## STATISTICS QUALIFYING EXAM - TAKE HOME Due Noon Friday August 21, 1998. Return to Math and Stat dept office

*Directions:* The answer to each problem should be word-processed, double spaced, and should be no longer than three pages. An appendix is allowed for each problem but will be examined only at the discretion of the graders. The better constructed your appendix, the more likely it is to get examined.

1. Does pollution kill people? Data in one early study designed to explore this issue came from 60 Standard Metropolitan Statistical Areas (SMSA) in the United States, obtained for the years 1959-1961. Total age adjusted mortality from all causes, in deaths per 100,000 population, is the response variable. The explanatory variables, listed in the table below, include mean annual precipitation (in inches), median number of school years completed for persons of age 25 or older, percentage of 1960 population that is non-white, relative pollution potential of oxides of nitrogen  $(NO_X)$ , and relative pollution potential of sulphur dioxide  $(SO_2)$ . "Relative pollution potential" is the product of the tons emitted per day per square kilometer and a factor correcting for SMSA dimension and exposure. The first three explanatory variables are a subset of climate and socioeconomic variables used in the original analysis. (Note: Two cities, Lancaster and York, are heavily populated by members of the Amish religion, who prefer to teach their children at home. The lower years of education for these two cities do not indicate a social climate similar to other cities with similar years of education.).

Is there evidence that mortality is associated either of the pollution variables, after the effect of the climate and socioeconomic variables are accounted for? Analyze the data, and write a report of your findings, including any important limitations of the study.

To get a copy of the entire data set in a file called pollute in your current AIX directory, type the following UNIX command:

% cp ~edbed/pollute pollute

Air pollution and mortality	data for 60 U.S. c	ities, 1959–1961

City	Mortality	Precipitation	Education	Nonwhite	NOx	SO <sub>2</sub>
San Jose, CA	790.733	13	12.2	3.0	32	
Wichita, KS	823.764	28	12.1	75	52	1
San Diego, CA	839.709	10	12.1	5.9	66	20
Lancaster, PA	<b>844</b> .053	43	9.5	2.9	7	32
Minneapolis, MN	857.622	25	12.1	2.0	11	26
Dallas, TX	860.101	35	11.8	14.8	1	1
Miami, FL	861.439	60	11.5	13.5	1	1
Los Angeles, CA	861.833	11	12.1	7.8	319	130
Grand Rapids, MI	871.338	31	10.9	5.1	3	10
Denver, CO	871.766	15	12.2	4.7	8	28
Kochester, NY	8/4.281	32	11.1	5.0	4	18
Fart Worth TY	88/.400	43	11.5	7.2	3	10
Portland OP	891./U8 802.001	31	11.4	11.5	1	1
Worcester MA	805 606	37	12.0	3.6	21	44
Seattle WA	800 264	43	11.1	1.0	3	8
Bridgenort CT	899 529	45	12.2	5.7		20
Springfield, MA	904 155	45	10.0	3.3	4	20
San Francisco, CA	911.701	18	12.2	13.4	171	20
York, PA	911.817	42	90	4.8	8	40
Utica, NY	912.202	40	10.3	2.5	2	11
Canton, OH	912.347	36	10.7	6.7	7	20
Kansas City, MO	919.729	35	12.0	12.6	4	4
Akron, OH	921.870	36	11.4	8.8	15	59
New Haven, CT	923.234	46	11.3	8.8	3	8
Milwaukee, WI	929.150	30	11.1	5.8	23	125
Doston, MA	934.700	43	12.1	3.5	32	62
Providence PI	930.234	30 ·	11.4	12.4	4	16
Flint, MI	941.181	30	10.1	2.2	4	18
Reading, PA	946.185	41	9.6	27	11	80
Syracuse, NY	<b>950.672</b>	38	11.4	3.8	5	25
Houston, TX	952.529	46	11.4	21.0	5	1
Saint Louis, MO	953.560	34	9.7	17.2	15	68
Youngstown, OH	954.442	38	10.7	11.7	13	39
Columbus, OH	958.839	. 37	11.9	13.1	9	15
Detroit, MI	959.221	31	10.8	15.8	35	124
Allentourn DA	901.009	45	10.1	21.0	14	78
Washington DC	902.334	44	9.8	0.8	6	33
Indianapolis IN	968 665	30	12.5	25.9	-28	102
Cincinnati, OH	970.467	40	10.2	13.0	26	35
Greensboro, NC	971.122	42	10.2	13.0 22 7	20	140
Toledo, OH	972.464	31	10.7	95	7	25
Atlanta, GA	982.291	47	11.1	27.1	8	23
Cleveland, OH	<b>985.95</b> 0	35	11.1	14.7	21	64
Louisville, KY	989.265	30	9.9	13.1	37	193
Pittsburgh, PA	991.290	36	10.6	8.1	59	263
New York, NY	994.648	42	10.7	11.3	26	108
Albany, NY	997.875	35	11.0	3.5	10	39
Wilmington DE	1001.902	30	10.5	8.1	12	37
Memphie TN	1005.502	40 50	11.3	12.1	11	42
Philadelphia PA	1015.032	42	10.4	30./ 17 5	18	54
Chattanooga, TN	1017.613	52	96	22.2	52	101
Chicago, IL	1024.885	33	10.9	16.3	63	278
Richmond, VA	1025.502	44	11.0	28.6	9 9	48
Birmingham, AL	1030.380	53	10.2	38.5	32	72
Baltimore, MD	1071.289	43	9.6	24.4	38	206
New Orleans, LA	1113.056	54	9.7	31.4	17	1

.

2. The table below gives the estimated numbers of 4 types of plankton (I,II,III,IV) caught in 12 successive days (hauls) in a fishing experiment. Analyze these data, focusing attention on potential differences in catches across days, and across plankton types. Be complete.

.

•

Haul	Ι	Π	III	IV
1	895	1520	43300	11000
2	540	1610	32800	8600
3	1020	1900	28800	8260
4	470	1350	34600	9830
5	428	980	27800	7600
6	620	1710	32800	9650
7	760	1930	28100	8900
8	537	1960	18900	6060
9	845	1840	31400	10200
10	1050	2410	39500	15500
11	387	1520	29000	9250
12	497	1685	22300	7900

·

Numbers of 4 types of plankton caught