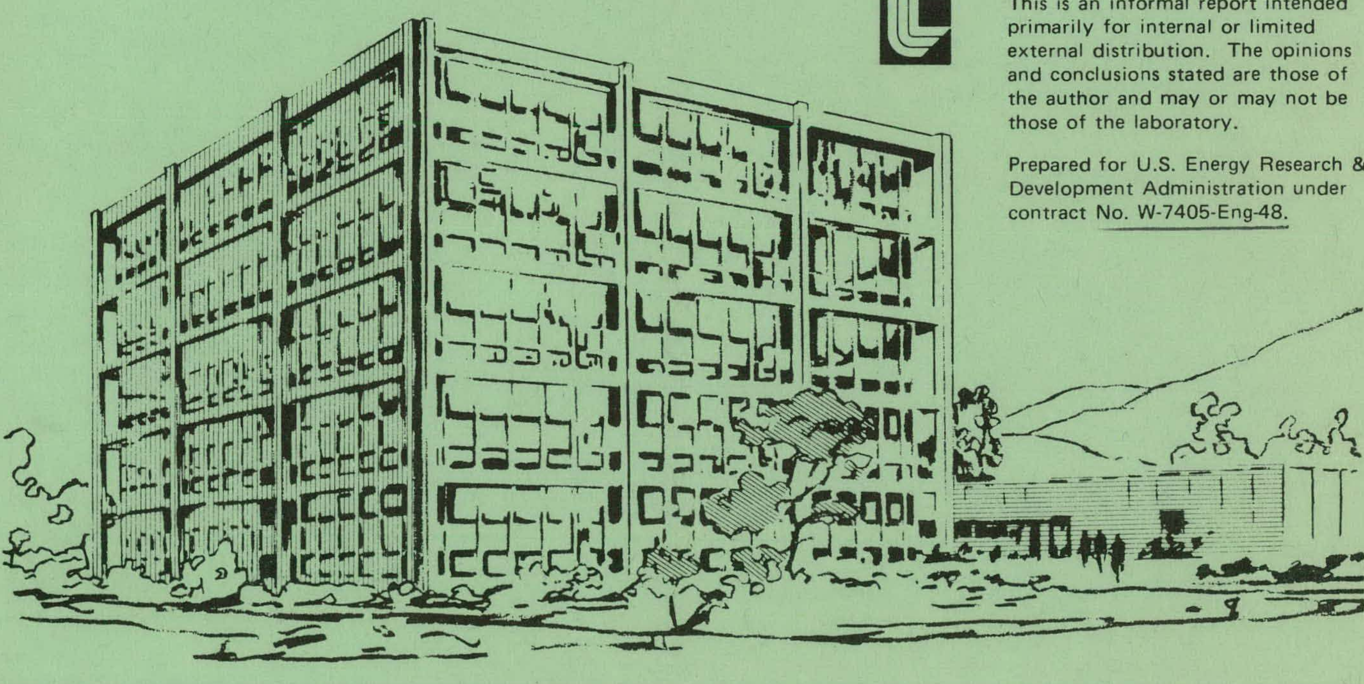
UCID- 17224

Lawrence Livermore Laboratory

An Application of Pattern Recognition Techniques to Crime Analysis

C. F. Bender, L. A. Cox, Jr., and G. A. Chappell

August 15, 1976



This is an informal report intended primarily for internal or limited external distribution. The opinions and conclusions stated are those of the author and may or may not be those of the laboratory.

Prepared for U.S. Energy Research & Development Administration under contract No. W-7405-Eng-48.

AN APPLICATION OF PATTERN RECOGNITION TECHNIQUES TO CRIME ANALYSIS*

C. F. Bender, L. A. Cox, Jr., and G. A. Chappell
University of California
Lawrence Livermore Laboratory
P. O. Box 808
Livermore, California 94550

*This work was performed under the auspices of the U. S. Energy Research and Development Administration under contract No. W-7405-Eng-48.

AN APPLICATION OF PATTERN RECOGNITION TECHNIQUES TO CRIME ANALYSIS

Introduction

In October of 1975 the San Diego Police Department received a grant from the Law Enforcement Assistance Administration for the construction of an Automated Regional Justice Information System (ARJIS), a computerized data processing system to serve all of San Diego County. One of the goals of this system was the implementation of an automated crime analysis capability. In late 1975 the Lawrence Livermore Laboratory was requested to assist the ARJIS project personnel under the auspices of the National Technology Transfer Program.

The initial goal of LLL's participation was to evaluate the capabilities of current pattern recognition techniques when applied to existing computerized crime data. Performance was to be evaluated both in terms of the system's capability to predict crimes and to optimize police manpower allocation.

The Data Analysis Project

The Data

In early 1976 we received a magnetic tape containing 21056 lines of data. Each data line represented one crime report and contained the following information fields:

Crime type (San Diego Police Department code), (see Figure 1)
Census tract number for the area of occurrence
Year of occurrence
Julian date
San Diego case number
Applicable penal code section
Street address where crime occurred
Hour of day
Status indicator reflecting status of the case (Figure 2)

Figure 1

Type of Crime Table

Code	Type of Crime
Ø1	Homicide
Ø1	Neg. Manslaughter
Ø2	Rape - Forcible
Ø2	Rape - Attempt
Ø3	Robbery - Firearm
Ø3	Robbery - Knife/Cut Inst.
Ø3	Robbery - Other Weapon
Ø3	Robbery - Strongarm
Ø4	Assault - Gun
Ø4	Assault - Knife
Ø4	Assault - Other Weapon
Ø4	Assault - Hands, Fists
Ø4	Assault - Battery
Ø4	Assault - Exhibit D/W
Ø5	Burglary - Residence (include garage)
Ø5	Burglary - Other Structure (include business)
Ø5	Burglary - Safe
Ø5	Burglary - Locked Vehicle
Ø5	Burglary - School (include parochial)
Ø5	Burglary - Church
Ø5	Burglary - Boat
Ø5	Burglary - Sporting Goods (include gun shops)
Ø6	Grand Theft - Person
Ø6	Grand Theft - Other
Ø6	Petty Theft - Other
Ø6	Petty Theft - Car Prowl
Ø6	Petty Theft - Bicycle
Ø6	Grand Theft - Bicycle
Ø7	Auto Theft - Autos
Ø7	Auto Theft - Trucks/Buses
Ø7	Auto Theft - Other

Type of Crime Table (Continued)

Code	Type of Crime
Ø8	Arson
Ø9	Forgery
1Ø	Fraudulent Checks
1Ø	Credit Card Frauds
1Ø	Other Fraud
1Ø	Embezzlement
11	Malicious Mischief
12	Deadly Weapons Act
13	Child Molest
13	Rape - Statutory
13	Perversion
13	Indecent Exposure
13	Other Sex Misdemeanors
14	Narcotics
15	Gambling - Bookmaking
16	Family & Children
17	Kidnap
18	Lost Article
19	Missing Adult
20	Death - Natural
21	Suicide
21	Attempted Suicide
22	Other Misdemeanors
23	Miscellaneous
24	Contrib. to Del. of Minor
25	Missing Juvenile

Figure 2

Case Status Indicators

Status	Meaning
1	Arrest (Adult)
2	Arrest (Juvenile)
3	Exception
4	Property Recovered Only (stolen locally, recovered locally)
5	Property Recovered Only (stolen locally, recovered elsewhere)
6	Closed
7	Unfounded
8	Property Recovered Only (stolen elsewhere, recovered locally)
L	Latent Prints are on File

Data were formatted and (1) tape anomalies, (2) cases whose status was "unfounded," and (3) cases with no certain time of day were eliminated. At this point 11739 cases remained. These data were reformatted into records with the following structure:

1. Hour of Day
2. Digitized Year/Day
3. East/West Coordinates
4. North/South Coordinates
5. Predicted Property: 1 = Case Closed
   Ø.5 = Property Recovered
   Ø = Case Open
6. Class Name (Open or Closed)

The geographical coordinates were obtained by digitizing the census tract map in arbitrary units {see Figure 3 (map)}. At this point, the data were in suitable format to begin pattern recognition analysis using the LLL PATTER* system.

The Role of Pattern Recognition

Pattern recognition functionally consists of a collection of techniques which have been developed to assist in the solution of general problems such as:

> ". . . given a set of objects and a list of measurements made on these objects, is it possible to find and/or predict a property of the objects which is not directly measurable but is known to be related by some unknown relationship?"**

The techniques employed draw heavily upon work in the fields of statistics, applied mathematics, and computer science. The PATTER system is a polyalgorithmic computer program which has been specifically designed to answer

*L. A. Cox, Jr. and C. F. Bender, "PATTER: A Polyalgorithm for the Analysis of Generalized Data Sets," LLL publication UCID-16915, October 1975.

**B. R. Kowalski and C. F. Bender, J. Am. Chem. Soc. 94, 3632 (1972).

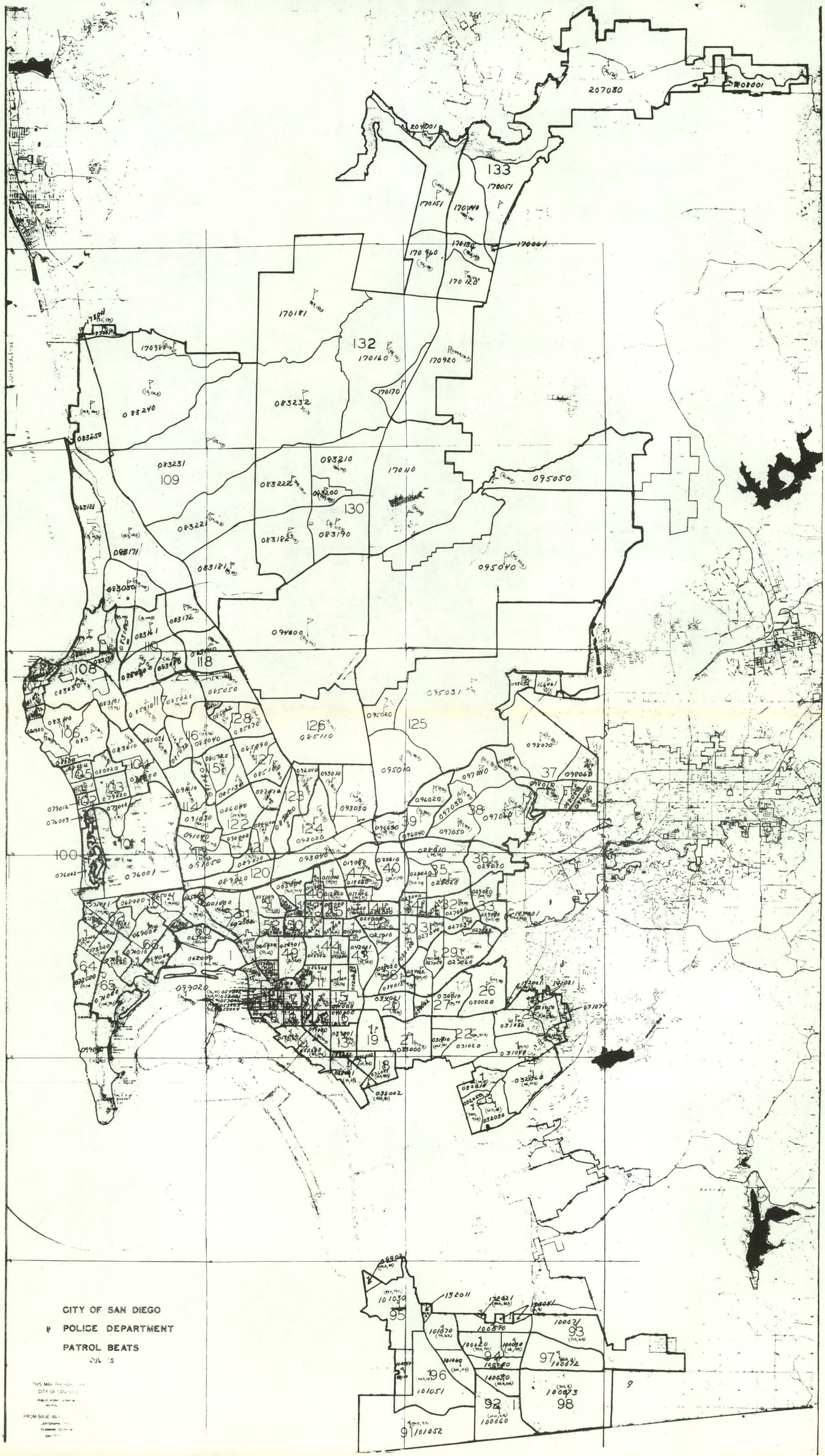


Figure 3

the above question for given sets of data. We have intentionally designed the system to be able to not only answer "yes" but to answer "no" when such a response is warranted.

The goal of our pattern recognition experiment was to first determine if some relationship existed which would allow us, based upon the knowledge of a crime's type, location, time, etc., to predict the crime's susceptibility to solution. If we found that such a relationship existed, our second goal was to attempt to determine or at least better understand this relationship. It was on this basis that we assigned the predicted property to each case as mentioned previously.

Simple Statistics

The initial analysis considered only the simple statistics inherent in the data. Figure 4 graphically shows the distribution of the data by type of crime. Figure 5 tabulates this distribution and shows the breakdown of crimes for each day of the week.

A study of the highest crime area in San Diego was performed, and the results are shown in Figure 6 and can be compared with the city totals in Figure 5.

The data, with digitized location coordinates, were plotted in three dimensions against an outline of the city itself. Two views of the cumulative total (all crimes) are shown in Figures 7 and 8. Individual plots for each crime type are contained in Appendix I. A computer-generated movie showing various views of the cumulative crimes was generated to fully display the geographical distribution.

An in-depth statistical analysis was performed on data from selected crimes which included assault, robbery, and burglary. Population histograms for each variable were generated. These graphics in addition to F-Test results, least squares curve fitting, and cross-correlation analysis output are shown in Appendix II.

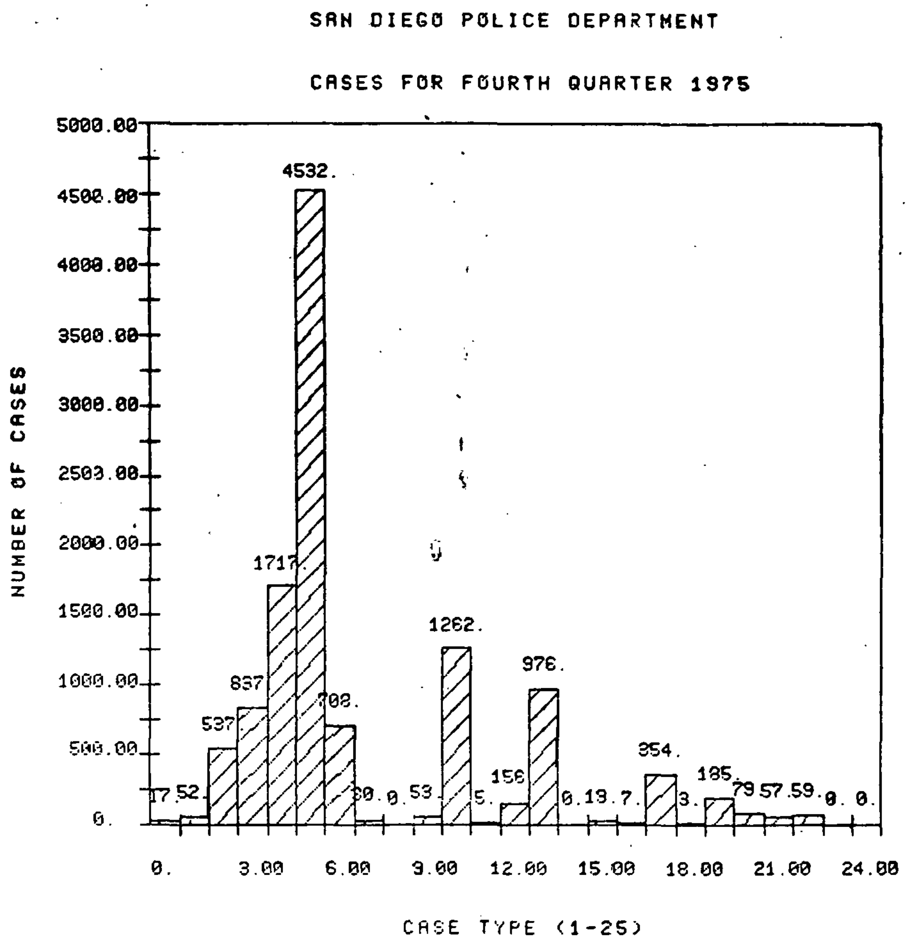

SAN DIEGO POLICE DEPARTMENT
CASES FOR FOURTH QUARTER 1975
NUMBER OF CASES
CASE TYPE (1-25)
5000.00
4500.00
4000.00
3500.00
3000.00
2500.00
2000.00
1500.00
1000.00
500.00
0.
0.
3.00
6.00
9.00
12.00
15.00
18.00
21.00
24.00
4532.
1717.
837
537.
708.
1262.
976.
354.
185.
156
79
57
59.
52.
17
53.

Figure 4

Criminal Cases
San Diego Police Department
Fourth Quarter 1975

Crime Type	Crime Description	Total No. of Cases	Day of Week M	T	W	Th	F	S	S
1	Homicide	17	1	2	3	3	4	1	3
2	Rape	52	7	7	6	8	6	11	7
3	Robbery	537	70	73	69	69	93	71	92
4	Assault	837	116	118	103	131	140	89	140
5	Burglary	1717	261	237	279	239	233	215	253
6	Grand Theft	4532	637	603	662	681	712	591	646
7	Grand Theft - Auto	708	91	95	97	105	123	93	104
8	Arson	30	1	5	8	2	2	8	4
9	Forgery	0	0	0	0	0	0	0	0
10	Fraud	53	4	12	8	4	9	8	8
11	Mischief	1262	156	189	187	180	179	183	188
12	Deadly Weapons Act	5	0	1	1	1	2	0	0
13	Sex Offenses	156	24	27	26	26	18	15	20
14	Narcotics	976	109	122	121	165	150	159	150
15	Gambling-Bookmaking	0	0	0	0	0	0	0	0
16	Family and Children	19	1	4	2	3	0	3	6
17	Kidnap	7	1	2	2	1	1	0	0
18	Lost Article	354	51	46	58	52	41	49	57
19	Missing Adult	3	0	1	1	0	1	0	0
20	Natural Death	185	30	29	28	22	26	23	27
21	Suicide	79	10	10	8	14	14	10	13
22	Other Misdemeanors	57	6	9	9	8	12	7	6
23	Miscellaneous	59	8	10	12	9	9	2	9
24	Contributing to Delinquency of Juvenile	0	0	0	0	0	0	0	0
25	Missing Juvenile	0	0	0	0	0	0	0	0
	Totals	11645	1584	1602	1690	1723	1775	1538	1733

Figure 5

Figure 6

Analysis of Highest Crime

Area in San Diego

Census Tracts Encompassed: 051001, 052000, 053003, 054001, 046000, 047000

Crime Type	Occurrences in Area	% of Total Occurrences	% of that Type of Crime
1	2	0.28	11.8
2	3	0.42	5.8
3	32	4.5	6.0
4	58	8.2	6.9
5	70	9.8	4.1
6	273	38.4	6.0
7	51	7.2	7.2
8	1	0.14	3.3
10	5	0.70	9.4
11	48	6.75	3.8
12	1	0.14	20.0
13	2	0.28	1.3
14	104	14.6	10.6
16	1	0.14	5.3
18	39	5.5	11.0
20	14	2.0	7.6
21	4	0.56	5.1
22	2	0.28	3.5
23	1	0.14	1.7

