NAME: OJA OLAWUNMI FLORENCE (500651567)

DATA AREA: VIOLENCE AGAINST WOMEN AND GIRLS

INFORMATION VISUALIZATION ASSESSMENT 3

Introduction:

Violence against women and girls is a serious issue affecting millions of people worldwide. As a result, it is important to understand the nature and scope of this problem in order to develop effective strategies for prevention and intervention. One way to do this is through the use of visualizations, which can help us see patterns and trends in the data, identify key factors that contribute to violence against women and girls, and communicate this information to others.

In this course, we have explored the analysis of this dataset, and did some in-depth descriptions such as; the types of data it is, and how is it stored, what are the ranges of each feature, what relationships are there in the data and so on. In assessment two of this module, we had the "Technical Design Plan" where alternate designs were explored, and critically thought through their suitability, using the Five Design-Sheet (FdS) method, and the Critique Design Survey (CDS) method which helped reflect on the topic. Several design ideas were featured basically to showcase some of the factors that contribute to people's beliefs in violence, such as the information or awareness limitation, which was shown using the radio-like sketch on the first sheet, to the impact of Education visualized through a line graph and legs with and without shoes on sheet 2, then location as another factor with a map sketch showing the rural and urban location with a bar chart comparing the violence rates in urban and rural areas. While analyzing the dataset, it was discovered that most females believe other females should be violated for the investigated reasons more than the males, so this was captured with a pie chart on sheet 4. Finally, the critique design sheet, which is sheet 5 featured the employment rate (unemployment) as one of the major factors that contribute to people perpetuating violence on women and girls and this was shown using a line graph and palm sketch which explains the unbalanced employment rate and other factors listed earlier.

However, the palm sketch will not be produced for the third assessment due to time constraints and some other limitations. A clustered bar chart and some choropleth maps were the final visualizations for this dataset.

The visualizations story using Python:

A critical reflection process was followed in this module to arrive at the visualizations in the poster.

1. **OBTAIN:** The dataset on violence against women and girls used in this module was gotten from the kaggle website **Violence Against Women & Girls | Kaggle,** it consists of 7 features and 2340 rows as shown on Table1.

Table 1. Sample of the dataset on Violence Against Women and Girls

```
In [272]: # Reading my dataset using pandas (OBTAIN)
          VAWG_df = pd.read_csv('VAWG_WestAfrica.csv')
          VAWG_df
```

Out[272]:

	Country	Gender	Demographics Question	Demographics Response	Question	Survey Year	Value
0	BEN	F	Education	Higher	... if she burns the food	2017	1.0
1	BEN	F	Education	Secondary	... if she burns the food	2017	9.5
2	BEN	F	Residence	Urban	... if she burns the food	2017	11.2
3	BEN	F	Employment	Unemployed	... if she burns the food	2017	11.7
4	BEN	F	Marital status	Never married	... if she burns the food	2017	11.8
...	...	...	...	...	...	...	...
2335	TGO	M	Residence	Urban	... if she goes out without telling him	2013	6.0
2336	TGO	M	Residence	Rural	... if she neglects the children	2013	15.6
2337	TGO	M	Residence	Urban	... if she neglects the children	2013	6.9
2338	TGO	M	Residence	Rural	... if she refuses to have sex with him	2013	7.1
2339	TGO	M	Residence	Urban	... if she refuses to have sex with him	2013	2.6

2. **DATA SCRUBBING:** The Question column of the dataset was transposed with the value column to make each of the violence cause a separate column with the 'value' which represent the percentage agreement to violence for each of the questions, as shown on Table 2.1

Table 2.1 Transposed Dataset on Violence Against Women and Girls

```
In [273]: #Transposing the dataset to make the Question values colums instead of rows
          VAWG_transposed_df = pd.pivot(VAWG_df, index=['Country', 'Gender', 'Demographics Question', 'Demographics Response', 'Survey Year'
                              columns='Question', values='Value').rename_axis().reset_index()
          VAWG_transposed_df = VAWG_transposed_df.rename_axis(index=None, columns=None)
          VAWG_transposed_df
```

Out[273]:

	Country	Gender	Demographics Question	Demographics Response	Survey Year	... for at least one specific reason	... if she argues with him	... if she burns the food	... if she goes out without telling him	... if she neglects the children	... if she refuses to have sex with him
0	BEN	F	Age	15-24	2017	29.8	19.1	13.2	18.7	20.5	11.4
1	BEN	F	Age	25-34	2017	32.7	21.7	13.4	22.2	23.5	12.9
2	BEN	F	Age	35-49	2017	33.7	22.3	15.5	23.7	24.5	14.5
3	BEN	F	Education	Higher	2017	4.7	3.0	1.0	2.4	2.8	1.4
4	BEN	F	Education	No education	2017	37.3	25.9	16.5	26.5	27.2	16.5
...	...	...	...	...	...	...	...	...	...	...	...
385	TGO	M	Marital status	Married or living together	2013	17.6	9.5	5.3	9.5	12.3	4.3
386	TGO	M	Marital status	Never married	2013	18.4	10.7	6.3	9.5	11.1	5.8
387	TGO	M	Marital status	Widowed, divorced, separated	2013	22.7	10.6	8.5	12.2	12.4	6.3
388	TGO	M	Residence	Rural	2013	23.4	13.2	8.5	12.5	15.6	7.1
389	TGO	M	Residence	Urban	2013	11.6	6.3	2.6	6.0	6.9	2.6

390 rows × 11 columns

2.2 Check for null values:

Table 2.2 checks for null values in the dataset

```
In [57]: print(VAWG_transposed_df.isnull().any().any())
         #this shows that there are no null values in the transposed

         False
```

2.3 Column Renames: The columns with longer names were renamed, based on the lecture on tables taught in class and to better manage the datasets. Table 2.3 shows the new structure of the data with the renamed columns.

Table 2.3 Dataset with renamed columns

```
In [6]: #Renaming the columns to shorter names
        VAWG = VAWG_transposed_df.rename(columns={'... if she argues with him': 'argues',
                                        '... if she burns the food': 'burns_food',
                                        '... if she neglects the children': 'neglect_children',
                                        '... if she refuses to have sex with him':'refuse_sex',
                                        '... if she goes out without telling him':'out_without_permission',
                                        'Demographics Response':'demographics',
                                        '... for at least one specific reason':'other',
                                        'Survey Year': 'year'})
        VAWG
```

Out[6]:

	Country	Gender	Demographics Question	demographics	year	other	argues	burns_food	out_without_permission	neglect_children	refuse_sex
0	BEN	F	Age	15-24	2017	29.8	19.1	13.2	18.7	20.5	11.4
1	BEN	F	Age	25-34	2017	32.7	21.7	13.4	22.2	23.5	12.9
2	BEN	F	Age	35-49	2017	33.7	22.3	15.5	23.7	24.5	14.5
3	BEN	F	Education	Higher	2017	4.7	3.0	1.0	2.4	2.8	1.4
4	BEN	F	Education	No education	2017	37.3	25.9	16.5	26.5	27.2	16.5
...	...	...	...	...	...	...	...	...	...	...	...
355	TGO	M	Marital status	Married or living together	2013	17.6	9.5	5.3	9.5	12.3	4.3
356	TGO	M	Marital status	Never married	2013	18.4	10.7	6.3	9.5	11.1	5.8
357	TGO	M	Marital status	Widowed, divorced, separated	2013	22.7	10.6	8.5	12.2	12.4	6.3
358	TGO	M	Residence	Rural	2013	23.4	13.2	8.5	12.5	15.6	7.1
359	TGO	M	Residence	Urban	2013	11.6	6.3	2.6	6.0	6.9	2.6

The demographic column was encoded to aid our computation for the bar chart visualization as shown in Table 2.4

Table 2.4 shows the structure of the dataset with a column for encoded demography

```
In [76]: #Encoding demographics column to numeric values
         VAWG_avg['demographics_encoded'] = VAWG_avg['demographics'].astype('category').cat.codes
         VAWG_avg
```

Out[76]:

	demographics	burns_food	argues	out_without_permission	neglect_children	refuse_sex	other	demographics_encoded
0	15-24	12.691667	27.625000	26.337500	27.575000	20.520833	40.516667	0
1	25-34	11.387500	26.362500	25.591667	26.345833	20.779167	38.816667	1
2	35-49	11.116667	25.487500	25.175000	25.725000	20.758333	37.329167	2
3	Employed for cash	11.100000	25.783333	24.804167	25.808333	20.258333	37.670833	3
4	Employed for kind	15.370833	31.391667	31.137500	32.200000	24.262500	46.325000	4
5	Higher	2.958333	8.475000	8.775000	10.504167	5.812500	17.091667	5
6	Married or living together	11.870833	27.008333	26.595833	27.183333	21.741667	39.275000	6
7	Never married	10.575000	24.025000	22.283333	24.037500	16.633333	36.316667	7
8	No education	14.483333	31.112500	31.045833	31.358333	26.391667	45.025000	8
9	Primary	12.666667	28.012500	26.829167	27.645833	21.429167	40.945833	9
10	Rural	14.808333	31.250000	30.595833	31.266667	25.012500	44.662500	10
11	Secondary	8.512500	21.400000	19.804167	22.116667	14.637500	33.470833	11
12	Unemployed	10.900000	24.425000	23.050000	24.183333	18.783333	35.958333	12
13	Urban	7.454167	20.020833	19.062500	20.358333	14.554167	31.641667	13
14	Widowed, divorced, separated	10.541667	23.945833	23.566667	24.441667	18.945833	36.404167	14

3. **ANALYSIS:** The average of the violence beliefs grouped by the different demographics was done as shown in Table 3.1. And average violence beliefs grouped by the country was captured Table 3.2

Table 3.1 shows the average violence beliefs by demographic

```
In [61]: VAWG_avg = VAWG.groupby('demographics')['burns_food','argues','out_without_permission','neglect_children',
                                    'refuse_sex', 'other'].mean().rename_axis('demographics').reset_index()
         VAWG_avg
```

C:\Users\Wunmi\AppData\Local\Temp\ipykernel_26436\1498402803.py:1: FutureWarning:

Indexing with multiple keys (implicitly converted to a tuple of keys) will be deprecated, use a list instead.

Out[61]:

	demographics	burns_food	argues	out_without_permission	neglect_children	refuse_sex	other
0	15-24	12.691667	27.625000	26.337500	27.575000	20.520833	40.516667
1	25-34	11.387500	26.362500	25.591667	26.345833	20.779167	38.816667
2	35-49	11.116667	25.487500	25.175000	25.725000	20.758333	37.329167
3	Employed for cash	11.100000	25.783333	24.804167	25.808333	20.258333	37.670833
4	Employed for kind	15.370833	31.391667	31.137500	32.200000	24.262500	46.325000
5	Higher	2.958333	8.475000	8.775000	10.504167	5.812500	17.091667
6	Married or living together	11.870833	27.008333	26.595833	27.183333	21.741667	39.275000
7	Never married	10.575000	24.025000	22.283333	24.037500	16.633333	36.316667
8	No education	14.483333	31.112500	31.045833	31.358333	26.391667	45.025000
9	Primary	12.666667	28.012500	26.829167	27.645833	21.429167	40.945833
10	Rural	14.808333	31.250000	30.595833	31.266667	25.012500	44.662500

Table 3.2 shows the average violence beliefs by country

```
In [25]: #Average acceptance of violence by country
         VAWG_country = VAWG.groupby('Country')['burns_food','out_without_permission','refuse_sex','argues','neglect_children',
                                    'other'].mean().rename_axis('country').reset_index()
         VAWG_country
```

C:\Users\Wunmi\AppData\Local\Temp\ipykernel_27700\535190797.py:2: FutureWarning:

Indexing with multiple keys (implicitly converted to a tuple of keys) will be deprecated, use a list instead.

