

Activity #17: State Data to practice t-tests, ANOVA, chi-square, and regression methods

Resources: cancer.sav datafile

We are going to conduct a series of statistical analyses on a dataset using SPSS. As we work with the data, we will develop some research questions. Think about what statistical methods would be appropriate to answer each of these questions. Some questions can be answered with graphs, some questions can be answered through statistical inference (t-tests, ANOVA, regression analyses), and some questions cannot be answered at all.

As we conduct our analyses, think about what assumptions must be met by the data. Also try to think of ways to determine if the assumptions are satisfied. If assumptions are not met by the data, what other options do we have?

My goal in this activity (and in the course as a whole) is to give you an idea of the actual practice of statistics. Outside the classroom, statisticians do not spend much time calculating statistics by hand – we use a pc. But we do spend a lot of time thinking about assumptions, appropriate analyses, and conclusions we can draw from the data. We also spend a lot of time thinking about the best way to present the information we learn.

Running SPSS and opening our data file

1. Open a web browser and go to: <http://homepage.mac.com/bradthiessen>
2. In the middle-right of the page, you will see a “**Cancer Data Regression Example**” link.
 - a. Right-click this link and save the **cancer.sav** file to your desktop
3. Exit the web browser and go to the Start menu. Locate **SPSS for Windows** and run the program
4. The program will load and you will see 2 error messages. Click **OK** for each of the error messages.
5. SPSS should now be running. You will see a blank grid on the screen.
 - a. Go to the **File Menu** and select **OPEN... DATA**
 - b. Navigate to the desktop and open the **cancer.sav** file that you downloaded from the website.
6. You should now see the data on your screen

Data Description:

The data for this project came from *The Data and Story Library* at: <http://lib.stat.cmu.edu/DASL/DataArchive.html>

Smoking and cancer data source:

Fraumeni, J.F. (1968). Cigarette smoking and cancers of the urinary tract: Geographic variations in the United States. *Journal of the National Cancer Institute*, 41(5), 1205-1211

Teacher pay and educational spending data source:

National Education Association, as reported by the Albuquerque Tribune, 11/07/1986

Age and economic data source:

U.S. Department of Commerce, Bureau of the Census, Government Finances in 1960, Census of Population, 1960, Census of Manufacturers, 1958, Statistical Abstract of the United States, 1961

Union and Right-to-Work Laws data source:

Meltz, N.M. (1989). Interstate and interprovincial differences in union density. *Industrial Relations*, 28:2, 142-158.

Presidential election data source:

Dave Leip's Atlas of U.S. Presidential Elections: <http://www.uselectionatlas.org/>

Raw Data:

State	Side	Region	Vote68	CigSales	Bladder	Lung	Kidney	Leuk	TeachPay	EdSpend	Over65	Econ	Upub	RTW	Upvt
AK	.	5	Rep	30.34	3.46	25.88	4.32	4.90	41.5	8.35	.	.	30.7	No	30
AL	West	1	Rep	18.20	2.90	17.05	1.59	6.15	22.9	2.73	8.0	69.4	32.4	RTW	14
AR	West	8	Rep	18.24	2.99	15.98	2.02	6.94	19.5	2.64	10.9	68.6	21.6	RTW	11
AZ	West	3	Rep	25.82	3.52	19.80	2.75	6.61	26.6	2.83	6.9	90.2	24.0	RTW	10
CA	West	5	Rep	28.60	4.46	22.07	2.66	7.06	29.1	3.61	8.8	112.6	32.9	No	24
CT	East	4	Dem	31.10	5.11	22.83	3.35	7.20	26.6	4.89	9.6	121.6	51.6	No	14
DC	.	6	Dem	40.46	5.60	27.27	3.13	7.08	34.0	5.02
DE	East	6	Rep	33.60	4.78	24.55	3.36	6.45	24.6	4.52	8.0	116.1	50.6	No	14
FL	East	6	Rep	28.27	4.46	23.57	2.41	6.07	22.3	3.73	11.2	80.9	21.4	RTW	7
IA	West	7	Rep	22.12	4.23	16.59	2.90	7.69	21.7	3.57	11.9	100.2	35.0	RTW	17
ID	West	3	Rep	20.10	3.08	13.58	2.46	6.62	21.0	2.51	8.7	93.9	26.1	No	13
IL	East	0	Rep	27.91	4.75	22.80	2.95	7.27	27.2	3.62	9.7	120.9	30.6	No	27
IN	East	0	Rep	26.18	4.09	20.30	2.81	7.00	24.3	3.16	9.6	110.8	28.6	No	24
KS	West	7	Rep	21.84	2.91	16.84	2.88	7.42	27.6	3.91	11.0	102.2	22.9	RTW	9
KY	East	1	Rep	23.44	2.86	17.71	2.13	6.41	20.9	2.85	9.6	76.8	32.8	No	18
LA	West	8	.	21.58	4.65	25.45	2.30	6.71	20.5	3.12	7.4	86.0	22.5	RTW	12
MA	East	4	Dem	26.92	4.69	22.04	3.03	6.89	26.8	4.64	11.1	107.5	52.5	No	14
MD	East	6	Dem	25.91	5.21	26.48	2.85	6.81	27.2	4.35	7.3	93.4	32.4	No	14
ME	East	4	Dem	28.92	4.79	20.94	3.22	6.24	19.6	3.35	11.0	85.5	47.7	No	11
MI	East	0	Dem	24.96	5.27	22.72	2.97	6.91	30.2	3.78	8.2	108.4	49.6	No	30
MN	West	7	.	22.06	3.72	14.20	3.54	8.28	27.4	3.98	10.4	95.7	43.5	No	21
MO	West	7	.	27.56	4.04	20.98	2.55	6.82	22.0	3.16	11.7	99.1	26.1	No	27
MS	West	1	Rep	16.08	3.06	15.60	1.77	6.08	18.4	2.31	8.7	57.4	13.4	RTW	8
MT	West	3	Rep	23.75	3.95	19.50	3.43	6.90	22.5	3.95	9.7	97.6	25.0	No	21
NB	West	7	Rep	23.32	3.72	16.70	2.92	7.80	20.9	3.29	11.6	99.1	20.8	RTW	12
ND	West	7	Rep	19.96	2.89	12.12	3.62	6.99	20.8	3.06	9.3	93.4	29.0	RTW	10
NJ	East	2	Rep	28.64	5.98	25.95	3.12	7.12	27.2	5.54	9.2	117.9	31.6	No	18
NM	West	3	Rep	21.16	2.90	14.59	2.52	5.95	22.6	3.40	5.4	86.2	17.4	No	11
NV	West	3	Rep	42.40	6.54	23.03	2.85	6.67	25.6	2.93	6.4	205.0	38.0	RTW	19
NY	East	2	Dem	29.14	5.30	25.02	3.10	7.23	30.7	5.71	10.1	111.5	56.9	No	31
OH	East	0	Rep	26.38	4.47	21.89	2.95	7.38	24.5	3.55	9.2	111.8	38.4	No	25
OK	West	8	Rep	23.44	2.93	19.45	2.45	7.46	21.4	2.75	10.7	84.9	25.7	No	10
PA	East	2	Dem	23.78	4.89	12.11	2.75	6.83	25.9	4.17	10.0	103.1	46.2	No	24
RI	East	4	Rep	29.18	4.99	23.68	2.84	6.35	29.5	4.67	10.4	94.9	61.1	No	12
SC	East	6	Rep	18.06	3.25	17.45	2.05	5.82	21.6	2.92	6.3	65.2	13.7	RTW	4
SD	West	7	Rep	20.94	3.64	14.11	3.11	8.15	18.1	2.97	10.5	88.2	23.1	RTW	6
TN	East	1	Rep	20.08	2.94	17.60	2.18	6.59	21.8	2.53	8.7	75.1	33.0	RTW	14
TX	West	8	Dem	22.57	3.21	20.74	2.69	7.02	25.2	3.43	7.8	98.8	27.3	RTW	10
UT	West	3	Rep	14.00	3.31	12.01	2.20	6.71	22.3	2.30	6.7	92.5	42.3	RTW	9
VT	East	4	Rep	25.89	4.63	21.22	3.17	6.56	20.3	3.55	11.2	87.0	37.3	No	7
WA	West	5	Dem	21.17	4.04	20.34	2.78	7.48	26.0	3.71	9.8	100.4	32.4	No	33
WI	East	0	Rep	21.25	5.14	20.55	2.34	6.73	26.5	4.25	10.2	104.3	44.8	No	20
WV	East	6	Dem	22.86	4.78	15.53	3.28	7.38	20.6	2.82	9.3	85.1	30.2	No	29
WY	West	3	Rep	28.04	3.20	15.92	2.66	5.78	27.2	5.44	7.8	125.8	31.4	RTW	12
AK	.	5	Rep	30.34	3.46	25.88	4.32	4.90	41.5	8.35	.	.	30.7	No	30
AL	West	1	Rep	18.20	2.90	17.05	1.59	6.15	22.9	2.73	8.0	69.4	32.4	RTW	14
AR	West	8	Rep	18.24	2.99	15.98	2.02	6.94	19.5	2.64	10.9	68.6	21.6	RTW	11
AZ	West	3	Rep	25.82	3.52	19.80	2.75	6.61	26.6	2.83	6.9	90.2	24.0	RTW	10
CA	West	5	Rep	28.60	4.46	22.07	2.66	7.06	29.1	3.61	8.8	112.6	32.9	No	24
CT	East	4	Dem	31.10	5.11	22.83	3.35	7.20	26.6	4.89	9.6	121.6	51.6	No	14
DC	.	6	Dem	40.46	5.60	27.27	3.13	7.08	34.0	5.02

Variables in this data set: (We will pretend as though this data all came from the same time period)

State: (1960) Name of state
Side: (1960) Is the state an Eastern state or Western state?
Region: (1960) States classified into 9 regions
Vote68: (1968) Did the state vote for a Republican or Democratic president in 1968?
CigSales: (1960) Number of cigarettes smoked (hds per capita)
Bladder: (1960) Deaths per 100K population from bladder cancer
Lung: (1960) Deaths per 100K population from lung cancer
Kidney: (1960) Deaths per 100K population from kidney cancer
Leuk: (1960) Deaths per 100K population from leukemia
TeachPay: (1985) Average public school teacher annual salary
EdSpend: (1985) Spending on public schools per pupil
Over65: (1960) Percent of population over 65 years of age
Economy: (1960) Economic ability index (income, retail sales, value of output)
UnionPub: (1982) Percent of public sector employees in unions
Ritowork: (1982) Does the state have right-to-work law?
UnionPvt: (1982) Percent of private sector employees in unions

Possible Research Questions:

- 1) Are the states that voted for a Republican presidential candidate more likely to have right-to-work laws?
- 2) Are the states that voted for a Democratic presidential candidate evenly distributed among the regions?
- 3) Do Republican states have lower teacher pay, lower educational spending, or better economies than Democratic states?
- 4) Is there a difference among regions in teacher pay, educational spending, cigarette sales, or cancer rates?
- 5) Is there a relationship between teacher pay and educational spending?
- 6) Is there a relationship between a state's economy and teacher pay?
- 7) Is there a relationship between educational spending and cigarette smoking in a state?
- 8) Does cigarette smoking lead to lung cancer?
- 9) Can we predict the likelihood that a state has right-to-work laws based on the percentage of public employees in unions?

For each research question, do the following:

- a) Identify an appropriate analysis to answer the question
- b) Determine if the necessary assumptions are met by the data
- c) Create some visual displays of the data
- d) Run the appropriate analyses and write your conclusions
- e) State how confident you are in your conclusions
- f) Provide some evidence that the computer calculated things correctly
- g) Identify any limitations to your study

