

# Territory Analysis Updates to the Traditional Methods

CAS RPM

March 22, 2011

Sandra Ross, FCAS, MAAA, CIC

# Agenda

---

- State of territory definitions today
- Reasons for modifying territories
- Considerations
- Processes
  - Data
    - Availability and collection
    - Capping
    - Smoothing
    - Combining
  - Clustering
  - Selecting

# Current Definitions

---

- Current sets
  - Often outdated
  - Uniform across product/policy
  - Less than optimal match of exposure
  - Developed in less than optimal ways
    - Technique
    - Basis for definitions
  - Tweaked over time
- Possibly leading to:
  - Misclassification
  - Misinterpretation of other factors
  - Adverse selection

# Changing Landscapes

---

- Anyone else notice where there used to be a crop planted there is now a subdivision or a strip-mall?
- Over a 20-year period (1970-1990), the 100 largest urbanized areas in the United States sprawled out over an additional 14,545 square miles. That is more than 9 million acres of natural habitats, farmland and rural areas that have been converted to subdivisions, shopping centers, etc.
- What has happened since 1990?
  - Increased population density
  - Increased vehicle density
  - More new homes
  - Less populations in cities, more abandoned homes

# Indianapolis

- 14 largest city in the U.S. according to 2010 Census
- 3<sup>rd</sup> largest in the Midwest
- One of the fastest growing regions in the Midwest.

<u>County</u>	<u>City</u>	<u>Population</u>	<u>Pop Chg</u> <u>4/1/00 -</u> <u>7/1/09</u>
Marion	Indianapolis	785,597	0.5%
	Remainder	105,282	33.2%
	Total	890,879	3.5%
Boone		56,287	22.1%
Hamilton		279,287	52.8%
Hancock		68,334	23.4%
Hendricks		140,606	35.1%
Morgan		70,876	6.3%
Johnson		141,501	22.8%
Shelby		44,503	2.4%
All Other		4,730,840	26.2%
Indiana		6,423,113	5.6%
http:\quickfacts.census.gov as of 3/3/11			

# Geographic Rating

---

- Goal is to isolate variables to explain risk
- Use variables to segment property insured, coverage selections and insured characteristics
- Territory is used to explain differentiation in risk not picked up by other rating variables and to explain geographic differences
- Geographic difference can be due to
  - Population and vehicle density
  - Theft/crime rates
  - Hazards
  - Differences in mix of business
    - Properties insured
    - Vehicles driven

# Deriving Territory Definitions

---

- Territory definition analysis is driven by a lot of numbers, analysis, statistical techniques, etc.
- However, there are still many areas where actuarial judgment plays an important roles



# Upfront Considerations

---

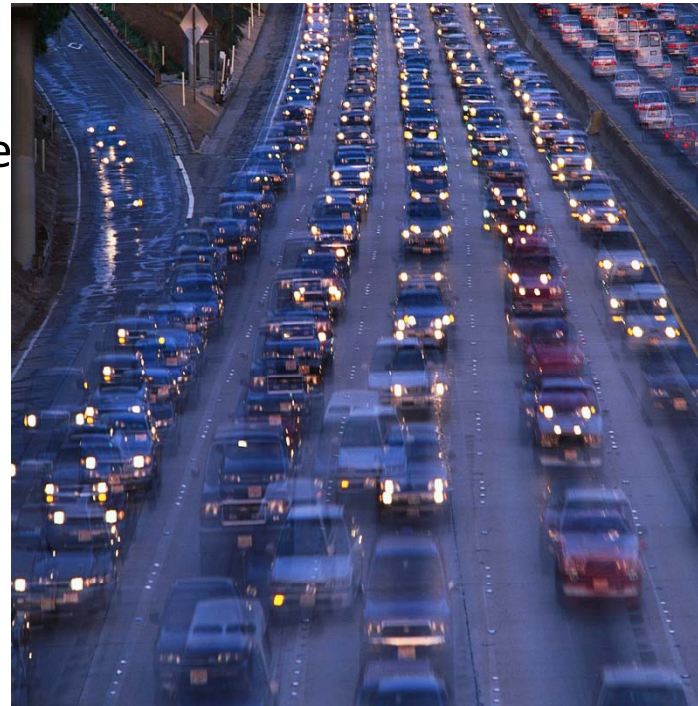
- State regulations
  - Ex. OH must rate by city
- Types of analysis
  - Total state/line
  - By coverage/peril
  - Contiguous or not
- Basis for analysis
  - Zip Code
  - Census Tract
  - Other
- System capabilities
- Available data
  - Internal
  - External
- Historical events
  - Desire to remove or adjust for them
- Specific concerns
  - Management
  - Sales
- Competitive pressures and competitor boundaries



# Increased Segmentation in Definitions

---

- Auto
  - Territories by coverage
  - Territories by coverage group
  - Territories by peril for Comprehensive
- Home
  - Territories by peril
  - Territories by peril group
  - Territories by coverage
- Loss Components
  - Pure Premium
  - Frequency
  - Severity



# External Data

- Historical Insurance Industry data
  - ISO
  - HLDI
- Hazard data providers
- Census and other governmental data
  - Housing density
  - Traffic density
  - Crime statistics
  - Accident statistics
  - Home values
- Catastrophe Model Output



# Basis for Data

---

- Statistics by
  - County
  - Zip Code
  - Census Block
  - Census Tract
  - Address
- Location
  - Longitude
  - Latitude
  - Adjacency



# Industry Data

---

## ■ ISO

- Auto
  - By coverage
  - Cat indicators
- Home
  - By cause of loss
  - By coverage
  - Cat indicators
- Data by zip

## ■ HLDI

- Auto
- Free to members
- More than 25 years
- By coverage
- Comprehensive broken into fire, theft, glass and other
- Data by zip

# How much data is necessary?

---

- Non-catastrophe
  - Generally 5-10 years depending on credibility of data
- Catastrophe
  - Much longer periods if available
  - HLDI provides over 25 years
- Cat Modelers
  - Represents hundred's of years of experience and forecast of future events

# Accounting for Catastrophes

---

- Company data
  - Usually cat and x-cat available
  - May not coincide with industry coding
- ISO
  - Cat and x-cat data
- HLDI
  - Comprehensive other than Fire, Theft and Glass
- Cat model data
  - AIR and RMS models
  - Wind/Hail models
  - Winter storm models
- Hazard data
  - Sinkholes
  - Distance to coast

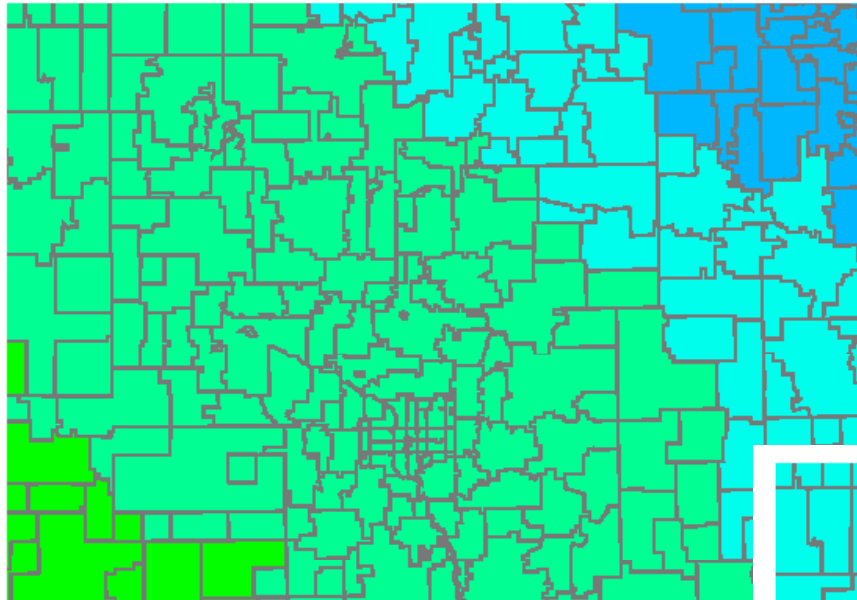


# Data Adjustments to Consider

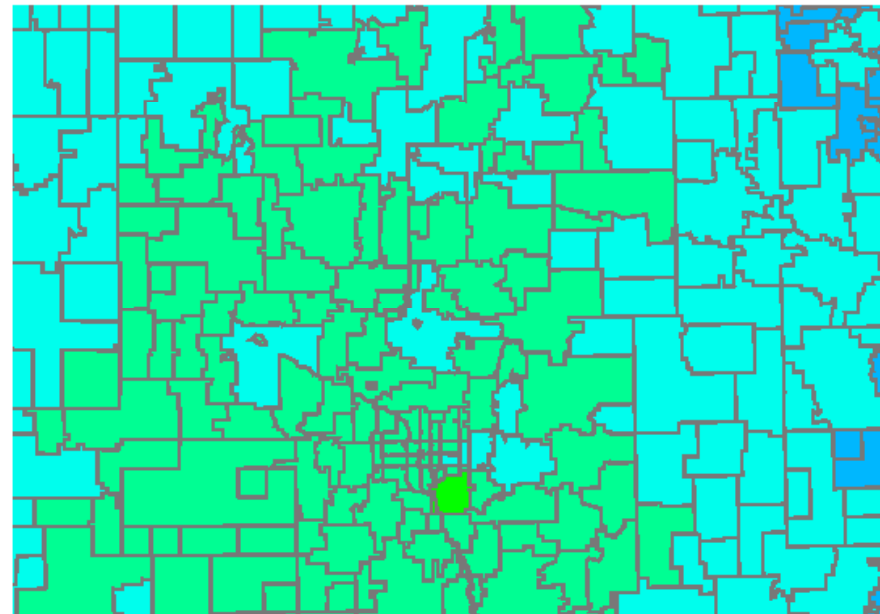
---

- Eliminate effect from all other rating variables
- Capping
- Smoothing
- Possible clustering of partial components to add further smoothing (i.e. cluster cat component before combining with non-cat)
- Normalizing
- Inflationary adjustments
- Weighting together of various data sources

