Disclaimer

This is a **practice notebook** that I did as a way to refresh myself with `numpy`, `pandas`, and `scipy.stats`. This is in **NO WAY A VALID STUDY**. The dataset was downloaded from Kaggle and as created by ruslankl. Methodology of the said survey is found here. Again, this entire notebook is **NO WAY A VALID STUDY**. This was msostly done for personal practice and upskilling.

Introduction

The LGBTQIA+ community continues to fight for their rights across parts of the world as it affects aspects of their lives. From Healthcare to Marriage, these sectors do not cater to the community at large. An unspoken aspect of queer and trans lives that is not spoken about is Employment. Many queer and trans people face discrimination in their lives. This is not new to countries inside the European Union (EU). Hence, this data report examined the following questions:

1. Is there any legal provision providing safety against discrimination in Employment for the LGBTQIA+ community within the EU?
2. Are the responses of the respondents correlated to their subset within the LGBTQIA+ community within the EU?
3. Are the responses of the respondents correlated to their country of origin within the EU?

Dataset and Libraries used

The main language used to analyze the data is Python. Using the libraries `numpy`, `pandas`, and `matplotlib.pyplot`, the writer can get a descriptive analysis of the dataset `LGBT_Survey_RightsAwareness.csv`. Likewise, the `researchpy` library and the `scipy.stats` library were imported to gather an inferential analysis of the dataset.

In [3]:
```
import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
import researchpy as rp
import scipy.stats as stats

df = pd.read_csv("LGBT_Survey_RightsAwareness.csv")

df.head()
```

Out[3]:

	CountryCode	subset	question_code	question_label	answer	percentage	notes
0	Austria	Lesbian	d1	In the country where you live, is there a law ...	Yes	53	NaN
1	Austria	Lesbian	d1	In the country where you live, is there a law ...	No	10	NaN
2	Austria	Lesbian	d1	In the country where you live, is there a law ...	Don`t know	37	NaN
3	Austria	Gay	d1	In the country where you live, is there a law ...	Yes	45	NaN
4	Austria	Gay	d1	In the country where you live, is there a law ...	No	16	NaN

Survey Questionnaire

The following is the complete questionnaire sent to the respondents.

In [4]:
```
questions = df.loc[:, 'question_label']
q_uni = questions.unique()

print(q_uni)
```

```
['In the country where you live, is there a law that forbids discrimination against p
ersons because of their sexual orientation when applying for a job?'
 'In the country where you live, is there a law that forbids discrimination against p
ersons because of their gender identity when applying for a job?'
 'Do you know of any organisation in the country where you live that can offer suppor
t or advice to people who have been discriminated against because they are Lesbian?'
 'Do you know of any organisation in the country where you live that can offer suppor
t or advice to people who have been discriminated against because they are Gay?'
 'Do you know of any organisation in the country where you live that can offer suppor
t or advice to people who have been discriminated against because they are Bisexual?'
 'Do you know of any organisation in the country where you live that can offer suppor
t or advice to people who have been discriminated against because they are Transgende
r?'
 'In the country where you live, have you ever seen any programme or awareness campai
gn by either the government or a non-governmental organisation addressing - Discrimin
ation against gay, lesbian and bisexual people?'
 'In the country where you live, have you ever seen any programme or awareness campai
gn by either the government or a non-governmental organisation addressing - Discrimin
ation against transgender people?'
 'In the country where you live, have you ever seen any programme or awareness campai
gn by either the government or a non-governmental organisation addressing - Discrimin
ation on the basis of gender?'
 'As far as you know, can same-sex couples legally marry and/or enter registered part
nerships in the country where you live?']
```

For brevity of this report, the writer has only focused on the first two questions (tagged as `d1` and `d2` in the dataset).

