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

In [2]:
```
# 1.a Load the dataset “prisoners.csv” using pandas and display the first and
 last five rows in the dataset- DATA LOADING
df = pd.read_csv("prisoners.csv")
df.head()
```

Out[2]:

	STATE/UT	YEAR	No. of Inmates benefitted by Elementary Education	No. of Inmates benefitted by Adult Education	No. of Inmates benefitted by Higher Education	No. of Inmates benefitted by Computer Course
0	Andhra Pradesh	2013	9480	13758	672	170
1	Arunachal Pradesh	2013	0	0	0	0
2	Assam	2013	676	750	14	30
3	Bihar	2013	1693	3013	125	1417
4	Chhatisgarh	2013	1664	1803	192	103

In [3]:
```
# 1.b Use describe method in pandas and find out the number of columns. Can yo
u say something about those rows who have zero inmates?
df.describe()
```

Out[3]:

	YEAR	No. of Inmates benefitted by Elementary Education	No. of Inmates benefitted by Adult Education	No. of Inmates benefitted by Higher Education	No. of Inmates benefitted by Computer Course
count	35.0	35.000000	35.000000	35.000000	35.000000
mean	2013.0	1057.914286	1534.857143	237.457143	210.171429
std	0.0	2078.196777	3022.110503	375.614191	359.117340
min	2013.0	0.000000	0.000000	0.000000	0.000000
25%	2013.0	0.000000	6.500000	2.000000	0.000000
50%	2013.0	167.000000	237.000000	33.000000	38.000000
75%	2013.0	1294.500000	1733.500000	234.500000	227.500000
max	2013.0	9480.000000	13758.000000	1353.000000	1417.000000

In [4]:
```
# 2.Data Manipulation
# a.Create a new column -'total_benefitted' that is a sum of inmates benefitte
d through all modes.
df["total_benefitted"] = df.sum(axis=1)
df["total_benefitted"].head()
```

Out[4]:
```
0    26093
1     2013
2     3483
3     8261
4     5775
Name: total_benefitted, dtype: int64
```

In [8]:
```
# b.Create a new row -"totals" that is the sum of all inmates benefitted throu
gh each mode across all states.
xlabels = prisoners['STATE/UT'].values
plot.figure(figsize=(20, 3))
plot.xticks(np.arange(xlabels.shape[0]), xlabels, rotation = 'vertical', fonts
ize = 18)
plot.xticks
plot.bar(np.arange(prisoners.values.shape[0]),prisoners['total_benefited'],ali
gn = 'edge')
```

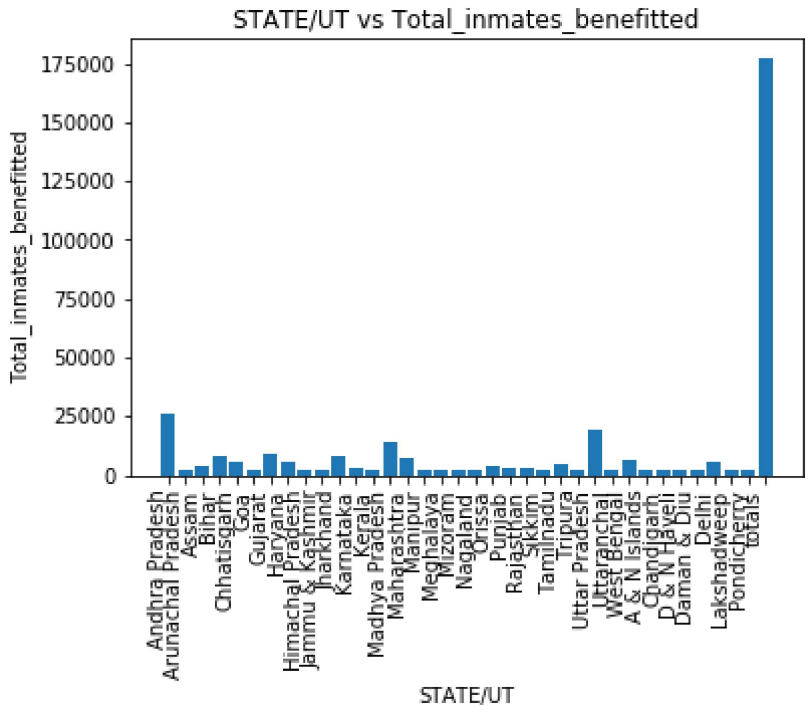
```
---------------------------------------------------------------------------
NameError                                 Traceback (most recent call last)
<ipython-input-8-c49935e4f354> in <module>
      1 # b.Create a new row -"totals" that is the sum of all inmates benefit
ted through each mode across all states.
----> 2 xlabels = prisoners['STATE/UT'].values
      3 plot.figure(figsize=(20, 3))
      4 plot.xticks(np.arange(xlabels.shape[0]), xlabels, rotation = 'vertica
l', fontsize = 18)
      5 plot.xticks

NameError: name 'prisoners' is not defined
```

In [6]:
```
# 3. Plotting
# a.Make a bar plot with each state name on the x -axis and their total benefi
tted inmates astheir bar heights.
# Which state has the maximum number of beneficiaries?
x = df["STATE/UT"]
y = df["total_benefitted"]

plt.bar(x,y)
plt.setp(plt.gca().get_xticklabels(), rotation=90, horizontalalignment='right'
)
plt.title("STATE/UT vs Total_inmates_benefitted")
plt.xlabel("STATE/UT")
plt.ylabel("Total_inmates_benefitted")
# Andgra Pradesh has the highest number of beneficiaries
```

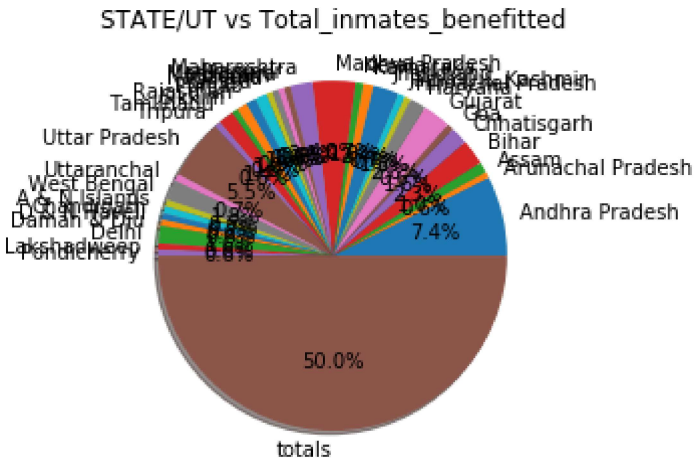
Out[6]: Text(0, 0.5, 'Total_inmates_benefitted')



In [9]:
```
# b.Make a pie chart that depicts the ratio among different modes of benefits.
labels = df["STATE/UT"]
values = df["total_benefitted"]

plt.pie(values, labels = labels, autopct='%1.1f%%', shadow=True)
plt.title("STATE/UT vs Total_inmates_benefitted")
```

Out[9]:
```
Text(0.5, 1.0, 'STATE/UT vs Total_inmates_benefitted')
```



In [ ]: