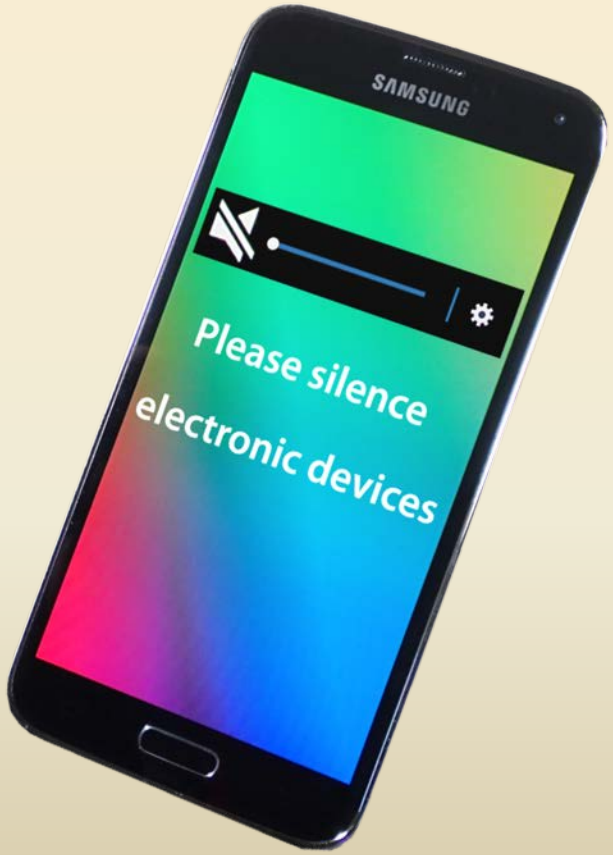
21st Annual GIS/CAMA Technologies Conference Chattanooga Convention Center

Continuing Education (CE) Credit

Recertification Credit forms for CE credit can be collected from the Registration Desk on <u>Thursday</u>.

Housekeeping

The conference proceedings will be available approximately 8 weeks after the conference.



