Course title:CISC3015 - Data and Information Visualization (2024-1 001)

Coursework item: Revised Project Proposal

Pair Number: 05

Student ID: DC229805; DC427182

Student name: Zhong Wanyi; Tailine

Q1. State the region and the domain category that your data sets are about.

# **City of Los Angeles in United States and crime.**

Q2. State a question about the domain category and region that you identified as being interesting.

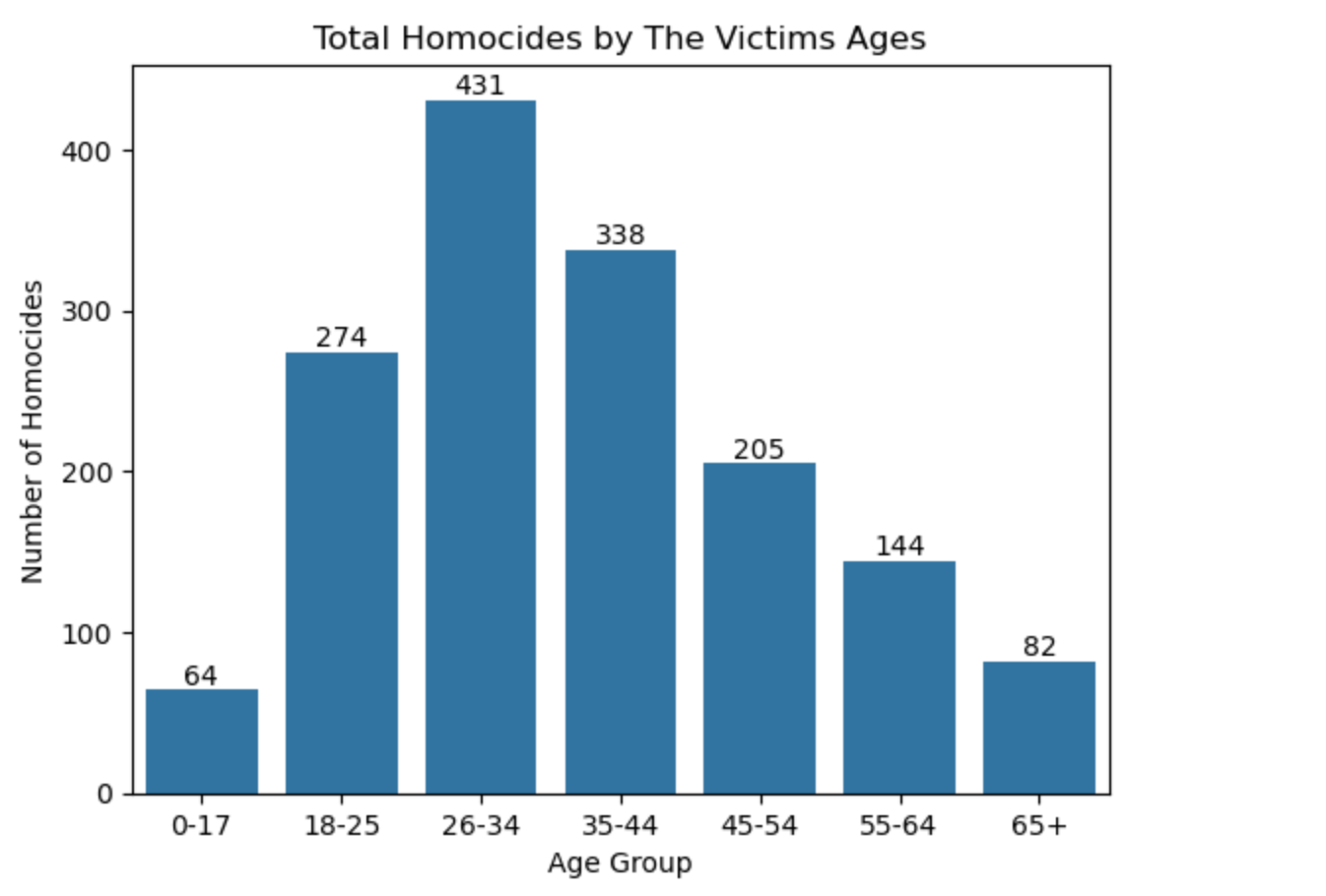
**Is criminal homicide related to gender or age of the victim in Los Angeles from 2020 to 2024 ?**

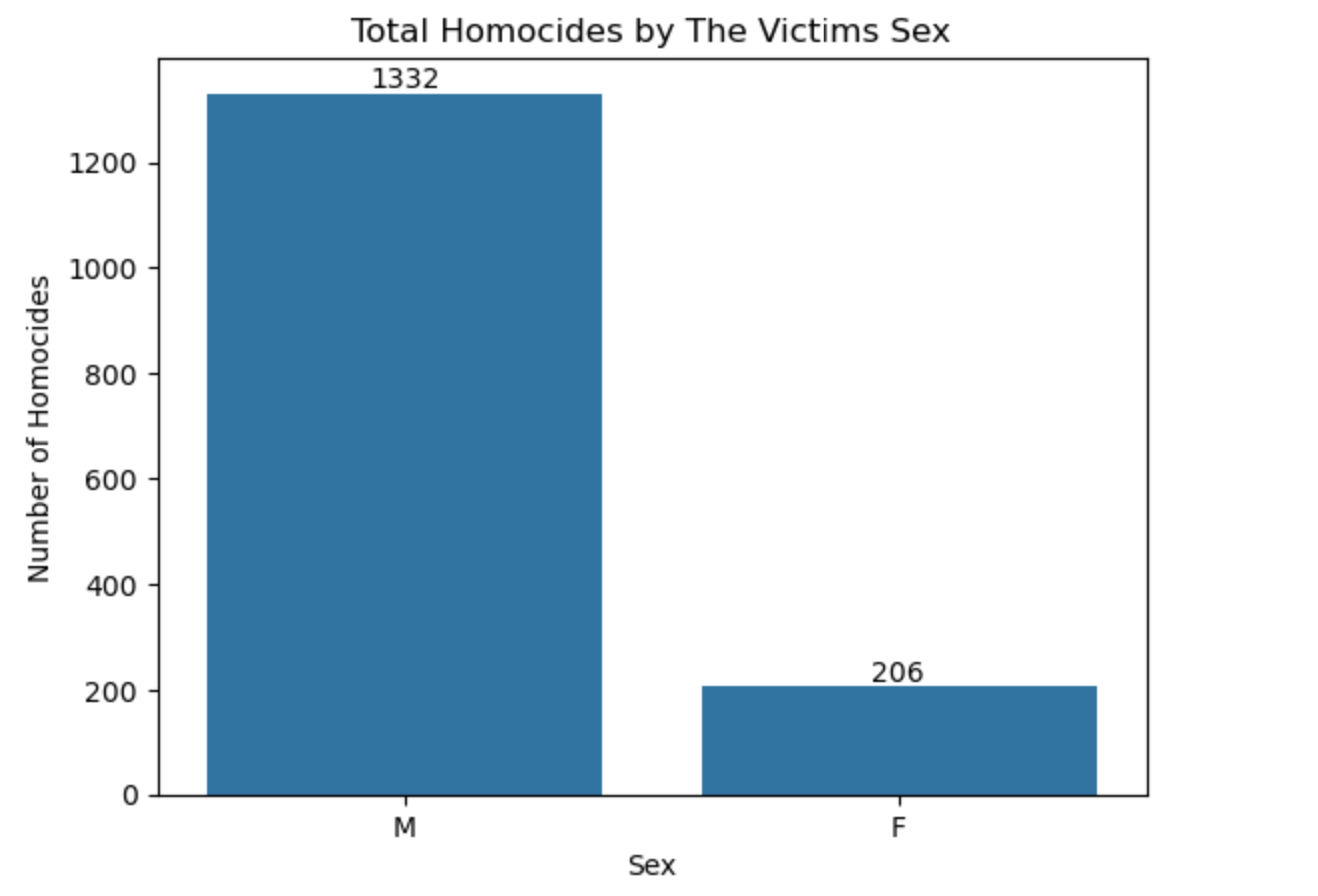
Q3. Provide at least two links to available datasets.

