|  |  |
| --- | --- |
| Gerb-BMSTU_01 | **Министерство науки и высшего образования Российской Федерации**  **Федеральное государственное бюджетное образовательное учреждение**  **высшего образования**  **«Московский государственный технический университет**  **имени Н.Э. Баумана**  **(национальный исследовательский университет)»**  **(МГТУ им. Н.Э. Баумана)** |

ФАКУЛЬТЕТ **Информатика, ИСКУССТВЕННЫЙ ИНТЕЛЛЕКТ и системы  
 управления**

КАФЕДРА **Компьютерные системы и сети (ИУ6)**

НАПРАВЛЕНИЕ ПОДГОТОВКИ **09.04.01 Информатика и вычислительная техника**

МАГИСТЕРСКАЯ ПРОГРАММА **09.04.01/07 Интеллектуальные системы анализа,**

**обработки и интерпретации больших данных**

**Отчет**

**по лабораторной работе №4**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Студент | ИУ6-32М |  |  | А.М. Панфилкин |
|  | (Группа) |  | (Подпись, дата) | (И.О. Фамилия) |
|  |  |  |  |  |
| Преподаватель |  |  |  | М.А. Захаров |
|  |  |  | (Подпись, дата) | (И.О. Фамилия) |

**Задание**

1. Построить тепловую карту. Оси - года и штаты. Значения – процент убийств от общего количества убийств в этот год
2. Создать Bar Chart, отражающий процент жертв по полу в каждом штате.
3. Отсортируйте штаты в предыдущем пункте по убыванию процентного соотношения мужчин

**Выполнение**

Листинг 1 – Выполнение задания

|  |
| --- |
| import pandas as pd  import seaborn as sns  import matplotlib.pyplot as plt  import squarify  # Построить тепловую карту. Оси - года и штаты. Значения - процент  # убийств от общего количества убийств в этот год  homicide\_data = pd.read\_csv('data/homicide-data.csv', encoding='ISO-8859-1')  homicide\_data\_b = homicide\_data[homicide\_data['victim\_race'] == 'Black']  homicide\_data\_b['year'] = homicide\_data\_b['reported\_date'].astype(str).str[:4]  homicides\_by\_year\_state = homicide\_data\_b.groupby(['year', 'state']).size().unstack(fill\_value=0)  total\_homicides\_per\_year = homicides\_by\_year\_state.sum(axis=1)  percentage\_homicides = (homicides\_by\_year\_state.T / total\_homicides\_per\_year).T \* 100  plt.figure(figsize=(12, 8))  sns.heatmap(percentage\_homicides, annot=True, fmt=".1f", cmap='coolwarm', cbar\_kws={'label': 'Percentage of Homicides'})  plt.yticks(rotation=0)  plt.title('Percentage of Homicides by State and Year')  plt.xlabel('State')  plt.ylabel('Year')  plt.show()  # Процент male female по годам  state\_gender\_counts = homicide\_data.groupby(['state', 'victim\_sex']).size().unstack(fill\_value=0)  state\_gender\_percentage = state\_gender\_counts.div(state\_gender\_counts.sum(axis=1), axis=0) \* 100  state\_gender\_percentage = state\_gender\_percentage.reset\_index().melt(id\_vars='state', var\_name='victim\_sex', value\_name='percentage')  # sort by males  state\_gender\_percentage.sort\_values()  plt.figure(figsize=(12, 8))  sns.barplot(x='state', y='percentage', hue='victim\_sex', data=state\_gender\_percentage, palette='coolwarm')  plt.title('Percentage of Men and Women in Each State')  plt.xlabel('State')  plt.ylabel('Percentage')  plt.xticks(rotation=45)  plt.legend(title='Victim Gender')  plt.tight\_layout()  plt.show()  grouped = homicide\_data.groupby(['state', 'victim\_sex']).size().unstack(fill\_value=0)  grouped['Total'] = grouped.sum(axis=1)  grouped['Male\_Percentage'] = (grouped['Male'] / grouped['Total']) \* 100  grouped['Female\_Percentage'] = (grouped['Female'] / grouped['Total']) \* 100  grouped['Unknown\_Percentage'] = (grouped['Unknown'] / grouped['Total']) \* 100  # Sorting states by the percentage of men  sorted\_states = grouped.sort\_values('Male\_Percentage', ascending=False).index  # Reordering the DataFrame according to sorted states  grouped = grouped.loc[sorted\_states]  # Reshape the data for plotting  stacked\_data = grouped[['Male\_Percentage', 'Female\_Percentage', 'Unknown\_Percentage']].stack().reset\_index()  stacked\_data.columns = ['State', 'Gender', 'Percentage']  # Create the bar plot using Seaborn  plt.figure(figsize=(12, 8))  sns.barplot(x='State', y='Percentage', hue='Gender', data=stacked\_data, palette='muted')  plt.title('Percentage of Men, Women, and Unknown Genders in Each State')  plt.xlabel('State')  plt.ylabel('Percentage')  plt.xticks(rotation=90)  plt.legend(title='Gender')  plt.tight\_layout()  plt.show() |

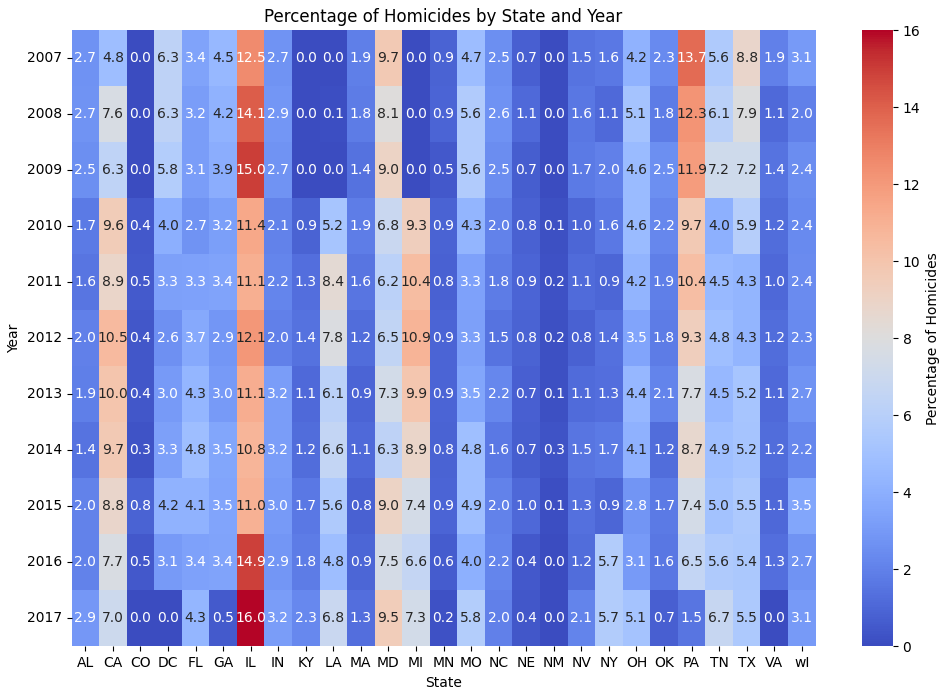


Рисунок 1 – Результат выполнение пункта 1 задания

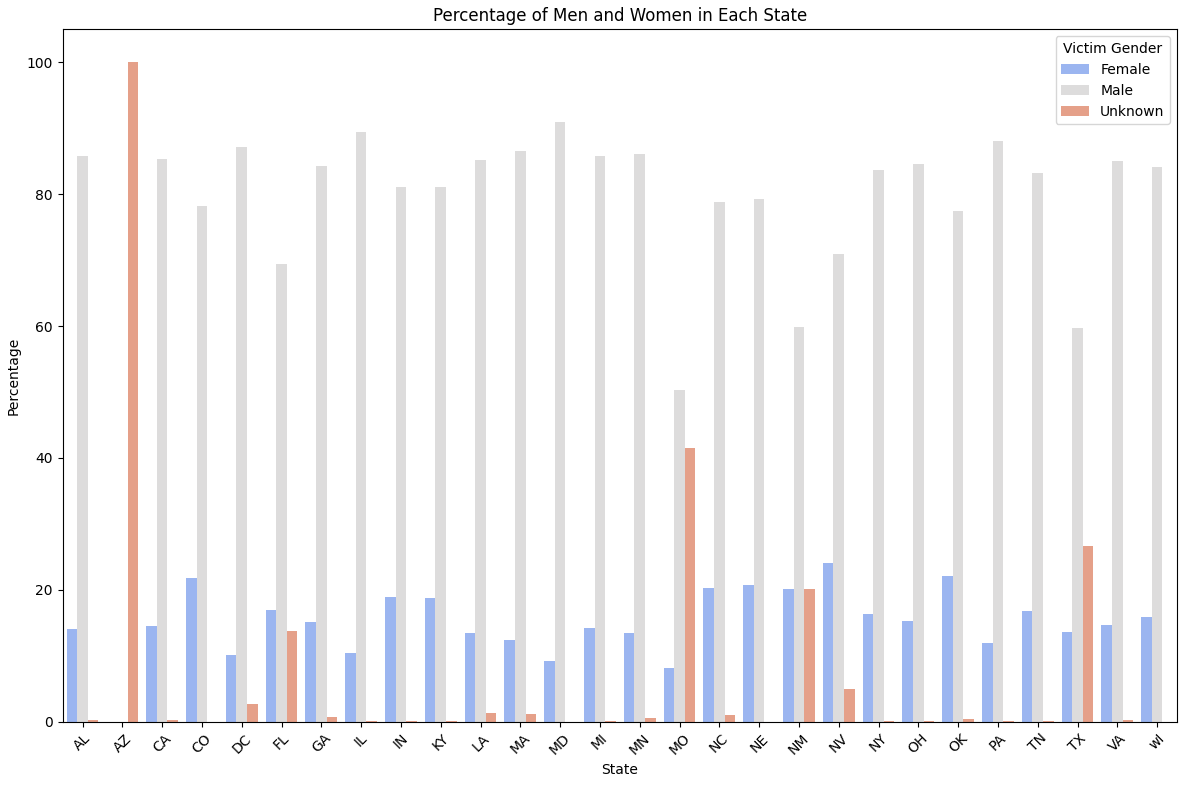


Рисунок 2 – Результат выполнения пункта 2 задания

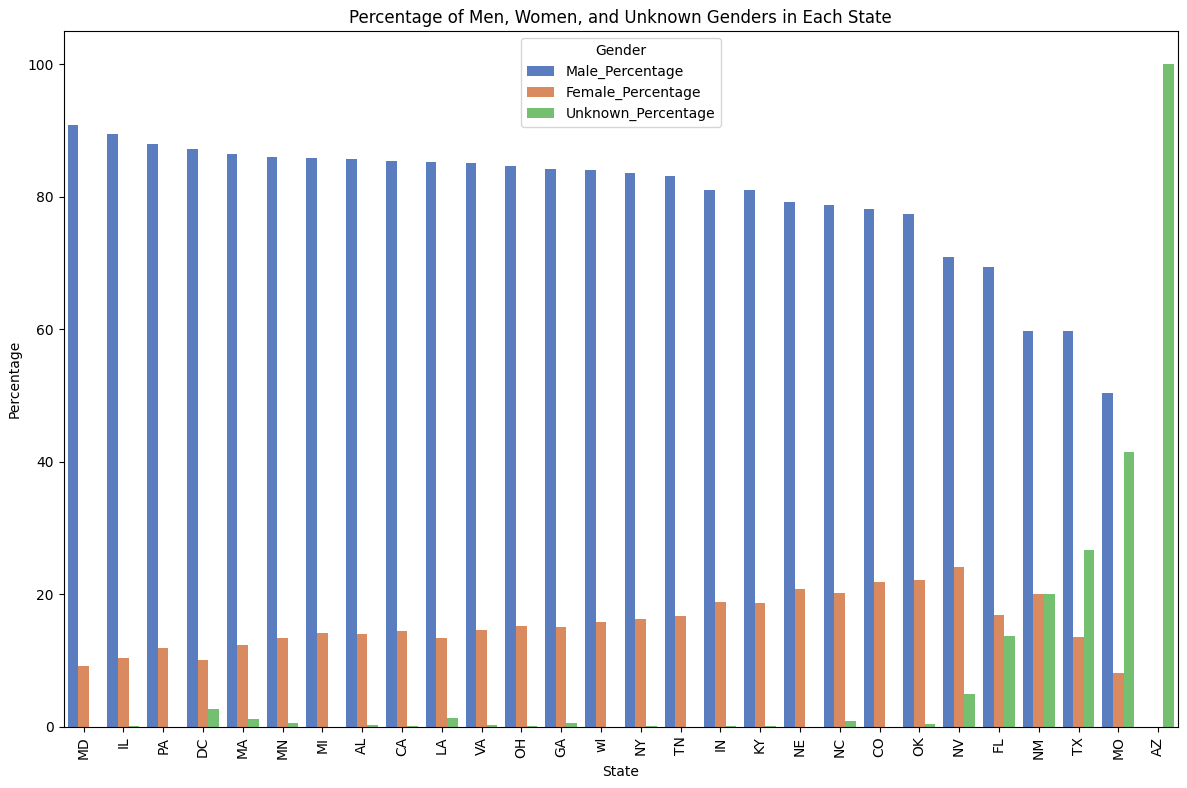


Рисунок 3 – Результат выполнения пункта 3 задания

**Вывод**

В ходе лабораторной работы мы углубились в работу с библиотекой seaborn, освоив методы построения тепловых карт и Bar Chart. Новые навыки включают в себя способность визуализации процентных соотношений убийств по штатам и годам, а также представление данных о проценте жертв по полу в каждом штате с использованием Bar Chart. Кроме того, мы научились эффективно сортировать данные визуализаций для более четкого восприятия информации.