



Manual Annotation Reliability

School of Information Studies
Syracuse University

| Where Does the Ground Truth Come From?

Convenience sample

- Star rating from customer reviews

Manual annotation

| How Trustworthy Is | Manual Annotation?

Reliability test

- If asking two or more people to mark the sentiment of a collection of tweets, to what extent will they agree with each other?

Inter-Coder Agreement

Raw agreement:

- $a = \frac{\text{count}(\text{agreed_items})}{\text{total_items}}$

Problem with raw agreement:

- Skewed categories: 90% raw agreement in both tables

	Coder A		
		Positive	Negative
Coder B	Positive	45	5
	Negative	5	45

	Coder A		
		Positive	Negative
Coder B	Positive	90	10
	Negative	0	0

Cohen's Kappa

a = raw_agreement

c = chance_agreement

$$K = \frac{a - c}{1 - c}$$

	Coder A		
		Positive	Negative
Coder B	Positive	45	5
	Negative	5	45

	Coder A		
		Positive	Negative
Coder B	Positive	90	10
	Negative	0	0

Cohen's Kappa

a = raw_agreement

c = chance_agreement

$$K = (a - c) / (1 - c)$$

~~K=0.8~~

Coder B	Coder A		
		Positive	Negative
	Positive	45	5
	Negative	5	45

K=0

Coder B	Coder A		
		Positive	Negative
	Positive	90	10
	Negative	0	0

How to Calculate Kappa?

Given a confusion matrix of two coders

	Coder A		
		Positive	Negative
	Coder B		
	Positive	45	5
	Negative	5	45

How to Calculate Kappa?

Calculate marginal distribution

	Coder A			
		Positive	Negative	
Coder B	Positive	45	5	50%
	Negative	5	45	50%
		50%	50%	

How to Calculate Kappa?

Calculate raw agreement ($a = 0.9$)

Calculate

- P (both A and B gives “positive” label) = 0.25
- P (both A and B gives “negative” label) = 0.25
- Chance_agreement: $c = 0.25 + 0.25 = 0.5$
- Kappa = $(a - c) / (1 - c) = (0.9 - 0.5) / (1 - 0.5) = 0.4 / 0.5 = 0.8$

What Is Kappa for This?

	Coder A		
		Positive	Negative
Coder B	Positive	89	9
	Negative	1	1

$$K = .139$$

Tools to Calculate Kappa

Online tool

- <http://vassarstats.net/kappa.html>

Statistical tools

- R package “irr”
- SPSS: crosstab function
- `sklearn.metrics.cohen_kappa_score`

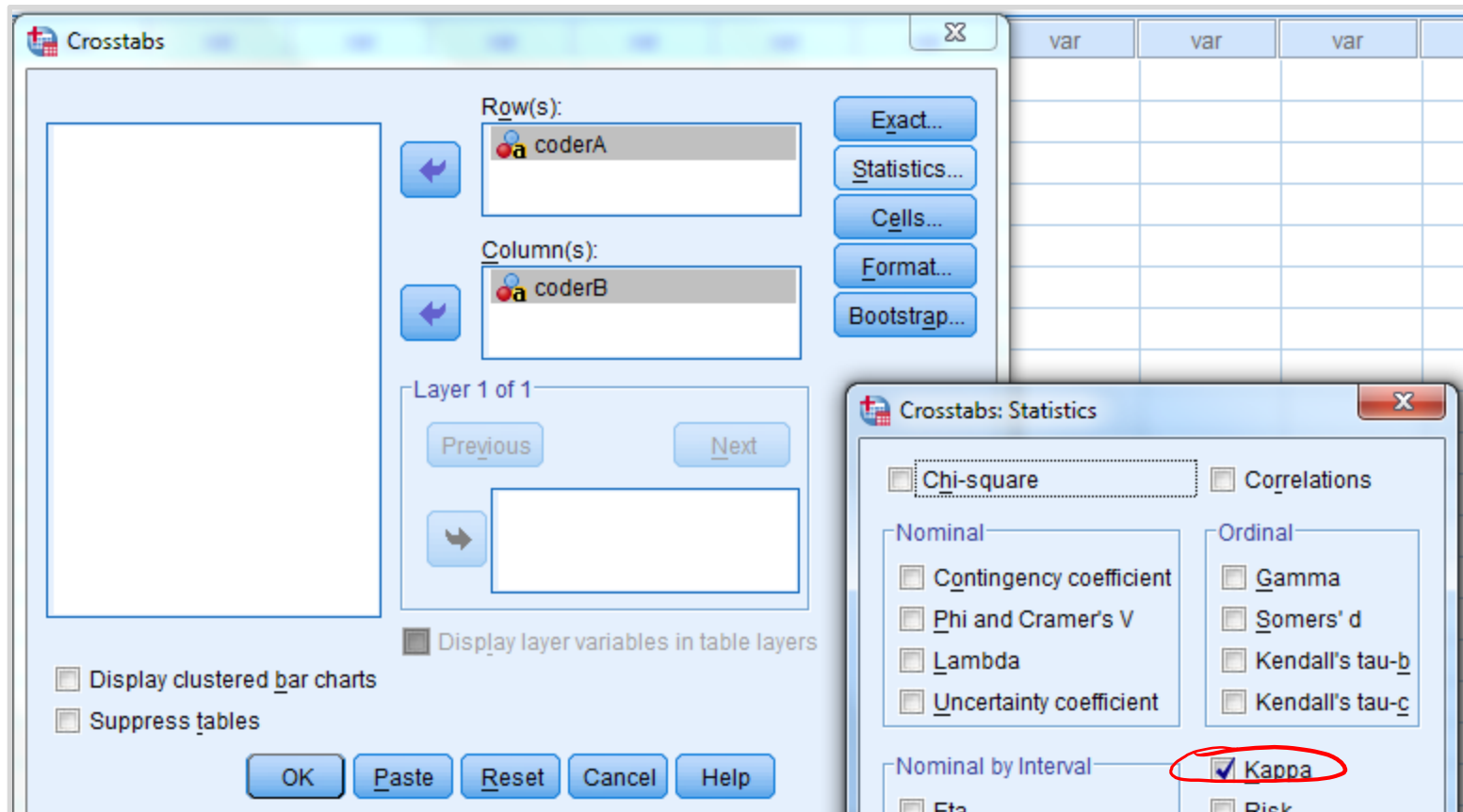
Kappa Calculation in SPSS

The screenshot shows the IBM SPSS Statistics Data Editor window with the file "week5-inter-coder-agreement-sample-data.sav [DataSet0]". The dataset has two columns: "coderA" and "coderB". The data is as follows:

	coderA	coderB
1	pos	neu
2	neg	neu
3	neu	neu
4	pos	pos
5	neg	neg
6	neu	pos
7	neg	neu
8	pos	neg
9	neu	neu
10	neu	neu
11		

The "Analyze" menu is open, and the "Crosstabs..." option is highlighted with a red circle. A red arrow points to the first row of data (row 1) in the dataset.

Kappa Calculation in SPSS



Kappa Calculation in SPSS

coderA * coderB Crosstabulation

Count

		coderB			Total
		neg	neu	pos	
coderA	neg	1	2	0	3
	neu	0	3	1	4
	pos	1	1	1	3
Total		2	6	2	10

Symmetric Measures

		Value	Asymp. Std. Error ^a	Approx. T ^b	Approx. Sig.
Measure of Agreement	Kappa	.219	.229	1.017	.309
N of Valid Cases		10			

a. Not assuming the null hypothesis.

b. Using the asymptotic standard error assuming the null hypothesis.

Multiple Coders

Average Kappa

Krippendorff's Alpha

- <http://afhayes.com/spss-sas-and-mplus-macros-and-code.html>
- Search “kappa”