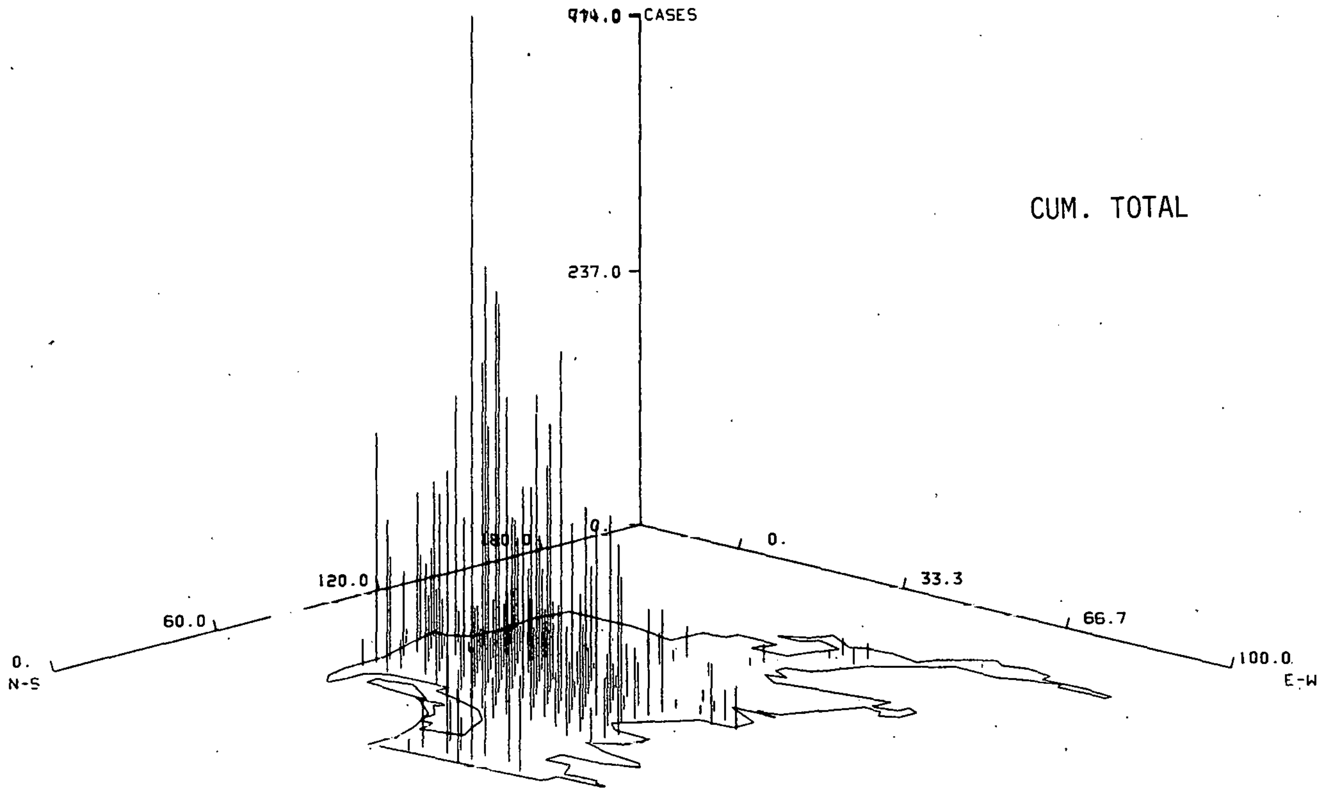
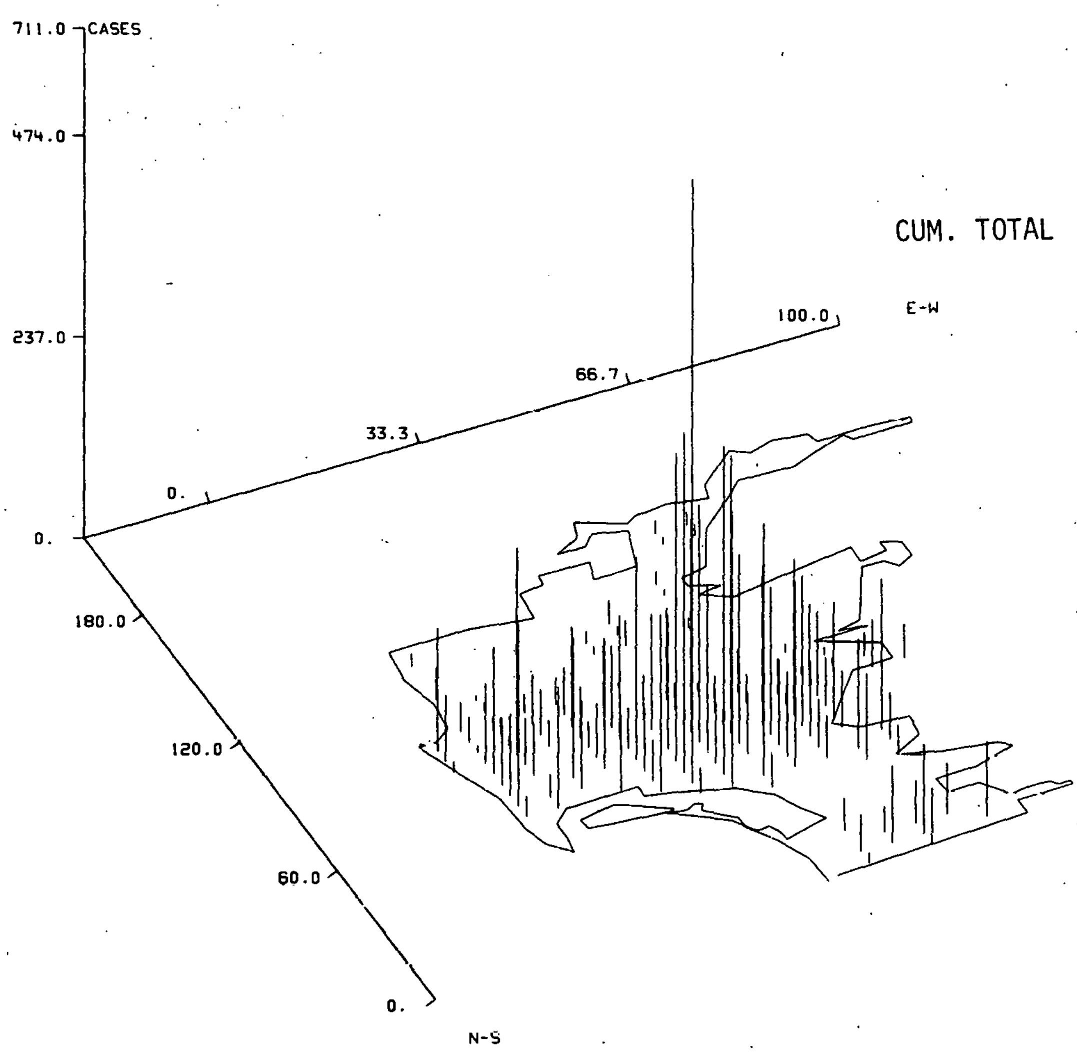


Figure 7



711.0
CASES
474.0
237.0
0.
CUM. TOTAL
100.0
E-W
66.7
33.3
0.
180.0
120.0
60.0
0.
N-S

Figure 8

Pattern Recognition

The initial efforts were designed to test the efficiencies of the various techniques in prediction of class membership of newly-reported crimes based on inferences from an historical "training set" which contained a representative sample of crimes whose solution status was known. Since the predicted property and the class name reflected similar information, both continuous property estimation techniques and conventional classification techniques were employed. The results of the first attempts in attempting to predict a crime's solution using the four basic variables of time of day, time of year, north/south, and east/west locations were unimpressive, and all techniques appeared to return essentially the statistics of the training set. Typical results are shown in Figure 9, which shows the results for Crime Type 3, Robbery.

A second experiment was designed to investigate the possibility of optimizing manpower allocations between crimes of two types. Two classes of crimes were considered together. The training set contained approximately equal numbers of each crime type and within each crime equal numbers of OPEN and CLOSED cases were included. An additional variable was included to denote crime class. The goal of this experiment was to test the feasibility of automatically determining which cases had the highest probability of closure.

Initial results of the second experiment are shown in Figure 10. As shown in part III, after autoscaling the variables with weighting proportional to the absolute value of the correlation coefficient, a 7% advantage could be gained over the sample statistics. While this showed promise, we realized that there were a number of useful factors which were implicit in our data which might significantly improve our predictive capability.

At the same time, we realized that while the X-Y coordinates had been included, they were not truly independent (since the shape of the city is not regular and the population distribution is not uniform) and that there was no direct measure of interpoint distances, i.e., a variable proportional to the cross product of the X and Y measures. A new data set was created which contained 100 cases. Of these 100 crimes, we included 20 each from rape, robbery, assault, burglary, and grand theft.

Figure 9

Initial Test Results

I. ROBBERY (Code Ø3) 266 cases considered, selected to give approximately equal class populations.
OPEN = 144 cases (54%)
CLOSED = 122 cases (46%)

a. K-Nearest Neighbor Technique

K	1	3	5	7
% Correct	56.4	54.5	53.4	55.7

b. MULTICLASS Linear Separation Technique

% Correct - 50.75

c. Select Optimized K-Near Technique (2 variables)

K	1	3	5	7
% Correct	46.2	51.1	53.0	51.5

d. MULTICLASS (2 variables selected data)

% Correct - 54.14

Figure 10

Manpower Allocation Test Results

Considering Robberies, Assaults, each individually balanced for approximately equal numbers of open and closed cases.

688 Cases % Closed = 50.6

I. K-Near Performance

K	1	3	5	7
% Correct	55.5	56.1	53.8	51.0

II. MULTICLASS

% Correct = 53.92

III. Autoscaling (Weight α|corr. coef.|)

K	1	3	5	7
% Correct	57.4	57.0	56.2	56.4

MULTICLASS

% Correct = 53.92

For each case, we included three additional variables to make a total of seven variables and one hundred patterns. The variables used in this test were:

Variable 1	hour of the day when crime was reported
Variable 2	time of the year
Variable 3	east/west coordinate (X)
Variable 4	north/south coordinate (Y)
Variable 5	X*Y (cross product of X and Y)
Variable 6	% of that type case which were closed in the original sample
Variable 7	% of that type case in the original sample

Results of pattern recognition analysis of this data set are shown in Figure 11.

Here we realize a significant improvement over the statistics. The MULTICLASS linear separation technique achieved a 65% prediction accuracy - 15% better than the input statistics. Continuous techniques such as BUTLER provided only 55% accuracy, which was significantly less than the conventional classification techniques.

The significance of the 65% accuracy of MULTICLASS is two-fold. First, in terms of operational impact, it is possible to predict the probability of a crime's solution with significantly greater accuracy than a random guess, even if this guess is based on the sample's statistics. This in effect implies that it is possible to optimize police investigative manpower assignments through the use of this predictive capability to assign manpower to those cases most susceptible to solution.

A second significant result of this performance is the improvement shown by the inclusion of the three additional variables. While variable 5, the cross product term, is relatively straightforward, it is clear that variables 6 and 7 are functions of a number of other variables which

Figure 11

Expanded Variables - Test Results

100 crimes, 20 each (10 open, 10 closed) from:

RAPE

ROBBERY

ASSAULT

BURGLARY

GRAND THEFT

Variable 1. Hour of Day

2. Time of Year

3. East/West (X)

4. North/South (Y)

5. XY

6. % Closed Originally of that Type

7. % of that Type in Original Sample

Number of Crimes Closed = 50

K-Near Results

K	1	3	5	7
% Correct	63	58	57	56

MULTICLASS - 65% correct

BUTLER - 55% correct

were not individually available in the initial data set. For example, the percentage of cases closed for a given type of crime (variable 6) depends on a number of other factors. If some of these could be isolated and quantized, we would expect greater increases in our predictive capability.

Conclusions and Recommendations

In late June 1976, a meeting was held between the authors and representatives of the San Diego Police Department to discuss the preliminary results presented in this report and to consider what further actions, if any, were appropriate. Several questions were raised which appear to be technically feasible and whose solutions are of practical importance.

Of primary importance, since it directly drives the design of justice information systems design, is the question of variables of significance. We have seen that with the limited variables presently available some degree of analysis is possible using pattern recognition techniques. We have also seen that performance increases as variables related to the probability of solution and crime type distribution are included. What then are the variables which are most important? We, in conjunction with the San Diego Police Department, intend to analyze the effect of adding additional variables currently collected but not regularly stored in automatic information retrieval systems.

Also of great practical significance is the question of manpower allocation. Crimes have been described as either "suppressible" or "non-suppressible." In those instances of non-suppressible crimes, the pattern recognition techniques could be used to optimize the assignment of investigative personnel in order to solve the greatest number of crimes. In cases of "suppressible" crimes, pattern recognition's predictive techniques can be used to assign patrol units in order to minimize actual occurrences of these crimes. Clearly this subject warrants further investigation.

Since the results of this preliminary study indicate that automatic crime analysis involving pattern recognition techniques is feasible, a further effort to determine optimum variables and techniques appears warranted. On the basis of follow-on research, an automated system could be implemented and controlled tests of these procedures in an operating environment could be attempted.

APPENDIX I

GEOGRAPHICAL PLOTS BY CRIME TYPE

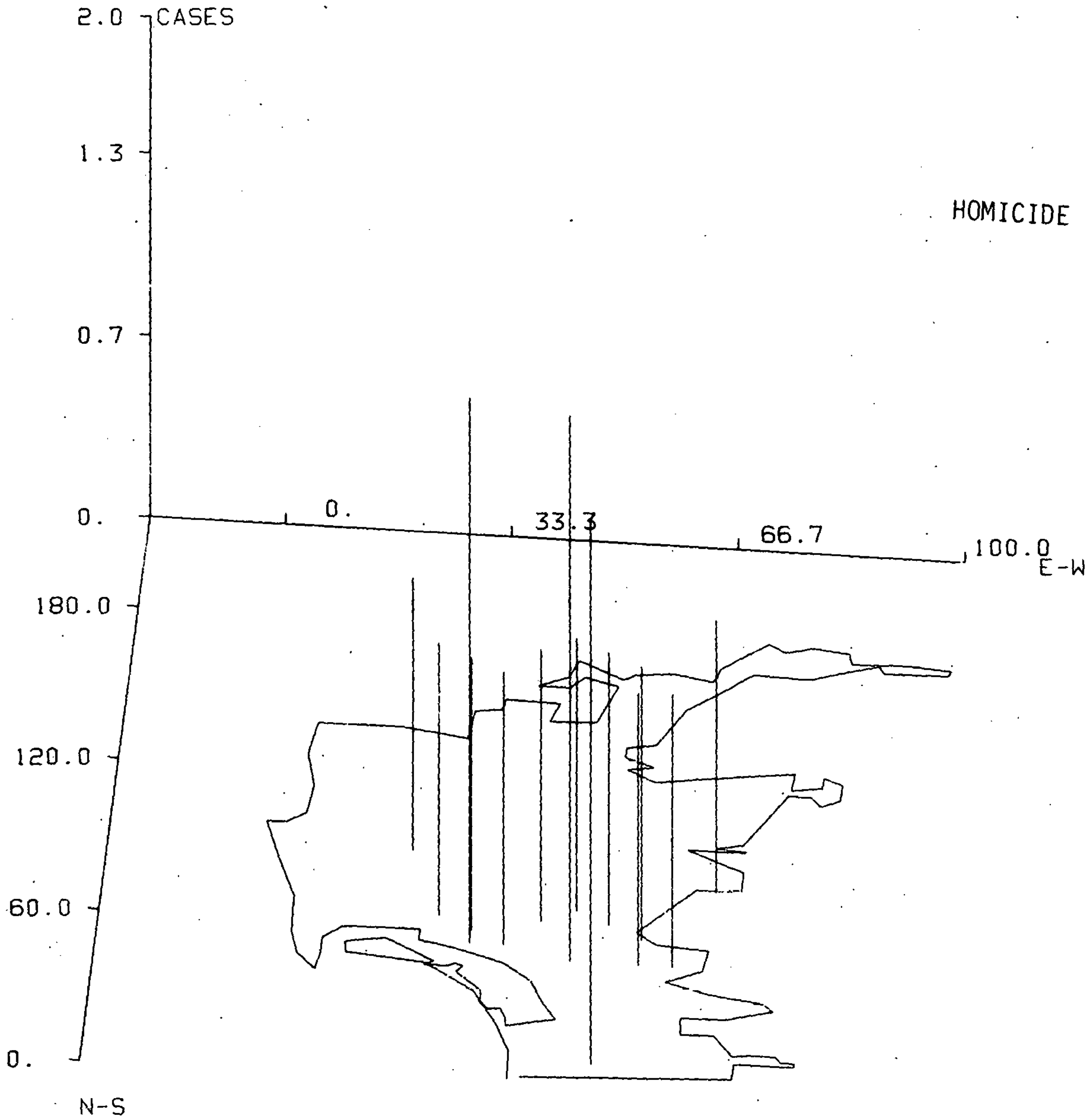


Figure A1-1

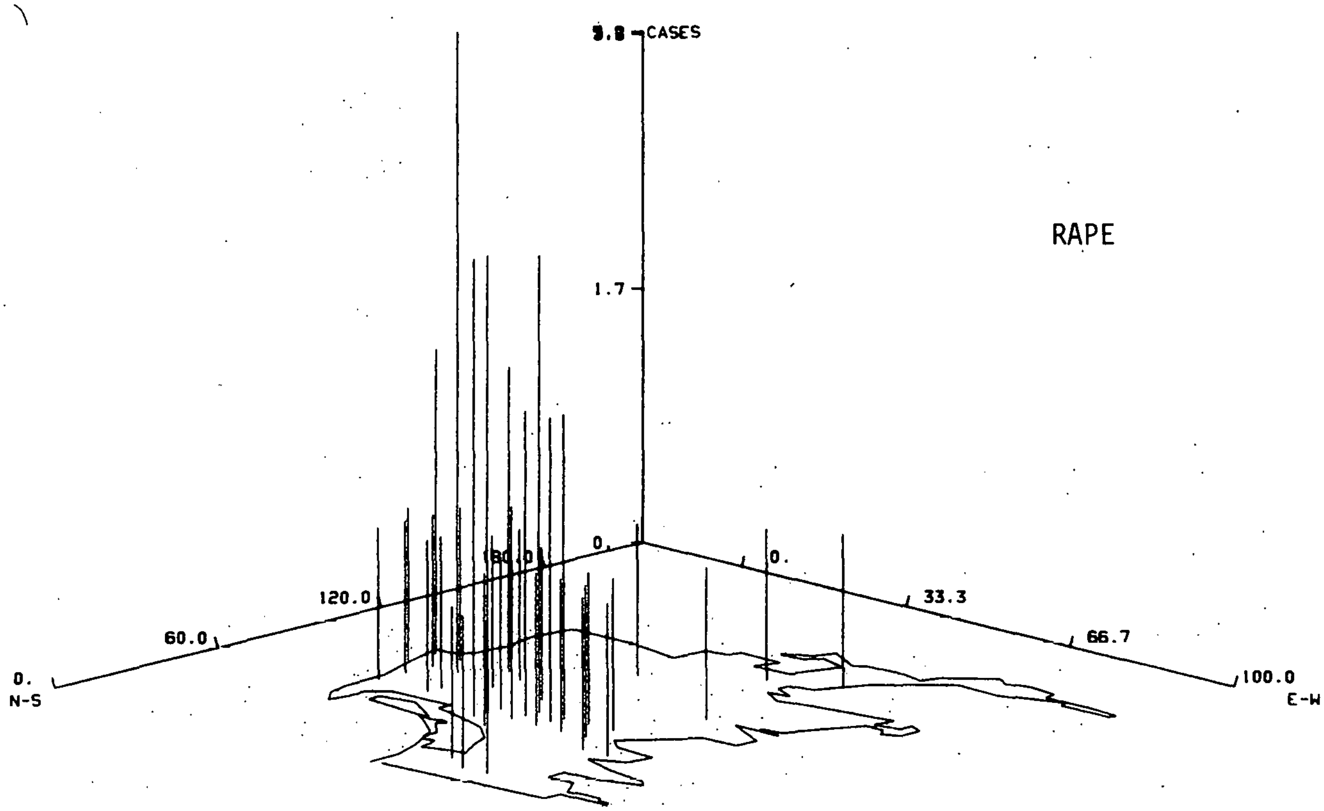


Figure A1-2

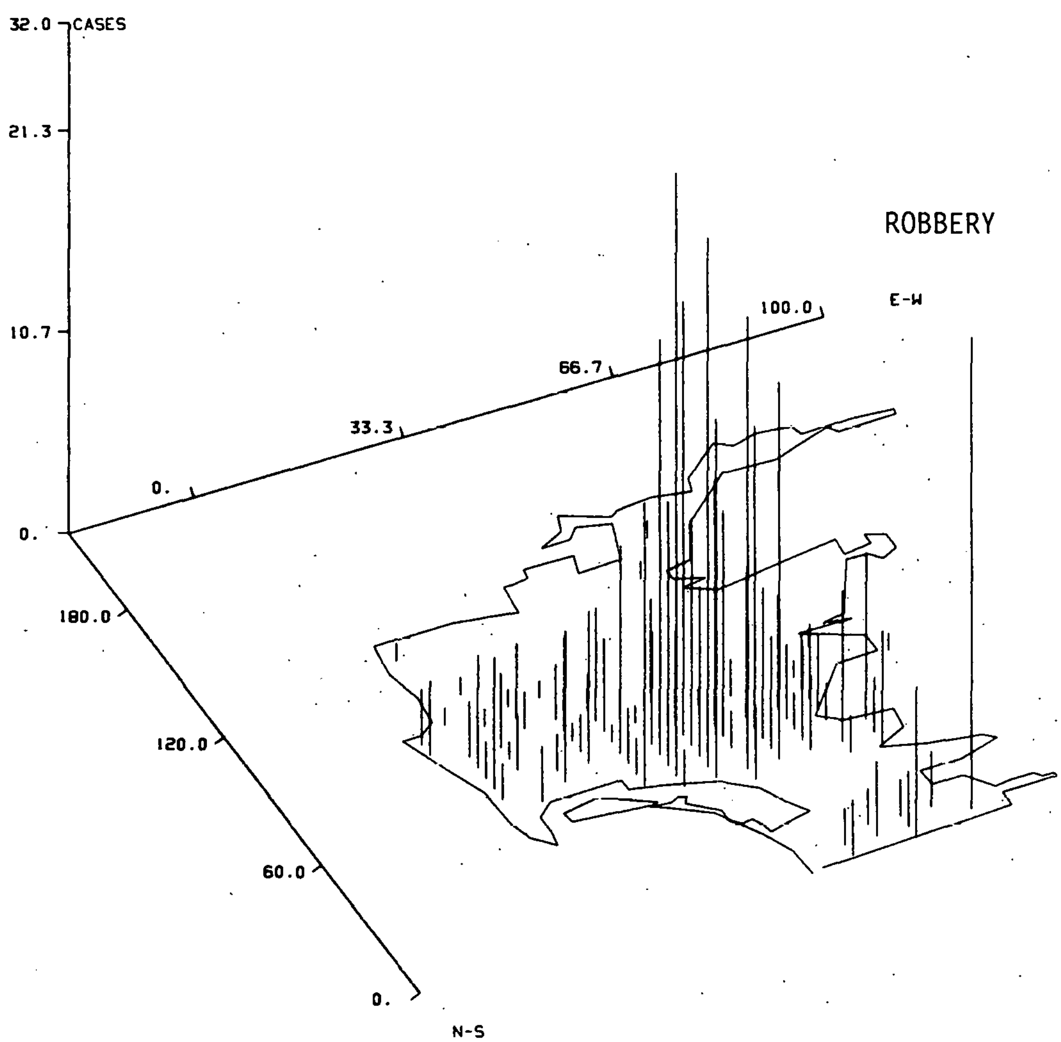


Figure A1-3

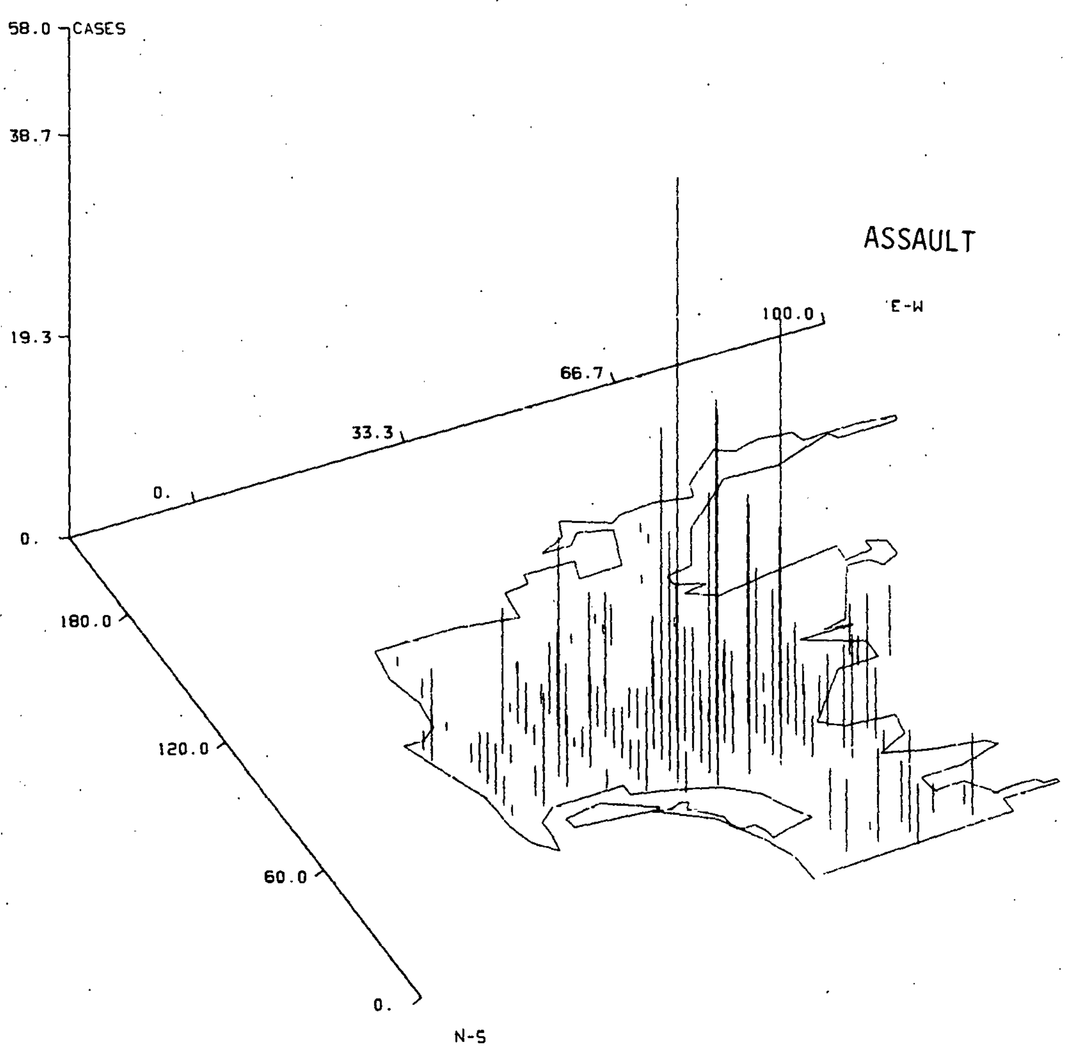


Figure A1-4

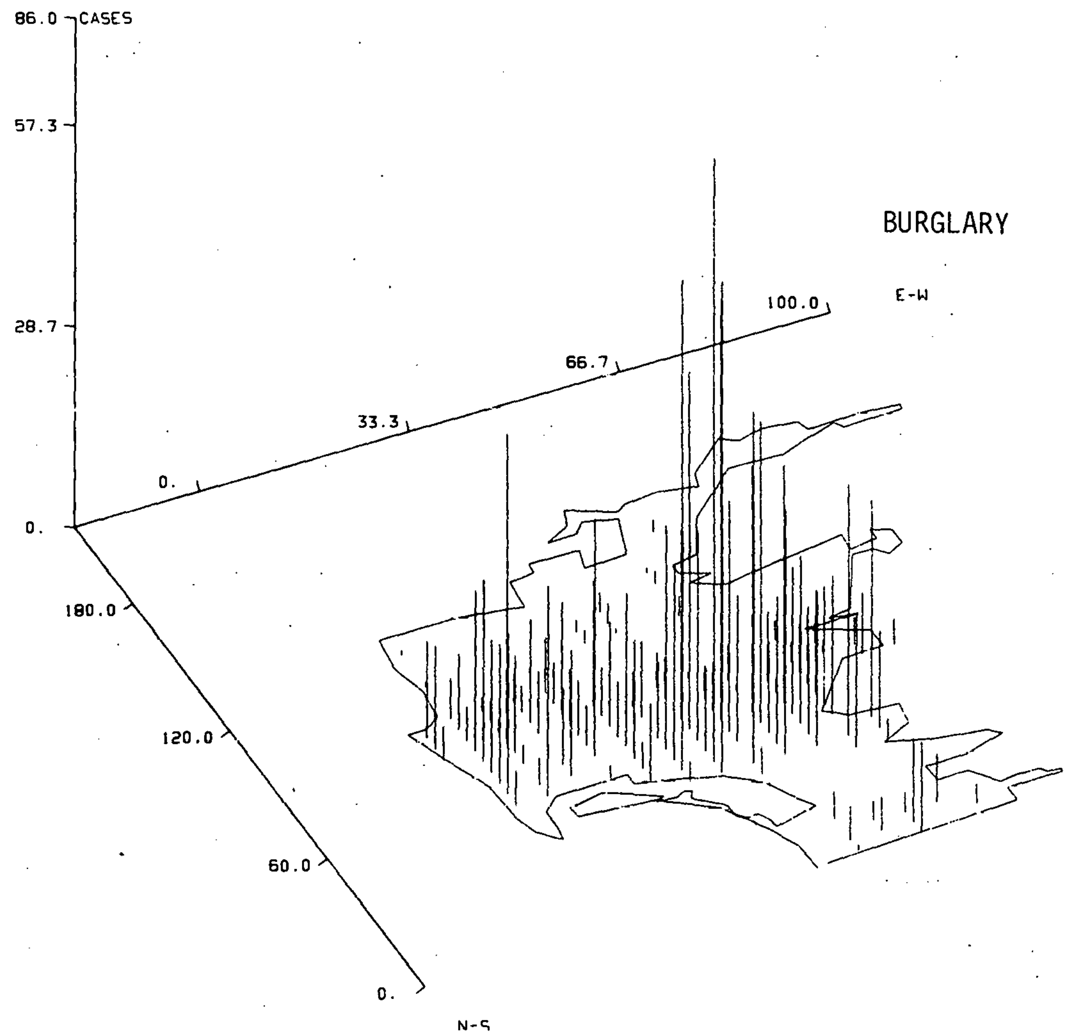
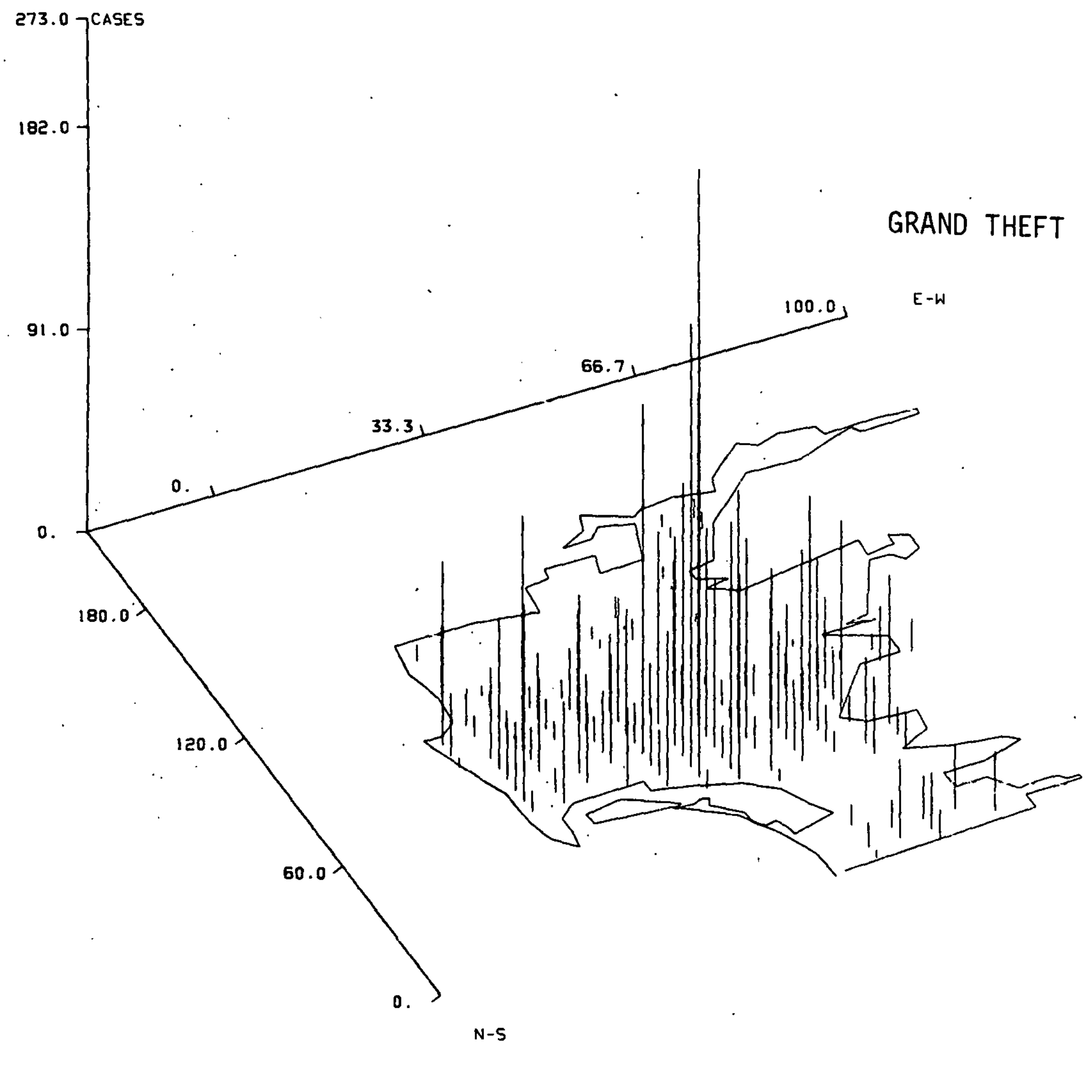

86.0
CASES
57.3
28.7
0.
BURGLARY
100.0
E-W
66.7
33.3
0.
180.0
120.0
60.0
0.
N-S

Figure A1-5



273.0
CASES
182.0
91.0
0.
GRAND THEFT
100.0
E-W
66.7
33.3
0.
180.0
120.0
60.0
0.
N-S

Figure A1-6

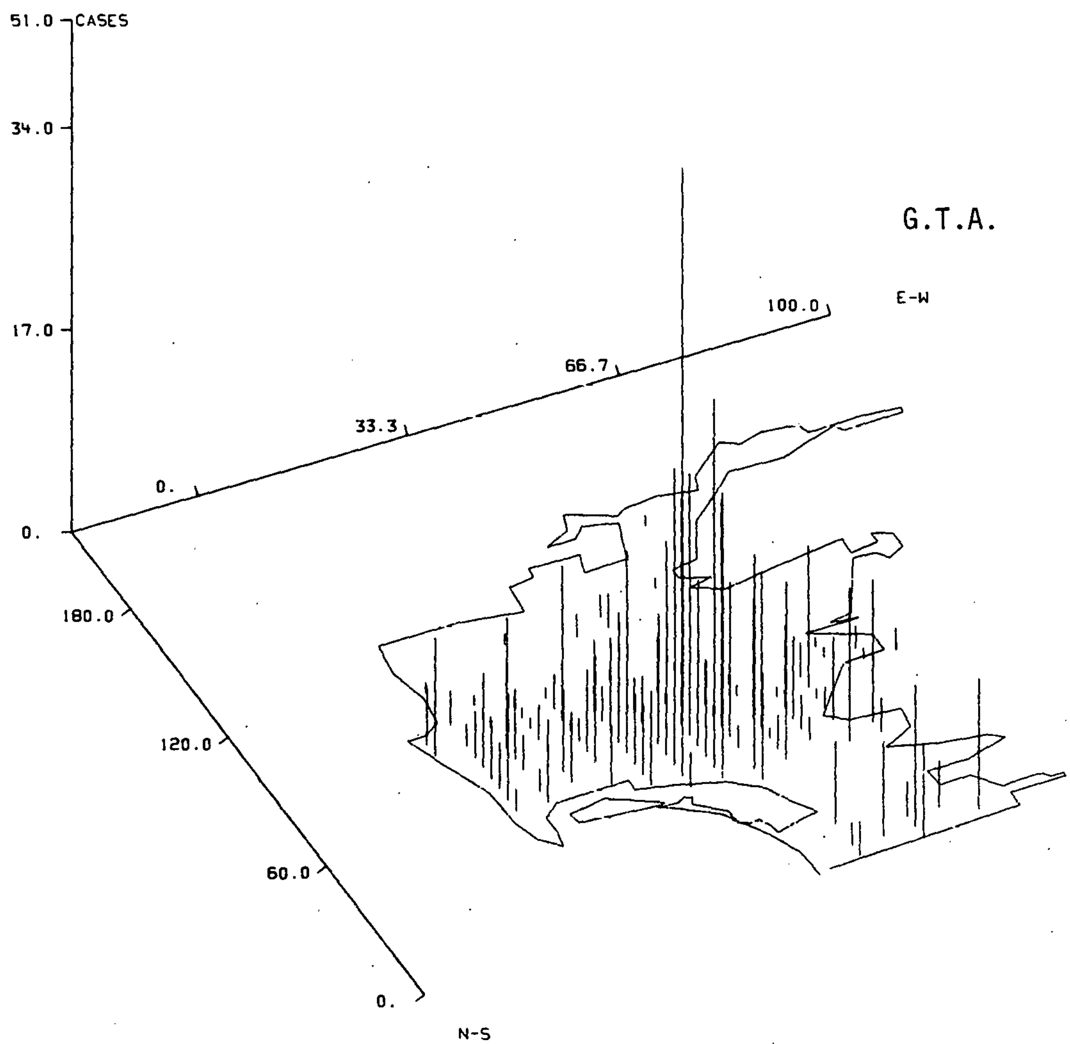


Figure A1-7

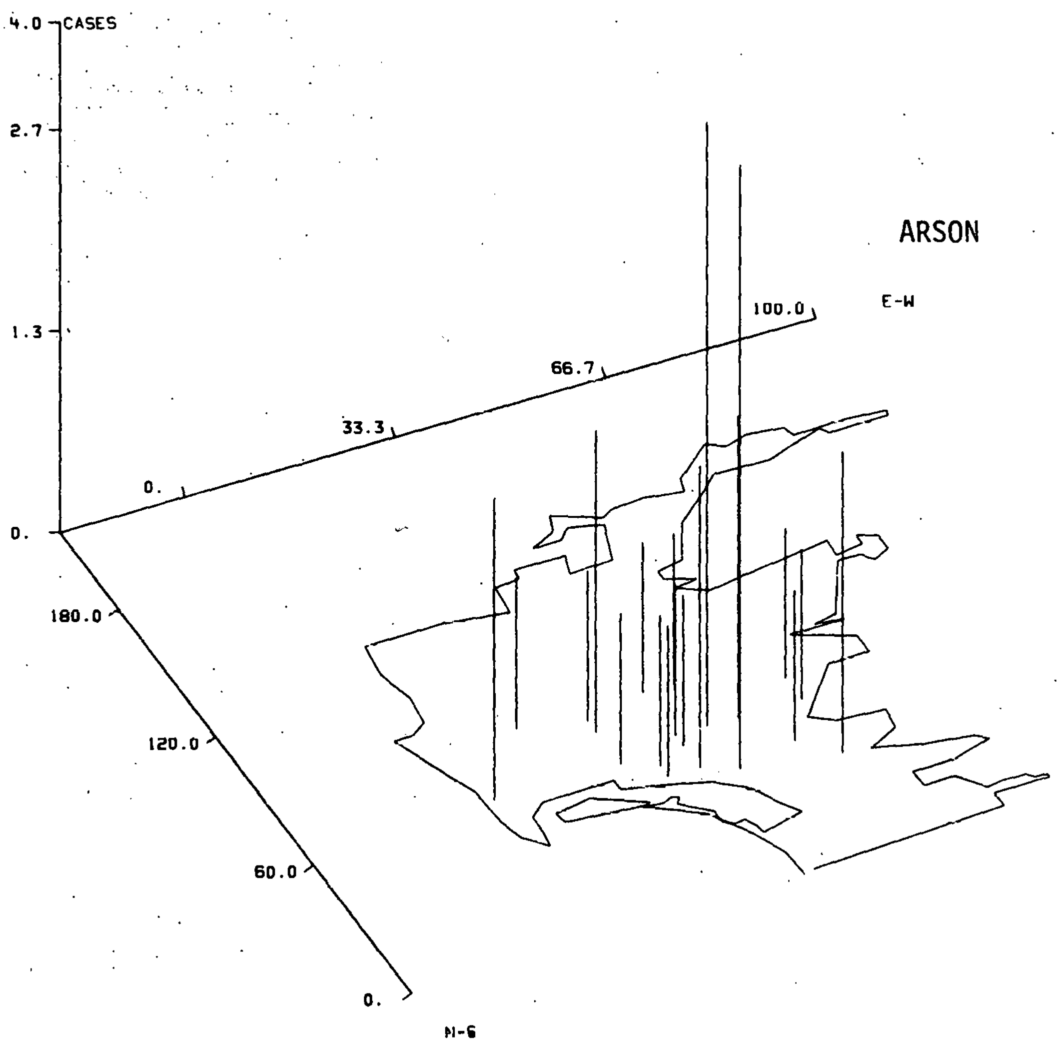


Figure A1-8

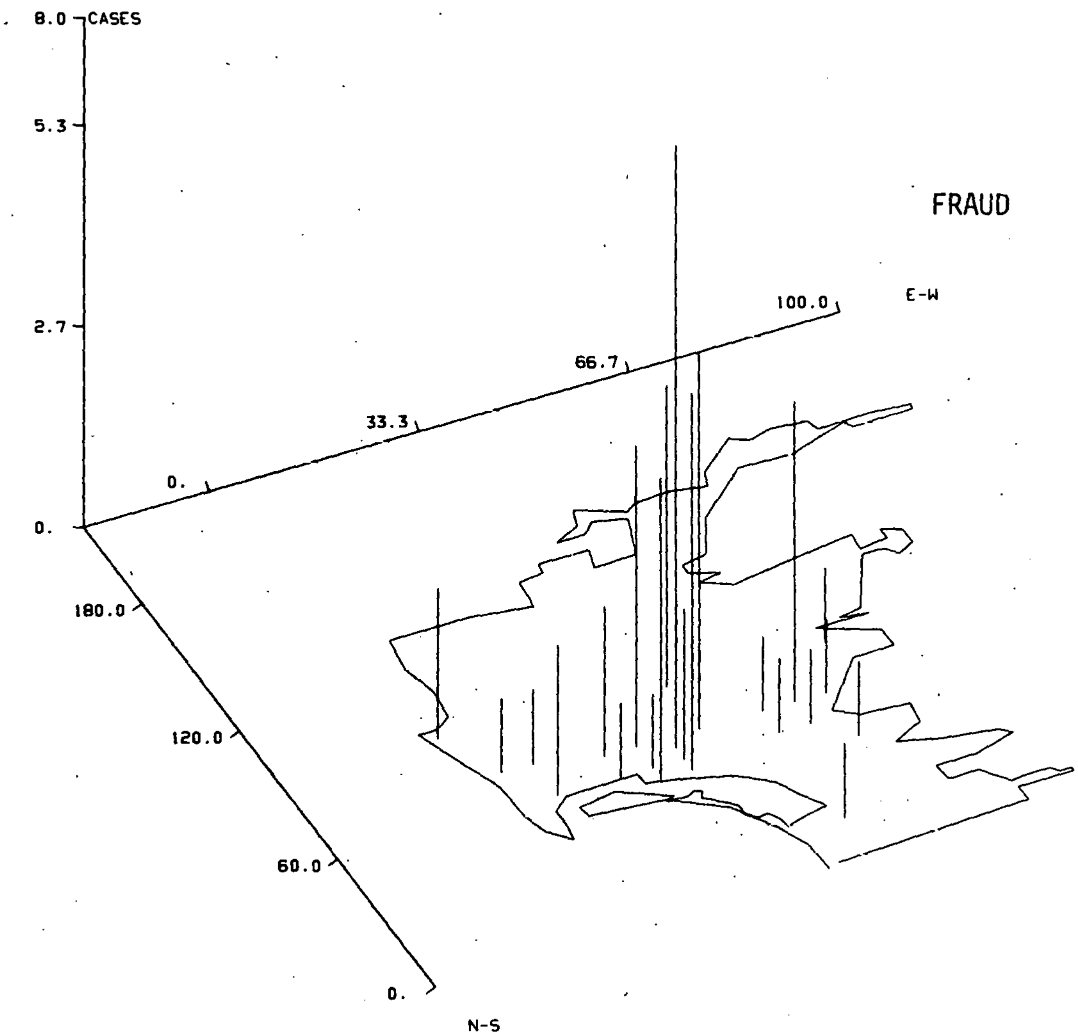


Figure A1-9

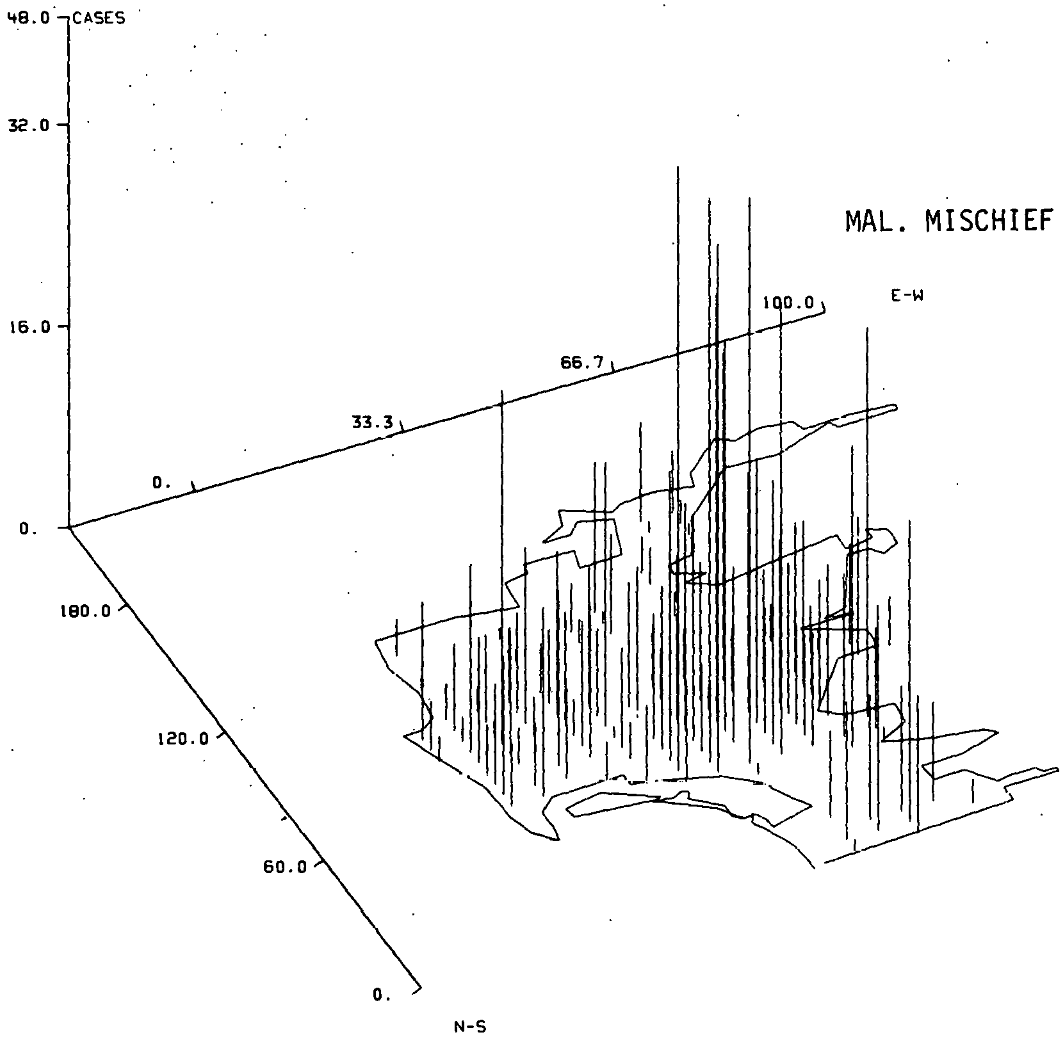
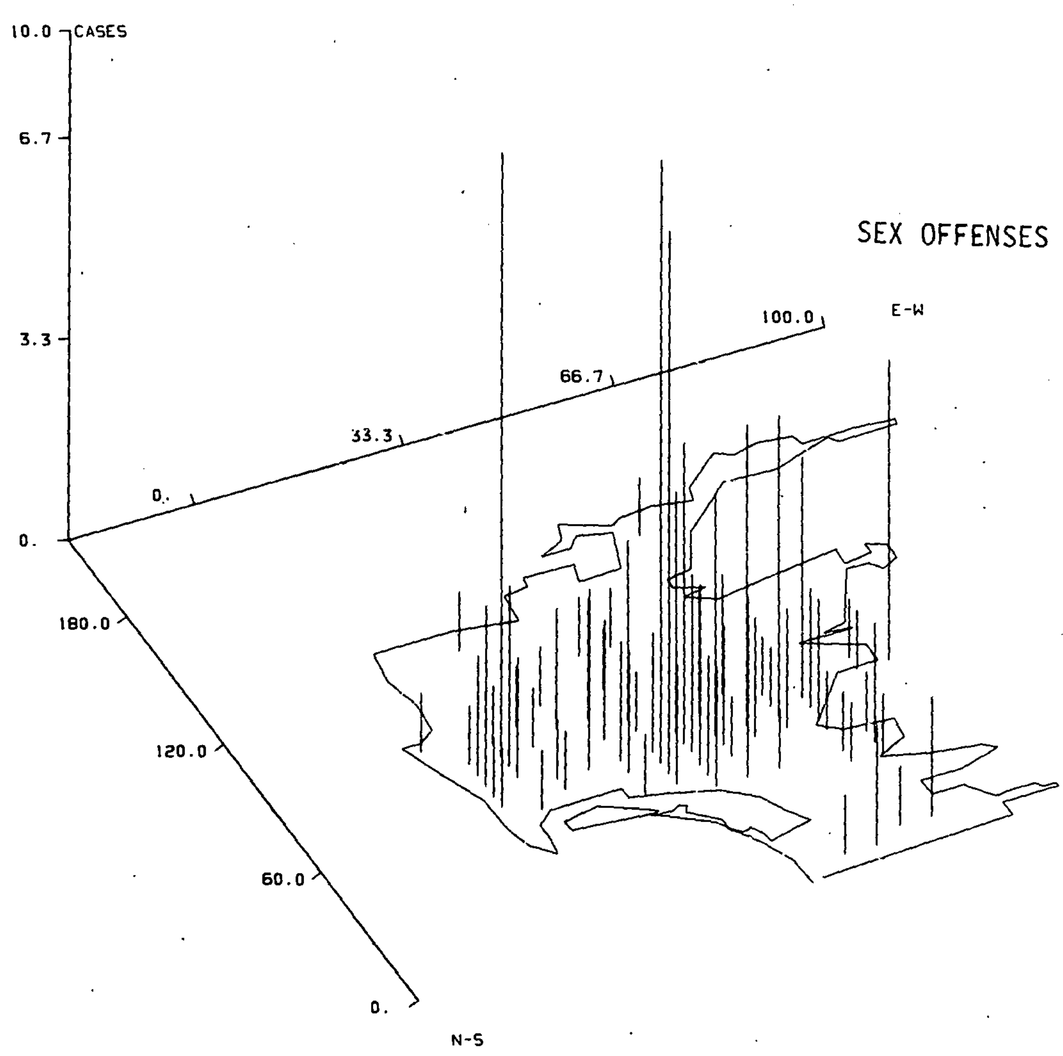

48.0
CASES
32.0
16.0
0.
MAL. MISCHIEF
E-W
0.
33.3
66.7
100.0
180.0
120.0
60.0
0.
N-S

Figure A1-10



10.0
CASES
6.7
3.3
0.
SEX OFFENSES
100.0
E-W
66.7
33.3
0.
180.0
120.0
60.0
0.
N-S

Figure A1-11

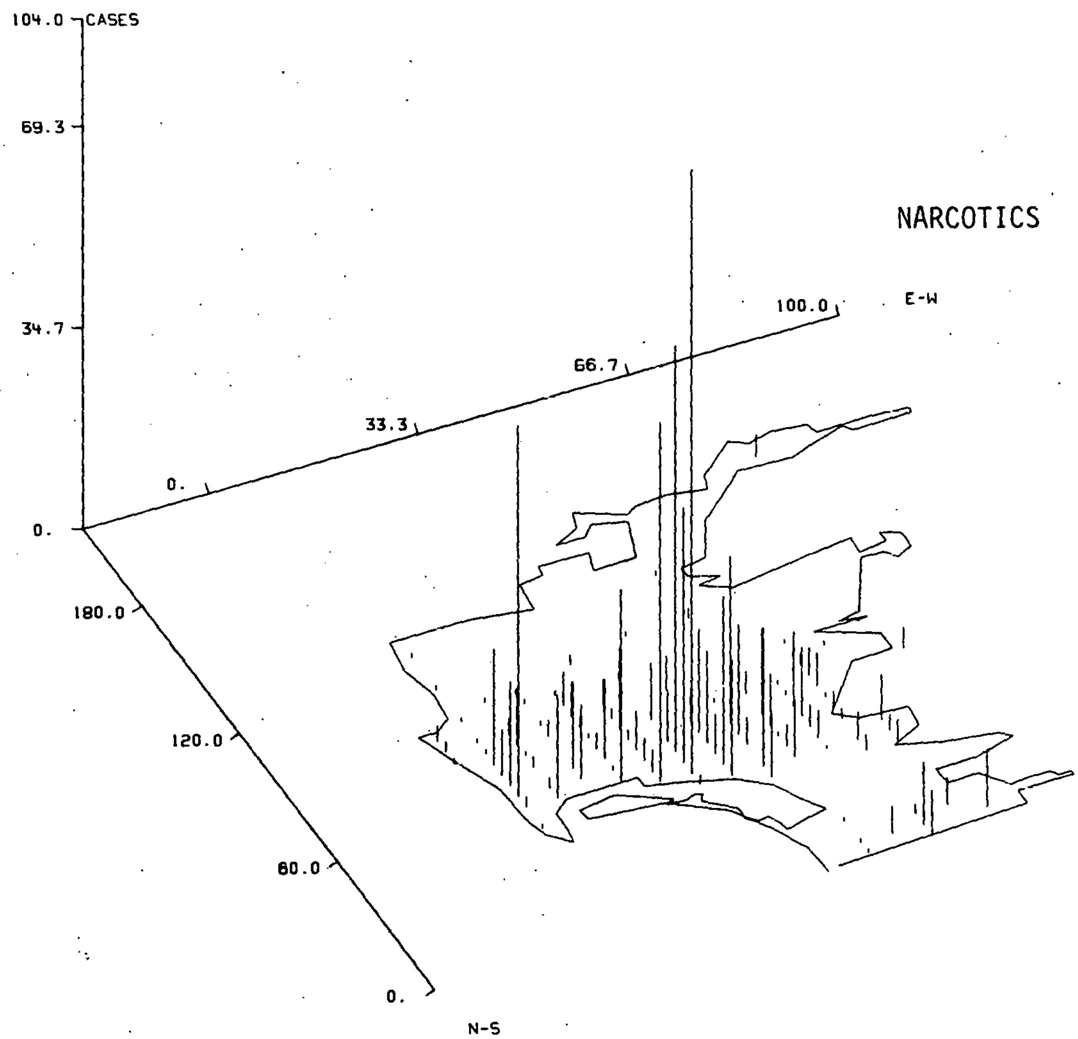
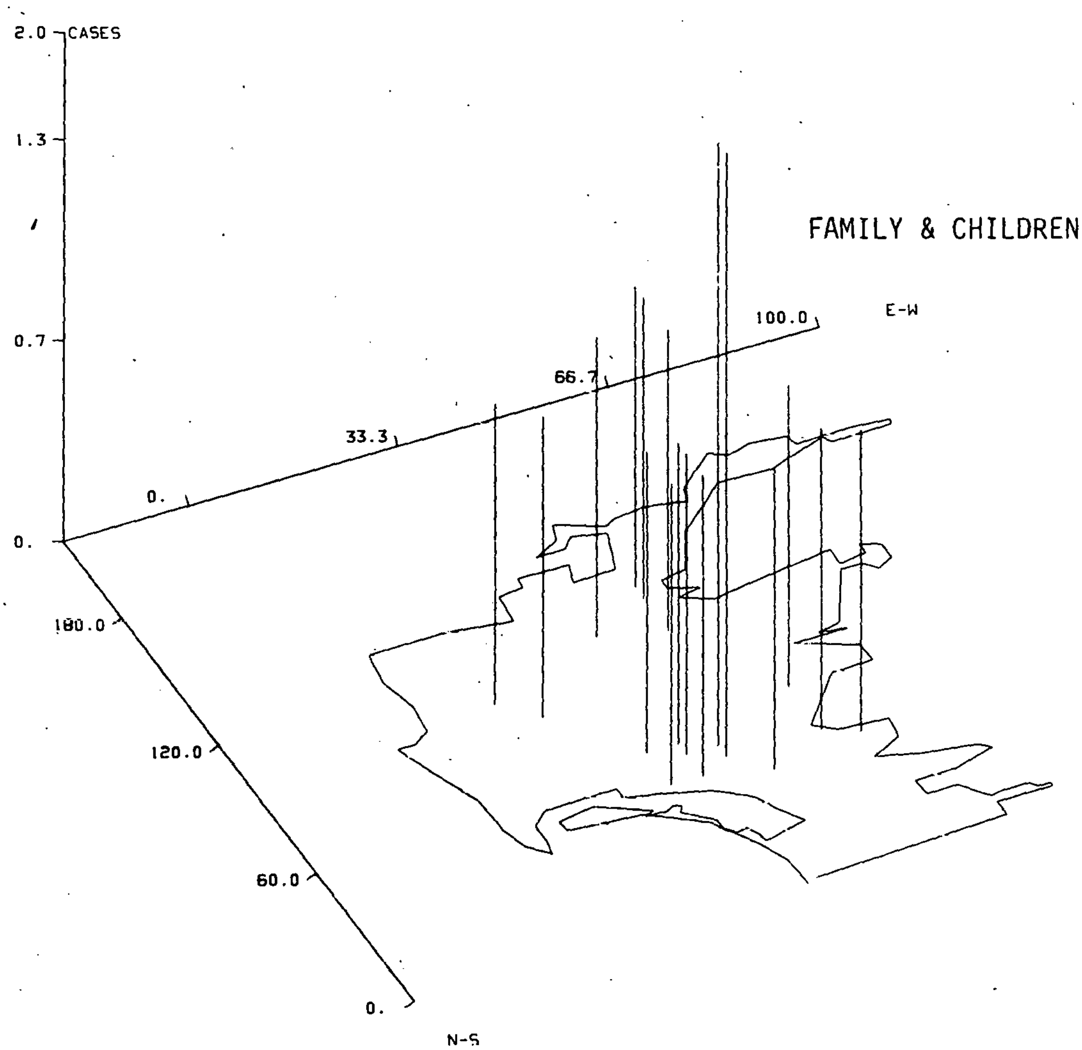

