Code git

DESCRIBING SEGMENTS

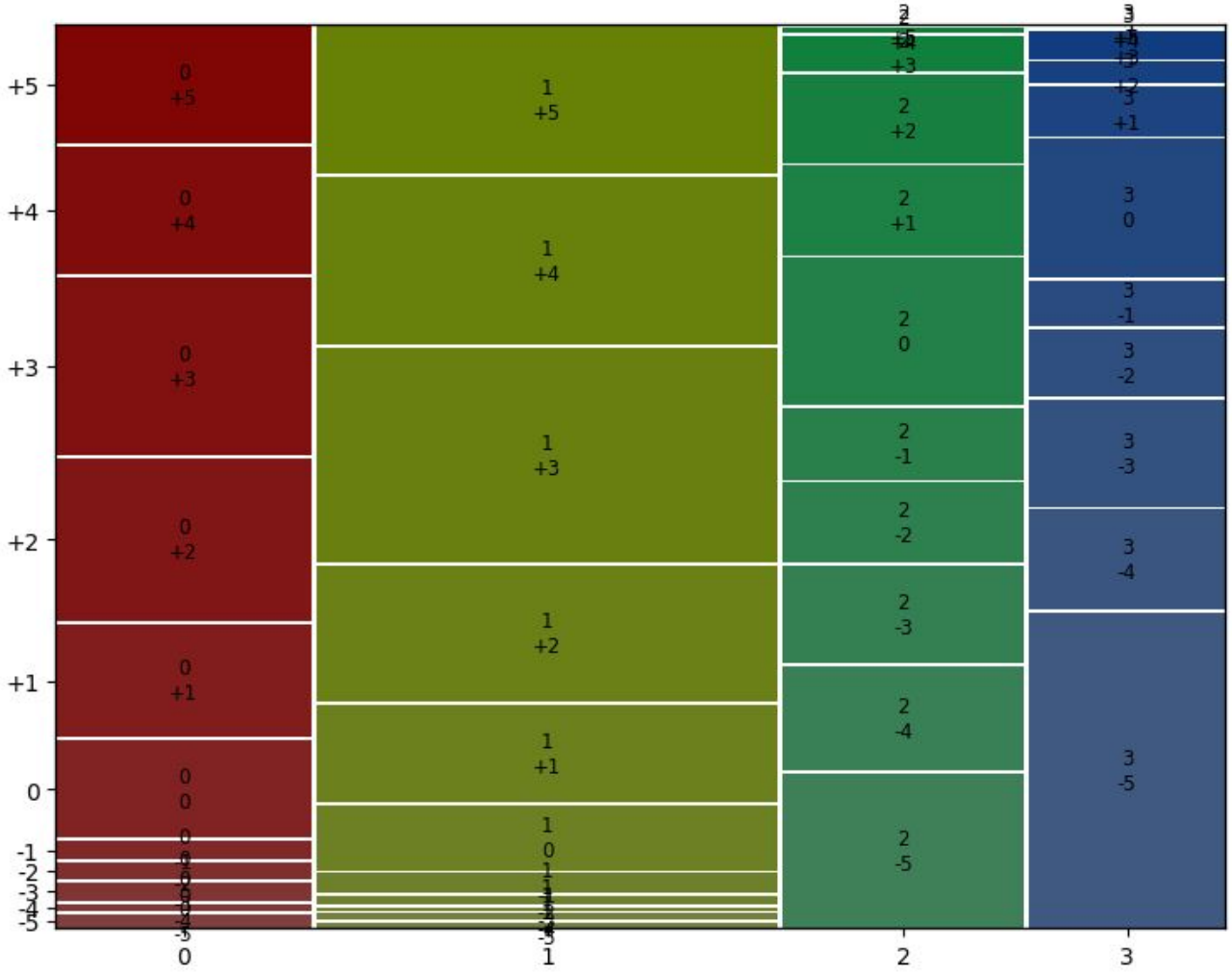
[23]:
```
#DESCRIBING SEGMENTS

from statsmodels.graphics.mosaicplot import mosaic
from itertools import product

crosstab =pd.crosstab(df['cluster_num'],df['Like'])
#Reordering cols
crosstab = crosstab[['-5','-4','-3','-2','-1','0','+1','+2','+3','+4','+5']]
crosstab
```

[23]:

Like	-5	-4	-3	-2	-1	0	+1	+2	+3	+4	+5
cluster_num											
0	5	3	7	6	7	36	42	60	66	47	44
1	4	4	2	6	13	43	65	90	143	111	99
2	54	36	34	28	25	51	31	31	12	2	0
3	89	28	30	19	13	39	14	6	8	0	0

