Visualizing

April 12, 2023

[47]:
```
import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
```

[48]:
```
df = pd.read_csv("Refined_Data.csv")
df.dtypes
```

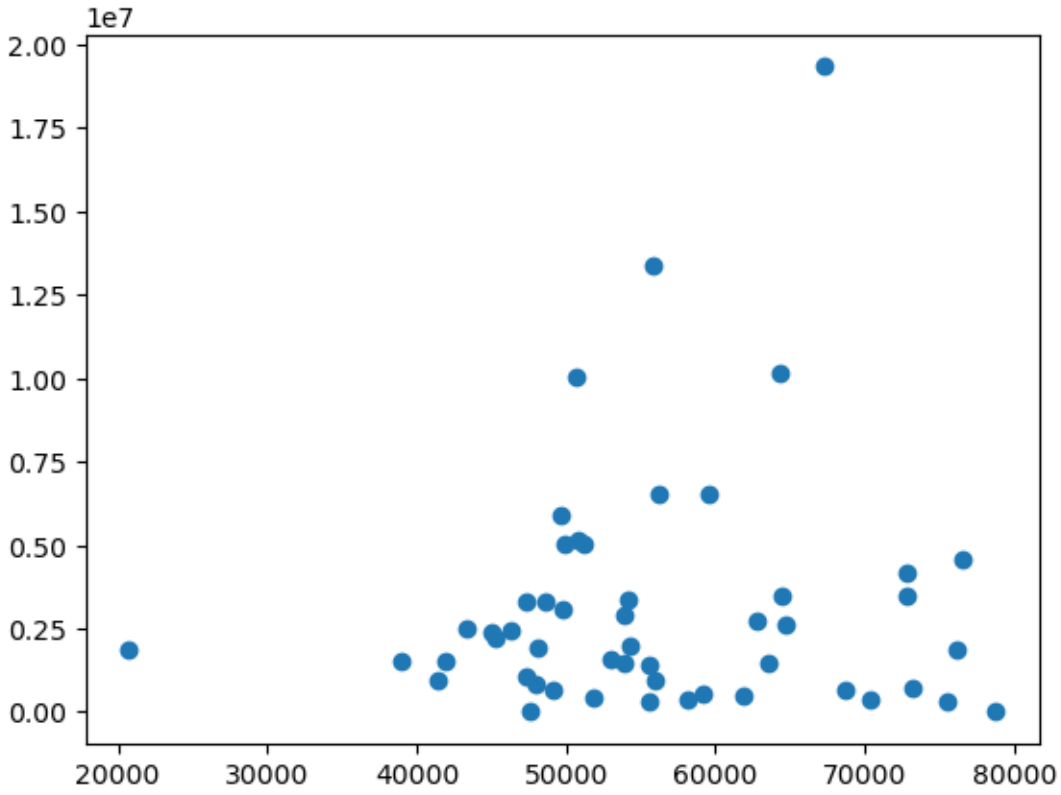
[48]:
```
Unnamed: 0      int64
State          object
TotalPop        int64
Hispanic      float64
White         float64
Black         float64
Native        float64
Asian         float64
Pacific       float64
Income        float64
males           int64
females         int64
dtype: object
```

[49]:
```
df.head()
```

[49]:

	Unnamed: 0	State	TotalPop	Hispanic	White	Black	Native	Asian	Pacific	Income	males	females
0	0	Alabama	4830620	3.751616	61.878656	31.252976	0.453231	1.050255	0.034354	43296.358603	2341093	2489527
1	1	Alaska	733375	5.909581	60.910180	2.848503	16.391018	5.450299	1.058683	70354.743902	384160	349215
2	2	Arizona	6641928	29.565921	57.120000	3.850987	4.355066	2.876579	0.167632	54207.820955	3299088	3342840
3	3	Arkansas	2958208	6.215474	71.137810	18.968759	0.522920	1.142336	0.146861	41935.633968	1451913	1506295
4	4	California	38421464	37.291875	40.215789	5.677396	0.405292	13.052234	0.351410	67264.782303	19087135	19334329

[50]:
```
plt.scatter(df["Income"],df["females"])
plt.show()
```



[15]:
```
# I already handled NULL Values so no error ;). I replaced those values with 0␣
 ↪in that data. Now replaceing it
# with what makes sense
```

[55]:
```
df['females'] = df['females'].mask(df['females'] == 0, df['TotalPop'] -␣
 ↪df['males'])
```

[56]:
```
df
```

[56]:

	Unnamed: 0	State	TotalPop	Hispanic	White	\
0	0	Alabama	4830620	3.751616	61.878656	
1	1	Alaska	733375	5.909581	60.910180	
2	2	Arizona	6641928	29.565921	57.120000	
3	3	Arkansas	2958208	6.215474	71.137810	
4	4	California	38421464	37.291875	40.215789	
5	5	Colorado	5278906	20.784380	69.895572	
6	1	Connecticut	3593222	15.604831	67.677053	
7	2	Delaware	926454	8.824766	64.632710	

8	3	District of Columbia	647484	9.165922	33.103911
9	4	Florida	19645772	21.338543	59.083749
10	5	Georgia	10006693	8.418242	54.286306
11	1	Hawaii	1406299	9.186709	25.032278
12	2	Idaho	1616547	11.505369	83.136242
13	3	Illinois	12873761	15.601734	60.859807
14	4	Indiana	6568645	6.536744	78.431894
15	5	Iowa	3093526	5.303645	87.719684
16	1	Kansas	2892987	11.644342	75.958289
17	2	Kentucky	4397353	3.222994	85.230748
18	3	Louisiana	4625253	4.866489	54.978546
19	4	Maine	1329100	1.431909	93.707407
20	5	Maryland	5930538	8.472498	52.679050
21	1	Massachusetts	6705586	11.461066	73.041052
22	2	Michigan	9900571	4.634993	72.381722
23	3	Minnesota	5419171	5.152924	81.427061
24	4	Mississippi	2988081	2.842401	53.286322
25	5	Missouri	6045448	4.037248	77.508069
26	1	Montana	1014699	3.268889	86.415556
27	2	Nebraska	1869365	9.203759	81.139474
28	3	Nevada	2798636	27.100884	53.239323
29	4	New Hampshire	1324201	3.321918	91.319178
30	5	New Jersey	8904413	18.749500	56.488761
31	1	New Mexico	2084117	45.282932	40.697992
32	2	New York	19673174	17.241425	56.470105
33	3	North Carolina	9845333	8.464763	64.597651
34	4	North Dakota	721640	2.832683	87.448293
35	5	Ohio	11575977	3.672084	75.903060
36	1	Oklahoma	3849733	10.079904	66.059426
37	2	Oregon	3939233	11.441212	78.395515
38	3	Pennsylvania	12779559	6.128014	77.383854
39	4	Puerto Rico	3583073	98.893574	0.773619
40	5	Rhode Island	1053661	13.356667	74.325417
41	1	South Carolina	4777576	5.056685	62.888736
42	2	South Dakota	843190	3.239640	82.500901
43	3	Tennessee	6499615	4.720027	73.490088
44	4	Texas	26538614	38.046738	44.687909
45	5	Utah	2903379	13.468376	79.406838
46	1	Vermont	626604	1.609290	93.983060
47	2	Virginia	8256630	8.011016	63.271048
48	3	Washington	6985464	11.140969	72.038408
49	4	West Virginia	1851420	1.290909	92.176240
50	5	Wisconsin	5742117	6.683333	79.864009

	Black	Native	Asian	Pacific	Income	males	\
0	31.252976	0.453231	1.050255	0.034354	43296.358603	2341093	
1	2.848503	16.391018	5.450299	1.058683	70354.743902	384160	

2	3.850987	4.355066	2.876579	0.167632	54207.820955	3299088
3	18.968759	0.522920	1.142336	0.146861	41935.633968	1451913
4	5.677396	0.405292	13.052234	0.351410	67264.782303	19087135
5	3.546377	0.573833	2.661997	NaN	64657.801787	2648667
6	10.348068	0.126208	4.021981	0.018599	76146.560588	1751607
7	20.743925	0.259813	3.268692	NaN	61827.976636	448413
8	51.776536	0.200559	3.383240	0.029609	75466.363636	306674
9	15.165676	0.210451	2.283174	0.051510	50690.194987	9600009
10	32.088298	0.187583	3.097649	0.046602	50811.082051	4883331
11	2.052848	0.144937	36.592089	8.758861	73264.426282	709871
12	0.566779	1.468121	1.135906	0.127181	48017.315436	810464
13	17.108411	0.118427	4.475377	0.020032	59587.048875	6316899
14	11.186977	0.194086	1.578272	0.032625	48616.227848	3235263
15	3.256987	0.289793	1.699392	0.055164	53017.753041	1534595
16	6.567895	0.733947	2.331053	NaN	53885.612648	1439862
17	8.272317	0.166637	1.129847	0.046438	45285.802536	2164208
18	36.326241	0.484309	1.669060	0.039184	44957.993761	2261156
19	1.134473	0.788319	0.965812	0.015670	49181.974359	650081
20	30.677754	0.203096	5.325414	0.036285	78765.400725	2872643
21	6.833128	0.128279	5.835656	0.019809	72838.936726	3249650
22	17.633103	0.484411	2.423110	0.019549	51201.830037	4861973
23	5.659820	1.069040	4.156072	0.032909	62820.833959	2692166
24	41.491945	0.389970	0.876444	0.015046	38909.919207	1451723
25	14.122118	0.363329	1.624496	0.101657	49763.987726	2964003
26	0.429259	7.060741	0.570370	0.072222	47645.682836	510163
27	4.956203	0.864474	1.859023	0.057143	55916.469697	929606
28	7.739617	1.087187	7.095729	0.574521	55526.525074	1407735
29	1.227740	0.142808	2.191438	0.016096	68728.859589	653484
30	14.387862	0.115335	8.159990	0.031319	76581.083417	4343027
31	1.755020	9.248594	1.234337	0.042771	47329.967871	1032414
32	15.668046	0.321639	7.897159	0.023451	64290.749113	9541801
33	21.395117	1.085491	2.317457	0.052326	49937.464137	4795408
34	1.284390	5.651220	0.961951	NaN	58188.112195	367963
35	16.207276	0.168888	1.621081	0.022645	49655.248466	5662893
36	8.314737	6.716842	1.801148	0.106220	48100.854267	1906944
37	1.730788	1.000242	3.594909	0.345333	54271.901818	1948453
38	11.633948	0.119269	2.797751	0.019394	56170.464510	6245344
39	0.092559	0.002818	0.075197	0.001240	20720.538286	1713860
40	5.682917	0.346250	3.247500	0.035833	59125.270833	510388
41	28.750916	0.292399	1.249176	0.046978	46296.807763	2322409
42	1.423874	9.417568	1.019369	0.041892	51805.405405	423477
43	18.283817	0.226635	1.407283	0.043156	47328.083617	3167756
44	11.650048	0.261144	3.669696	0.068816	55874.522601	13171316
45	1.017949	1.081368	2.196068	0.825983	63488.917808	1459229
46	0.980874	0.301639	1.238798	0.030601	55602.967213	308573
47	20.175998	0.212453	5.455242	0.064715	72866.013412	4060948
48	3.384429	1.410727	7.022007	0.609896	64493.767684	3487725