Demographic of the Sample

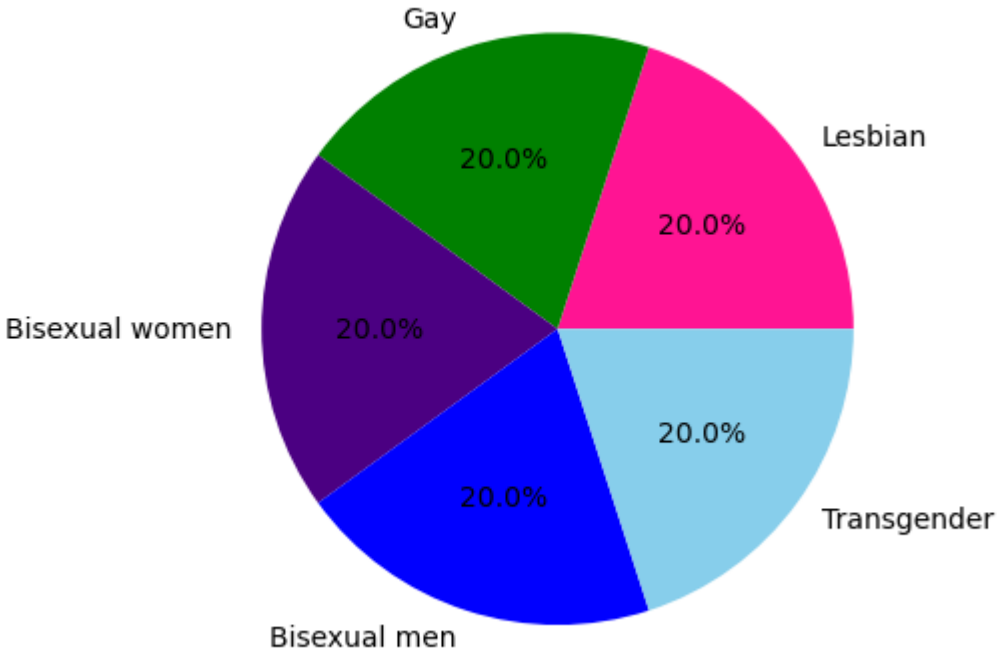
The following are the total respondents based on their subsets within the LGBTQIA+ community.

In [27]:
```
comm_df = df.loc[df['question_code'] == 'd1']
comm_x = comm_df.loc[:,'subset'].unique()
comm_y = comm_df.loc[:,'subset'].value_counts()
comm_colors = ["deeppink","g","indigo","blue","skyblue"]

def comm_pie():
    plt.pie(comm_y, labels=comm_x, autopct='%1.1f%%', colors=comm_colors)

print(comm_y)
comm_pie()
```

```
Lesbian           87
Gay               87
Bisexual women    87
Bisexual men      87
Transgender       87
Name: subset, dtype: int64
```



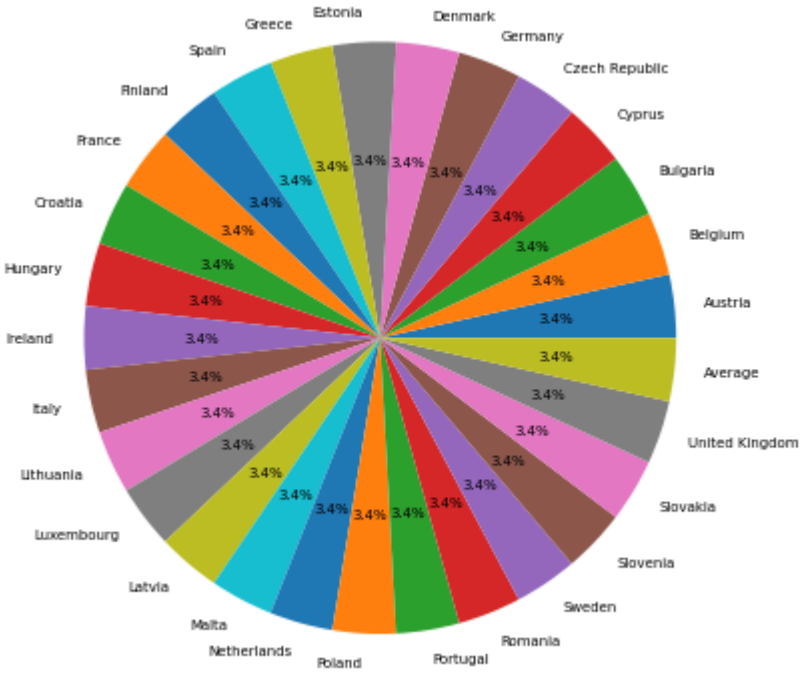
The following are the countries of origin of the respondents.

In [26]:
```
country_df = df.loc[df['question_code'] == 'd1']
country_x = country_df.loc[:,'CountryCode'].unique()
country_y = country_df.loc[:,'CountryCode'].value_counts()

def country_pie():
    plt.pie(country_y, labels=country_x, autopct='%1.1f%%', textprops={'fontsize': 5})

print(country_y)
country_pie()
```

```
Austria           15
Italy             15
United Kingdom    15
Slovakia          15
Slovenia          15
Sweden            15
Romania           15
Portugal          15
Poland            15
Netherlands       15
Malta             15
Latvia            15
Luxembourg        15
Lithuania         15
Ireland           15
Belgium           15
Hungary           15
Croatia           15
France            15
Finland           15
Spain             15
Greece            15
Estonia           15
Denmark           15
Germany           15
Czech Republic    15
Cyprus            15
Bulgaria          15
Average           15
Name: CountryCode, dtype: int64
```



On Workplace Discrimination

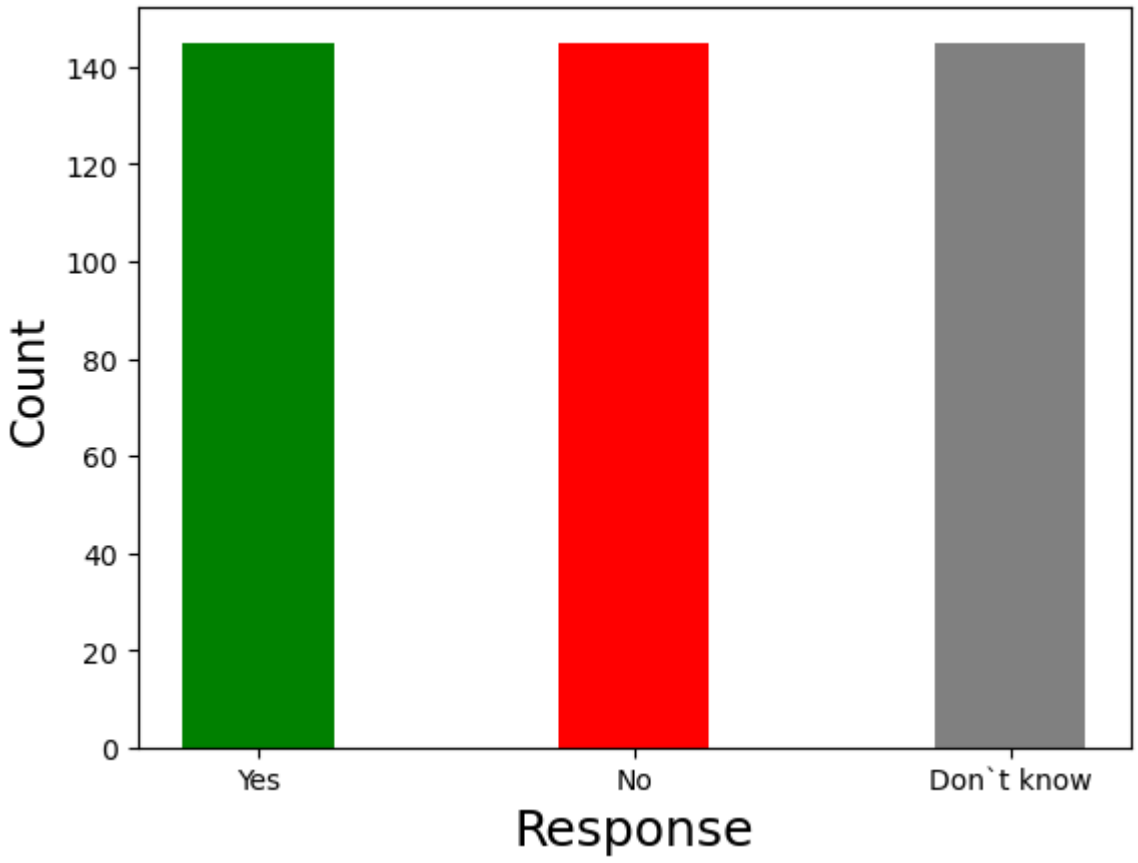
When asked about laws that forbids discrimination against persons because of their sexual orientation when applying for a job, this is the response of the respondents:

In [29]:
```
job_disc_2 = df.loc[df['question_code'] == 'd1']
job_disc_a2 = job_disc_2.loc[:,'answer'].value_counts()
job_disc_l2 = job_disc_2.loc[:,'answer'].unique()
colors = ["Green", "Red", "Grey"]

def job_disc1() :
    plt.bar(job_disc_l2, job_disc_a2, color=colors, width=0.4)
    plt.xlabel("Response", fontsize=18)
    plt.ylabel("Count", fontsize=16)
    plt.show()

print(job_disc_a2)
job_disc1()
```

```
Yes           145
No            145
Don`t know    145
Name: answer, dtype: int64
```



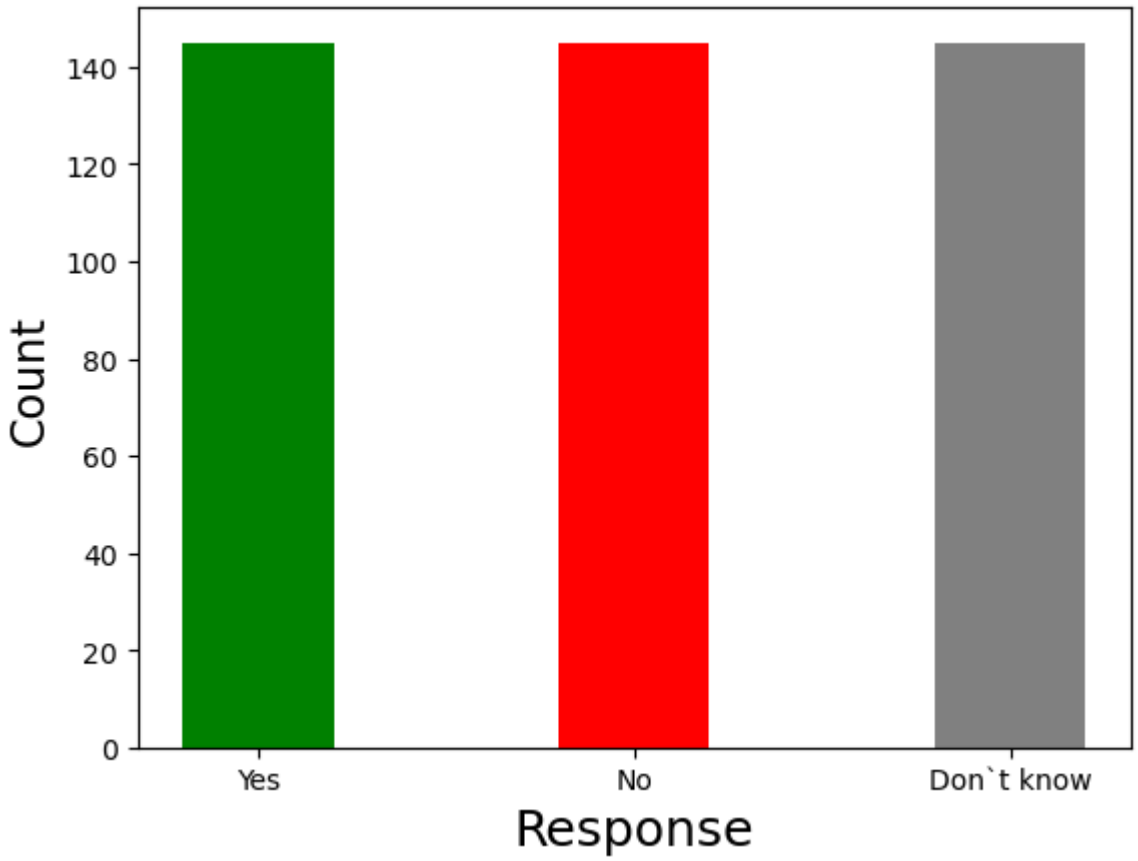
When asked about laws that forbids discrimination against persons because of their gender identity when applying for a job, this is the response of the respondents:

In [32]:
```
job_disc_3 = df.loc[df['question_code'] == 'd2']
job_disc_a3 = job_disc_2.loc[:,'answer'].value_counts()
job_disc_l3 = job_disc_2.loc[:,'answer'].unique()
colors = ["Green", "Red", "Grey"]
```

```
def job_disc2() :
    plt.bar(job_disc_l3, job_disc_a3, color=colors, width=0.4)
    plt.xlabel("Response", fontsize=18)
    plt.ylabel("Count", fontsize=16)
    plt.show()

print(job_disc_a3)
job_disc2()
```

```
Yes           145
No            145
Don`t know    145
Name: answer, dtype: int64
```



Correlation with Subset within LGBTQIA+ Community

After performing a cross-tabulation of the `subset` with the `answer` column of the dataframe (filtered only to answers under the question code `d1`), the writer has performed a Chi-Square test of Independence.

Hypotheses

H_0 = There is no correlation between the perception of the respondents regarding laws protecting discrimination against their sexual orientation and their subset within the LGBTQIA+ community.

H_a = There is a correlation between the perception of the respondents regarding laws protecting discrimination against their sexual orientation and their subset within the LGBTQIA+

community.
$p - value = 0.05$

In [33]:
```
d1_df = df.loc[df['question_code'] == 'd1']
ctab1 = pd.crosstab(d1_df.subset, d1_df.answer)
ctab1_cs = stats.chi2_contingency(ctab1)

print(ctab1)
print(" ")
print(ctab1_cs)
```

```
answer          Don`t know  No  Yes
subset
Bisexual men            29  29   29
Bisexual women          29  29   29
Gay                     29  29   29
Lesbian                 29  29   29
Transgender             29  29   29

Chi2ContingencyResult(statistic=0.0, pvalue=1.0, dof=8, expected_freq=array([[29., 2
9., 29.],
       [29., 29., 29.],
       [29., 29., 29.],
       [29., 29., 29.],
       [29., 29., 29.]]))
```

As findings show that the responses in `d1` have a $p - value$ of 1.0, the null hypothesis is not rejected. Therefore, there is no correlation between the perception of the respondents regarding laws protecting discrimination against their sexual orientation and their subset withing the LGBTQIA+ community.

We then performed a cross-tabulation of the `subset` with the `answer` column of the dataframe filtered only to answers under the question code `d2` and performed the same procedures.

Hypotheses

H_0 = There is no correlation between the perception of the respondents regarding laws protecting discrimination against their gender identity and their subset within the LGBTQIA+ community.
H_a = There is a correlation between the perception of the respondents regarding laws protecting discrimination against their gender identity and their subset within the LGBTQIA+ community.
$p - value = 0.05$

In [34]:
```
d2_df = df.loc[df['question_code'] == 'd2']
ctab2 = pd.crosstab(d2_df.subset, d2_df.answer)
ctab2_cs = stats.chi2_contingency(ctab1)

print(ctab2)
print(" ")
print(ctab2_cs)
```

```
answer          Don`t know  No  Yes
subset
Bisexual men            29  29   29
Bisexual women          29  29   29
Gay                     29  29   29
Lesbian                 29  29   29
Transgender             29  29   29

Chi2ContingencyResult(statistic=0.0, pvalue=1.0, dof=8, expected_freq=array([[29., 2
9., 29.],
       [29., 29., 29.],
       [29., 29., 29.],
       [29., 29., 29.],
       [29., 29., 29.]]))
```

As findings show that the responses in `d2` have a $p-value$ of 1.00, the null hypothesis is thus not rejected. Therefore, there is a correlation between the perception of the respondents regarding laws protecting discrimination against their gender identity and their subset withing the LGBTQIA+ community.

Correlation with Country of Origin

After performing a cross-tabulation of the `CountryCode` with the `answer` column of the dataframe (filtered only to answers under the question code `d1`), the writer has performed a Chi-Square test of Independence.

Hypotheses

H_0 = There is no correlation between the perception of the respondents regarding laws protecting discrimination against their sexual orientation and their country of origin.
H_a = There is a correlation between the perception of the respondents regarding laws protecting discrimination against their sexual orientation and their country of origin.
$p-value = 0.05$

In [35]:
```
d1_df = df.loc[df['question_code'] == 'd1']
ctab1a = pd.crosstab(d1_df.CountryCode, d1_df.answer)
ctab1a_cs = stats.chi2_contingency(ctab1a)

print(ctab1a)
print(" ")
print(ctab1a_cs)
```

```
answer          Don`t know  No  Yes
CountryCode
Austria                  5   5    5
Average                  5   5    5
Belgium                  5   5    5
Bulgaria                 5   5    5
Croatia                  5   5    5
Cyprus                   5   5    5
Czech Republic           5   5    5
Denmark                  5   5    5
Estonia                  5   5    5
Finland                  5   5    5
France                   5   5    5
Germany                  5   5    5
Greece                   5   5    5
Hungary                  5   5    5
Ireland                  5   5    5
Italy                    5   5    5
Latvia                   5   5    5
Lithuania                5   5    5
Luxembourg               5   5    5
Malta                    5   5    5
Netherlands              5   5    5
Poland                   5   5    5
Portugal                 5   5    5
Romania                  5   5    5
Slovakia                 5   5    5
Slovenia                 5   5    5
Spain                    5   5    5
Sweden                   5   5    5
United Kingdom           5   5    5

Chi2ContingencyResult(statistic=0.0, pvalue=1.0, dof=56, expected_freq=array([[5.,
5., 5.],
       [5., 5., 5.],
       [5., 5., 5.],
       [5., 5., 5.],
       [5., 5., 5.],
       [5., 5., 5.],
       [5., 5., 5.],
       [5., 5., 5.],
       [5., 5., 5.],
       [5., 5., 5.],
       [5., 5., 5.],
       [5., 5., 5.],
       [5., 5., 5.],
       [5., 5., 5.],
       [5., 5., 5.],
       [5., 5., 5.],
       [5., 5., 5.],
       [5., 5., 5.],
       [5., 5., 5.],
       [5., 5., 5.],
       [5., 5., 5.],
       [5., 5., 5.],
       [5., 5., 5.],
       [5., 5., 5.],
       [5., 5., 5.],
       [5., 5., 5.],
```

```
[5., 5., 5.],
[5., 5., 5.]]))
```

As findings show that the responses in `d1` have a $p-value$ of 1.00, the null hypothesis is thus not rejected. Therefore, there is a correlation between the perception of the respondents regarding laws protecting discrimination against their sexual orientation and their country of origin.

We then performed a cross-tabulation of the `CountryCode` with the `answer` column of the dataframe filtered only to answers under the question code `d2` and performed the same procedures.

Hypotheses

H_0 = There is no correlation between the perception of the respondents regarding laws protecting discrimination against their gender identity and their country of origin.
H_a = There is a correlation between the perception of the respondents regarding laws protecting discrimination against their gender identity and their country of origin.
$p-value = 0.05$

In [36]:
```
d2_df = df.loc[df['question_code'] == 'd2']
ctab2a = pd.crosstab(d2_df.CountryCode, d2_df.answer)
ctab2a_cs = stats.chi2_contingency(ctab2a)

print(ctab2a)
print(" ")
print(ctab2a_cs)
```

```
answer          Don`t know  No  Yes
CountryCode
Austria                  5   5    5
Average                  5   5    5
Belgium                  5   5    5
Bulgaria                 5   5    5
Croatia                  5   5    5
Cyprus                   5   5    5
Czech Republic           5   5    5
Denmark                  5   5    5
Estonia                  5   5    5
Finland                  5   5    5
France                   5   5    5
Germany                  5   5    5
Greece                   5   5    5
Hungary                  5   5    5
Ireland                  5   5    5
Italy                    5   5    5
Latvia                   5   5    5
Lithuania                5   5    5
Luxembourg               5   5    5
Malta                    5   5    5
Netherlands              5   5    5
Poland                   5   5    5
Portugal                 5   5    5
Romania                  5   5    5
Slovakia                 5   5    5
Slovenia                 5   5    5
Spain                    5   5    5
Sweden                   5   5    5
United Kingdom           5   5    5

Chi2ContingencyResult(statistic=0.0, pvalue=1.0, dof=56, expected_freq=array([[5.,
5., 5.],
       [5., 5., 5.],
       [5., 5., 5.],
       [5., 5., 5.],
       [5., 5., 5.],
       [5., 5., 5.],
       [5., 5., 5.],
       [5., 5., 5.],
       [5., 5., 5.],
       [5., 5., 5.],
       [5., 5., 5.],
       [5., 5., 5.],
       [5., 5., 5.],
       [5., 5., 5.],
       [5., 5., 5.],
       [5., 5., 5.],
       [5., 5., 5.],
       [5., 5., 5.],
       [5., 5., 5.],
       [5., 5., 5.],
       [5., 5., 5.],
       [5., 5., 5.],
       [5., 5., 5.],
       [5., 5., 5.],
       [5., 5., 5.],
       [5., 5., 5.],
```

```
       [5., 5., 5.],
       [5., 5., 5.]]))
```

As findings show that the responses in `d2` have a $p - value$ of 1.0, the null hypothesis is thus rejected. Therefore, there is a correlation between the perception of the respondents regarding laws protecting discrimination against their gender identity and their country of origin.

Conclusion

Selected LGBTQIA+ residents within the EU have expressed that there is a lack of legal protection for them when it comes to the workplace discrimination. However, these perceptions were not influenced depending on their subset within the LGBTQIA+ community and their country of origin.