Code git

[24]:
```
#MOSAIC PLOT
plt.rcParams['figure.figsize'] = (9,7)
mosaic(crosstab.stack())
plt.show()
```



Code git

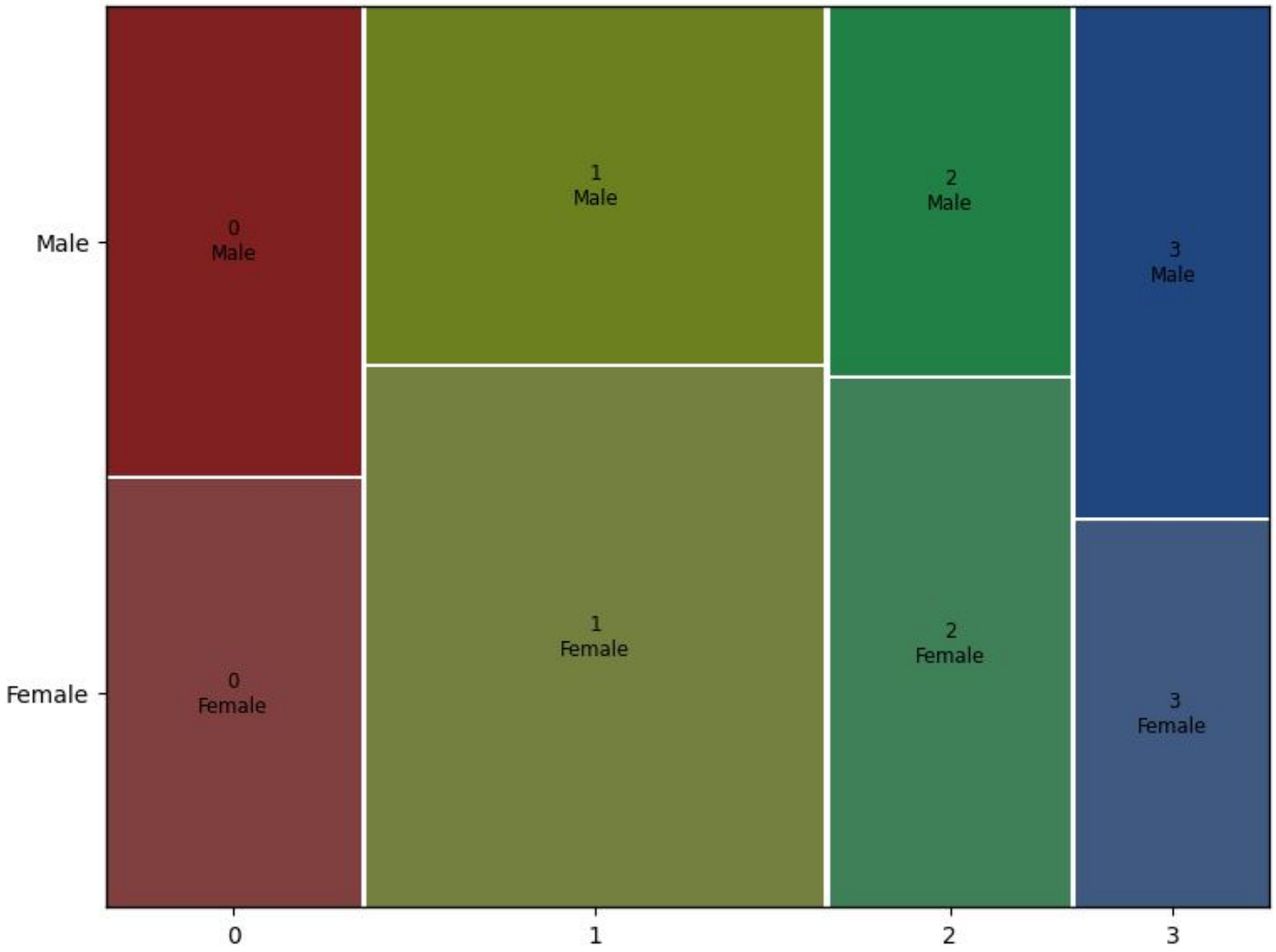
[25]:
```
#Mosaic plot gender vs segment
crosstab_gender =pd.crosstab(df['cluster_num'],df['Gender'])
crosstab_gender
```

[25]:

Gender	Female	Male
cluster_num		
0	154	169
1	349	231
2	179	125
3	106	140

Code git

[26]:
```
plt.rcParams['figure.figsize'] = (9,7)
mosaic(crosstab_gender.stack())
plt.show()
```



Code git

[27]:
```
#box plot for age

sns.boxplot(x="cluster_num", y="Age", data=df)
```

[27]:
```
<AxesSubplot:xlabel='cluster_num', ylabel='Age'>
```