Why is India a significant case study (in comparison to Iran & Indonesia)?

In [1]:
```
# Import pandas, numPy, & dataset
import pandas as pd
import numpy as np
path = r"C:\Users\berns\OneDrive\Desktop\2019Report.csv"
df = pd.read_csv(path)
df.head(10)
```

Out[1]:

	ID	start_date	end_date	duration	Info_source	news_link
0	1	3/28/2018	7/13/2019	472	News Media Article	https://qz.com/1247234/chad-has-blocked-social...
1	2	10/17/2018	3/17/2019	151	News Media Article	https://almushahid.net/41057/
2	3	12/20/2018	1/10/2019	21	News Media Article	https://jordanopensource.org/technical-report/...
3	4	12/21/2018	2/26/2019	67	News Media Article	https://www.reuters.com/article/us-sudan-prote...
4	5	12/31/2018	1/20/2019	20	News Media Article	https://www.theguardian.com/world/2019/jan/01/...
5	6	1/1/2019	1/2/2019	1	News Media Article	https://www.indiatoday.in/india/video/bhima-ko...
6	7	1/2/2019	NaN	NaN	News Media Article	https://www.newsx.com/national/pulwama-encount...
7	8	1/3/2019	NaN	NaN	News media Article	https://www.newsx.com/national/pulwama-encount...
8	9	1/5/2019	NaN	NaN	News Media Article	https://freepresskashmir.com/2019/01/05/search...
9	10	1/7/2019	1/8/2019	1	CSO KIO Partners	https://ooni.torproject.org/post/gabon-interne...

10 rows × 43 columns

In [2]:
```
df.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 213 entries, 0 to 212
Data columns (total 43 columns):
 #   Column                              Non-Null Count  Dtype  
---  ------                              --------------  -----  
 0   ID                                  213 non-null    int64  
 1   start_date                          213 non-null    object 
 2   end_date                            145 non-null    object 
 3   duration                            93 non-null     object 
 4   Info_source                         213 non-null    object 
 5   news_link                           212 non-null    object 
 6   continent                           213 non-null    object 
 7   sub-region                          0 non-null      float64
 8   country                             213 non-null    object 
 9   geo_scope                           207 non-null    object 
 10  area_name                           207 non-null    object 
 11  ordered_by                          209 non-null    object 
 12  decision_maker                      129 non-null    object 
 13  shutdown_type_new                   213 non-null    object 
 14  affected_network                    213 non-null    object 
 15  full or service-based               213 non-null    object 
 16  Facebook_affected                   184 non-null    object 
 17  Twitter_affected                    182 non-null    object 
 18  WhatsApp_affected                   182 non-null    object 
 19  Instagram_affected                  182 non-null    object 
 20  Telegram_affected                   182 non-null    object 
 21  other_service_details (specify)     23 non-null     object 
 22  SMS_and_phone_call_affected         160 non-null    object 
 23  telcos_involved                     62 non-null     object 
 24  gov_ack                             206 non-null    object 
 25  official_just                       213 non-null    object 
 26  other_just_details                  79 non-null     object 
 27  off_statement                       13 non-null     object 
 28  actual_cause                        213 non-null    object 
 29  other_cause_details                 90 non-null     object 
 30  election                            208 non-null    object 
 31  violence                            211 non-null    object 
 32  hr_abuse_reported                   209 non-null    object 
 33  users_notified                      206 non-null    object 
 34  users_affected/targetted            160 non-null    object 
 35  legal_justif                        204 non-null    object 
 36  legal_method                        31 non-null     object 
 37  telco_resp                          200 non-null    object 
 38  telco_ ack                          14 non-null     object 
 39  econ_impact                         0 non-null      float64
 40  event                               144 non-null    object 
 41  an_link                             3 non-null      object 
 42  notes                               6 non-null      object 
dtypes: float64(2), int64(1), object(40)
memory usage: 71.7+ KB
```

In [3]:
```
df.describe()
```

Out[3]:

	ID	sub-region	econ_impact
count	213.000000	0.0	0.0
mean	107.000000	NaN	NaN
std	61.631972	NaN	NaN
min	1.000000	NaN	NaN
25%	54.000000	NaN	NaN
50%	107.000000	NaN	NaN
75%	160.000000	NaN	NaN
max	213.000000	NaN	NaN

In [4]:
```
df.columns
```

Out[4]:
```
Index(['ID', 'start_date', 'end_date ', 'duration', 'Info_source ',
       'news_link ', 'continent', 'sub-region', 'country ', 'geo_scope',
       'area_name ', 'ordered_by ', 'decision_maker', 'shutdown_type_new',
       'affected_network', 'full or service-based', 'Facebook_affected',
       'Twitter_affected', 'WhatsApp_affected', 'Instagram_affected',
       'Telegram_affected', 'other_service_details (specify) ',
       'SMS_and_phone_call_affected', 'telcos_involved ', 'gov_ack',
       'official_just', 'other_just_details', 'off_statement', 'actual_caus
e',
       'other_cause_details', 'election', 'violence', 'hr_abuse_reported',
       'users_notified', 'users_affected/targetted ', 'legal_justif',
       'legal_method ', 'telco_resp', 'telco_ ack', 'econ_impact', 'event ',
       'an_link ', 'notes'],
      dtype='object')
```

In [5]:
```
df['country '].value_counts()
```

Out[5]:
```
India                                121
Venezuela, Bolivarian Republic of     12
Yemen                                 11
Iraq                                   8
Algeria                                6
Pakistan                               5
Ethiopia                               4
SriLanka                               3
Sudan                                  3
Indonesia                              3
Bangladesh                             3
Iran                                   3
Russia                                 3
Kazakhstan                             3
Ecuador                                2
Syria                                  2
Myanmar                                2
Tajikistan                             2
Turkey                                 2
Benin                                  2
Chad                                   1
Liberia                                1
Gabon                                  1
China                                  1
Cameroon                               1
Malawi                                 1
Eritrea                                1
Egypt                                  1
DemocraticRepublicCongo                1
Mauritania                             1
United Kingdom                         1
Jordan                                 1
Zimbabwe                               1
Name: country , dtype: int64
```

In [6]:
```
from matplotlib import pyplot as plt
#make the plots show up inline
%matplotlib inline
import seaborn as sns
# set style
plt.style.use('_classic_test')
```

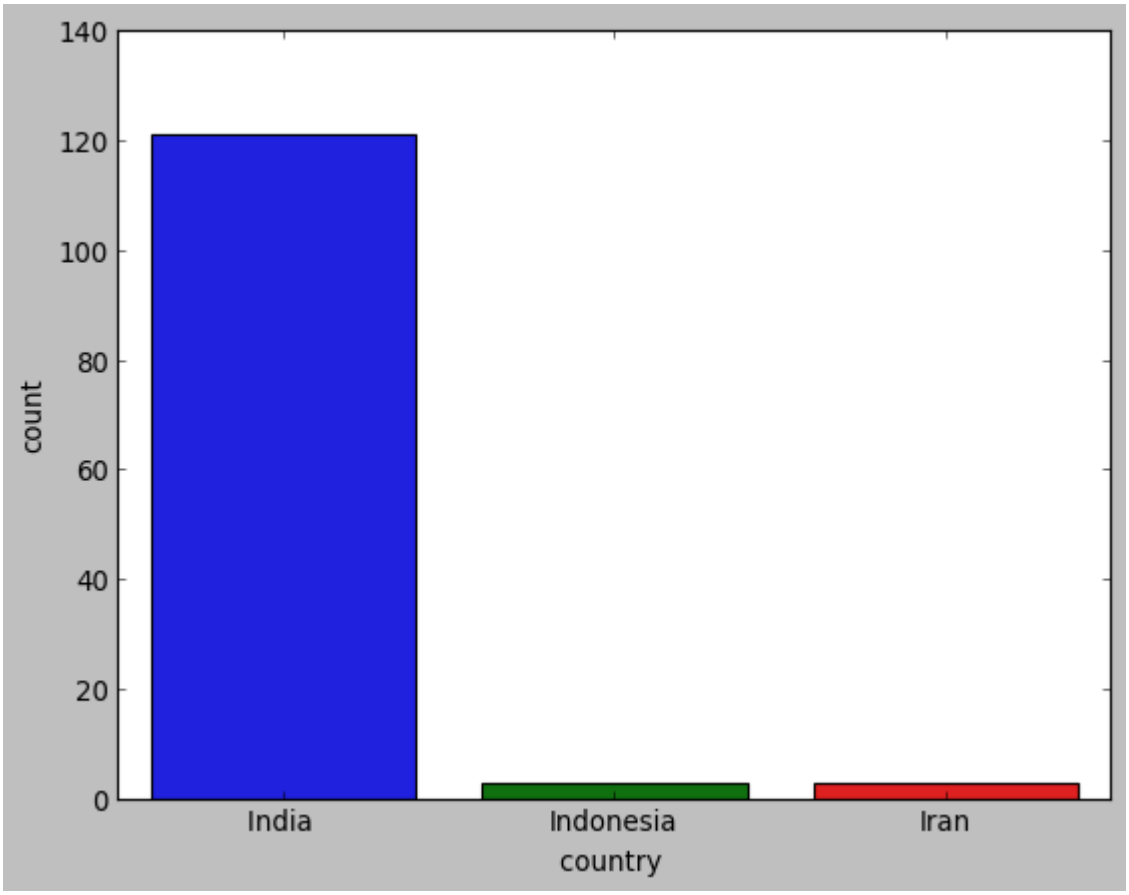
In [7]:
```
df['country '] = df['country '].astype(str)
df['country '].unique()
```

Out[7]:
```
array(['Chad', 'Yemen ', 'Jordan', 'Sudan', 'DemocraticRepublicCongo',
       'India ', 'Gabon', 'Zimbabwe ',
       'Venezuela, Bolivarian Republic of', 'Algeria', 'Russia',
       'Pakistan', 'SriLanka', 'Tajikistan', 'Benin ', 'Kazakhstan',
       'Iraq ', 'United Kingdom', 'China', 'Indonesia', 'Malawi ',
       'Liberia', 'Ethiopia ', 'Syria ', 'Eritrea', 'Cameroon', 'Myanmar',
       'Mauritania', 'Iran', 'Bangladesh', 'Egypt ', 'Ecuador', 'Turkey'],
      dtype=object)
```

In [8]:
```
df = df[df['country '].isin(['India ', 'Indonesia','Iran'])]
df['country '].unique()
```

Out[8]:
```
array(['India ', 'Indonesia', 'Iran'], dtype=object)
```

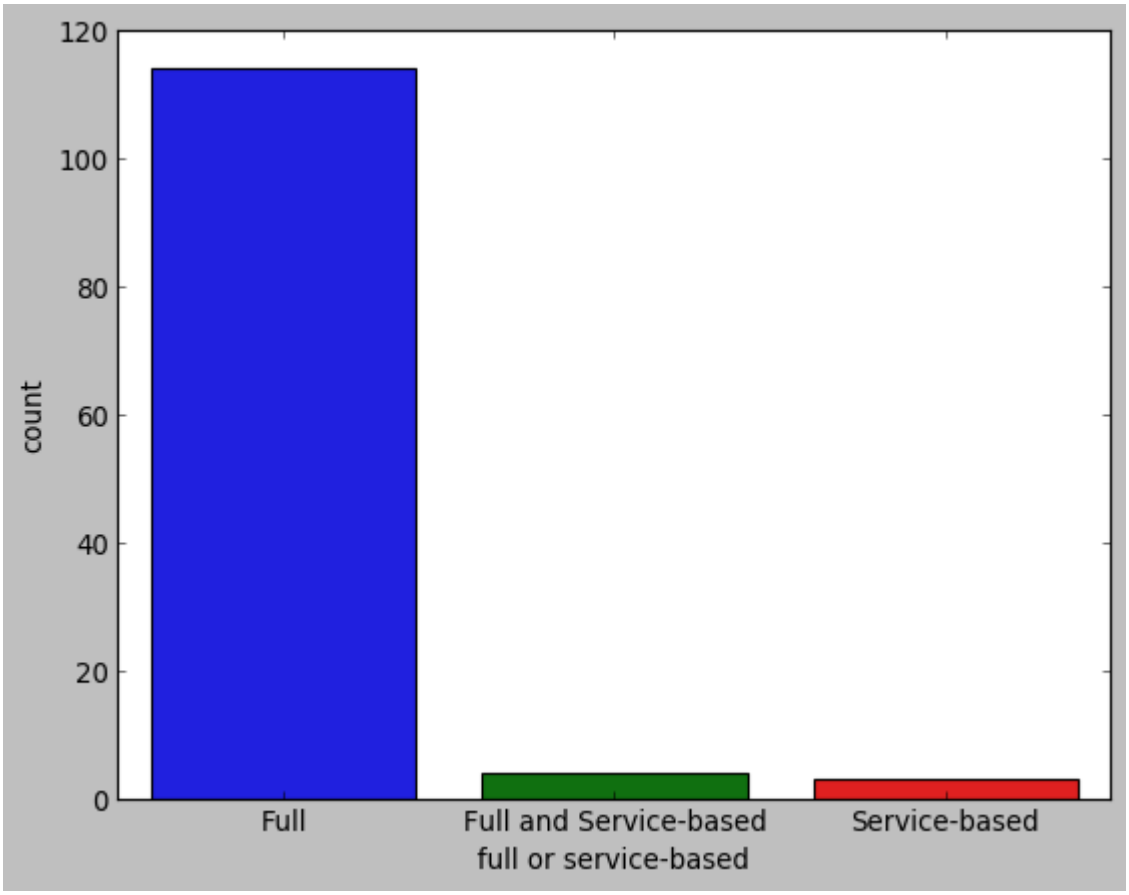
In [9]:
```
# Comparison of number of internet shutdowns in 2019 in countries of interest, depicts India with the most shutdowns
sns.countplot(x = 'country ',data = df)
plt.show()
```



In [10]:
```
df = df[df['country '].isin(['India '])]
df['country '].unique()
```

Out[10]:
```
array(['India '], dtype=object)
```

In [11]:
```
# Comparison of full internet shutdowns, full and/or service-based shutdowns,
 and service-based shutdowns
# Full internet shutdowns are more common
sns.countplot(x = 'full or service-based', data = df)
plt.show()
```



