

# Assignment 3

February 2, 2023

```
[1]: import geopandas as gpd
```

```
[ ]: #This notebook documents median household income in Tulare County in 2021 by
      ↪census tract. Mapping levels of income
      #helps us understand the demographic profile of the county where oil drilling
      ↪is so prevelant.
```

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

```
[17]: #difficulty reading file because .geojson was not saved as part of file name in
      ↪the directory
      tcincome = gpd.read_file('tulareincome.geojson')
```

```
[18]: tcincome.dtypes
```

```
[18]: geoid          object
      name          object
      B19013001      float64
      B19013001, Error  float64
      geometry       geometry
      dtype: object
```

```
[19]: tcincome.head(5)
```

```
[19]:
```

	geoid	name	B19013001	\
0	05000US06107	Tulare County, CA	57394.0	
1	14000US06107000101	Census Tract 1.01, Tulare, CA	73582.0	
2	14000US06107000102	Census Tract 1.02, Tulare, CA	63796.0	
3	14000US06107000202	Census Tract 2.02, Tulare, CA	31445.0	
4	14000US06107000203	Census Tract 2.03, Tulare, CA	61154.0	

	B19013001, Error	geometry
0	1402.0	MULTIPOLYGON (((-118.80244 35.79017, -118.8027...
1	4699.0	MULTIPOLYGON (((-119.07932 36.32554, -119.0787...
2	28561.0	MULTIPOLYGON (((-119.24651 36.51340, -119.2462...
3	5786.0	MULTIPOLYGON (((-119.32270 36.54527, -119.3218...
4	18328.0	MULTIPOLYGON (((-119.30507 36.54962, -119.3050...

```
[20]: tcincome.shape
```

```
[20]: (104, 5)
```

```
[22]: #need to rename SE column
tcincome1 = tcincome.rename(columns={'B19013001': 'Median Household Income',
↳ 'geoid': 'Census Tract'})
tcincome1
```

```
[22]:
```

	Census Tract	name \
0	05000US06107	Tulare County, CA
1	14000US06107000101	Census Tract 1.01, Tulare, CA
2	14000US06107000102	Census Tract 1.02, Tulare, CA
3	14000US06107000202	Census Tract 2.02, Tulare, CA
4	14000US06107000203	Census Tract 2.03, Tulare, CA
..	...	...
99	14000US06107004301	Census Tract 43.01, Tulare, CA
100	14000US06107004302	Census Tract 43.02, Tulare, CA
101	14000US06107004401	Census Tract 44.01, Tulare, CA
102	14000US06107004402	Census Tract 44.02, Tulare, CA
103	14000US06107004500	Census Tract 45, Tulare, CA

	Median Household Income	B19013001, Error \
0	57394.0	1402.0
1	73582.0	4699.0
2	63796.0	28561.0
3	31445.0	5786.0
4	61154.0	18328.0
..	...	...
99	32500.0	16924.0
100	43938.0	6195.0
101	56741.0	11940.0
102	37886.0	7927.0
103	42819.0	10448.0

	geometry
0	MULTIPOLYGON (((-118.80244 35.79017, -118.8027...
1	MULTIPOLYGON (((-119.07932 36.32554, -119.0787...
2	MULTIPOLYGON (((-119.24651 36.51340, -119.2462...
3	MULTIPOLYGON (((-119.32270 36.54527, -119.3218...
4	MULTIPOLYGON (((-119.30507 36.54962, -119.3050...
..	...
99	MULTIPOLYGON (((-119.25041 35.89190, -119.2486...
100	MULTIPOLYGON (((-119.53806 35.79706, -119.5380...
101	MULTIPOLYGON (((-119.28591 35.88358, -119.2859...
102	MULTIPOLYGON (((-119.28593 35.87828, -119.2859...
103	MULTIPOLYGON (((-119.10783 35.89467, -119.1078...

[104 rows x 5 columns]

```
[25]: #getting ride of margin of error column
tcincome2 = tcincome1[['Census Tract', 'name', 'Median Household Income',
↪ 'geometry']]
tcincome2
```

```
[25]:
```

	Census Tract	name \
0	05000US06107	Tulare County, CA
1	14000US06107000101	Census Tract 1.01, Tulare, CA
2	14000US06107000102	Census Tract 1.02, Tulare, CA
3	14000US06107000202	Census Tract 2.02, Tulare, CA
4	14000US06107000203	Census Tract 2.03, Tulare, CA
..	...	...
99	14000US06107004301	Census Tract 43.01, Tulare, CA
100	14000US06107004302	Census Tract 43.02, Tulare, CA
101	14000US06107004401	Census Tract 44.01, Tulare, CA
102	14000US06107004402	Census Tract 44.02, Tulare, CA
103	14000US06107004500	Census Tract 45, Tulare, CA

	Median Household Income \
0	57394.0
1	73582.0
2	63796.0
3	31445.0
4	61154.0
..	...
99	32500.0
100	43938.0
101	56741.0
102	37886.0
103	42819.0

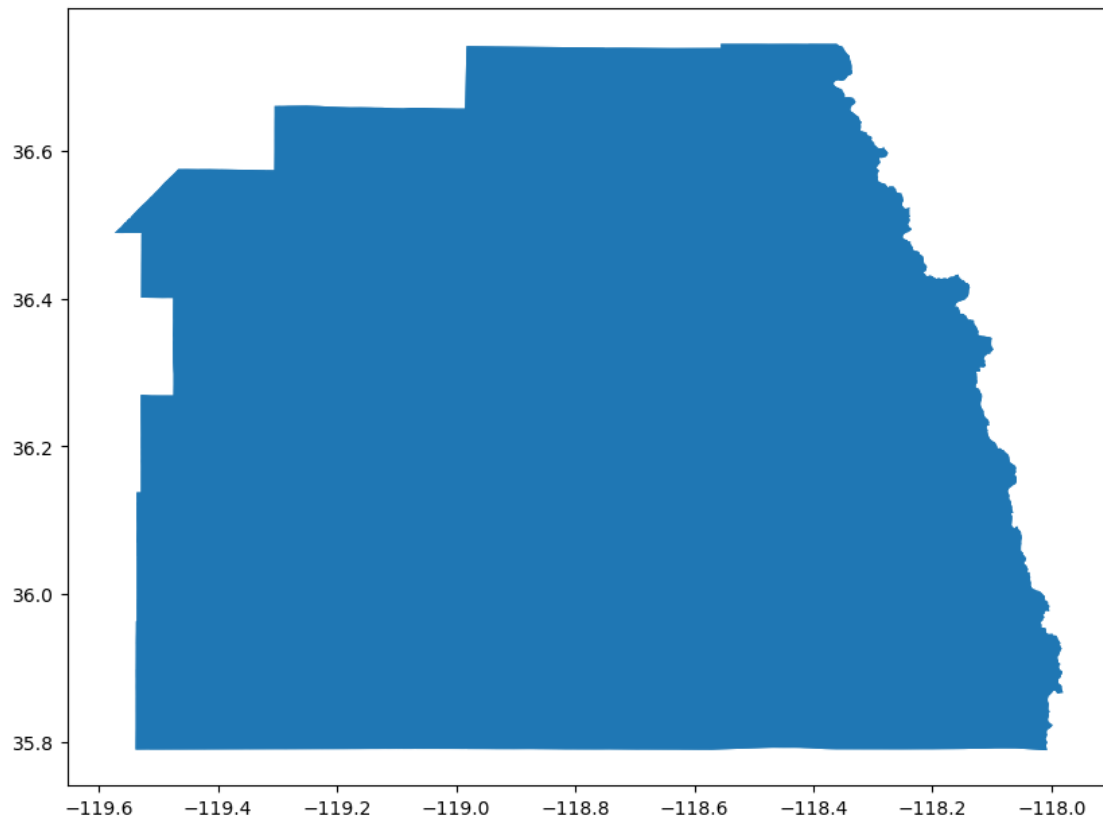
  

	geometry
0	MULTIPOLYGON (((-118.80244 35.79017, -118.8027...
1	MULTIPOLYGON (((-119.07932 36.32554, -119.0787...
2	MULTIPOLYGON (((-119.24651 36.51340, -119.2462...
3	MULTIPOLYGON (((-119.32270 36.54527, -119.3218...
4	MULTIPOLYGON (((-119.30507 36.54962, -119.3050...
..	...
99	MULTIPOLYGON (((-119.25041 35.89190, -119.2486...
100	MULTIPOLYGON (((-119.53806 35.79706, -119.5380...
101	MULTIPOLYGON (((-119.28591 35.88358, -119.2859...
102	MULTIPOLYGON (((-119.28593 35.87828, -119.2859...
103	MULTIPOLYGON (((-119.10783 35.89467, -119.1078...

[104 rows x 4 columns]

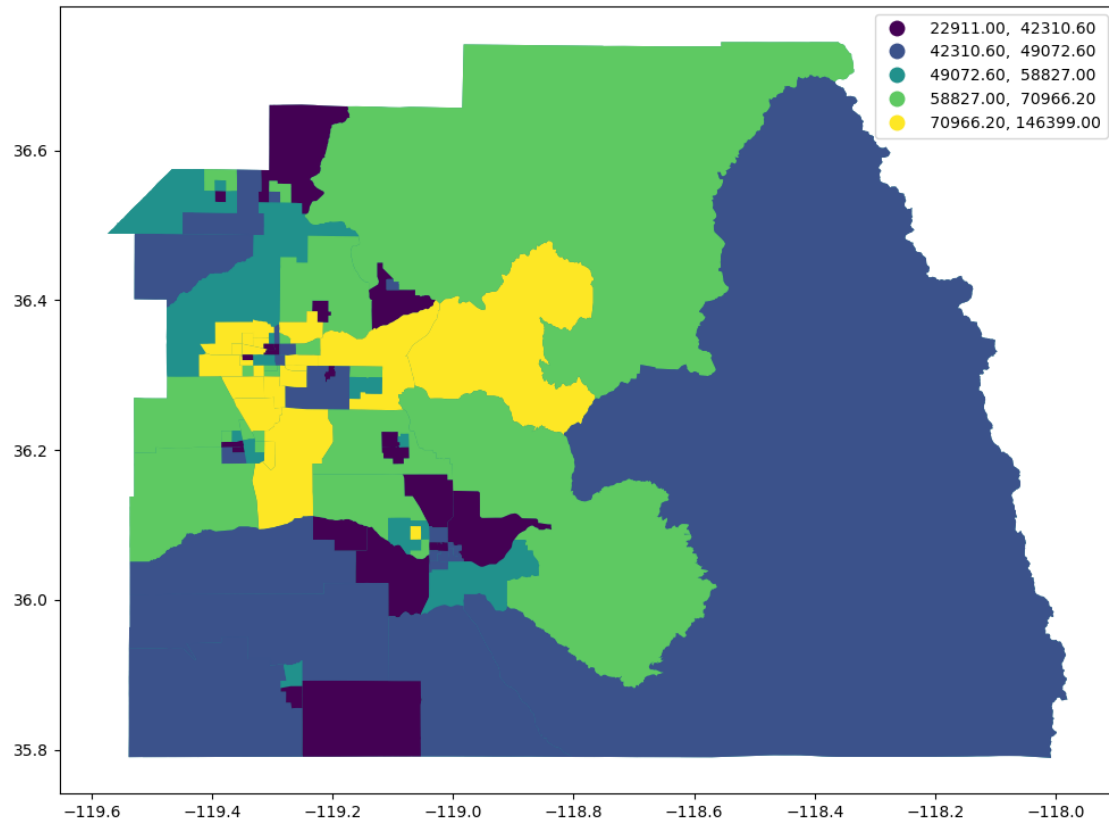
```
[27]: tcincome2.plot(figsize = (10,10))
```

[27]: <AxesSubplot: >



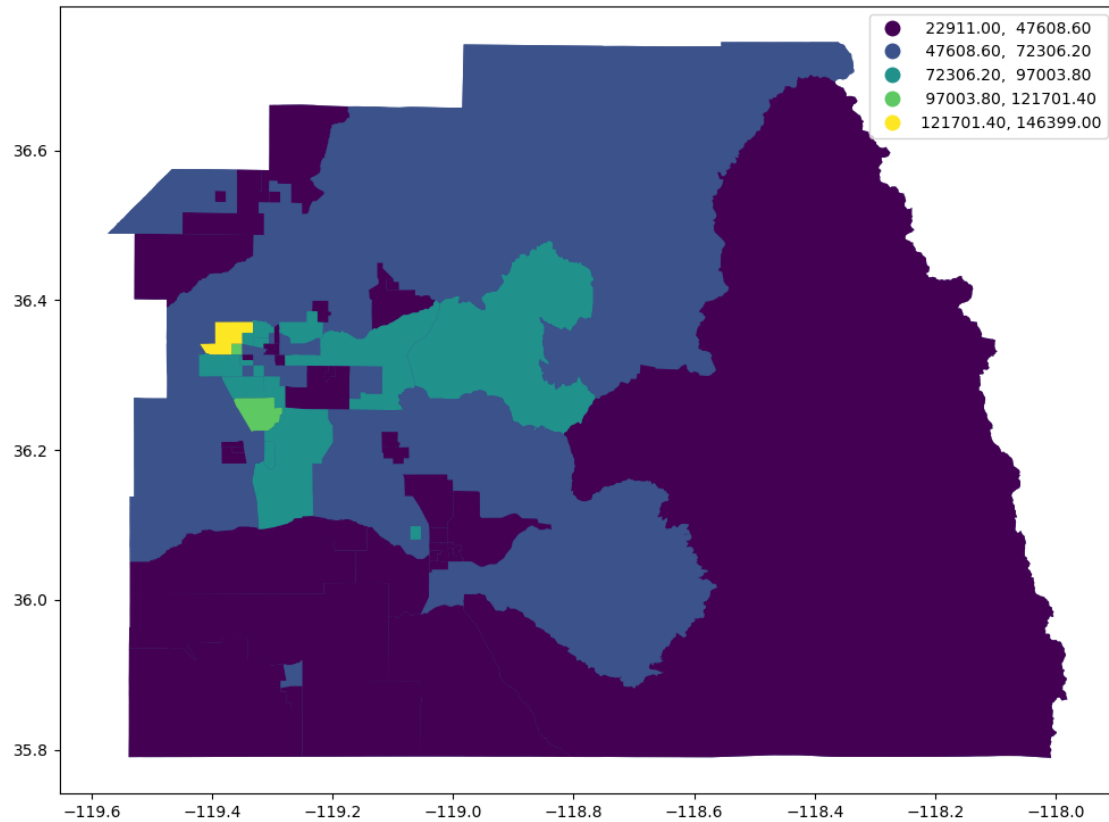
```
[28]: #This provides an even categorization of the data but does not actually provide  
↪an accurate sense of how the data is clustered.  
tcincome2.plot(figsize=(12,10),  
               column='Median Household Income',  
               legend=True,  
               scheme='quantiles')
```

[28]: <AxesSubplot: >



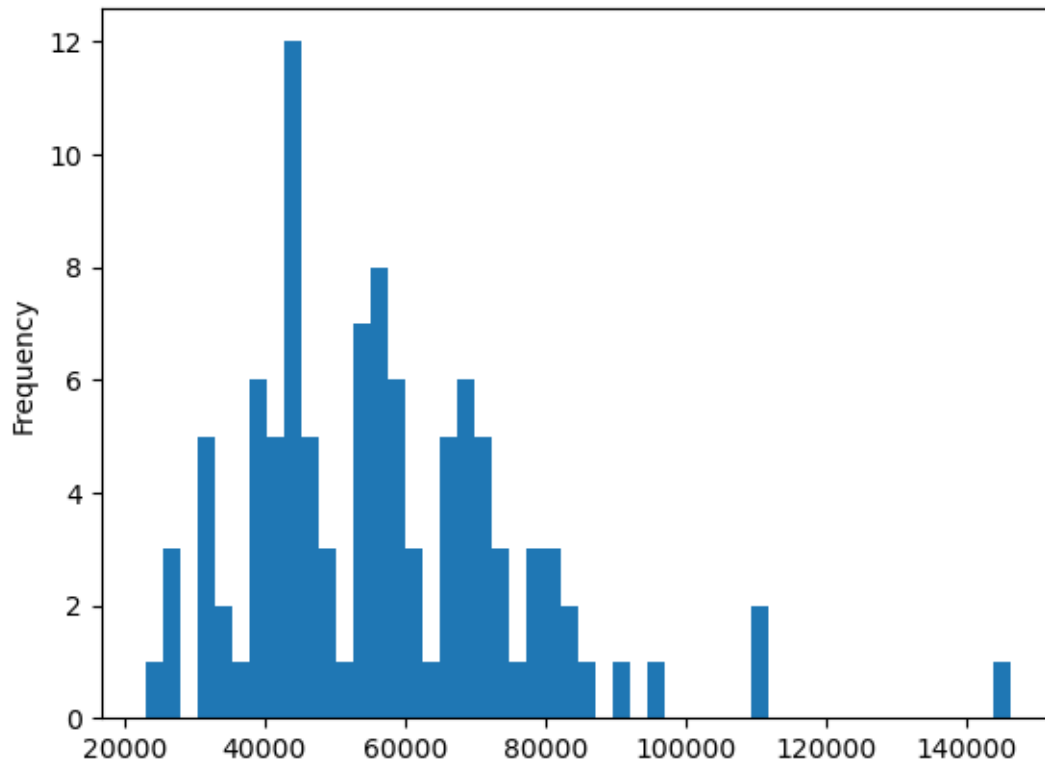
```
[30]: #plotting by natural breaks shows that the periphery of the county is much
      ↪ lower income than the center and west tracts.
tcincome2.plot(figsize=(12,10),
               column='Median Household Income',
               legend=True,
               scheme='equal_interval')
```

```
[30]: <AxesSubplot: >
```



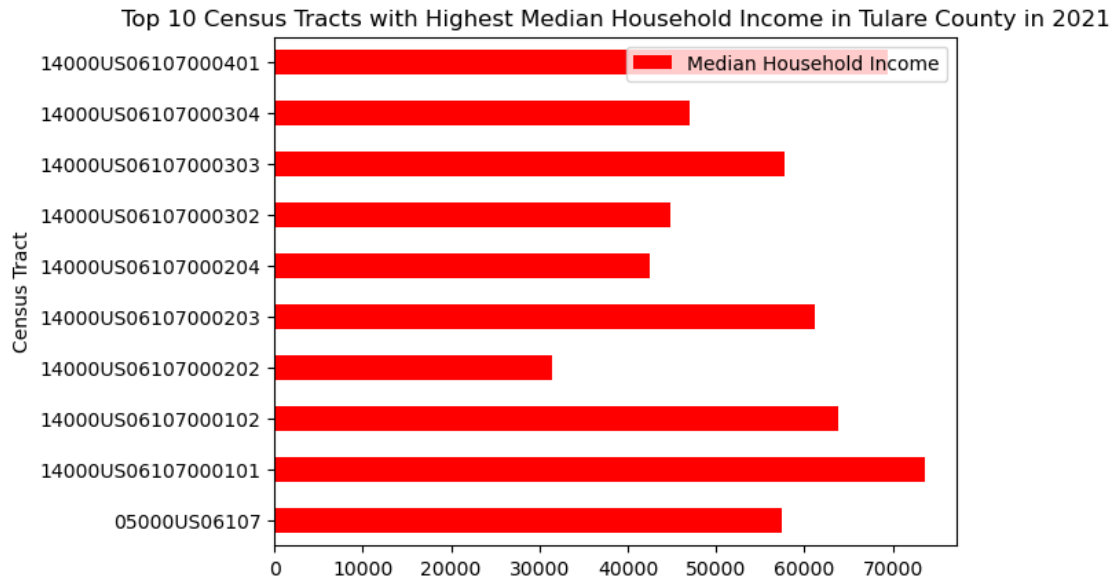
```
[31]: #this histogram shows the distribution of census tract median household income.
      ↪ There are a few outliers above $100,000
      #but the greatest number of census tracts have incomes between 40,000 and
      ↪ 60,000.
      tcincome2['Median Household Income'].plot.hist(bins=50)
```

```
[31]: <AxesSubplot: ylabel='Frequency'>
```



```
[48]: tcincome2.head(10).plot.barh(x='Census Tract',
                                   y='Median Household Income',
                                   title='Top 10 Census Tracts with Highest Median_
Household Income in Tulare County in 2021',
                                   color='red')
```

```
[48]: <AxesSubplot: title={'center': 'Top 10 Census Tracts with Highest Median
Household Income in Tulare County in 2021'}, ylabel='Census Tract'>
```

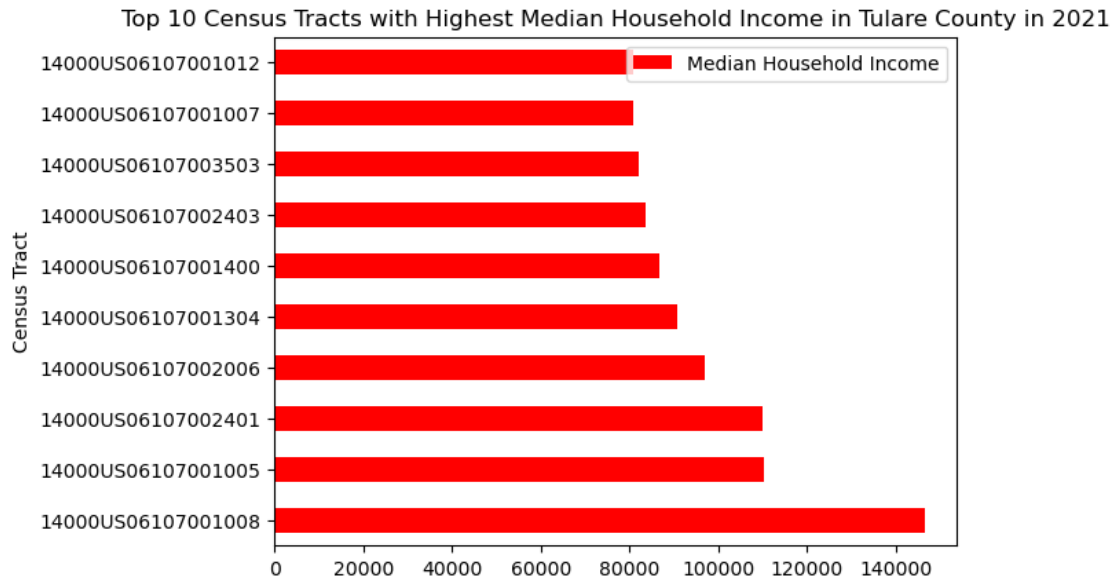


- ```
[35]: #sort data so it can be plotted more cleanly on the bar graph
tcincome3 = tcincome2.sort_values(by='Median Household Income',ascending =False)

[36]: #this graph shows the highest median household income tracts, a few reaching
      ↪beyond the median household income for California
      #which was $81,575 in 2021 (Statista)
tcincome3.head(10).plot.barh(x='Census Tract',
                             y='Median Household Income',
                             title='Top 10 Census Tracts with Highest Median
      ↪Household Income in Tulare County in 2021',
                             color='red')

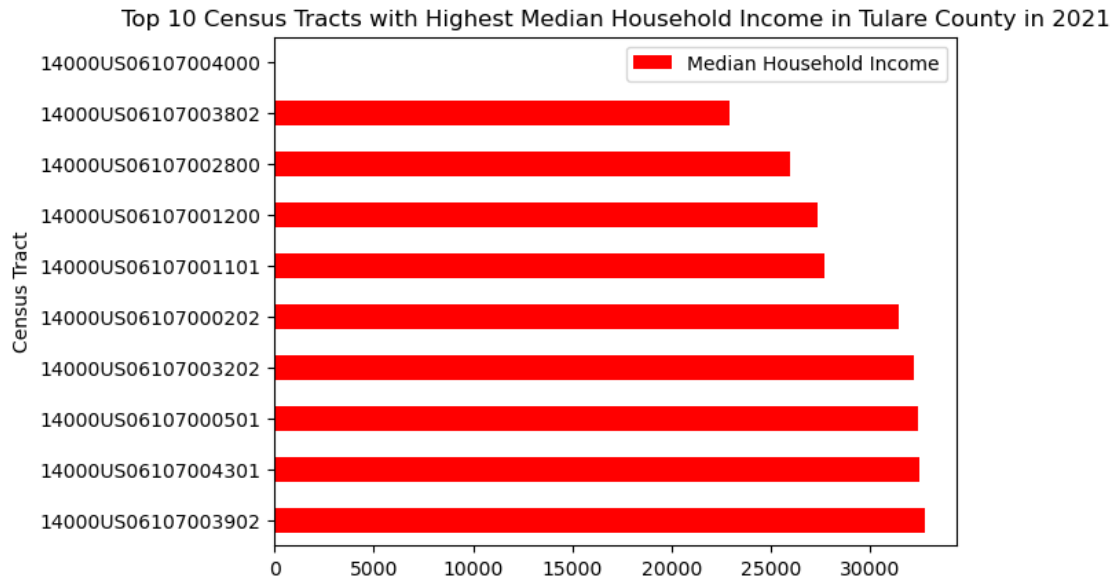
[36]: <AxesSubplot: title={'center': 'Top 10 Census Tracts with Highest Median
      Household Income in Tulare County in 2021'}, ylabel='Census Tract'>
```





```
[37]: #using tail to get lower income level of the county. The state poverty line was
      ↪$36,900 for a family of four in 2021
      # (Public Policy Institute of California, Poverty Fact Sheet). All of these
      ↪tracts except one fall below that line
      #The one tract without data suggests either an issue with data collection or
      ↪more likely that given the land use in the county
      #which is heavily agricultural or used for oil production, that this tract has
      ↪no residents.
      tcincome3.tail(10).plot.barh(x='Census Tract',
                                   y='Median Household Income',
                                   title='Top 10 Census Tracts with Highest Median
      ↪Household Income in Tulare County in 2021',
                                   color='red')
```

```
[37]: <AxesSubplot: title={'center': 'Top 10 Census Tracts with Highest Median
      Household Income in Tulare County in 2021'}, ylabel='Census Tract'>
```



```
[49]: tcincome3.tail()
```

```
[49]:
```

|    | Census Tract       | name \                         | Median Household Income | geometry                                          |
|----|--------------------|--------------------------------|-------------------------|---------------------------------------------------|
| 28 | 14000US06107001101 | Census Tract 11.01, Tulare, CA | 27731.0                 | MULTIPOLYGON (((-119.29668 36.34182, -119.2965... |
| 30 | 14000US06107001200 | Census Tract 12, Tulare, CA    | 27375.0                 | MULTIPOLYGON (((-119.31403 36.32924, -119.3140... |
| 70 | 14000US06107002800 | Census Tract 28, Tulare, CA    | 25956.0                 | MULTIPOLYGON (((-119.11793 36.20350, -119.1174... |
| 90 | 14000US06107003802 | Census Tract 38.02, Tulare, CA | 22911.0                 | MULTIPOLYGON (((-119.02535 36.07336, -119.0253... |
| 94 | 14000US06107004000 | Census Tract 40, Tulare, CA    | NaN                     | MULTIPOLYGON (((-118.99095 36.04037, -118.9906... |

```
[ ]:
```