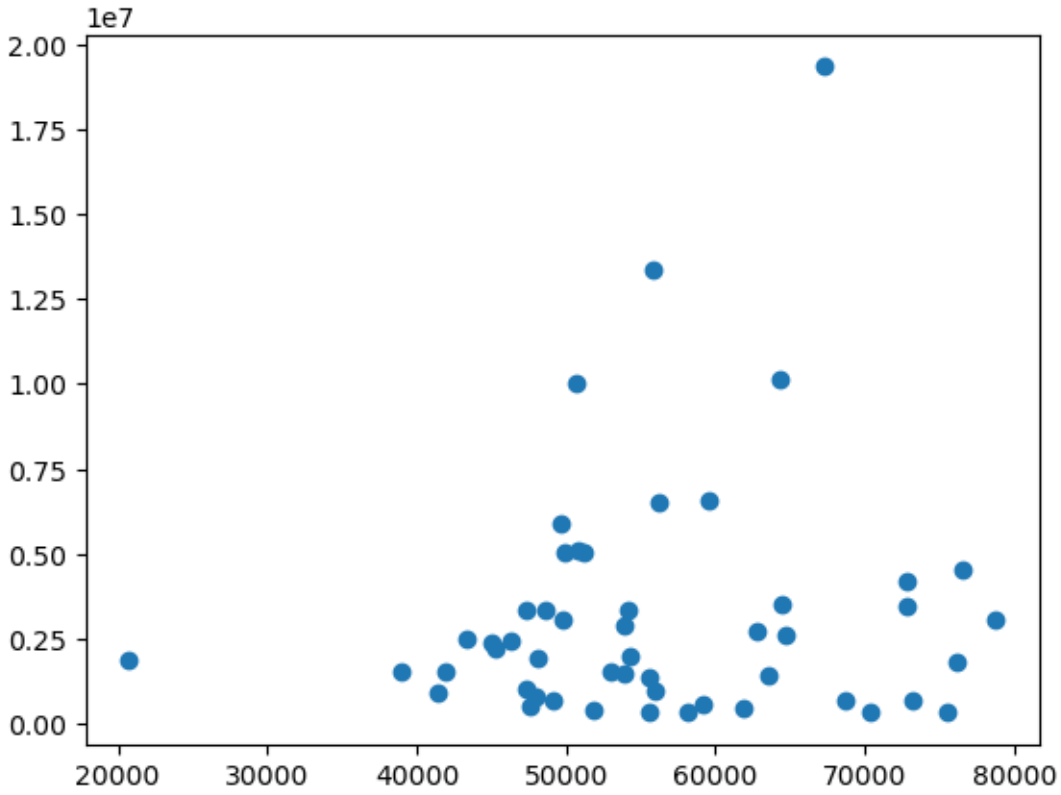
49	3.662810	0.152686	0.682438	0.026446	41437.111570	913631
50	8.195187	0.953664	2.404239	0.020833	53898.889209	2851385

	females
0	2489527
1	349215
2	3342840
3	1506295
4	19334329
5	2630239
6	1841615
7	478041
8	340810
9	10045763
10	5123362
11	696428
12	806083
13	6556862
14	3333382
15	1558931
16	1453125
17	2233145
18	2364097
19	679019
20	3057895
21	3455936
22	5038598
23	2727005
24	1536358
25	3081445
26	504536
27	939759
28	1390901
29	670717
30	4561386
31	1051703
32	10131373
33	5049925
34	353677
35	5913084
36	1942789
37	1990780
38	6534215
39	1869213
40	543273
41	2455167
42	419713

```
43   3331859
44  13367298
45   1444150
46    318031
47   4195682
48   3497739
49    937789
50   2890732
```

[15]:
```
# All females Column Filled
```

[57]:
```
plt.scatter(df["Income"],df["females"])
plt.show()
```



[58]:
```
df.head()
```

[58]:

	Unnamed: 0	State	TotalPop	Hispanic	White	Black	\
0	0	Alabama	4830620	3.751616	61.878656	31.252976	
1	1	Alaska	733375	5.909581	60.910180	2.848503	
2	2	Arizona	6641928	29.565921	57.120000	3.850987	
3	3	Arkansas	2958208	6.215474	71.137810	18.968759	

```
4           4  California  38421464  37.291875  40.215789   5.677396

      Native      Asian   Pacific        Income     males   females
0   0.453231   1.050255  0.034354  43296.358603   2341093   2489527
1  16.391018   5.450299  1.058683  70354.743902    384160    349215
2   4.355066   2.876579  0.167632  54207.820955   3299088   3342840
3   0.522920   1.142336  0.146861  41935.633968   1451913   1506295
4   0.405292  13.052234  0.351410  67264.782303  19087135  19334329
```

[59]:
```
df.isnull().sum()
```

[59]:
```
Unnamed: 0    0
State         0
TotalPop      0
Hispanic      0
White         0
Black         0
Native        0
Asian         0
Pacific       4
Income        0
males         0
females       0
dtype: int64
```

[60]:
```
df['Pacific'].fillna(100 - (df['Hispanic'] + df['White'] + df['Black'] +␣
↪df['Native'] + df['Asian']), inplace=True)
```

[61]:
```
df
```

[61]:
```
    Unnamed: 0                State  TotalPop   Hispanic      White  \
0            0              Alabama   4830620   3.751616  61.878656
1            1               Alaska    733375   5.909581  60.910180
2            2              Arizona   6641928  29.565921  57.120000
3            3             Arkansas   2958208   6.215474  71.137810
4            4           California  38421464  37.291875  40.215789
5            5             Colorado   5278906  20.784380  69.895572
6            1          Connecticut   3593222  15.604831  67.677053
7            2             Delaware    926454   8.824766  64.632710
8            3  District of Columbia    647484   9.165922  33.103911
9            4              Florida  19645772  21.338543  59.083749
10           5              Georgia  10006693   8.418242  54.286306
11           1               Hawaii   1406299   9.186709  25.032278
12           2                Idaho   1616547  11.505369  83.136242
13           3             Illinois  12873761  15.601734  60.859807
14           4              Indiana   6568645   6.536744  78.431894
15           5                 Iowa   3093526   5.303645  87.719684
```

16	1	Kansas	2892987	11.644342	75.958289	
17	2	Kentucky	4397353	3.222994	85.230748	
18	3	Louisiana	4625253	4.866489	54.978546	
19	4	Maine	1329100	1.431909	93.707407	
20	5	Maryland	5930538	8.472498	52.679050	
21	1	Massachusetts	6705586	11.461066	73.041052	
22	2	Michigan	9900571	4.634993	72.381722	
23	3	Minnesota	5419171	5.152924	81.427061	
24	4	Mississippi	2988081	2.842401	53.286322	
25	5	Missouri	6045448	4.037248	77.508069	
26	1	Montana	1014699	3.268889	86.415556	
27	2	Nebraska	1869365	9.203759	81.139474	
28	3	Nevada	2798636	27.100884	53.239323	
29	4	New Hampshire	1324201	3.321918	91.319178	
30	5	New Jersey	8904413	18.749500	56.488761	
31	1	New Mexico	2084117	45.282932	40.697992	
32	2	New York	19673174	17.241425	56.470105	
33	3	North Carolina	9845333	8.464763	64.597651	
34	4	North Dakota	721640	2.832683	87.448293	
35	5	Ohio	11575977	3.672084	75.903060	
36	1	Oklahoma	3849733	10.079904	66.059426	
37	2	Oregon	3939233	11.441212	78.395515	
38	3	Pennsylvania	12779559	6.128014	77.383854	
39	4	Puerto Rico	3583073	98.893574	0.773619	
40	5	Rhode Island	1053661	13.356667	74.325417	
41	1	South Carolina	4777576	5.056685	62.888736	
42	2	South Dakota	843190	3.239640	82.500901	
43	3	Tennessee	6499615	4.720027	73.490088	
44	4	Texas	26538614	38.046738	44.687909	
45	5	Utah	2903379	13.468376	79.406838	
46	1	Vermont	626604	1.609290	93.983060	
47	2	Virginia	8256630	8.011016	63.271048	
48	3	Washington	6985464	11.140969	72.038408	
49	4	West Virginia	1851420	1.290909	92.176240	
50	5	Wisconsin	5742117	6.683333	79.864009	

	Black	Native	Asian	Pacific	Income	males	\
0	31.252976	0.453231	1.050255	0.034354	43296.358603	2341093	
1	2.848503	16.391018	5.450299	1.058683	70354.743902	384160	
2	3.850987	4.355066	2.876579	0.167632	54207.820955	3299088	
3	18.968759	0.522920	1.142336	0.146861	41935.633968	1451913	
4	5.677396	0.405292	13.052234	0.351410	67264.782303	19087135	
5	3.546377	0.573833	2.661997	2.537842	64657.801787	2648667	
6	10.348068	0.126208	4.021981	0.018599	76146.560588	1751607	
7	20.743925	0.259813	3.268692	2.270093	61827.976636	448413	
8	51.776536	0.200559	3.383240	0.029609	75466.363636	306674	
9	15.165676	0.210451	2.283174	0.051510	50690.194987	9600009	

10	32.088298	0.187583	3.097649	0.046602	50811.082051	4883331
11	2.052848	0.144937	36.592089	8.758861	73264.426282	709871
12	0.566779	1.468121	1.135906	0.127181	48017.315436	810464
13	17.108411	0.118427	4.475377	0.020032	59587.048875	6316899
14	11.186977	0.194086	1.578272	0.032625	48616.227848	3235263
15	3.256987	0.289793	1.699392	0.055164	53017.753041	1534595
16	6.567895	0.733947	2.331053	2.764474	53885.612648	1439862
17	8.272317	0.166637	1.129847	0.046438	45285.802536	2164208
18	36.326241	0.484309	1.669060	0.039184	44957.993761	2261156
19	1.134473	0.788319	0.965812	0.015670	49181.974359	650081
20	30.677754	0.203096	5.325414	0.036285	78765.400725	2872643
21	6.833128	0.128279	5.835656	0.019809	72838.936726	3249650
22	17.633103	0.484411	2.423110	0.019549	51201.830037	4861973
23	5.659820	1.069040	4.156072	0.032909	62820.833959	2692166
24	41.491945	0.389970	0.876444	0.015046	38909.919207	1451723
25	14.122118	0.363329	1.624496	0.101657	49763.987726	2964003
26	0.429259	7.060741	0.570370	0.072222	47645.682836	510163
27	4.956203	0.864474	1.859023	0.057143	55916.469697	929606
28	7.739617	1.087187	7.095729	0.574521	55526.525074	1407735
29	1.227740	0.142808	2.191438	0.016096	68728.859589	653484
30	14.387862	0.115335	8.159990	0.031319	76581.083417	4343027
31	1.755020	9.248594	1.234337	0.042771	47329.967871	1032414
32	15.668046	0.321639	7.897159	0.023451	64290.749113	9541801
33	21.395117	1.085491	2.317457	0.052326	49937.464137	4795408
34	1.284390	5.651220	0.961951	1.821463	58188.112195	367963
35	16.207276	0.168888	1.621081	0.022645	49655.248466	5662893
36	8.314737	6.716842	1.801148	0.106220	48100.854267	1906944
37	1.730788	1.000242	3.594909	0.345333	54271.901818	1948453
38	11.633948	0.119269	2.797751	0.019394	56170.464510	6245344
39	0.092559	0.002818	0.075197	0.001240	20720.538286	1713860
40	5.682917	0.346250	3.247500	0.035833	59125.270833	510388
41	28.750916	0.292399	1.249176	0.046978	46296.807763	2322409
42	1.423874	9.417568	1.019369	0.041892	51805.405405	423477
43	18.283817	0.226635	1.407283	0.043156	47328.083617	3167756
44	11.650048	0.261144	3.669696	0.068816	55874.522601	13171316
45	1.017949	1.081368	2.196068	0.825983	63488.917808	1459229
46	0.980874	0.301639	1.238798	0.030601	55602.967213	308573
47	20.175998	0.212453	5.455242	0.064715	72866.013412	4060948
48	3.384429	1.410727	7.022007	0.609896	64493.767684	3487725
49	3.662810	0.152686	0.682438	0.026446	41437.111570	913631
50	8.195187	0.953664	2.404239	0.020833	53898.889209	2851385

	females
0	2489527
1	349215
2	3342840
3	1506295

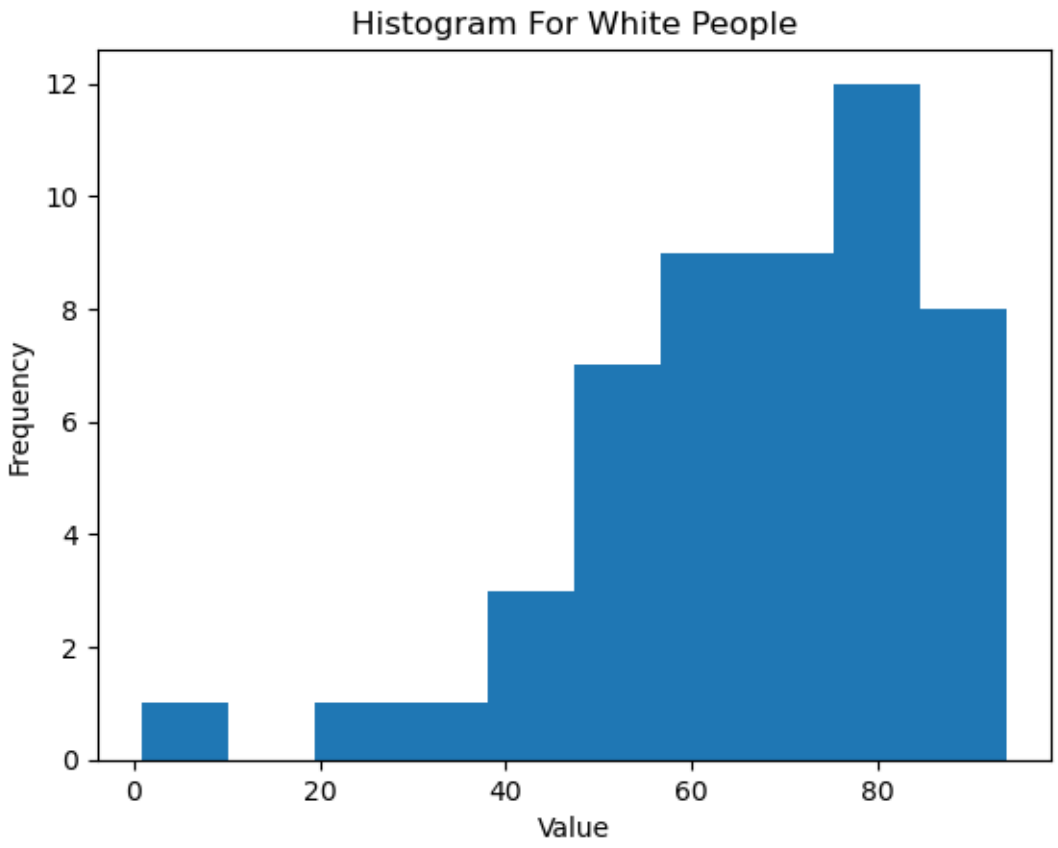
4	19334329
5	2630239
6	1841615
7	478041
8	340810
9	10045763
10	5123362
11	696428
12	806083
13	6556862
14	3333382
15	1558931
16	1453125
17	2233145
18	2364097
19	679019
20	3057895
21	3455936
22	5038598
23	2727005
24	1536358
25	3081445
26	504536
27	939759
28	1390901
29	670717
30	4561386
31	1051703
32	10131373
33	5049925
34	353677
35	5913084
36	1942789
37	1990780
38	6534215
39	1869213
40	543273
41	2455167
42	419713
43	3331859
44	13367298
45	1444150
46	318031
47	4195682
48	3497739
49	937789
50	2890732

[62]:
```
df.to_csv("FinalData.csv")
```

[63]:
```
df['White'].plot.hist(bins=10)

plt.title('Histogram For White People')
plt.xlabel('Value')
plt.ylabel('Frequency')
plt.show
```

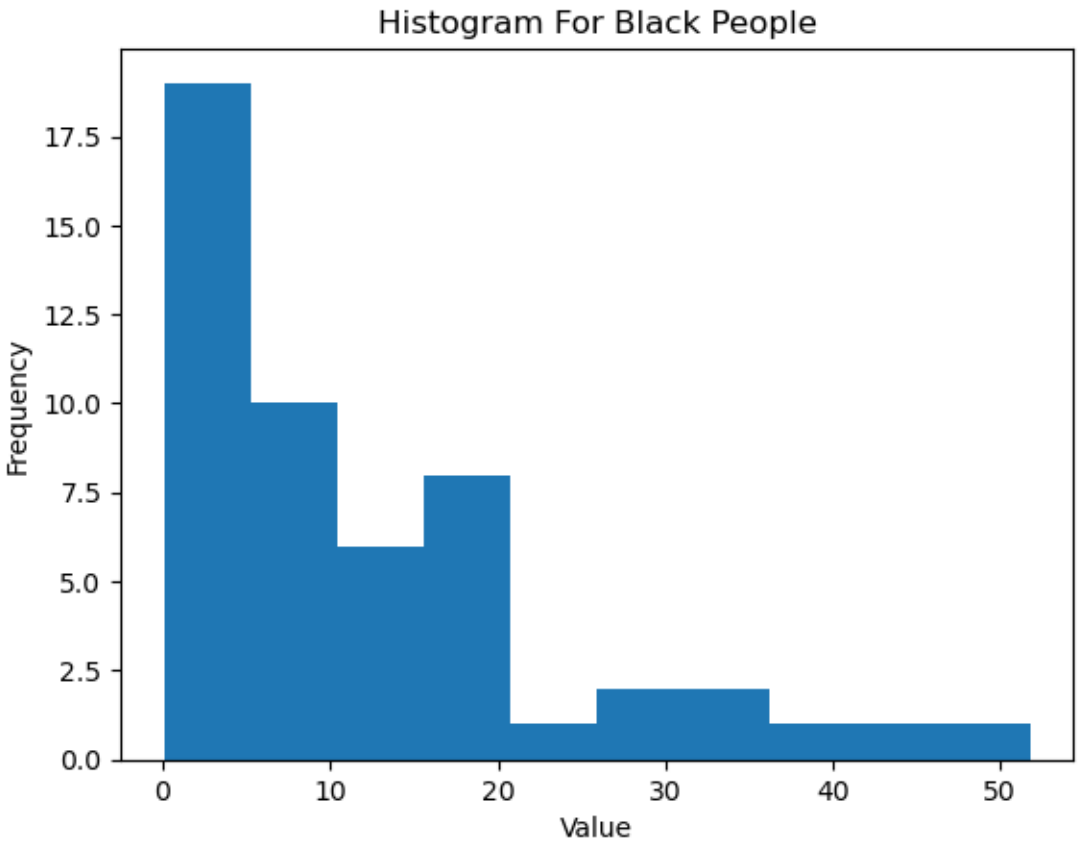
[63]:
```
<function matplotlib.pyplot.show(close=None, block=None)>
```



[64]:
```
df['Black'].plot.hist(bins=10)

plt.title('Histogram For Black People')
plt.xlabel('Value')
plt.ylabel('Frequency')
plt.show
```

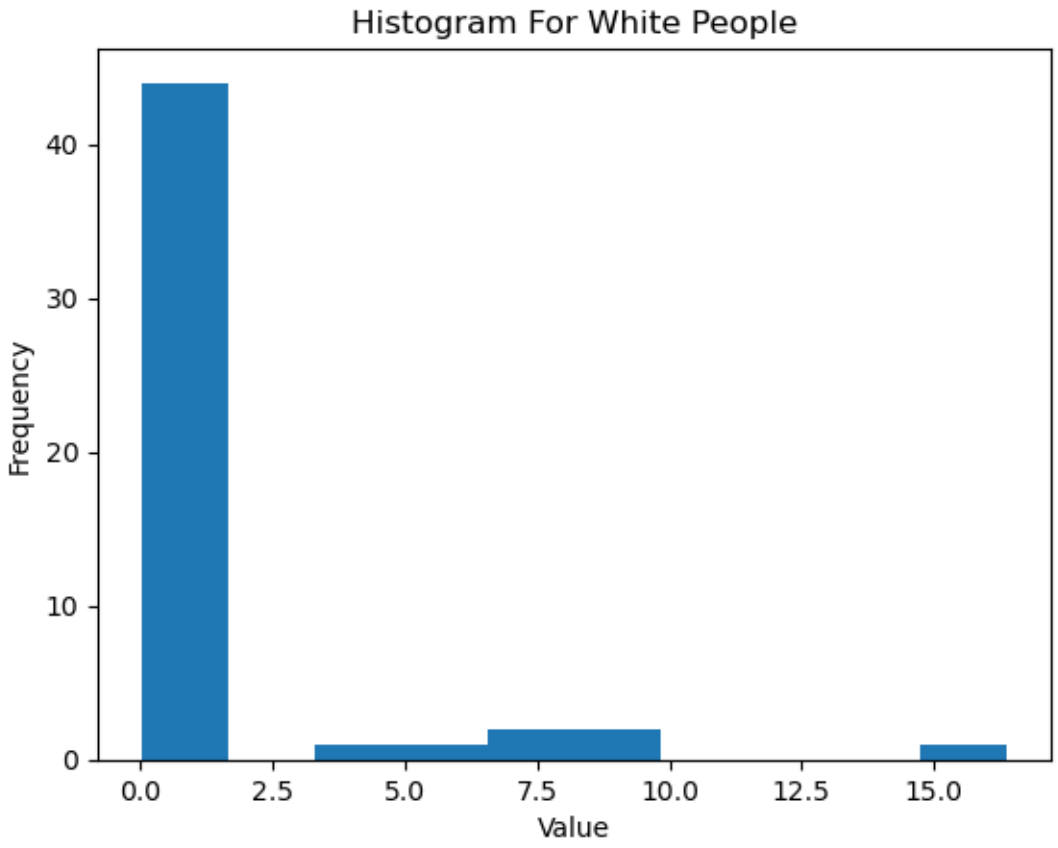
[64]:
```
<function matplotlib.pyplot.show(close=None, block=None)>
```



