

```
In [1]: import pandas as pd
```

```
In [2]: # import, read, and clean health data
health_xlsx = "County Health Rankings-Final.xlsx"
health_df = pd.read_excel(health_xlsx)
health_df = health_df.rename(columns={"overall county rank": "health rank"})
health_df = health_df.drop(["state", "county"], axis=1)
health_df
```

Out[2]:

	fips	health rank
0	1001	2
1	1003	3
2	1005	0
3	1007	1
4	1009	2
...	...	...
3103	56037	4
3104	56039	6
3105	56041	4
3106	56043	5
3107	56045	6

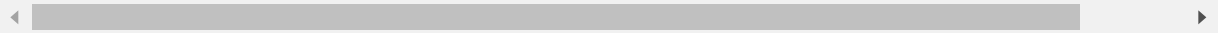
3108 rows × 2 columns

```
In [3]: # import, read, and clean population density data
density_csv = "pop_density_ranked.csv"
density_df = pd.read_csv(density_csv)
density_df = density_df.rename(columns={"county_fips": "fips"})
density_df = density_df.drop(["county_name", "state_id", "county_name"], axis=
1)
density_df
```

Out[3]:

	fips	population	pop_area_est	density	pop_ranked	pop_scaled	density_ranked	densi
<b>0</b>	1001	57046	1512.03	37.73	2216	8	2191	
<b>1</b>	1003	207694	4077.05	50.94	2783	9	2367	
<b>2</b>	1005	25278	1821.81	13.88	1511	5	1287	
<b>3</b>	1007	22859	1415.86	16.14	1432	5	1437	
<b>4</b>	1009	45030	1311.39	34.34	2051	7	2122	
...	...	...	...	...	...	...	...	...
<b>3095</b>	56037	44117	6322.20	6.98	2021	7	759	
<b>3096</b>	56039	22196	6268.00	3.54	1401	5	469	
<b>3097</b>	56041	20630	3833.92	5.38	1328	5	621	
<b>3098</b>	56043	8295	3262.25	2.54	581	2	373	
<b>3099</b>	56045	7263	6741.91	1.08	503	2	160	

3100 rows × 8 columns



```
In [4]: # import, read, and clean gender data
gender_csv = "Gender_Final_ranked.csv"
gender_df = pd.read_csv(gender_csv)
gender_df = gender_df.drop(["state", "county"], axis=1)
gender_df
```

Out[4]:

	TOT_MALE	TOT_FEMALE	% TOT_MALE	% TOT_FEMALE	fips	maleness_ranked	maleness_:
0	26995	28606	48.551285	51.448715	1001	475	
1	105657	112365	48.461623	51.538377	1003	430	
2	13133	11748	52.783248	47.216752	1005	2874	
3	11921	10479	53.218750	46.781250	1007	2908	
4	28500	29340	49.273859	50.726141	1009	1070	
...	...	...	...	...	...	...	...
3104	22152	20899	51.455251	48.544749	56037	2671	
3105	11909	11172	51.596551	48.403449	56039	2710	
3106	10285	10014	50.667521	49.332479	56041	2391	
3107	3988	3897	50.577045	49.422955	56043	2341	
3108	3689	3278	52.949620	47.050380	56045	2890	

3109 rows × 7 columns



```
In [5]: # import, read, and clean age data 1 of 2
age_csv = "Age_Final_ranked.csv"
age_df = pd.read_csv(age_csv)
age_df = age_df.drop(["state", "county"], axis=1)
age_df = age_df.rename(columns={"over 70": "% over 70", "Rank": "Rank over 70",
, "Scale": "scale over 70"})
age_df
```

Out[5]:

	fips	% over 70	Rank over 70	scale over 70
0	1001	10.906279	763	3
1	1003	13.899515	1967	7
2	1005	13.230979	1677	6
3	1007	11.196429	851	3
4	1009	12.562241	1428	5
...	...	...	...	...
3104	56037	7.309935	80	1
3105	56039	9.644296	426	2
3106	56041	8.571851	230	1
3107	56043	15.129994	2359	8
3108	56045	13.721831	1879	7

3109 rows × 4 columns

```
In [6]: # import, read, and clean age data 2 of 2
allages_csv = "Age_Final.csv"
allages_df = pd.read_csv(allages_csv)
allages_df = allages_df.drop(["state", "county"], axis=1)
allages_df
```

Out[6]:

	% Age 0 to 29 years	% Age 30 to 39	% Age 40 to 49 years	% Age 50 to 59 years	% Age 60 to 69 years	% Age 70 to 79	% Age 80 to 85+	fips
<b>0</b>	38.312261	12.749771	13.636445	13.917016	10.478229	7.284042	3.622237	1001
<b>1</b>	34.514407	11.564429	12.539102	13.753658	13.728890	9.602701	4.296814	1003
<b>2</b>	36.011414	12.865239	12.354809	13.174712	12.362847	9.018930	4.212049	1005
<b>3</b>	36.075893	13.897321	13.562500	13.906250	11.361607	7.691964	3.504464	1007
<b>4</b>	37.010719	11.901798	12.921853	13.561549	12.041840	8.525242	4.036999	1009
...	...	...	...	...	...	...	...	...
<b>3104</b>	41.320759	15.277229	12.389956	12.220390	11.481731	5.105572	2.204362	56037
<b>3105</b>	32.650232	17.733200	14.830380	12.685759	12.456133	6.797799	2.846497	56039
<b>3106</b>	41.770531	13.311001	11.773979	11.990738	12.581901	6.074191	2.497660	56041
<b>3107</b>	34.064680	11.452124	11.414077	13.760304	14.178821	9.397590	5.732403	56043
<b>3108</b>	32.151572	13.004162	10.535381	14.152433	16.434620	8.037893	5.683939	56045

3109 rows × 8 columns

```
In [7]: # import, read, and clean poverty data
poverty_csv = "Poverty_ranked.csv"
poverty_df = pd.read_csv(poverty_csv)
poverty_df = poverty_df.rename(columns={"FIPStxt": "fips"})
poverty_df = poverty_df.drop(["Stabr", "Area_name"], axis=1)
poverty_df
```

Out[7]:

	fips	Poverty, total	Poverty, %	Poverty_ranked	Poverty_scaled
<b>0</b>	1001	7,587	13.8	1462	5
<b>1</b>	1003	21,069	9.8	533	2
<b>2</b>	1005	6,788	30.9	3039	10
<b>3</b>	1007	4,400	21.8	2696	9
<b>4</b>	1009	7,527	13.2	1328	5
...	...	...	...	...	...
<b>3103</b>	56037	3,540	8.4	275	1
<b>3104</b>	56039	1,443	6.3	73	1
<b>3105</b>	56041	2,010	10.0	571	2
<b>3106</b>	56043	918	11.9	1002	4
<b>3107</b>	56045	777	11.8	982	4

3108 rows × 5 columns

```
In [8]: # import, read, and clean income data
income_csv = "Median_income_ranked.csv"
income_df = pd.read_csv(income_csv)
income_df = income_df.rename(columns={"FIPStxt": "fips"})
income_df = income_df.drop(["Stabr", "Area_name"], axis=1)
income_df
```

Out[8]:

	fips	Median_Income	Income_ranked	Income_scaled
<b>0</b>	1001	59,338	2380	8
<b>1</b>	1003	57,588	2262	8
<b>2</b>	1005	34,382	124	1
<b>3</b>	1007	46,064	1033	4
<b>4</b>	1009	50,412	1542	5
...	...	...	...	...
<b>3103</b>	56037	73,315	2879	10
<b>3104</b>	56039	99,087	3068	10
<b>3105</b>	56041	63,401	2614	9
<b>3106</b>	56043	55,190	2050	7
<b>3107</b>	56045	54,319	1979	7

3108 rows × 4 columns

```
In [9]: # import, read, and clean transportation data
transportation_csv = "Transportation_May_4.csv"
transportation_df = pd.read_csv(transportation_csv)
transportation_df = transportation_df.rename(columns={"Percent": "% using transportation", "county_fips": "fips"})
transportation_df = transportation_df.drop(["county_name", "state_id"], axis=1)
transportation_df
```

Out[9]:

	fips	% using transportation
0	1001.0	NaN
1	1003.0	NaN
2	1005.0	NaN
3	1007.0	NaN
4	1009.0	NaN
...	...	...
3101	NaN	NaN
3102	NaN	NaN
3103	NaN	NaN
3104	NaN	NaN
3105	NaN	NaN

3106 rows × 2 columns



```

In [10]: # import, read, and clean covid cases and deaths from April 11 data
covid411_csv = "April_11_data_ranked.csv"
covid411_df = pd.read_csv(covid411_csv)
covid411_df = covid411_df.drop(["County", "State"], axis=1)
covid411_df = covid411_df.dropna(how='any')
covid411_df = covid411_df.rename(columns={"FIPS": "fips", "Cases_rank": "Apr11
_cases_rank",
                                         "Cases_scale": "Apr11_cases_scale",
                                         "Deaths_rank": "Apr11_deaths_rank",
                                         "Deaths_scale": "Apr11_deaths_scale"
})
covid411_df

```

Out[10]:

	fips	Apr11_Cases	Apr11_Deaths	Apr11_cases_rank	Apr11_cases_scale	Apr11_deaths_r
0	1001	19	1	1577	7	1
1	1003	66	1	2077	8	1
2	1005	10	0	1204	5	
3	1007	13	0	1369	6	
4	1009	12	0	1319	6	
...	...	...	...	...	...	
2610	56035	1	0	1	1	
2611	56037	7	0	1000	4	
2612	56039	53	0	2013	8	
2613	56041	4	0	662	3	
2614	56043	4	0	662	3	

2615 rows × 7 columns

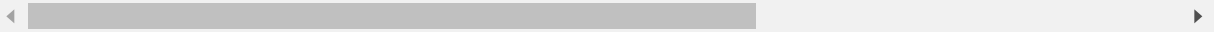


```
In [11]: # import, read, and clean covid cases and deaths from April 18 data
covid418_csv = "April_18_data_ranked.csv"
covid418_df = pd.read_csv(covid418_csv)
covid418_df = covid418_df.dropna(how='any')
covid418_df = covid418_df.rename(columns={"FIPS": "fips"})
covid418_df
```

Out[11]:

	County	State	fips	Apr18_Cases	Apr18_Deaths	Apr18_cases_ranked	Apr_18_ca
0	Autauga	Alabama	1001	25	2	1621	
1	Baldwin	Alabama	1003	109	2	2226	
2	Barbour	Alabama	1005	18	0	1431	
3	Bibb	Alabama	1007	26	0	1643	
4	Blount	Alabama	1009	20	0	1487	
...	...	...	...	...	...	...	
2708	Sublette	Wyoming	56035	1	0	1	
2709	Sweetwater	Wyoming	56037	10	0	1096	
2710	Teton	Wyoming	56039	62	0	2028	
2711	Uinta	Wyoming	56041	6	0	808	
2712	Washakie	Wyoming	56043	4	0	572	

2713 rows × 9 columns



```
In [12]: # import, read, and clean fips data
latlong_csv = "uszip.csv"
latlong_df = pd.read_csv(latlong_csv)
latlong_df
```

Out[12]:

	lat	lng	fips
0	32.45322	-86.72385	1001
1	30.86319	-87.73904	1003
2	31.85209	-85.42935	1005
3	32.91192	-87.26985	1007
4	33.90169	-86.51787	1009
...	...	...	...
3207	18.42132	-66.40012	72145
3208	18.12851	-65.44093	72147
3209	18.13299	-66.47703	72149
3210	18.07053	-65.89604	72151
3211	18.06728	-66.85273	72153

3212 rows × 4 columns

```
In [13]: # merge all 11 dataframes together 1 of 10
merge12_df = pd.merge(health_df, density_df, how='left', on = ['fips'])
merge12_df
```

Out[13]:

	fips	health rank	population	pop_area_est	density	pop_ranked	pop_scaled	density_ranked
0	1001	2	57046.0	1512.03	37.73	2216.0	8.0	2191.0
1	1003	3	207694.0	4077.05	50.94	2783.0	9.0	2367.0
2	1005	0	25278.0	1821.81	13.88	1511.0	5.0	1287.0
3	1007	1	22859.0	1415.86	16.14	1432.0	5.0	1437.0
4	1009	2	45030.0	1311.39	34.34	2051.0	7.0	2122.0
...	...	...	...	...	...	...	...	...
3103	56037	4	44117.0	6322.20	6.98	2021.0	7.0	759.0
3104	56039	6	22196.0	6268.00	3.54	1401.0	5.0	469.0
3105	56041	4	20630.0	3833.92	5.38	1328.0	5.0	621.0
3106	56043	5	8295.0	3262.25	2.54	581.0	2.0	373.0
3107	56045	6	7263.0	6741.91	1.08	503.0	2.0	160.0

3108 rows × 9 columns

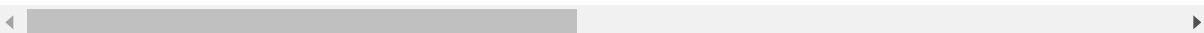


```
In [14]: # merge all 11 dataframes together 2 of 10
merge123_df = pd.merge(merge12_df, gender_df, how='left', on = ['fips'])
merge123_df
```

Out[14]:

	fips	health rank	population	pop_area_est	density	pop_ranked	pop_scaled	density_ranked
0	1001	2	57046.0	1512.03	37.73	2216.0	8.0	2191.0
1	1003	3	207694.0	4077.05	50.94	2783.0	9.0	2367.0
2	1005	0	25278.0	1821.81	13.88	1511.0	5.0	1287.0
3	1007	1	22859.0	1415.86	16.14	1432.0	5.0	1437.0
4	1009	2	45030.0	1311.39	34.34	2051.0	7.0	2122.0
...	...	...	...	...	...	...	...	...
3104	56037	4	44117.0	6322.20	6.98	2021.0	7.0	759.0
3105	56039	6	22196.0	6268.00	3.54	1401.0	5.0	469.0
3106	56041	4	20630.0	3833.92	5.38	1328.0	5.0	621.0
3107	56043	5	8295.0	3262.25	2.54	581.0	2.0	373.0
3108	56045	6	7263.0	6741.91	1.08	503.0	2.0	160.0

3109 rows × 15 columns

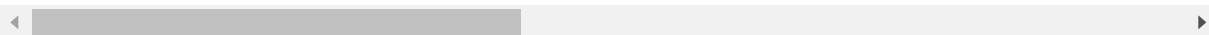


```
In [15]: # merge all 11 dataframes together 3 of 10
merge1234_df = pd.merge(merge123_df, age_df, how='left', on = ['fips'])
merge1234_df
```

Out[15]:

	fips	health rank	population	pop_area_est	density	pop_ranked	pop_scaled	density_ranked
0	1001	2	57046.0	1512.03	37.73	2216.0	8.0	2191.0
1	1003	3	207694.0	4077.05	50.94	2783.0	9.0	2367.0
2	1005	0	25278.0	1821.81	13.88	1511.0	5.0	1287.0
3	1007	1	22859.0	1415.86	16.14	1432.0	5.0	1437.0
4	1009	2	45030.0	1311.39	34.34	2051.0	7.0	2122.0
...	...	...	...	...	...	...	...	...
3106	56037	4	44117.0	6322.20	6.98	2021.0	7.0	759.0
3107	56039	6	22196.0	6268.00	3.54	1401.0	5.0	469.0
3108	56041	4	20630.0	3833.92	5.38	1328.0	5.0	621.0
3109	56043	5	8295.0	3262.25	2.54	581.0	2.0	373.0
3110	56045	6	7263.0	6741.91	1.08	503.0	2.0	160.0

3111 rows × 18 columns



```
In [16]: # merge all 11 dataframes together 4 of 10
merge12345_df = pd.merge(merge1234_df, allages_df, how='left', on = ['fips'])
merge12345_df
```

Out[16]:

	fips	health rank	population	pop_area_est	density	pop_ranked	pop_scaled	density_ranked
0	1001	2	57046.0	1512.03	37.73	2216.0	8.0	2191.0
1	1003	3	207694.0	4077.05	50.94	2783.0	9.0	2367.0
2	1005	0	25278.0	1821.81	13.88	1511.0	5.0	1287.0
3	1007	1	22859.0	1415.86	16.14	1432.0	5.0	1437.0
4	1009	2	45030.0	1311.39	34.34	2051.0	7.0	2122.0
...	...	...	...	...	...	...	...	...
3110	56037	4	44117.0	6322.20	6.98	2021.0	7.0	759.0
3111	56039	6	22196.0	6268.00	3.54	1401.0	5.0	469.0
3112	56041	4	20630.0	3833.92	5.38	1328.0	5.0	621.0
3113	56043	5	8295.0	3262.25	2.54	581.0	2.0	373.0
3114	56045	6	7263.0	6741.91	1.08	503.0	2.0	160.0

3115 rows × 25 columns

```
In [17]: # merge all 11 dataframes together 5 of 10
merge123456_df = pd.merge(merge12345_df, poverty_df, how='left', on = ['fips'
])
merge123456_df
```

Out[17]:

	fips	health rank	population	pop_area_est	density	pop_ranked	pop_scaled	density_ranked
0	1001	2	57046.0	1512.03	37.73	2216.0	8.0	2191.0
1	1003	3	207694.0	4077.05	50.94	2783.0	9.0	2367.0
2	1005	0	25278.0	1821.81	13.88	1511.0	5.0	1287.0
3	1007	1	22859.0	1415.86	16.14	1432.0	5.0	1437.0
4	1009	2	45030.0	1311.39	34.34	2051.0	7.0	2122.0
...	...	...	...	...	...	...	...	...
3110	56037	4	44117.0	6322.20	6.98	2021.0	7.0	759.0
3111	56039	6	22196.0	6268.00	3.54	1401.0	5.0	469.0
3112	56041	4	20630.0	3833.92	5.38	1328.0	5.0	621.0
3113	56043	5	8295.0	3262.25	2.54	581.0	2.0	373.0
3114	56045	6	7263.0	6741.91	1.08	503.0	2.0	160.0

3115 rows × 29 columns

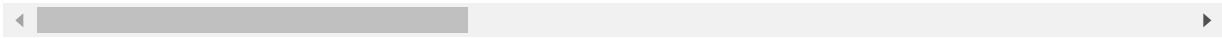


```
In [18]: # merge all 11 dataframes together 6 of 10
merge1234567_df = pd.merge(merge123456_df, income_df, how='left', on = ['fips'
])
merge1234567_df
```

Out[18]:

	fips	health rank	population	pop_area_est	density	pop_ranked	pop_scaled	density_ranked
0	1001	2	57046.0	1512.03	37.73	2216.0	8.0	2191.0
1	1003	3	207694.0	4077.05	50.94	2783.0	9.0	2367.0
2	1005	0	25278.0	1821.81	13.88	1511.0	5.0	1287.0
3	1007	1	22859.0	1415.86	16.14	1432.0	5.0	1437.0
4	1009	2	45030.0	1311.39	34.34	2051.0	7.0	2122.0
...	...	...	...	...	...	...	...	...
3110	56037	4	44117.0	6322.20	6.98	2021.0	7.0	759.0
3111	56039	6	22196.0	6268.00	3.54	1401.0	5.0	469.0
3112	56041	4	20630.0	3833.92	5.38	1328.0	5.0	621.0
3113	56043	5	8295.0	3262.25	2.54	581.0	2.0	373.0
3114	56045	6	7263.0	6741.91	1.08	503.0	2.0	160.0

3115 rows × 32 columns

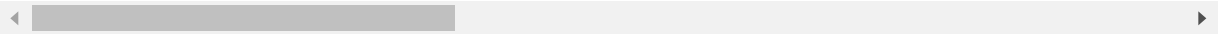


```
In [19]: # merge all 11 dataframes together 7 of 10
merge12345678_df = pd.merge(merge1234567_df, transportation_df, how='left', on
= ['fips'])
merge12345678_df
```

Out[19]:

	fips	health rank	population	pop_area_est	density	pop_ranked	pop_scaled	density_ranked
0	1001	2	57046.0	1512.03	37.73	2216.0	8.0	2191.0
1	1003	3	207694.0	4077.05	50.94	2783.0	9.0	2367.0
2	1005	0	25278.0	1821.81	13.88	1511.0	5.0	1287.0
3	1007	1	22859.0	1415.86	16.14	1432.0	5.0	1437.0
4	1009	2	45030.0	1311.39	34.34	2051.0	7.0	2122.0
...	...	...	...	...	...	...	...	...
3110	56037	4	44117.0	6322.20	6.98	2021.0	7.0	759.0
3111	56039	6	22196.0	6268.00	3.54	1401.0	5.0	469.0
3112	56041	4	20630.0	3833.92	5.38	1328.0	5.0	621.0
3113	56043	5	8295.0	3262.25	2.54	581.0	2.0	373.0
3114	56045	6	7263.0	6741.91	1.08	503.0	2.0	160.0

3115 rows × 33 columns



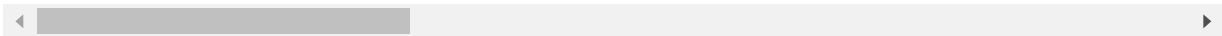


```
In [20]: # merge all 11 dataframes together 8 of 10
merge123456789_df = pd.merge(merge12345678_df, covid411_df, how='left', on = [
'fips'])
merge123456789_df
```

Out[20]:

	fips	health rank	population	pop_area_est	density	pop_ranked	pop_scaled	density_ranked
0	1001	2	57046.0	1512.03	37.73	2216.0	8.0	2191.0
1	1003	3	207694.0	4077.05	50.94	2783.0	9.0	2367.0
2	1005	0	25278.0	1821.81	13.88	1511.0	5.0	1287.0
3	1007	1	22859.0	1415.86	16.14	1432.0	5.0	1437.0
4	1009	2	45030.0	1311.39	34.34	2051.0	7.0	2122.0
...	...	...	...	...	...	...	...	...
3110	56037	4	44117.0	6322.20	6.98	2021.0	7.0	759.0
3111	56039	6	22196.0	6268.00	3.54	1401.0	5.0	469.0
3112	56041	4	20630.0	3833.92	5.38	1328.0	5.0	621.0
3113	56043	5	8295.0	3262.25	2.54	581.0	2.0	373.0
3114	56045	6	7263.0	6741.91	1.08	503.0	2.0	160.0

3115 rows × 39 columns

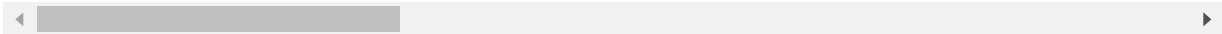


```
In [21]: # merge all 11 dataframes together 9 of 10
merge12345678910_df = pd.merge(merge123456789_df, covid418_df, how='left', on
= ['fips'])
merge12345678910_df
```

Out[21]:

	fips	health rank	population	pop_area_est	density	pop_ranked	pop_scaled	density_ranked
0	1001	2	57046.0	1512.03	37.73	2216.0	8.0	2191.0
1	1003	3	207694.0	4077.05	50.94	2783.0	9.0	2367.0
2	1005	0	25278.0	1821.81	13.88	1511.0	5.0	1287.0
3	1007	1	22859.0	1415.86	16.14	1432.0	5.0	1437.0
4	1009	2	45030.0	1311.39	34.34	2051.0	7.0	2122.0
...	...	...	...	...	...	...	...	...
3110	56037	4	44117.0	6322.20	6.98	2021.0	7.0	759.0
3111	56039	6	22196.0	6268.00	3.54	1401.0	5.0	469.0
3112	56041	4	20630.0	3833.92	5.38	1328.0	5.0	621.0
3113	56043	5	8295.0	3262.25	2.54	581.0	2.0	373.0
3114	56045	6	7263.0	6741.91	1.08	503.0	2.0	160.0

3115 rows × 47 columns

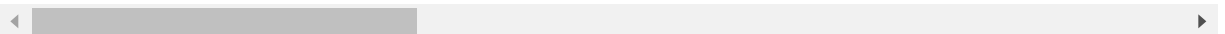


```
In [22]: # merge all 11 dataframes together 10 of 10
merge1234567891011_df = pd.merge(merge12345678910_df, latlong_df, how='left',
on = ['fips'])
merge1234567891011_df
```

Out[22]:

	fips	health rank	population	pop_area_est	density	pop_ranked	pop_scaled	density_ranked
<b>0</b>	1001	2	57046.0	1512.03	37.73	2216.0	8.0	2191.0
<b>1</b>	1003	3	207694.0	4077.05	50.94	2783.0	9.0	2367.0
<b>2</b>	1005	0	25278.0	1821.81	13.88	1511.0	5.0	1287.0
<b>3</b>	1007	1	22859.0	1415.86	16.14	1432.0	5.0	1437.0
<b>4</b>	1009	2	45030.0	1311.39	34.34	2051.0	7.0	2122.0
...	...	...	...	...	...	...	...	..
<b>3110</b>	56037	4	44117.0	6322.20	6.98	2021.0	7.0	759.0
<b>3111</b>	56039	6	22196.0	6268.00	3.54	1401.0	5.0	469.0
<b>3112</b>	56041	4	20630.0	3833.92	5.38	1328.0	5.0	621.0
<b>3113</b>	56043	5	8295.0	3262.25	2.54	581.0	2.0	373.0
<b>3114</b>	56045	6	7263.0	6741.91	1.08	503.0	2.0	160.0

3115 rows × 49 columns



```
In [23]: # write merged data to a csv file
totalmerged_df = merge1234567891011_df
totalmerged_df.to_csv("totalmerged.csv", index=False, header=True)
```