Abstract:

When we talk about data, we often think of numbers and letters, and sometimes this data contains hundreds of values that are challenging to interpret. In today's world, where data comes from many sources; such as our phones, tablets, or AI generated images, visualizations have become essential tools for data scientists. They help us make vast amounts of information more interactive and engaging for human interpretation.

The goal of this report is to explore various visualizations using tools like Power BI and Python to answer interesting questions about our dataset.

Some questions like:

What are the yearly trends in visa applications?

Are there peak or low years for visa applications?

How do visa approval rates compare to refusal?

Is there a relationship between visa applications and approvals by nationality?

In this process, we import, modify, calculate, and project data values into clear and useful visualizations. By the end of this report, we aim to understand the importance of visualization and its usefulness in future projects.

About the data set:

This dataset relates to Domestic Residence & Permissions applications and decisions by year and nationality. This dataset has been generated at a specific point in time and is subject to revision. Values in the range 1 to 3 are suppressed using * for statistical non-disclosure reasons.

Date released   2024-09-13

Date updated   2024-09-10

Data of subtraction 2024-10-31

Update frequency   Monthly

Published by: Department of Justice

Licensed under: Creative Commons Attribution 4.0

Web: https://data.gov.ie/dataset/domestic-residence-permissions-applications-and-decisions-year-and-nationality

Type	Status	Nationality	2017	2018	2019	2020	2021	2022	2023	2024
Domestic Residence & Permissions applications	Received	Afghanistan	13	104	84	82	222	376	830	471
Domestic Residence & Permissions applications	Received	Åland Islands	0	0	0	0	0	0	0	0
Domestic Residence & Permissions applications	Received	Albania	32	307	303	195	305	364	213	97
Domestic Residence & Permissions applications	Received	Algeria	31	134	102	112	224	1030	2305	1289
Domestic Residence & Permissions applications	Received	American Samoa	0	0	0	0	0	0	0	0
Domestic Residence & Permissions applications	Received	Andorra	0	0	0	0	0	0	0	0
Domestic Residence & Permissions applications	Received	Angola	1	16	11	9	20	35	32	23

Values that were suppressed have been adjusted to represent a value of one. Given that the original suppressed values range between 1 and 3, I have chosen a value of one to indicate that at least one visa application was submitted in these cases. This decision is important, as any potential bias introduced by this adjustment should be considered during our analysis

Understanding the Dataset

In Ireland, Domestic Residence and Permissions applications refer to requests made by non-Irish nationals seeking authorization to reside in the country for various reasons, including employment, study, family reunification, or humanitarian needs. These permissions allow individuals to live legally and, in some cases, work in Ireland under specific conditions defined by Irish immigration laws.

The Irish Naturalisation and Immigration Service (INIS) is responsible for processing these applications, evaluating each one based on eligibility requirements and the applicant's specific circumstances. (*Long Term Residency - Immigration Service Delivery*, 2024)

We can analyze and describe this dataset effectively using the Python package pandas, which provides tools for data manipulation and exploration.

	2017	2018	2019	2020	2021	2022	2023	2024
count	660	660	660	660	660	660	660	660
mean	18.20606	29.21364	31.25	25.39242	35.87273	74.01364	79.23485	60.27879
std	74.38641	100.9772	107.3677	86.06779	127.7103	286.2006	292.5077	223.4505
min	0	0	0	0	0	0	0	0
25%	0	0	0	0	0	0	0	0
50%	0	1	1	1	1	1	1	1
75%	5	10	11	10	11	17	18	15
max	916	1105	1052	816	1176	3149	2401	2211

Our dataset contains 660 rows, covering a time period from 2017 to September 2024 (the month of the last update). It includes 11 columns:

'Type': Represents the type of visa, with our dataset focusing exclusively on Domestic Residence and Permissions.

'Status': Contains three different statuses—Received, Granted, and Refused.

'Nationality': Includes visa data for 202 unique nationalities from around the world.

2017-2024: Each column represents a year, with the corresponding total number of visas per nationality.(See Index I for code exploration)

Based on the table, we observe a significant increase in visa applications over the period from 2017 to 2024, with over a 200% rise and a peak in applications in 2022. The mean number of applications per year has also grown, indicating a consistent upward trend in demand for this visa type.

The minimum value of zero suggests that there are nationalities which did not apply for this visa in certain years, while the standard deviation values highlight variability in application counts, particularly in 2022 and 2023, which saw higher fluctuations.

From 2018 onward, at least 50% of nationalities in the dataset applied for this visa type at least once each year. Additionally, the maximum values for each year, such as the peak of 3,149 applications in 2022, suggest that certain nationalities had a notably high demand in specific year.

These patterns provide insights into potential factors affecting visa demand, such as policy changes, economic conditions, or geopolitical factors, which could be explored further.

Visa applications tend to fluctuate over the years for various reasons—political, social, or, as in 2020, due to the pandemic. This brings us to the question:

What are the yearly trends in visa applications from 2017 to 2024?

To help answer this, let's take a look at the bar plot below
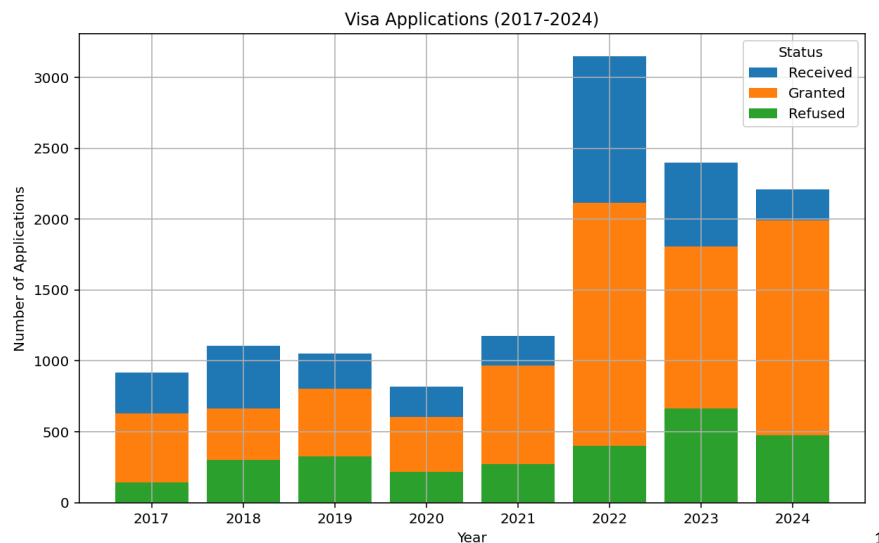
Based on our visualization, we can observe several important points. For instance, in 2022, there was a significant peak in the number of applications, with Received applications standing out above all others. This is important to note because, according to Citizens Information, a new scheme allowing people without immigration status in Ireland to apply for legal residence opened for applications between January 31 and July 31, 2022. A separate scheme for international protection applicants was also open until August 7, 2022 (CitizensInformation.ie, n.d.), this may be the reason of this fluctuation.

We can also see a general trend of increasing applications over this period, with a slight drop in 2020. This decrease aligns with the pandemic period, which led to border closures and likely impacted application numbers for that specific year. Interestingly, even with the lockdown lasting over a year, the reduction in applications was relatively minimal.

Another interesting observation is for 2023, where we notice a slightly higher rate of Refused applications compared to other years. This trend might be worth investigating further to understand the reasons behind the increase in refusals. It's worth mentioning that in 2023, Ireland, specifically Dublin, experienced tension involving foreign nationals. "The Dublin riot took place on the evening of November 23, 2023, involving multiple incidents of vandalism, arson, and looting in the city center, as well as assaults on Gardaí (Irish police) and members of the public" (Wikipedia contributors, 2024).

Finally, we observe that in 2024, Granted applications are leading over both Received and Refused statuses. However, it's essential to note that the data for this year only covers January to September, so additional decisions later in the year could impact these figures by year's end.

---

[1] Plot generated by matplotlib code index 2

As mention before to many events can happen to modify the trend in visas application another question, I like to answer is How do visa approval rates compare to refusal?
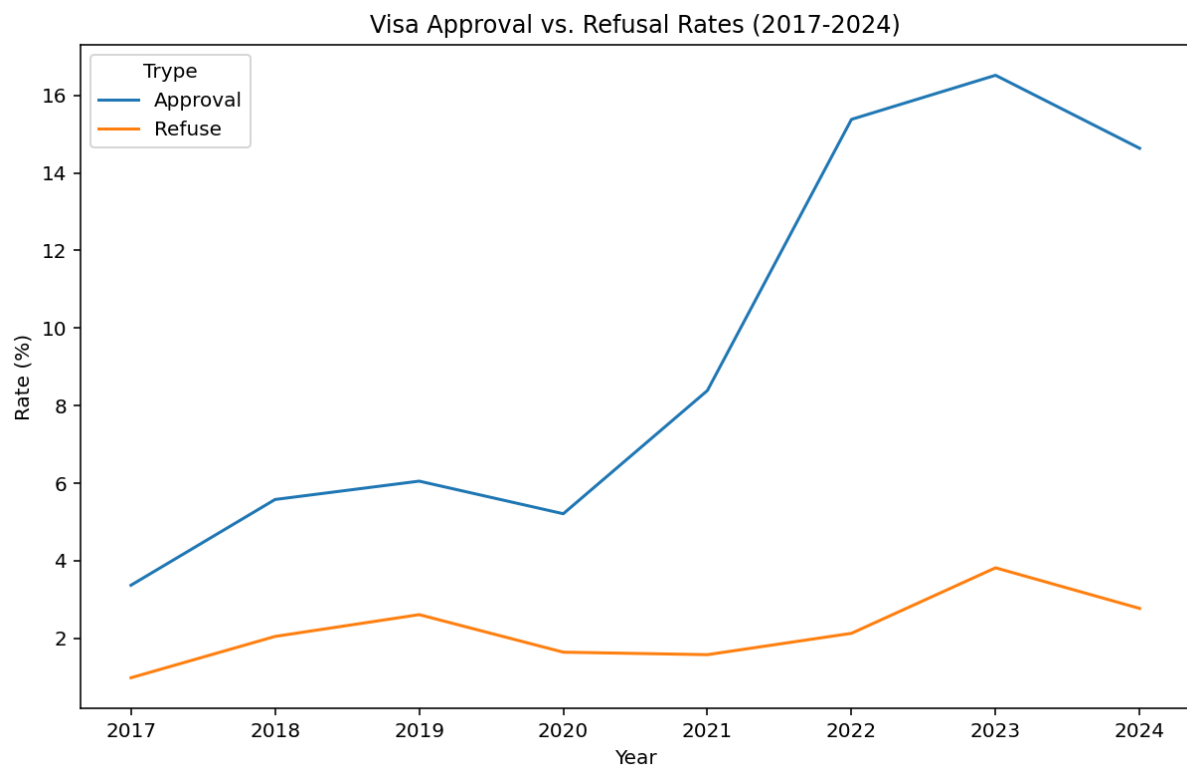
As mentioned before, we observe that the Approval rate has generally maintained a strong increase over the time with a pick from 2021 until now, while the Refusal rate remains minimal by comparison. This suggests that Ireland has a relatively flexible approach to this type of visa application, with significant increases in approvals, particularly in recent years. The only noticeable peaks in Refusals are in 2023, as previously discussed may be for Dublin riot that took place this same year, and a smaller increase in 2019 which may imply that those years the government may be slightly stricter.

---

[2] Plot generated by matplotlib code index 2

Another interesting question is whether visa approvals depend on the continent from which the applications originate and if there are any notable differences between continents.

Scatter Plot of Visa Applications by Continent and Status
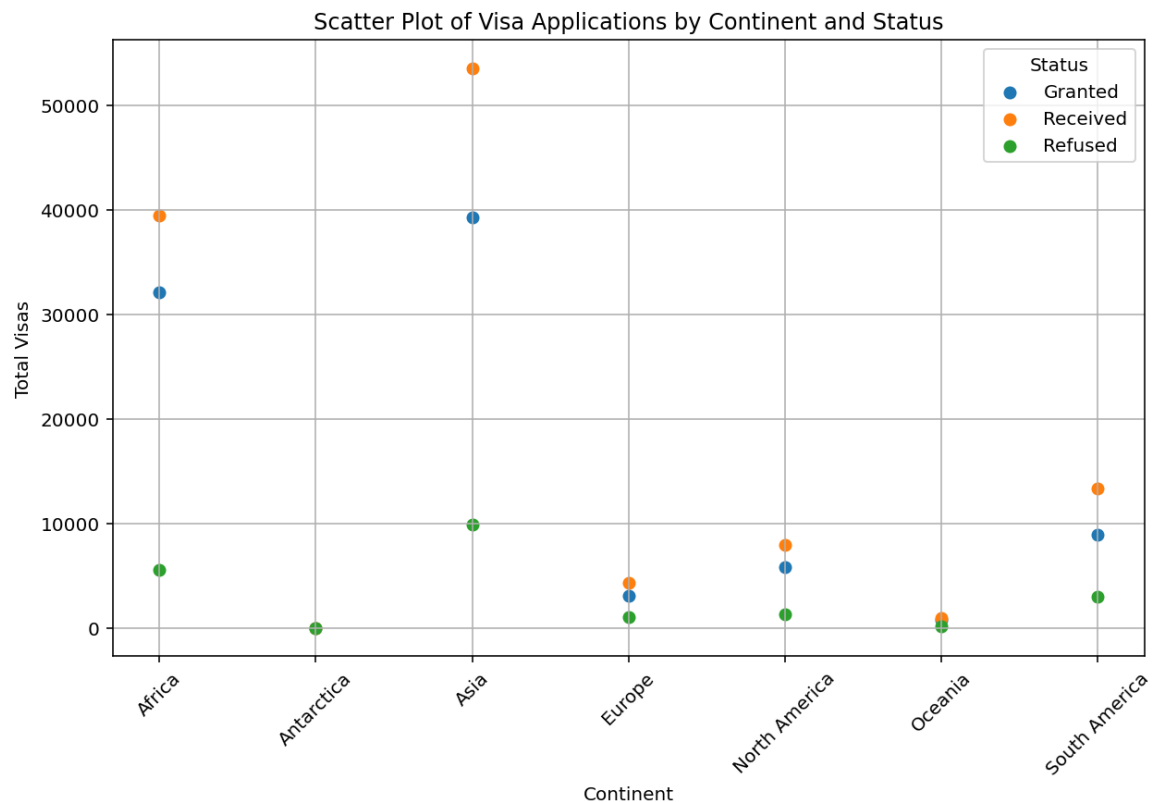
Asia, for instance, has received over 50,000 applications, with a rejection rate of about 20%, leaving nearly 80% of applications approved. On the other hand, African countries have submitted close to 40,000 applications, with approximately 7,000 refusals, meaning only 12.5% of applications were rejected.

South America and North America are relatively close, but we can notice that South America the rejection and the approval are almost equals, North America is slightly better in obtaining approbation in comparation the south.

Finaly Europe has a slow application, this understandable because Ireland is part of the European Union and must of the countries do no required visa to remind in the state.

---

[3] Plot generated by matplotlib code index 3

Finally, with the help of an interactive dashboard, we can potentially analyze individual cases, allowing the user to interact with and focus on specific countries
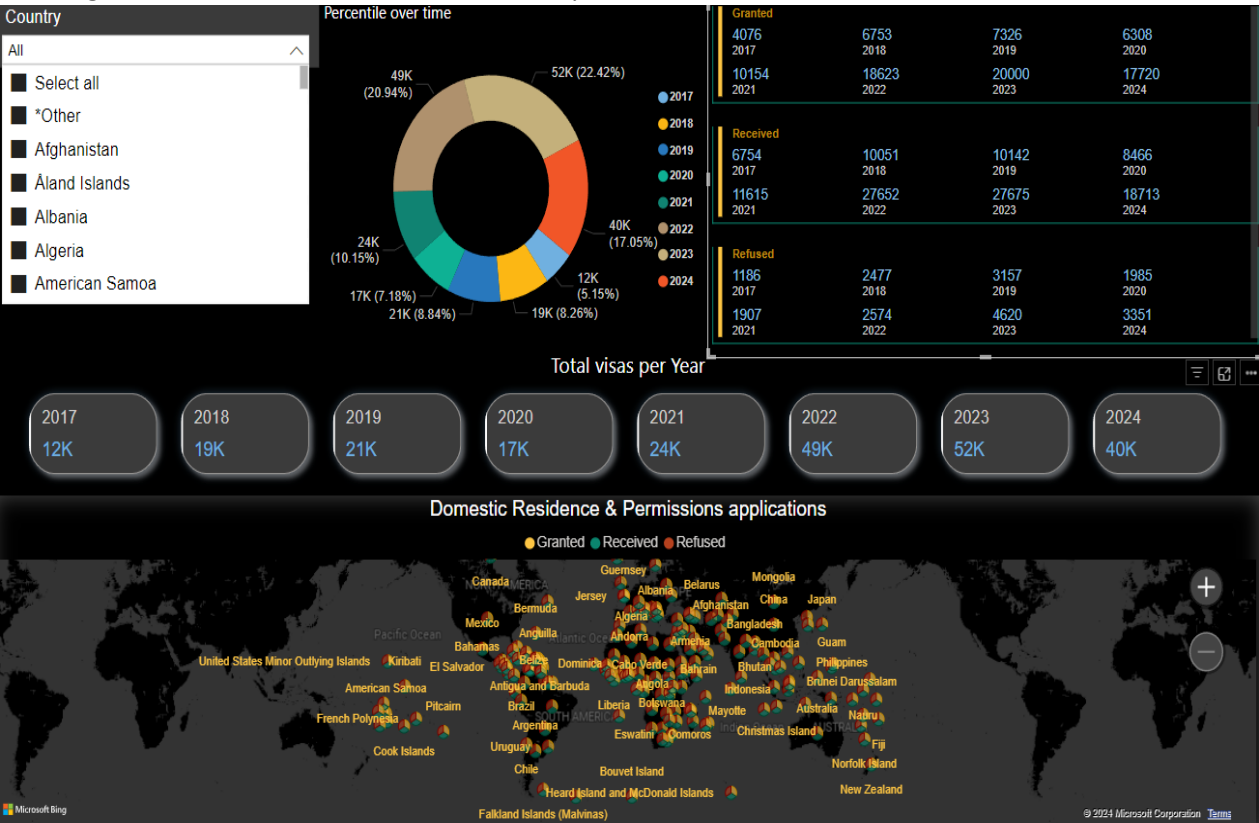
This feature is helpful for further analysis, enabling us to answer questions such as: Has the percentage of rejections increased or decreased over the years?
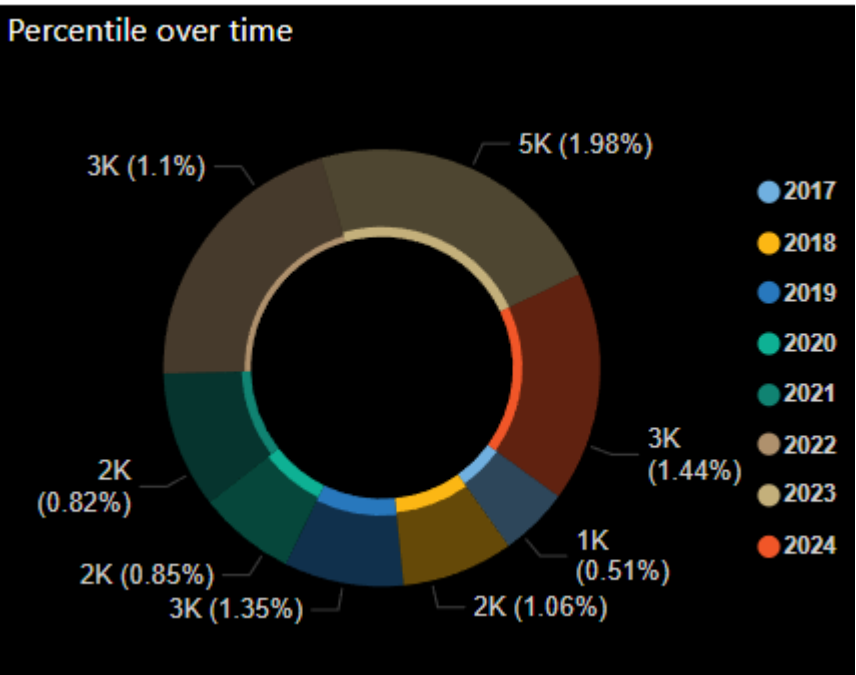
We can quickly view these values using the donut chart and selecting the 'Refused' category.

From this, we can conclude that the percentage of refusals has increased compared to 2017.

Conclusion:

Visualizations are a powerful tool for data analysis. In this case, with our dataset on Domestic Residence and Permissions applications, we have revealed significant information about refused, granted, and received applications. We can clearly see highlights in years where external events may have affected the trend, such as COVID-19 or the conflict in 2023. Further analysis would provide a deeper understanding of how these events influenced visa trends; however, for the scope of our assignment, we limited our focus to showing general insights from our dataset.

Additionally, we created interactive features in the dashboard that allow users to drill down into specific countries and years, providing a deeper understanding of regional and temporal patterns. This interactivity enables quick insights into questions like the fluctuation of rejection rates over time. Overall, the dashboard serves as a powerful tool for visualizing trends, identifying anomalies, and informing decisions related to visa policies and resource allocation

Reference:

Long Term Residency - Immigration Service Delivery. (2024, October 24). Immigration Service Delivery. https://www.irishimmigration.ie/my-situation-has-changed-since-i-arrived-in-ireland/long-term-residency/

Citizensinformation.Ie. (n.d.). *Regularisation of Long-Term Undocumented Migrants Scheme*. https://www.citizensinformation.ie/en/moving-country/moving-to-ireland/rights-of-residence-in-ireland/permission-to-remain-for-undocumented-noneea-nationals-in-ireland/

Wikipedia contributors. (2024, October 5). Irish anti-immigration protests (2022–present). Wikipedia. https://en.wikipedia.org/wiki/Irish_anti-immigration_protests_(2022%E2%80%93present)#:~:text=The%202023%20Dublin%20riot%20took,and%20members%20of%20the%20public.