[65]:
```
df['Native'].plot.hist(bins=10)

plt.title('Histogram For White People')
plt.xlabel('Value')
plt.ylabel('Frequency')
plt.show
```

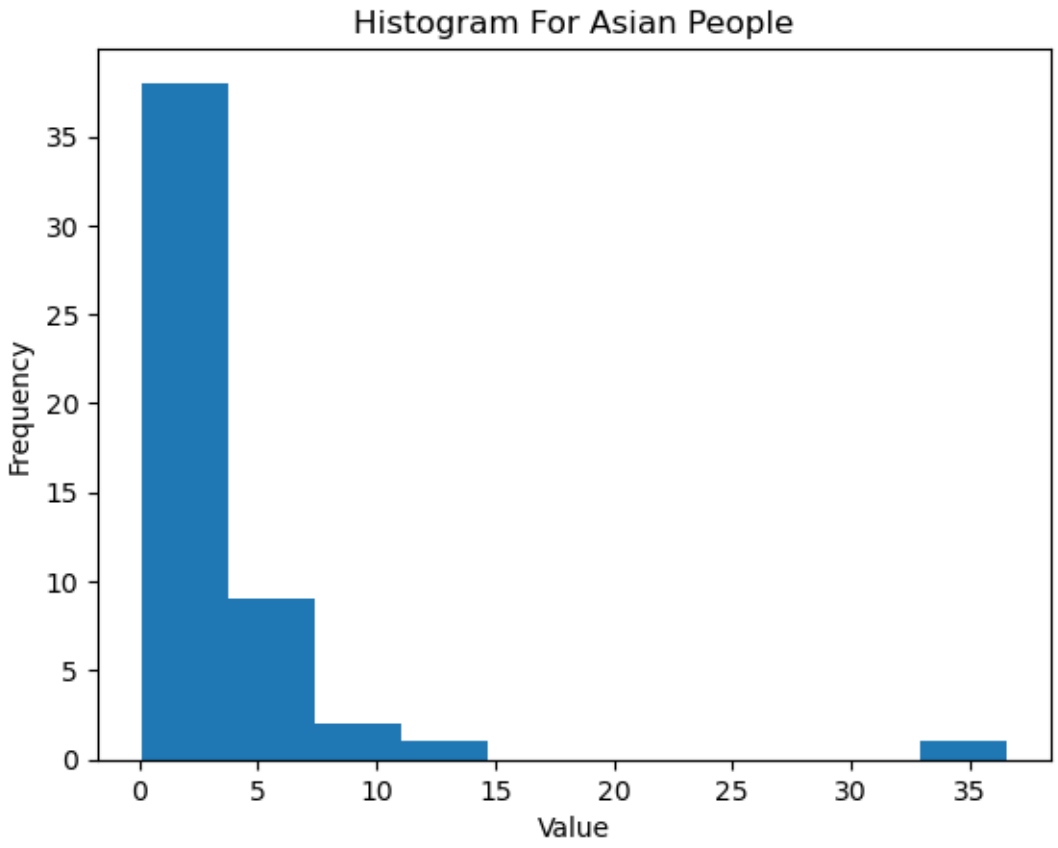
[65]:
```
<function matplotlib.pyplot.show(close=None, block=None)>
```



[44]:
```
df['Asian'].plot.hist(bins=10)

plt.title('Histogram For Asian People')
plt.xlabel('Value')
plt.ylabel('Frequency')
plt.show
```

