We chose to examine a data set from the U.S. Census Bureau website, titled *Means of Transportation to Work by Vehicles Available*. We felt that this data set was relevant to the part of our model that will attempt to maximize accessibility to the residents of Spartanburg. This is because we feel that individuals without vehicle access (or access to few vehicles) are the residents in Spartanburg that are most likely to need public transportation, so we hope that this data set will give us insight into the targeted audience of our model. As we can see, this data set splits the population of Spartanburg County into categories based on how individuals travel to work, and shows how many personal vehicles the people in each category have access to:

```
import matplotlib.pyplot as plt
import numpy as np
import pandas as pd

meansOfTransport = pd.read_csv('TRANSPORT/meansoftransportation.csv')
meansOfTransport.columns = ['Category', 'Estimate', 'Margin of Error']

meansOfTransport.head(15)
```

	Category	Estimate	Margin of Error
0	Total:	159,556	±4,516
1	No vehicle available	2,655	±1,277
2	1 vehicle available	30,583	±3,533
3	2 vehicles available	67,515	±5,250
4	3 or more vehicles available	58,803	±4,433
5	Car, truck, or van - drove alone:	128,123	±5,016
6	No vehicle available	1,046	±717
7	1 vehicle available	22,755	±2,833
8	2 vehicles available	55,945	±4,913
9	3 or more vehicles available	48,377	±4,297
10	Car, truck, or van - carpooled:	14,660	±2,580
11	No vehicle available	517	±420
12	1 vehicle available	4,104	±1,495
13	2 vehicles available	4,527	±1,055
14	3 or more vehicles available	5,512	±1,884

Next steps: Generate code with `meansOfTransport` View recommended plots

As you can see in the above portion of our dataset, the overarching categories focus on the means of transportation to work - whether that be through a personal car, public transportation, or walking. Then, in each category, the data shows how many cars that each individual had available to them. Now, onto a few summary statistics:

```
'''

In this first block of code, we want to determine the average amount of cars that each individual
travelling to work has available to them. To be able to do this, we will consider each person
with three or more cars available to them to have exactly three cars available to them.

'''

numWithZeroCars = int(meansOfTransport.iloc[1,1].replace(",",""))
numWithOneCar = int(meansOfTransport.iloc[2,1].replace(",",""))
numWithTwoCars = int(meansOfTransport.iloc[3,1].replace(",",""))
numWithThreeCars = int(meansOfTransport.iloc[4,1].replace(",",""))

meanCarsAvailable = (numWithOneCar + (2 * numWithTwoCars) + (3 * numWithThreeCars)) / (numWithZeroCars + numWithOneCar + numWithTwoCars + n
print(f'On average, each working individual in Spartanburg County has {str(round(meanCarsAvailable, 3))} cars available to them.')


totalWorkersEstimate = int(meansOfTransport.iloc[0,1].replace(",",""))
targetCategories = ['    Car, truck, or van - carpooled:', '    Public transportation (excluding taxicab):', '    Walked:', '    Taxicab, m
nonDrivingWorkersData = meansOfTransport[meansOfTransport['Category'].isin(targetCategories)]
nonDrivingWorkersSum = 0
for row in nonDrivingWorkersData.iterrows():
  temp = row[1][1].replace(",","")
  nonDrivingWorkersSum += int(temp)
nonDrivingWorkersPercent = round((nonDrivingWorkersSum / totalWorkersEstimate) * 100, 3)

print(f'There are an estimated {nonDrivingWorkersSum} working individuals in Spartanburg county that travel for work but choose not to driv

oneOrLessCarsPercent = round(((numWithZeroCars + numWithOneCar) / totalWorkersEstimate) * 100, 3)
print(f'There are an estimated {numWithZeroCars + numWithOneCar} working individuals in Spartanburg County with access to one or less perso
print(f'There are an estimated {numWithZeroCars} working individuals in Spartanburg County with access to zero personal vehicles. These cou
```

```
On average, each working individual in Spartanburg County has 2.144 cars available to them.
There are an estimated 16610 working individuals in Spartanburg county that travel for work but choose not to drive on their own in a p
There are an estimated 33238 working individuals in Spartanburg County with access to one or less personal vehicles. This is estimated
There are an estimated 2655 working individuals in Spartanburg County with access to zero personal vehicles. These could be considered
```

As these summary statistics show, working individuals in Spartanburg have approximately 2 cars available for them to drive; however, this statistic does not consider how many other people in their households that they may be sharing those cars with. Furthermore, the population of Spartanburg County that seems to be at a disadvantage (whether or not due to having no vehicle or not having the means to use one to get to work) is relatively small proportionately, but still consists of several thousands of people. Based on research we've done, have decided to consider working individuals with one or less cars available to them as individuals who are likely to benefit from improved transportation. So, based on these numbers, we can say that approximately 20.8% of workers in Spartanburg could benefit from an improved bus system, and about 2650 workers could be considered as individuals with critical need for improved public transportation, since that's how many are estimated to have 0 cars available to them. Now, onto some data visualization:

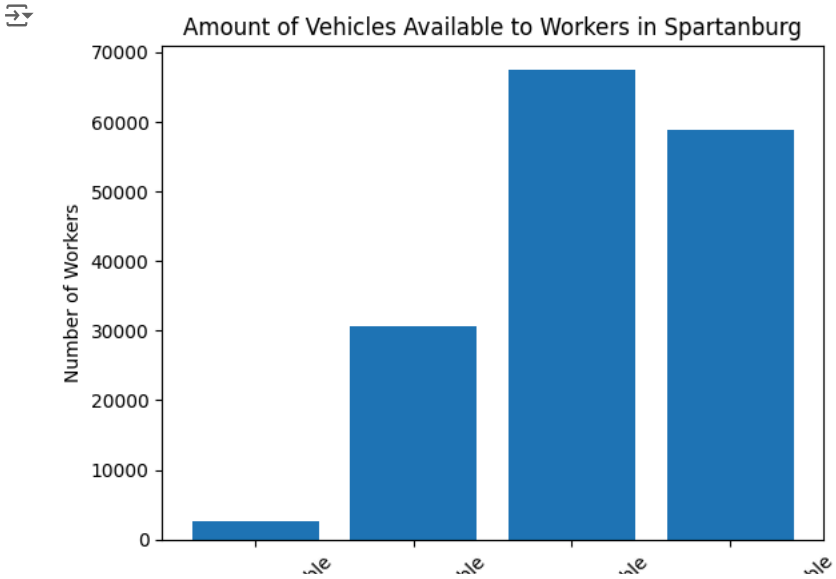
Double-click (or enter) to edit

```
totalVehicleEstimates = meansOfTransport.iloc[1:5]
categories = totalVehicleEstimates['Category']
estimates = [int(row.replace(",","")) for row in totalVehicleEstimates['Estimate']]
plt.bar(categories, estimates)
plt.title("Amount of Vehicles Available to Workers in Spartanburg")
plt.xlabel("Amount of Vehicles Available")
plt.ylabel("Number of Workers")
plt.xticks(rotation=45)
plt.show()

totalDrivingAlone = int(meansOfTransport.iloc[5,1].replace(",",""))
totalWorkingFromHome = int(meansOfTransport.iloc[30,1].replace(",",""))
barX = ['Drives Personal Vehicle Alone', 'Uses Other Transportation Means', 'Works From Home']
barY = [totalDrivingAlone, nonDrivingWorkersSum, totalWorkingFromHome]
plt.bar(barX, barY)
plt.xticks(rotation=45)
plt.title("Comparison of Transportation Means to Work in Spartanburg County")
plt.show()

walkingEstimates = meansOfTransport.iloc[21:24]
categories = walkingEstimates['Category']
estimates = [int(row.replace(",","")) for row in walkingEstimates['Estimate']]
plt.bar(categories, estimates)
```

```
plt.title("Amount of Vehicles Available to Workers that Walk in Spartanburg")
plt.xlabel("Amount of Vehicles Available")
plt.ylabel("Number of Workers")
plt.xticks(rotation=45)
plt.show()
```



The first bar graph, Amount of Vehicles Available to workers in Spartanburg, shows that most people have at least one car to available to them. The second bar graph, Comparison of Transportation Means to Work in Spartanburg County, shows that most people (about 80.3%), drive a personal vehicle alone to work while about 10.4% use other transportation means. Lastly, in the third bar graph, Amount of Vehicles Avaialble to Workers that Walk in Spartanburg, we can see that the vast majority of workers that walk to work are individuals that don't have a personal vehicle available to them. This isn't surprising and we believe that this is indicative of an increased need for efficient buses in the city, especially in areas where these workers live. From these data visualizations, we can determine that a relatively small section of the population relies on transportation that is not a car. However, this data is from Spartanburg County as a whole, so it would be beneficial to find data specific to the city of Spartanburg. Additionally, understanding that there are only 10.4% of the population that rely on other forms of transportation can help us narrow our search in finding communities in the city of Spartanburg who rely on the bus system.

Overall, we feel that this data is important because it cues us in on approximately how many people in the Spartanburg area could benefit from and improved public transportation infrastructure. Not only does it show us how many people already choose not to use a personal vehicle and could therefore benefit from cheap public transport, but it also shows us how many people outright don't have access to a personal vehicle, or may have to share one with a whole family. We feel this kind of data could be used in our problem to understand how much to modify/expand the bus system within reason, and also justify why this kind of problem is important.