104.0
CASES
69.3
34.7
0.
NARCOTICS
100.0
E-W
66.7
33.3
0.
180.0
120.0
60.0
0.
N-S

Figure A1-12



2.0
CASES
1.3
0.7
0.
FAMILY & CHILDREN
100.0
E-W
66.7
33.3
0.
180.0
120.0
60.0
0.
N-S

Figure A1-13

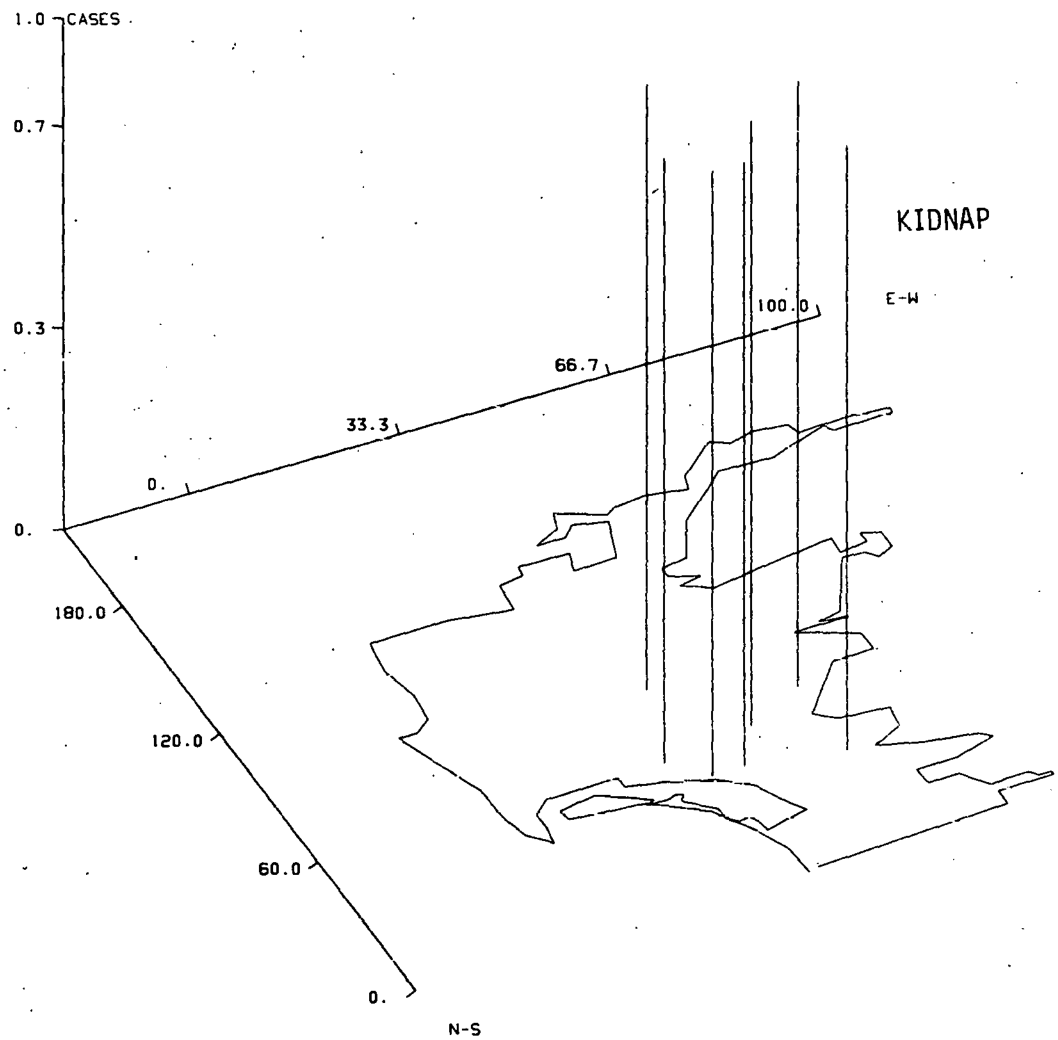


Figure A1-14

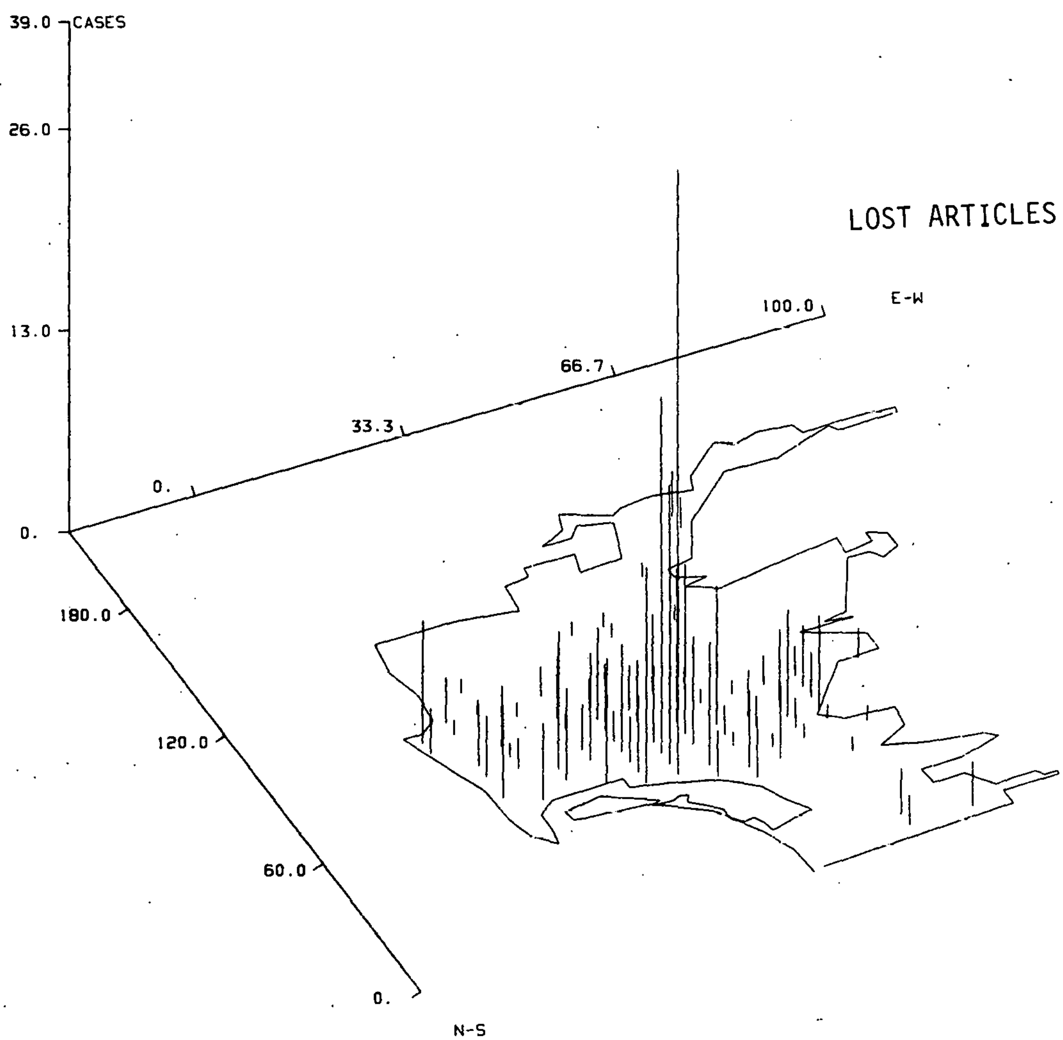

39.0
CASES
26.0
13.0
0.
LOST ARTICLES
100.0
E-W
66.7
33.3
0.
180.0
120.0
60.0
0.
N-S

Figure A1-15

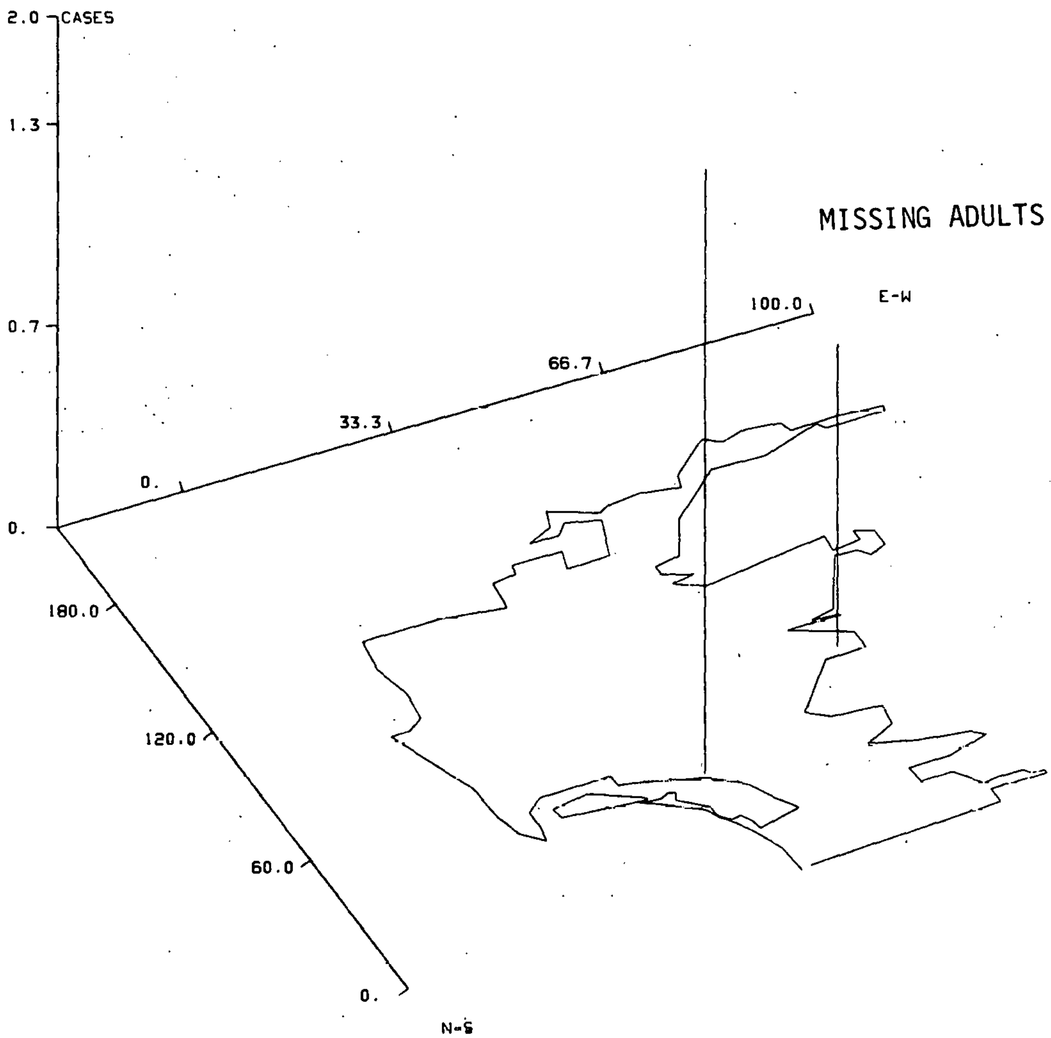

2.0 CASES
1.3
0.7
0.
MISSING ADULTS
100.0
E-W
66.7
33.3
0.
180.0
120.0
60.0
0.
N-S

Figure A1-16

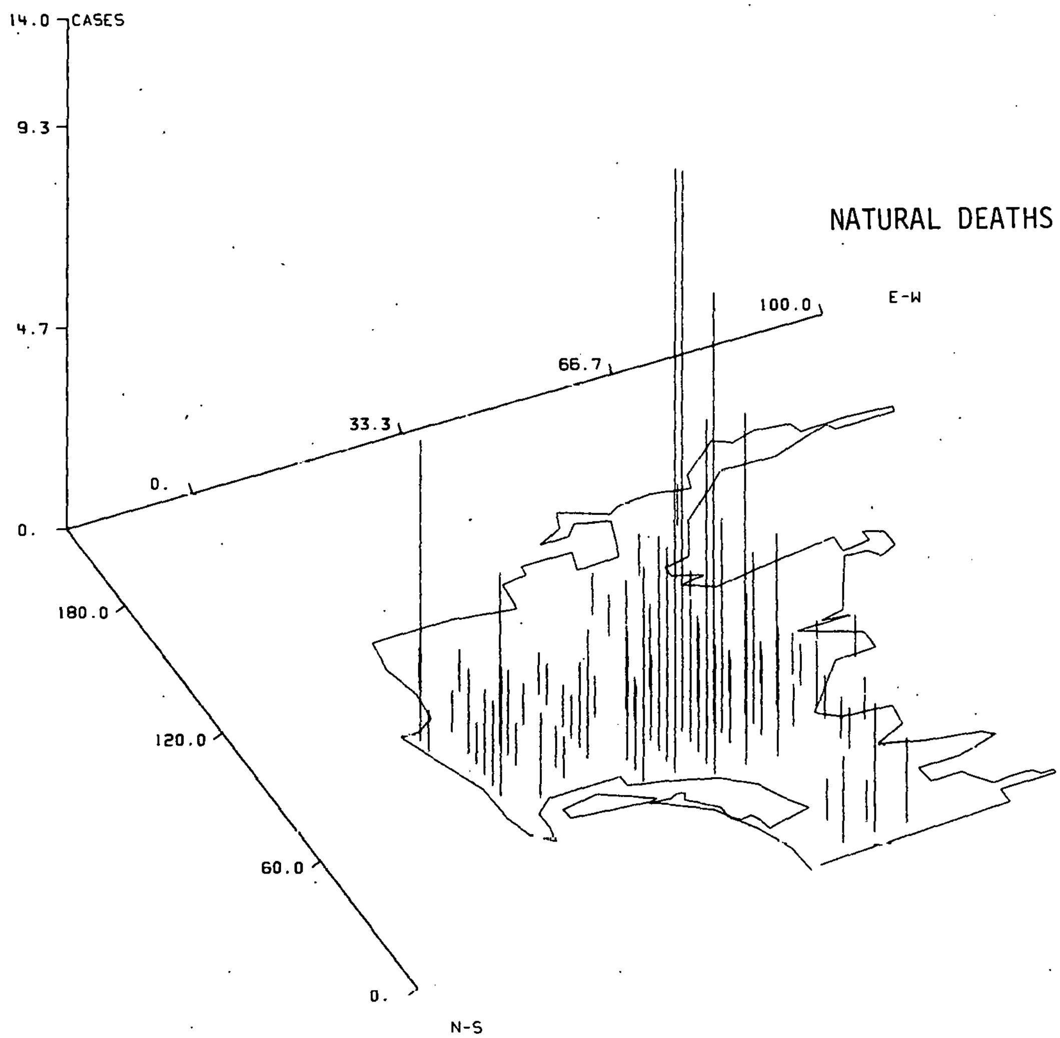

14.0 CASES
9.3
4.7
0.
NATURAL DEATHS
100.0
E-W
66.7
33.3
0.
180.0
120.0
60.0
0.
N-S

Figure A1-17

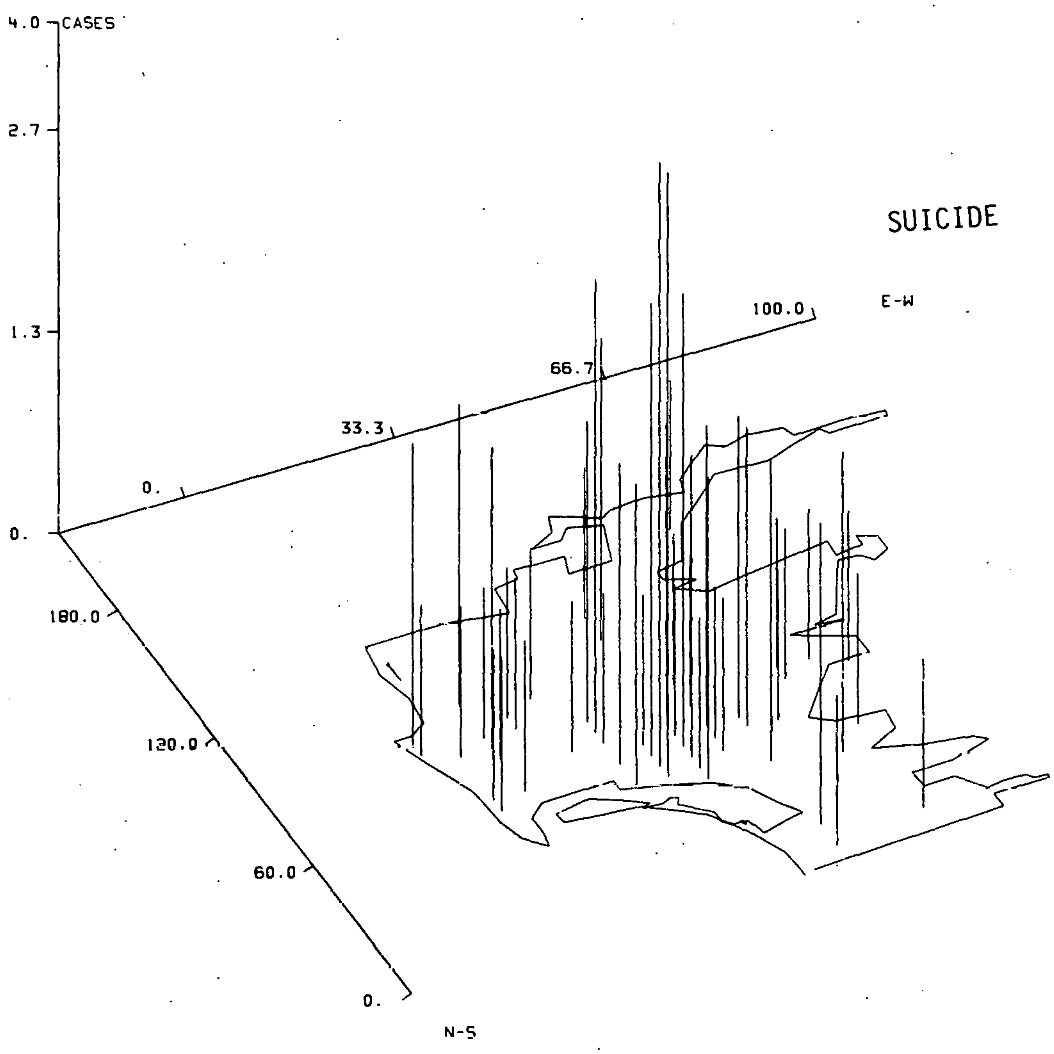


Figure A1-18

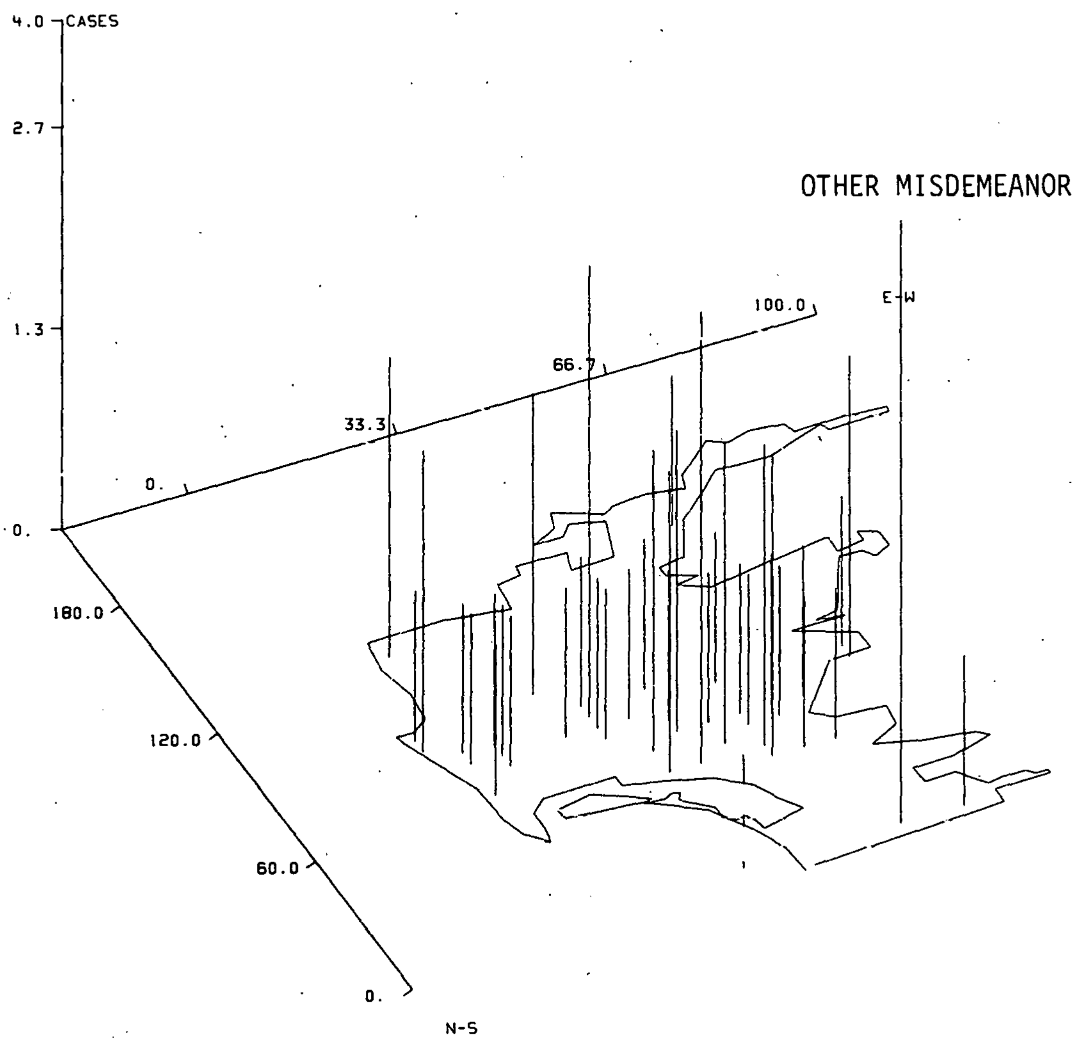

4.0
CASES
2.7
1.3
0.
OTHER MISDEMEANOR
E-W
0.
33.3
66.7
100.0
180.0
120.0
60.0
0.
N-S

Figure A1-19

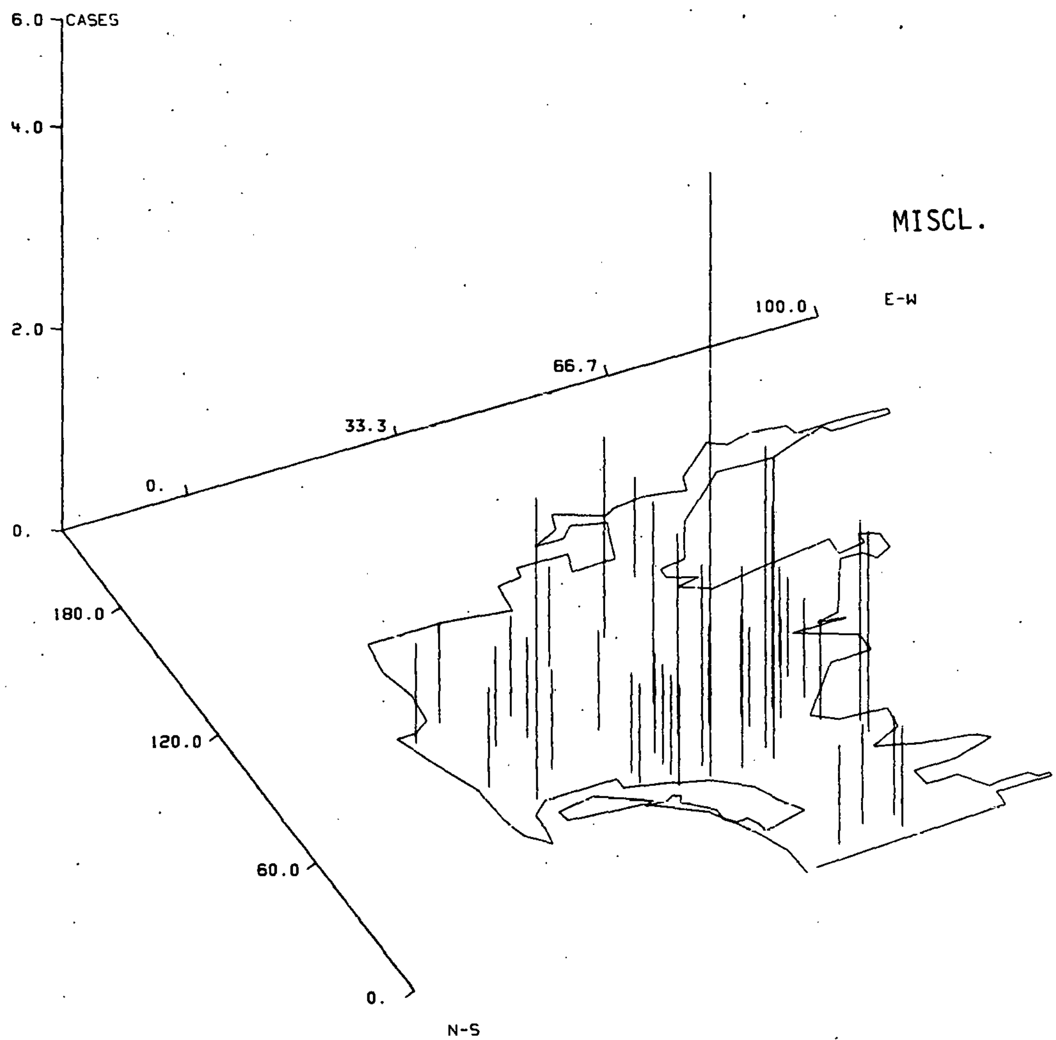


Figure A1-20

APPENDIX II

STATISTICS FOR SELECTED CRIME TYPES

PATTER VERSION 1.0B
INPUT COMPLETE IN 0.35 SECONDS
KEEPED COMPLETE IN 0.00 SECONDS
TESTING: DATA

Figure A2-1

PREDICTED PROPERTY:
MEAN = 3.07263E-01
VARIATION = 2.12852E-01

— BEGIN ANALYSIS OF ROBBERY DATA —

F-TEST RESULTS

WITH ALL VARIABLES:
RESIDUAL SUM OF SQUARES (RSS)= 7.41272E+05
MULT. CORREL. COEF.SQD (MCCS)= -6.48422E+03
F-VALUE (FVAL)= -1.32979E+02
RESIDUAL MEAN SQR (RMS)= 1.39337E+03

BEGIN -LEAVE 1 OUT- ANALYSIS

VARIABLE	VARIANCE	RSS	MCCS	FVAL	RMS
1	5.22272E+01	7.76646E+05	-6.79370E+03	-1.77641E+02	1.45712E+03
2	5.39177E-03	1.50679E+02	-4.05745E-01	-5.12805E+01	3.21462E-01
3	2.24558E+02	7.42187E+05	-6.49223E+03	-1.77639E+02	1.39247E+03
4	4.25018E+02	7.59141E+05	-6.64056E+03	-1.77640E+02	1.42428E+03

FTEST COMPLETE IN 0.17 SECONDS

AUTOSCALED DATA WILL BE STORED IN A1 WEIGHT CODE 0
FOR VARIABLE 1 XB= 1.4662942E+01 X2= 1.6746338E+02 WEIGHT= 1.0000000E+00
FOR VARIABLE 2 XB= 7.5869395E+01 X2= 1.7015823E+00 WEIGHT= 1.0000000E+00
FOR VARIABLE 3 XB= 4.1823091E+01 X2= 3.4725739E+02 WEIGHT= 1.0000000E+00
FOR VARIABLE 4 XB= 5.9432961E+01 X2= 4.7773904E+02 WEIGHT= 1.0000000E+00
AUTO COMPLETE IN 0.02 SECONDS

LEAST SQUARES CURVE FIT FOR AUTOSCALED DATA

CURVE FITTING ON PATTERN SET A1
WEIGHT VECTOR FOLLOWS
-6.5815E-01 -8.2650E-01 2.4649E-02 5.0375E-01 3.0726E-01
PB= 3.073E-01 P2= 1.069E+01 SD= 4.585E-01 PSD= 1.492E+02PRD= 0.616
LEAST COMPLETE IN 0.04 SECONDS

CORRELATION COEFFICIENTS FOR THE VARIABLES

VARIABLE CORRELATION COEFFICIENT ANALYSI
A1

1	1	1.00000									
2	1	0.03266	2	2	1.00000						
3	1	-0.01716	3	2	-0.01161	3	3	1.00000			
4	1	-0.00680	4	2	-0.00838	4	3	-0.64096	4	4	1.00000

VARIABLES ORDERED ON CORRELATION
4 3 -0.64096
CORR COMPLETE IN 0.03 SECONDS

CORRELATION COEFFICIENTS, VARIABLES VS. PREDICTED PROPERTY

VARIABLE-PREDICTED PROPERTY ANALYSIS FOR A1
1 1 -0.06445 2 2 -0.07974 3 3 -0.02594 4 4 0.04671
ORDERED ON ABSOLUTE VALUE OF CC.
1 2 -0.07974 2 1 -0.06445 3 4 0.04671 4 3 -0.02594
CORREP COMPLETE IN 0.01 SECONDS

PLOTVS FOR DATA SET...DATA
PLOTVS COMPLETE IN 0.16 SECONDS
HISTOGRAMS FOR PATTERN SET DATA
HISTO COMPLETE IN 0.07 SECONDS
HISTOGRAMS FOR PATTERN SET DATA
HISTO COMPLETE IN 0.09 SECONDS
HISTOGRAMS FOR PATTERN SET DATA
HISTO COMPLETE IN 0.12 SECONDS
HISTOGRAMS FOR PATTERN SET DATA
HISTO COMPLETE IN 0.08 SECONDS

Figure A2-2

```
2-DIMENSIONAL PLOTTING OF A1
PERCENTAGE RETAINED    66.859
2DPLOT      COMPLETE IN        0.65 SECONDS
NEW         COMPLETE IN        0.00 SECONDS
INPUT       COMPLETE IN        0.55 SECONDS
KEEP80      COMPLETE IN        0.00 SECONDS
TESTING: DATA
```

— BEGIN ANALYSIS FOR ASSAULT DATA —

```
PREDICTED PROPERTY:
MEAN =      7.41933E-01
VARIATION =      1.91467E-01
```

F-TEST RESULTS

```
WITH ALL VARIABLES:
         RESIDUAL SUM OF SQUARES (RSS)=      1.15885E+04
         MULT. CORREL. COEF.SQD (MCCS)=     -7.13115E+01
         F-VALUE (FVAL)=                   -2.05124E+02
         RESIDUAL MEAN SQR (RMS)=           1.39285E+01

BEGIN -LEAVE 1 OUT- ANALYSIS

VARIABLE        VARIANCE          RSS              MCCS              FVAL              RMS
--------------------------------------------------------------------------------------------------

   1           4.78855E+01     1.27066E+04     -7.82882E+01     -2.74165E+02     1.52540E+01
   2           5.38592E-03     5.72840E+02     -2.57449E+00     -1.99986E+02     6.87683E-01
   3           2.19739E+02     1.26823E+04     -7.81367E+01     -2.74158E+02     1.52249E+01
   4           4.09249E+02     1.13954E+04     -7.01065E+01     -2.73762E+02     1.36799E+01
FTEST       COMPLETE IN        0.26 SECONDS
AUTOSCALED DATA        WILL BE STORED IN A1          WEIGHT CODE      0
FOR VARIABLE      1 XB=  1.3787336E+01 X2=  2.0020026E+02 WEIGHT=  1.0000000E+00
FOR VARIABLE      2 XB=  7.5864578E+01 X2=  2.1232087E+00 WEIGHT=  1.0000000E+00
FOR VARIABLE      3 XB=  4.1151135E+01 X2=  4.2886085E+02 WEIGHT=  1.0000000E+00
FOR VARIABLE      4 XB=  6.3063892E+01 X2=  5.8527016E+02 WEIGHT=  1.0000000E+00
AUTO        COMPLETE IN        0.03 SECONDS
```

LEAST SQUARES CURVE FIT FOR AUTOSCALED DATA

```
CURVE FITTING ON PATTERN SET A1
WEIGHT VECTOR FOLLOWS
 4.0035E-01  1.2306E-01 -1.5229E-01  2.5020E-01  7.4194E-01
    PB= 7.419E-01      P2= 1.266E+01      SD= 4.372E-01      PSD= 5.893E+01PRD=    0.086
LEAST       COMPLETE IN        0.06 SECONDS
```

CORRELATION COEFFICIENTS FOR THE VARIABLES

```
VARIABLE CORRELATION COEFFICIENT ANALYSI
A1
    1    1    1.00000
    2    1    0.00939    2    2    1.00000
    3    1    0.12809    3    2   -0.03483    3    3    1.00000
    4    1   -0.04813    4    2    0.01721    4    3   -0.46781    4    4    1.00000
VARIABLES ORDERED ON CORRELATION
CORR        COMPLETE IN        0.04 SECONDS
```

CORRELATION COEFFICIENTS, VARIABLES VS. PREDICTED PROPERTY

```
VARIABLE-PREDICTED PROPERTY ANALYSIS FOR A1
    1    1    0.02922    2    2    0.01078    3    3   -0.01756    4    4    0.02404
ORDERED ON ABSOLUTE VALUE OF CC.
    1    1    0.02922    2    4    0.02404    3    3   -0.01756    4    2    0.01078
CORRPP      COMPLETE IN        0.01 SECONDS
PLOTVS FOR DATA SET...DATA
PLOTVS      COMPLETE IN        0.25 SECONDS
HISTOGRAMS FOR PATTERN SET DATA
HISTO       COMPLETE IN        0.07 SECONDS
HISTOGRAMS FOR PATTERN SET DATA
HISTO       COMPLETE IN        0.09 SECONDS
HISTOGRAMS FOR PATTERN SET DATA
HISTO       COMPLETE IN        0.11 SECONDS
```

Figure A2-3

```
HISTOGRAMS FOR PATTERN SET DATA
HISTO       COMPLETE IN        0.08 SECONDS
2-DIMENSIONAL PLOTTING OF A1
PERCENTAGE RETAINED     62.705
2DPLOT      COMPLETE IN        1.01 SECONDS
NEW         COMPLETE IN        0.00 SECONDS
INPUT       COMPLETE IN        1.12 SECONDS
KEEPSC      COMPLETE IN        0.00 SECONDS
TESTING: DATA

PREDICTED PROPERTY:
MEAN =       2.85381E-01
VARIATION =       2.03939E-01
```

— BEGIN ANALYSIS FOR BURGLARY DATA —

F-TEST RESULTS

```
WITH ALL VARIABLES:
         RESIDUAL SUM OF SQUARES (RSS)=         3.23839E+05
         MULT. CORREL. COEF.SQD (MCCS)=        -9.23824E+02
         F-VALUE (FVAL)=                       -4.27537E+02
         RESIDUAL MEAN SQR (RMS)=               1.89158E+02

BEGIN -LEAVE 1 OUT- ANALYSIS
```

VARIABLE	VARIANCE	RSS	MCCS	FVAL	RMS
1	4.00335E+01	4.41090E+05	-1.25867E+03	-5.70547E+02	2.57495E+02
2	5.16078E-03	4.43654E+02	-2.66993E-01	-1.20327E+02	2.58992E-01
3	2.70460E+02	2.58573E+05	-7.37436E+02	-5.70227E+02	1.50947E+02
4	3.08302E+02	2.55270E+05	-7.28004E+02	-5.70217E+02	1.49019E+02

```
FTEST       COMPLETE IN        0.52 SECONDS
AUTOSCALED DATA       WILL BE STORED IN A1           WEIGHT CODE      0
FOR VARIABLE      1 XB=  1.4319744E+01 X2=  2.6217830E+02 WEIGHT=  1.0000000E+00
FOR VARIABLE      2 XB=  7.5876413E+01 X2=  2.9767556E+00 WEIGHT=  1.0000000E+00
FOR VARIABLE      3 XB=  3.8965440E+01 X2=  6.8145369E+02 WEIGHT=  1.0000000E+00
FOR VARIABLE      4 XB=  6.5944671E+01 X2=  7.2756700E+02 WEIGHT=  1.0000000E+00
AUTO        COMPLETE IN        0.06 SECONDS
```

LEAST SQUARES CURVE FIT FOR AUTOSCALED DATA

```
CURVE FITTING ON PATTERN SET A1
WEIGHT VECTOR FOLLOWS
  0.         0.         0.         0.          2.8538E-01
    PB= 2.854E-01     P2= 1.871E+01    SD= 4.516E-01     PSD= 1.582E+02PRD=    0.000
LEAST       COMPLETE IN        0.28 SECONDS
```

CORRELATION COEFFICIENTS FOR THE VARIABLES

```
VARIABLE CORRELATION COEFFICIENT ANALYSI
A1
   1    1   0.
   2    1   0.          2    2   0.
   3    1   0.          3    2   0.          3    3   0.
   4    1   0.          4    2   0.          4    3   0.          4    4   0.
VARIABLES ORDERED ON CORRELATION
CORR        COMPLETE IN        0.15 SECONDS
```

CORRELATION COEFFICIENTS, VARIABLES VS. PREDICTED PROPERTY

```
VARIABLE-PREDICTED PROPERTY ANALYSIS FOR A1
    1    1  -0.12264     2    2  -0.02896     3    3   0.03946     4    4   0.02256
ORDERED ON ABSOLUTE VALUE OF CC.
    1    1  -0.12264     2    3   0.03946     3    2  -0.02896     4    4   0.02256
CORRPP      COMPLETE IN        0.02 SECONDS
PLOTVS FOR DATA SET...DATA
PLOTVS      COMPLETE IN        0.51 SECONDS
HISTOGRAMS FOR PATTERN SET DATA
HISTO       COMPLETE IN        0.07 SECONDS
HISTOGRAMS FOR PATTERN SET DATA
HISTO       COMPLETE IN        0.09 SECONDS
```

```
HISTOGRAMS FOR PATTERN SET DATA
HISTO      COMPLETE IN       0.12 SECONDS
HISTOGRAMS FOR PATTERN SET DATA
HISTO      COMPLETE IN       0.09 SECONDS
PATTER     COMPLETE IN       7.40 SECONDS
```

Figure A2-4

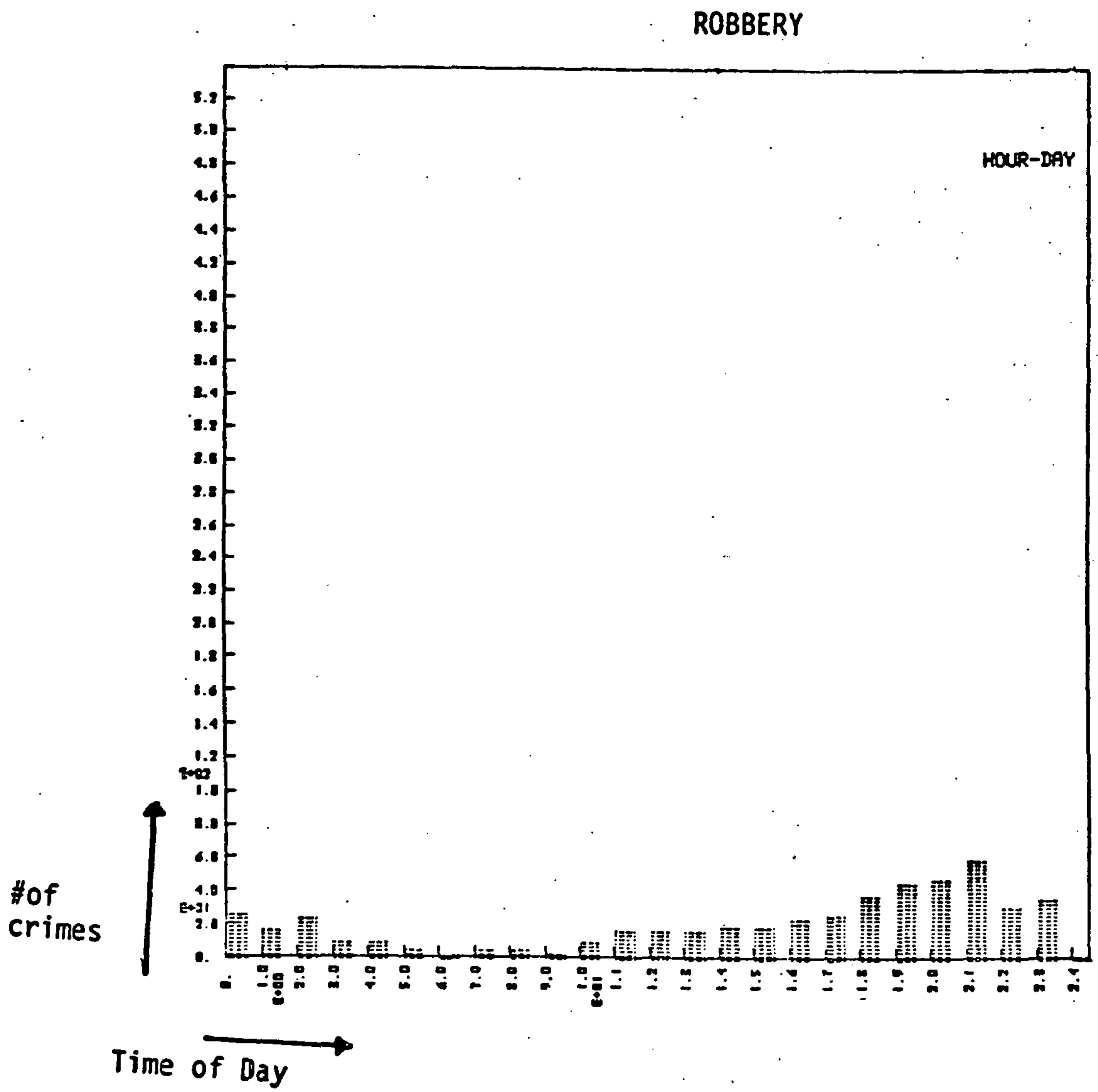


Figure A2-5

ROBBERY

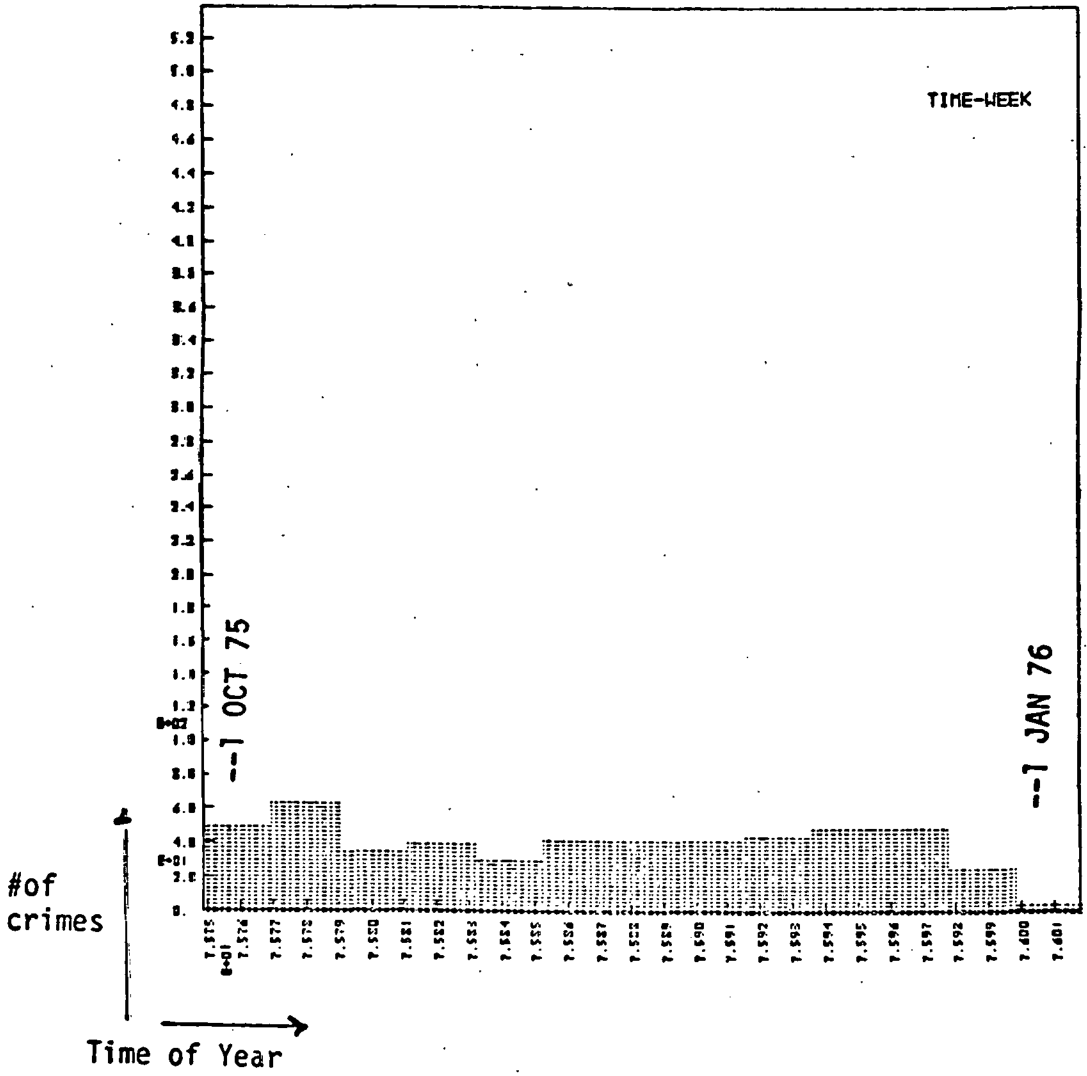


Figure A2-6

ROBBERY

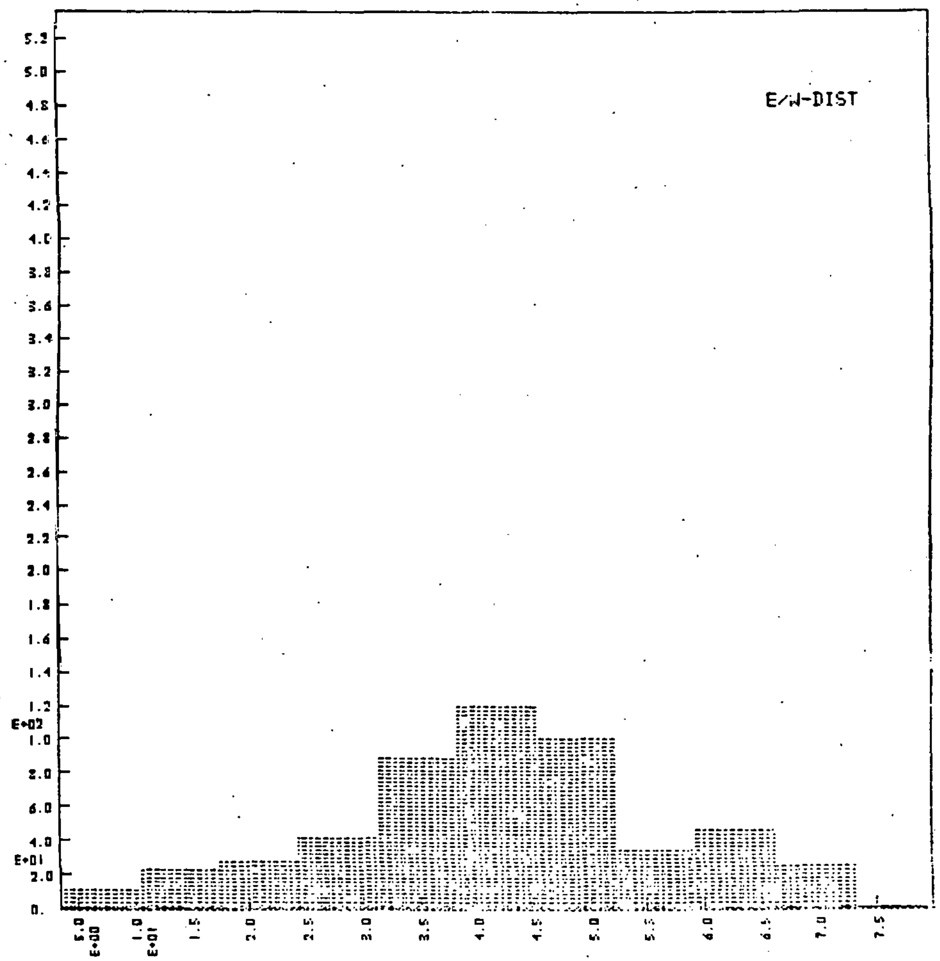


Figure A2-7

ROBBERY

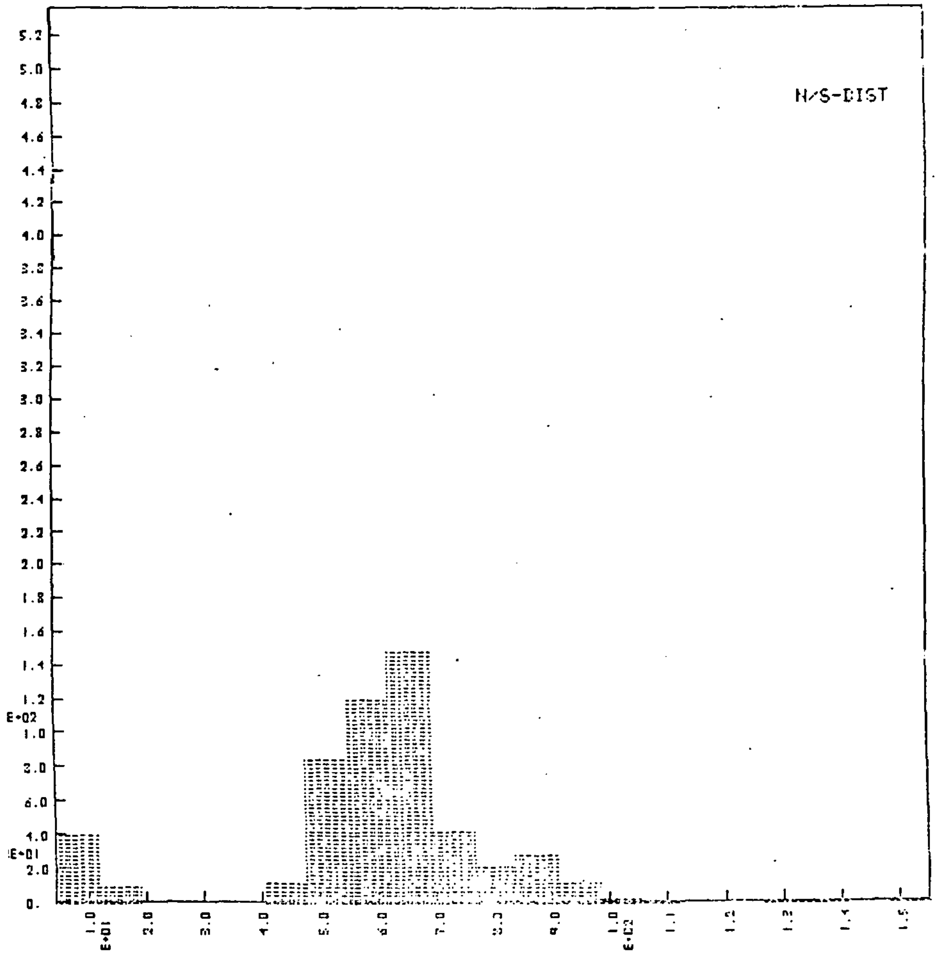


Figure A2-8

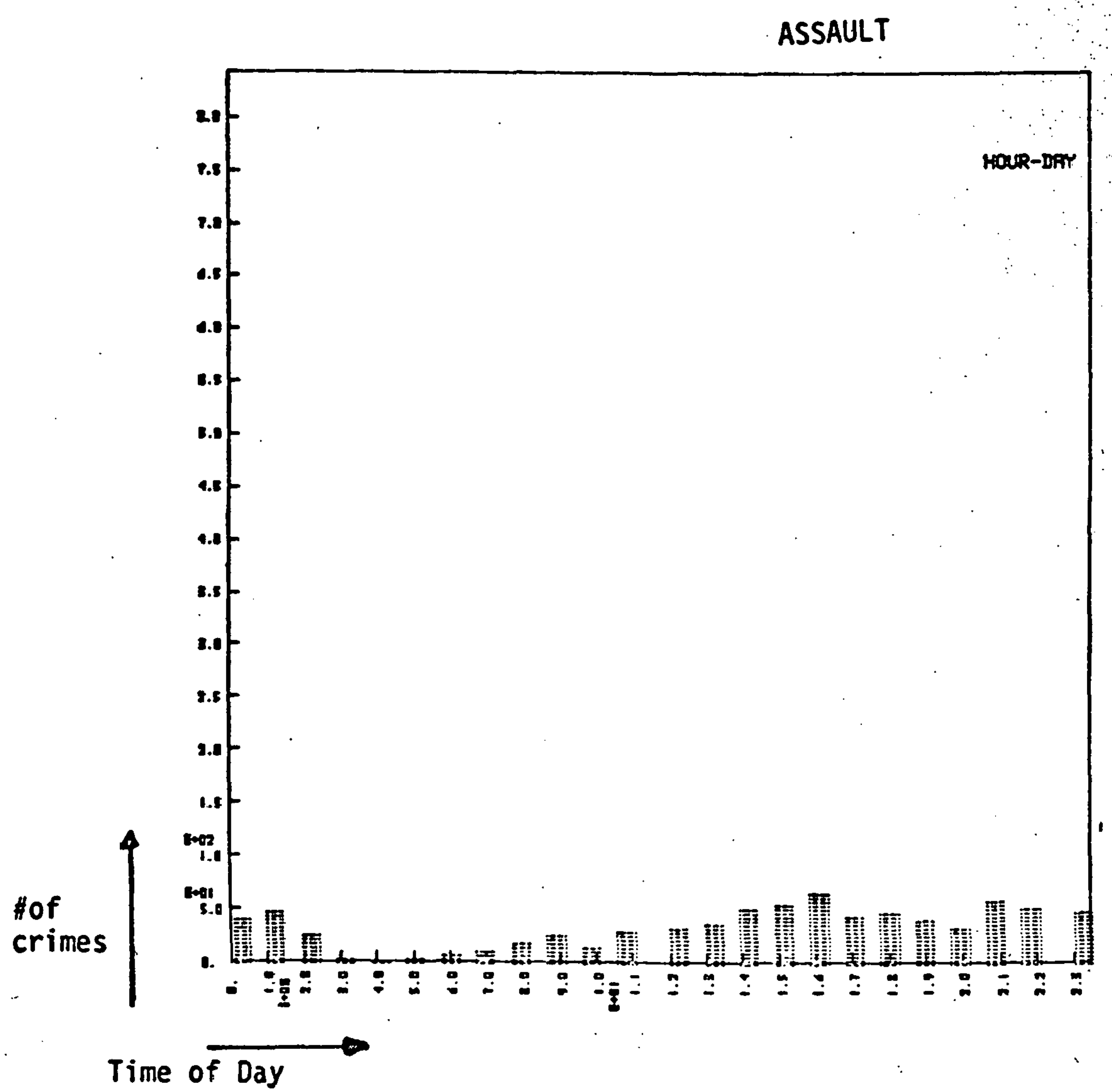


Figure A2-9

ASSAULT

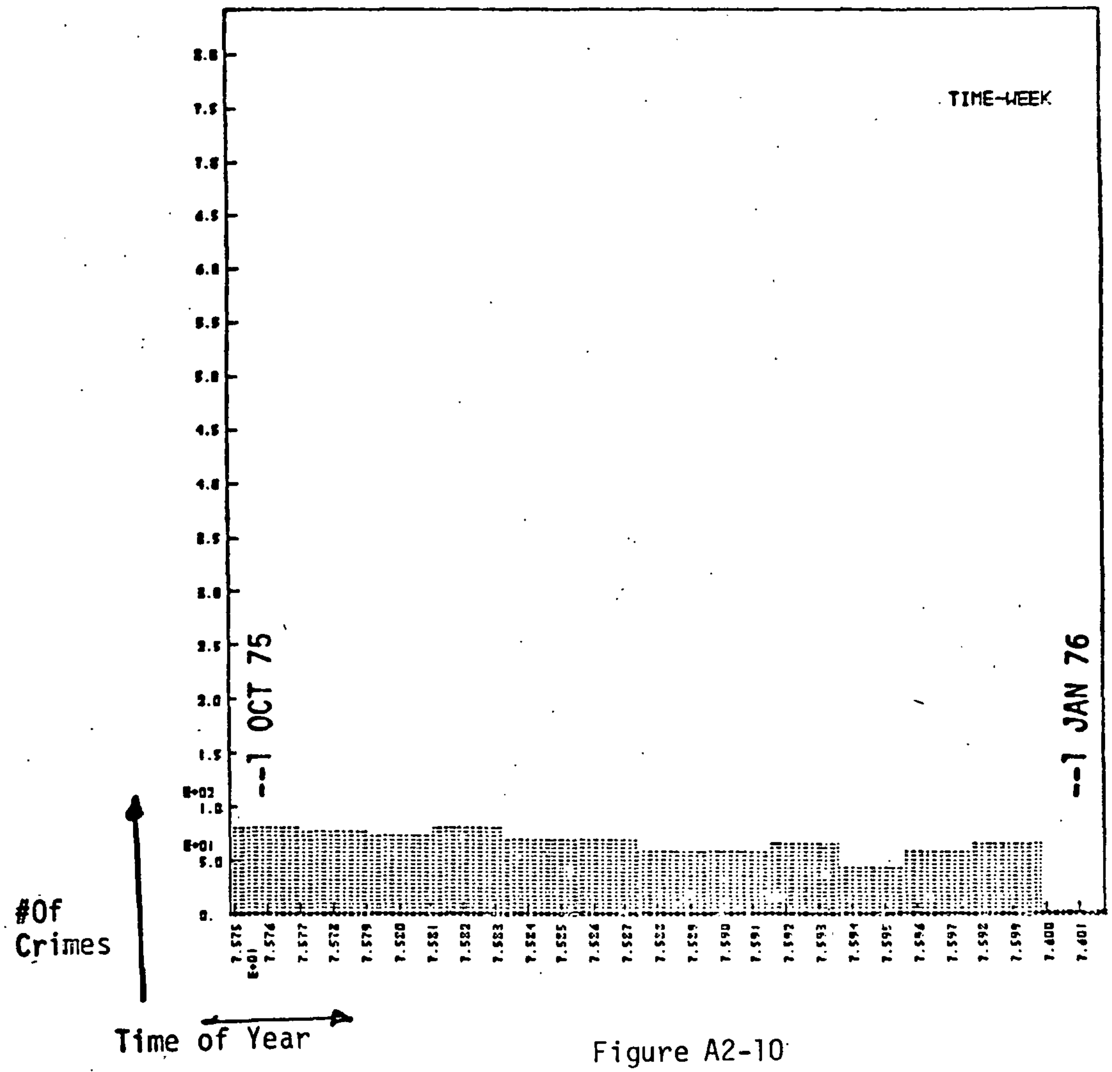


Figure A2-10

ASSAULT

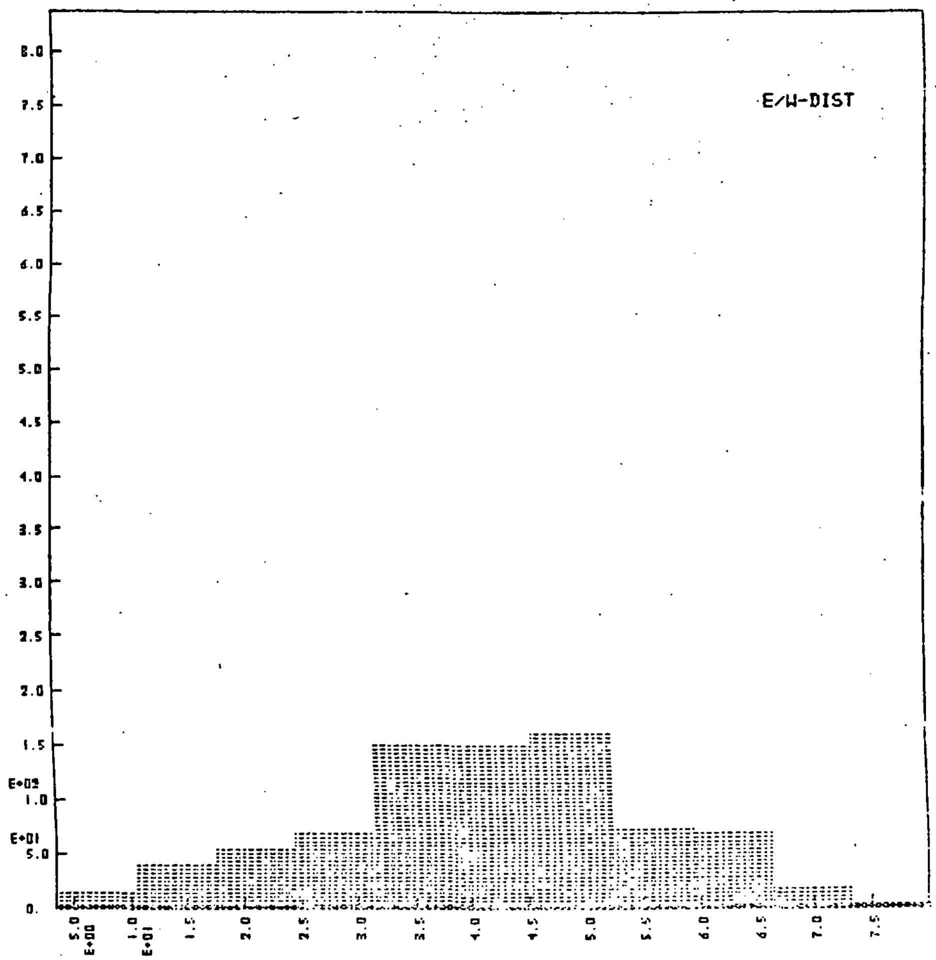


Figure A2-11

ASSAULT

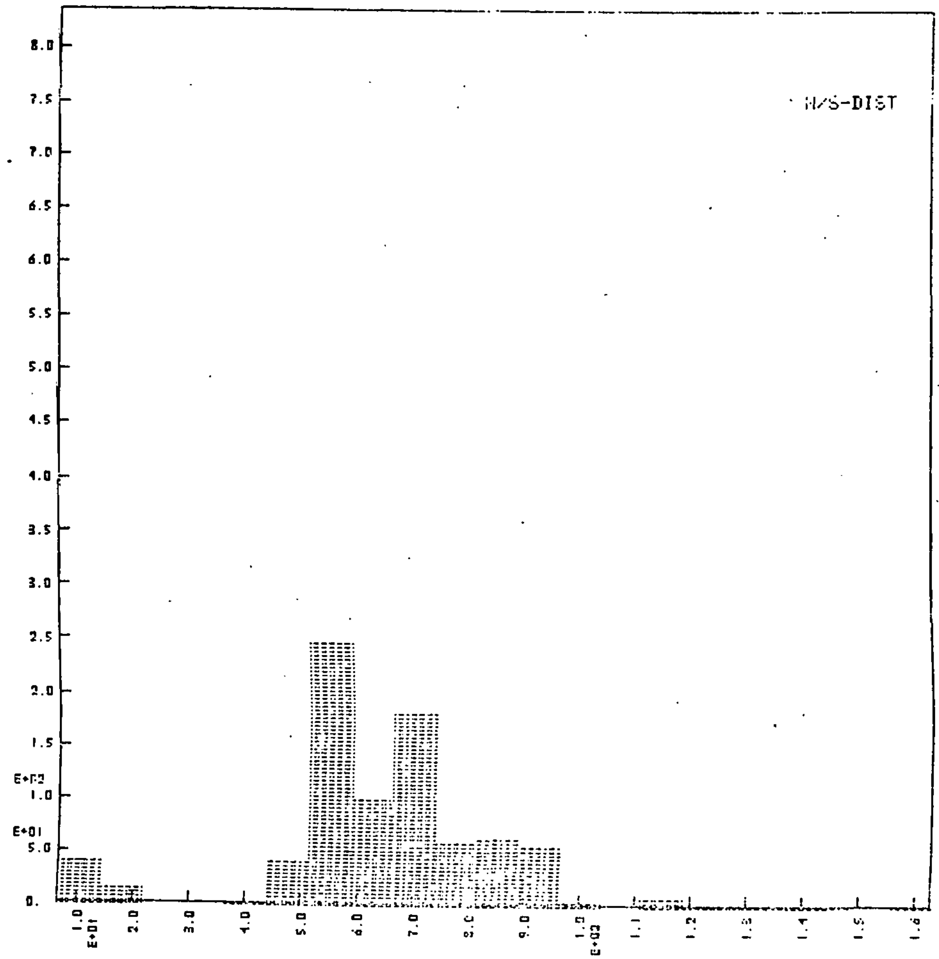


Figure A2-12

BURGLARY

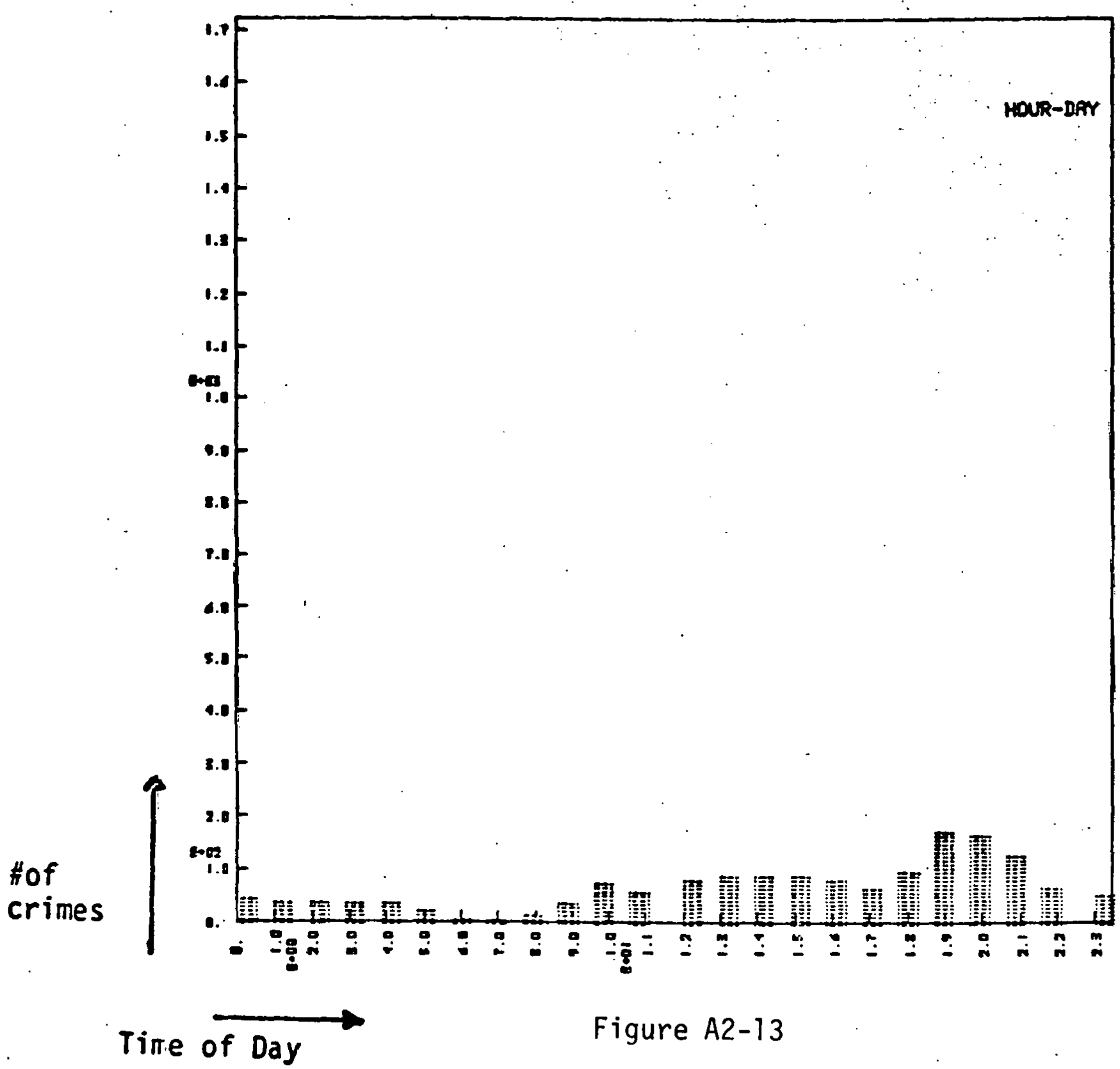


Figure A2-13

BURGLARY

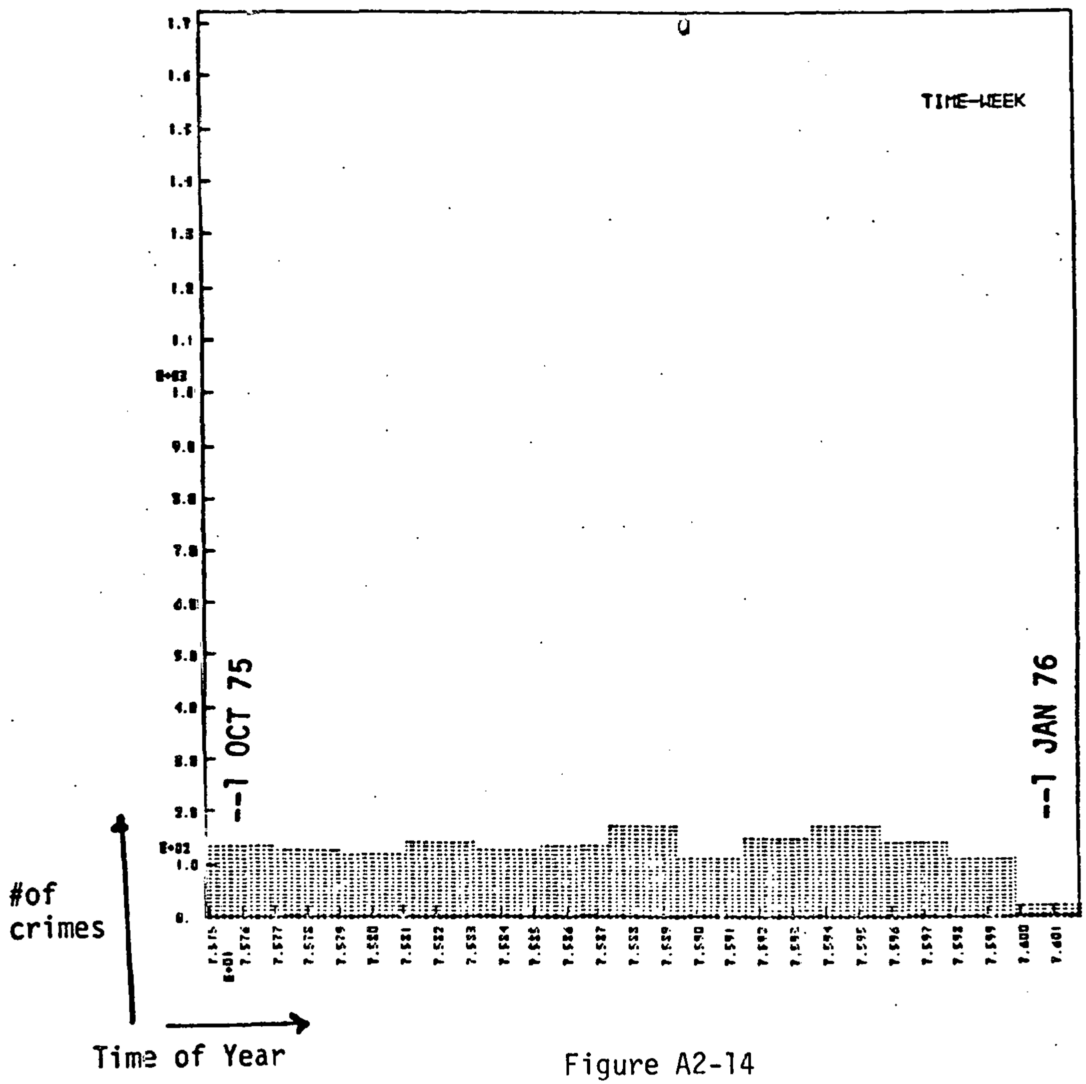


Figure A2-14

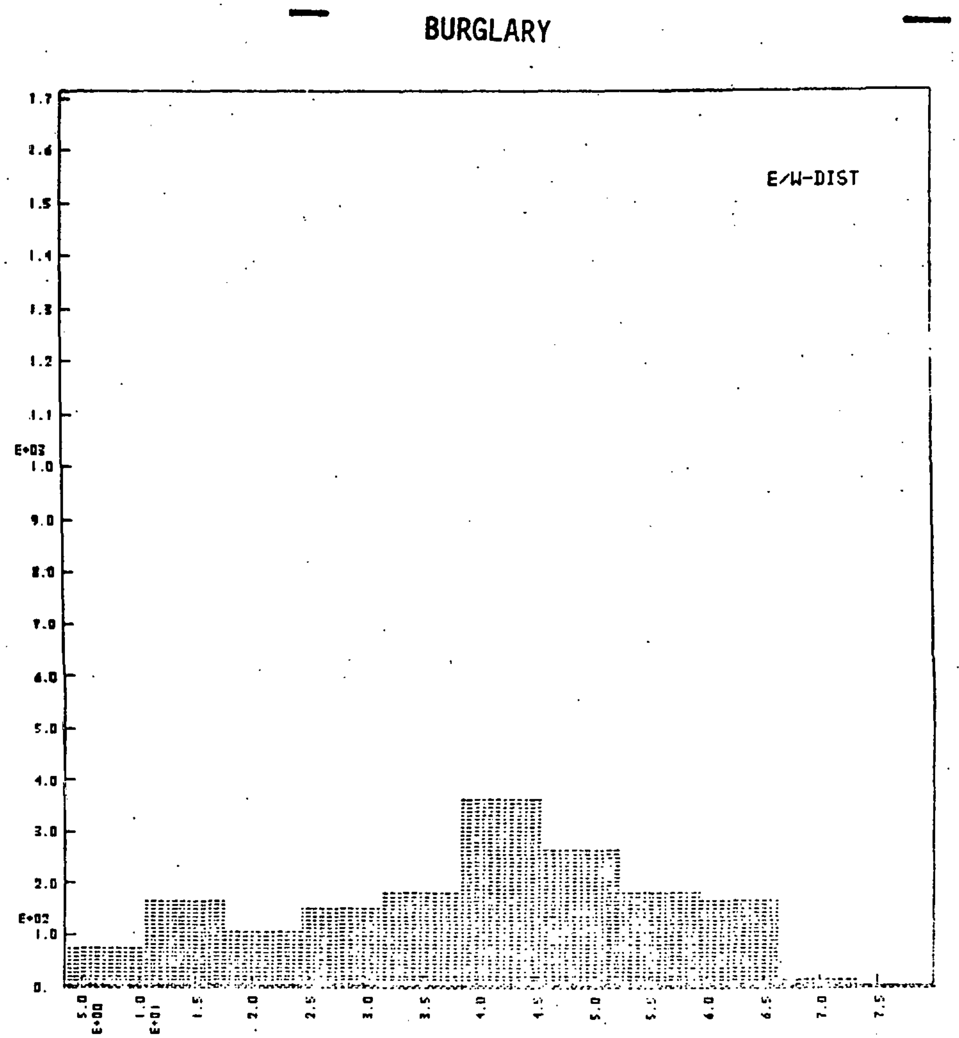


Figure A2-15

BURGLARY

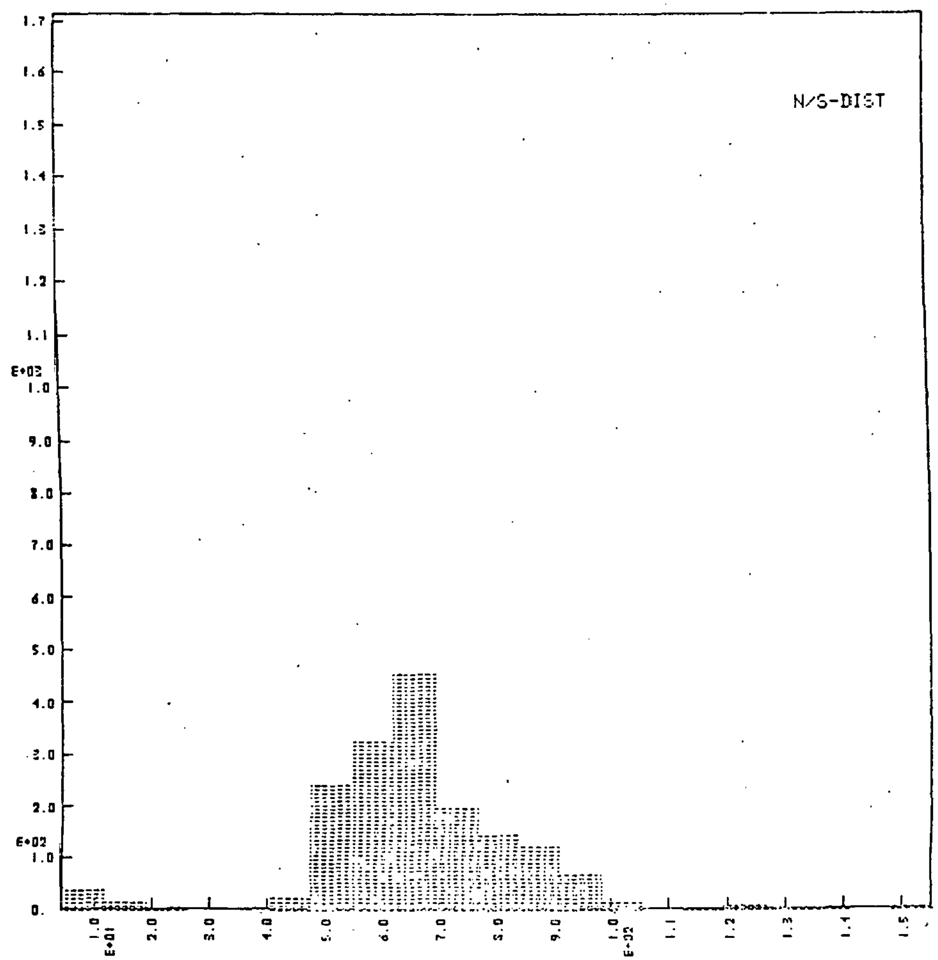


Figure A2-16

Acknowledgement

This report was sponsored by the U. S. Energy Research and Development Administration as a contribution to the National Technology Transfer Program. The assistance of the San Diego Police Department, and particularly Police Captain C. Ecklund and Police Lieutenant J. McQueeney, is gratefully acknowledged.

Printed in the United States of America
Available from
National Technical Information Service
U.S. Department of Commerce
5285 Port Royal Road
Springfield, VA 22161
Price: Printed Copy $; Microfiche $2.25

Page Range	Domestic Price	Page Range	Domestic Price
001–025	$ 3.50	326–350	10.00
026–050	4.00	351–375	10.50
051–075	4.50	376–400	10.75
076–100	5.00	401–425	11.00
101–125	5.25	426–450	11.75
126–150	5.50	451–475	12.00
151–175	6.00	476–500	12.50
176–200	7.50	501–525	12.75
201–225	7.75	526–550	13.00
226–250	8.00	551–575	13.50
251–275	9.00	576–600	13.75
276–300	9.25	601–up	*
301–325	9.75		

*Add $2.50 for each additional 100 page increment from 601 to 1,000 pages: add $4.50 for each additional 100 page increment over 1,000 pages.

Technical Information Department
LAWRENCE LIVERMORE LABORATORY
University of California | Livermore, California | 94550