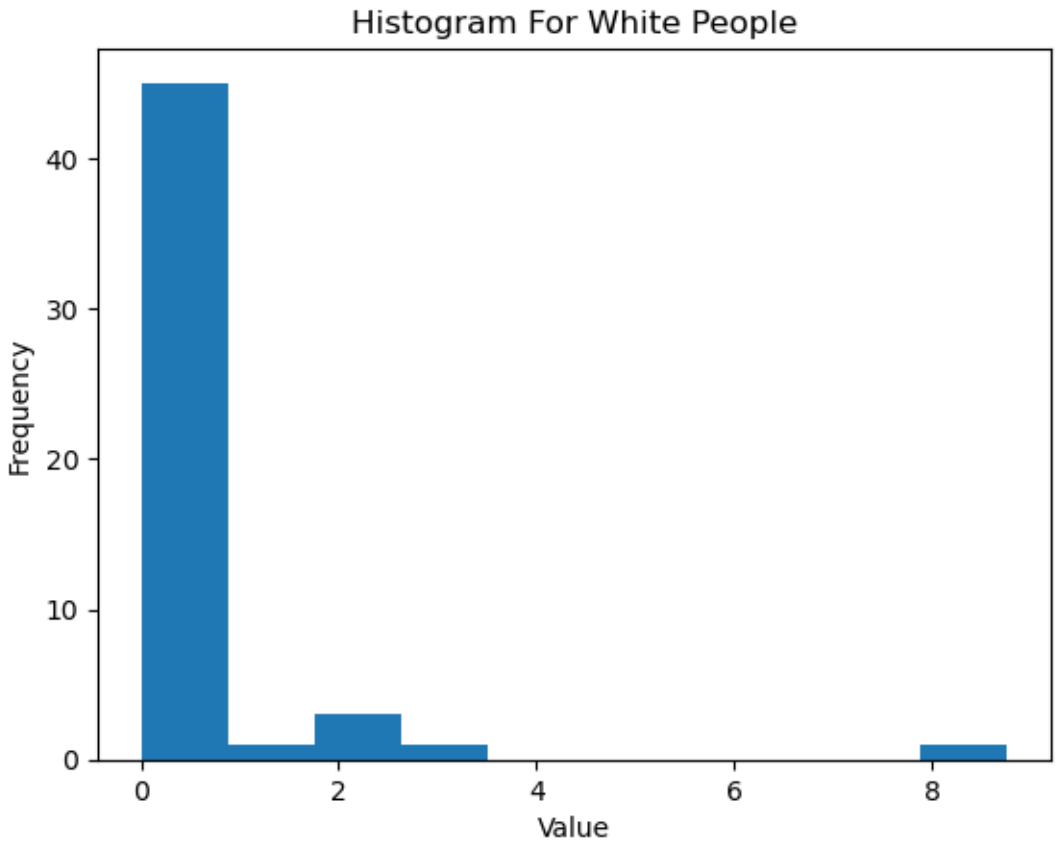
[44]: <function matplotlib.pyplot.show(close=None, block=None)>



[45]:
```
df['Pacific'].plot.hist(bins=10)

plt.title('Histogram For White People')
plt.xlabel('Value')
plt.ylabel('Frequency')
plt.show
```

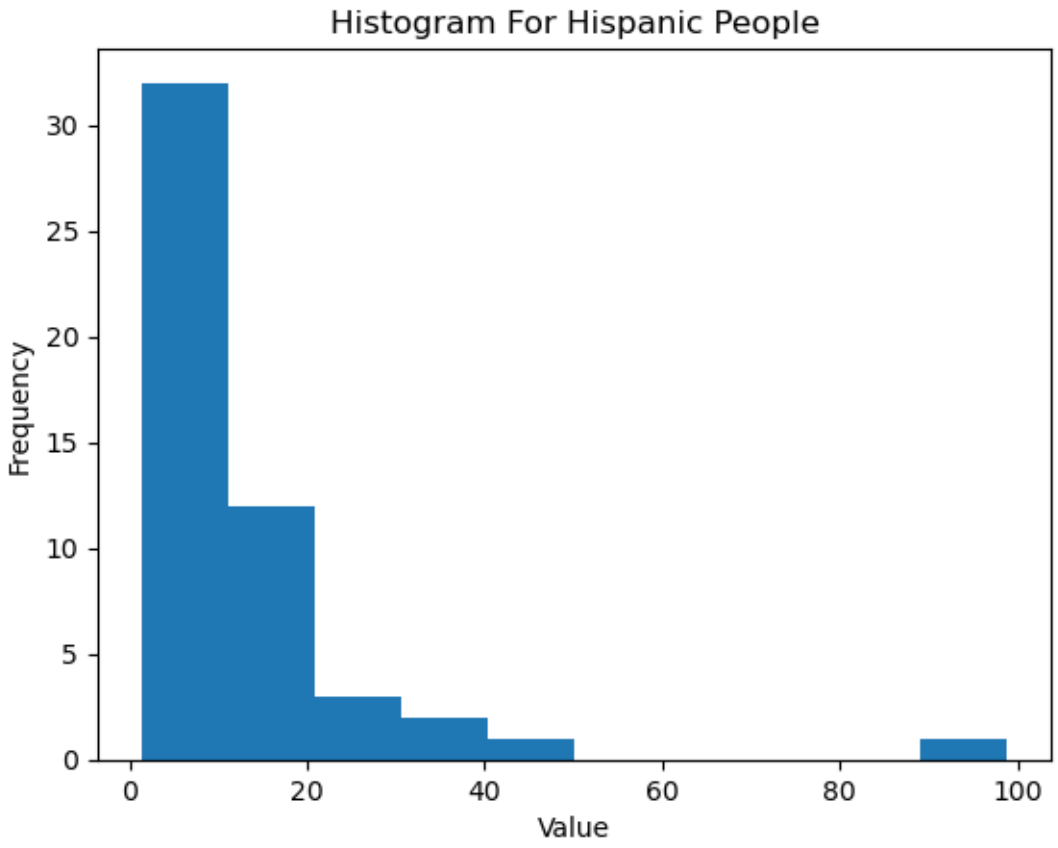
[45]:
```
<function matplotlib.pyplot.show(close=None, block=None)>
```



[46]:
```
df['Hispanic'].plot.hist(bins=10)

plt.title('Histogram For Hispanic People')
plt.xlabel('Value')
plt.ylabel('Frequency')
plt.show
```

[46]: <function matplotlib.pyplot.show(close=None, block=None)>


Histogram For Hispanic People
Frequency
Value
0
5
10
15
20
25
30
0
20
40
60
80
100

[ ]: