Experiment-2
Introduction to SAS programming language

Tools used: SAS Online Studio

Code:

```
DATA outdata;
SET SASHELP.birthwgt;
proc print data=outdata(obs=10);

proc univariate data=outdata;
var AgeGroup;
histogram;

proc sgplot data=outdata;
vbar Race;
run;

proc freq data=outdata order=data;
   tables Drinking*Smoking / chisq;
run;

ods graphics on;
proc freq data=outdata;
   tables Race*Smoking /
          chisq cmh plots(only)=freqplot;
run;
ods graphics off;
```

Results:

Obs	LowBirthWgt	Married	AgeGroup	Race	Drinking	Death	Smoking	SomeCollege
1	No	No	3	Asian	No	No	No	Yes
2	No	No	2	White	No	No	No	No
3	Yes	Yes	2	Native	No	Yes	No	No
4	No	No	2	White	No	No	No	No
5	No	No	2	White	No	No	No	Yes
6	No	No	2	White	No	No	No	
7	No	No	2	Asian	No	No	No	Yes
8	No	No	3	White	No	No	No	Yes
9	No	Yes	1	Black	No	No	No	No
10	No	No	2	Native	No	No	No	Yes

The UNIVARIATE Procedure
Variable: AgeGroup

Moments			
N	100000	Sum Weights	100000
Mean	2.03877	Sum Observations	203877
Std Deviation	0.49210701	Variance	0.24216931
Skewness	0.08848977	Kurtosis	1.08996298
Uncorrected SS	439875	Corrected SS	24216.6887
Coeff Variation	24.1374461	Std Error Mean	0.00155618

Basic Statistical Measures			
Location		Variability	
Mean	2.038770	Std Deviation	0.49211
Median	2.000000	Variance	0.24217
Mode	2.000000	Range	2.00000
		Interquartile Range	0

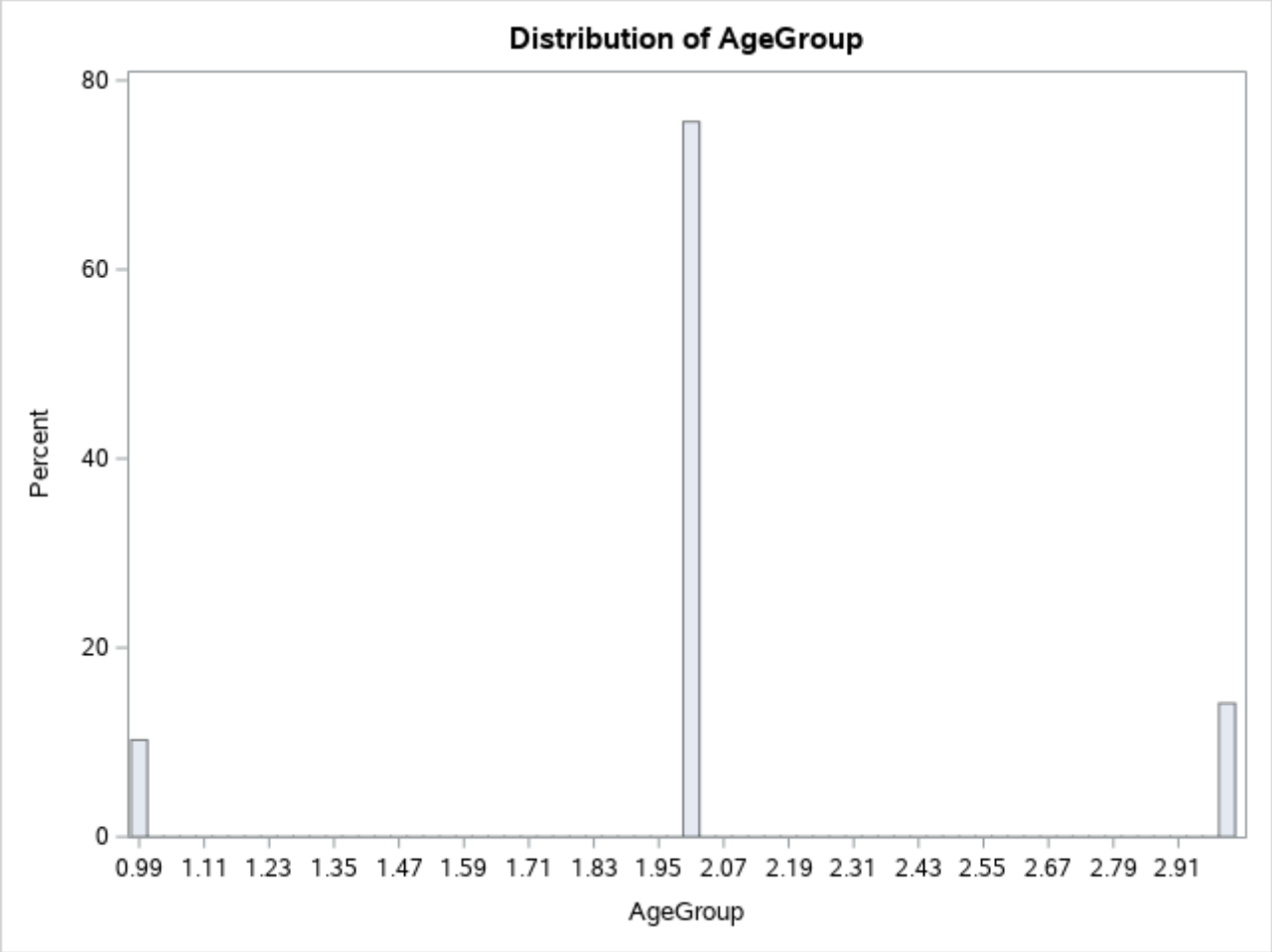
Tests for Location: Mu0=0				
Test	Statistic		p Value	
Student's t	t	1310.113	Pr > |t|	<.0001
Sign	M	50000	Pr >= |M|	<.0001
Signed Rank	S	2.5E9	Pr >= |S|	<.0001

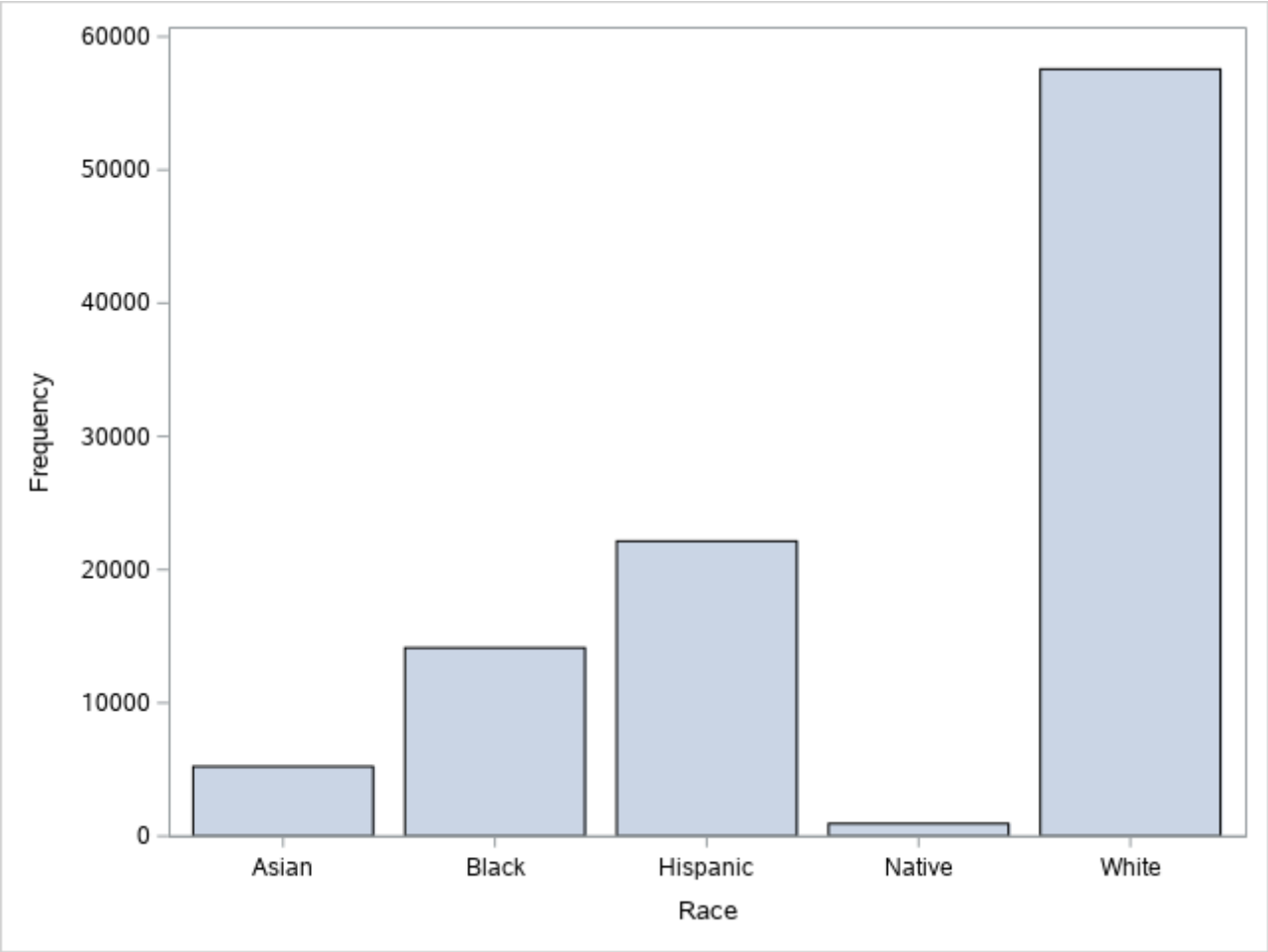
Quantiles (Definition 5)	
Level	Quantile
100% Max	3
99%	3
95%	3
90%	3
75% Q3	2
50% Median	2
25% Q1	2
10%	1
5%	1
1%	1
0% Min	1

Extreme Observations			
Lowest		Highest	
Value	Obs	Value	Obs

Extreme Observations			
Lowest		Highest	
Value	Obs	Value	Obs
1	99971	3	99979
1	99967	3	99982
1	99953	3	99983
1	99949	3	99990
1	99937	3	99992

The UNIVARIATE Procedure





The FREQ Procedure

Frequency Percent Row Pct Col Pct	Table of Drinking by Smoking		
	Smoking		
Drinking	No	Yes	Total
No	72679 77.00 89.82 99.50	8235 8.72 10.18 38.58	80914 85.72
Yes	364 0.39 2.70 0.50	13110 13.89 97.30 61.42	13474 14.28
Total	73043 77.39	21345 22.61	94388 100.00
Frequency Missing = 5612			

Statistics for Table of Drinking by Smoking

Statistic	DF	Value	Prob
Chi-Square	1	50096.5968	<.0001
Likelihood Ratio Chi-Square	1	44331.3293	<.0001
Continuity Adj. Chi-Square	1	50091.6186	<.0001
Mantel-Haenszel Chi-Square	1	50096.0660	<.0001
Phi Coefficient		0.7285	
Contingency Coefficient		0.5888	
Cramer's V		0.7285	

Fisher's Exact Test	
Cell (1,1) Frequency (F)	72679

Fisher's Exact Test	
Left-sided Pr <= F	1.0000
Right-sided Pr >= F	<.0001
Table Probability (P)	<.0001
Two-sided Pr <= P	<.0001

