CAPSTONE PROJECT

-Fabrizio Yair Oyarzabal

INTRODUCTION

- **The present Data Science project aims to conduct a comprehensive analysis of data related to the number of homicides in Argentina. The analysis will focus on the records of homicides that occurred in the country during the period between 2015 and 2021. The increase in violence and homicides is a constant concern for society and authorities. Understanding the trends and patterns surrounding these events is essential to implement effective policies and strategies that contribute to the safety and well-being of the population.**

BUSINESS PROBLEM

- **The Government of the Argentine Republic is seeking to address a critical concern related to public safety and crime rates within the country. As part of their efforts to ensure the well-being and security of the population, they are interested in understanding the current state of homicides in Argentina and identifying any potential trends or improvements.**

DATA – ARGENTINA

Region	Subregion	Country	Category	Age	Sex	Source	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021
Americas	Latin America and the Caribbean	Argentina	Total	Total	Female	CTS						387		408	396	408	371	308
					Male	CTS						2,450		1,909	1,988	1,895	2,044	1,785
					Total	MoS/CTS	2,384	2,511	2,648	3,072	3,228	2,837	2,625	2,317	2,384	2,307	2,417	2,093

METHODOLOGY

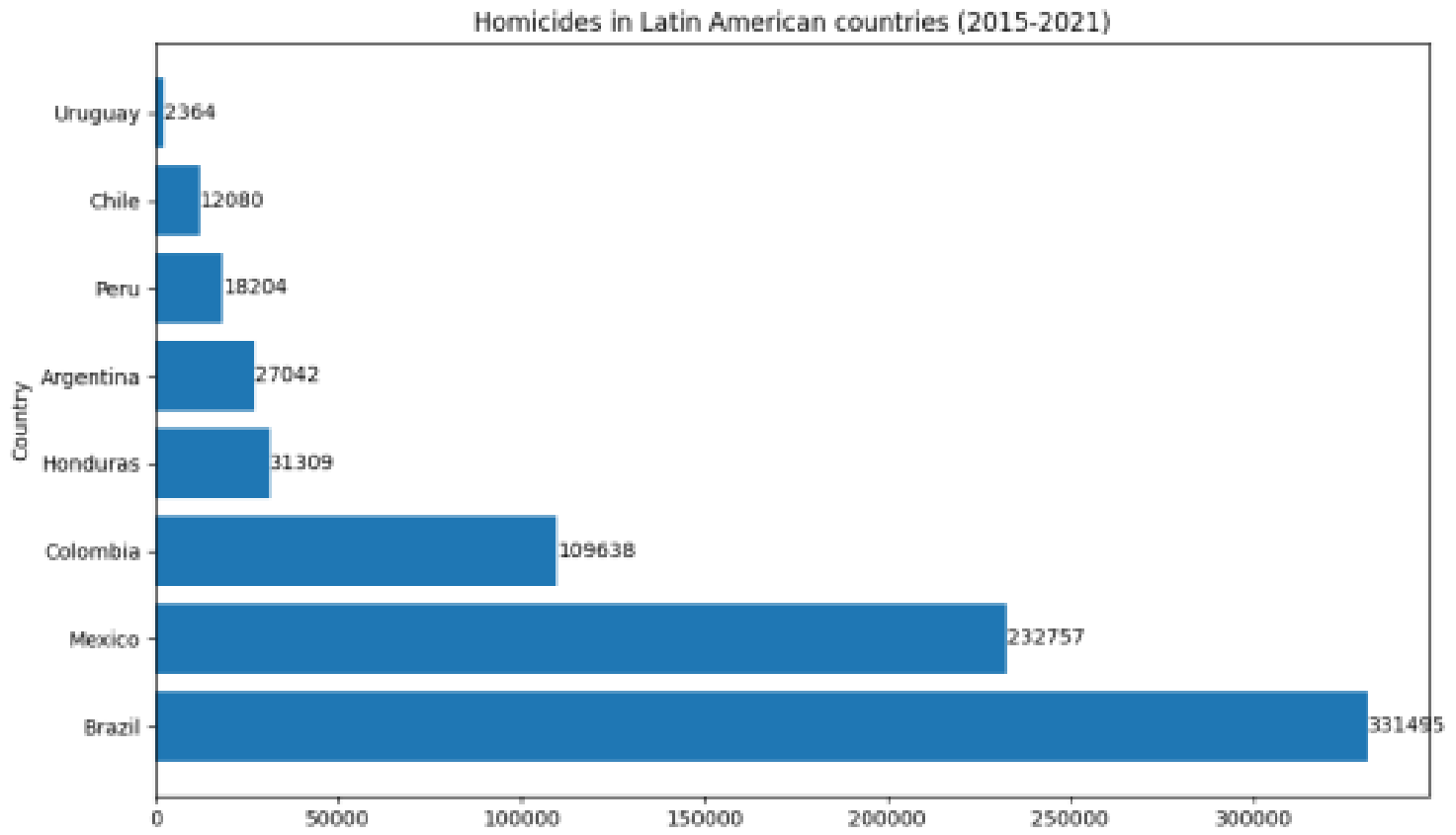
- **Data Collection:** Data is loaded from an external URL using the pandas library to read the Excel file and store the data in a DataFrame.
- **Data Wrangling:** Data filtering is performed using conditions based on specific columns of the DataFrame to select relevant data for each graph.
- **Exploratory Data Analysis (EDA) with Visualization:** The matplotlib and pandas libraries are used to create different graphs that help explore and visualize the data. Horizontal bar charts, bidirectional bar charts, line charts, and pie charts are generated to display information about homicides in Latin American countries, homicides by age and gender in Argentina, and the distribution of homicides by category in Argentina.
- **Dash Dashboard:** The Dash library is utilized to create an interactive panel that showcases the pie chart for the distribution of homicides by category in Argentina. Dash is a Python library for building interactive web applications based on data.

DATA COLLECTION

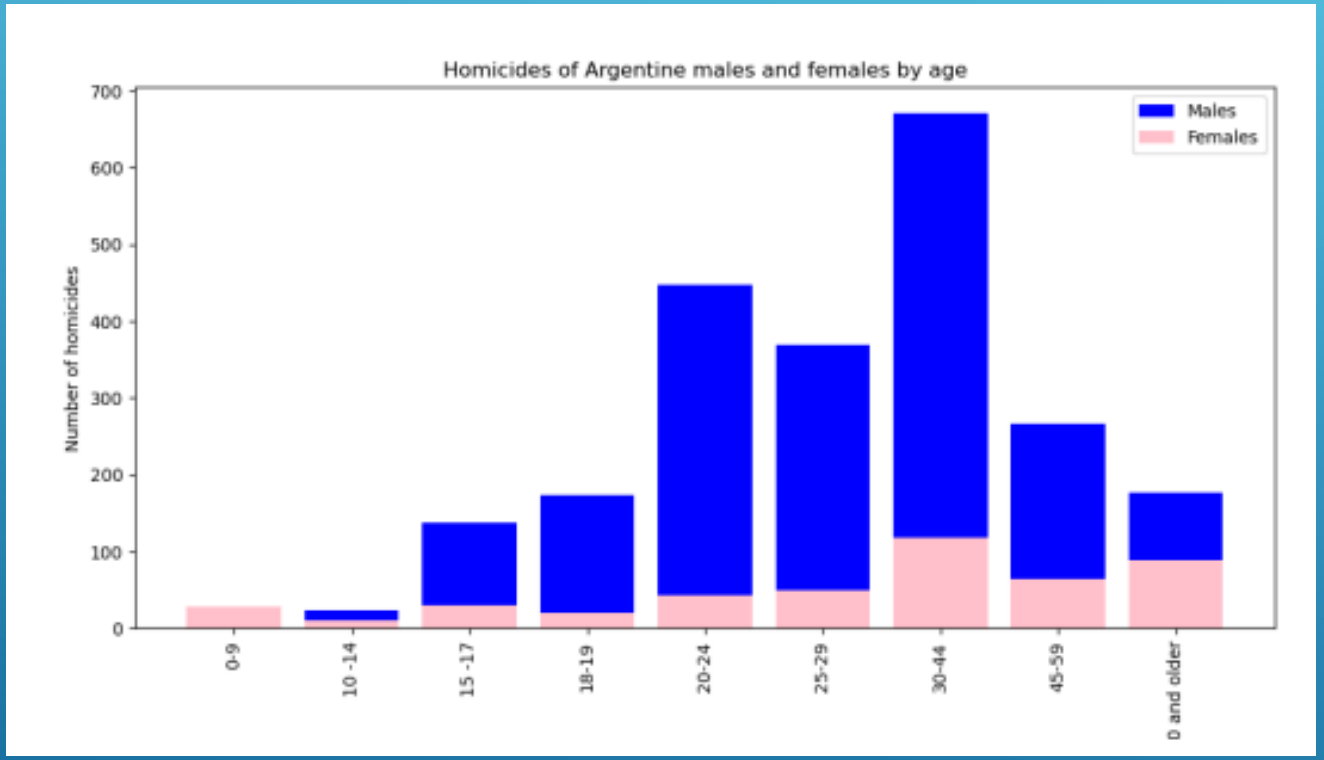
```
import pandas as pd
import matplotlib.pyplot as plt
import ssl

ssl._create_default_https_context = ssl._create_default_https_context = ssl._create_unverified_context

url = 'https://dataunodc.un.org/sites/dataunodc.un.org/files/data_cts_intentional_homicide.xlsx'
df = pd.read_excel(url, header=2)
```

DATA WRANGLING

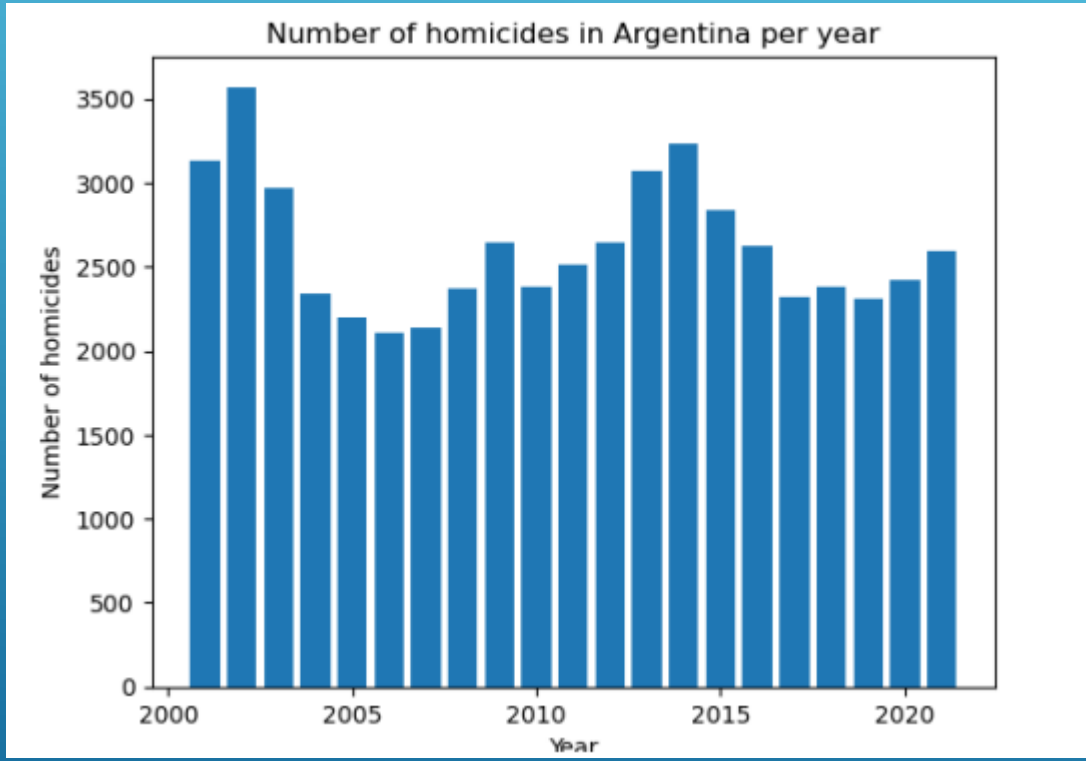
```
# Filter the data for the total of Year, Dimension, Category, and Sex
df_filtered = df[(df['Year'].between(2015, 2021)) & (df['DImension'] == 'Total') & (df['Category'] == 'Total') & (df['Sex'] == 'Total') & (df['Unit of measu

# Select the desired countries
selected_countries = ['Argentina', 'Brazil', 'Mexico', 'Colombia', 'Chile', 'Peru', 'Honduras', 'Uruguay']
df_selected = df_filtered[df_filtered['Country'].isin(selected_countries)]
```



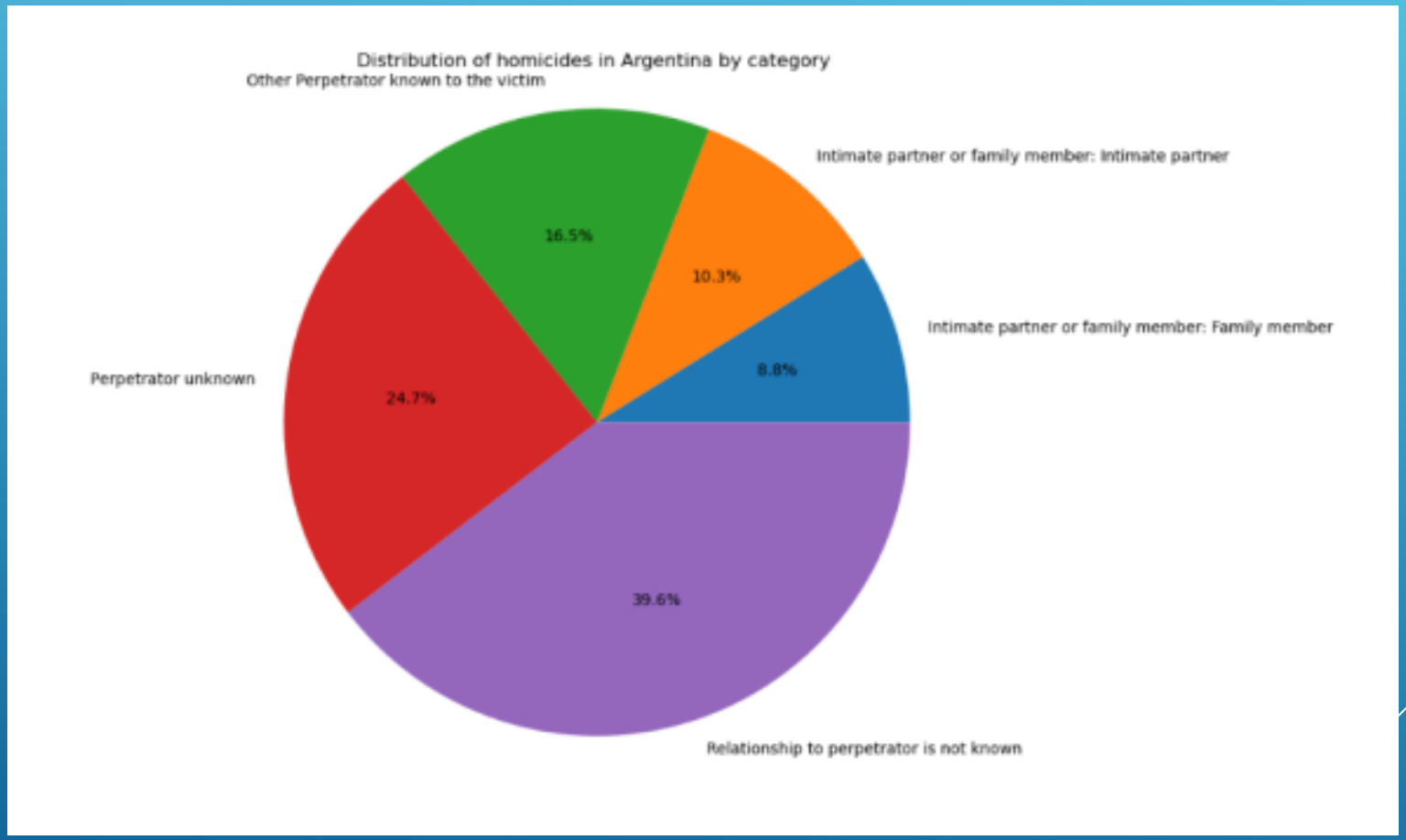
EXPLORATORY DATA ANALYSIS (EDA) WITH VISUALIZATION - 1



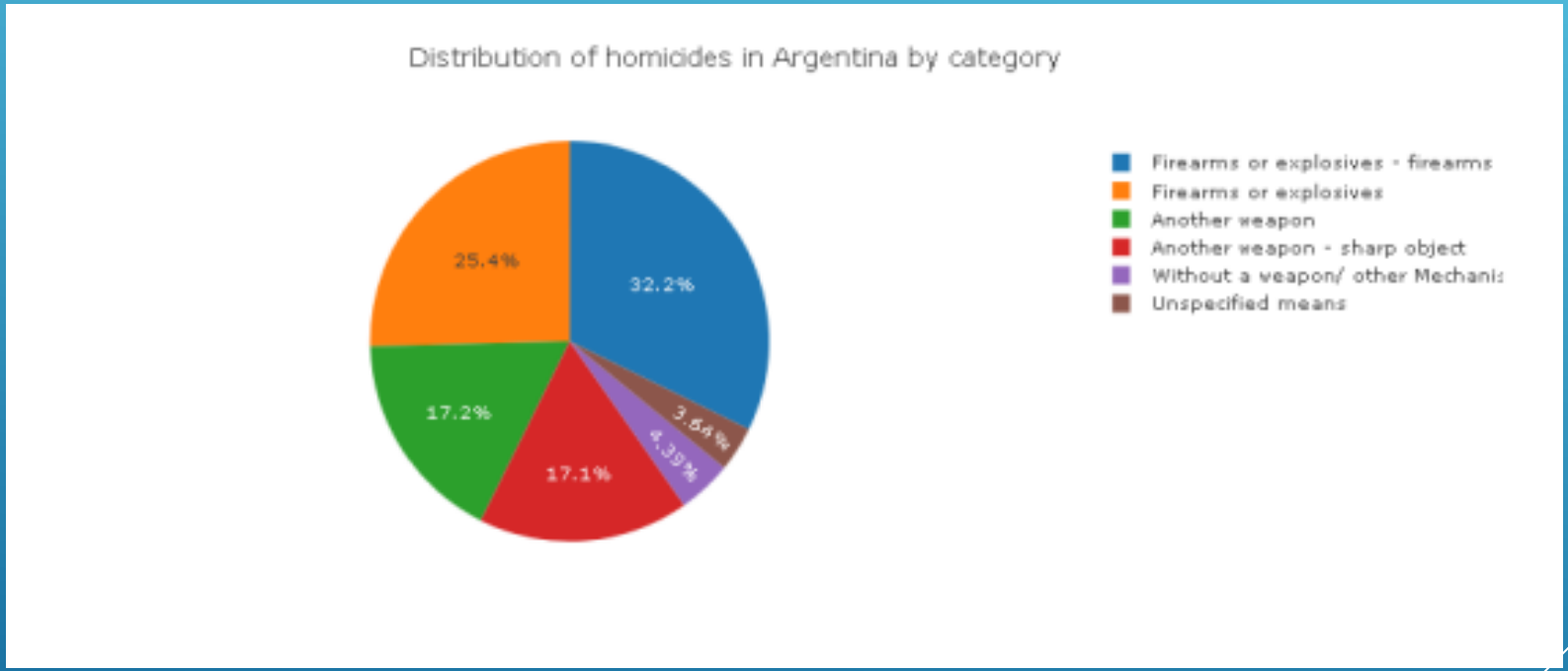
EXPLORATORY DATA ANALYSIS (EDA) WITH VISUALIZATION - 2



EXPLORATORY DATA ANALYSIS (EDA) WITH VISUALIZATION - 3



DASH DASHBOARD



DISCUSSION

- After a comprehensive analysis of homicide data in Argentina, significant conclusions have been drawn, offering valuable insights for policy and preventive measures. The findings indicate a decline in homicides from 2015, suggesting the effectiveness of certain policies or measures. The analysis also highlights the correlation between male perpetrators and the age range of 19 to 44, calling for targeted strategies to address male-related violent behavior. Moreover, the prevalent use of firearms and dangerous weapons in homicides emphasizes the need for stricter regulations. These insights provide a basis for the government of Argentina to design effective crime prevention strategies and prioritize measures to enhance citizen safety and well-being. Further efforts are required to continue reducing homicides, with a focus on addressing male violence and controlling weapon use. By leveraging data-driven analysis, Argentina can make informed decisions and ensure a safer environment for its citizens.

CONCLUSION

- In conclusion, this Data Science project has successfully conducted a comprehensive analysis of homicide data in Argentina, revealing a decline in homicides since 2015 and a significant correlation between male perpetrators and the age group of 19 to 44. The prevalent use of firearms and other dangerous weapons in the commission of homicides was also identified. These findings highlight the need for targeted strategies to address male-related violent behavior and stricter measures to control weapon-related crimes. By leveraging these insights, the government of Argentina can make informed decisions and implement effective crime prevention policies, ensuring the safety and well-being of its citizens and fostering a safer environment for the nation. The use of data analysis and Python's powerful libraries has been instrumental in providing valuable information for data-informed decision-making.