2_run_report

October 8, 2021

[1]:
```
%matplotlib inline
from matplotlib import pyplot as plt
import pandas as pd
```

[2]:
```
dfp = pd.read_csv('model_output.csv')
```

[3]:
```
dfp.columns
```

[3]:
```
Index(['GLOBALEVENTID', 'SQLDATE', 'MonthYear', 'Year', 'FractionDate',
       'Actor1Code', 'Actor1Name', 'Actor1CountryCode', 'Actor1KnownGroupCode',
       'Actor1EthnicCode', 'Actor1Religion1Code', 'Actor1Religion2Code',
       'Actor1Type1Code', 'Actor1Type2Code', 'Actor1Type3Code', 'Actor2Code',
       'Actor2Name', 'Actor2CountryCode', 'Actor2KnownGroupCode',
       'Actor2EthnicCode', 'Actor2Religion1Code', 'Actor2Religion2Code',
       'Actor2Type1Code', 'Actor2Type2Code', 'Actor2Type3Code', 'IsRootEvent',
       'EventCode', 'EventBaseCode', 'EventRootCode', 'QuadClass',
       'GoldsteinScale', 'NumMentions', 'NumSources', 'NumArticles', 'AvgTone',
       'Actor1Geo_Type', 'Actor1Geo_FullName', 'Actor1Geo_CountryCode',
       'Actor1Geo_ADM1Code', 'Actor1Geo_ADM2Code', 'Actor1Geo_Lat',
       'Actor1Geo_Long', 'Actor1Geo_FeatureID', 'Actor2Geo_Type',
       'Actor2Geo_FullName', 'Actor2Geo_CountryCode', 'Actor2Geo_ADM1Code',
       'Actor2Geo_ADM2Code', 'Actor2Geo_Lat', 'Actor2Geo_Long',
       'Actor2Geo_FeatureID', 'ActionGeo_Type', 'ActionGeo_FullName',
       'ActionGeo_CountryCode', 'ActionGeo_ADM1Code', 'ActionGeo_ADM2Code',
       'ActionGeo_Lat', 'ActionGeo_Long', 'ActionGeo_FeatureID', 'DATEADDED',
       'SOURCEURL', 'Actor1__model_time_in_ms',
       'Actor1_release_harness_version', 'Actor1_release_model_version',
       'Actor1_release_model_version_number', 'Actor1_request_id',
       'Actor1_result_class1', 'Actor1_result_class2', 'Actor1_timing',
       'Actor2__model_time_in_ms', 'Actor2_release_harness_version',
       'Actor2_release_model_version', 'Actor2_release_model_version_number',
       'Actor2_request_id', 'Actor2_result_class1', 'Actor2_result_class2',
       'Actor2_timing'],
      dtype='object')
```

[4]:
```
dfp.head()
```

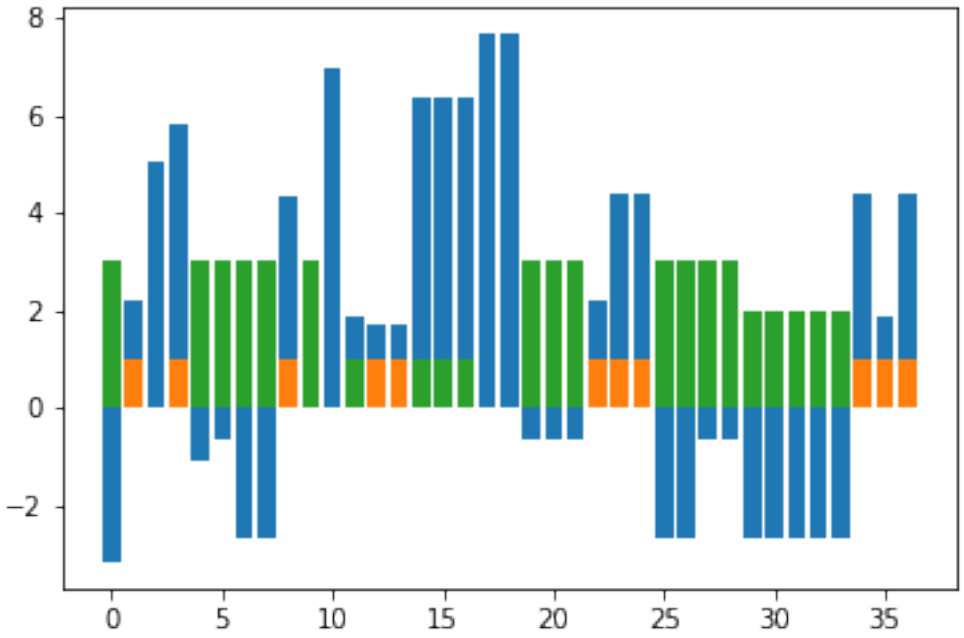
[4]:
```
   GLOBALEVENTID    SQLDATE  MonthYear  Year  FractionDate Actor1Code  \
0      838788881  4/16/2018     201804  2018     2018.2904        EDU
1      838788882  4/16/2018     201804  2018     2018.2904        EDU
2      838788884  4/16/2018     201804  2018     2018.2904        GOV
3      838788885  4/16/2018     201804  2018     2018.2904        GOV
4      838788886  4/16/2018     201804  2018     2018.2904        GOV

   Actor1Name Actor1CountryCode Actor1KnownGroupCode Actor1EthnicCode  …  \
0   ECONOMIST               NaN                  NaN              NaN  …
1     STUDENT               NaN                  NaN              NaN  …
2  GOVERNMENT               NaN                  NaN              NaN  …
3  GOVERNMENT               NaN                  NaN              NaN  …
4      MINIST               NaN                  NaN              NaN  …

  Actor1_result_class2 Actor1_timing Actor2__model_time_in_ms  \
0                    3      0.078201                     1001
1                    0      0.079155                        0
2                    3      0.120401                     1001
3                    3      0.074387                        0
4                    3      0.069141                        0

  Actor2_release_harness_version  Actor2_release_model_version  \
0                            0.1      5ec427ae4cedfd0008830f07
1                            0.1      5ec427ae4cedfd0008830f07
2                            0.1      5ec427ae4cedfd0008830f07
3                            0.1      5ec427ae4cedfd0008830f07
4                            0.1      5ec427ae4cedfd0008830f07

  Actor2_release_model_version_number Actor2_request_id Actor2_result_class1  \
0                                   4  RSFCLN4EK35XOU0V                 True
1                                   4  RHC58LU0X41VKWE2                 True
2                                   4  FHZU2BCO10AY9LIS                 True
3                                   4  3FQ158RWS97IJCLH                 True
4                                   4  BMWNI3P4LV6FSJUR                False

  Actor2_result_class2 Actor2_timing
0                    3   1001.194239
1                    4      0.066280
2                    3   1001.168489
3                    3      0.053406
4                    3      0.045300

[5 rows x 77 columns]
```

[5]:
```
actor1_name = 'EDU'
y_avgTone = list(dfp[dfp['Actor1Code'] == actor1_name]['AvgTone'])
y_cls1 = list(dfp[dfp['Actor1Code'] == actor1_name]['Actor1_result_class1'])
```

```
y_cls2 = list(dfp[dfp['Actor1Code'] == actor1_name]['Actor1_result_class2'])
```

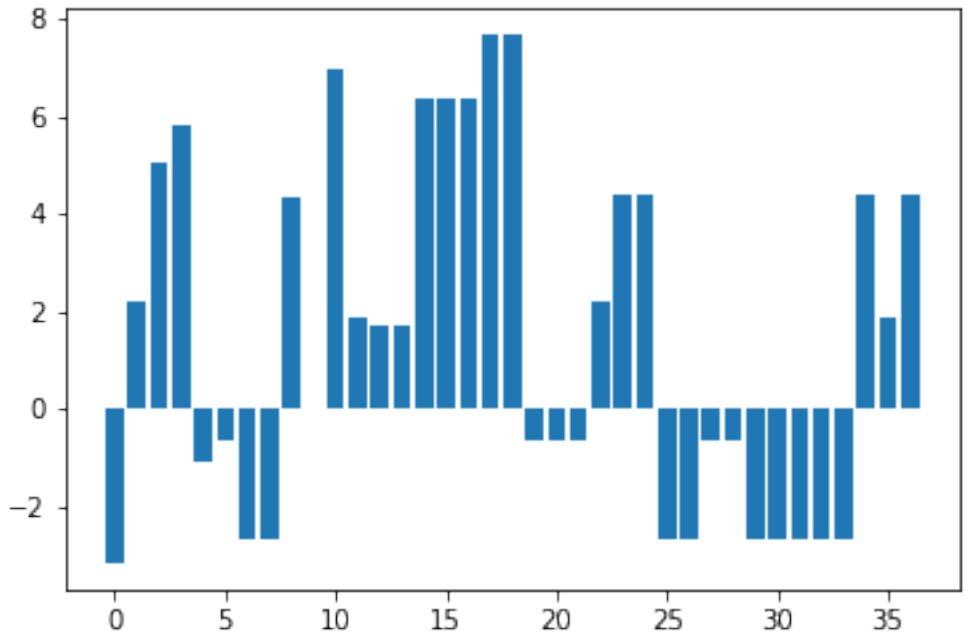
[6]:
```
plt.bar(range(len(y_avgTone)), y_avgTone)
plt.bar(range(len(y_avgTone)), y_cls1)
plt.bar(range(len(y_avgTone)), y_cls2)
```

[6]: <BarContainer object of 37 artists>



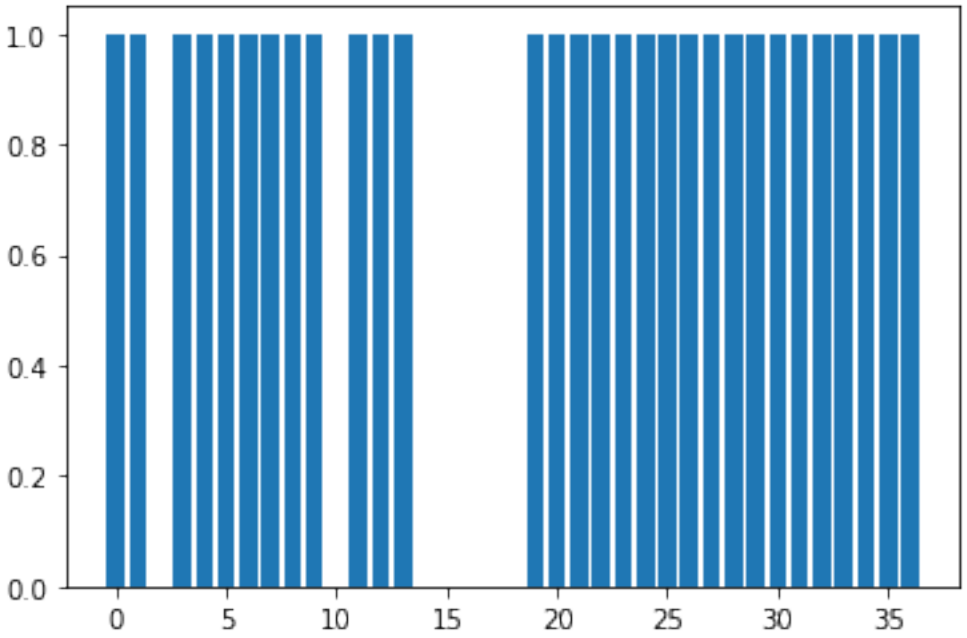
[7]:
```
plt.bar(range(len(y_avgTone)), y_avgTone)
```

