Assignment 5

Michael Entwistle, Momiana Fazal, Yosef Aryeh Kahan

July 24, 2023

1. To calculate the effect of the variables relative to each other, we can use two pieces of information, the correlation coefficient and the standard deviation. By seeing how a change of one standard deviation affect the dependent variable for each, we can compare strength of effect. In the case of x_1, the change for one standard deviation is β_1 times σ_1 which is 12 times 0.5, or six. For x_2, the change is β_2 times σ_2 which 1.5 times 5, or 7.5. zx_3 is a standardized regressor. This means it tells us what the change in the dependent variable is for a change of σ_3 in x_3 as opposed to for a change of one unit in x_3. This means we don't need to adjust the coefficient and can simply say a change of σ_3 produces a change of 4 in the dependent variable. Using this information we can say that x_2 has the greatest effect and x_3 has the weakest effect.

2.

Table 1: Table of Descrpitive Statistics of Data on Income Inequality in Urban Areas(2009)

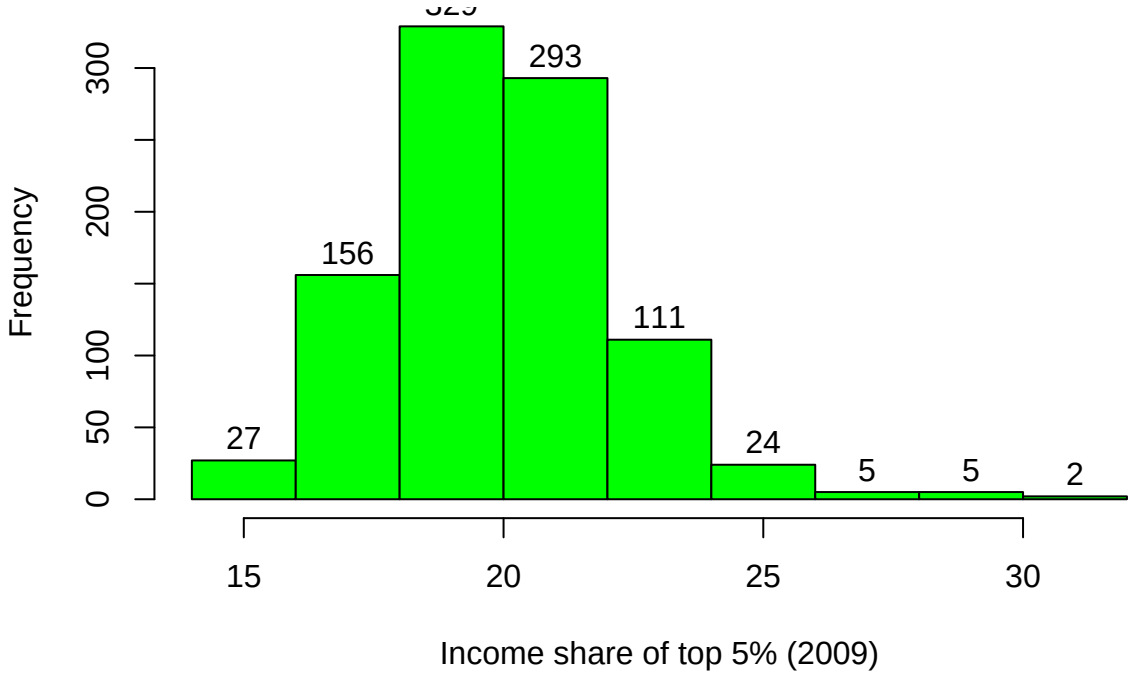
Statistic	Mean	Median	Min	Max	Pctl(25)	Pctl(75)	St. Dev.
Income share of the top 5%	19.948	19.900	14.400	31.400	18.500	21.300	2.256
Gini Index	44.017	44	35	56	42	46	3.195
Labor Force Participation Rate	62.007	62.500	31.000	83.400	58.300	66.400	6.549
Unemployment Rate	43.467	43.200	20.500	72.600	38.600	47.500	7.054
Number of households	111,470.600	28,297	3,911	6,750,902	15,128	65,746	362,874.400
Households with Income in the past 12 months below poverty level	15.812	14.330	4.170	60.270	11.500	18.170	6.952
Household with public assistance income	2.441	2.210	0.130	10.770	1.630	2.920	1.210
Total Population in thousands	30.029	7.346	1.11	18,912.644	3.947	170.049	1,020.203
Per capita income in the past 12 months (in 2009 inflation-adjusted dollars)	22,332.350	21,900.5	6,022	48,394	19,581.8	24,790.5	4,745.286
Population 16 Years and Older	78.781	78.910	65.750	87.730	77.220	80.560	2.904
Population between 25 and 64 years old	50.913	51.540	33.010	64.630	49.500	53.150	3.602
Population with a High school diploma	16.961	16.670	5.280	29.090	14.410	19.590	3.854
Population with some college degree or associate's degree	15.663	15.620	7.160	24.910	13.750	17.550	2.830
Population with a Bachelor's degree or higher	11.219	10.240	3.680	36.500	8.280	13.580	4.280
Male	49.466	49.190	45.070	65.260	48.580	49.960	1.650
Female	50.534	50.810	34.740	54.930	50.040	51.420	1.650
White alone (population)	82.328	86.650	15.500	98.710	75.360	93.310	14.207
Black or African American alone (population)	9.403	3.600	0.000	82.350	1.100	12.020	13.019
American Indian and Alaska Native alone (population)	1.278	0.360	0.000	72.690	0.210	0.780	4.239
Asian alone (population)	1.555	0.840	0.000	43.400	0.500	1.670	2.879
Hispanic or Latino (population)	10.955	4.170	0.170	99.870	1.840	10.560	17.918
Married (population)	41.555	41.650	24.610	54.670	38.940	44.630	4.387
Not Married (population)	37.226	37.030	23.000	55.280	33.930	39.820	4.759

Table 2: Table of Descrpitive Statistics of Data on Income Inequality in Urban Areas, 2014

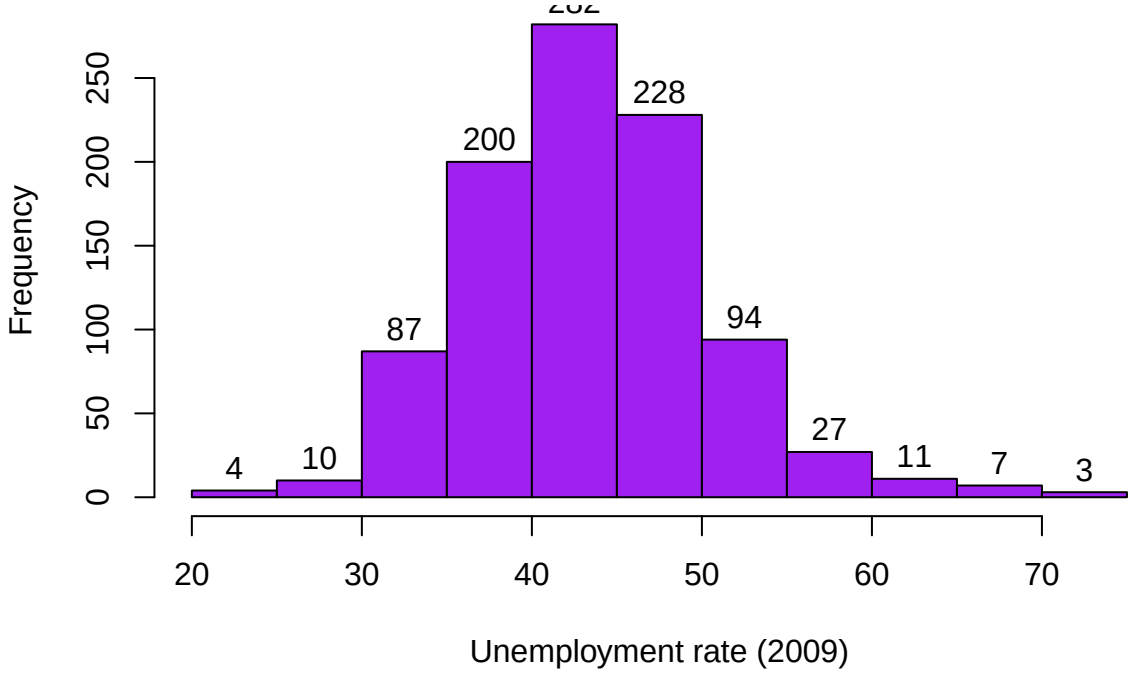
Statistic	Mean	Median	Min	Max	Pctl(25)	Pctl(75)	St. Dev.
Year	2,014.000	2,014	2,014	2,014	2,014	2,014	0.000
MSACode	29,749.730	29,780	10,100	49,820	19,940	39,460	11,370.330
Income share of the top 5%	20.509	20.430	14.480	30.820	19.020	21.830	2.167
Gini Index	44.780	44.610	37.180	60.230	42.780	46.620	3.003
Labor Force Participation Rate	60.376	60.900	24.300	80.000	56.700	65.000	6.841
Unemployment Rate	8.359	7.900	1.700	36.500	6.400	9.700	3.138
Number of households	118,445.100	29,558	3,705	7,105,003	15,361	68,502	383,833.900
Households with Income in the past 12 months below poverty level	16.812	15.530	4.760	63.690	12.930	18.960	6.574
Household with public assistance income	2.781	2.480	0.240	19.620	1.860	3.340	1.448
Total Population	321,505.000	76,237	13,265	19,865,045	40,496	184,652	1,079,669.000
Per capita income in the past 12 months (in 2009 inflation-adjusted dollars)	23,965.850	23,510	7,042	50,723	20,964	26,497	5,028.046
Population 16 Years and Older	79.848	79.890	69.150	97.400	78.240	81.580	2.896
Population between 25 and 64 years old	66.220	66.760	46.380	88.380	64.170	68.940	4.637
Population with a High school diploma	18.304	18.140	6.450	32.740	15.210	21.170	4.591
Population with some college degree or associate's degree	5.557	5.510	1.590	11.860	4.530	6.460	1.476
Population with a Bachelor's degree or higher	13.855	12.800	5.230	44.290	10.190	16.760	5.062
Male	49.702	49.410	45.680	67.030	48.730	50.200	1.700
Female	50.298	50.590	32.970	54.320	49.800	51.270	1.700
White alone (population)	82.539	87.150	16.660	98.260	76.240	92.690	13.961
Black or African American alone (population)	8.937	3.610	0.010	75.920	1.270	11.140	12.359
American Indian and Alaska Native alone (population)	1.323	0.360	0.000	74.370	0.200	0.820	4.306
Asian alone (population)	1.761	0.970	0.000	43.170	0.550	1.870	3.031
Hispanic or Latino (population)	12.604	5.520	0.160	99.920	2.650	12.810	18.586
Married (population)	40.142	40.570	22.440	59.220	37.680	42.860	4.184

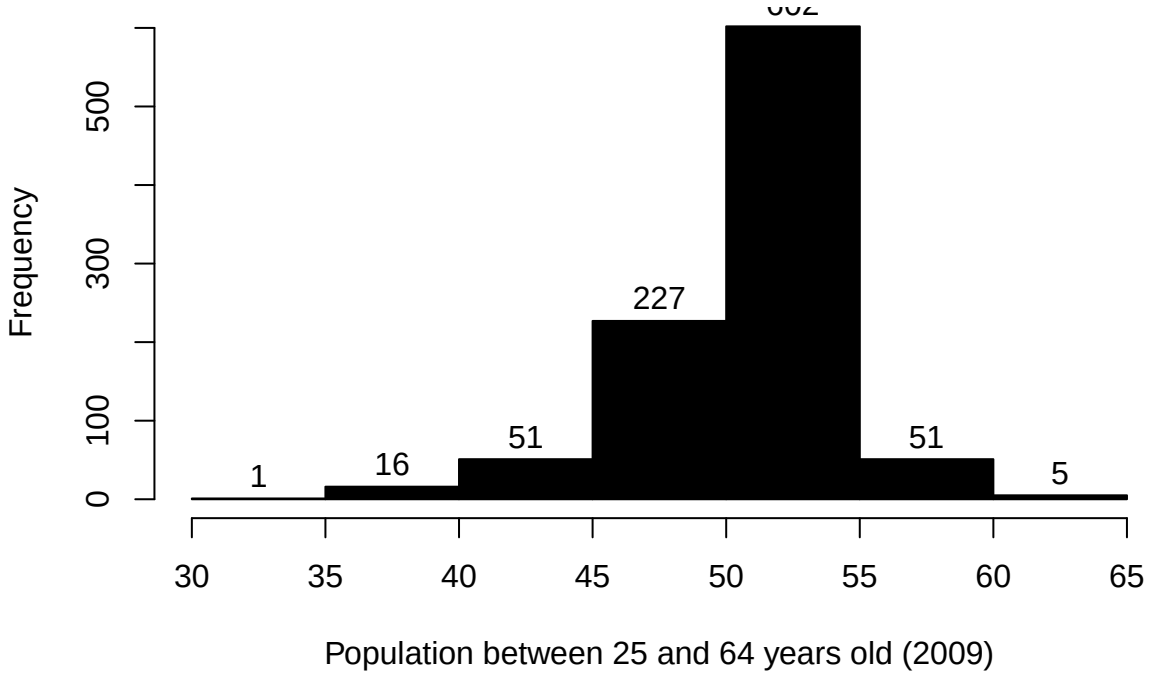
Table 3: Table of Descrpitive Statistics of Data on Income Inequality in Urban Areas, 2019

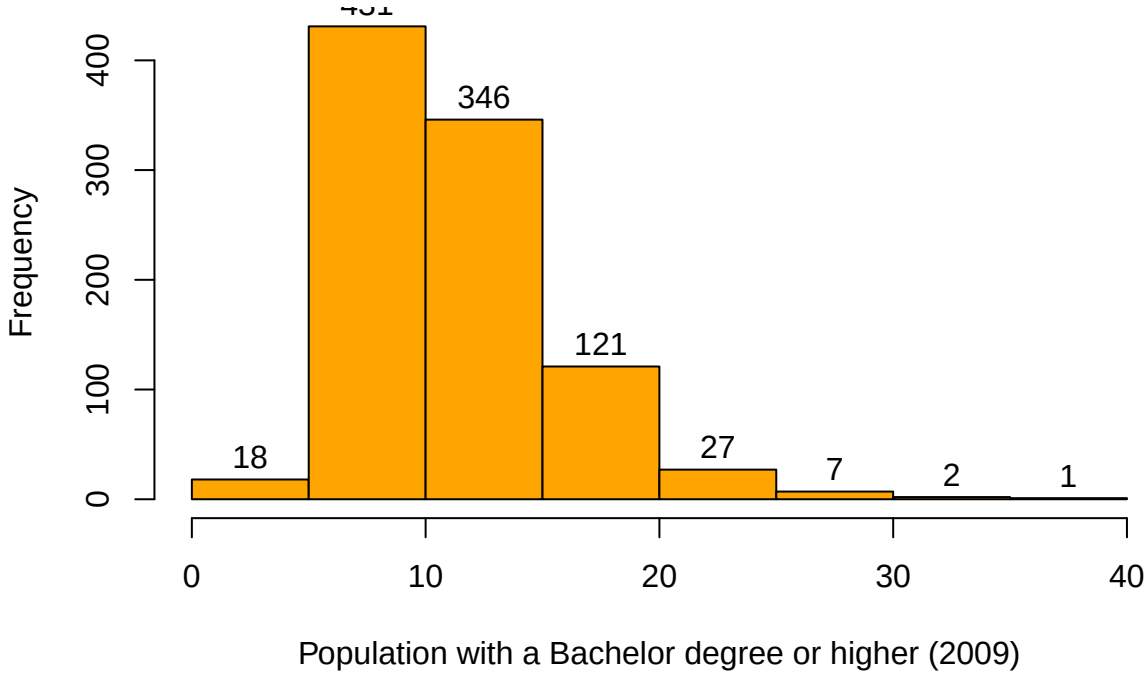
Statistic	Mean	Median	Min	Max	Pctl(25)	Pctl(75)	St. Dev.
Year	2,019.000	2,019	2,019	2,019	2,019	2,019	0.000
MSACode	29,820.330	29,800	10,100	49,820	20,030	39,640	11,383.78
Income share of the top 5%	21.130	21.040	14.510	31.040	19.593	22.518	2.353
Gini Index	45.353	45	38	59	43	47	3.058
Labor Force Participation Rate	59.609	60.150	22.500	79.100	55.600	64.300	6.923
Unemployment Rate	5.641	5.300	1.400	24.300	4.200	6.500	2.305
Number of households	122,375.500	29,376.5	3,548	6,997,001	15,317.2	71,361.8	391,641.90
Households with Income in the past 12 months below poverty level	15.527	14.280	4.520	61.550	11.610	17.925	6.436
Household with public assistance income	2.433	2.085	0.210	25.310	1.532	2.908	1.709
Total Population	330,282.300	75,436	12,813	19,294,236	39,463	182,121.2	1,093,044.0
Per capita income in the past 12 months (in 2009 inflation-adjusted dollars)	28,096.890	27,466	7,308	60,746	24,289	31,190.8	6,162.138
Population 16 Years and Older	80.162	80.315	68.340	93.590	78.625	81.940	2.880
Population between 25 and 64 years old	67.469	67.955	47.500	89.740	65.333	70.138	4.606
Population with a High school diploma	18.264	18.055	5.680	33.040	15.250	20.950	4.557
Population with some college degree or associate's degree	6.239	6.115	1.980	14.020	5.102	7.280	1.597
Population with a Bachelor's degree or higher	16.224	14.940	4.560	47.990	11.925	19.565	5.995
Male	49.732	49.450	46.080	65.970	48.790	50.310	1.706
Female	50.268	50.550	34.030	53.920	49.690	51.210	1.706
White alone (population)	81.592	86.380	14.390	97.770	75.353	92.230	14.395
Black or African American alone (population)	9.276	3.775	0.000	77.620	1.385	11.518	12.703
American Indian and Alaska Native alone (population)	1.381	0.380	0.000	75.900	0.220	0.860	4.465
Asian alone (population)	1.919	1.025	0.000	42.660	0.610	2.047	3.159
Hispanic or Latino (population)	13.493	6.225	0.070	99.750	3.043	13.958	19.001
Married (population)	39.807	40.080	22.250	58.120	37.452	42.580	4.257
Not Married (population)	40.356	40.080	27.950	57.560	37.130	42.978	4.782

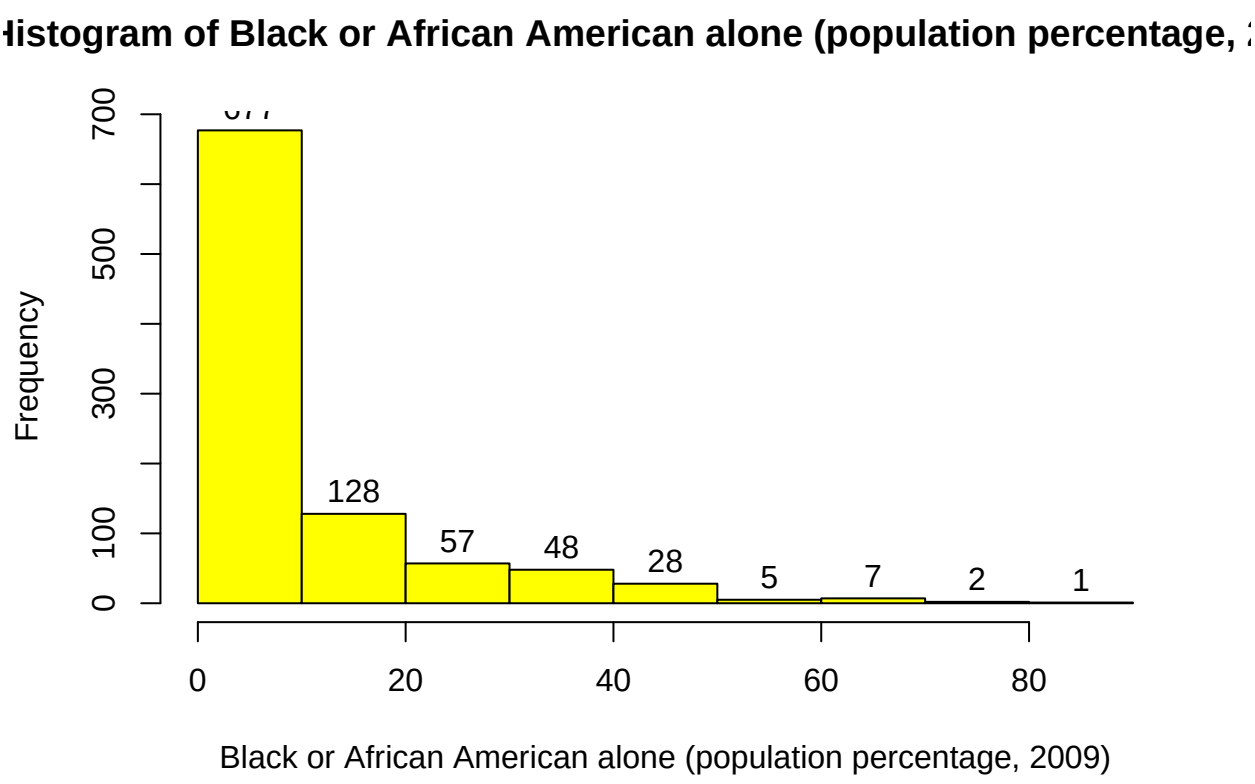

Histogram of income share of top 5%
Frequency
0
50
100
200
300
27
156
329
293
111
24
5
5
2
15
20
25
30
Income share of top 5% (2009)

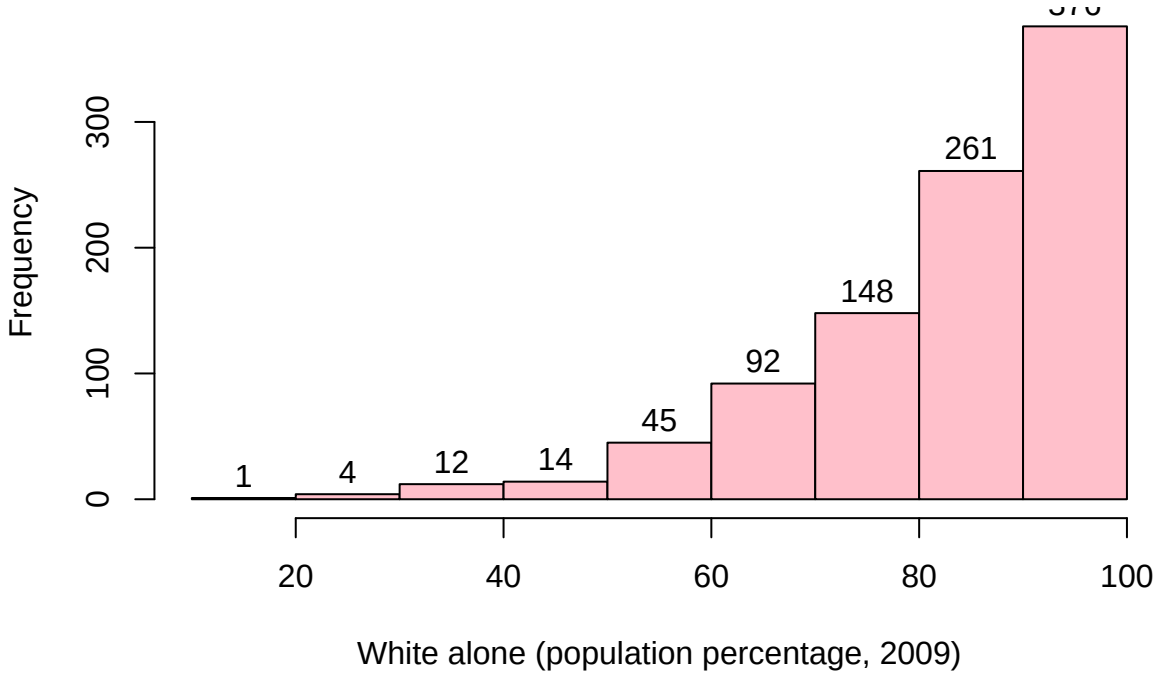

Histogram of Gini index (2009)
Frequency
0
50
100
150
200
250
6
17
93
190
211
117
33
25
7
3
35
40
45
50
55
Gini index (2009)

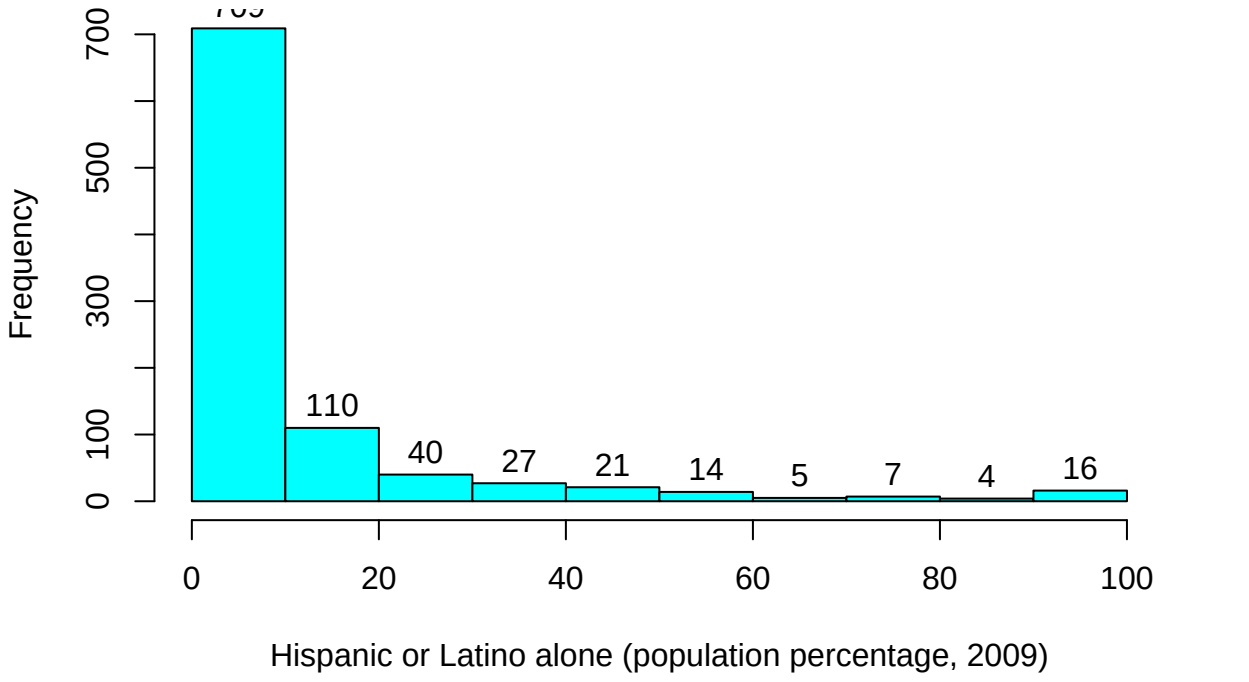

Histogram of unemployment rate (2009)
Frequency
0
50
100
150
200
250
4
10
87
200
252
228
94
27
11
7
3
20
30
40
50
60
70
Unemployment rate (2009)

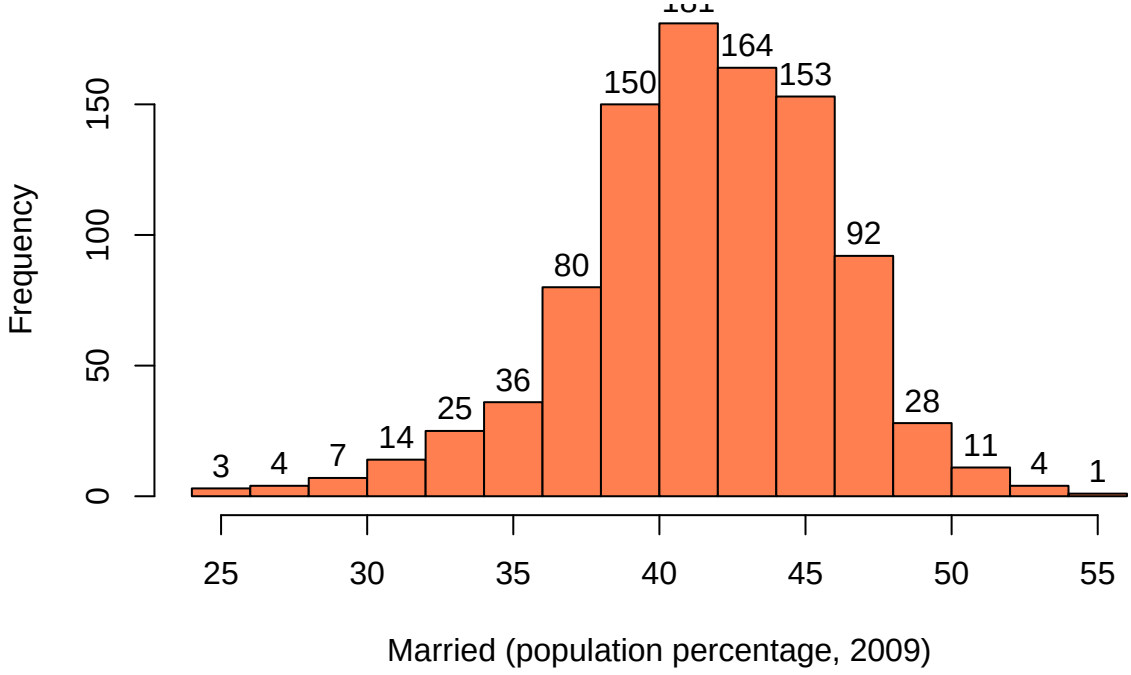

Histogram of population between 25 and 64 years old (2009)
Frequency
0
100
300
500
1
16
51
227
51
5
30
35
40
45
50
55
60
65
Population between 25 and 64 years old (2009)

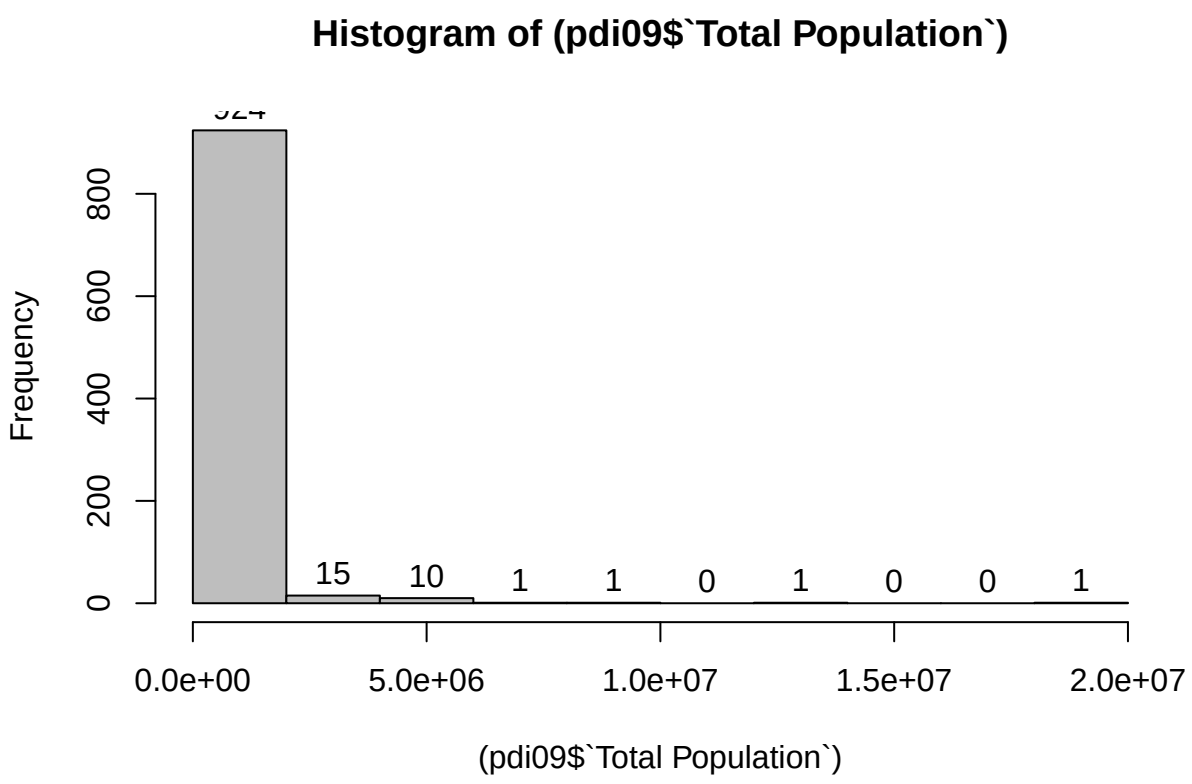

Histogram of population with a Bachelor degree or higher (2009)
Frequency
0
100
200
300
400
18
346
121
27
7
2
1
0
10
20
30
40
Population with a Bachelor degree or higher (2009)


Histogram of Black or African American alone (population percentage,
Frequency
0
100
300
500
700
128
57
48
28
5
7
2
1
0
20
40
60
80
Black or African American alone (population percentage, 2009)

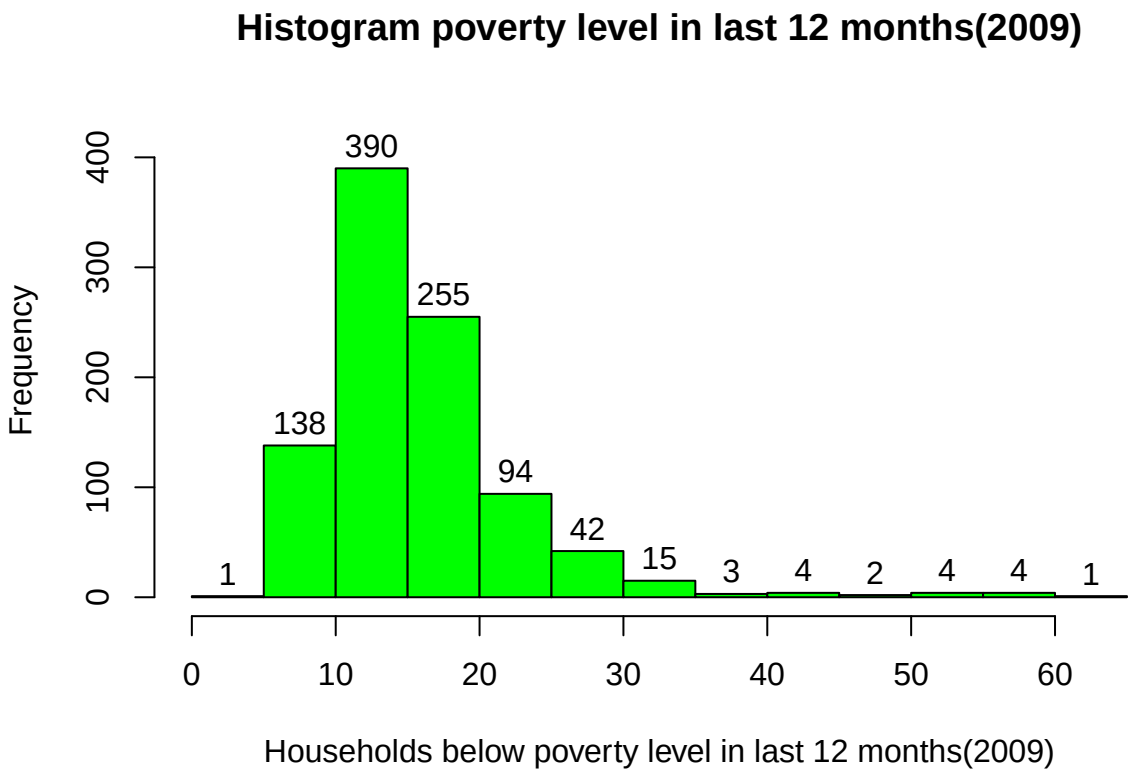

Histogram of White alone (population percentage, 2009)
Frequency
0
100
200
300
1
4
12
14
45
92
148
261
20
40
60
80
100
White alone (population percentage, 2009)

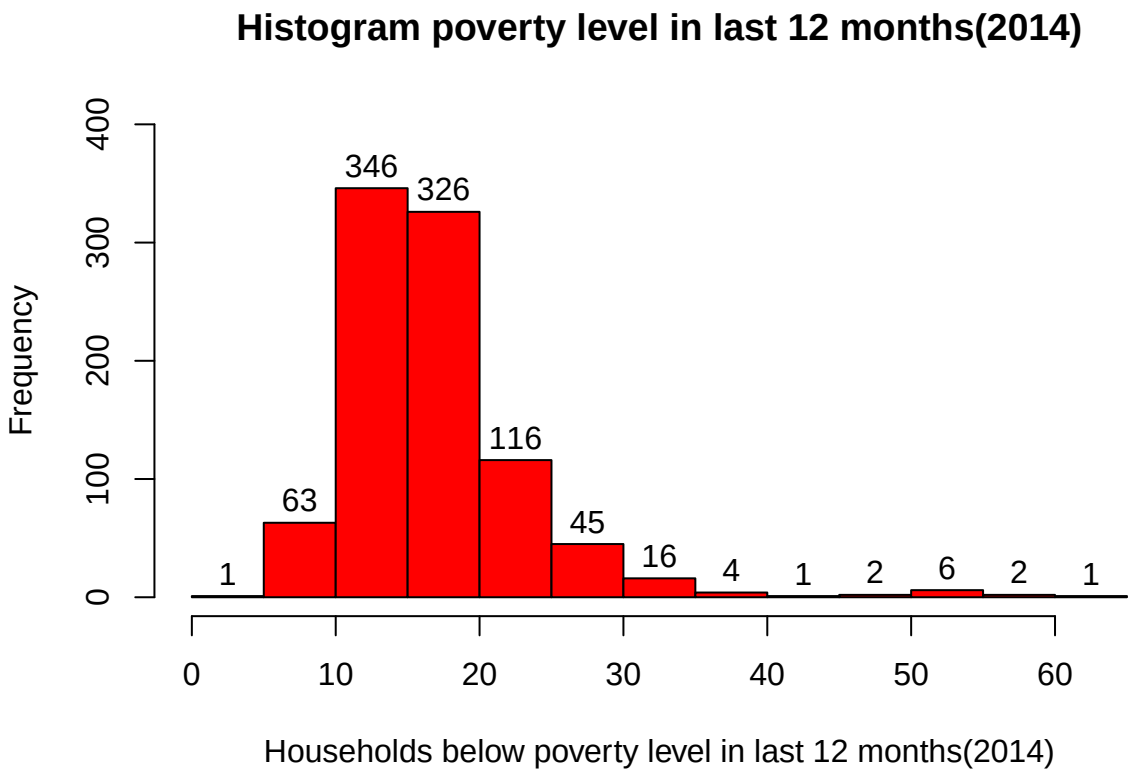

Histogram of Hispanic or Latino alone (population percentage, 2009
Frequency
0
100
300
500
700
110
40
27
21
14
5
7
4
16
0
20
40
60
80
100
Hispanic or Latino alone (population percentage, 2009)

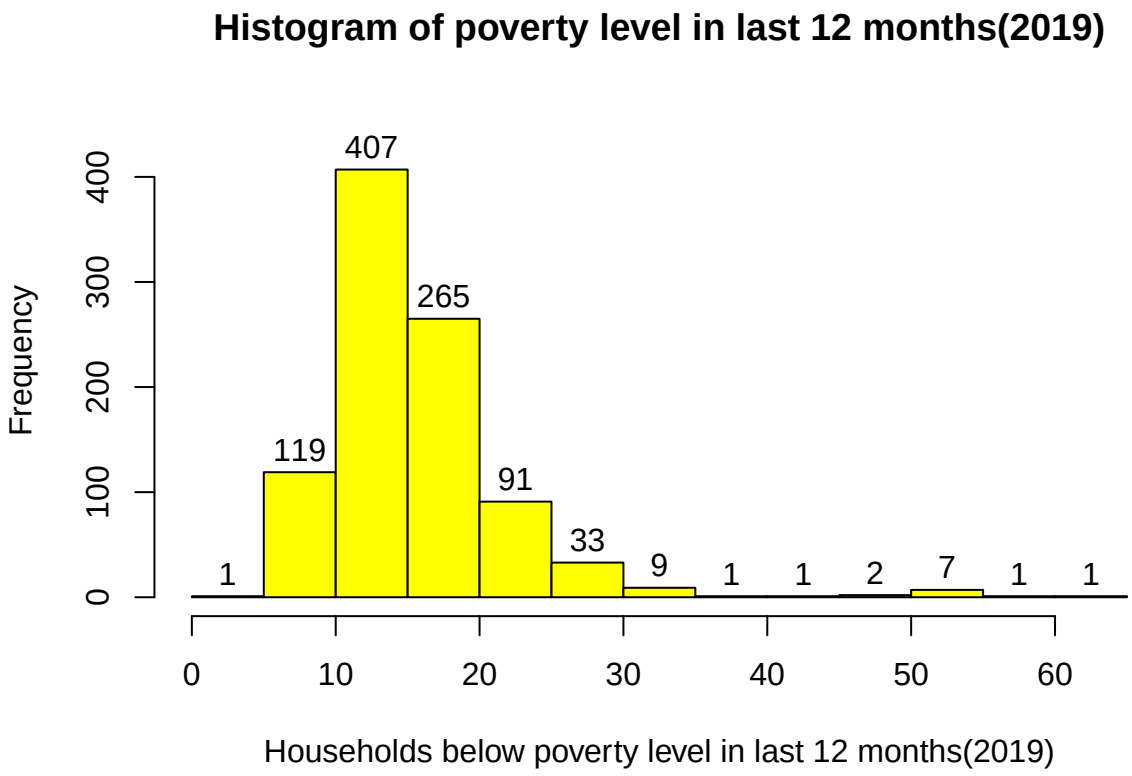

Histogram of Married (population percentage, 2009)
Frequency
0
50
100
150
3
4
7
14
25
36
80
150
164
153
92
28
11
4
1
25
30
35
40
45
50
55
Married (population percentage, 2009)



1.Poverty and wealth inequality

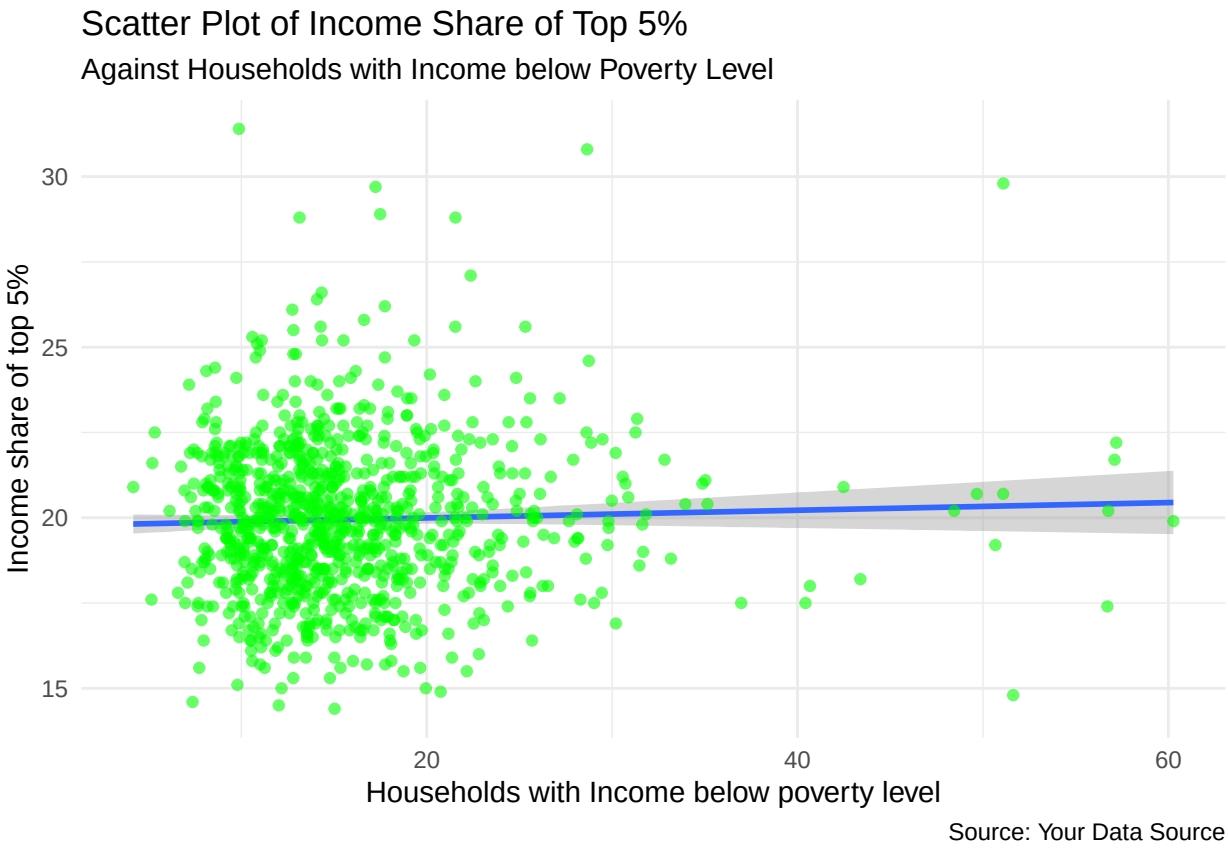

Histogram poverty level in last 12 months(2009)
Frequency
0
100
200
300
400
1
138
390
255
94
42
15
3
4
2
4
4
1
0
10
20
30
40
50
60
Households below poverty level in last 12 months(2009)

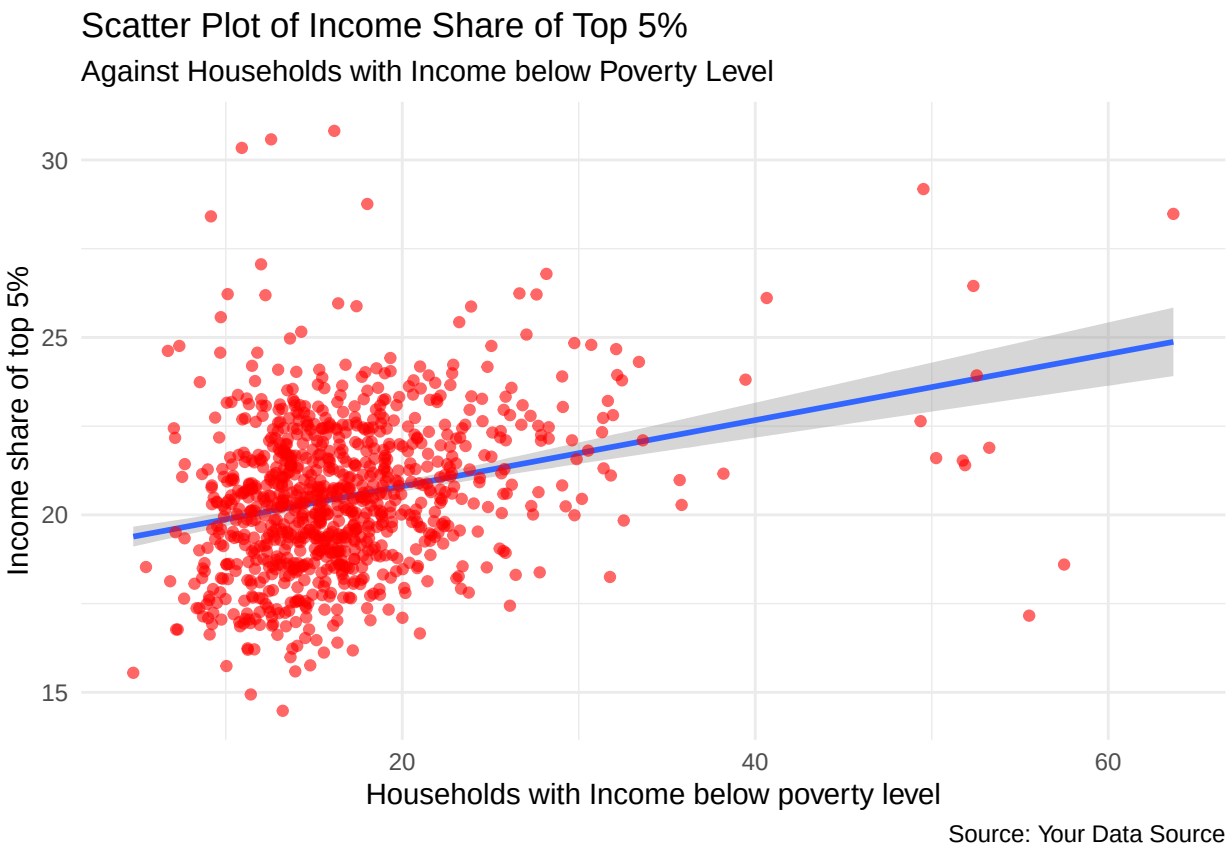

Histogram poverty level in last 12 months(2014)
Frequency
0
100
200
300
400
1
63
346
326
116
45
16
4
1
2
6
2
1
0
10
20
30
40
50
60
Households below poverty level in last 12 months(2014)

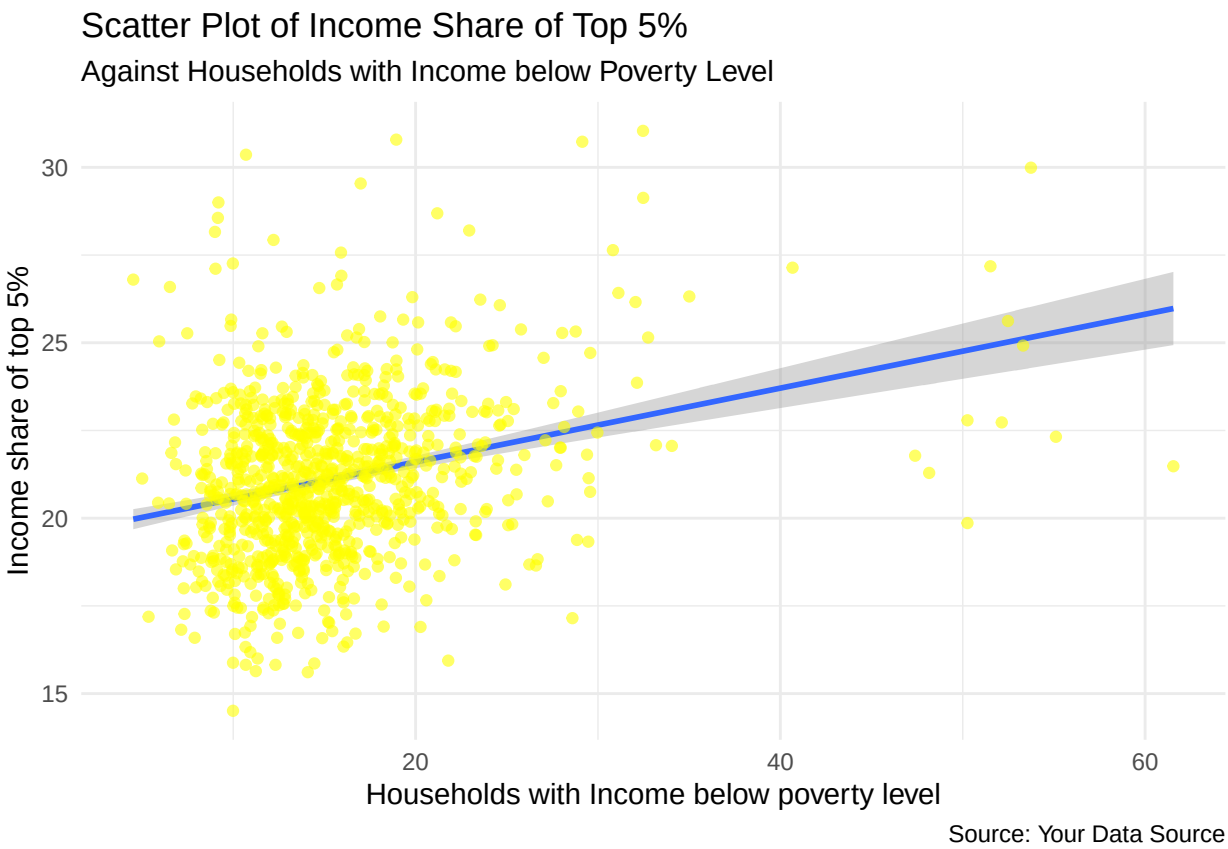



```
## Warning: Removed 1 rows containing non-finite values (‘stat_smooth()‘).
```

```
## Warning: Removed 1 rows containing missing values (‘geom_point()‘).
```


Scatter Plot of Income Share of Top 5%
Against Households with Income below Poverty Level
Income share of top 5%
30
25
20
15
20
40
60
Households with Income below poverty level
Source: Your Data Source


Scatter Plot of Income Share of Top 5%
Against Households with Income below Poverty Level
Income share of top 5%
30
25
20
15
20
40
60
Households with Income below poverty level
Source: Your Data Source



An examination of the data: The 2009 scatter plot shows almost no correlation between Households with income in the past 12 months below poverty level and the income share of the top 5%. However, the plots for 2014 and 2019 do show some degree of correlation. To get more information on this, we will need to run a regression.

Table 4:

	Dependent variable:
	'Income share of the top 5%'
'Households with Income in the past 12 months below poverty level'	0.011
	p = 0.286
	(0.011)
Constant	19.770***
	(0.182)
Observations	952
R^2	0.001
Adjusted R^2	0.0001
Residual Std. Error	2.256 (df = 950)
F Statistic	1.142 (df = 1; 950)
Note:	*p<0.1; **p<0.05; ***p<0.01

As can be seen from the regression tables, 2009's regression coefficient lacks both statistical significance and

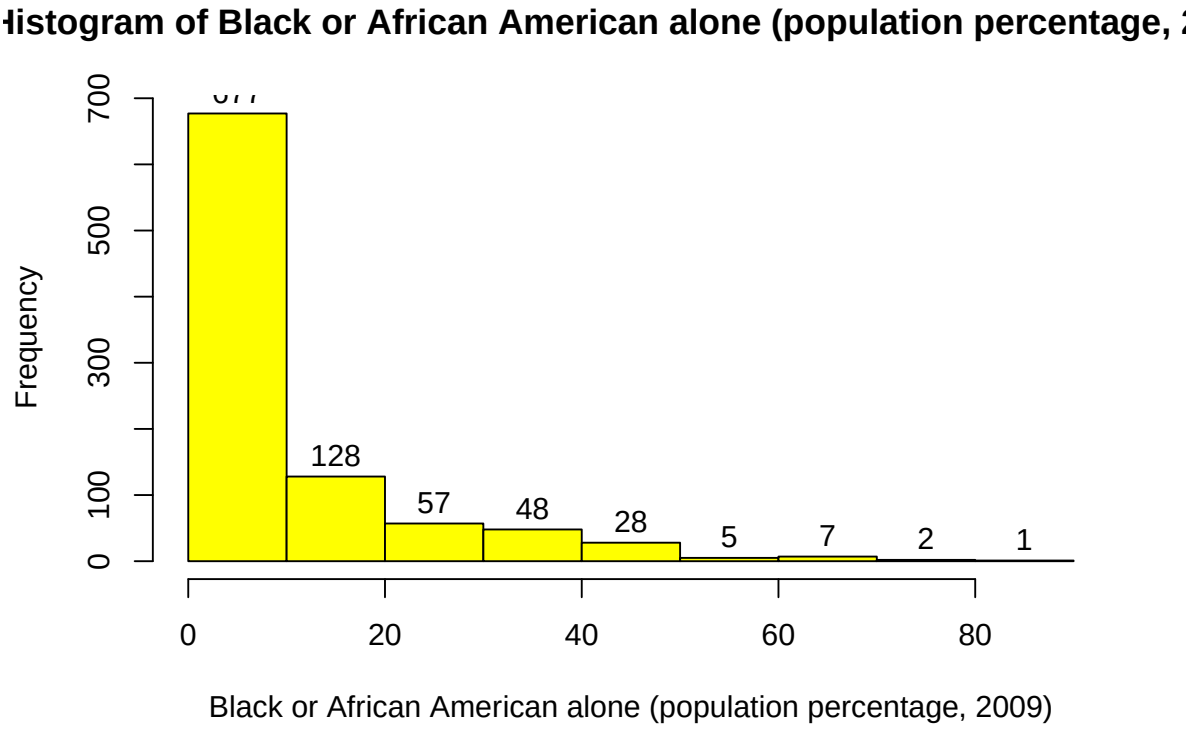
Table 5:

	Dependent variable:
	'Income share of the top 5%'
'Households with Income in the past 12 months below poverty level'	0.093***
	p = 0.000
	(0.010)
Constant	18.944***
	(0.188)
Observations	929
R^2	0.080
Adjusted R^2	0.079
Residual Std. Error	2.080 (df = 927)
F Statistic	80.361*** (df = 1; 927)
Note:	*p<0.1; **p<0.05; ***p<0.01

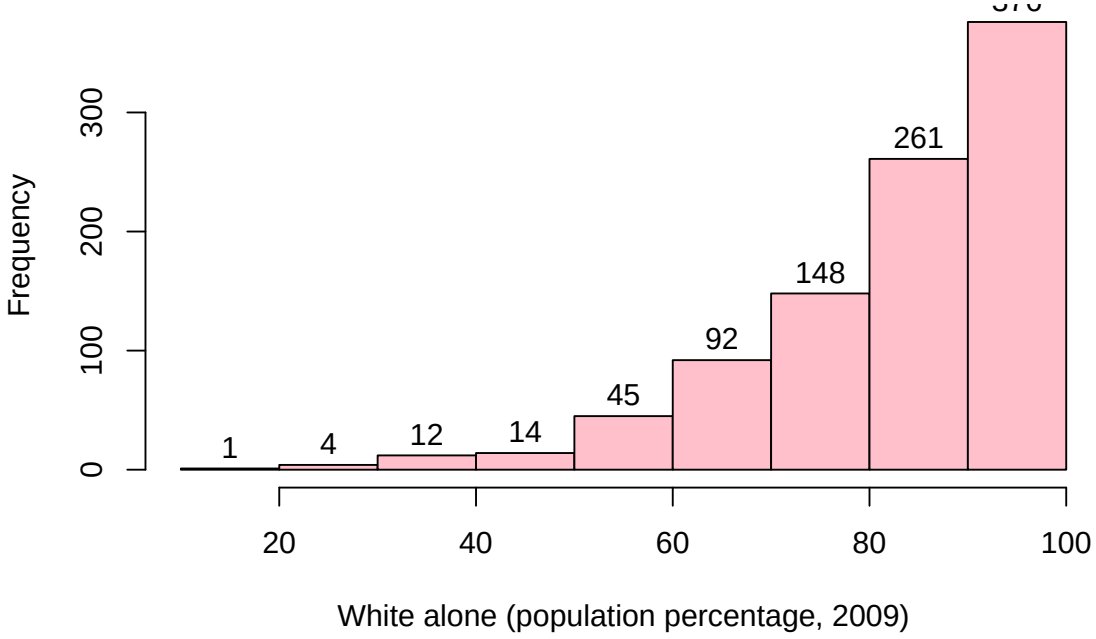
Table 6:

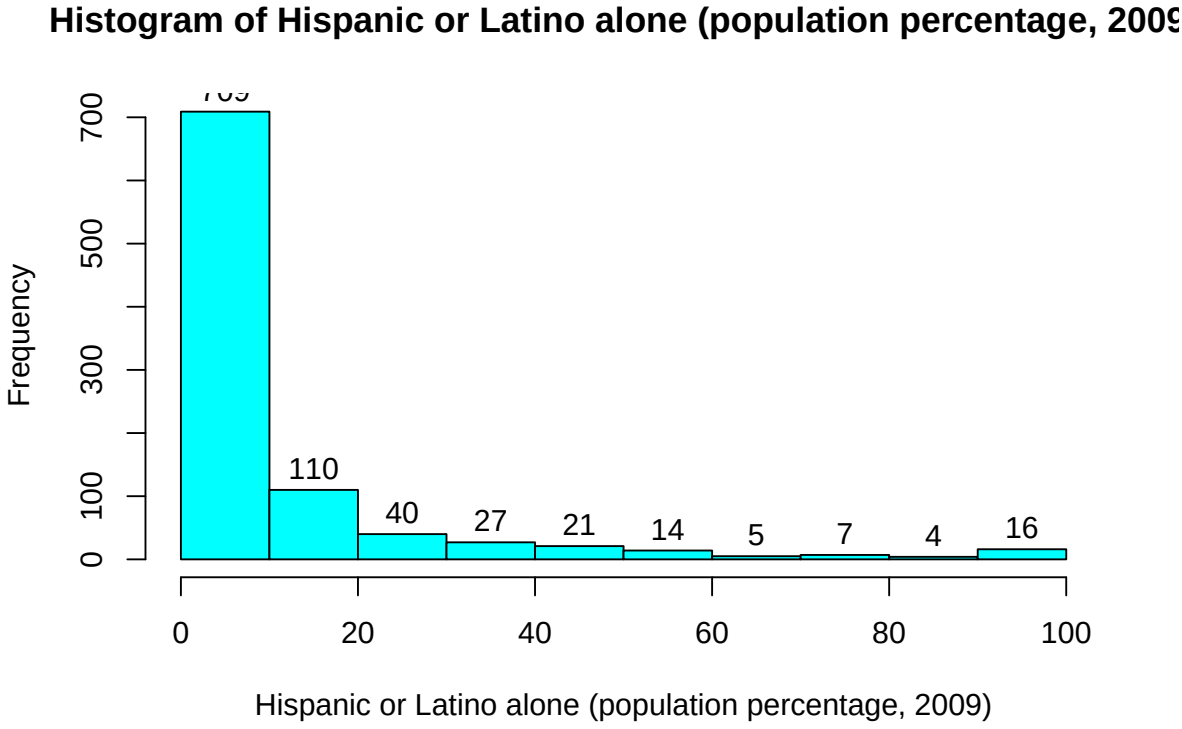
	Dependent variable:
	'Income share of the top 5%'
'Households with Income in the past 12 months below poverty level'	0.105***
	p = 0.000
	(0.011)
Constant	19.495***
	(0.192)
Observations	938
R^2	0.083
Adjusted R^2	0.082
Residual Std. Error	2.254 (df = 936)
F Statistic	84.686*** (df = 1; 936)
Note:	*p<0.1; **p<0.05; ***p<0.01

practical significance. It's p value is 0.286 and a β of 0.011 means that an increase in one percentage point in households below the poverty level is correlated to a 0.011 percentage point increase in the income share of the top 5%. An increase of 90 percentage points would be needed for a one On the other hand, the coefficients for 2014 and 2019 are statistically significant with three stars. A change of one σ_x in the 2014 poverty level correlates with an increase of 0.611 percentage points in income share by the top 5%. That is about a 0.2821 σ_y change. For 2019 the results are similar with a change of σ_x correlating to a 0.2872 σ_y change. These would seem to be practically significant changes. Of course, there are almost certainly cases of OVB here as these are single variable regressions. Once we have finished discussing the variables by themselves, we will try looking at them together to get a more accurate picture of the situation. It is possible that the reason the 2009 numbers don't display any significant correlation is because there are other variables at play that essentially cancel out the effect of the correlation and perhaps the effect would become viable if we could isolate the variables. It is also possible that poverty only correlates with income inequality insofar as it correlates with other things, like education level, and isolating the variables would show almost no correlation. It is important to establish if a link exists between income inequality and poverty in part to establish what exactly income inequality is. There are two sides in the debate about income inequality. One side argues that income inequality is a product of society as a whole getting wealthier and the gains are just not distributed equally, although everyone does gain. In this understanding, income inequality is not a major problem, even if it is not ideal. The other side argues that income inequality is a result sectors of society getting left behind, and not gaining from the advances that society makes. In this understanding, income inequality is a major problem, and should be correlated with high poverty levels. Examining the correlation between poverty and inequality is therefore important in drafting policy to deal with the issue.



2. Race

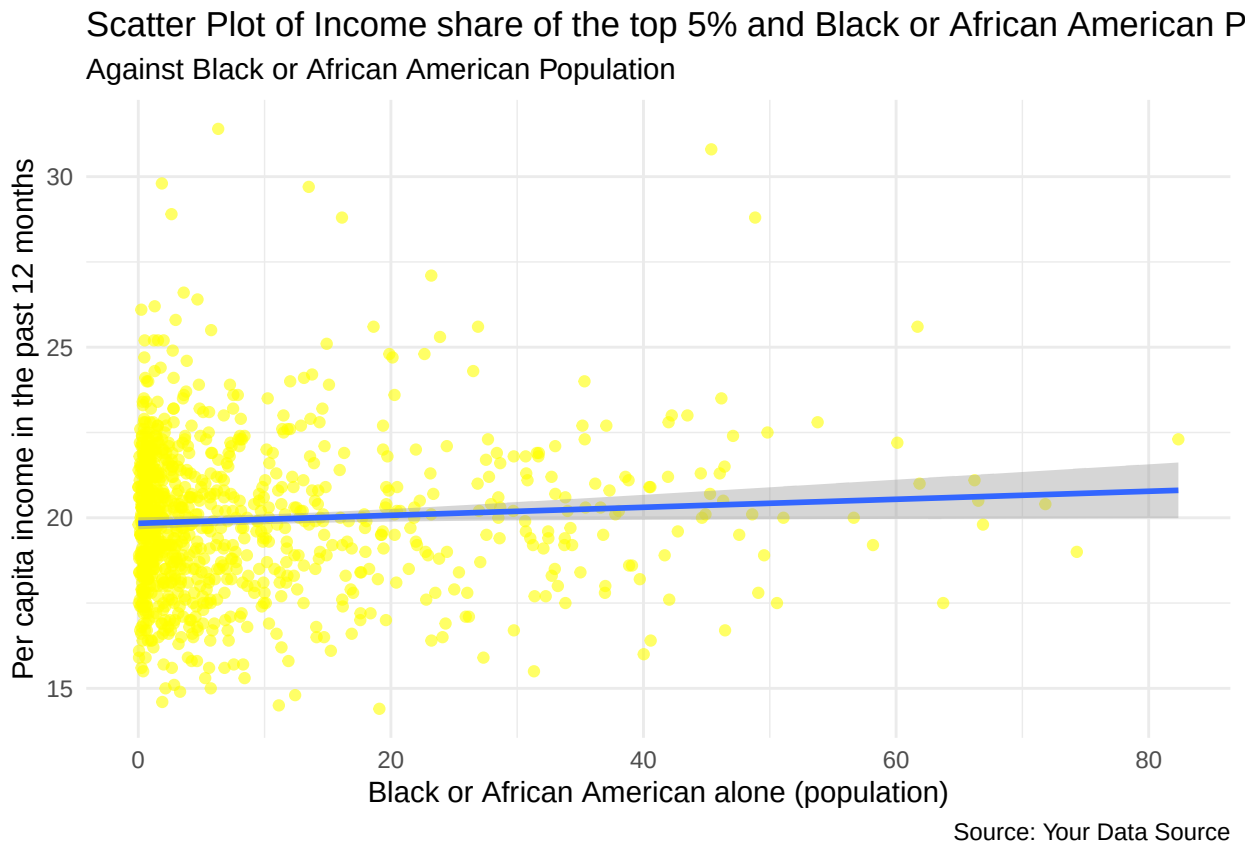





```
## `geom_smooth()` using formula = 'y ~ x'
```

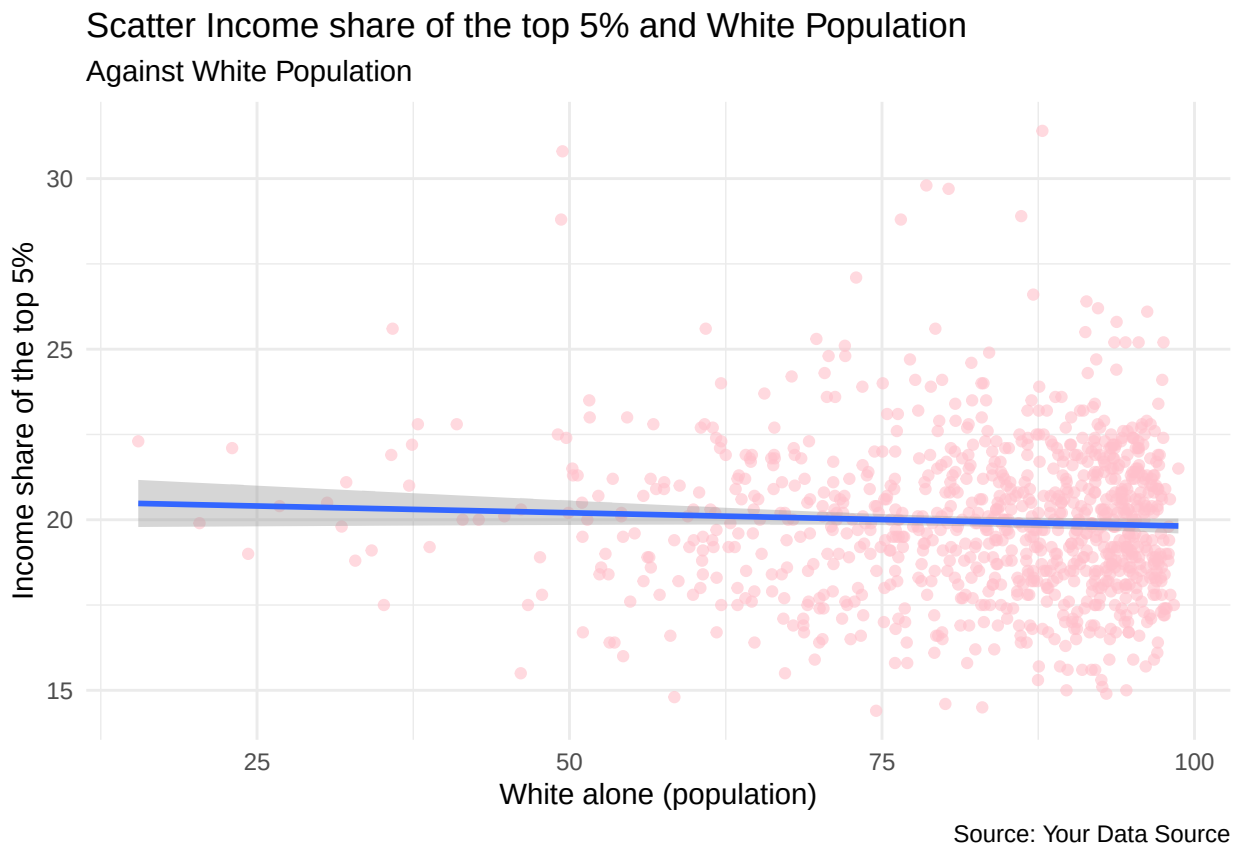
```
## Warning: Removed 1 rows containing non-finite values (`stat_smooth()`).
```

```
## Warning: Removed 1 rows containing missing values (`geom_point()`).
```

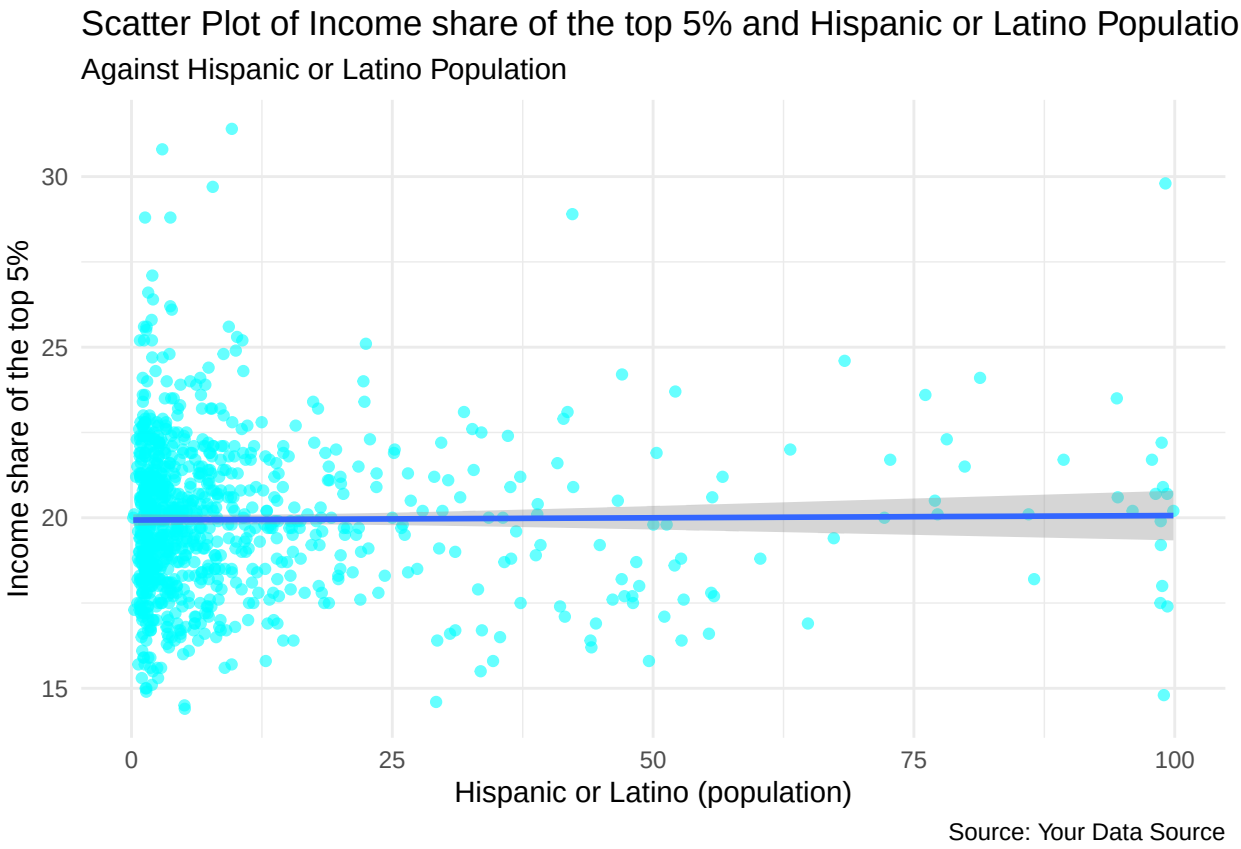


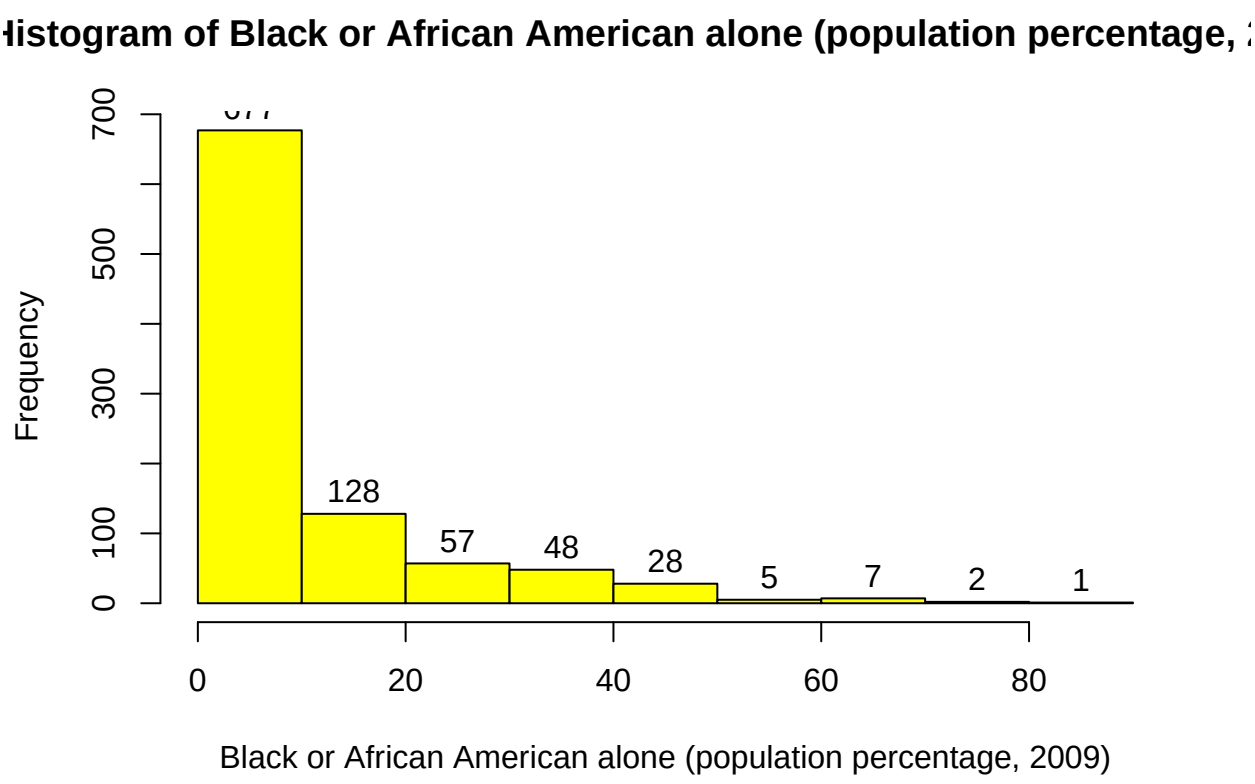

```
## `geom_smooth()` using formula = 'y ~ x'
```