In [12]:
```
df.replace(False, 0, inplace=True)
```

In [13]:
```
df['Facebook_affected'].value_counts()
```

Out[13]:
```
0       95
True     7
Name: Facebook_affected, dtype: int64
```

In [14]:
```
df['Twitter_affected'].value_counts()
```

Out[14]:
```
0       98
True     3
Name: Twitter_affected, dtype: int64
```

In [15]:
```
df['WhatsApp_affected'].value_counts()
```

Out[15]:
```
0       97
True     4
Name: WhatsApp_affected, dtype: int64
```

In [16]:
```
df['Instagram_affected'].value_counts()
```

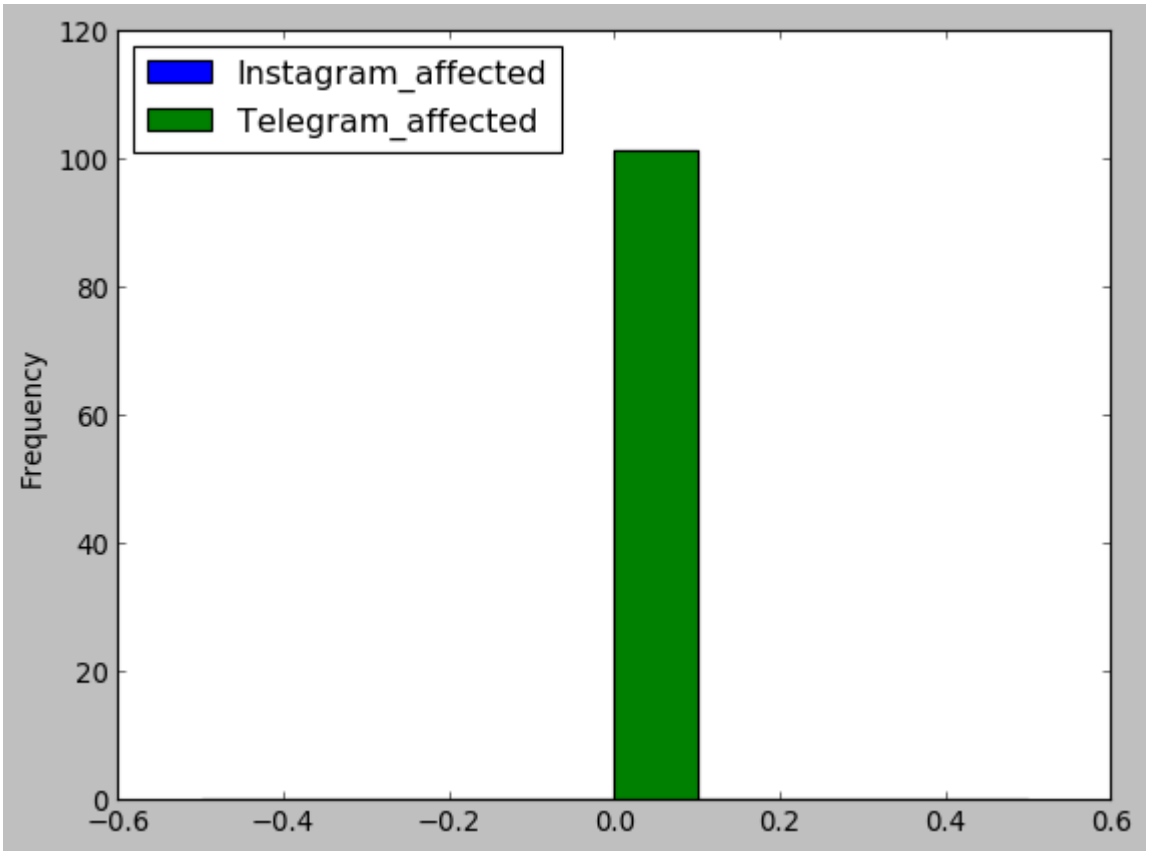
Out[16]:
```
0.0    101
Name: Instagram_affected, dtype: int64
```

In [17]:
```
df['Telegram_affected'].value_counts()
```

Out[17]:
```
0.0    101
Name: Telegram_affected, dtype: int64
```

In [19]:
```
df[['Facebook_affected', 'Twitter_affected', 'WhatsApp_affected', 'Instagram_affected', 'Telegram_affected']].plot(kind= 'hist')
```

Out[19]:
```
<matplotlib.axes._subplots.AxesSubplot at 0x220f0c76348>
```



In [23]:
```
# Visualizing when spikes in shutdown occured
import datetime
df['end_date '] = pd.to_datetime(df['end_date '], format= "%m/%d/%Y")
sns.distplot(df['end_date '])
plt.show()
```