[7]: <BarContainer object of 37 artists>



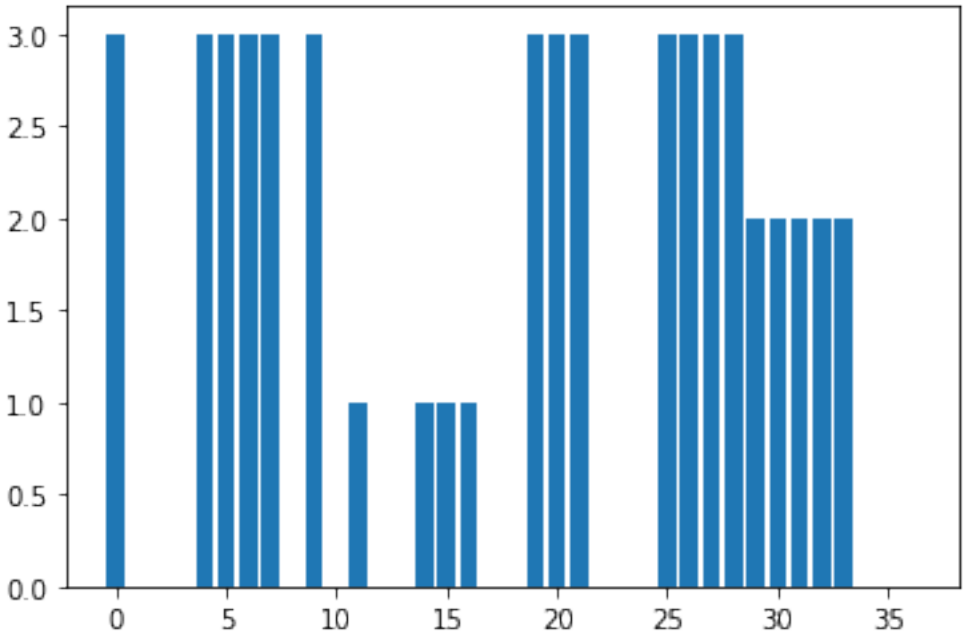
[8]:
```
plt.bar(range(len(y_avgTone)), y_cls1)
```

[8]: <BarContainer object of 37 artists>



[9]:
```
plt.bar(range(len(y_avgTone)), y_cls2)
```

[9]:
```
<BarContainer object of 37 artists>
```



1 Most Common Actors

[11]:
```
from collections import Counter
counter = Counter(dfp[dfp['Actor1Code'] == actor1_name]['Actor2Code'])
counter.most_common(n=10)
```

[11]:
```
[('USA', 16),
 ('CAN', 3),
 ('GOV', 2),
 ('BUS', 2),
 ('LEG', 2),
 ('PAKEDU', 2),
 ('AUS', 1),
 ('CRM', 1),
 ('FRA', 1),
 ('GBR', 1)]
```

2 Actor Correlations

[12]:
```
actor_code = 'USA'

actor_df = dfp[(
    dfp['Actor1Code'] == actor_code) |
    (dfp['Actor2Code'] == actor_code)
][['GLOBALEVENTID', 'Actor1Code', 'Actor2Code', 'AvgTone',␣
 ↪'Actor1_result_class1', 'Actor1_result_class2', 'Actor2_result_class1',␣
 ↪'Actor2_result_class2']]
```