Detecting Biased Samples in Assessment

Representative Samples for both Ratio Studies and CAMA Valuation Model Building

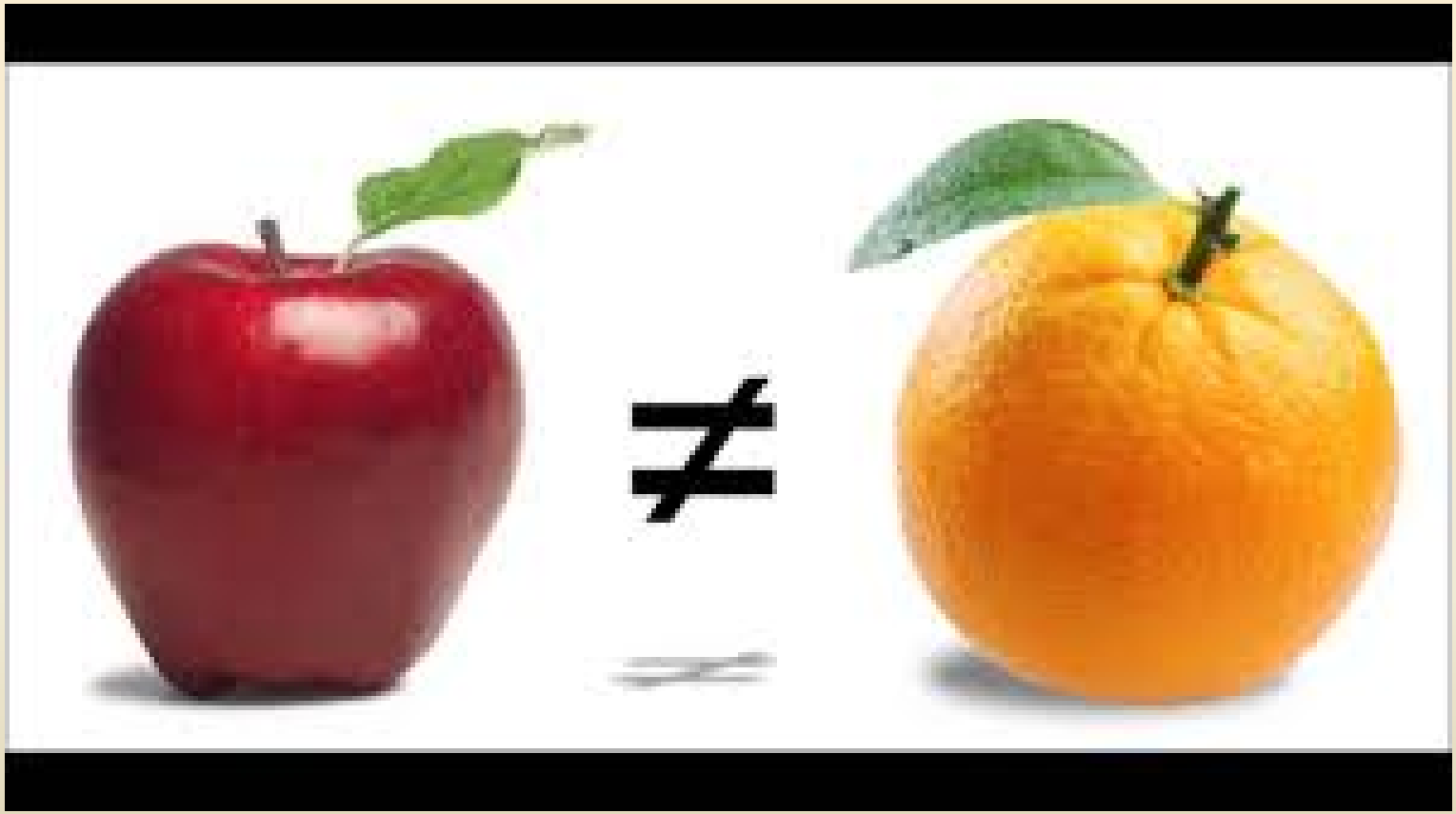
By

Robert Denne, partner, AGJD

At

GIS/CAMA 2017

Obligatory Cheesy Slide



IAAO Standard on Ratio Studies, 2013

- Section 4.5 discussion of representativeness
- Also Section 4.2 in part 2
- 83 appearances of the word representative in std
- Appendix E (2 pages) on Sales Chasing Detection
- Terminology: sale motivated adjustment of value, by means of changed characteristic data (or price)
- Appendix E cites 5 methods re Sales Chasing

Standard Appendix E Methods re SC

1. Comparison of Average Value Changes
2. Comparison of Average Unit Values
3. Split Sample Technique
4. Comparison of Observed versus Expected Distribution of Ratios
5. Mass Appraisal Techniques

SC: practice of changing incommensurately

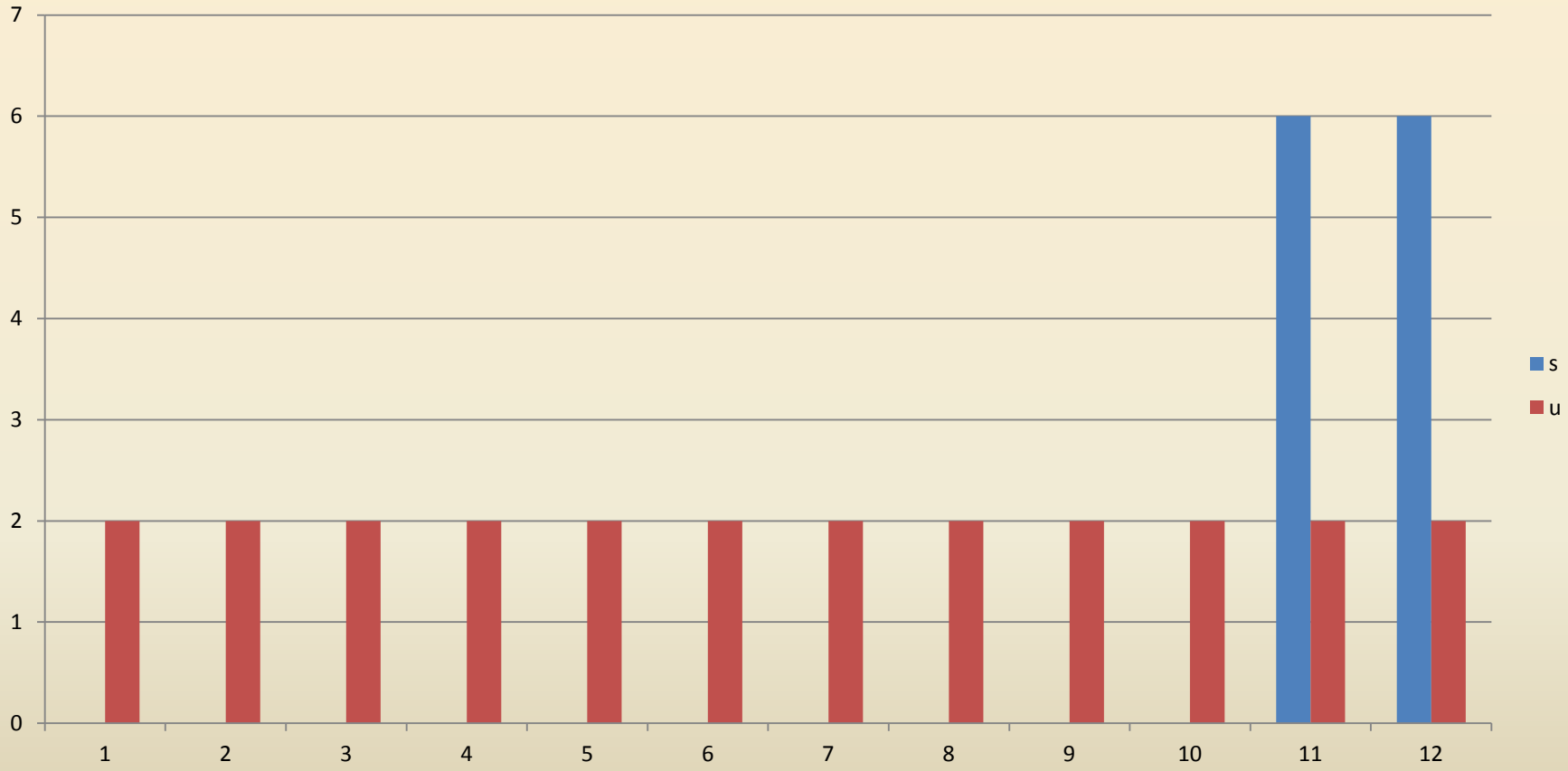
Distinguish (current) *practice* from *results* of past (?) practice

Power of test is affected by level of effort devoted to eliminating noise (exogenously justified changes)

If physical changes are eliminated by noting changed property characteristics year of over year the remaining (endogenous) changes can be attributable to personal responsibility -> retraining

If chgs eliminated due to permits worked, then failures may be more systemically based, or inadvertent.

Distributions vs. Averages/Medians



Issues

Issues

- **Average or Median vs. Entire Distribution**
- **Dollars per Unit vs Percentage Changes**
- **Sale-Motivated Adjustments vs. Total Neglect (vars.)**

Usages known

- **4-R case in early 90s**
- **Status today: 1 large oversight agency and 3-4 original-assessment operations: 3 large, 1 medium**
- **not just for residential properties nor just S Chasers**

Tutorial & Disclaimer

SPSS (IAAO teaching resource & I know) but not Excel

NCSS(the other IAAO teaching resource) also suitable

SAS, R, Python, etc. also possible, but I will demonstrate what I use most.

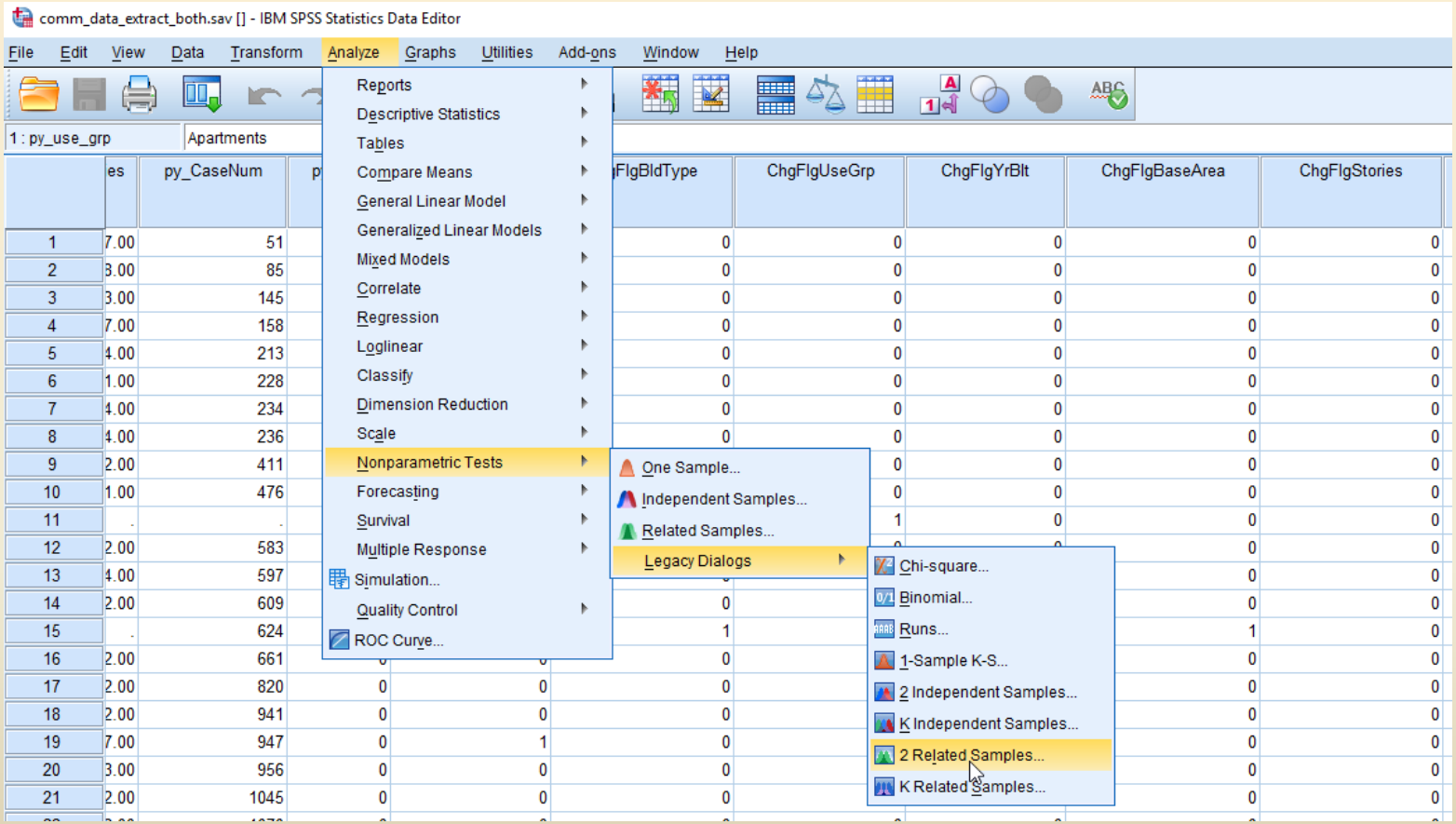
NPAR TESTS

/M-W= PctChgNoticedMktVal1314 BY TestControl(0 1)

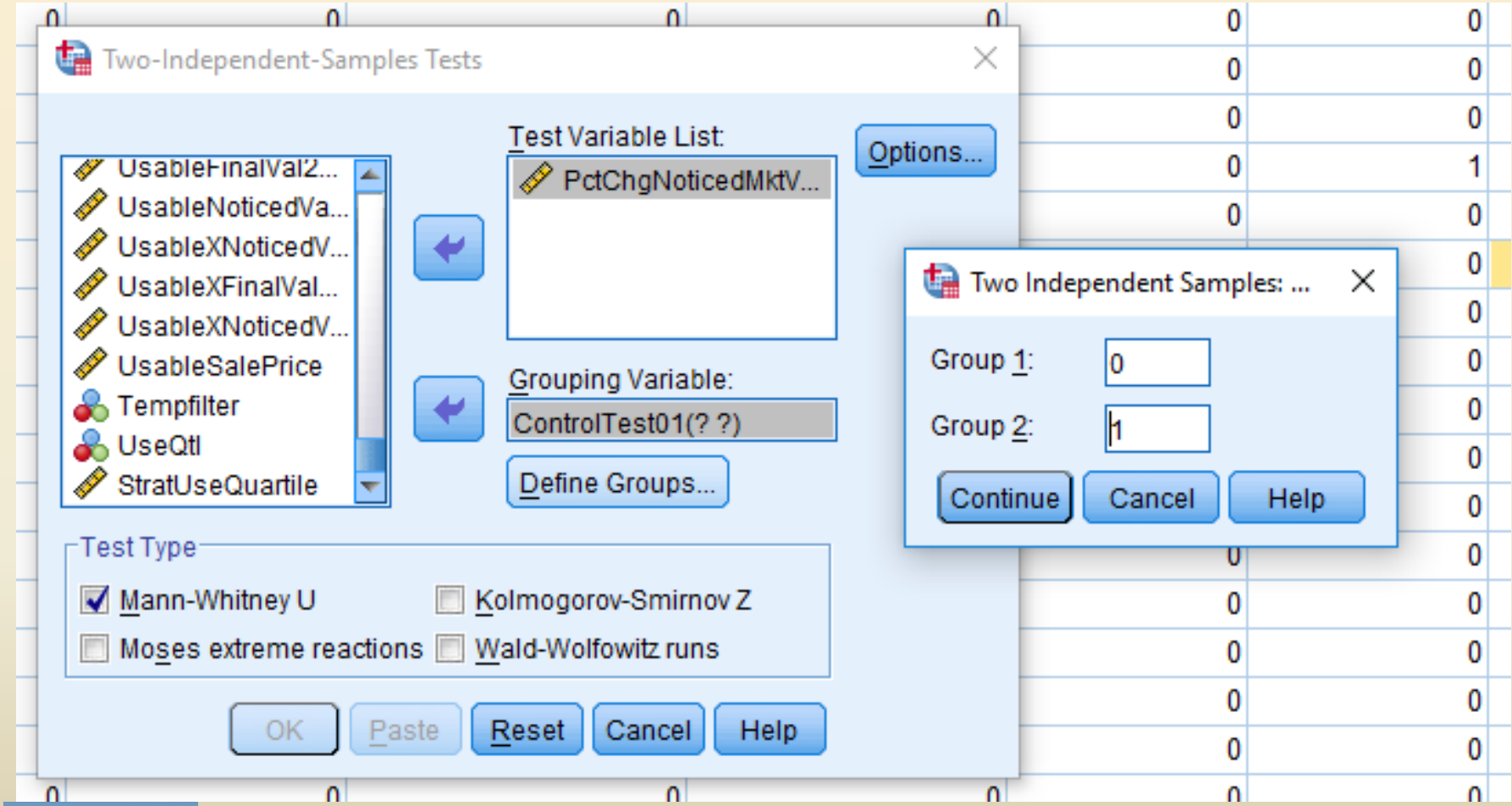
/MISSING ANALYSIS.

*** Highly significant for Apartments, Office Bldgs, Retail, and Warehouses; not signif for Medical Office; inadequate data for vacant.**

GUI-1


comm_data_extract_both.sav [] - IBM SPSS Statistics Data Editor
File Edit View Data Transform Analyze Graphs Utilities Add-ons Window Help
1 : py_use_grp
Apartments
Reports
Descriptive Statistics
Tables
Compare Means
General Linear Model
Generalized Linear Models
Mixed Models
Correlate
Regression
Loglinear
Classify
Dimension Reduction
Scale
Nonparametric Tests
Forecasting
Survival
Multiple Response
Simulation...
Quality Control
ROC Curve...
One Sample...
Independent Samples...
Related Samples...
Legacy Dialogs
Chi-square...
Binomial...
Runs...
1-Sample K-S...
2 Independent Samples...
K Independent Samples...
2 Related Samples...
K Related Samples...
py_CaseNum
ChgFlgUseGrp
ChgFlgYrBlt
ChgFlgBaseArea
ChgFlgStories

GUI-2


Two-Independent-Samples Tests
UsableFinalVal2...
UsableNoticedVa...
UsableXNoticedV...
UsableXFinalVal...
UsableXNoticedV...
UsableSalePrice
Tempfilter
UseQtl
StratUseQuartile
Test Variable List:
PctChgNoticedMktV...
Options...
Grouping Variable:
ControlTest01(? ?)
Define Groups...
Test Type
Mann-Whitney U
Kolmogorov-Smirnov Z
Moses extreme reactions
Wald-Wolfowitz runs
OK
Paste
Reset
Cancel
Help
Two Independent Samples: ...
Group 1:
0
Group 2:
1
Continue
Cancel
Help

Crosstabulation Output

UseGrpN * TestControl Crosstabulation

Count

		TestControl			
		0 Not recently sold	1 Sold in Study Period	9 Exogenous chg; Excluded from test	Total
UseGrpN	1 Apartments	4943	94	721	5758
	3 Office Bldgs	3690	59	293	4042
	4 Retail	26519	205	2491	29215
	6 Warehouse	15780	149	1743	17672
Total		50932	507	5248	56687

Mann-Whitney Output (Raw)

			Sum of Ranks
PctChgNoticedMktVal1314			

UseGrpN = 1 Apartments

	PctChgNoticedMktVal1314

UseGrpN = 1 Apartments

Grouping Variable: TestControl

Simpson's Paradox: the context of the test

Read Wikipedia article (aka Agglomeration paradox)

Famously arose re sex discrimination in admissions at UC Berkely

Similar problems can arise in assessment when confounding (residential) neighborhoods that differ from one another, as by age of construction or rates of turnover.

Solution: perform test at lowest, most specific level possible.

Problem not unique to our use, shared by all stats.

Switch to Excel

Line#	Neighborhood within Township & Major Class	Median Percent Change In Assessment For Sold Properties	Sample Size For Parcels Sold In Study Period	Median Percent Change In Assessment For Unsold Properties	Sample Size For Parcels Not Sold In Study Period	Diffference in Median Percentage Changes: Sold - Unsold Parcels	Mann-Whitney Likelihood Of Such Differences In Percentage Changes in the Absence of Different Treatments of the Groups
1	ResidImpr,200150	6	14	6	73	0	0.786
2	ResidImpr,200300	-7	6	-7	107	0	0.573
3	ResidImpr,201000	2	9	2	167	0	0.278
4	ResidImpr,201600	7	8	7	150	0	0.152
5	ResidImpr,202000	1	11	1	284	0	0.010
6	ResidImpr,202300	2	8	2	169	0	0.409
7	ResidImpr,202600	0	10	1	191	-1	0.147
8	ResidImpr,202650	4	5	4	47	0	0.525
9	ResidImpr,202850	3	11	3	98	0	0.478
10	ResidImpr,202900	7	7	7	75	0	0.784
11	ResidImpr,203650	9	6	9	83	0	0.768
12	ResidImpr,203655	-1	14	-1	177	0	0.591
13	ResidImpr,203656	5	8	5	131	0	0.426
14	ResidImpr,204500	0	9	0	164	0	0.000
15	ResidImpr,204900	0	5	0	122	0	1.000

Many tests, Nicely formatted, but …

Line#	Neighborhood within Township & Major Class	Median Percent Change In Assessment For Sold Properties	Sample Size For Parcels Sold In Study Period	Median Percent Change In Assessment For Unsold Properties	Sample Size For Parcels Not Sold In Study Period	Diffference in Median Percentage Changes: Sold - Unsold Parcels	Mann-Whitney Likelihood Of Such Differences In Percentage Changes in the Absence of Different Treatments of the Groups
1	DemoTwp,ResidImpr,200150	6	14	6	73	0	0.786
2	DemoTwp,ResidImpr,200300	-7	6	-7	107	0	0.573
3	DemoTwp,ResidImpr,201000	2	9	2	167	0	0.278
4	DemoTwp,ResidImpr,201600	7	8	7	150	0	0.152
5	DemoTwp,ResidImpr,202000	1	11	1	284	0	0.010
6	DemoTwp,ResidImpr,202300	2	8	2	169	0	0.409
7	DemoTwp,ResidImpr,202600	0	10	1	191	-1	0.147
8	DemoTwp,ResidImpr,202650	4	5	4	47	0	0.525
9	DemoTwp,ResidImpr,202850	3	11	3	98	0	0.478
10	DemoTwp,ResidImpr,202900	7	7	7	75	0	0.784
11	DemoTwp,ResidImpr,203650	9	6	9	83	0	0.768
12	DemoTwp,ResidImpr,203655	-1	14	-1	177	0	0.591
13	DemoTwp,ResidImpr,203656	5	8	5	131	0	0.426
14	DemoTwp,ResidImpr,204500	0	9	0	164	0	0.000
15	DemoTwp,ResidImpr,204900	0	5	0	122	0	1.000

Multiple Tests Require Special Handling

Benjamini-Hochberg False Discovery Control Calculations

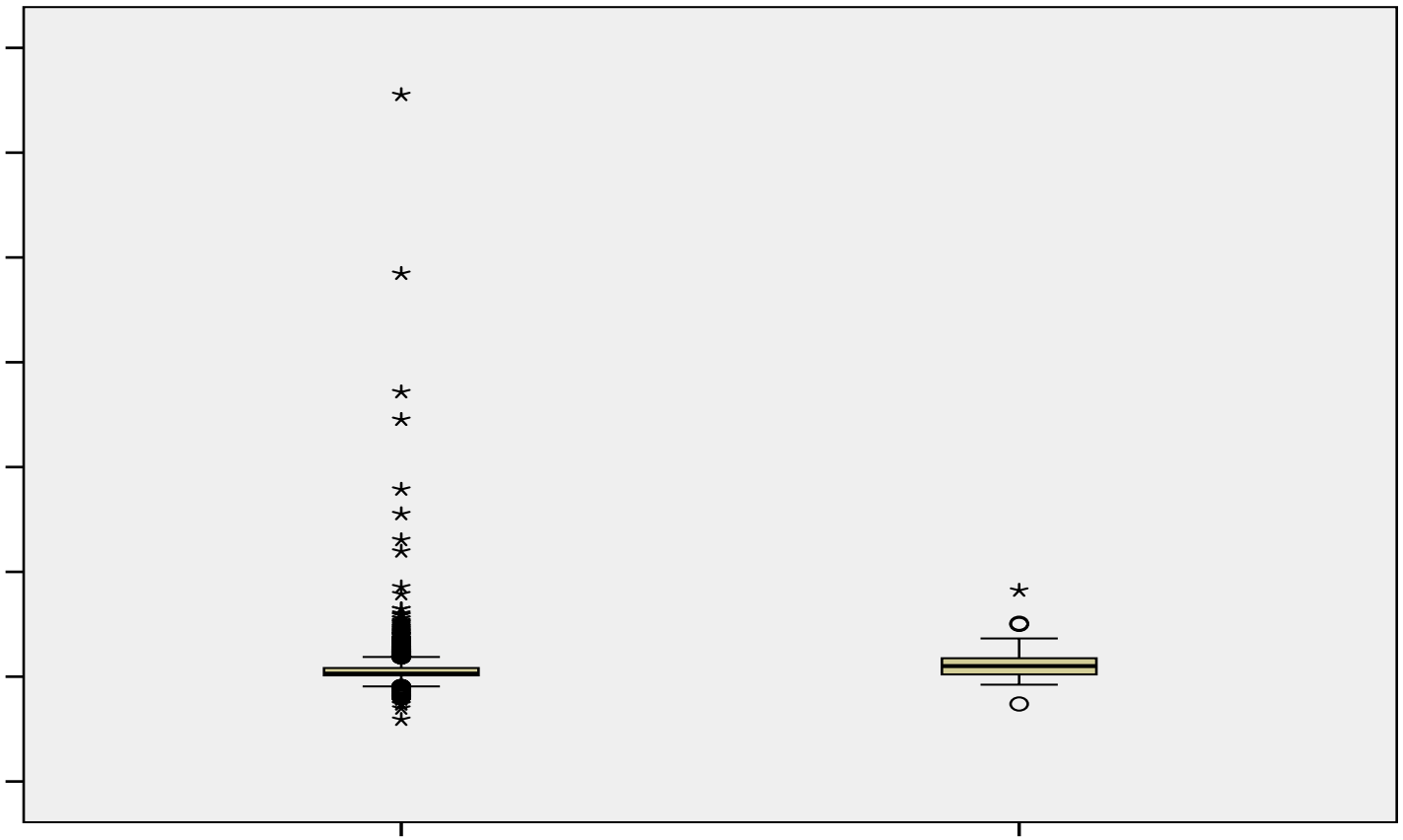
Input Line#	Input P-Value	Sorted by P-Value: i	Correction formula: iq/m	False Discovery?
26	0.00000005	1	0.0004274	0
83	0.00007517	2	0.0008547	0
14	0.00012384	3	0.0012821	0
67	0.00019967	4	0.0017094	0
60	0.00754256	5	0.0021368	1
5	0.00974602	6	0.0025641	1
66	0.01052653	7	0.0029915	1
39	0.01725026	8	0.0034188	1

Note 1	Note 2
0.05000000	q: false discovery rate
117	m: number of tests
	4 significant results given
Conclusion:	multiple tests

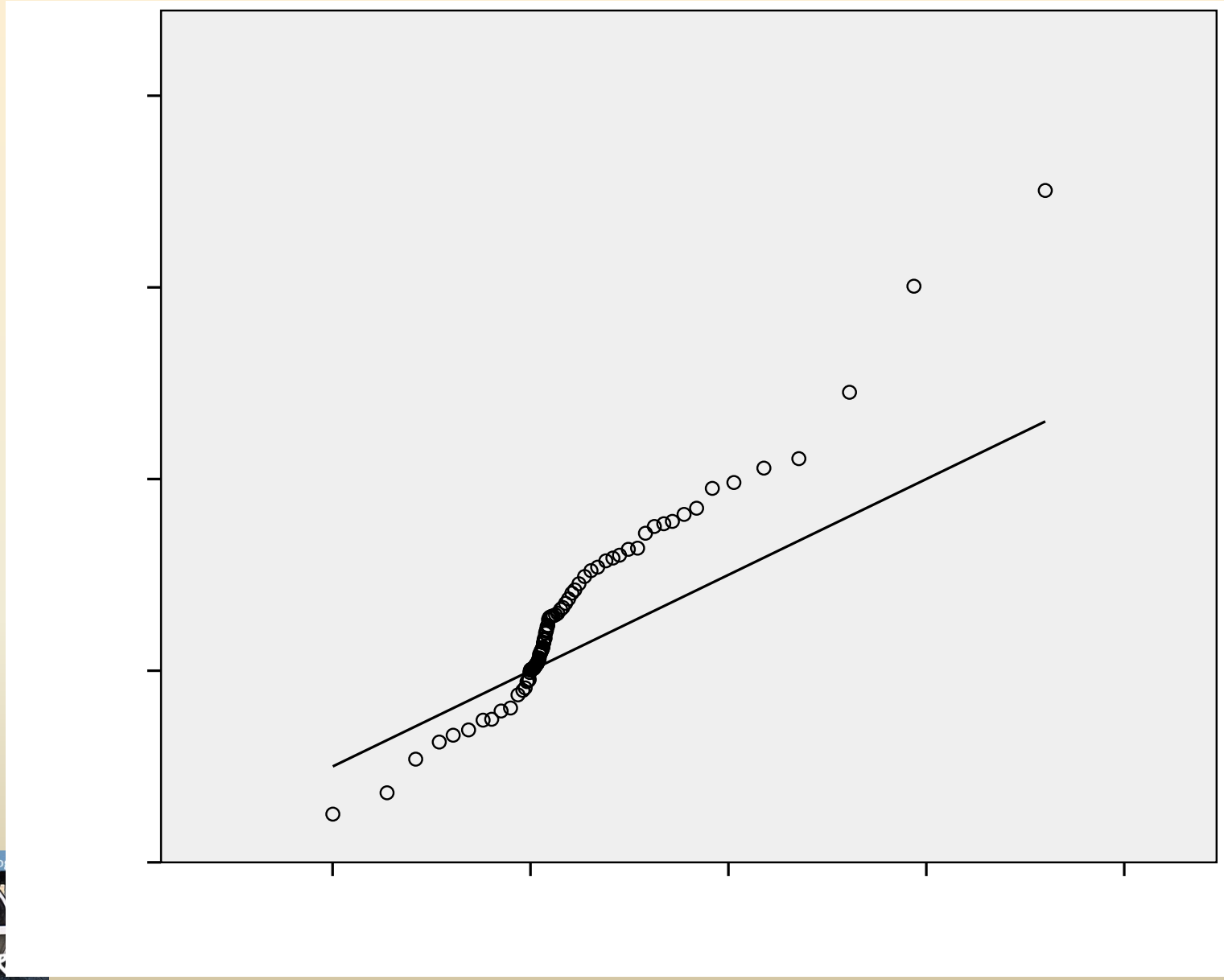
Benjamini-Hochberg False Discovery Control Calculations

InLine	InPValue	i	iq/m		
26	0.00000005	1	0.0004274	0.05	q: false discovery rate
83	0.00007517	2	0.0008547	117	m: number of tests
14	0.00012384	3	0.0012821	Conclusion	4 significant results given multiple tests
67	0.00019967	4	0.0017094		
60	0.00754256	5	0.0021368		
5	0.00974602	6	0.0025641		
66	0.01052653	7	0.0029915		
39	0.01725026	8	0.0034188		

Boxplot Graphics Encourage Understanding



Quantile-Quantile Graphic



Biases from cherry picking

Limited applicability of statistical tests to detect, but an important problem to guard against

Can be discouraged by policies that maximize "sunshine": controlled deed recordation, info to both local and oversight agency, posted on internet, filing requires completion of form with certain minimal sale/qualification questions, etc. Deed-image access would be ideal.

Biases from inclusion of new construction

Often appropriately excluded

If included, should be weighted to reflect only themselves, not (nonexistent) other(s) unsold

Biases in model building: omitted variables or combinations, especially w.r.t. location

This is a big problem with post reassessment reviews

Recall results from GIS-CAMA contest nine years ago won by users of 20/25 years of sales, esp. to model location in a triumph of time adjust + primitive GWR

Recommendations

1. Testing pctg changes in assessment, with statistical significance judged by a M-W test, is well worth trying. Recall level- of-detail and false-discovery considerations. Easily repeatable once set up (not hard) and produces quantifiable objective results.
2. Involvement of 3rd parties (oversight and public) can diminish incidence of cherry picking. Also value in documentation and possibility of audit
3. Multi-parcel sales are invaluable and should receive special care. CAMA systems not always capable…

Recommendations continued

4. Not all sales equally important, e.g. high vs low values & new construction esp. wrt residential (weighting?). CAMA systems should provide multiple qualification alternatives & address transaction and timing issues in qualification re purposes.
5. Larger samples and longer sample periods increase the value of the sample, albeit at a diminishing rate. Sales qualification even in the absence of routine reassessment is invaluable. Reliable matching via unique parcel identifiers is essential.
6. Performance criteria for reassessment contracts should specify appropriate sampling periods for the model building datasets and should include provision for subsequent testing against sales not included in the model-building dataset.

Recommendations concluded

7. **Ratio study trimming s/b based on IQR per IAAO standard, but consider deviations as a percentage rather than a raw-number basis. Thus if the median ratio were 1.0, two ratios of 4.00 and 0.25 should be recognized as being equally distant from the median, rather than the first being considered much farther from it than the latter. This is most easily implemented by trimming using the logarithms of the ratios, rather than the ratios themselves. If logarithms are not used, the effect of trimming will potentially be to bias the sample by preferentially eliminating over assessments and retaining most, if not all under assessments. Not everyone agrees.**

Thanks for your attention!

Questions ?

Contact info:

rcdenne@uchicago.edu

agjd.com


GIS-Pro 2017
October 23-26, 2017
Jacksonville, Florida


Assessment Leadership
Beyond All Limits
IAAO LAS VEGAS
ANNUAL CONFERENCE
& Exhibition
September 24-27
2017
MGM