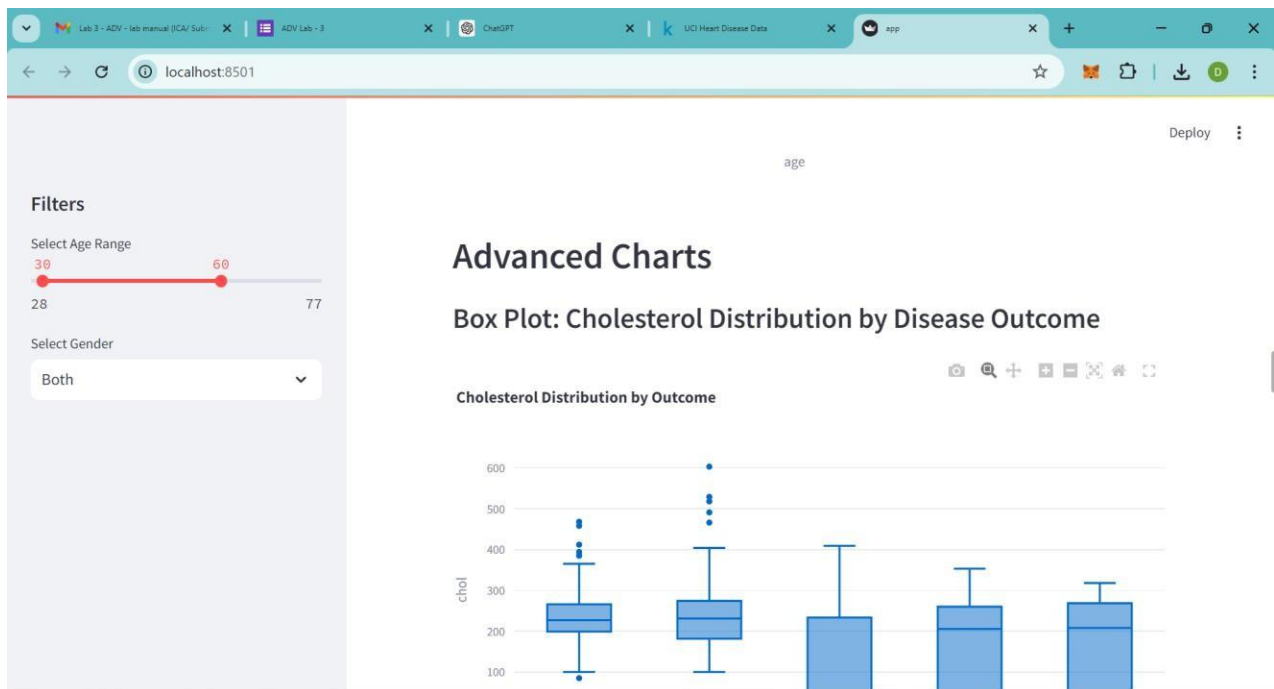
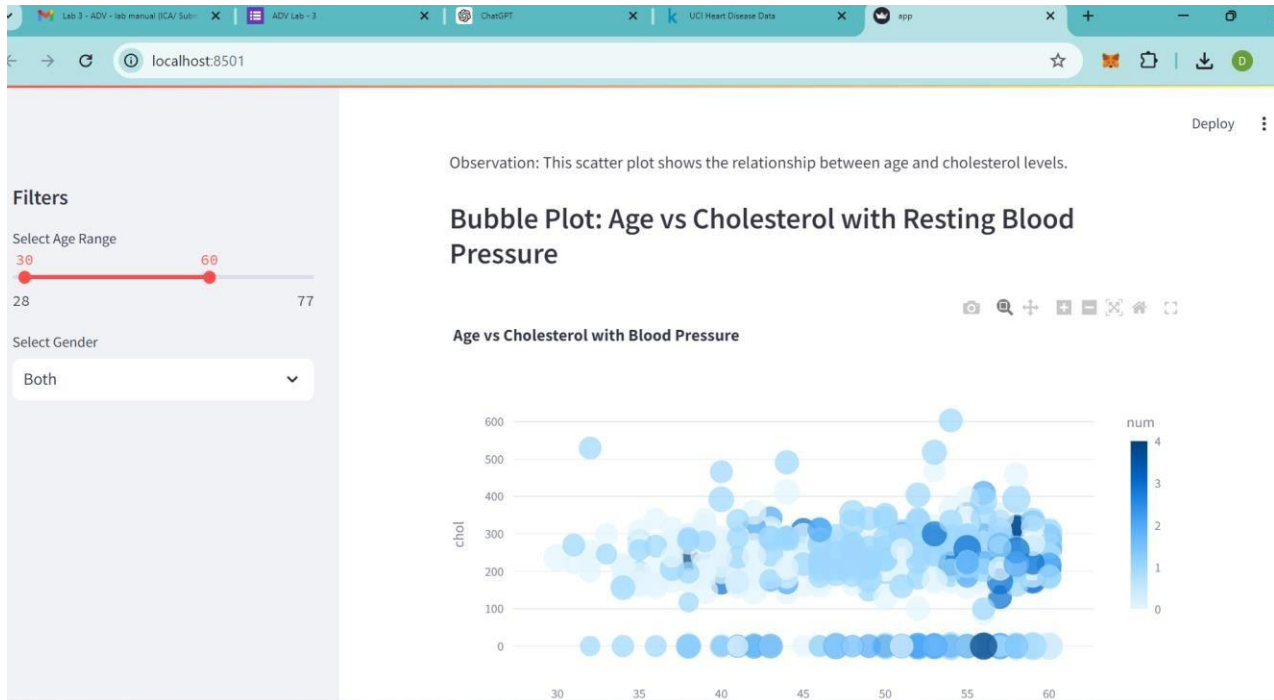


Bharatiya Vidya Bhavan's  
**SARDAR PATEL INSTITUTE OF TECHNOLOGY**



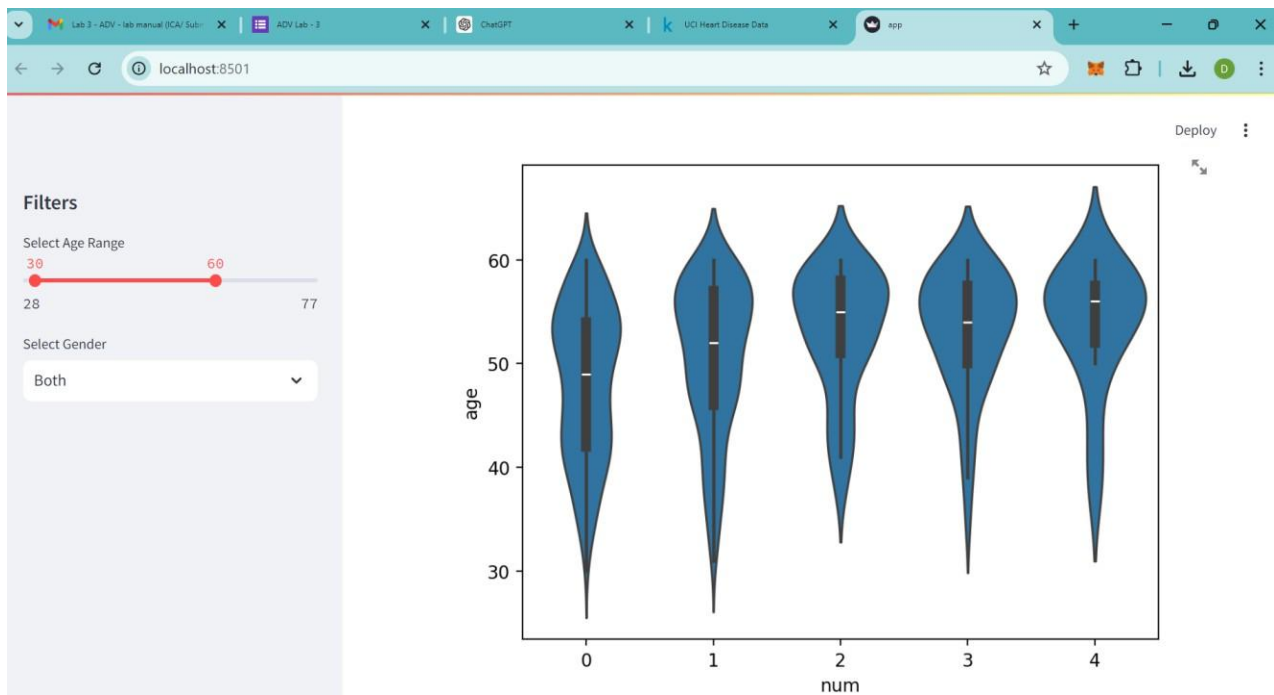
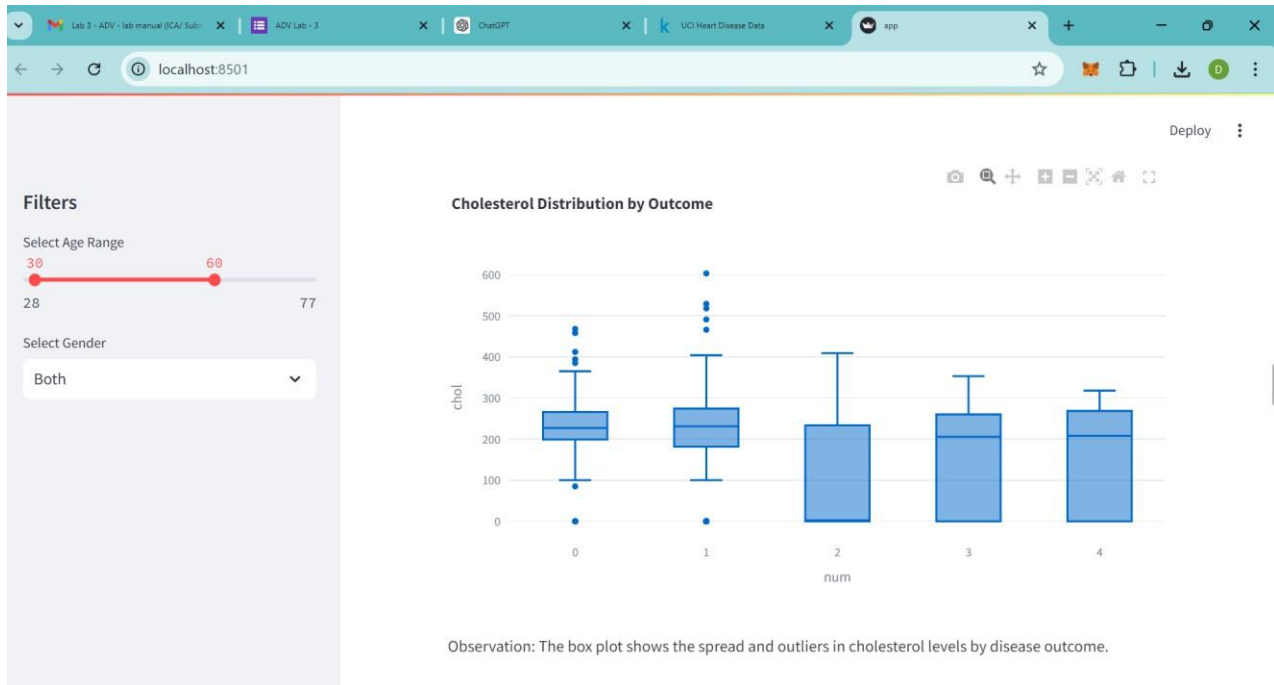


Bharatiya Vidya Bhavan's  