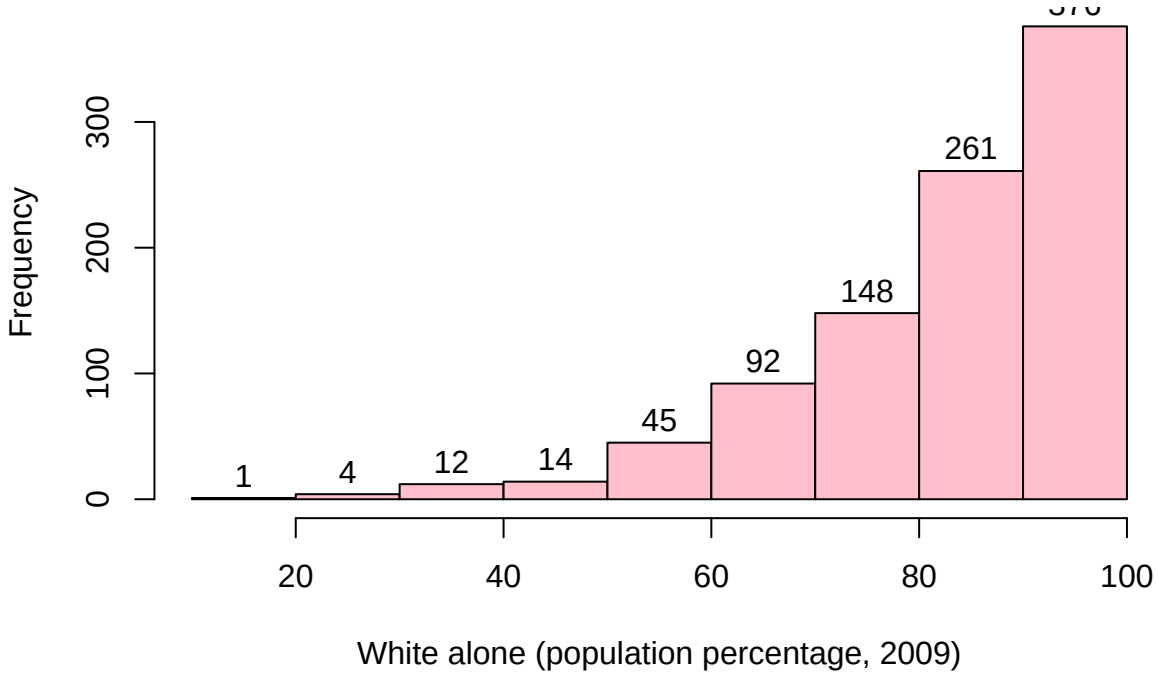
```
## Warning: Removed 1 rows containing non-finite values (`stat_smooth()`).
## Removed 1 rows containing missing values (`geom_point()`).
```

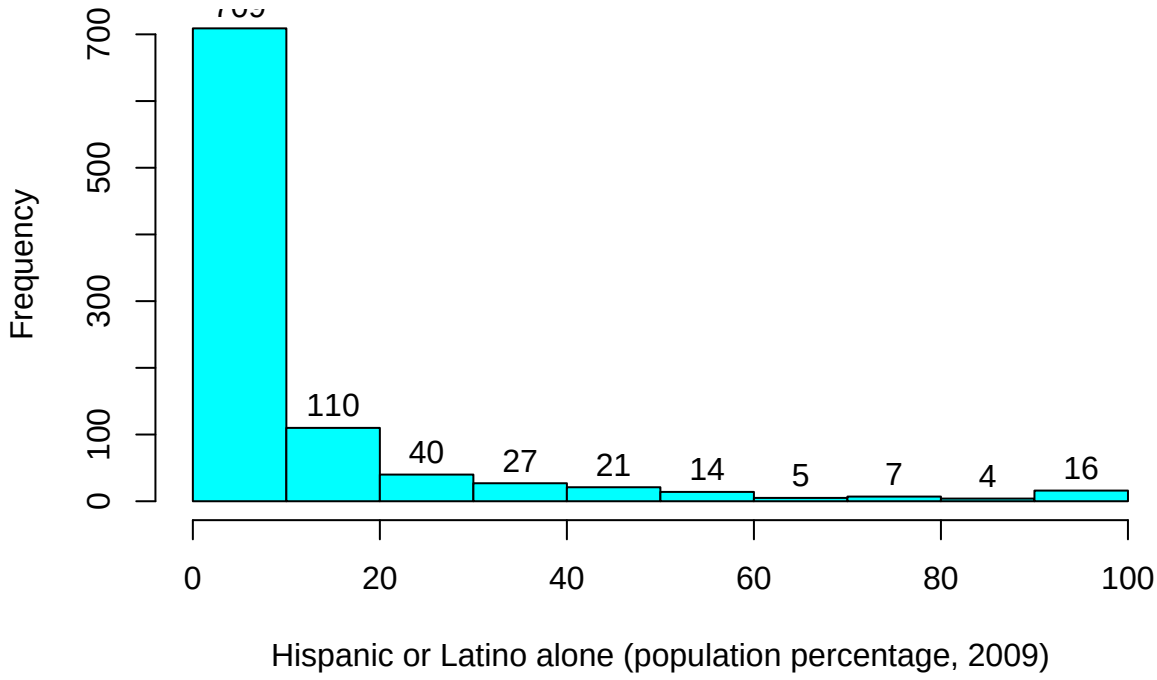



```
## `geom_smooth()` using formula = 'y ~ x'

## Warning: Removed 1 rows containing non-finite values (`stat_smooth()`).
## Removed 1 rows containing missing values (`geom_point()`).
```

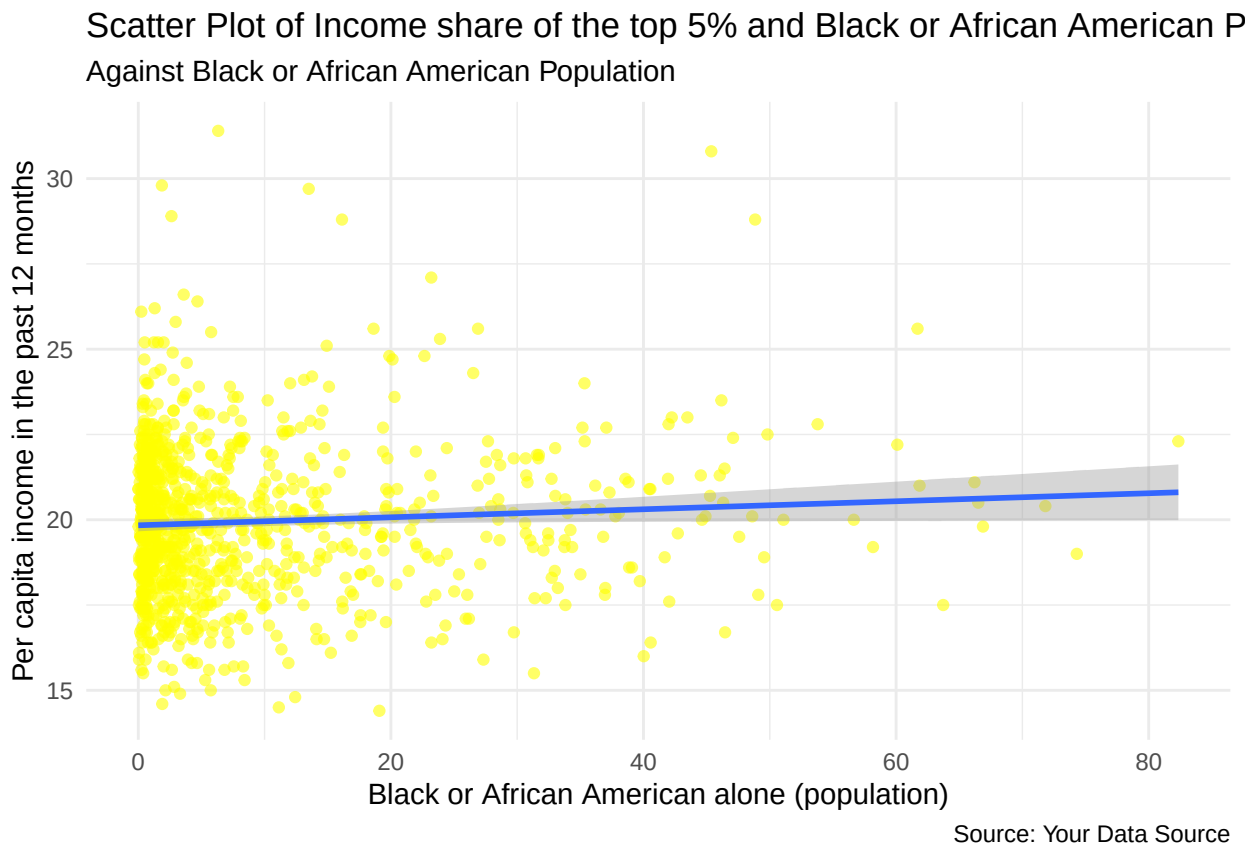

Scatter Plot of Income share of the top 5% and Hispanic or Latino Populatio
Against Hispanic or Latino Population
Income share of the top 5%
30
25
20
15
0
25
50
75
100
Hispanic or Latino (population)
Source: Your Data Source


Histogram of Black or African American alone (population percentage,
Frequency
0
100
300
500
700
128
57
48
28
5
7
2
1
0
20
40
60
80
Black or African American alone (population percentage, 2009)

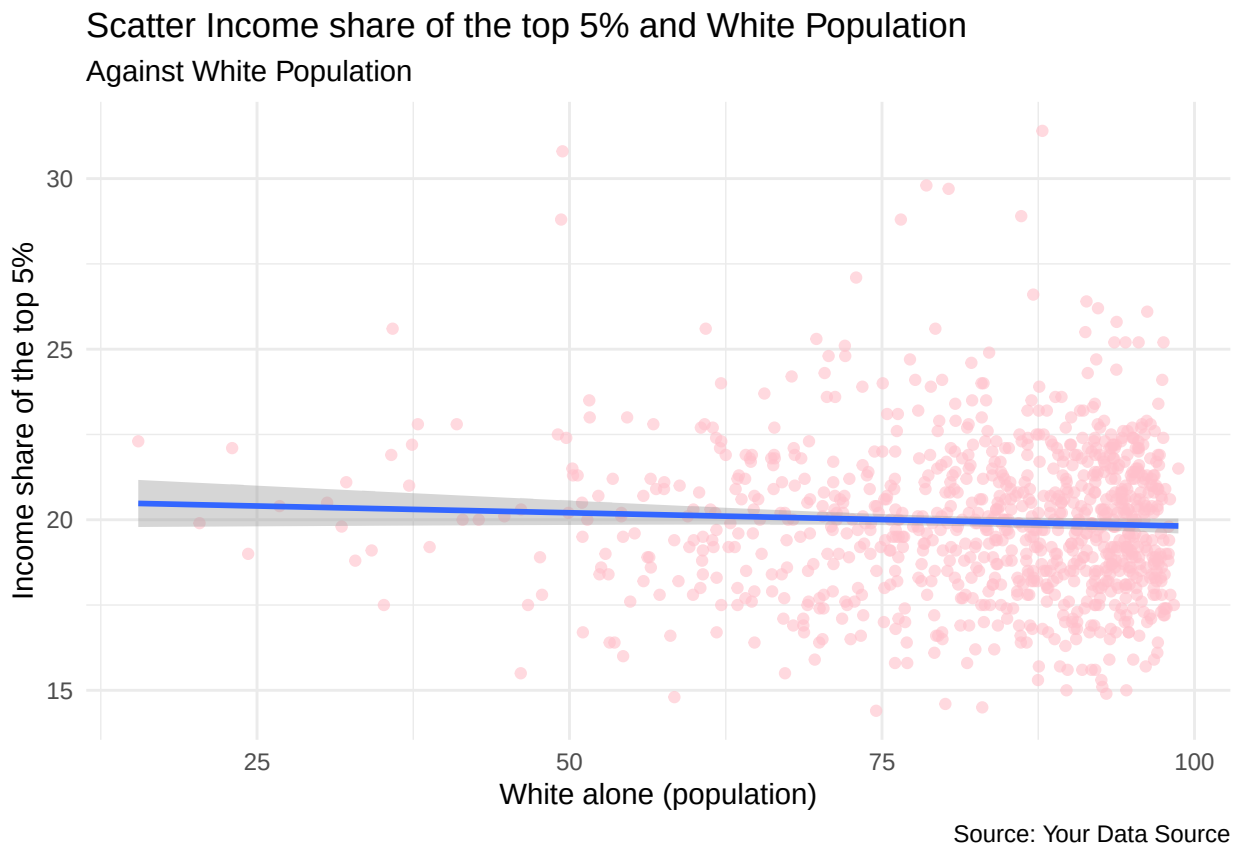

Histogram of White alone (population percentage, 2009)
Frequency
0
100
200
300
1
4
12
14
45
92
148
261
20
40
60
80
100
White alone (population percentage, 2009)



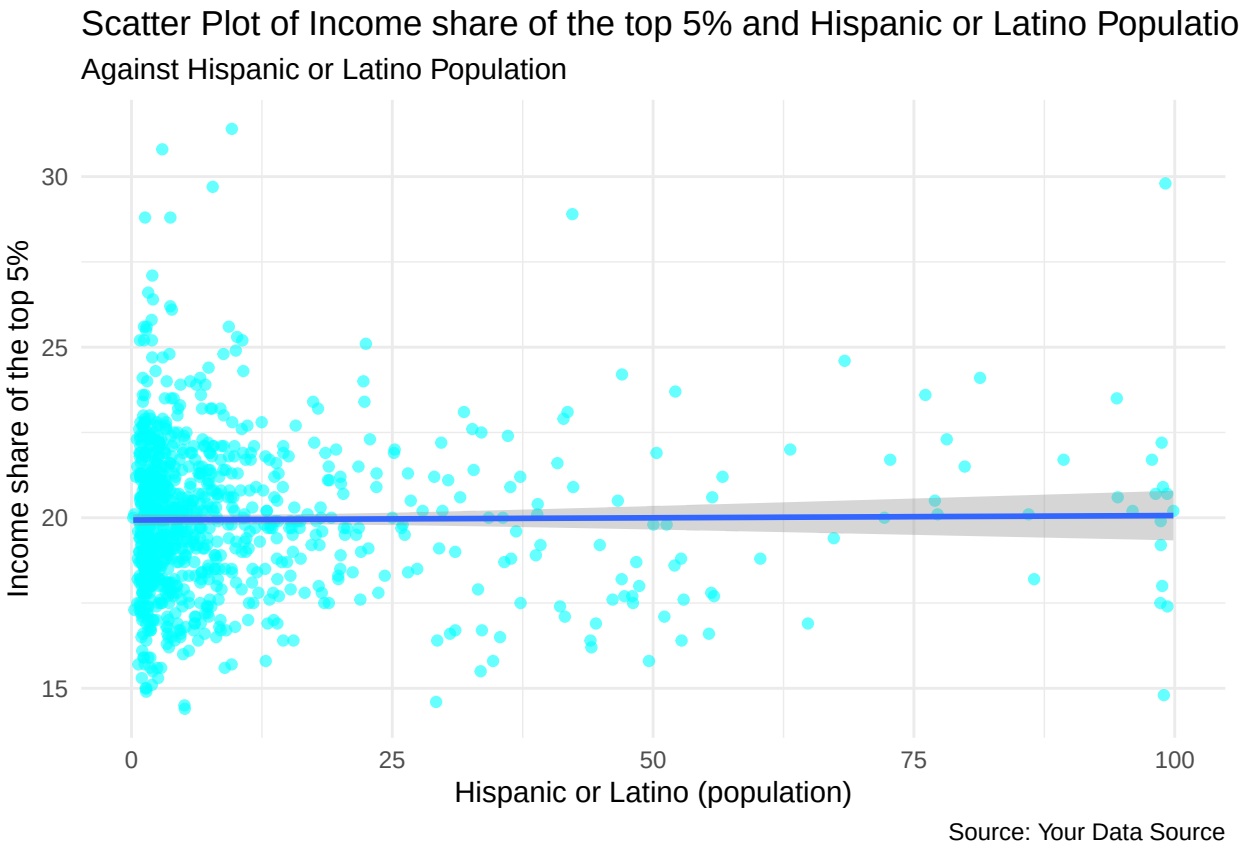

```
## `geom_smooth()` using formula = 'y ~ x'

## Warning: Removed 1 rows containing non-finite values (`stat_smooth()`).
## Removed 1 rows containing missing values (`geom_point()`).
```

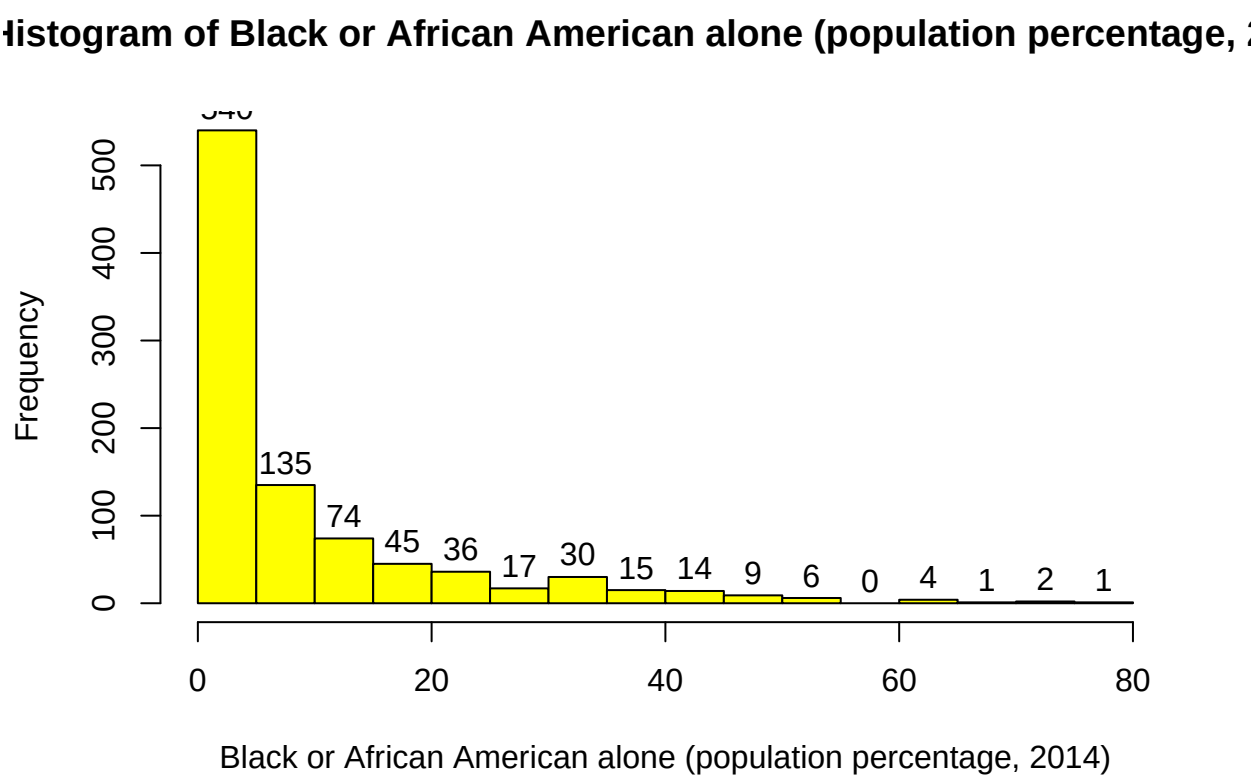


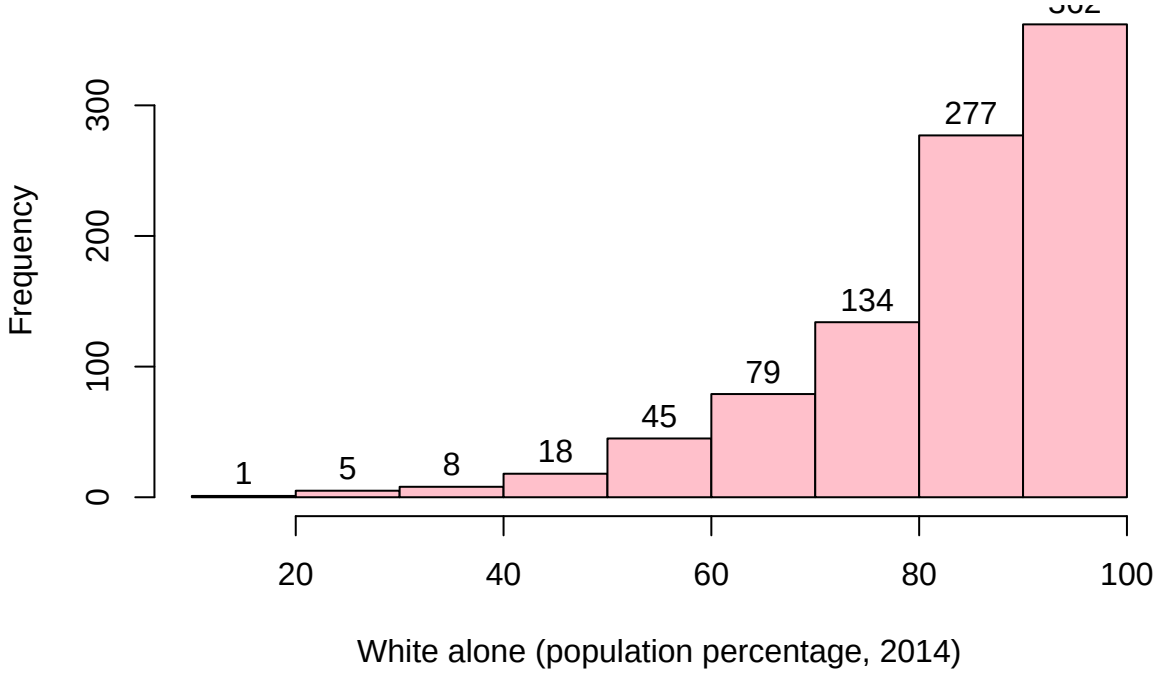

```
## `geom_smooth()` using formula = 'y ~ x'

## Warning: Removed 1 rows containing non-finite values (`stat_smooth()`).
## Removed 1 rows containing missing values (`geom_point()`).
```

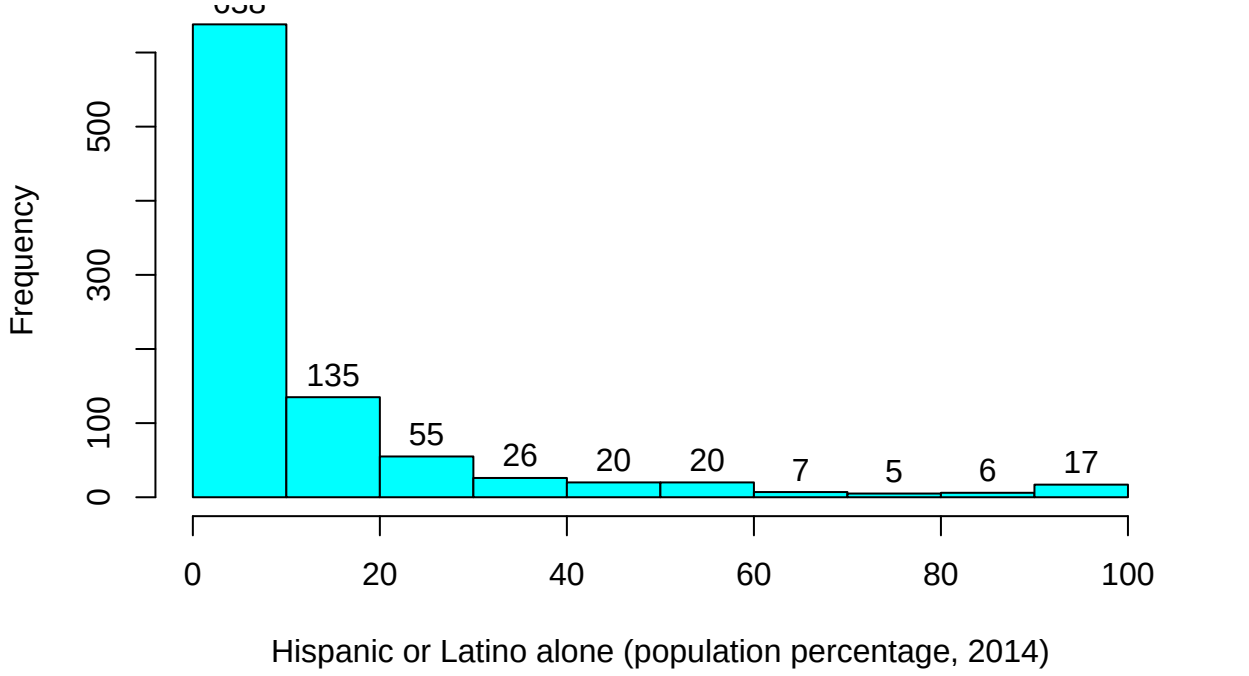



```
## `geom_smooth()` using formula = 'y ~ x'

## Warning: Removed 1 rows containing non-finite values (`stat_smooth()`).
## Removed 1 rows containing missing values (`geom_point()`).
```

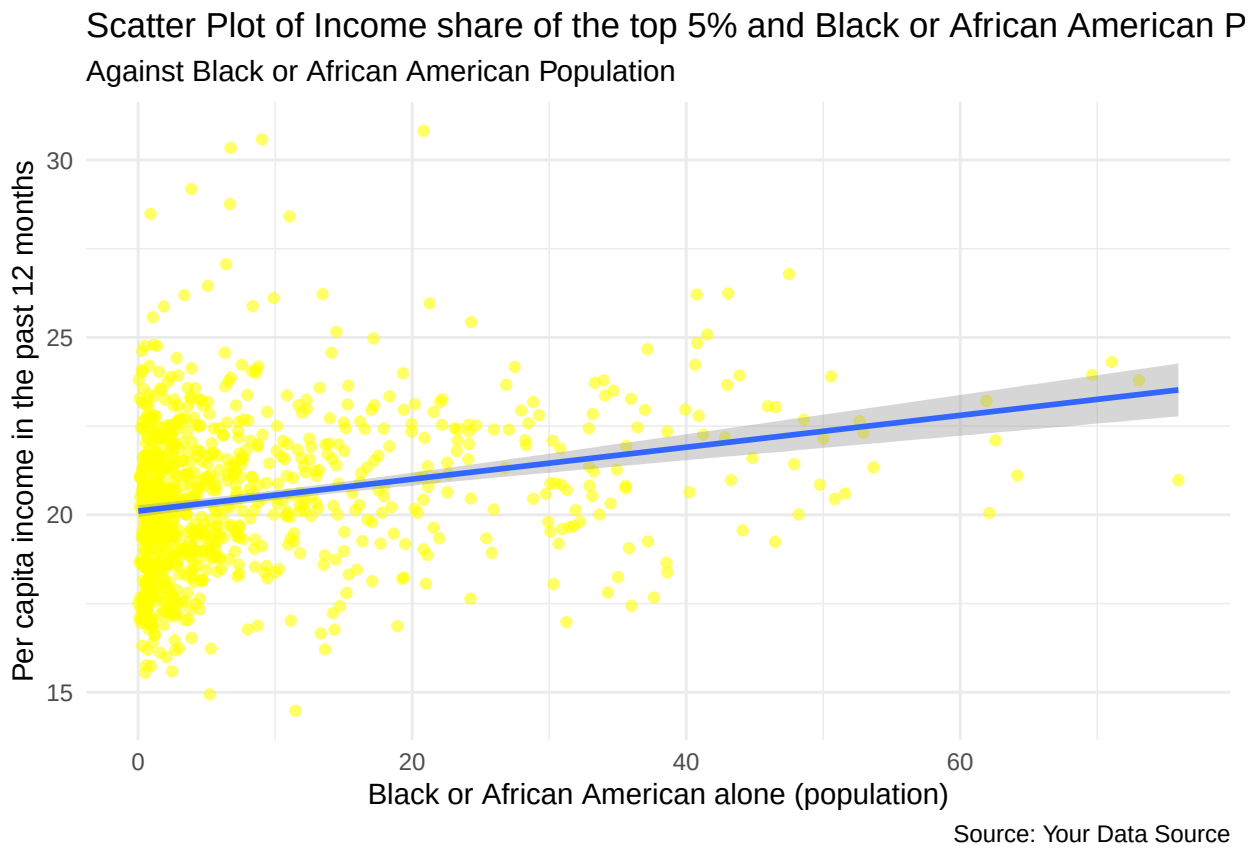

Scatter Plot of Income share of the top 5% and Hispanic or Latino Populatio
Against Hispanic or Latino Population
Income share of the top 5%
30
25
20
15
0
25
50
75
100
Hispanic or Latino (population)
Source: Your Data Source


Histogram of Black or African American alone (population percentage,
Frequency
0
100
200
300
400
500
540
135
74
45
36
17
30
15
14
9
6
0
4
1
2
1
0
20
40
60
80
Black or African American alone (population percentage, 2014)

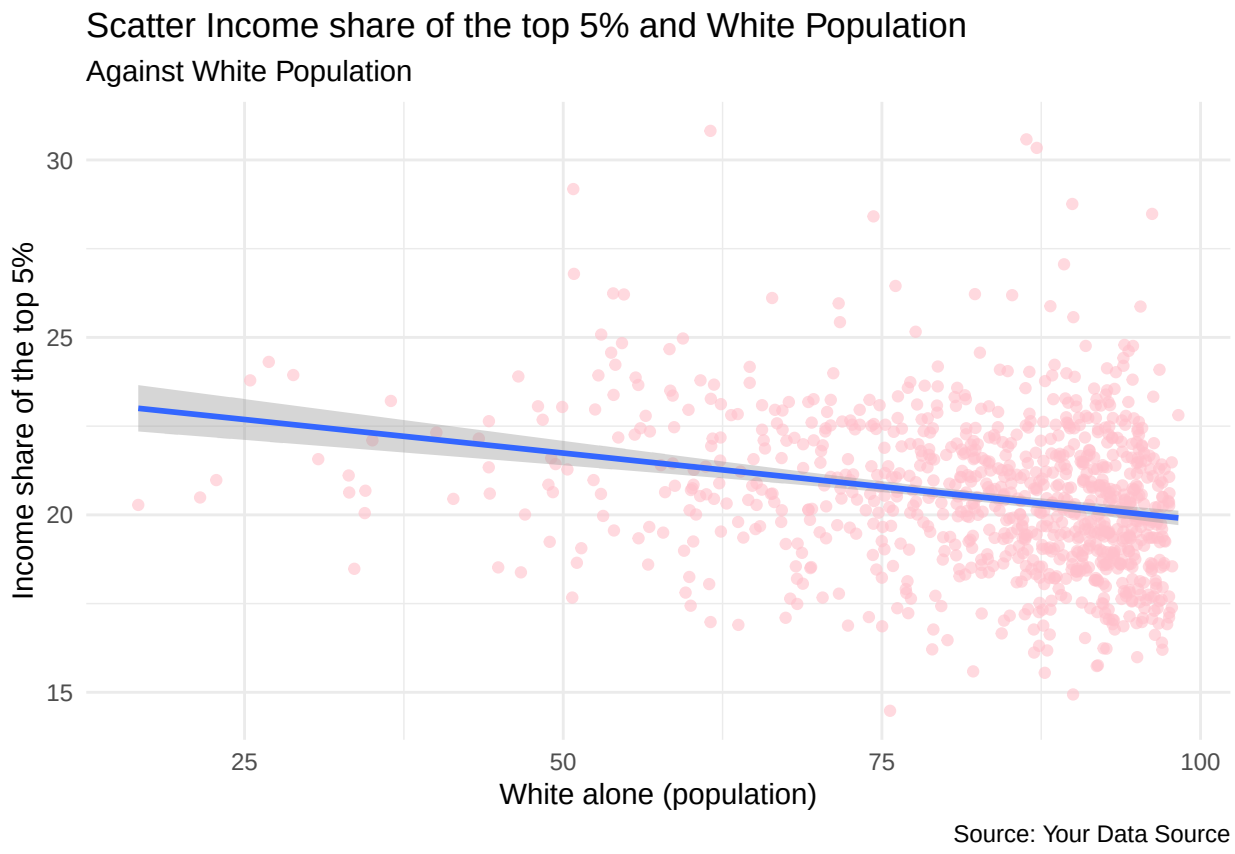

Histogram of White alone (population percentage, 2014)
Frequency
0
100
200
300
1
5
8
18
45
79
134
277
20
40
60
80
100
White alone (population percentage, 2014)



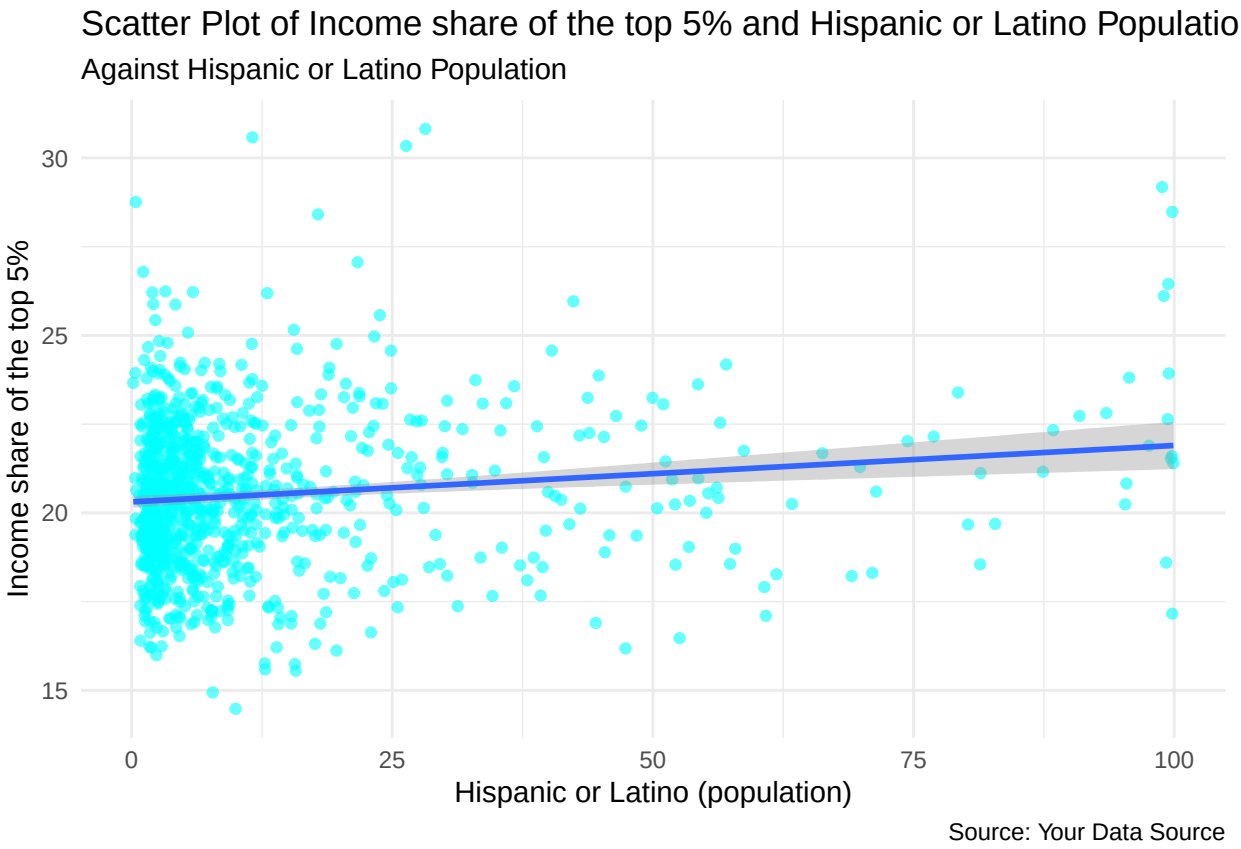

```
## `geom_smooth()` using formula = 'y ~ x'
```

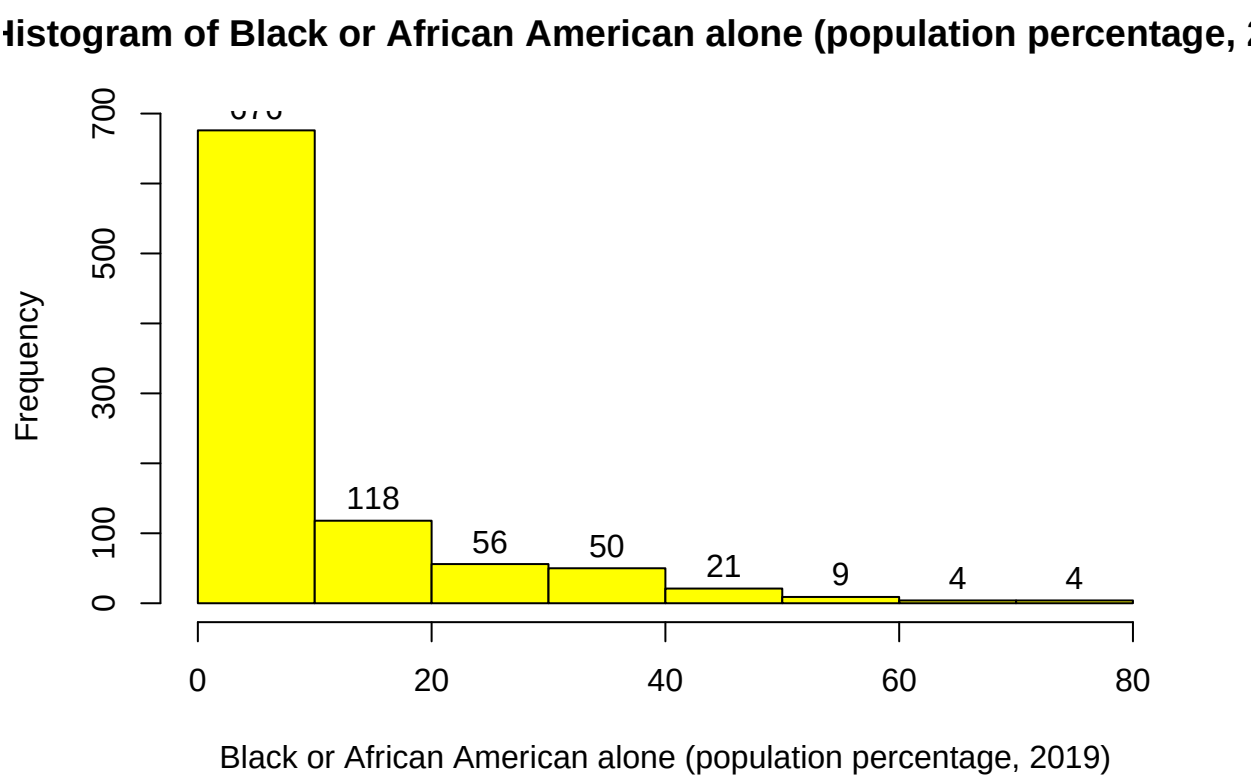


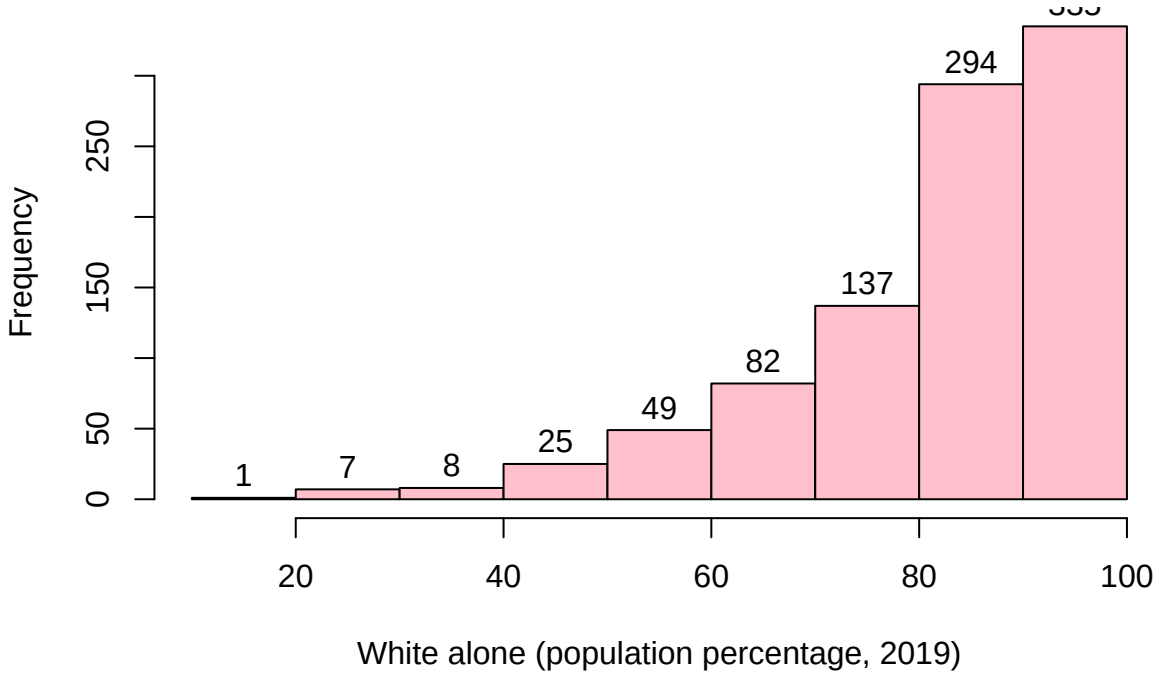

```
## `geom_smooth()` using formula = 'y ~ x'
```

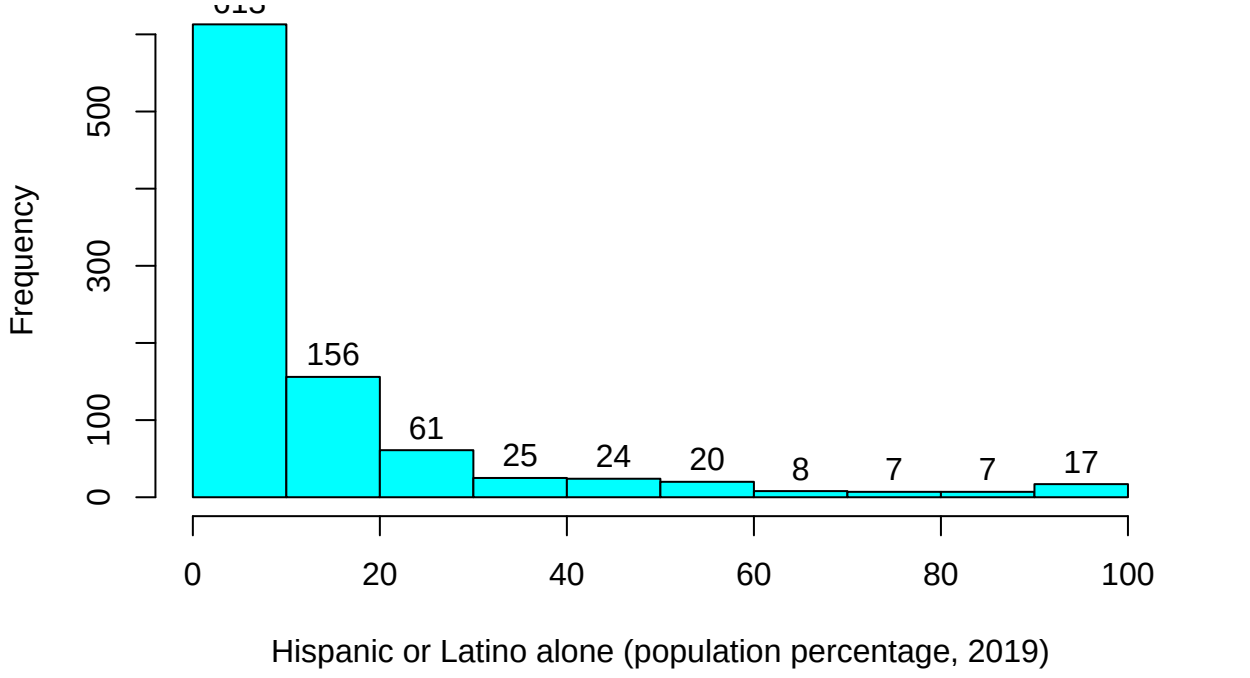



```
## `geom_smooth()` using formula = 'y ~ x'
```

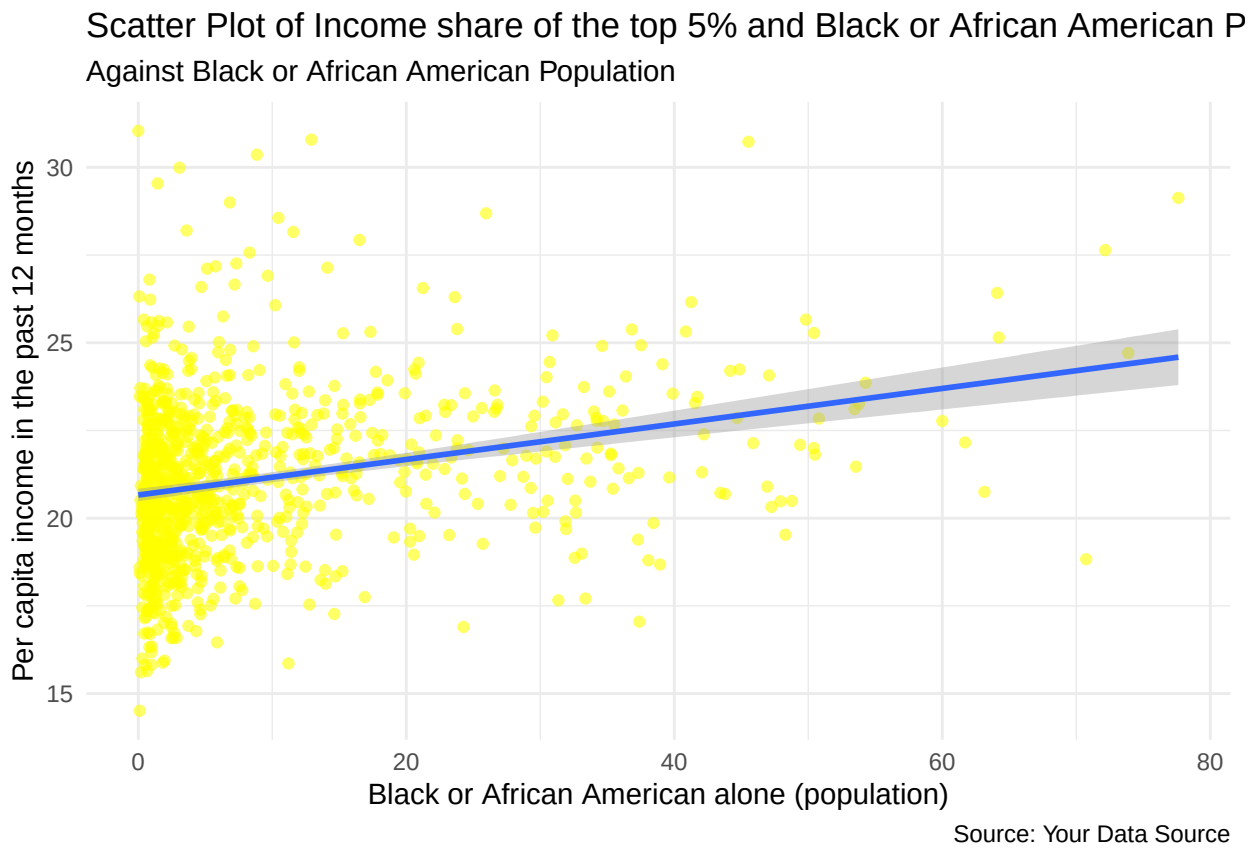

Scatter Plot of Income share of the top 5% and Hispanic or Latino Populatio
Against Hispanic or Latino Population
Income share of the top 5%
30
25
20
15
0
25
50
75
100
Hispanic or Latino (population)
Source: Your Data Source


Histogram of Black or African American alone (population percentage,
Frequency
0
100
300
500
700
118
56
50
21
9
4
4
0
20
40
60
80
Black or African American alone (population percentage, 2019)

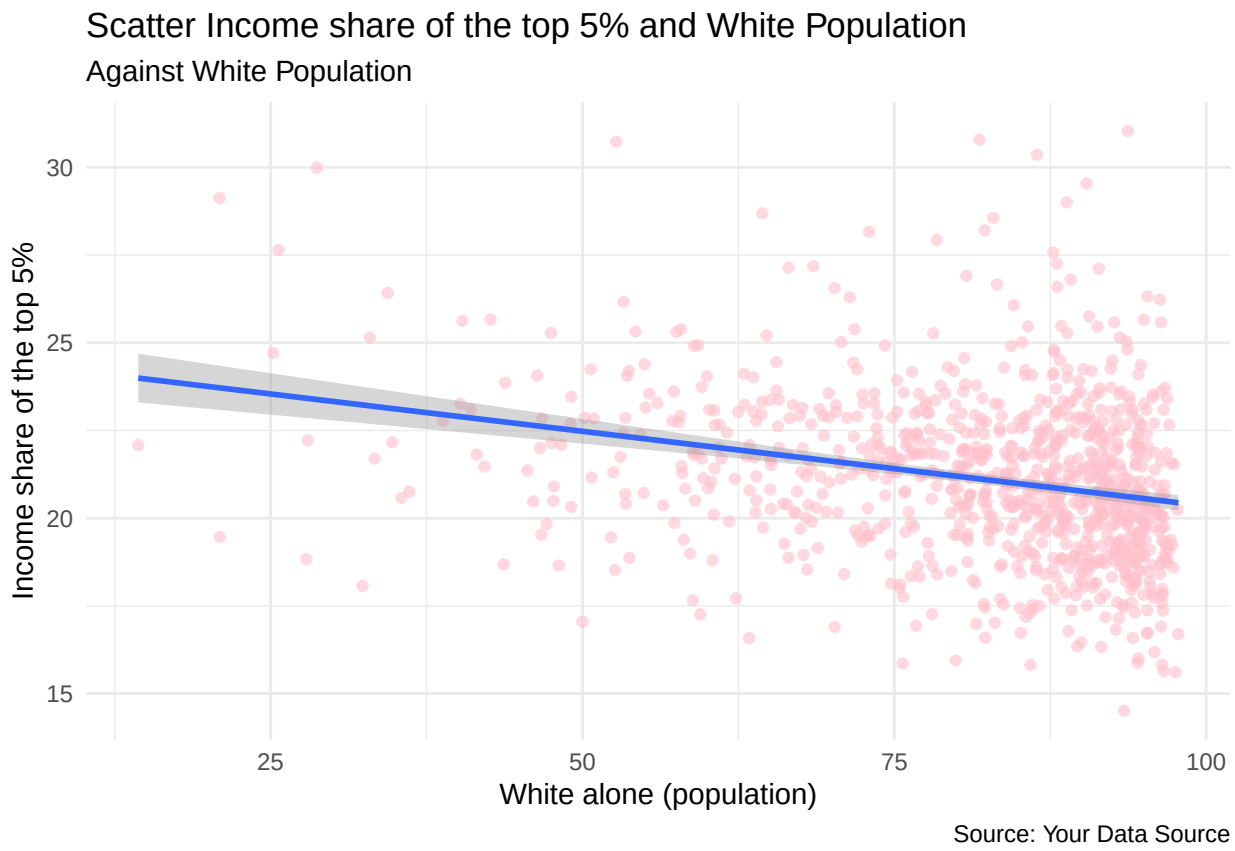

Histogram of White alone (population percentage, 2019)
Frequency
0
50
150
250
1
7
8
25
49
82
137
294
20
40
60
80
100
White alone (population percentage, 2019)



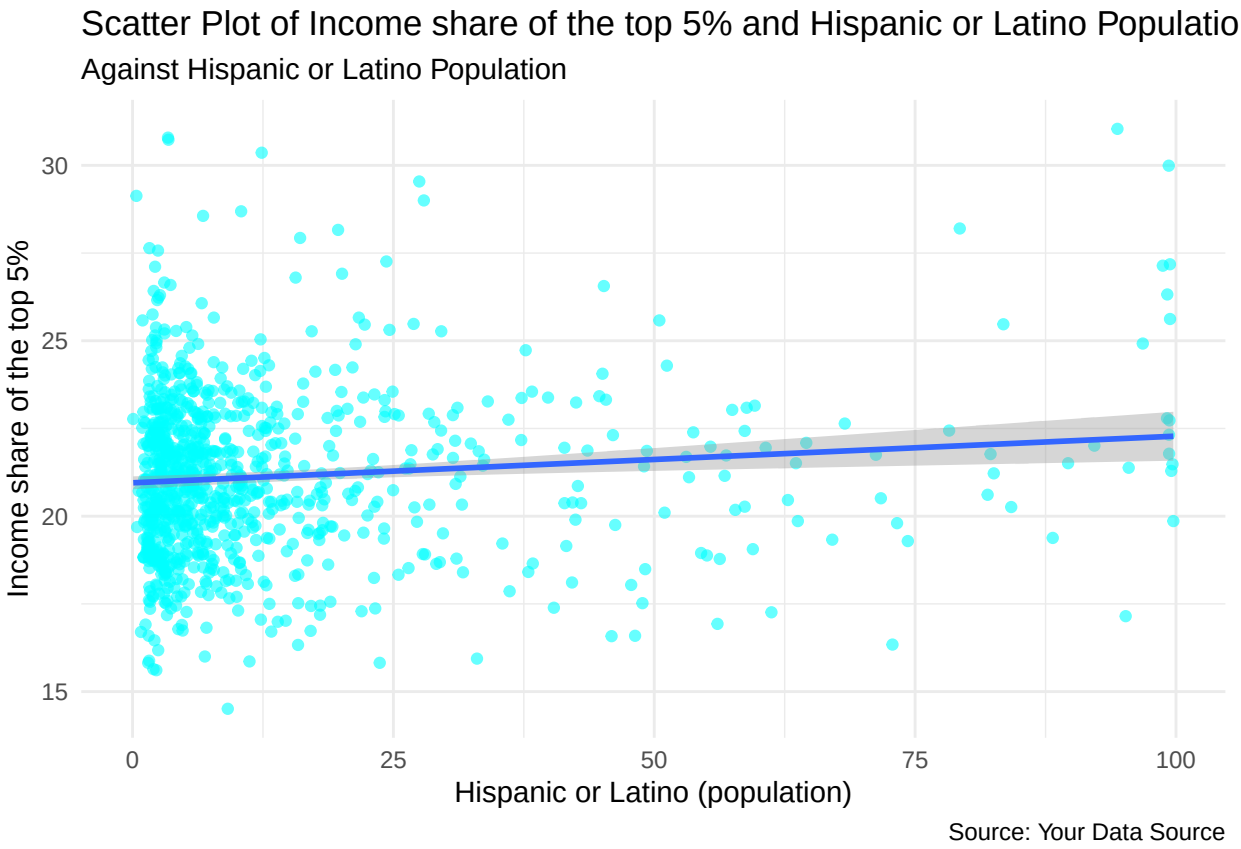

```
## `geom_smooth()` using formula = 'y ~ x'
```



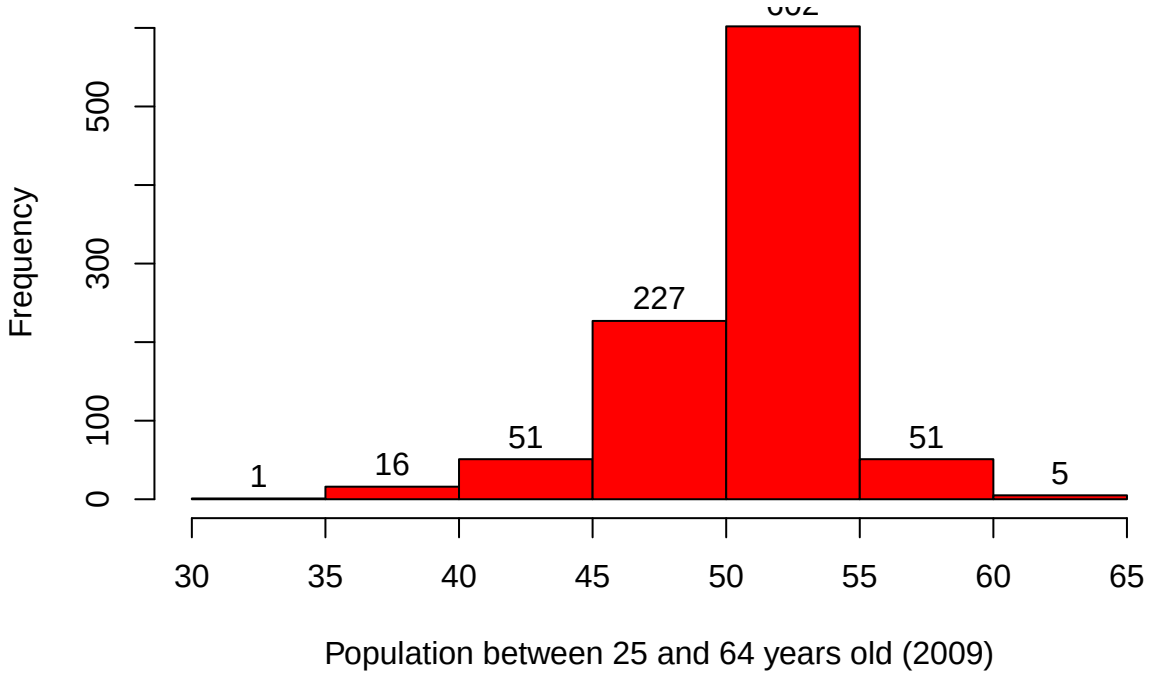

```
## `geom_smooth()` using formula = 'y ~ x'
```

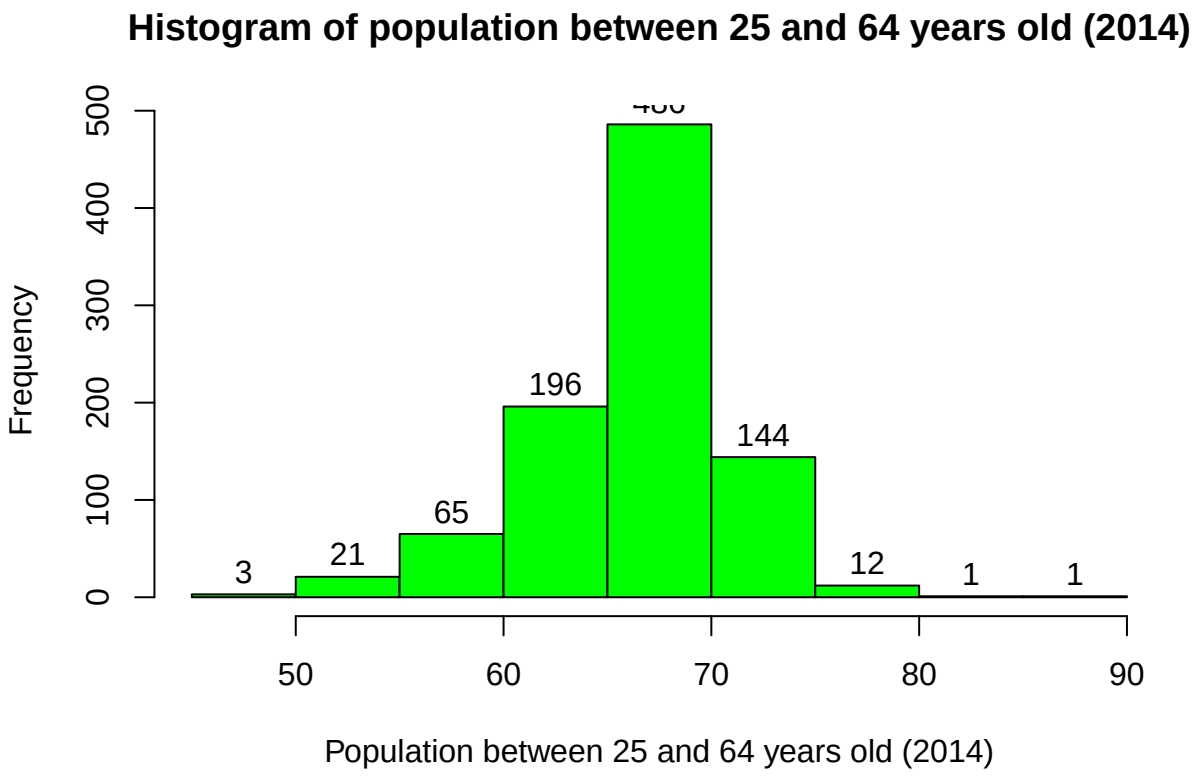


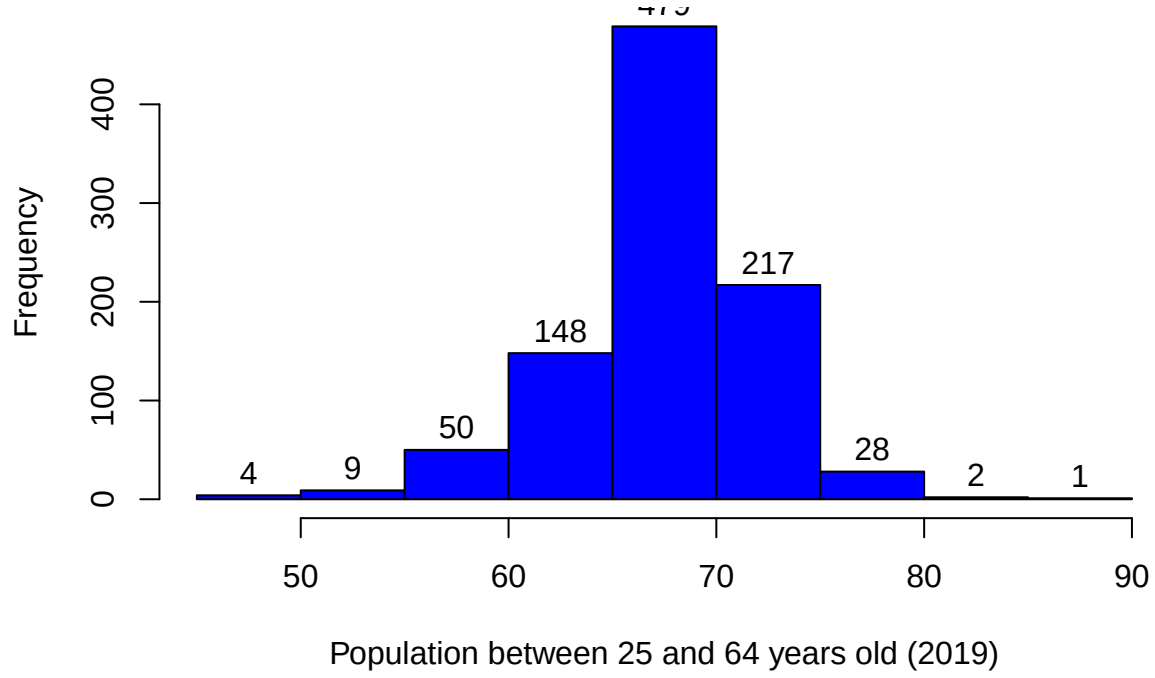

```
## `geom_smooth()` using formula = 'y ~ x'
```



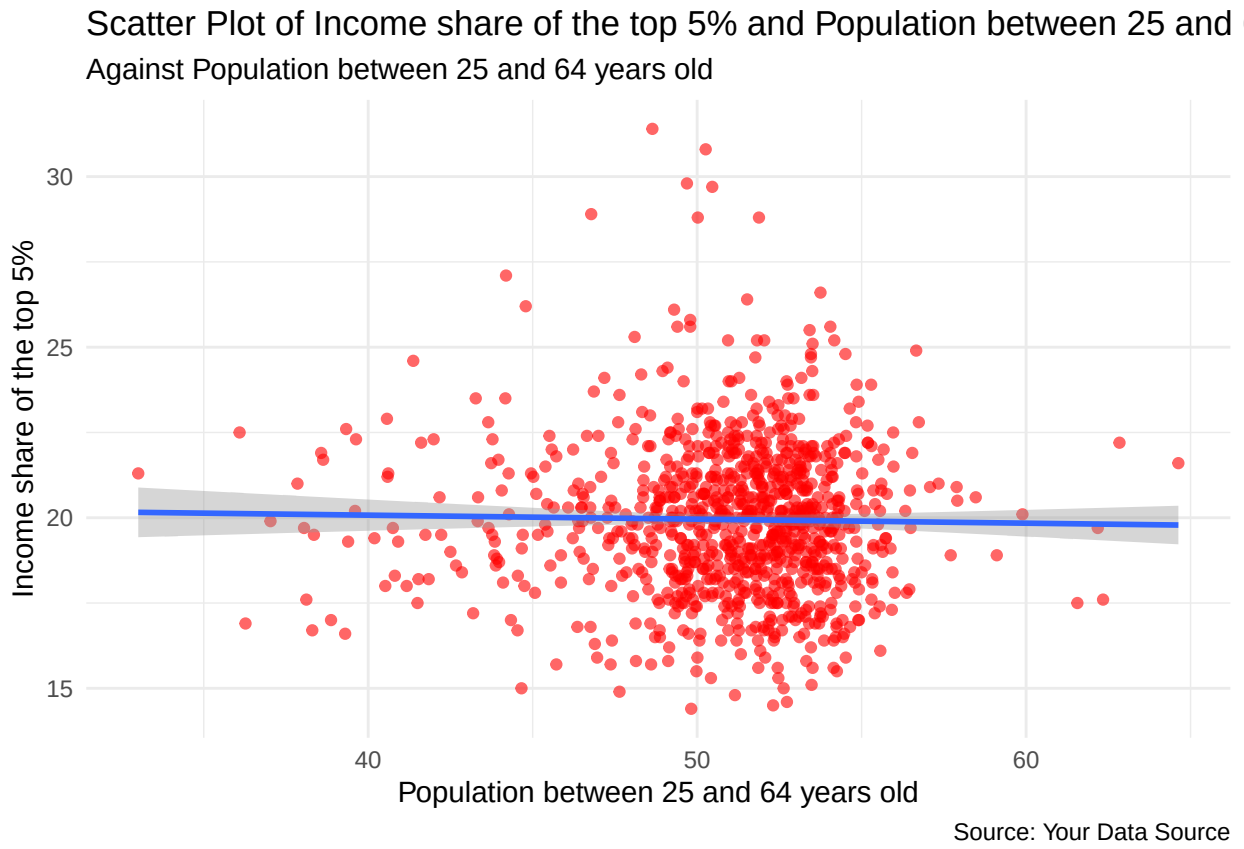
3. Working age


Histogram of population between 25 and 64 years old (2009)
Frequency
0
100
300
500
1
16
51
227
51
5
30
35
40
45
50
55
60
65
Population between 25 and 64 years old (2009)

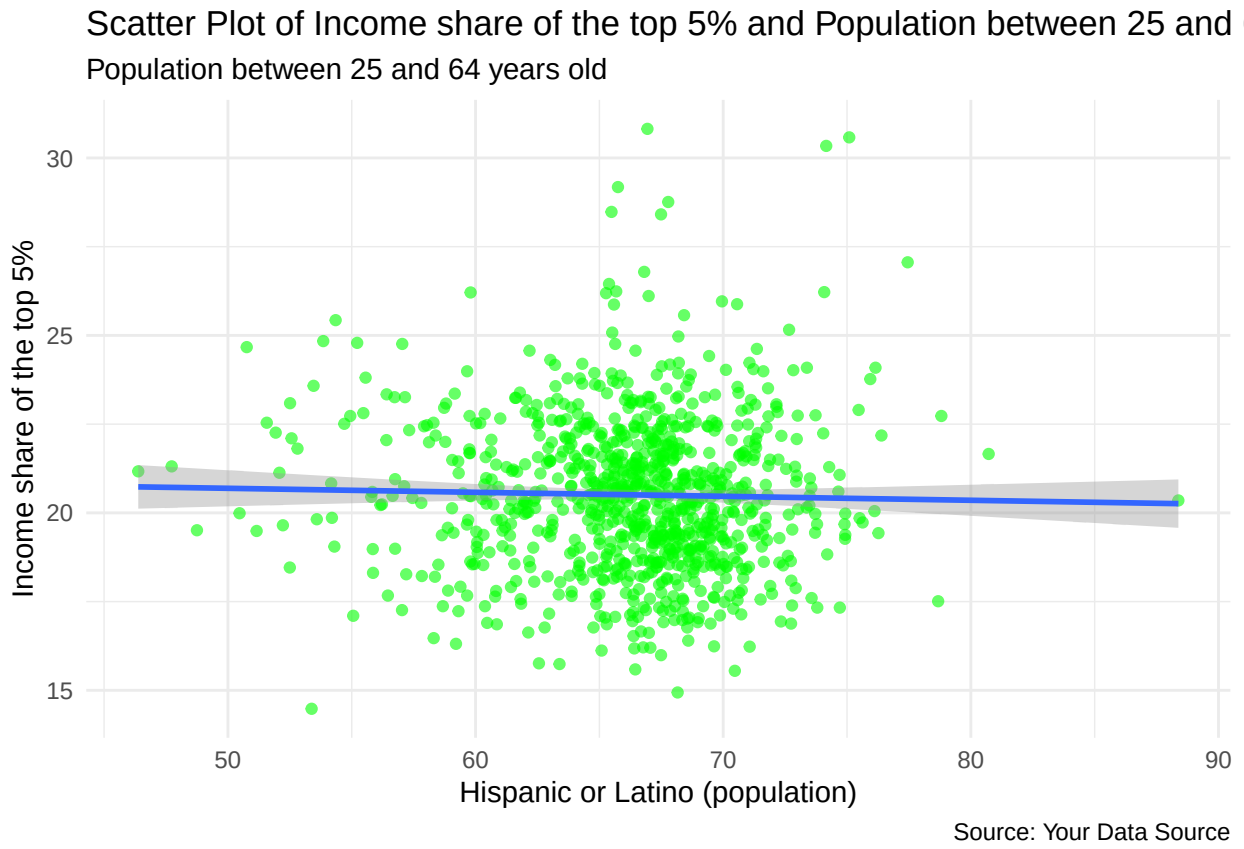

Histogram of population between 25 and 64 years old (2014)
Frequency
0
100
200
300
400
500
3
21
65
196
144
12
1
1
50
60
70
80
90
Population between 25 and 64 years old (2014)



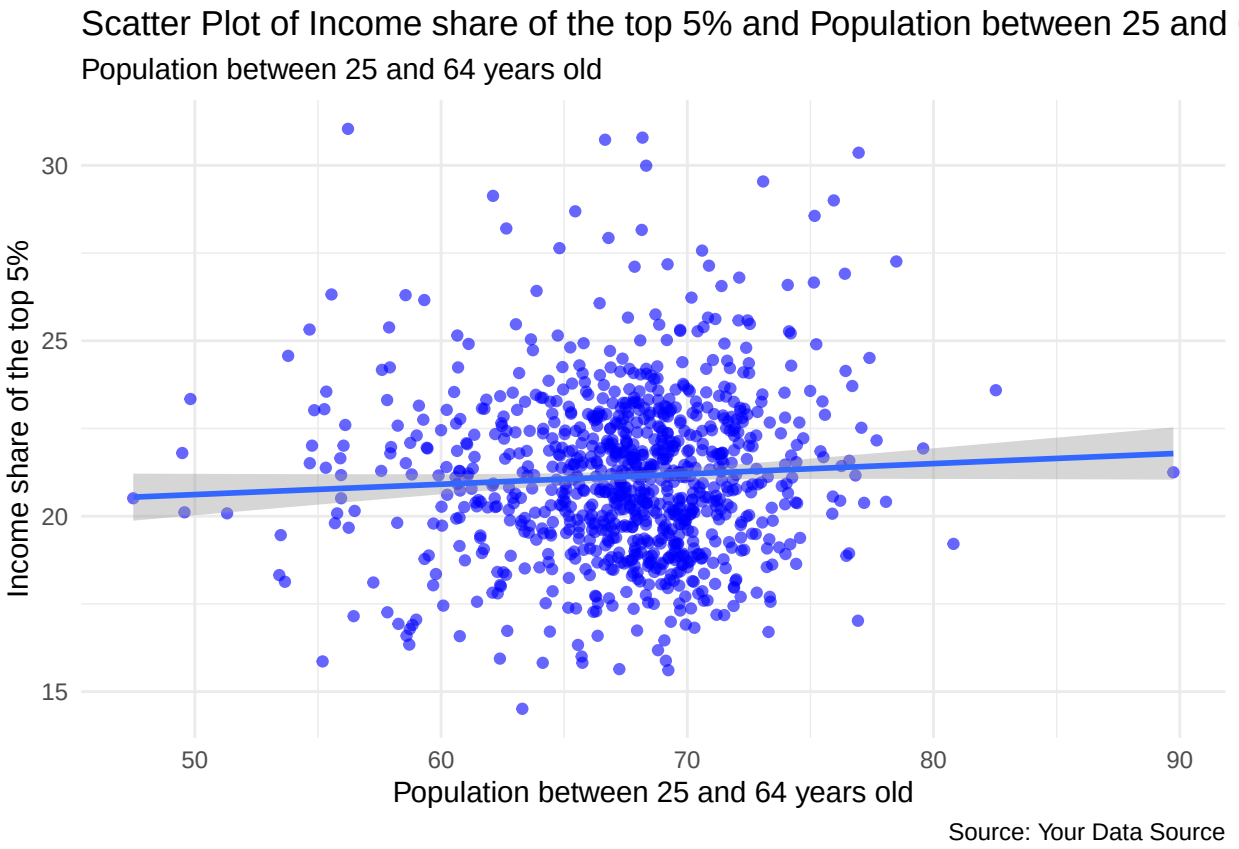

```
## `geom_smooth()` using formula = 'y ~ x'

## Warning: Removed 1 rows containing non-finite values (`stat_smooth()`).

## Warning: Removed 1 rows containing missing values (`geom_point()`).
```




```
## `geom_smooth()` using formula = 'y ~ x'
```




```
## `geom_smooth()` using formula = 'y ~ x'
```


Scatter Plot of Income share of the top 5% and Population between 25 and
Population between 25 and 64 years old
Income share of the top 5%
30
25
20
15
50
60
70
80
90
Population between 25 and 64 years old
Source: Your Data Source