CIS 5400 – Topics in Comput Info Sys. Topic: Data Analysis Methods

Spring 2020

Instructor: Fitzroy Nembhard

Homework 9: Visualization Tools (Matplotlib and Seaborn))

Total Points: 35

Date Assigned: Wednesday, Apr 22, 2020

Due Date: Wednesday, Apr 29, 2020

Submission Instructions: Please submit your work on Canvas as a Jupyter Notebook `ipynb` file named `yourname_cis5400hw9.ipynb`. Make sure to add comments/headings to your code showing especially the question numbers.

Key Data Analysis Methods Demonstrated

- Summarizing and Visualizing Data
  - Grouping data by a certain feature
  - Creating charts using Matplotlib and Seaborn

Locate the dataset named **guns.csv** on Canvas and use it to answer the following questions. This dataset is from FiveThirtyEight's Gun Deaths in America project[*], which includes information about gun-death in the US from 2012-2014.

Data Analysis & Visualization

1. (5 points) Given the following key for **educational status** of the victim, compute the number of female victims that graduated from college:
   1: Less than High School
   2: Graduated from High School or equivalent
   3: Some College
   4: At least graduated from College
   5: Not available

2. (5 points) Create a Matplotlib bar chart with Seaborn darkgrid style showing the total number of males vs females who died by gun-related suicide.

3. (10 points) Create a pie chart showing the percentage of deaths by race per 100,000 people. Remove the percentage label from slices with less than 8%.
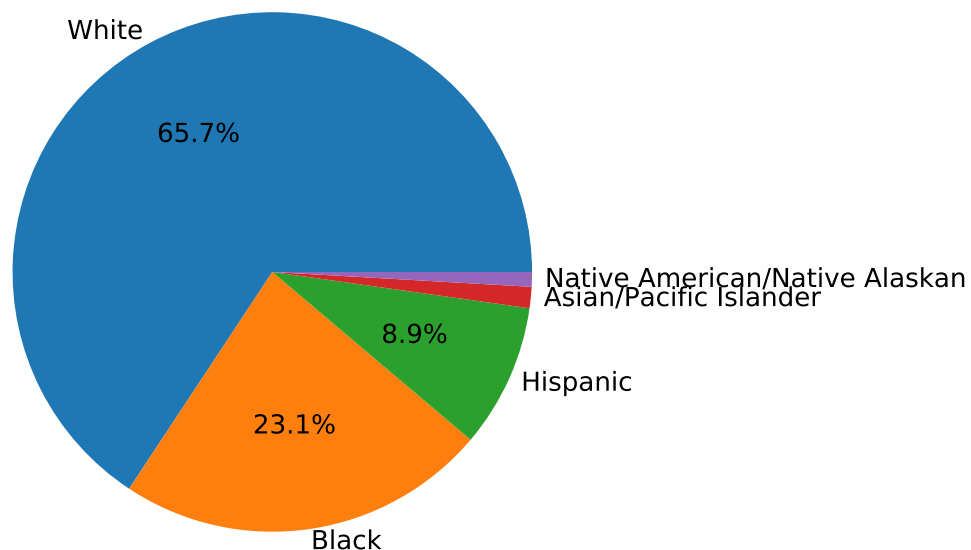

Figure 1: Expected pie chart for Question 3

4. (5 points) Plot a Seaborn bar chart with magma colored palette showing the number of gun deaths in the United States from 2012 to 2014 based on intent.
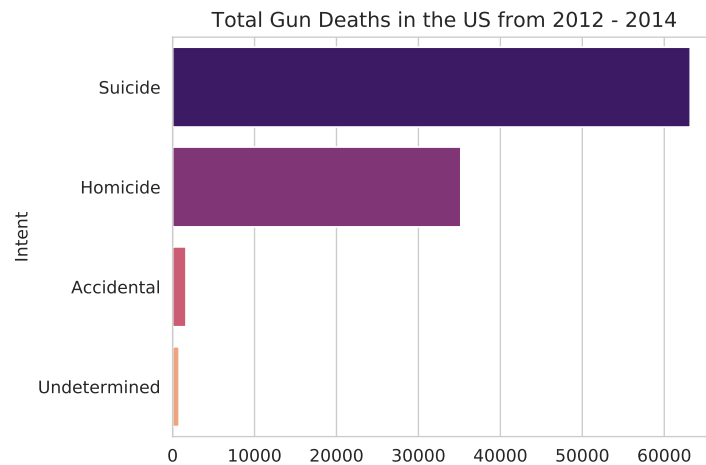

Figure 2: Expected bar chart for Question 4

5. (10 points) Plot an exploded donut (pie chart) showing the percentage of gun deaths in the United States per 100,000 people from 2012 to 2014 based on intent. (Hint: a donut is a pie chart with a circle added in the midde.)
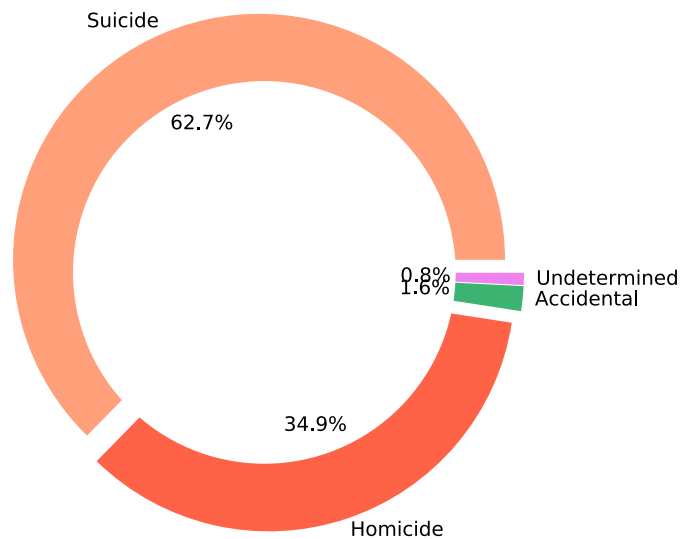

Figure 3: Expected donut for Question 5