Index:

1.0.-Cleaning data set:

```
path_file = "C:/Users/MAAP/Documents/55a55b1e-cf64-4ca3-b34c-9be3c0ae8b70.xlsx"

df = pd.read_excel(path_file)

df["Total_visas"] = df.loc[:,2017:2024].sum(axis=1)

df["Average"]=df.loc[:,2017:2024].sum(axis=1)/8

##check for correct name

country_list = [country.name for country in pycountry.countries]

def correct_name(name):

    match = get_close_matches(name, country_list, n=1, cutoff=0.5)  # n=1 for best match, cutoff for accuracy

    return match[0] if match else name

df["Nationality"] = df["Nationality"].apply(correct_name)

##check for continent

def country_continent(country):

    ##Handle countries not in .country_name_to_country_alpha2

    continents = {"Europe":['Holy See (Vatican City State)','Kosovo'],"Oceania" :['Pitcairn'],

            "North America":['Sint Maarten (Dutch part)','United States Minor Outlying Islands'],

            "Asia" : ["Timor-Leste"],"Africa" : ['Western Sahara'],"Antarctica":["Antarctica"]}

    try:

        country_code = pc.country_name_to_country_alpha2(country)

        continent_code = pc.country_alpha2_to_continent_code(country_code)

        continent_name = pc.convert_continent_code_to_continent_name(continent_code)

        return continent_name

    except KeyError:

        for i in continents:

            if country in continents[i]:

                return i

    else:

        return "Other"

df["Continent"] = df["Nationality"].apply(country_continent)


```

1.1-Pandas code for exploration:

```
##explore dataset before manipulation

print(df.describe())

print(df.columns)
```

```python
df1 = df["Nationality"].nunique()

print(df1)
```

2.-Bar plot code.

```python
##modify dataframe for plotting

df_melted = df2.melt(id_vars='Status', var_name='Year', value_name='Count')

df_melted['Year'] = pd.to_numeric(df_melted['Year'])

# Plotting

plt.figure(figsize=(10, 6))

for status in df_melted['Status'].unique():

    subset = df_melted[df_melted['Status'] == status]

    plt.bar(subset['Year'], subset['Count'], label=status)

plt.title("Visa Applications (2017-2024)")

plt.xlabel("Year")

plt.ylabel("Number of Applications")

plt.legend(title="Status")

plt.grid(True)

plt.show()
```

2.-Trend plot code.

```python
approval_rate = []

refuse_rate = []

years= [2017,2018,2019,2020,2021,2022,2023,2024]

for i in years:

    x= float((df[i][df["Status"]== "Granted"].sum()/df["Total_visas"][df["Status"]== "Received"].sum())*100)

    y= float((df[i][df["Status"]== "Refused"].sum()/df["Total_visas"][df["Status"]== "Received"].sum())*100)

    approval_rate.append(x)

    refuse_rate.append(y)

plt.figure(figsize=(10,6))

plt.plot(years,approval_rate,label="Approval")

plt.plot(years,refuse_rate,label="Refuse")

plt.title("Visa Approval vs. Refusal Rates (2017-2024)")

plt.xlabel("Year")

plt.ylabel("Rate (%)")

plt.legend(title="Trype")

plt.show()

plt.clf()
```

3.- Continental plot code

```python
df_grouped = df.groupby(["Continent","Status"])["Total_visas"].sum().reset_index()

df_grouped['Continent_Index'] = pd.factorize(df_grouped['Continent'])[0]

# Scatter plot with different colors for each Status

plt.figure(figsize=(10, 6))

for status in df_grouped['Status'].unique():

    subset = df_grouped[df_grouped['Status'] == status]

    plt.scatter(subset['Continent_Index'], subset['Total_visas'], label=status)

# Customizing the plot

plt.xticks(df_grouped['Continent_Index'].unique(), df_grouped['Continent'].unique(), rotation=45)

plt.xlabel("Continent")

plt.ylabel("Total Visas")

plt.title("Scatter Plot of Visa Applications by Continent and Status")

plt.legend(title="Status")

plt.grid(True)
```