

Chapter 1: social data analysis basics

1) What is social statistics?

It is a science of using statistical methods to analyze social data.

2) What is the social data?

Examples are GSS (General Social Surveys) from 1972 to current.

CPS (Current Population Survey), PSID, and others

3) What are those statistical software?

SAS, STATA, SPSS, R, and others

4) Data structure

V1 V2 V3 VN

C1

C2

C3

CM

5) Variables: attributes that vary from one case to another

6) Types of variables

A) Nominal variable: variables with groups that cannot be ranked

Gender (man, woman), race, ethnicity, region, religion

B) Ordinal variables: variables with groups that can be ranked, but can not be numerically defined

Social class, satisfaction surveys

C) Interval/ratio variable: variables with groups that can be ranked and numerically measured

Temperature, height, weight, income, education, etc.

D) Dummy variables: variables with two groups.

Gender, you can dichotomize a non-dummy variable