# Modeling Output

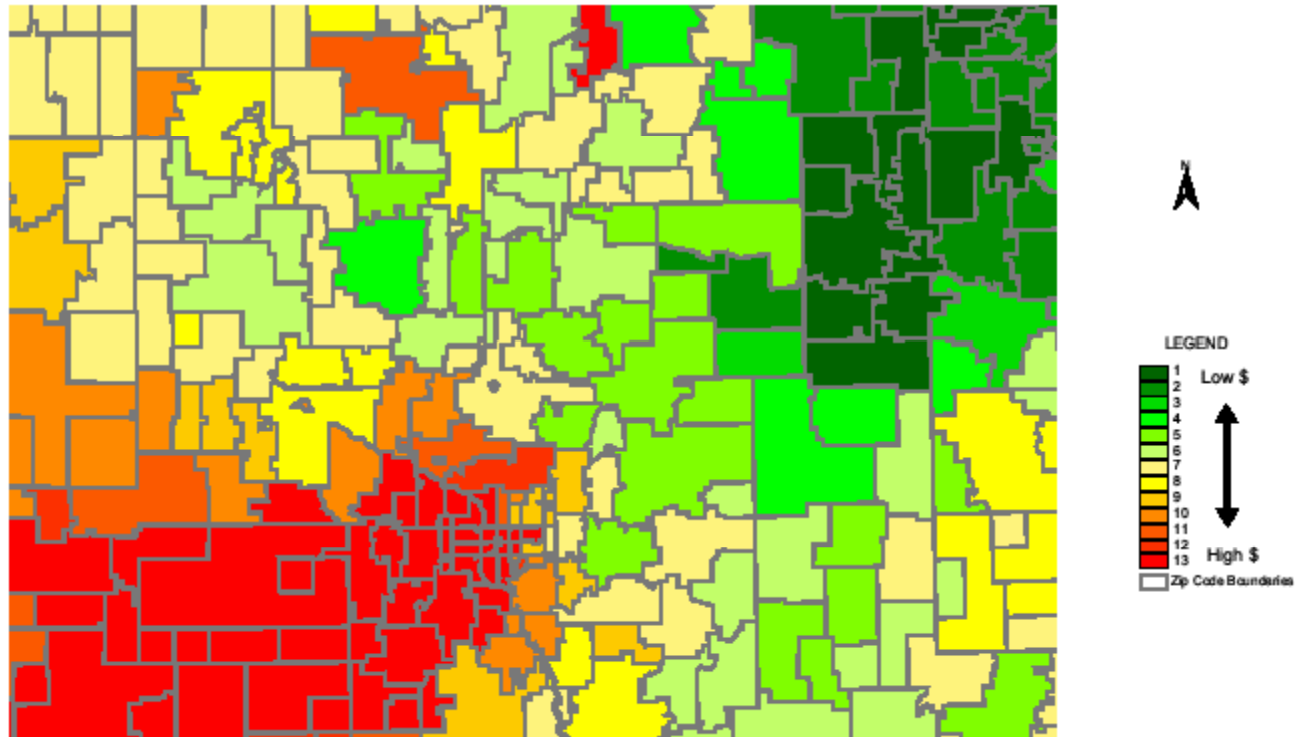


- Cat modeler output can provide very different results





# Historical Experience



- Model results can also be quite different from historical experience

# Capping

---

- Used at various places in process
- Average rating factors
- Large individual losses
- Large events or catastrophes

# Territories by Coverage and Peril

---

- Separate definition sets by coverage or peril provide more optimal rate classification and factors
  - Geographic location may not uniformly impact coverage or peril
- Similar process for frequency/severity separate analysis
- There are ways to develop territory sets by coverage or peril and combine the sets into one consolidated set
  - May ease systems implementation

# Auto Components

---

Liability &  
Collision

Comprehensive

Company

Industry

Non-Cat

Cat



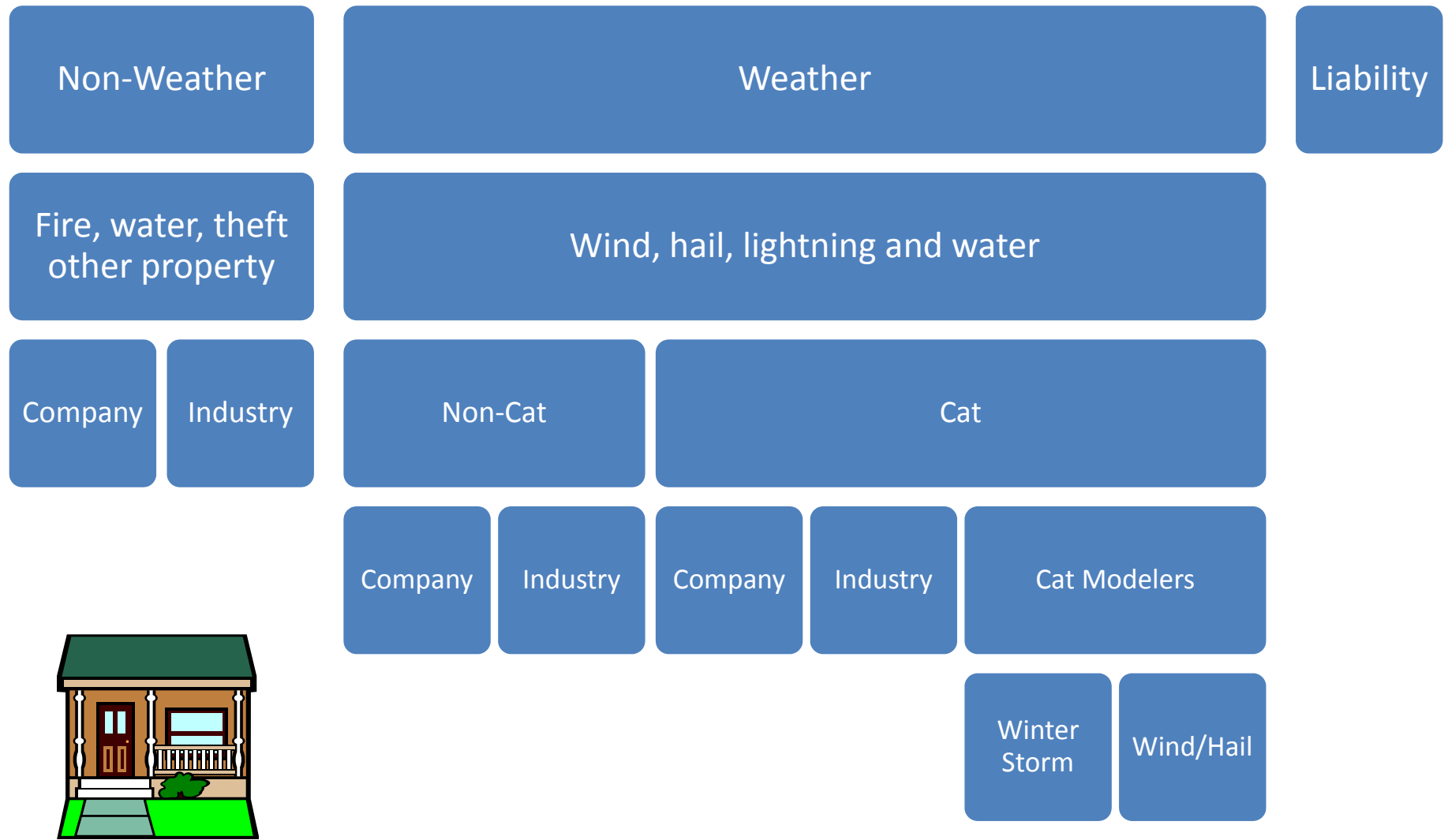
Company

Industry

Company

Industry

# Home Components



# Average Rating Plan Factors

---

- Adjustment of historical experience to a common level
- Removes distributional biases from the underlying data
- Assisted by generalized linear models
- Rating variables such as:
  - Age of driver
  - Insured Value of Homes
  - Protection Class
  - Deductible
  - Discounts
  - Claims surcharge

# Smoothing

---

- Data at the basic element level lacks “credibility”
- Smoothing process allows inclusion of more localized data rather than statewide information
- Results in a rate or rate relativity for each individual zip code based upon the data within that zip code modified as necessary to include a significant number of observations



# Smoothing

---

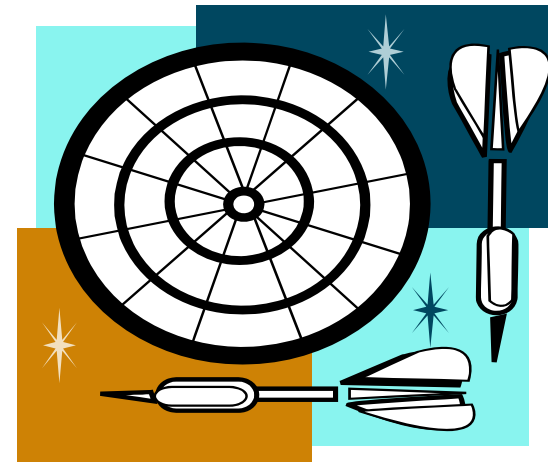
- Key smoothing variables
  - Predictive value of local data
  - Identification of complement data
  - How many observations are required to smooth
  - How far to allow smoothing search to continue
- Many equations are available to combine local data with surrounding information
  - Exposure Weighted Average
  - Straight Line Declining Distance formula
  - Squared Declining Distance formula
  - Werland-Christopherson Method



# Smoothing Considerations

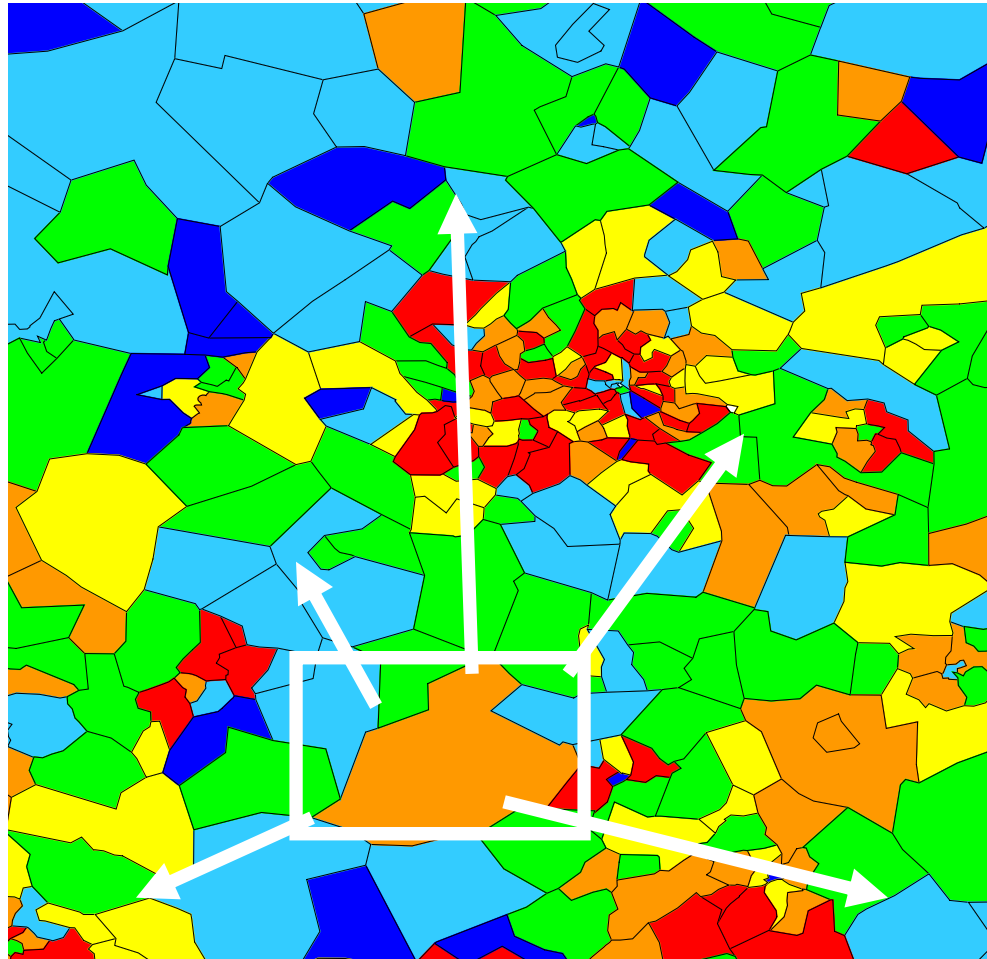
---

- State Borders and Corners
- Use of smoothing across state boundaries
- Potential separate smoothing of urban and rural areas
- Distance based smoothing process or contiguous based smoothing process



# “Neighboring”

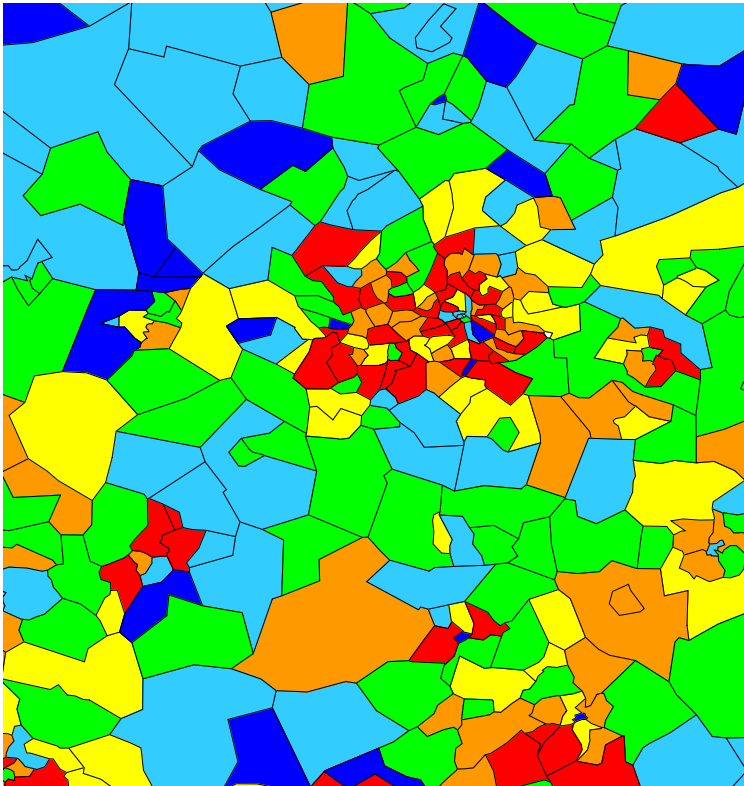
---



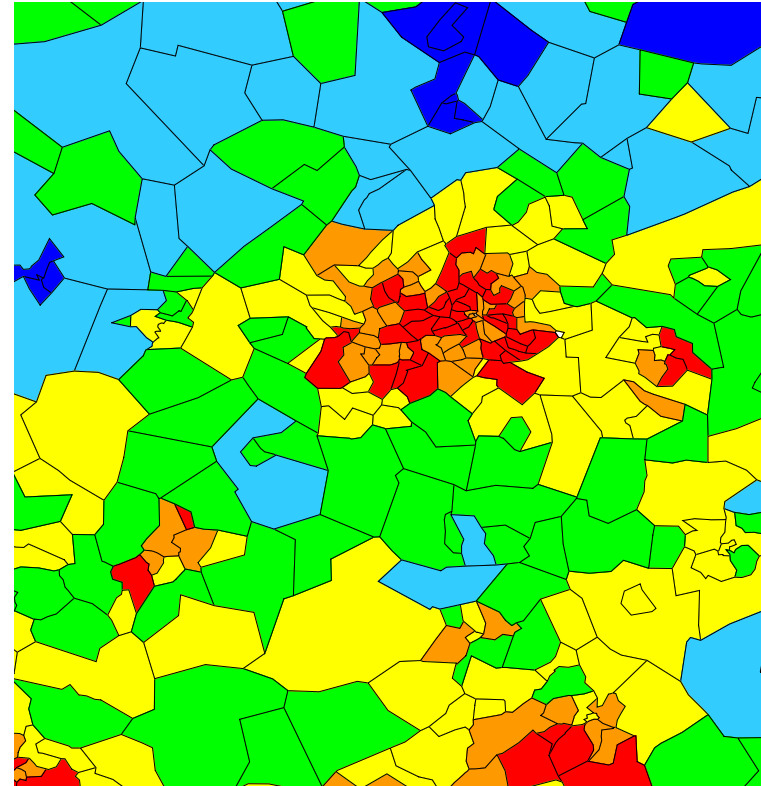
# Smoothing Impact

---

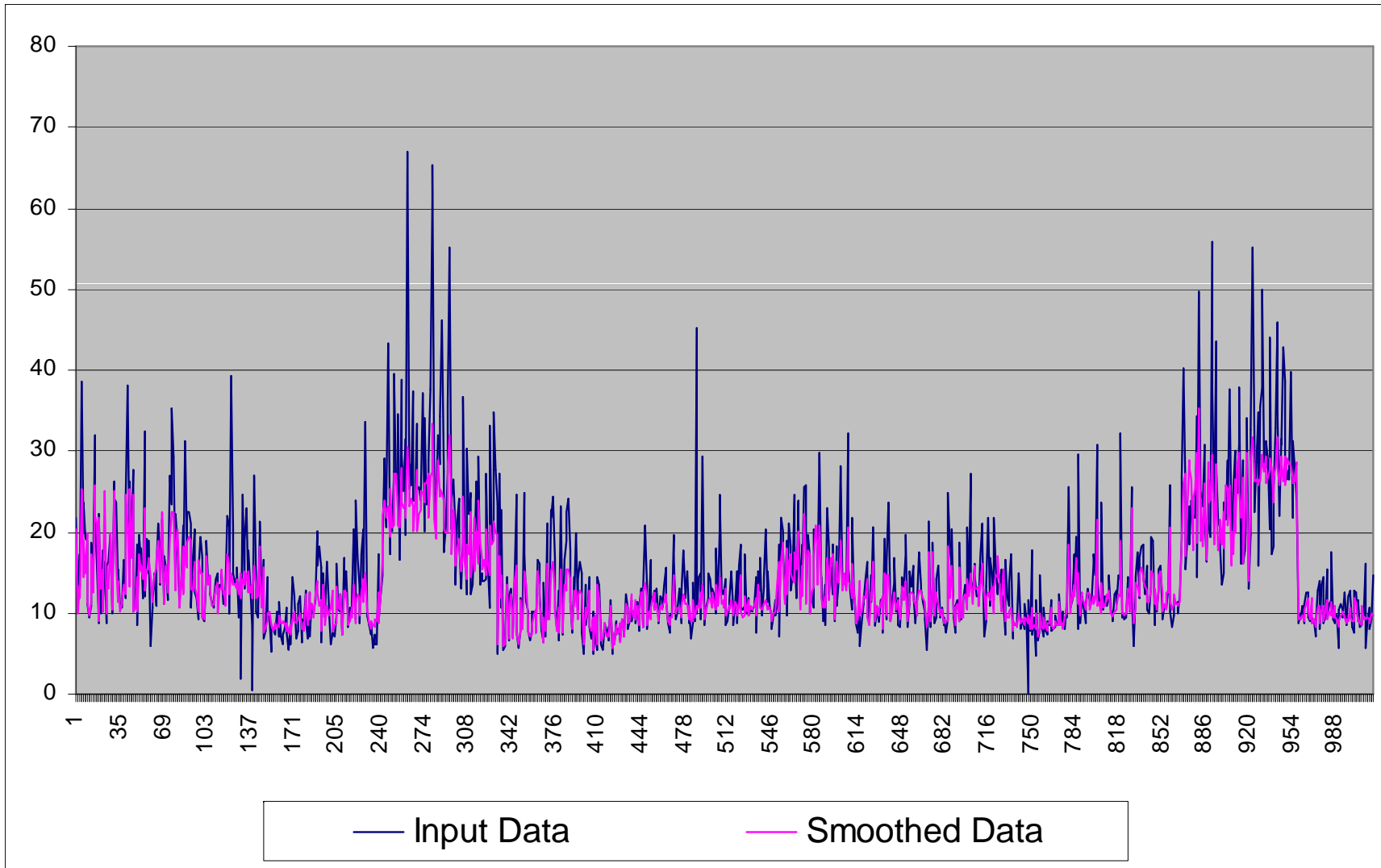
Unsmoothed data



Smoothed data



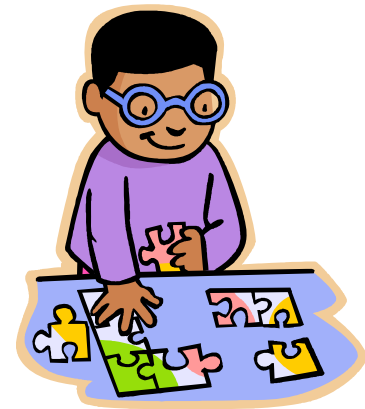
# Smoothing Impact



# Clustering Process

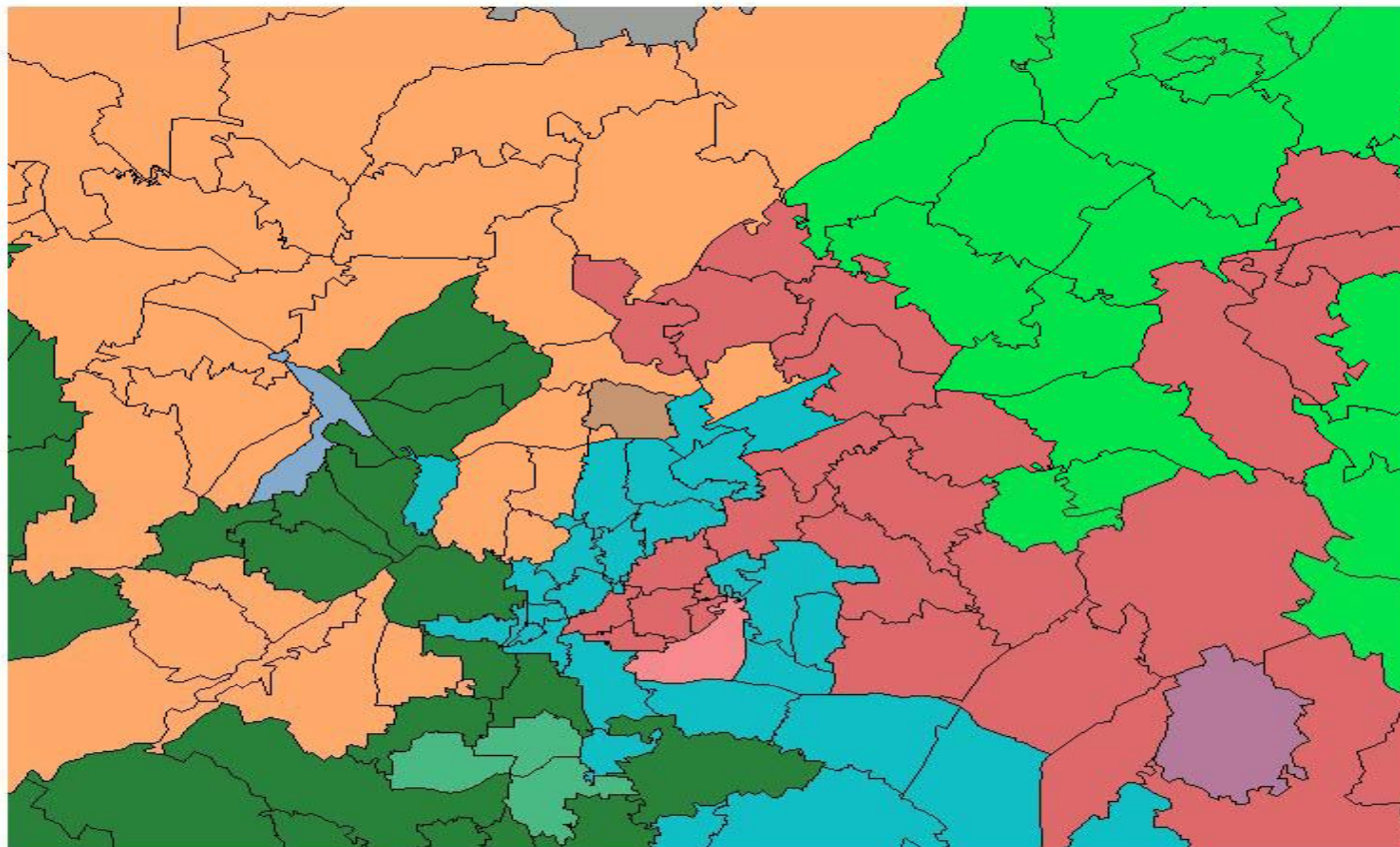
---

- Grouping of areas based on similarity of statistics
- Begin with most detailed data and combine – bottom up approach
- Comparison can be based on percentage or value differences
- Contiguity can be a constraint
- Summitt™



# Contiguous Clusters

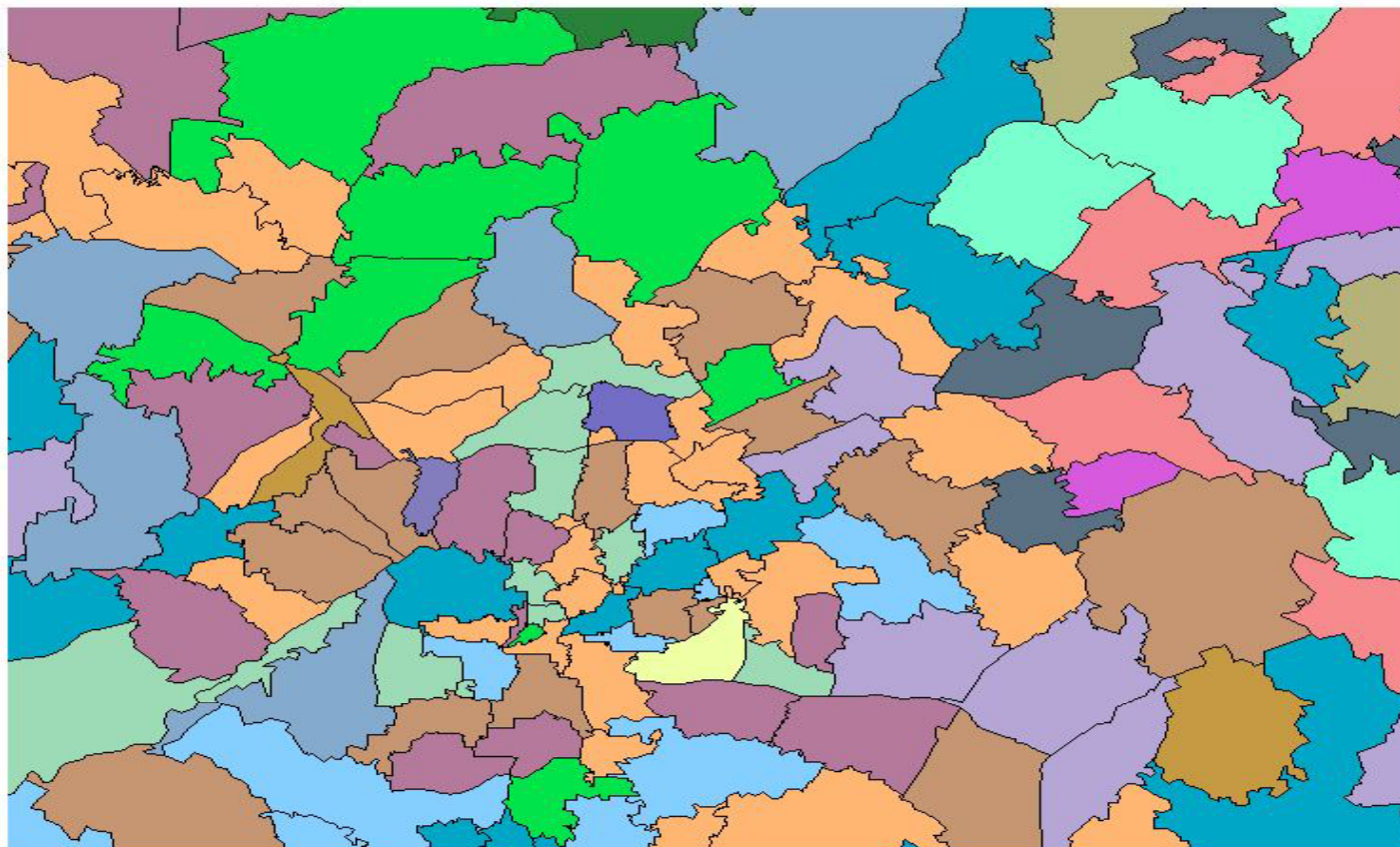
Contiguous Clustering



Pinnacle Actuarial Resources, Inc. 2005

# Non-Contiguous Clusters

Noncontiguous Clustering



Pinnacle Actuarial Resources, Inc. 2005

# Optimal Cluster Evaluation

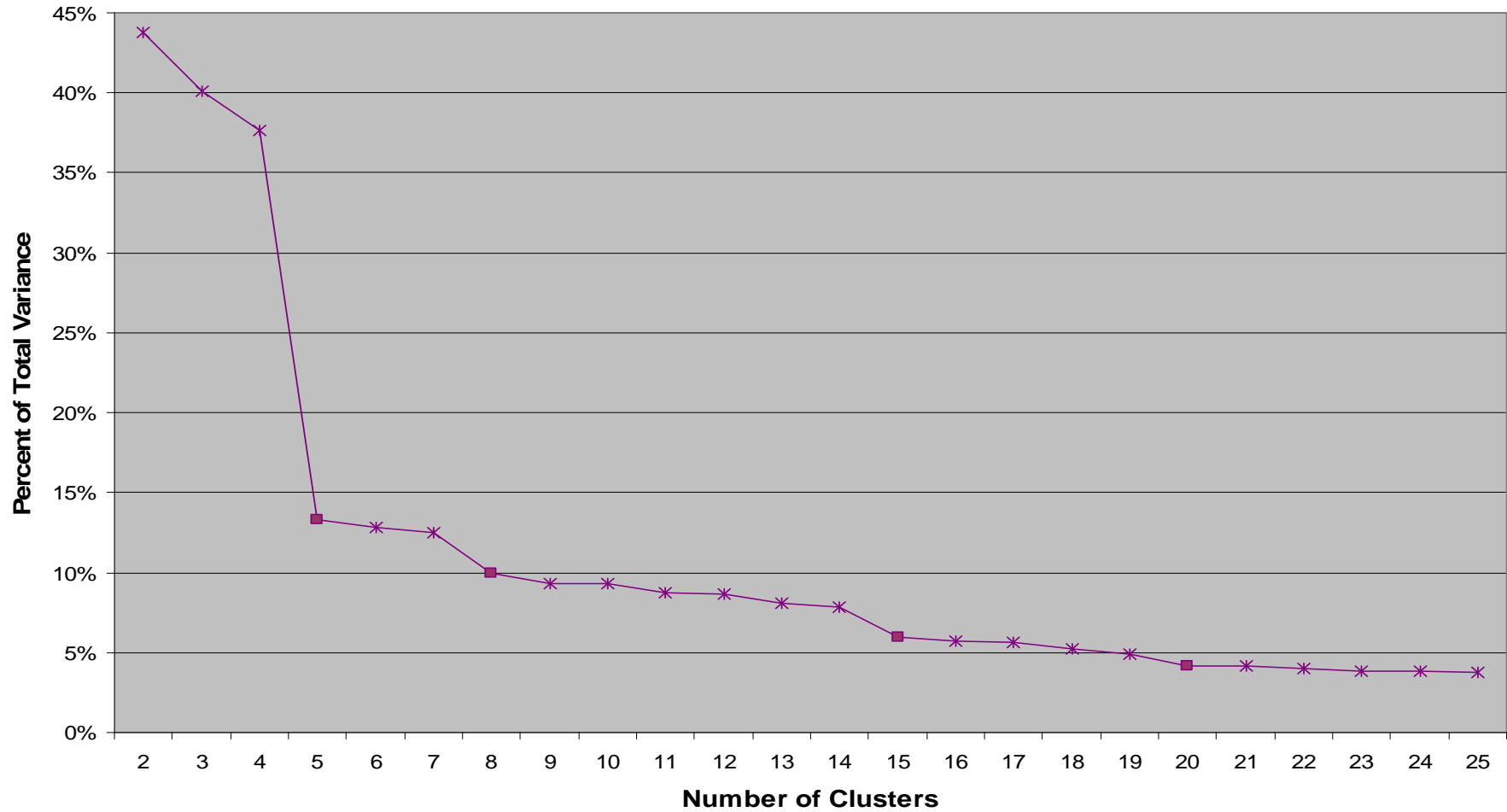
---

- Selection of Target Optimal Clusters for use in establishing territories based on analysis of variance data
- Goal
  - Risks within territory very similar to each other
    - Minimize within variance
  - Risks outside territory different from those within
    - Maximize between variance



# Optimal Cluster Evaluation

*Within Variance / Total Variance*



# Selection of Territories for Rating Purposes

---

- Smoothed data
- Clustered data
- Combination of Smoothed and Clustered
- Additional Judgment

# Territory Definition Selections

- Helpful to look at a variety of cluster sets to provide guidance when making judgmental changes

Cluster To Review	14			15			16		
15	Exposure Weighted PP	Exposure	Zip Count	Exposure Weighted PP	Exposure	Zip Count	Exposure Weighted PP	Exposure	Zip Count
Proposed Terr:									
1	385	16396	4	385	16396	4	400	7262	2
2	353	4929	3	353	4929	3	373	9134	2
3	317	3665	3	317	3665	3	353	4929	3
4	297	9170	9	297	9170	9	317	3665	3
5	266	10391	9	278	4670	4	297	9170	9
6	229	44776	42	255	5721	5	278	4670	4
7	197	71087	49	229	44776	42	255	5721	5
8	181	63994	62	197	71087	49	229	44776	42
9	165	120410	133	181	63994	62	197	71087	49
10	150	82311	118	165	120410	133	181	63994	62
11	139	61094	58	150	82311	118	165	120410	133
12	130	54651	47	139	61094	58	150	82311	118
13	117	69135	33	130	54651	47	139	61094	58
14	103	4261	3	117	69135	33	130	54651	47
15		0		103	4261	3	117	69135	33
16		0			0		103	4261	3

# Territory Definition Selections

- Judgmental selections need to be made with consideration of several variables, for example:
  - Size of resulting territories
  - Past events distorting results
  - Competitive considerations

15	15			15adj			
	Clu15			Clu15adj			
FBM Proposed Terr:	Exposure Weighted PP	Exposure	Zip Count	Exposure Weighted PP	Exposure	Zip Count	Terr Exp/ Tot Exp
1	385	16396	4	369	24,990	10	4.06%
2	353	4929	3				
3	317	3665	3				
4	297	9170	9	280	19,561	18	3.17%
5	278	4670	4				
6	255	5721	5				
7	229	44776	42	229	44,776	42	7.27%
8	197	71087	49	197	71,087	49	11.53%
9	181	63994	62	181	63,994	62	10.38%
10	165	120410	133	165	120,410	133	19.54%
11	150	82311	118	150	82,311	118	13.36%
12	139	61094	58	139	61,094	58	9.91%
13	130	54651	47	130	54,651	47	8.87%
14	117	69135	33	116	73,396	36	11.91%
15	103	4261	3				

# Comparison of Predictive Value

- By comparing the within variance statistics for the current definition set with the indicated and proposed (after any manual adjustments), a measure of the potential benefit or “lift” is helpful in understanding the benefit to rate equity

Definitions	# of Territories	Liability Coverages	# of Territories	Physical Damage
Current Set	28	23.1%	28	56.1%
Indicated Set	16	0.4%	9	1.8%
Proposed Set	15	0.5%	7	14.5%
Indicated “Lift”		98.3%		96.8%
Proposed “Lift”		97.8%		74.2%

# Why Re-Discover Territories

---

- Better match of rate with exposure
- Action to avoid adverse selection
- Greater availability of external data
- More companies are developing territories based upon their experience rather than using industry or competitor territories
- Desire for greater segmentation
- Tools now readily available to easily analyze data and develop indicated definitions based on your historical experience

## Thank You for Your Attention

Visit us at [www.pinnacleactuaries.com](http://www.pinnacleactuaries.com)

Sandra Ross, FCAS, MAAA, CIC

734-927-5103

[sross@pinnacleactuaries.com](mailto:sross@pinnacleactuaries.com)