FinalTutorial

December 18, 2021

1 Introduction: (talk about what we will be doing)

2 About our Data:

The dataset that we obtained contains data on the living wages in 100 different U.S. cities. The dataset provides information such as the population of the city in years 2010 and 2020, the area of land (in sq miles), and the density of the city. Most importantly, the dataset contains data on the living wages for each city, based on the number of individuals (adults and kids) in the household and whether the adults are working. Households may consist of: 1) one adult working and no kids 2) one adult working and one kid 3) one adult working and two kids 4) one adult working and three kids 5) two adults (only one adult working) and no kids 6) two adults (only one adult working) and one kid 7) two adults (only one adult working) and two kids 8) two adults (only one adult working) and three kids 9) two adults (both adults working) and no kids 10) two adults (both adults working) and one kid 11) two adults (both adults working) and two kids 12) two adults (both adults working) and three kids

3 Table of Contents:

[3]:
```
import pandas as pd
```

4 Getting our Data:

First, we need to read in our data from the CSV file and store it into a dataframe.

[6]:
```
url = 'https://www.kaggle.com/brandonconrady/living-wage-top-100-cities'
df = pd.read_csv('livingwage.csv')
df.head()
```

[6]:

	rank_2020	city	state	population_2020	population_2010	land_area_sqmi	density	one_adult_no_kids_living_wage
0	1	New York	New York	8804190	8175133	300.5	29298	20.00
1	2	Los Angeles	California	3898747	3792621			
2	3	Chicago	Illinois	2746388	2695598			
3	4	Houston	Texas	2304580	2099451			
4	5	Phoenix	Arizona	1608139	1445632			

```
1          469.5     8304                          19.22
2          227.7    12061                          16.08
3          640.4     3599                          14.29
4          518.0     3105                          15.41

   one_adult_one_kid_living_wage  one_adult_two_kids_living_wage  \
0                          38.99                           49.18
1                          41.55                           51.30
2                          32.25                           40.20
3                          29.09                           35.73
4                          29.44                           35.40

   one_adult_three_kids_living_wage  \
0                             64.30
1                             67.56
2                             51.69
3                             46.02
4                             46.01

   two_adults_one_working_no_kids_living_wage  \
0                                       29.16
1                                       31.01
2                                       25.41
3                                       22.88
4                                       24.85

   two_adults_one_working_one_kid_living_wage  \
0                                       34.35
1                                       37.96
2                                       29.73
3                                       27.54
4                                       29.25

   two_adults_one_working_two_kids_living_wage  \
0                                        38.13
1                                        41.94
2                                        33.38
3                                        30.93
4                                        32.98

   two_adults_one_working_three_kids_living_wage  \
0                                          42.42
1                                          47.75
2                                          35.93
3                                          33.69
4                                          36.95
```

```
   two_adults_both_working_no_kids_living_wage  \
0                                        14.03
1                                        14.62
2                                        12.28
3                                        11.14
4                                        12.03

   two_adults_both_working_one_kid_living_wage  \
0                                        20.95
1                                        22.36
2                                        17.52
3                                        15.85
4                                        16.21

   two_adults_both_working_two_kids_living_wage  \
0                                         26.53
1                                         27.73
2                                         22.01
3                                         19.63
4                                         19.65

   two_adults_both_working_three_kids_living_wage
0                                           32.40
1                                           34.01
2                                           25.94
3                                           23.09
4                                           23.23
```

From the CSV file, we noticed that we have data for multiple cities in the same state. For example, for the state of California, we have data for cities like Bakersfield, Anaheim, Stockton, Riverside, etc. Therefore, first we will organize our data for different cities by state.

[9]:
```
#array that stores the name of each state in our dataframe (no repeats)
array_of_states = []

#iterates through the dataframe and generates a list of all of the states with␣
 ↪no repeats
for index, row in df.iterrows():
    if row['state'] not in array_of_states:
        array_of_states.append(row["state"])

for state in array_of_states:
```

[9]:
```
['New York',
 'California',
 'Illinois',
 'Texas',
```

```
'Arizona',
'Pennsylvania',
'Florida',
'Ohio',
'Indiana',
'North Carolina',
'Washington',
'Colorado',
'District of Columbia',
'Tennessee',
'Oklahoma',
'Massachusetts',
'Oregon',
'Nevada',
'Michigan',
'Kentucky',
'Maryland',
'Wisconsin',
'New Mexico',
'Missouri',
'Georgia',
'Nebraska',
'Virginia',
'Minnesota',
'Kansas',
'Louisiana',
'Hawaii',
'New Jersey',
'Alaska',
'Idaho']
```

According to Statista, the average family size in the US in 2020 was approximately 3.15 persons. We can estimate that this is around two adults and one child per family.

Source: https://www.statista.com/statistics/183657/average-size-of-a-family-in-the-us/

Next, we will like to

[ ]: