

Introduction to Data Analysis: Coding the Data

6/1/2006

1



- Types of Political Analysis
 - Informal
 - Qualitative
 - Quantitative
 - Data Analysis
 - Formal Theory

6/1/2006

2



- Data Analysis
 - Measurement
 - Code Data

6/1/2006

3

NES Measure of PID



"Generally speaking, do you usually think of yourself as a Republican, a Democrat, an Independent, or what?"

(IF REPUBLICAN OR DEMOCRAT)

"Would you call yourself a strong ... or a not very strong ...?"

(IF INDEPENDENT, OTHER, OR NO PREFERENCE)

"Do you think of yourself as closer to the Republican or Democratic party?"

6/1/2006

4



Code Data, Cont.

e.g.,

or ...

- | | |
|----------------------|----------------------|
| 1 Strong Democrat | 1 Strong Republican |
| 2 Weak Democrat | 2 Weak Republican |
| 3 Leaning Democrat | 3 Leaning Republican |
| 4 Pure Independent | 4 Pure Independent |
| 5 Leaning Republican | 5 Leaning Democrat |
| 6 Weak Republican | 6 Weak Democrat |
| 7 Strong Republican | 7 Strong Republican |



Code Data, Cont.

e.g., Sex

or ...

- | | |
|----------|----------|
| 0 Male | 0 Female |
| 1 Female | 1 Male |



Code Data, Cont.

From the Georgia Poll

To begin, I'd like to ask you a few questions about national and local issues.

First, what would you say is the single biggest problem facing the state of Georgia today?

How Categorize, Code?



most important problem	Freq.	Percent	Cum.
aids	1	0.23	0.23
breakdown of family	2	0.46	0.69
crime	16	3.69	4.38
delinquency	1	0.23	4.61
drugs	11	2.53	7.14
economy	53	12.21	19.35
education/poor public education	114	26.27	45.62
environmental quality	6	1.38	47.00
health insurance/medical insurance	18	4.15	51.15
housing/poor public housing	1	0.23	51.38
inflation/high prices	9	2.07	53.46
medical care/medicare	11	2.53	55.99
poverty/gap between rich and poor	7	1.61	57.60
racial conflict/racism	6	1.38	58.99
terrorism	6	1.38	60.37
unemployment	64	14.75	75.12
other	108	24.88	100.00
Total	434	100.00	



- Code Data, Cont.
 - Coding and Levels of Measurement
 - Nominal
 - Dummy Code: Empirically Arbitrary, Theoretically Meaningful
 - Ordinal
 - Usually Assumes Equal Distances
 - Interval/Ratio
 - May (Not) Need to Reduce Data
 - e.g. Income
 - 0-\$10,000, \$10,000-\$20,000, etc.



- Code Data, Cont.
 - The Codebook