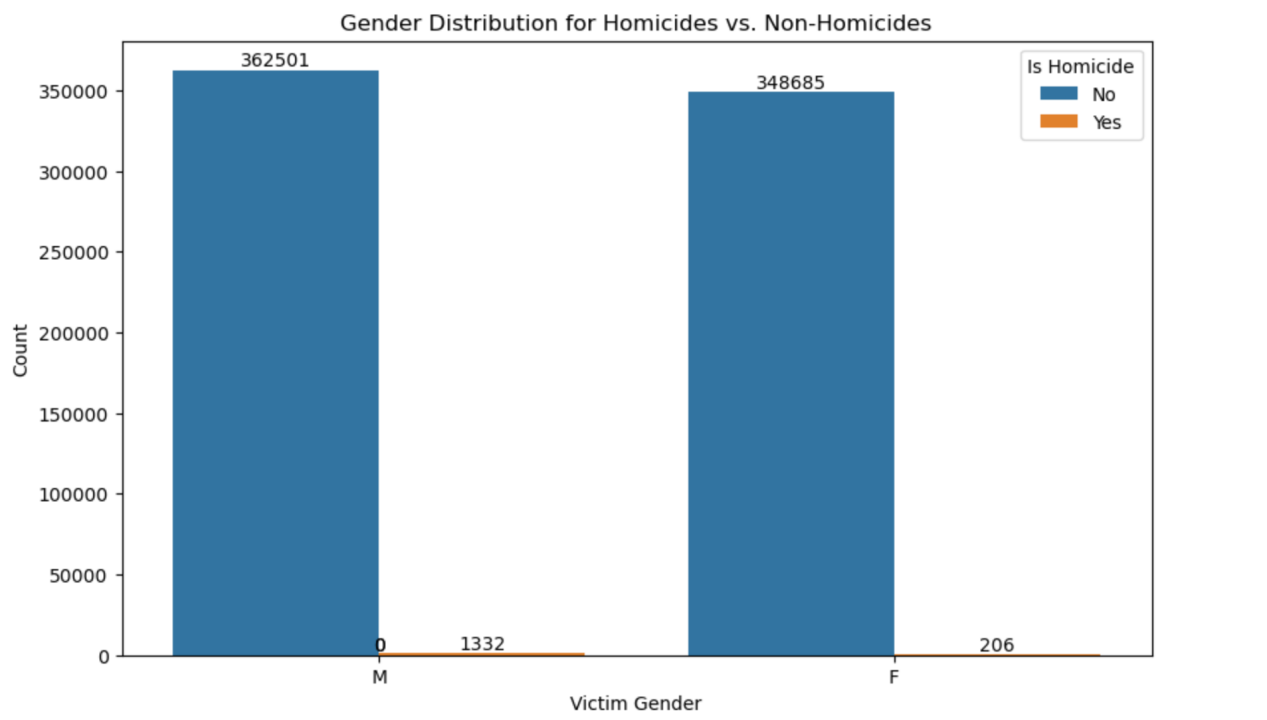
**CSVFile**:<https://catalog.data.gov/dataset/crime-data-from-2020-to-present/resource/5eb6507e-fa82-4595-a604-023f8a326099>

**RDFFile**:<https://catalog.data.gov/dataset/crime-data-from-2020-to-present/resource/2957fec8-d080-46df-995b-d55287d9c24b>

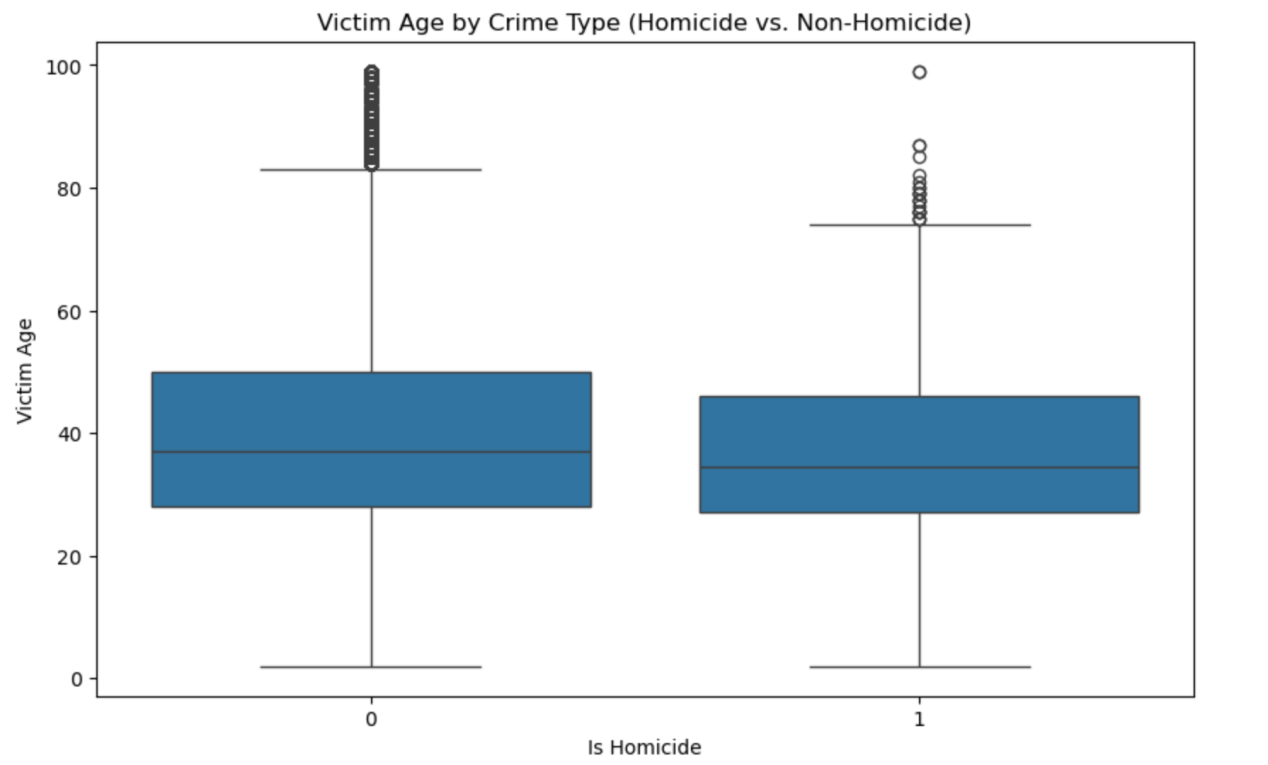
Q4.

**Countplot Visualization:**  






**Boxplot Visualization:**



Q5. Short (1-2 paragraph) written justification of how your visualization addresses your stated research question.

We choose **countplot and boxplot visualizations** to tackle our issue. Countplot is a simple and effective way to show the distribution of categorical data such as gender and age, by dividing them to different gender groups (male group vs female group) and age groups (age 0-17, age 18-25, age 26-34, age 35-44, age 45-54, age 55-64 and age above 65), it can obvious to make an significant comparison across different groups visually. For example, according to above visualizations, we can know that **the age group of 26-34** has the highest number of victims, with a significantly higher number of **male** **victims** compared to female victims. At the same time, boxplot can also provide a concise summary of the distribution of being a homicide or not among age.