Lab 05: Data App
CS3300 Data Science

Learning Outcomes

1. Understand the basic process of data science and exploratory data analysis including modes of inquiry (hypothesis driven, data driven, and methods driven).

2. Identify, access, load, and prepare (clean) a data set for a given problem.

3. Select, apply, and interpret appropriate visual and statistical methods to analyze distributions of individual variables and relationships between pairs of variables.

4. Communicate findings through generated data visualizations and reports.

Overview

As part of their roles, data scientists may create "data apps" that enable non-technical users like business analysists to analyze data on their own. Data apps use interactive visualizations and simple widgets to support basic queries and filtering. Dashboards are a special type of data app that sources data from a database or other live source enabling real-time monitoring of business data.

Bokeh (Python) and Shiny (R) are open-source tools for creating data apps that are commonly used by data scientists. Commercially-supported dashboard tools include Tableau and Microsoft Power BI. Bokeh apps often generate single-page web apps saved as a HTML file. The logic for constructing the app is written in Python and converted to HTML and JavaScript by the Bokeh library.

In this lab, you are going to use Bokeh to create a data app to explore the real estate data you worked with previously.

Bokeh's documentation:
https://bokeh.pydata.org/en/latest/index.html

Instructions

Your app should be developed as a single Python script. The script should take the CSV file as input and return a single-page web app (HTML file). You will not use a Jupyter notebook for this.

Your application should:

1. Display Real Estate on a Scatter Plot

a. Use the latitudes and longitudes to create a scatter plot. Label the axes.

b. Color the points by residence type.

c. When the user puts their mouse over a point, display the address, price, square footage, number of beds, and number of baths as a tooltip.

2. Add Widgets to Filter Records

a. Add widgets to filter records by ranges of price, square footage, number of beds, and number of baths

b. Add a widget to filter records by type

c. Link these widgets to the scatter plot so that only records that match the filters are displayed

3. Display Records Data

a. Create either a table of the records or an area to view a single selected record

b. If using the table, only display the records that match the filters

c. If using a single record, allow the user to click on a point in the scatter plot to select that record for display.

4. Use the `output_file()` function to save the app as a HTML file.

Submission Instructions

Submit a single zip file to Blackboard containing your .py script and your generated HTML file.

Rubric

Followed submission instructions (especially HTML file)	10%
Real Estate Location Plot	25%
Widgets and Correct Filtering Logic	50%
Display of Record Data	15%