Analysis of Crime Rate in USA

| SAS |

Table of Contents

Problem Statement

To determine the percentage of each type of crime committed in the United States of America from 1960 through 2016.

Data Description

The data provided has the following components:

Variable	Description
Year	The time frame span over 57 years, beginning from 1960 to 2016
Population	The population of each year has been assumed to be recorded at the end of the year
Property	The number of property crimes of each year have been recorded
Murder	The number of murders of each year have been recorded
Rape	The number of rapes of each year have been recorded
Robbery	The number of robberies of each year have been recorded
Aggravated Assault	The number of assaults of each year have been recorded
Burglary	The number of burglaries of each year have been recorded
Larceny Theft	The number of larceny thefts of each year have been recorded
Vehicle Theft	The number of vehicle thefts of each year have been recorded

Data Exploration

Assigning library name:

```
libname crime "/folders/myshortcuts/my_folder";
```

Data Input

```
/*Importing dataset*/
data crime.crimeRate;
	infile '/folders/myshortcuts/my_folder/crime data 1960-
2016.csv' delimiter=',' dsd missover firstobs=2;
	input Year Population Property Murder Rape Robbery
AggrevatedAssault Burglary LarcenyTheft VehicleTheft;
run;
```

Proc Contents

```
/*To check columns types, names and formats present in the dataset*/
```

```
proc contents data= crime.crimerate;
run;
```
Output:

The CONTENTS Procedure

Data Set Name	CRIME.CRIMERATE	**Observations**	57
Member Type	DATA	**Variables**	10
Engine	V9	**Indexes**	0
Created	11/13/2017 19:04:22	**Observation Length**	80
Last Modified	11/13/2017 19:04:22	**Deleted Observations**	0
Protection		**Compressed**	NO
Data Set Type		**Sorted**	NO
Label			
Data Representation	SOLARIS_X86_64, LINUX_X86_64, ALPHA_TRU64, LINUX_IA64		
Encoding	utf-8 Unicode (UTF-8)		

Engine/Host Dependent Information

Data Set Page Size	65536
Number of Data Set Pages	1
First Data Page	1
Max Obs per Page	817
Obs in First Data Page	57
Number of Data Set Repairs	0
Filename	/folders/myshortcuts/my_folder/crimerate.sas7bdat
Release Created	9.0401M4

Engine/Host Dependent Information	
Host Created	Linux
Inode Number	685
Access Permission	rwxrwx---
Owner Name	Root
File Size	128KB
File Size (bytes)	131072

Alphabetic List of Variables and Attributes

#	Variable	Type	Len
7	AggrevatedAssault	Num	8
8	Burglary	Num	8
9	LarcenyTheft	Num	8
4	Murder	Num	8
2	Population	Num	8
3	Property	Num	8
5	Rape	Num	8
6	Robbery	Num	8
10	VehicleTheft	Num	8
1	Year	Num	8

NMiss

```
/*To check for missing values*/
proc means data= crime.crimerate nmiss;
run;
```

Analysis:
There are no missing values in the dataset. We can therefore move forward.

Proc Means

```
/*To get minimum, maximum, standard deviation, mean, median and number
of missing values from
each column (for each variable)*/
proc means data=crime.crimerate;
      var Population Property Murder Rape Robbery AggrevatedAssault
Burglary LarcenyTheft VehicleTheft;
run;
```

Output:

The MEANS Procedure

Variable	Minimum	Maximum	Std Dev	Mean	Median	N Miss
Population	179323175	323127513	43269069.24	248818100	245807000	0
Property	3095700.00	12961100.00	2702962.57	9360732.44	10174754.00	0
Murder	8530.00	24700.00	4195.97	17289.40	17030.00	0
Rape	17190.00	109060.00	28044.08	73458.39	85141.00	0
Robbery	106670.00	687730.00	148691.52	413371.12	420806.00	0
AggrevatedAssault	154320.00	1135610.00	291302.59	684594.05	752423.00	0
Burglary	912100.00	3795200.00	731302.85	2393104.60	2228474.00	0
LarcenyTheft	1855400.00	8142200.00	1813943.93	5949291.39	6588046.00	0
VehicleTheft	328200.00	1661700.00	343649.37	1018413.44	1009600.00	0

Proc Freq

```
/*To get total levels for each variable*/
proc freq data= crime.crimerate nlevels;
      tables Population Property Murder Rape Robbery AggrevatedAssault
Burglary LarcenyTheft VehicleTheft /nocum nopercent noprint;
run;
```

Output:

The FREQ Procedure

Number of Variable Levels	
Variable	**Levels**
Population	57
Property	57
Murder	57
Rape	57
Robbery	57
AggrevatedAssault	57
Burglary	57
LarcenyTheft	57
VehicleTheft	57

Proc Univariate

```
/*To get Moments, Basic Statistical Measures, Tests for Location:
Mu0=0, Quantiles (Definition 5) and Extreme Lowest/highest Observations
for each column(variable)*/

proc univariate data=crime.crimerate;

      var Population Property Murder Rape Robbery AggrevatedAssault
Burglary LarcenyTheft VehicleTheft;

run;
```

Output:

The UNIVARIATE Procedure
Variable: Population

Moments			
N	57	**Sum Weights**	57
Mean	248818100	**Sum Observations**	1.41826E10
Std Deviation	43269069.2	**Variance**	1.87221E15

Moments			
Skewness	0.1740623	Kurtosis	-1.2306141
Uncorrected SS	3.63374E18	Corrected SS	1.04844E17
Coeff Variation	17.3898399	Std Error Mean	5731128.23

Basic Statistical Measures			
Location		Variability	
Mean	2.4882E8	Std Deviation	43269069
Median	2.4581E8	Variance	1.87221E15
Mode	.	Range	143804338
		Interquartile Range	76581924

Tests for Location: Mu0=0				
Test	Statistic		p Value	
Student's t	t	43.4152	Pr> |t|	<.0001
Sign	M	28.5	Pr>= |M|	<.0001
Signed Rank	S	826.5	Pr>= |S|	<.0001

Quantiles (Definition 5)	
Level	Quantile
100% Max	323127513
99%	323127513
95%	318907401
90%	311587816
75% Q3	287973924

Quantiles (Definition 5)	
Level	**Quantile**
50% Median	245807000
25% Q1	211392000
10%	193526000
5%	185771000
1%	179323175
0% Min	179323175

Extreme Observations			
Lowest		**Highest**	
Value	**Obs**	**Value**	**Obs**
179323175	1	313873685	53
182992000	2	316497531	54
185771000	3	318907401	55
188483000	4	320896618	56
191141000	5	323127513	57

The UNIVARIATE Procedure
Variable: Property

Moments			
N	57	**Sum Weights**	57
Mean	9360732.44	**Sum Observations**	533561749

Moments			
Std Deviation	2702962.57	Variance	7.30601E12
Skewness	-0.9231482	Kurtosis	-0.0176557
Uncorrected SS	5.40367E15	Corrected SS	4.09136E14
Coeff Variation	28.8755456	Std Error Mean	358016.138

Basic Statistical Measures			
Location		Variability	
Mean	9360732	Std Deviation	2702963
Median	10174754	Variance	7.30601E12
Mode	.	Range	9865400
		Interquartile Range	3639140

Tests for Location: Mu0=0				
Test	Statistic		p Value	
Student's t	t	26.14612	Pr> |t|	<.0001
Sign	M	28.5	Pr>= |M|	<.0001
Signed Rank	S	826.5	Pr>= |S|	<.0001

Quantiles (Definition 5)	
Level	Quantile
100% Max	12961100
99%	12961100
95%	12605400
90%	12218800

Quantiles (Definition 5)	
Level	**Quantile**
75% Q3	11558175
50% Median	10174754
25% Q1	7919035
10%	4352000
5%	3450700
1%	3095700
0% Min	3095700

Extreme Observations			
Lowest		**Highest**	
Value	**Obs**	**Value**	**Obs**
3095700	1	12356900	29
3198600	2	12505900	33
3450700	3	12605400	30
3792500	4	12655500	31
4200400	5	12961100	32

The UNIVARIATE Procedure
Variable: Murder

Moments			
N	57	**Sum Weights**	57

Moments			
Mean	17289.4035	Sum Observations	985496
Std Deviation	4195.9671	Variance	17606139.9
Skewness	-0.3976125	Kurtosis	-0.2933924
Uncorrected SS	1.80246E10	Corrected SS	985943836
Coeff Variation	24.2690102	Std Error Mean	555.76942

Basic Statistical Measures			
Location		Variability	
Mean	17289.40	Std Deviation	4196
Median	17030.00	Variance	17606140
Mode	.	Range	16170
		Interquartile Range	5644

Tests for Location: Mu0=0				
Test	Statistic		p Value	
Student's t	t	31.10895	Pr> |t|	<.0001
Sign	M	28.5	Pr>= |M|	<.0001
Signed Rank	S	826.5	Pr>= |S|	<.0001

Quantiles (Definition 5)	
Level	Quantile
100% Max	24700
99%	24700
95%	23760

Quantiles (Definition 5)	
Level	Quantile
90%	23040
75% Q3	20510
50% Median	17030
25% Q1	14866
10%	9960
5%	8740
1%	8530
0% Min	8530

Extreme Observations			
Lowest		Highest	
Value	Obs	Value	Obs
8530	3	23330	35
8640	4	23440	31
8740	2	23760	33
9110	1	24530	34
9360	5	24700	32

The UNIVARIATE Procedure
Variable: Rape

Moments

Moments			
N	57	Sum Weights	57
Mean	73458.386	Sum Observations	4187128
Std Deviation	28044.076	Variance	786470199
Skewness	-0.9038882	Kurtosis	-0.603183
Uncorrected SS	3.51622E11	Corrected SS	4.40423E10
Coeff Variation	38.1768203	Std Error Mean	3714.52861

Basic Statistical Measures			
Location		Variability	
Mean	73458.39	Std Deviation	28044
Median	85141.00	Variance	786470199
Mode	.	Range	91870
		Interquartile Range	37703

Tests for Location: Mu0=0				
Test	Statistic		p Value	
Student's t	t	19.77596	Pr> |t|	<.0001
Sign	M	28.5	Pr>= |M|	<.0001
Signed Rank	S	826.5	Pr>= |S|	<.0001

Quantiles (Definition 5)	
Level	Quantile
100% Max	109060
99%	109060

Quantiles (Definition 5)	
Level	**Quantile**
95%	106010
90%	97470
75% Q3	93103
50% Median	85141
25% Q1	55400
10%	23410
5%	17550
1%	17190
0% Min	17190

Extreme Observations			
Lowest		**Highest**	
Value	**Obs**	**Value**	**Obs**
17190	1	102220	35
17220	2	102560	31
17550	3	106010	34
17650	4	106590	32
21420	5	109060	33

The UNIVARIATE Procedure
Variable: Robbery

Moments			
N	57	Sum Weights	57
Mean	413371.123	Sum Observations	23562154
Std Deviation	148691.519	Variance	2.21092E10
Skewness	-0.4713573	Kurtosis	-0.0841579
Uncorrected SS	1.0978E13	Corrected SS	1.23811E12
Coeff Variation	35.9704659	Std Error Mean	19694.6728

Basic Statistical Measures			
Location		Variability	
Mean	413371.1	Std Deviation	148692
Median	420806.0	Variance	2.21092E10
Mode	.	Range	581060
		Interquartile Range	156710

Tests for Location: Mu0=0				
Test	Statistic		p Value	
Student's t	t	20.98898	Pr> |t|	<.0001
Sign	M	28.5	Pr>= |M|	<.0001
Signed Rank	S	826.5	Pr>= |S|	<.0001

Quantiles (Definition 5)	
Level	Quantile
100% Max	687730
99%	687730

Quantiles (Definition 5)	
Level	**Quantile**
95%	659870
90%	592910
75% Q3	506570
50% Median	420806
25% Q1	349860
10%	138690
5%	110860
1%	106670
0% Min	106670

Extreme Observations			
Lowest		**Highest**	
Value	**Obs**	**Value**	**Obs**
106670	2	618950	35
107840	1	639270	31
110860	3	659870	34
116470	4	672480	33
130390	5	687730	32

The UNIVARIATE Procedure
Variable: AggrevatedAssault

Moments			
N	57	Sum Weights	57
Mean	684594.053	Sum Observations	39021861
Std Deviation	291302.587	Variance	8.48572E10
Skewness	-0.4073039	Kurtosis	-0.9288253
Uncorrected SS	3.14661E13	Corrected SS	4.752E12
Coeff Variation	42.5511419	Std Error Mean	38583.9702

Basic Statistical Measures			
Location		Variability	
Mean	684594.1	Std Deviation	291303
Median	752423.0	Variance	8.48572E10
Mode	.	Range	981290
		Interquartile Range	435197

Tests for Location: Mu0=0				
Test	Statistic		p Value	
Student's t	t	17.74297	Pr> |t|	<.0001
Sign	M	28.5	Pr>= |M|	<.0001
Signed Rank	S	826.5	Pr>= |S|	<.0001

Quantiles (Definition 5)	
Level	Quantile
100% Max	1135610
99%	1135610

Quantiles (Definition 5)	
Level	**Quantile**
95%	1113180
90%	1054860
75% Q3	891407
50% Median	752423
25% Q1	456210
10%	215330
5%	164570
1%	154320
0% Min	154320

Extreme Observations			
Lowest		**Highest**	
Value	**Obs**	**Value**	**Obs**
154320	1	1092740	32
156760	2	1099210	36
164570	3	1113180	35
174210	4	1126970	33
203050	5	1135610	34

The UNIVARIATE Procedure
Variable: Burglary

Moments			
N	57	Sum Weights	57
Mean	2393104.6	Sum Observations	136406962
Std Deviation	731302.852	Variance	5.34804E11
Skewness	-0.1897614	Kurtosis	-0.6414656
Uncorrected SS	3.56385E14	Corrected SS	2.9949E13
Coeff Variation	30.55875	Std Error Mean	96863.429

Basic Statistical Measures			
Location		Variability	
Mean	2393105	Std Deviation	731303
Median	2228474	Variance	5.34804E11
Mode	.	Range	2883100
		Interquartile Range	1022308

Tests for Location: Mu0=0				
Test	Statistic		p Value	
Student's t	t	24.70597	Pr> |t|	<.0001
Sign	M	28.5	Pr>= |M|	<.0001
Signed Rank	S	826.5	Pr>= |S|	<.0001

Quantiles (Definition 5)	
Level	Quantile
100% Max	3795200
99%	3795200

Quantiles (Definition 5)	
Level	**Quantile**
95%	3447100
90%	3241410
75% Q3	3073300
50% Median	2228474
25% Q1	2050992
10%	1282500
5%	994300
1%	912100
0% Min	912100

Extreme Observations			
Lowest		**Highest**	
Value	**Obs**	**Value**	**Obs**
912100	1	3265300	16
949600	2	3327700	20
994300	3	3447100	23
1086400	4	3779700	22
1213200	5	3795200	21

The UNIVARIATE Procedure
Variable: LarcenyTheft

Moments			
N	57	Sum Weights	57
Mean	5949291.39	Sum Observations	339109609
Std Deviation	1813943.93	Variance	3.29039E12
Skewness	-0.9821954	Kurtosis	-0.1609742
Uncorrected SS	2.20172E15	Corrected SS	1.84262E14
Coeff Variation	30.4900838	Std Error Mean	240262.743

Basic Statistical Measures			
Location		Variability	
Mean	5949291	Std Deviation	1813944
Median	6588046	Variance	3.29039E12
Mode	.	Range	6286800
		Interquartile Range	1880000

Tests for Location: Mu0=0				
Test	Statistic		p Value	
Student's t	t	24.76161	Pr> |t|	<.0001
Sign	M	28.5	Pr>= |M|	<.0001
Signed Rank	S	826.5	Pr>= |S|	<.0001

Quantiles (Definition 5)	
Level	Quantile
100% Max	8142200
99%	8142200

Quantiles (Definition 5)	
Level	**Quantile**
95%	7945700
90%	7879800
75% Q3	7142500
50% Median	6588046
25% Q1	5262500
10%	2572600
5%	2089600
1%	1855400
0% Min	1855400

Extreme Observations			
Lowest		**Highest**	
Value	**Obs**	**Value**	**Obs**
1855400	1	7904700	37
1913000	2	7915200	33
2089600	3	7945700	31
2297800	4	7997700	36
2514400	5	8142200	32

The UNIVARIATE Procedure
Variable: VehicleTheft

Moments			
N	57	Sum Weights	57
Mean	1018413.44	Sum Observations	58049566
Std Deviation	343649.374	Variance	1.18095E11
Skewness	-0.1252003	Kurtosis	-0.5047485
Uncorrected SS	6.57318E13	Corrected SS	6.61331E12
Coeff Variation	33.7436017	Std Error Mean	45517.4716

Basic Statistical Measures			
Location		Variability	
Mean	1018413	Std Deviation	343649
Median	1009600	Variance	1.18095E11
Mode	.	Range	1333500
		Interquartile Range	472367

Tests for Location: Mu0=0				
Test	Statistic		p Value	
Student's t	t	22.37412	Pr> |t|	<.0001
Sign	M	28.5	Pr>= |M|	<.0001
Signed Rank	S	826.5	Pr>= |S|	<.0001

Quantiles (Definition 5)	
Level	Quantile
100% Max	1661700
99%	1661700

Quantiles (Definition 5)	
Level	Quantile
95%	1610800
90%	1539300
75% Q3	1237851
50% Median	1009600
25% Q1	765484
10%	496900
5%	366800
1%	328200
0% Min	328200

Extreme Observations			
Lowest		Highest	
Value	Obs	Value	Obs
328200	1	1563100	34
336000	2	1564800	30
366800	3	1610800	33
408300	4	1635900	31
472800	5	1661700	32

Analysis:

This step gives us the basic descriptive statistics for each variable.

Proc Sort

```
/*To sort data by year and store it in a new dataset to work on*/
```

```
proc sort data=crime.crimerate
              out=work.maindata;
              by year;
run;
```

Data Transformation

New Column Addition

```
/*Creating column for total number of crimes per year.
creating column to find out minimum crime per year when compare to all
types of crime.
creating column to find out maximum crime per year when compare to all
types of crime.
to get value of a variable Population in the previous observation.
to calculate Rate of population.
to get value of a variable Total in the previous observation.
to calculate Rate of Total crime for each year. */

data work.Total;
      set work.maindata;

      Total = SUM(of Property--VehicleTheft);
      MinCrime = MIN(of Property--VehicleTheft);
      MaxCrime = MAX(of Property--VehicleTheft);

      lagPopulation = lag(Population);
      rate_Population = (Population-lagPopulation)/lagPopulation;
      lagTotal = lag(Total);
      rate_Total = (Total-lagTotal)/lagTotal;

      Label rate_Population = "Rate of Population" rate_Total = "Rate
      of Total"
              MinCrime = " Minimum crime in a year" MaxCrime = "Maximum
              crime in a year";
run;
```

Proc Print

```
/*Printing dataset total*/
proc print data= work.total (drop = lagPopulation lagTotal) Label;
run;
```

Output:

Obs	Year	Population	Property	Murder	Rape	Robbery	AggrevatedAssault	Burglary	LarcenyTheft	VehicleTheft	Total	Minimum crime in a year	Maximum crime in a year	Rate of Population	Rate of Total
1	1960	179323175	3095700	9110	17190	107840	154320	912100	1855400	328200	6479860	9110	3095700	.	.
2	1961	182992000	3198600	8740	17220	106670	156760	949600	1913000	336000	6686590	8740	3198600	0.02045	0.03190
3	1962	185771000	3450700	8530	17550	110860	164570	994300	2089600	366800	7202910	8530	3450700	0.01518	0.07722
4	1963	188483000	3792500	8640	17650	116470	174210	1086400	2297800	408300	7901970	8640	3792500	0.01459	0.09705
5	1964	191141000	4200400	9360	21420	130390	203050	1213200	2514400	472800	8765020	9360	4200400	0.01410	0.10922
6	1965	193526000	4352000	9960	23410	138690	215330	1282500	2572600	496900	9091390	9960	4352000	0.01247	0.03724
7	1966	195576000	4793300	11040	25820	157990	235330	1410100	2822000	561200	10016780	11040	4793300	0.01059	0.10179
8	1967	197457000	5403500	12240	27620	202910	257160	1632100	3111600	659800	11306930	12240	5403500	0.00961	0.12880
9	1968	199399000	6125200	13800	31670	262840	286700	1858900	3482700	783600	12845410	13800	6125200	0.00983	0.13607
10	1969	201385000	6749000	14760	37170	298850	311090	1981900	3888600	878500	14159870	14760	6749000	0.00960	0.10233
11	1970	203235298	7359200	16000	37990	349860	334970	2205000	4225800	928400	15457220	16000	7359200	0.00918	0.09162
12	1971	206212000	7771700	17780	42260	387700	368760	2399300	4424200	948200	16359900	17780	7771700	0.01464	0.05840
13	1972	208230000	7413900	18670	46850	376290	393090	2375500	4151200	887200	15662700	18670	7413900	0.00978	-0.04262
14	1973	209851000	7842200	19640	51400	384220	420650	2565500	4347900	928800	16560310	19640	7842200	0.00785	0.05731
15	1974	211392000	9278700	20710	55400	442400	456210	3039200	5262500	977100	19532220	20710	9278700	0.00734	0.17946

Obs	Year	Population	Property	Murder	Rape	Robbery	AggrevatedAssault	Burglary	LarcenyTheft	VehicleTheft	Total	Minimum crime in a year	Maximum crime in a year	Rate of Population	Rate of Total
16	1975	213124000	10252700	20510	56090	470500	492620	3265300	5977700	1009600	21545020	20510	10252700	0.008193	0.10305
17	1976	214659000	10345500	18780	57080	427810	500530	3108700	6270800	966000	21695200	18780	10345500	0.007202	0.00697
18	1977	216332000	9955000	19120	63500	412610	534350	3071500	5905700	977700	20939480	19120	9955000	0.007794	-0.03483
19	1978	218059000	10123400	19560	67610	426930	571460	3128300	5991000	1004100	21332360	19560	10123400	0.007983	0.01876
20	1979	220099000	11041500	21460	76390	480700	629480	3327700	6601000	1112800	23291030	21460	11041500	0.009355	0.09182
21	1980	225349264	12063700	23040	82990	565840	672650	3795200	7136900	1131700	25472020	23040	12063700	0.023854	0.09364
22	1981	229146000	12061900	22520	82500	592910	663900	3779700	7194400	1087800	25485630	22520	12061900	0.016848	0.00053
23	1982	231534000	11652000	21010	78770	553130	669480	3447100	7142500	1062400	24626390	21010	11652000	0.010421	-0.03371
24	1983	233981000	10850500	19310	78920	506570	653290	3129900	6712800	1007900	22959190	19310	10850500	0.010569	-0.06770
25	1984	236158000	10608500	18690	84230	485010	685350	2984400	6591900	1032200	22490280	18690	10608500	0.009304	-0.02042
26	1985	238740000	11102600	18980	88670	497870	723250	3073300	6926400	1102900	23533970	18980	11102600	0.010933	0.04641
27	1986	240132887	11722700	20613	91459	542775	834322	3241410	7257153	1224137	24934569	20613	11722700	0.005834	0.05951
28	1987	242282918	12024700	20096	91110	517704	855088	3236184	7499900	1288674	25533456	20096	12024700	0.008954	0.02402
29	1988	245807000	12356900	20680	92490	542970	910090	3218100	7705900	1432900	26280030	20680	12356900	0.014545	0.02924

Obs	Year	Population	Property	Murder	Rape	Robbery	AggrevatedAssault	Burglary	LarcenyTheft	VehicleTheft	Total	Minimum crime in a year	Maximum crime in a year	Rate of Population	Rate of Total
30	1989	248239000	12605400	21500	94500	578330	951710	3168200	7872400	1564800	26856840	21500	12605400	0.00989894	0.02195
31	1990	248709873	12655500	23440	102560	639270	1054860	3073900	7945700	1635900	27131130	23440	12655500	0.00101897	0.01021
32	1991	252177000	12961100	24700	106590	687730	1092740	3157200	8142200	1661700	27833960	24700	12961100	0.00913940	0.02590
33	1992	255082000	12505900	23760	109060	672480	1126970	2979900	7915200	1610800	26944070	23760	12505900	0.00911520	-0.03197
34	1993	257908000	12218800	24530	106010	659870	1135610	2834800	7820900	1563100	26363620	24530	12218800	0.00911079	-0.02154
35	1994	260341000	12131900	23330	102220	618950	1113180	2712800	7879800	1539300	26121480	23330	12131900	0.00909434	-0.00918
36	1995	262755000	12063900	21610	97470	580510	1099210	2593800	7997700	1472400	25926600	21610	12063900	0.00909272	-0.00746
37	1996	265228572	11805300	19650	96250	535590	1037050	2506400	7904700	1394200	25299140	19650	11805300	0.00909414	-0.02420
38	1997	267637000	11558175	18208	96153	498534	1023201	2460526	7743760	1354189	24752746	18208	11558175	0.00909081	-0.02160
39	1998	270296000	10944590	16914	93103	446625	974402	2329950	7373886	1240754	23420224	16914	10944590	0.00909935	-0.05383
40	1999	272690813	10208334	15522	89411	409371	911740	2100739	6955520	1152075	21842712	15522	10208334	0.00808860	-0.06736
41	2000	281421906	10182586	15586	90178	408016	911706	2050992	6971590	1160002	21790656	15586	10182586	0.03202018	-0.00238
42	2001	285317559	10437480	16037	90863	423557	909023	2116531	7092267	1228391	22314149	16037	10437480	0.01303843	0.02402

Obs	Year	Population	Property	Murder	Rape	Robbery	AggrevatedAssault	Burglary	LarcenyTheft	VehicleTheft	Total	Minimum crime in a year	Maximum crime in a year	Rate of Population	Rate of Total
43	2002	287973924	10455277	16229	95235	420806	891407	2151252	7057370	1246646	22334222	16229	10455277	0.009310	0.00090
44	2003	290690788	10442862	16528	93883	414235	859030	2154834	7026802	1261226	22269400	16528	10442862	0.009434	-0.00290
45	2004	293656842	10319386	16148	95089	401470	847381	2144446	6937089	1237851	21998860	16148	10319386	0.010203	-0.01215
46	2005	296507061	10174754	16740	94347	417438	862220	2155448	6783447	1235859	21740253	16740	10174754	0.009706	-0.01176
47	2006	299398484	9983568	17030	92757	447403	860853	2183746	6607013	1192809	21385179	17030	9983568	0.009752	-0.01633
48	2007	301621157	9843481	16929	90427	445125	855856	2176140	6568572	1095769	21092299	16929	9843481	0.007424	-0.01370
49	2008	304374846	9767915	16442	90479	443574	842134	2228474	6588046	958629	20935693	16442	9767915	0.009130	-0.00742
50	2009	307006550	9337060	15399	89241	408742	812514	2203313	6338095	795652	20000016	15399	9337060	0.008646	-0.04469
51	2010	309330219	9112625	14772	85593	369089	781844	2168457	6204601	739565	19476546	14772	9112625	0.007569	-0.02617
52	2011	311587816	9052743	14661	84175	354772	752423	2185140	6151095	716508	19311517	14661	9052743	0.007298	-0.00847
53	2012	313873685	9001992	14866	85141	355051	762009	2109932	6168874	723186	19221051	14866	9001992	0.007336	-0.00468
54	2013	316497531	8650761	14319	82109	345095	726575	1931835	6018632	700294	18469620	14319	8650761	0.008360	-0.03909

Obs	Year	Population	Property	Murder	Rape	Robbery	AggrevatedAssault	Burglary	LarcenyTheft	VehicleTheft	Total	Minimum crime in a year	Maximum crime in a year	Rate of Population	Rate of Total
55	2014	318907401	8209010	14164	84864	322905	731089	1713153	5809054	686803	17571042	14164	8209010	0.007614	-0.04865
56	2015	320896618	8024115	15883	91261	328109	764057	1587564	5723488	713063	17247540	15883	8024115	0.006238	-0.01841
57	2016	323127513	7919035	17250	95730	332198	803007	1515096	5638455	765484	17086255	17250	7919035	0.006952	-0.00935

Analysis:

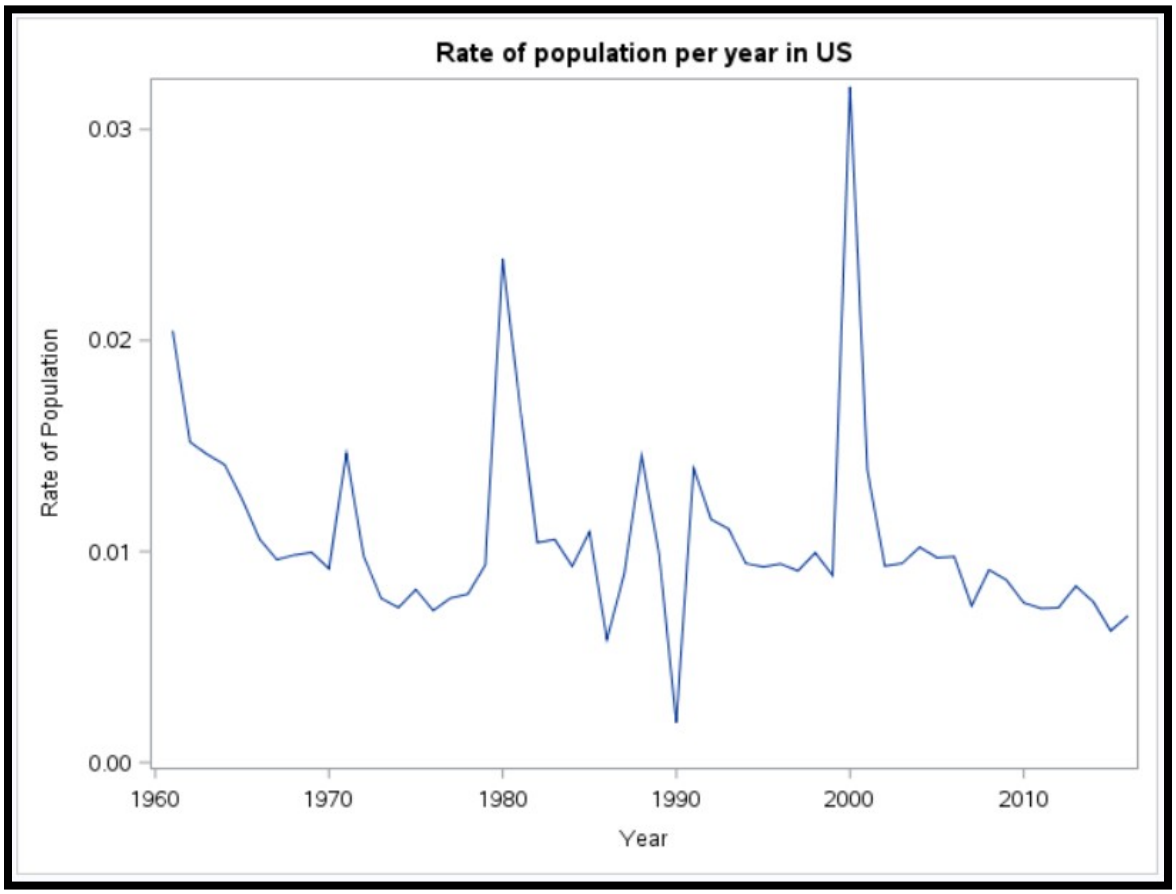
Lag(variableName) basically creates a column with values equating to variableName-1, which needs to be used to calculate the rate of population.

Data Visualization

Proc SGPlot

```
/*To plot line graph for rate of population vs year rate of total crime
vs year*/
title 'Rate of population per year in US';
proc sgplot data= Total;
      series X = Year Y = rate_Population;
run;
title;
```

Output:

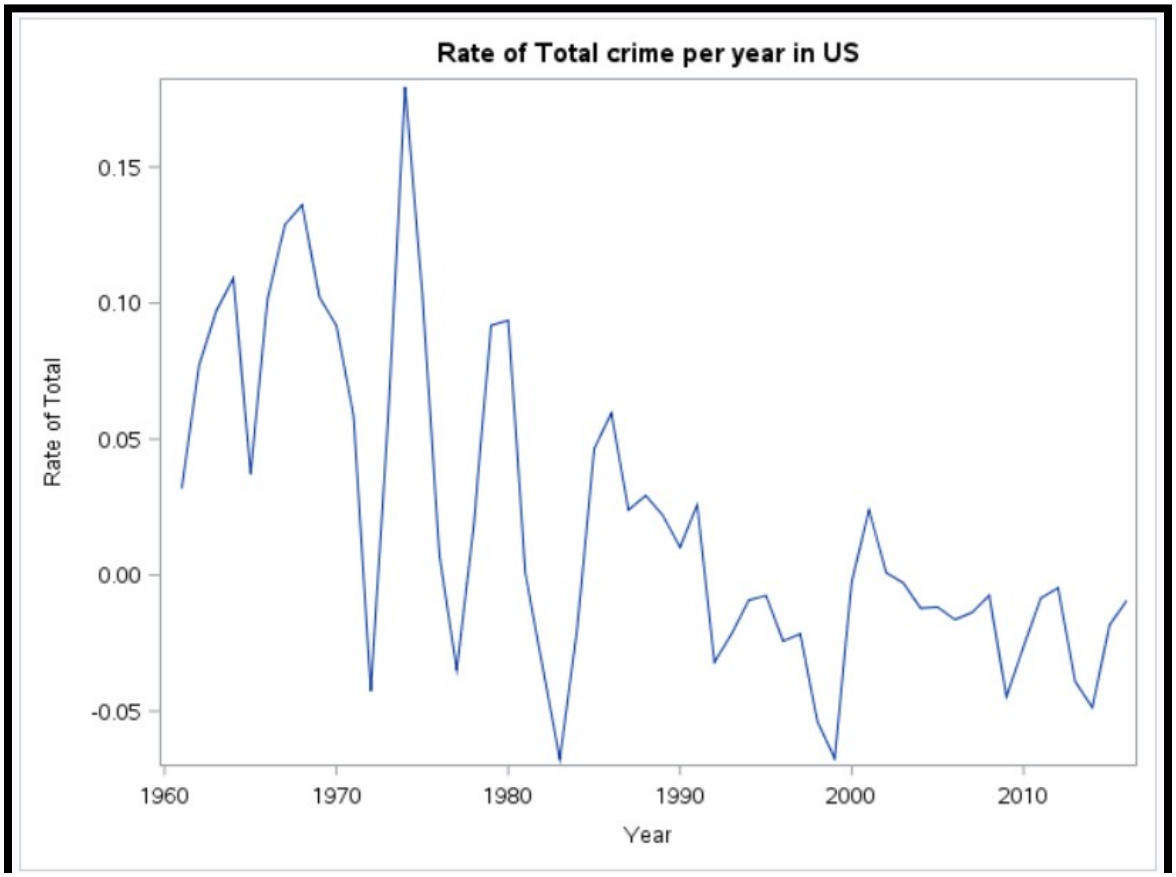


Analysis:
This plots a line graph, which tells us the rate of population of US over the past 57 years has not been a steady rise as is an expected trend for population.

```
/*To plot line graph for rate of total crime vs year*/
title 'Rate of Total crime per year in US';
proc sgplot data= Total;
      series X = Year Y = rate_Total;

run;
title;
```

Output:



Analysis:

This graph shows that there is high variability in the rate of crimes perpetrated in the US.

Data Preparation

Retain

```
/*To calculate accumulating total for each type of crime (column wise
total) and store it in a variable */
data work.PerTotal;
      set work.Total;

      retain T_Property 0 T_Murder 0 T_Rape 0 T_Robbery 0
            T_AggrevatedAssault 0 T_Burglary 0
            T_LarcenyTheft 0 T_VehicleTheft 0 T_Total 0;

      T_Property+Property;
      T_Murder+Murder;
      T_Rape+Rape;
      T_Robbery+Robbery;
      T_AggrevatedAssault+AggrevatedAssault;
```

```
      T_Burglary+Burglary;
      T_LarcenyTheft+LarcenyTheft;
      T_VehicleTheft+VehicleTheft;
      T_Total+Total;
run;
```

Analysis:
To calculate the percentage of crimes in US across the years, we would have to aggregate all the number of crimes committed each year and then deduce the percentage value of each. This step is to sum up all individual crimes since 1960.

Percentage Calculation

```
/*To find out the percentage of overall crime for each crime type*/
data work.CrimePer;
      set PerTotal;
      where Year=2016;

      Per_Property = T_Property/T_Total*100 ;
      Per_Murder = T_Murder/T_Total*100;
      Per_Rape = T_Rape/T_Total*100;
      Per_Robbery = T_Robbery/T_Total*100;
      Per_AggrevatedAssault = T_AggrevatedAssault/T_Total*100;
      Per_Burglary = T_Burglary/T_Total*100;
      Per_LarcenyTheft = T_LarcenyTheft/T_Total*100;
      Per_VehicleTheft = T_VehicleTheft/T_Total*100;
run;
```

Analysis:
The above snippet of code calculates the percentage of each crime out of the total.

Proc Format

```
/*Creating user define format*/
proc format;
      value populationPercent 0 - 10 = "low"
                              11 - 30 = "medium"
                              31 - 100 = "high";

run;
```

Proc Transpose

```
/*Transposed total percentage of all crimes(columns)
We can see in all these years percentage of property crime is highest*/
```

```
proc transpose data=work.CrimePer(keep= Per_Property Per_Murder
Per_Rape Per_Robbery Per_AggrevatedAssault Per_Burglary
Per_LarcenyTheft Per_VehicleTheft)
                        out=work.transposeCrimePer
(rename=(col1=PercentageOfCrimes)) name=TypesOfCrime ;
      format PercentageOfCrime populationPercent.;
run;
```

Analysis:

By using transpose, we are clearly able to get a table of all percentages with respect to the crime. This can be further used to pick the highest number of crimes committed in US across all crime types.

```
/*To check the details of variables*/
proc contents data=work.transposeCrimePer;
run;
```

Proc Print with Label and Title

```
/*Printing Total percentage of each crime*/
title "Total percentage of each crime";
proc print data= work.transposeCrimePer label;
      Label TypesOfCrime="Types of Crimes"
PercentageOfCrimes="Percentage of Crimes";
run;
title;
```

Proc Sort

```
/*To sort data to determine top 3 highest crimes committed in US*/
proc sort data= work.transposecrimeper;
      by descending PercentageOfCrimes;
run;
```

Conclusion

The three highest number of crime types committed come out to be Property, Larceny Theft and Burglary. All three are crimes that are perpetrated to gain from the victim, in terms of value acquisition, rather than just causing physical harm. It can be deduced that the primary reason for committing a crime is to accomplish an ulterior motive.

These are closely followed by Vehicle Theft, which is another example of delinquency for personal gain.