Sample Size = 94388
Frequency Missing = 5612

The FREQ Procedure

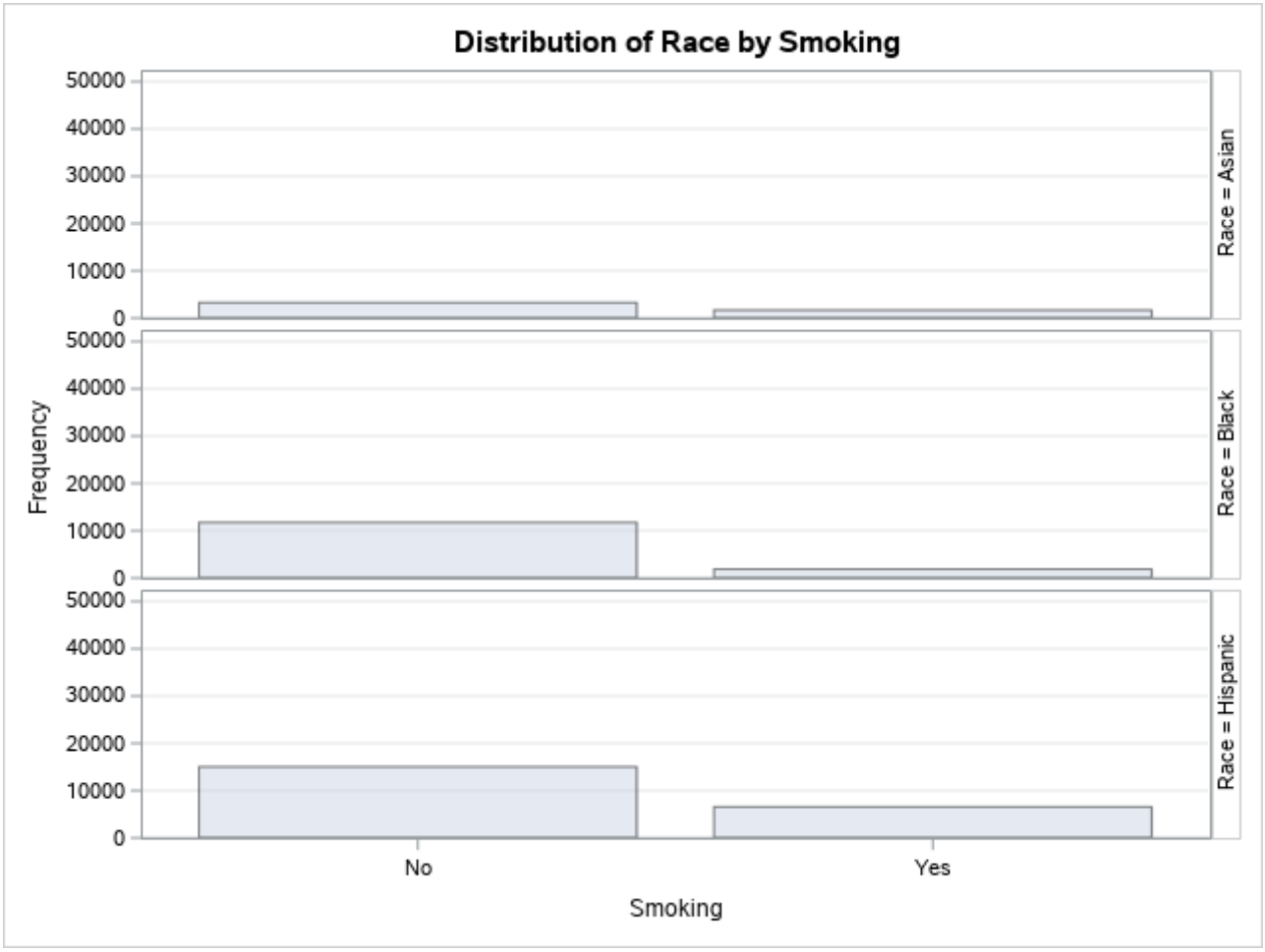
Frequency Percent Row Pct Col Pct

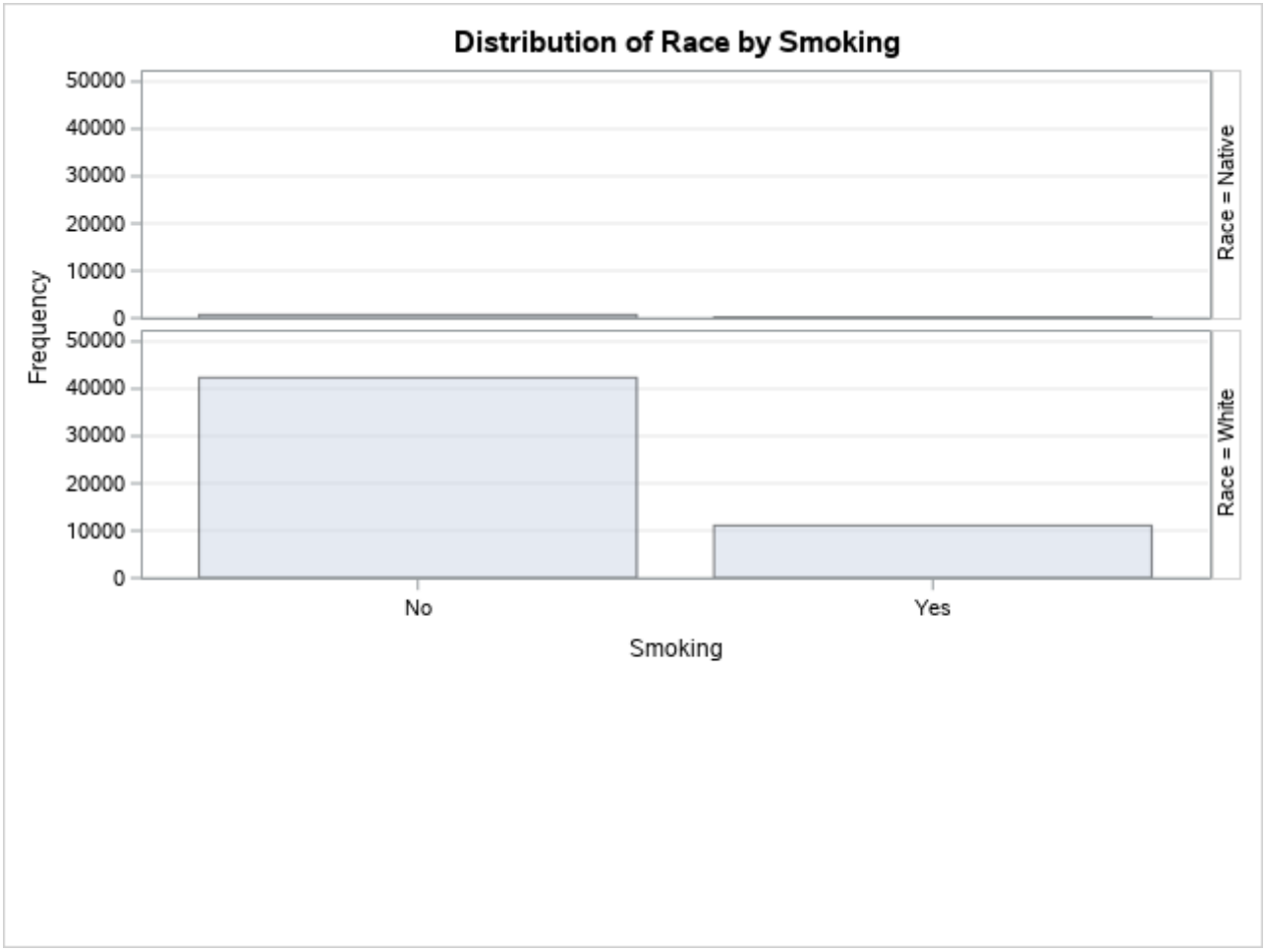
Table of Race by Smoking

Race	Smoking		
	No	Yes	Total
Asian	3251 3.44 66.10 4.45	1667 1.77 33.90 7.81	4918 5.21
Black	11723 12.42 86.39 16.05	1847 1.96 13.61 8.65	13570 14.38
Hispanic	15015 15.91 69.64 20.56	6547 6.94 30.36 30.67	21562 22.84
Native	697 0.74 78.49 0.95	191 0.20 21.51 0.89	888 0.94
White	42357 44.88 79.25 57.99	11093 11.75 20.75 51.97	53450 56.63
Total	73043 77.39	21345 22.61	94388 100.00

Frequency Missing = 5612

The FREQ Procedure





Statistics for Table of Race by Smoking

Statistic	DF	Value	Prob

Statistic	DF	Value	Prob
Chi-Square	4	1832.4632	<.0001
Likelihood Ratio Chi-Square	4	1827.9307	<.0001
Mantel-Haenszel Chi-Square	1	117.8617	<.0001
Phi Coefficient		0.1393	
Contingency Coefficient		0.1380	
Cramer's V		0.1393	

Sample Size = 94388
Frequency Missing = 5612

Summary Statistics for Race by Smoking

Cochran-Mantel-Haenszel Statistics (Based on Table Scores)				
Statistic	Alternative Hypothesis	DF	Value	Prob
1	Nonzero Correlation	1	117.8617	<.0001
2	Row Mean Scores Differ	4	1832.4438	<.0001
3	General Association	4	1832.4438	<.0001

Sample Size = 94388
Frequency Missing = 5612