Out[25]:

	country	burns_food	out_without_permission	refuse_sex	argues	neglect_children	other
0	BEN	9.656667	13.766667	9.016667	14.023333	14.620000	21.990000
1	BFA	7.893333	20.933333	13.070000	22.433333	22.950000	35.080000
2	CIV	13.596667	22.450000	15.260000	28.046667	29.453333	43.126667
3	GHA	5.210000	11.933333	8.883333	10.913333	14.943333	20.943333
4	GIN	19.226667	43.273333	34.896667	38.556667	46.280000	59.303333
5	GMB	8.140000	30.723333	31.023333	17.043333	26.616667	44.696667
6	LBR	4.810000	20.760000	7.163333	25.150000	22.866667	32.766667
7	MLI	15.450000	34.853333	40.313333	50.340000	35.633333	59.930000
8	NER	19.820000	26.840000	28.856667	32.390000	26.473333	40.886667
9	NGA	11.026667	19.276667	14.783333	17.323333	19.433333	29.896667
10	SEN	12.893333	21.693333	21.206667	24.873333	24.200000	34.313333
11	SLE	10.550000	34.376667	16.086667	33.760000	35.526667	46.346667
12	TGO	8.473333	13.293333	7.116667	14.253333	15.056667	23.006667

4. **VISUALIZATIONS:** A clustered bar chart and two choropleth maps were used to tell the story of violence against women and girls. Showcasing some key factors of violence and the regions of dominance in West Africa.

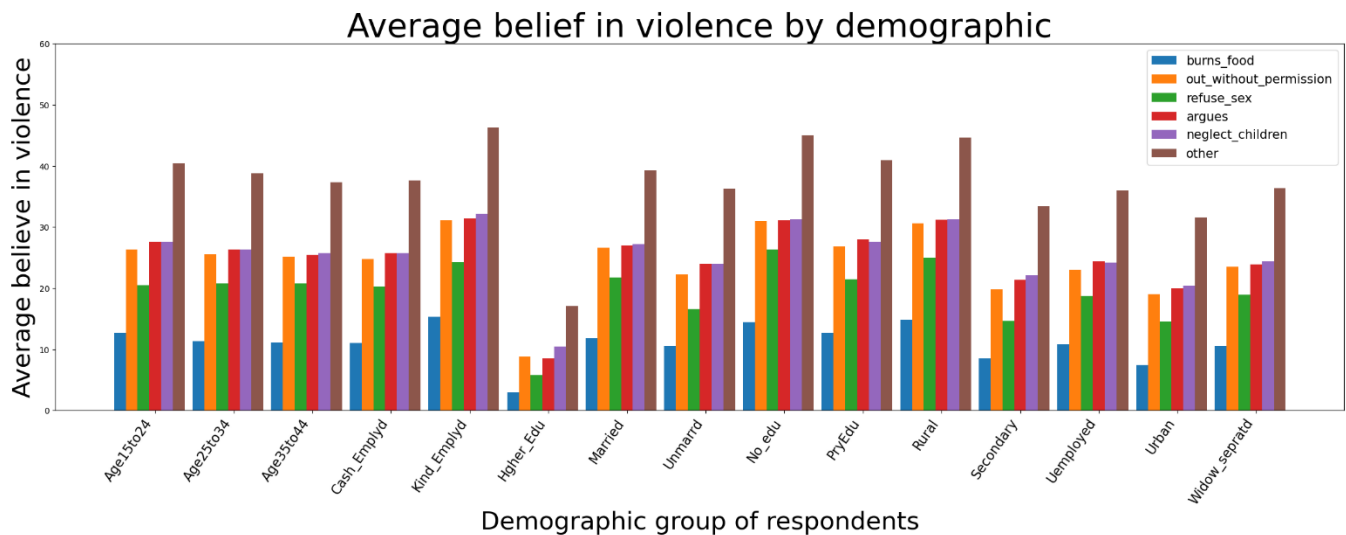## 4.1 First Visualization: Clustered Bar chart

Fig 4.1 shows the average believe in violence across West Africa when she by demographic groups

Interpretation:

The barchart compares the different demographics in the dataset such as Age, Education level, Residence, Marital level and their rate of beliefs in the different violence reasons investigated in the survey. From the analysis and as captured on the chart, across all the demographic groups, 'other' had the highest violence rate, which makes sense because they could be more serious causes of violence which were not captured in this experiment. Next to others are 'neglect children', 'arguing', and 'going out without permission', which maintained the same level across the demographic groups. However, the higher education demographic group and the urban group presented the least scale of all these causes of violence investigated. This analysis shows the impact of Education and Location on people's beliefs in violence.

Code:

```
# A barchart showing average belief in violence across different demographic
x1= VAWG_avg['demographics_encoded']
a= VAWG_avg['burns_food']
b= VAWG_avg['out_without_permission']
c= VAWG_avg['refuse_sex']
d= VAWG_avg['argues']
e= VAWG_avg['neglect_children']
f= VAWG_avg['other']
w= 0.15
demo = ['Age15to24','Age25to34','Age35to44','Cash_Emplyd','Kind_Emplyd','Hgher_Edu','Married','Unmarrd','No_edu','PryEdu','Rural','Secondary','Uemployed','Urban', 'Widow_sepratd']
mp.figure(figsize=(28,8))
mp.bar(x1, a, w, label = 'burns_food')
```

```
mp.bar(x1+w, b, w, label = 'out_without_permission')
mp.bar(x1+w*2, c, w, label = 'refuse_sex')
mp.bar(x1+w*3, d, w, label = 'argues')
mp.bar(x1+w*4, e, w, label = 'neglect_children')
mp.bar(x1+w*5, f, w, label = 'other')
mp.xticks(np.unique(x1+w*3), demo, ha='right', rotation=55, fontsize=16)
mp.xlabel("Demographic group of respondents", fontsize=20)
mp.ylabel("Average believe in violence", fontsize=20)
mp.title("Average belief in violence by demographic", fontsize=20)
mp.ylim(0,100)
mp.legend(fontsize=20)
mp.show()
```

4.2 Second Visualization: Choropleth Maps

From the bar graph above, the two top possible causes of violence after 'other' in this dataset are neglecting children and arguing. A choropleth map to show the part of West Africa where they believed in violence for the two identified reasons are shown below.

I initially thought of a rainbow-colored map as shown in Fig 4.2, for the first choropleth map to display the violence acceptance rate when she neglects children, however, I remembered the issues with rainbow colors taught in class.

First, it can be difficult for the human eye to distinguish between small differences in color intensity, particularly in the yellow-green region. Secondly, the color sequence does not follow a logical progression in hue or intensity, making it confusing for viewers to interpret data. Also, it may be problematic for people with colorblindness, who may have difficulty distinguishing between colors in the yellow-green region. As a result, I used a plasma, Fig 4.3 and a red Fig 4.4 for the two choropleths

Average violence acceptance rate in West African countries when she neglects children
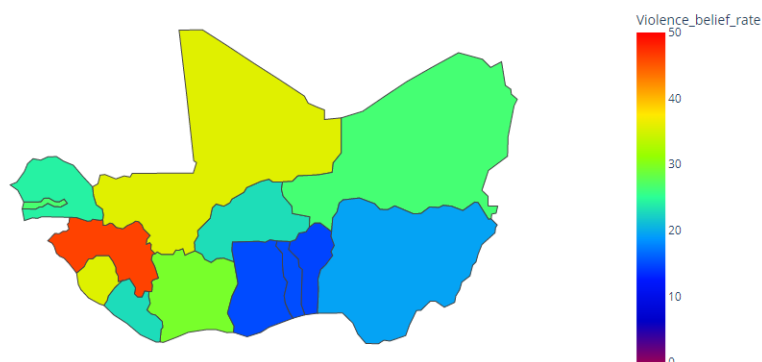
Fig 4.2 shows the average belief in violence across West Africa when she neglects children using a rainbow color scale, not a good color scale for visualization

Average violence acceptance rate in West African countries when she neglects children
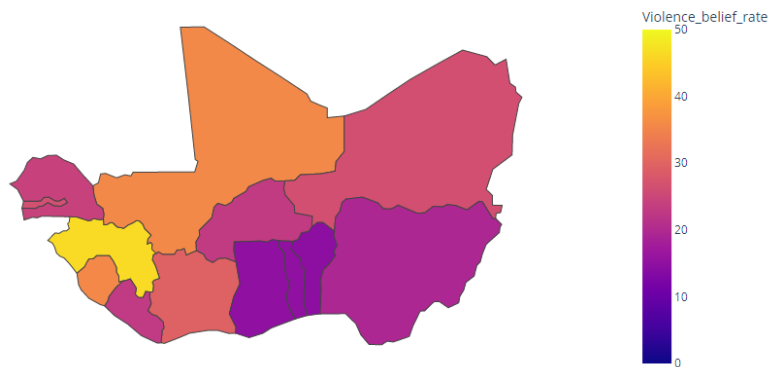

Violence_belief_rate

Fig 4.3 shows the average belief in violence across West Africa when she neglects children using a plasma color scale

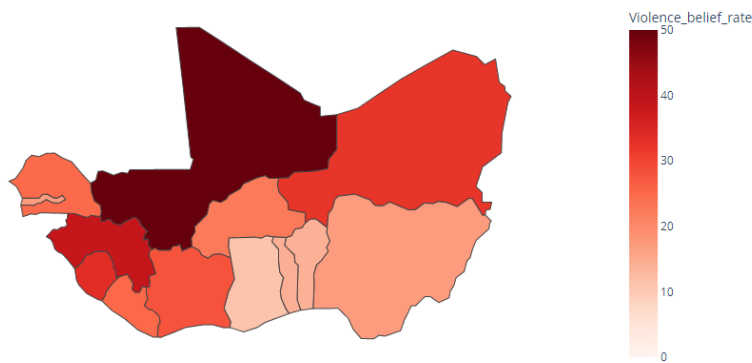Average violence acceptance rate in West African countries when she argues

Violence_belief_rate

Fig 4.4 shows the average belief in violence across West Africa when she argues using a red color scale

Interpretation:

The maps show Guinea presents the highest belief in violence when she neglects children with an average of 46% compared to other countries in West Africa, while Mali presents the highest belief in violence when she argues with him with an average of 50%.

Code:

```
import plotly.express as px
fig = px.choropleth(VAWG_country,
            locations="country",
            color="neglect_children",
            hover_name="country",
            color_continuous_scale="viridis",
            range_color=(0, 50))
fig.update_geos(fitbounds="locations", visible=False)
fig.update_layout(
    title={
        'text': "Average violence acceptance rate in West African countries when she neglects children",
```

```
        'font': {'size': 20, 'family': 'Arial'}
    },
    coloraxis_colorbar_title="Violence_belief_rate"
)
fig.show()
```

Code:
```
import plotly.express as px
fig = px.choropleth(VAWG_country, locations="country",
            color="argues",
            hover_name="country",
            color_continuous_scale="reds",
            range_color=(0, 50))
fig.update_geos(fitbounds="locations", visible=False)
fig.update_layout(
    title={
        'text': "Average violence acceptance rate in West African countries when she argues",
        'font': {'size': 20}
    },
    coloraxis_colorbar_title="Violence_belief_rate"
)fig.show()
```

5. CONCLUSION:

This assessment has provided valuable insights into the nature and scope of violence against women and girls, how violence beliefs vary across the different parts of West Africa and the role of education in this. From the analysis and visualization, we were able to draw the following insight;

- Mali has the highest acceptance of violence when a woman or girl argues with a man
- People from Guinea believed most compared to other parts of West Africa, that a woman should be hurt for neglecting children
- Education is a positive influence on violence, higher education correlates to a drop in the acceptance of violence

The use of visualizations has been effective in communicating this information and can inform the development of effective strategies for prevention and intervention. It is important that efforts are made to address this serious issue and ensure that women and girls are protected from violence, and help create a culture of non-violence and promote the human rights and dignity of all women and girls. We all have a role to play in ending violence against women and girls, and by working together, we can make a positive difference in the world.