

01_TwiBot_20_histograms

August 19, 2023

```
[1]: import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
```

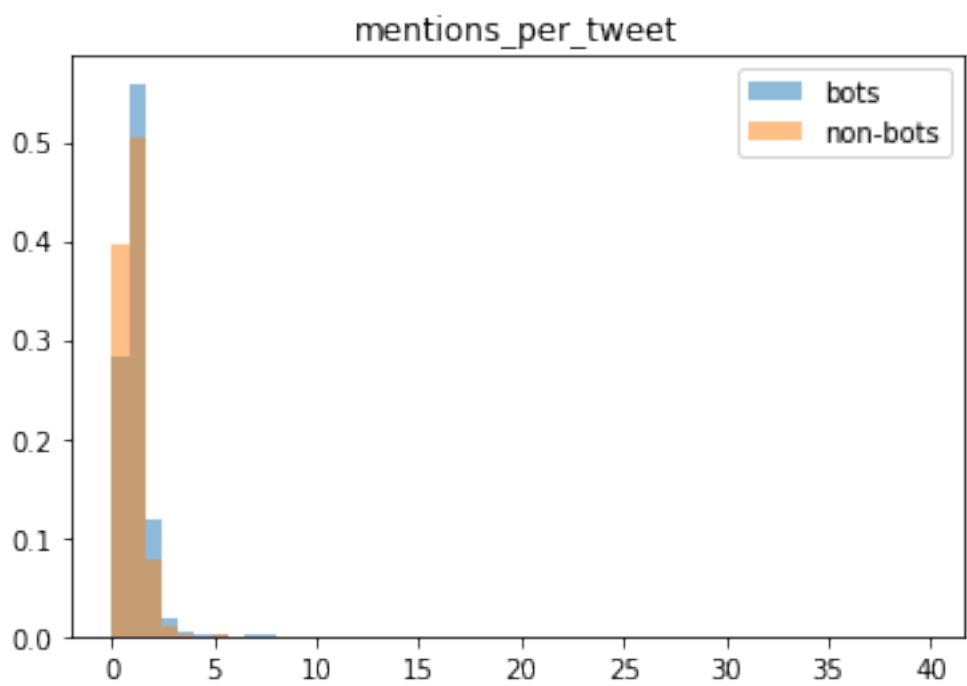
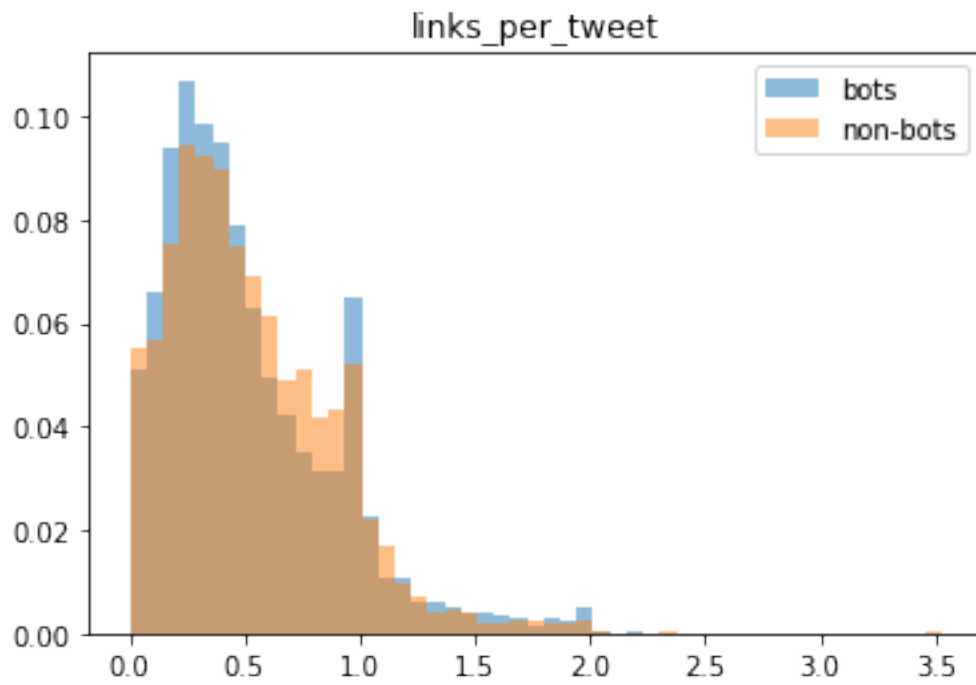
0.1 TwiBot-20

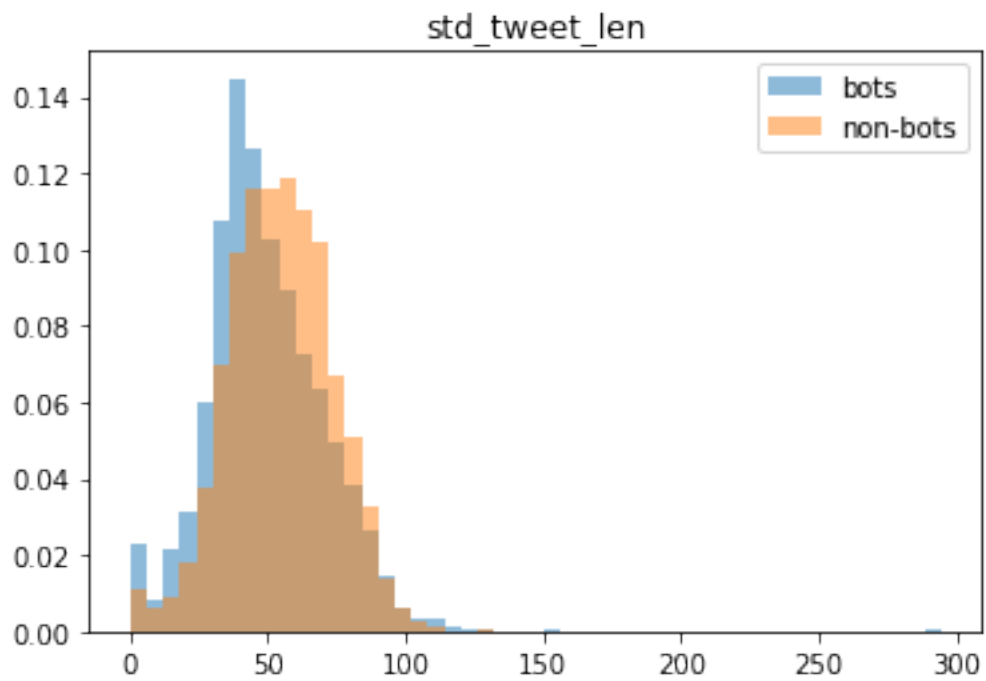
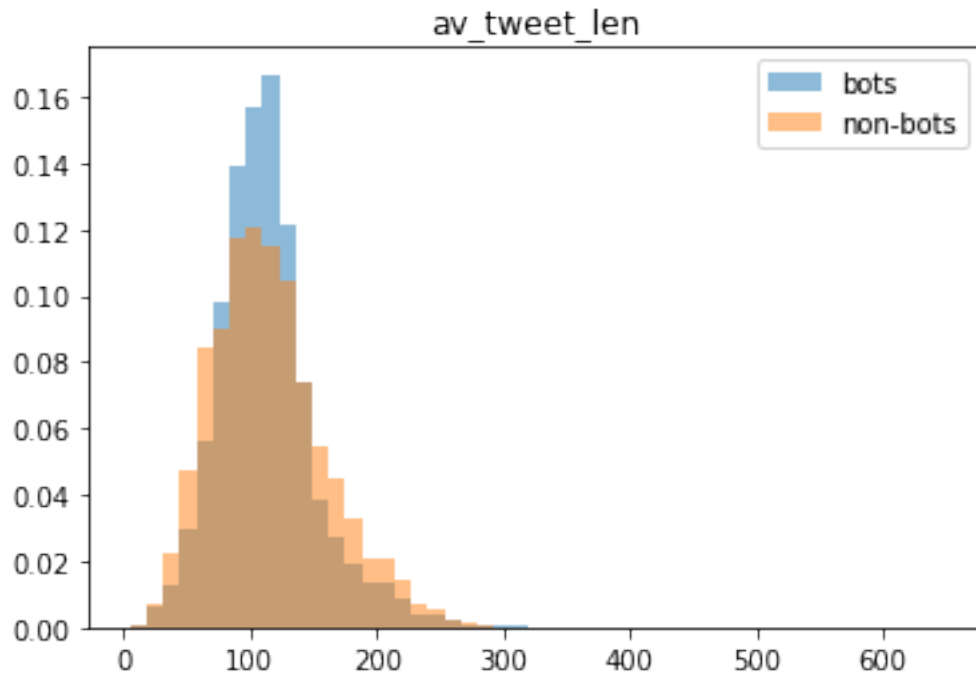
0.1.1 NLP data

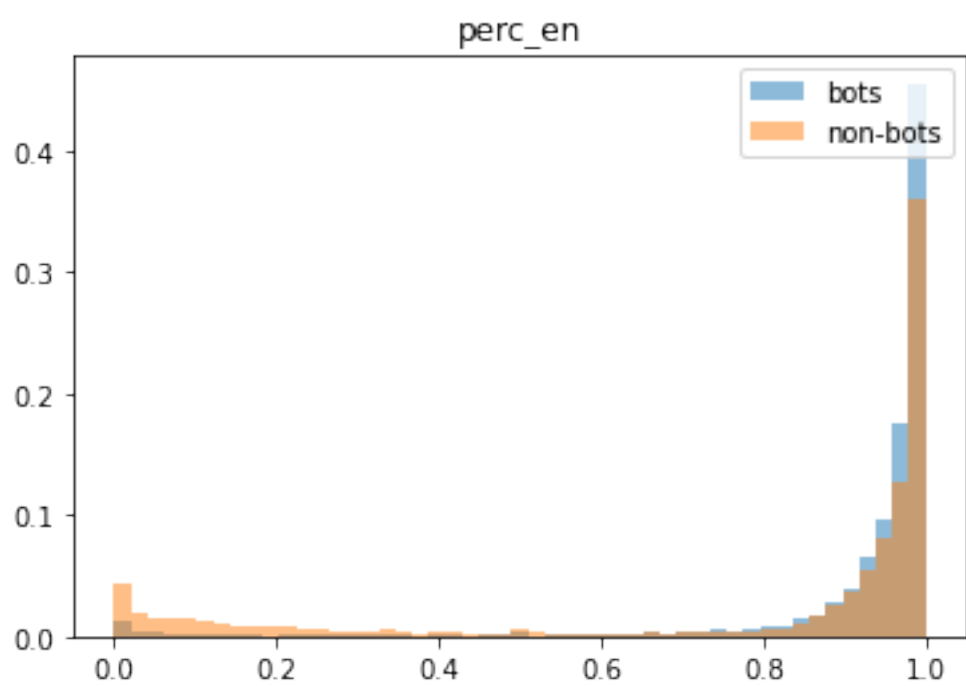