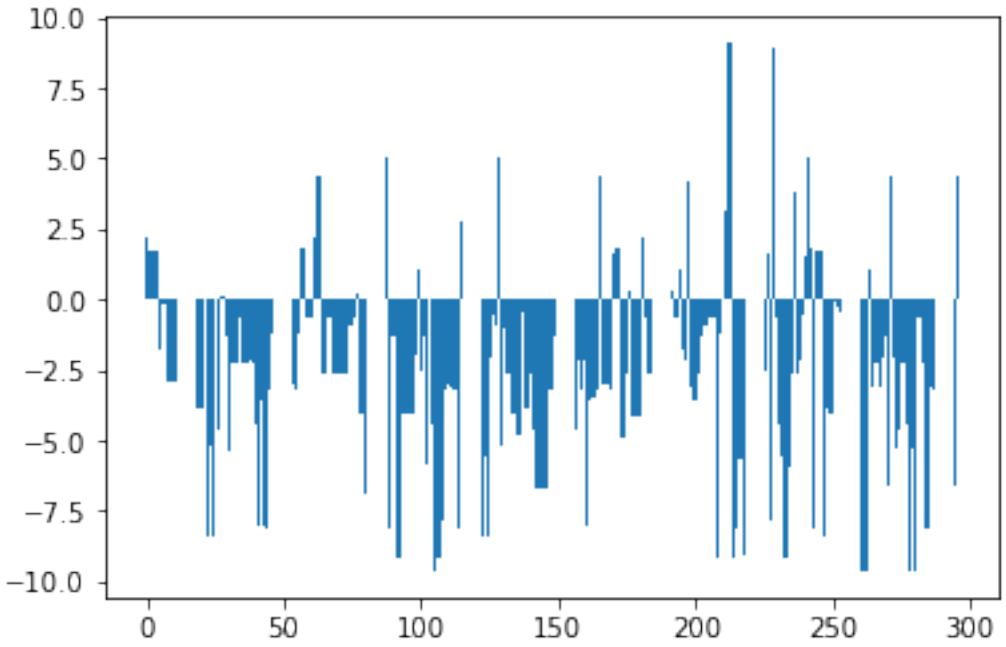
[13]:
```
actor_df.head(5)
```

[13]:

	GLOBALEVENTID	Actor1Code	Actor2Code	AvgTone	Actor1_result_class1	Actor1_result_class2	Actor2_result_class1	Actor2_result_class2
1	838788882	EDU	USA	2.214022	True	0	True	4
8	838788896	USA	OPP	1.692748	True	4	True	1
9	838788897	USA	OPP	1.692748	True	4	True	1
10	838788898	USA	OPP	1.692748	True	4	True	1
11	838788899	USA	OPP	1.692748	True	4	True	1

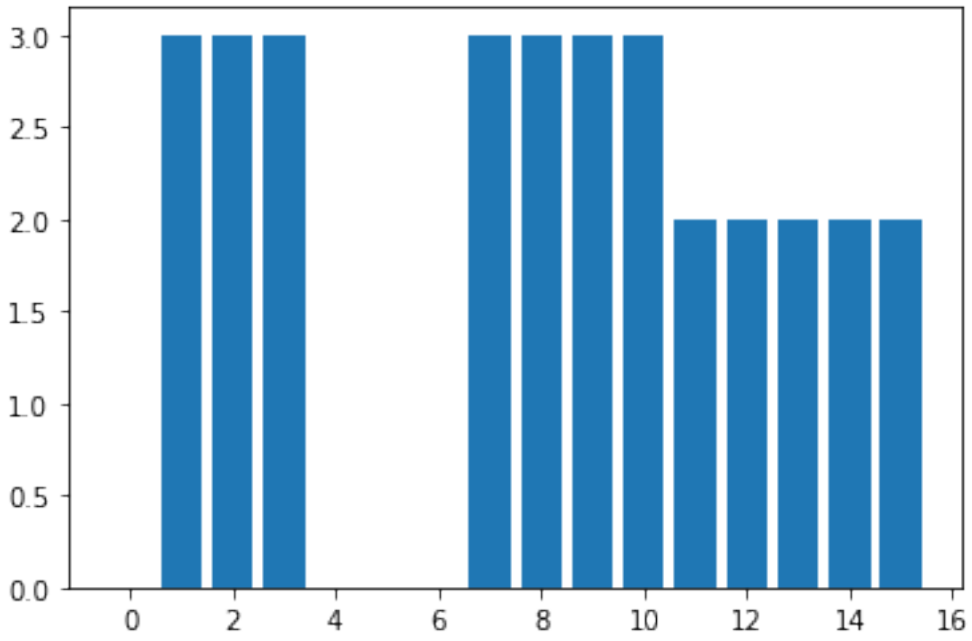
[14]:
```
y = list(actor_df['AvgTone'])
plt.bar(range(len(y)), y)
```

[14]:
```
<BarContainer object of 296 artists>
```



[16]:
```
y = list(actor_df[actor_df['Actor1Code'] ==␣
 ↪actor1_name]['Actor1_result_class2'])
plt.bar(range(len(y)), y)
```

[16]: <BarContainer object of 16 artists>



[17]:
```
y = list(actor_df[actor_df['Actor2Code'] ==␣
 ↪actor1_name]['Actor2_result_class2'])
plt.bar(range(len(y)), y)
```

[17]: <BarContainer object of 22 artists>