**SARDAR PATEL INSTITUTE OF TECHNOLOGY**



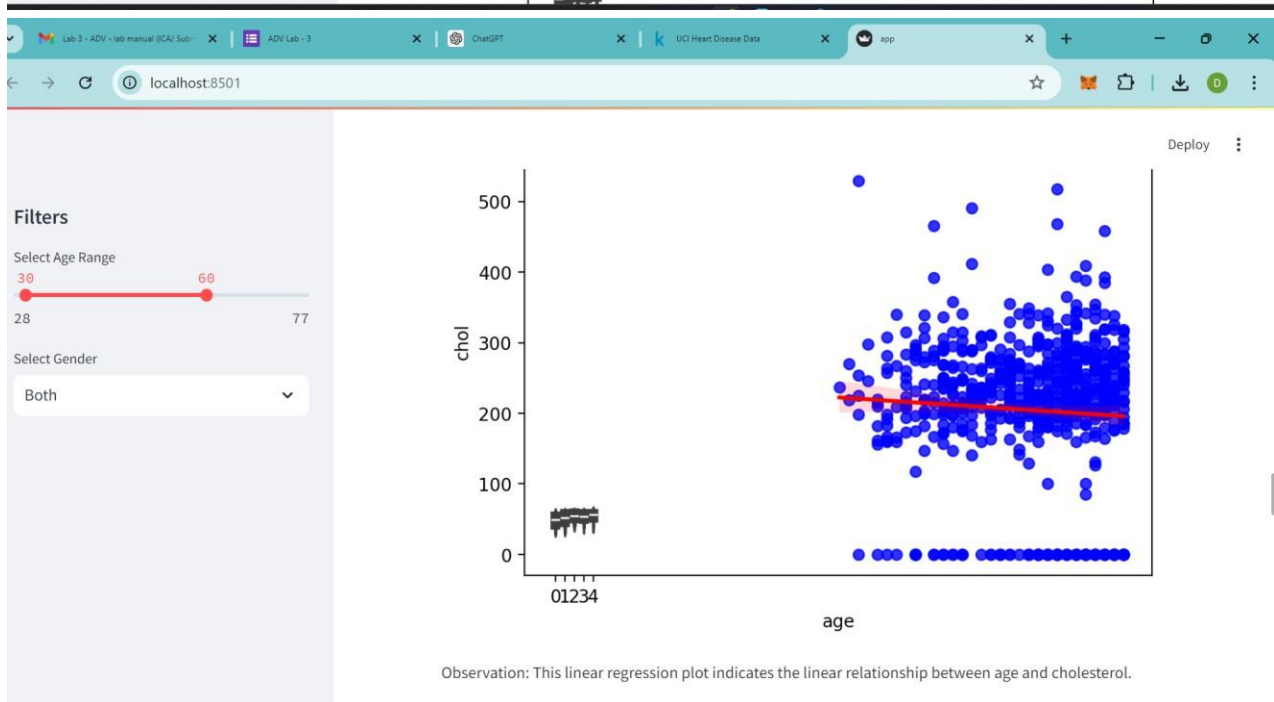


# Bharatiya Vidya Bhavan's SARDAR PATEL INSTITUTE OF TECHNOLOGY

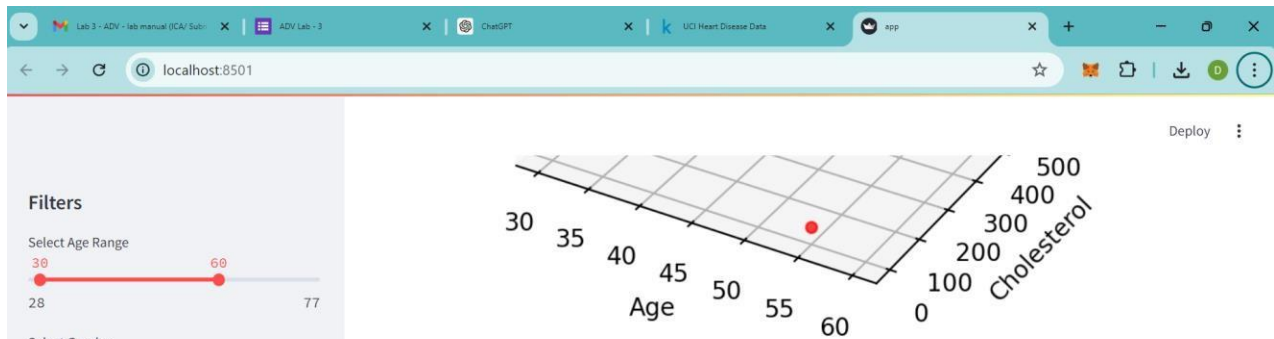
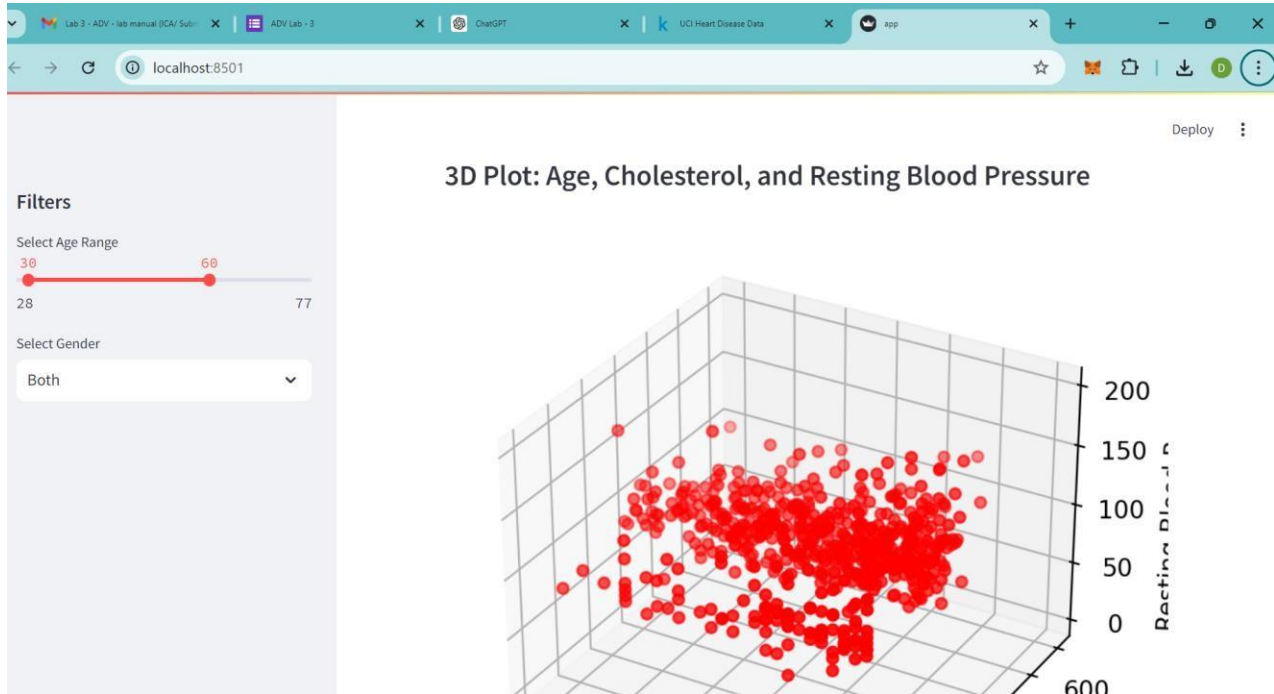




Bharatiya Vidya Bhavan's  
**SARDAR PATEL INSTITUTE OF TECHNOLOGY**





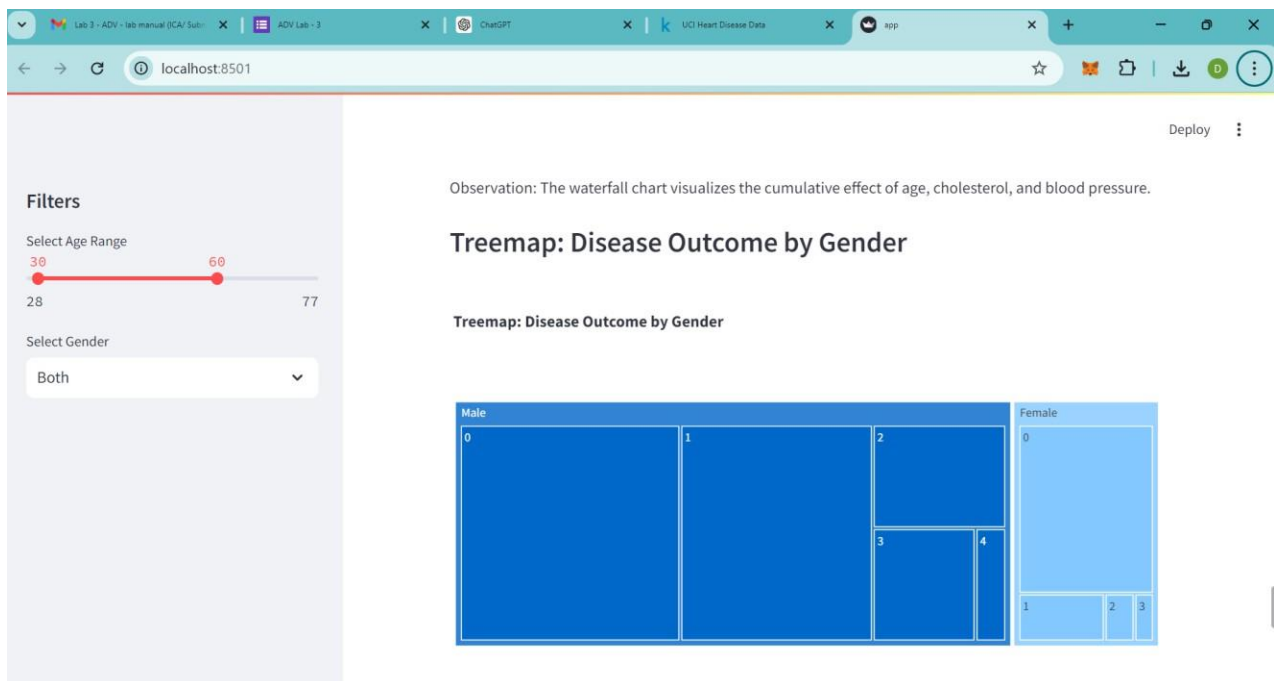
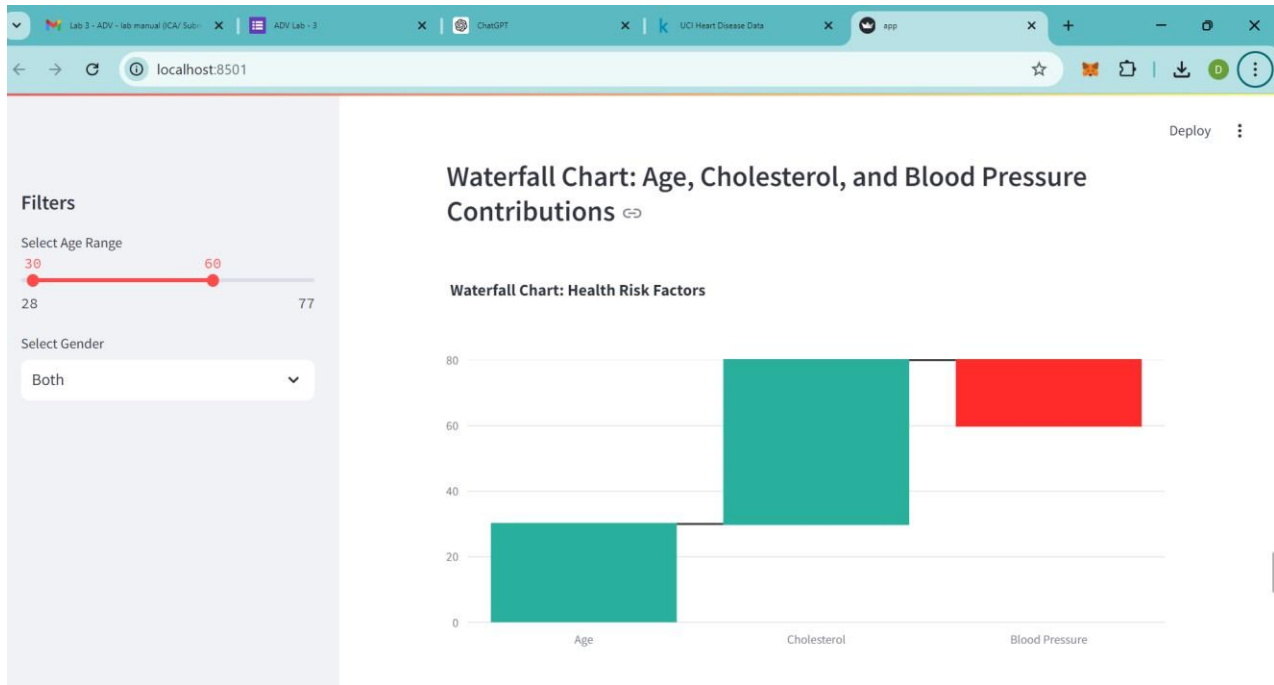


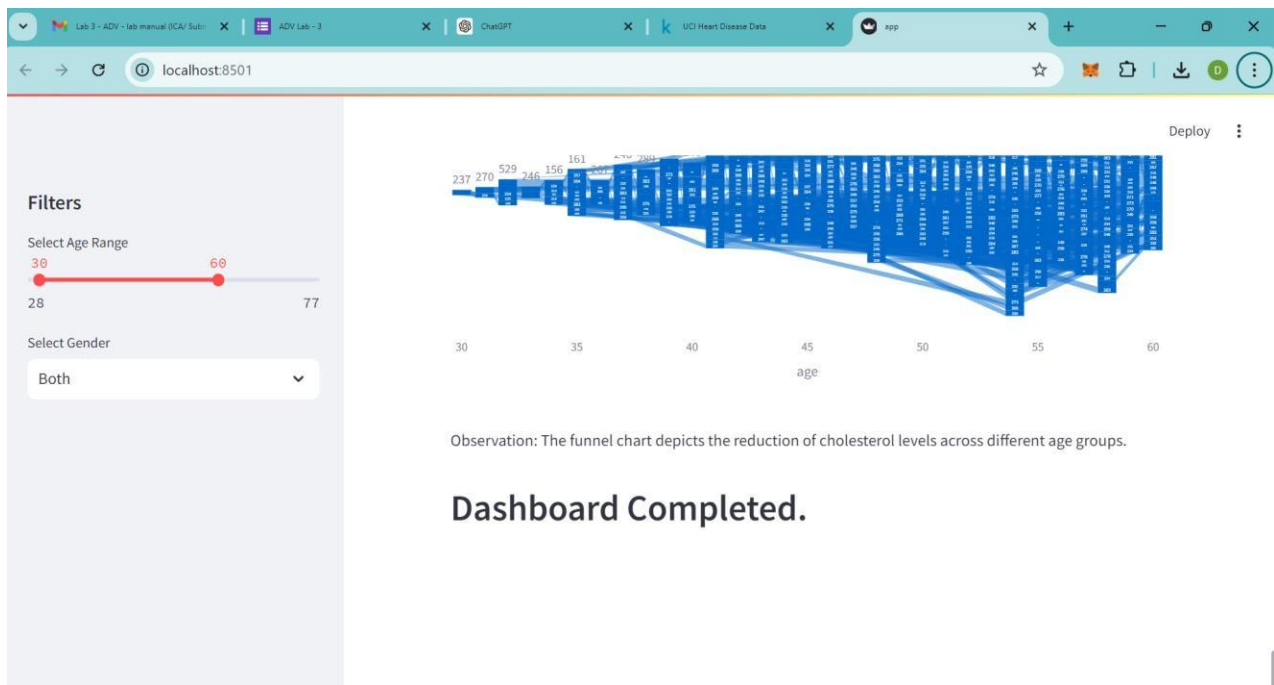
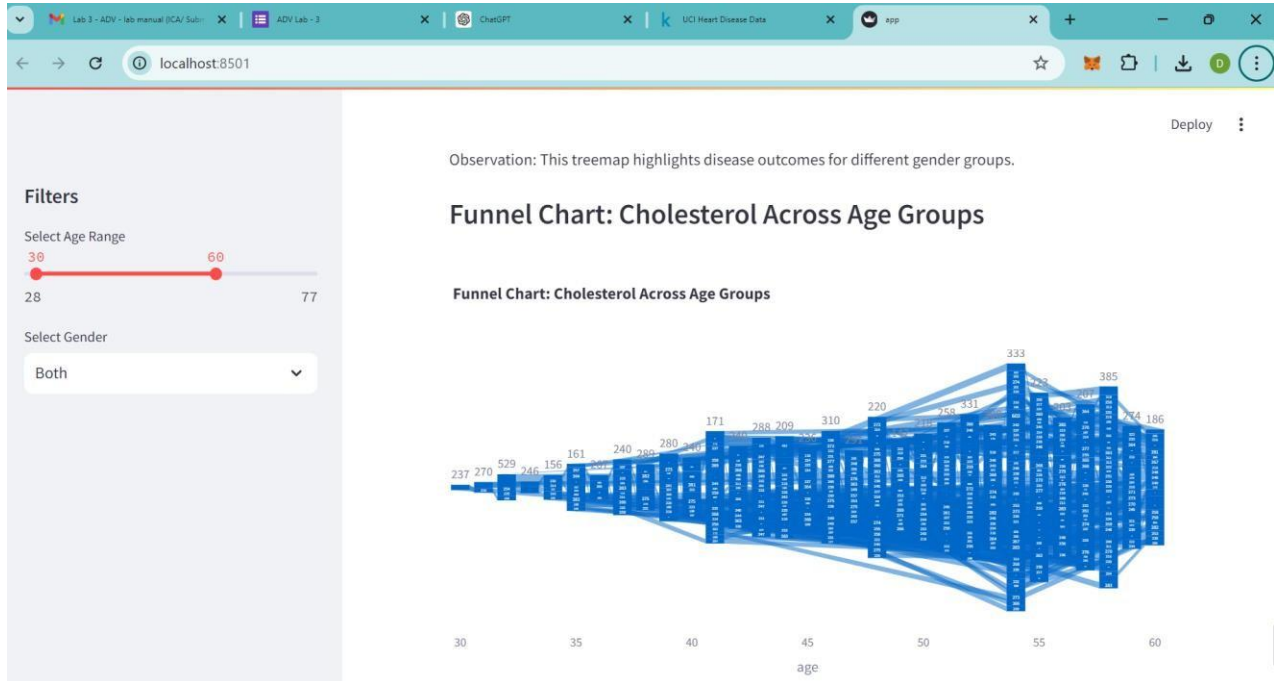
Observation: The 3D chart helps visualize the relationship between age, cholesterol, and blood pressure.

### Waterfall Chart: Age, Cholesterol, and Blood Pressure Contributions

**Waterfall Chart: Health Risk Factors**







## 5. Source Code

[https://github.com/dmo-27/ADV\\_experiments/blob/main/app.py](https://github.com/dmo-27/ADV_experiments/blob/main/app.py)