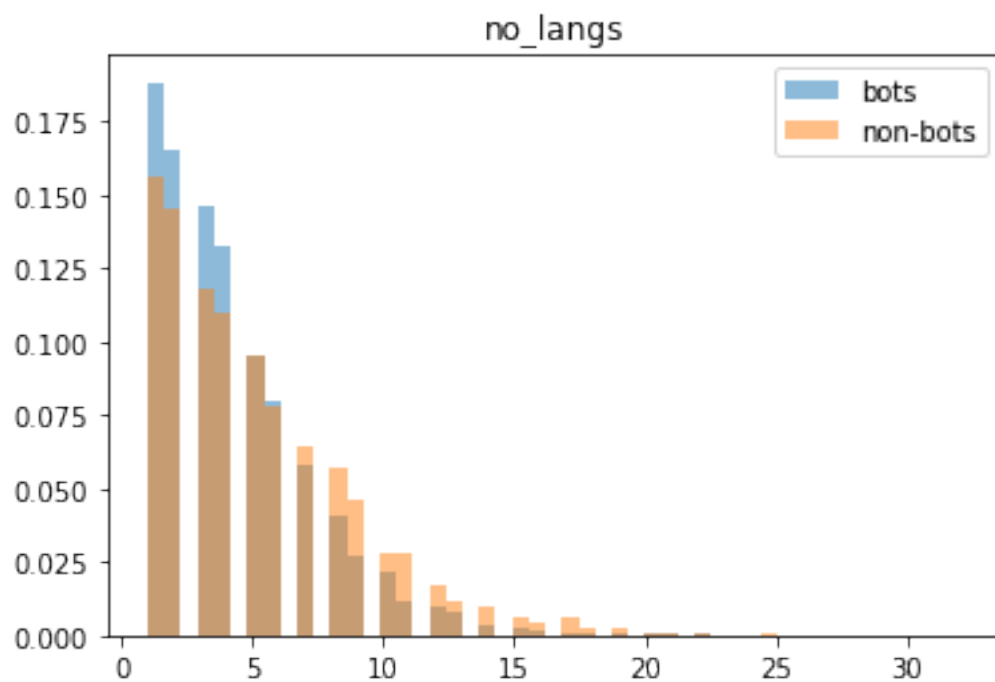
```
[2]: nlp = pd.read_csv("twibot_20_nlp.csv")
for col in nlp.columns[1:-1]:
    x = nlp.loc[nlp["label"] == 1, col]
    y = nlp.loc[nlp["label"] == 0, col]

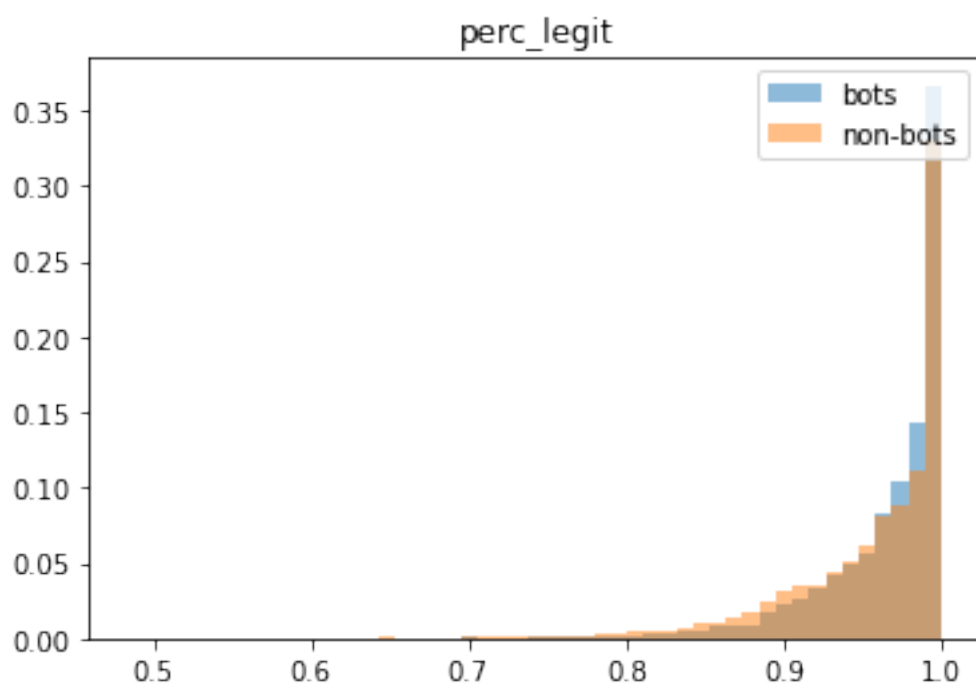
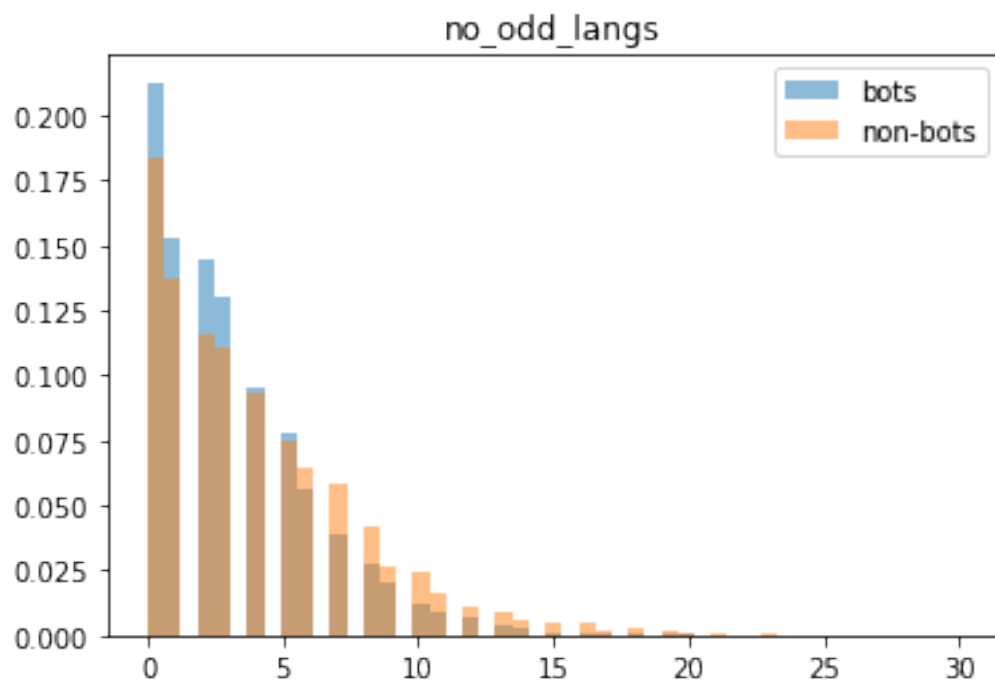
    bins = np.linspace(min(nlp[col]), max(nlp[col]), 50)

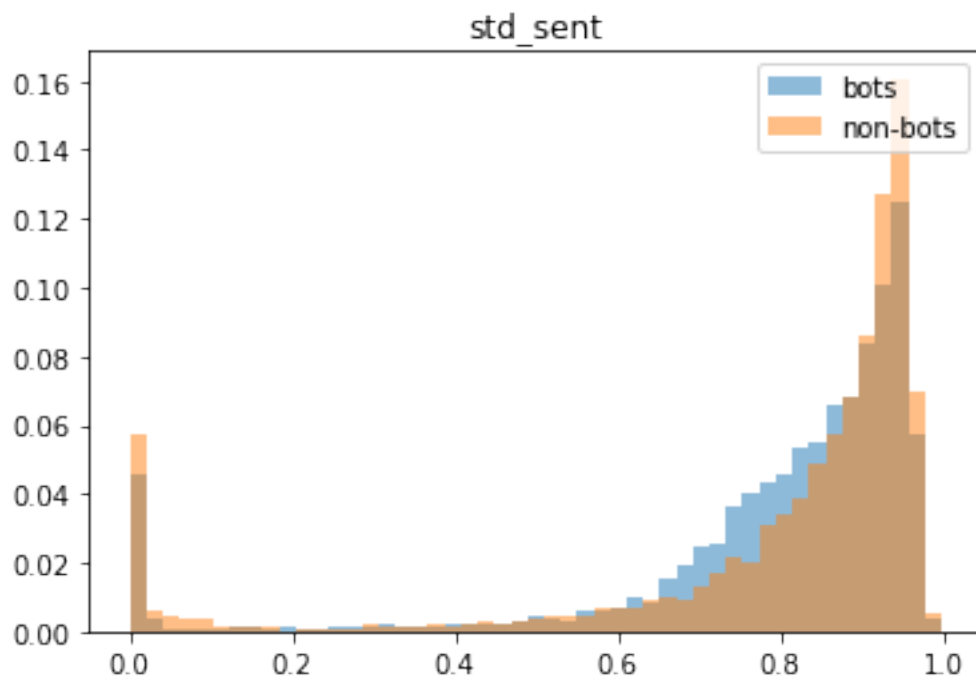
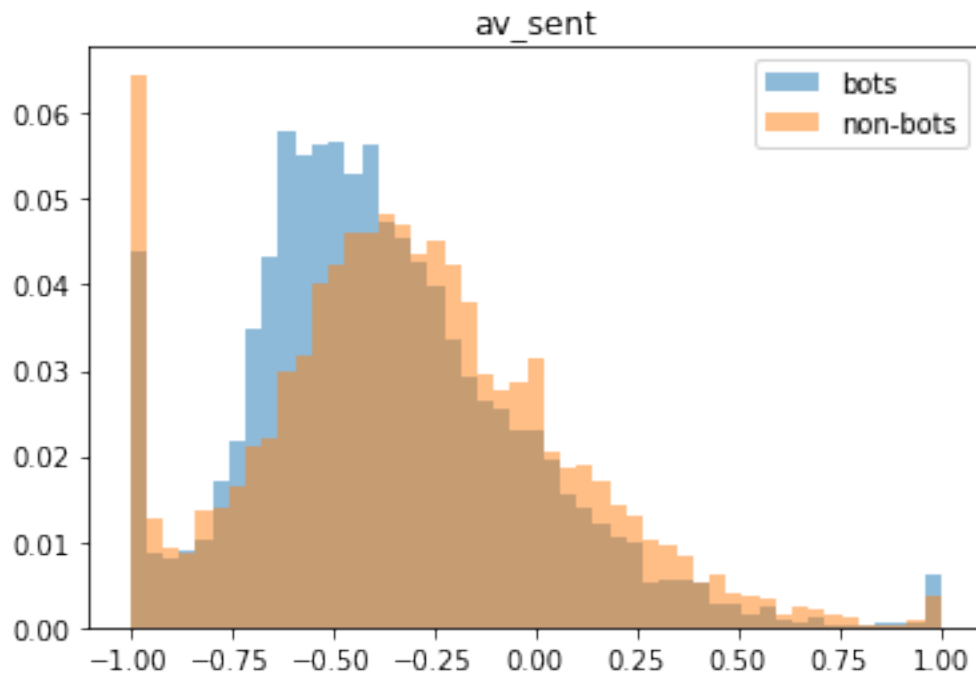
    plt.hist(round(x, 5), bins, alpha=0.5, label='bots', weights=np.
↪ones(len(x))/len(x))
    plt.hist(y, bins, alpha=0.5, label='non-bots', weights=np.ones(len(y))/
↪len(y))
    plt.legend(loc='upper right')
    plt.title(col)
    #plt.savefig(f'./charts/{col}.pdf')
    plt.show()
```

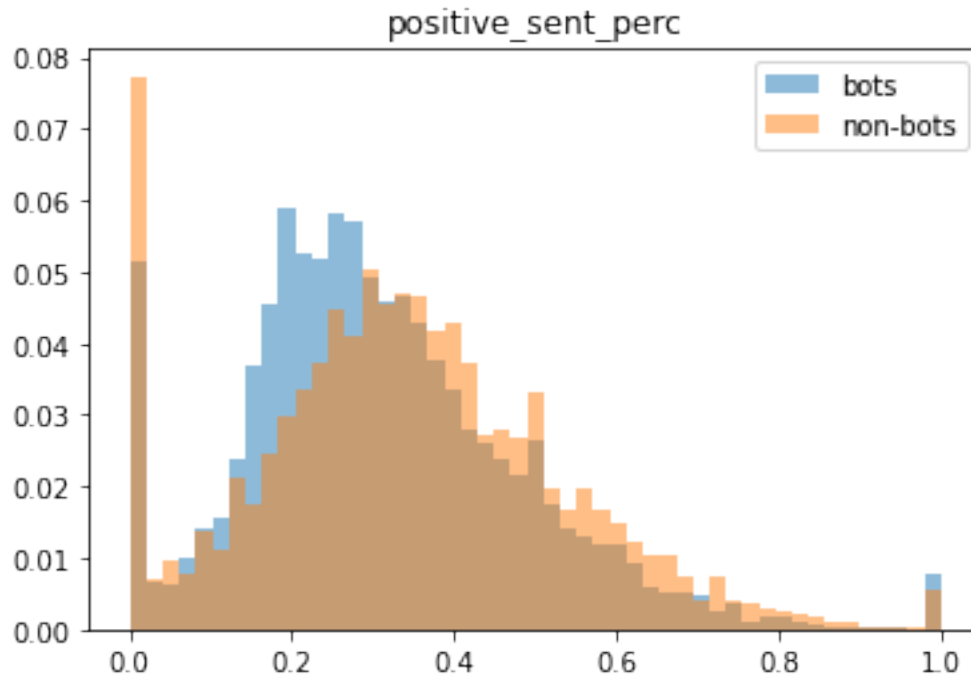












0.1.2 node statistics

```
[3]: twibot_gf = pd.read_csv("twibot_20_gf.csv")
```

```
[4]: twibot_gf = pd.read_csv("twibot_20_gf.csv")
for col in twibot_gf.columns[1:-1]:
    x = twibot_gf.loc[twibot_gf["label"] == 1, col]
    y = twibot_gf.loc[twibot_gf["label"] == 0, col]

    bins = np.linspace(min(twibot_gf[col]), max(twibot_gf[col]), 50)

    plt.hist(round(x, 5), bins, alpha=0.5, label='bots', weights=np.
    ↪ ones(len(x))/len(x))
    plt.hist(y, bins, alpha=0.5, label='non-bots', weights=np.ones(len(y))/
    ↪ len(y))
    plt.legend(loc='upper right')
    plt.title(col)
    #plt.savefig(f'./charts/{col}.pdf')
    plt.show()
```

