

Which bank should you buy?

A prospective of operational similarity



Data Analysis Project

BUDT 733

Eric Grant

Yasuo Mutsuura

Xiaomin Shi

Omarr Tobias

Xiaoyu Zhang

Agenda

A decorative graphic at the top of the slide consists of two groups of circles. The first group on the left has a solid light purple circle behind the word 'Agenda' and an empty light purple circle outline to its right. The second group on the right has three light purple circles: a solid one on the left, an empty outline in the middle, and another solid one on the right.

- Situation
- Data
- Exploration
- Analysis
- Recommendations/Conclusions



Situation

- Yatograxx Consulting intends to help banks that intend to grow through acquisition determine which banks to purchase.
- By explaining operational characteristics of different banks, we can assess functional similarity.

Data



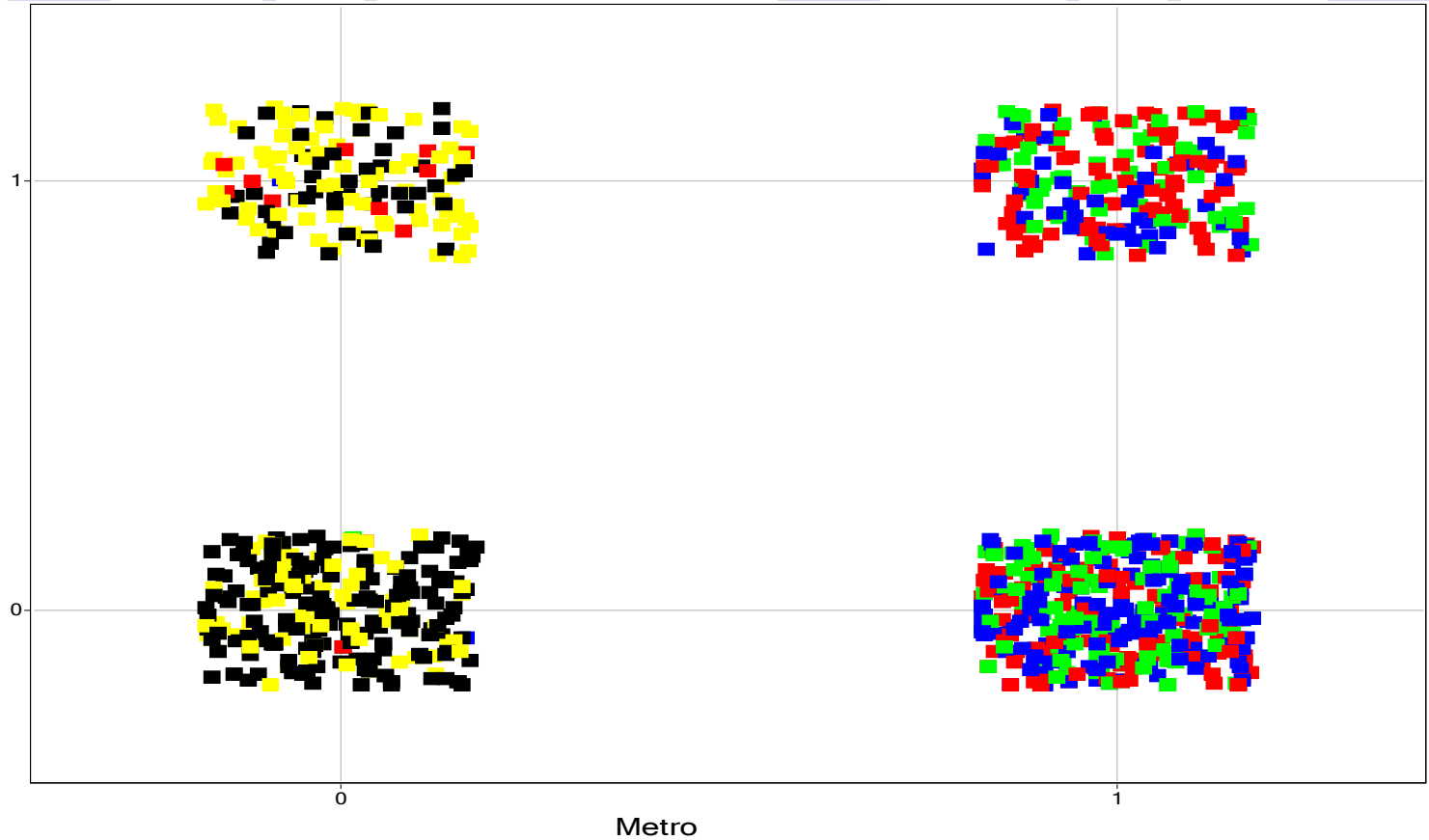
- Over 1000 branches of the top five banks (based on the number of banks) in the DC/MD/VA area
- Original data contained over 100 variables describing every aspect of the locations for several different bank branches
- We combined like variables based on domain knowledge and common sense For example:
 - Combined major downtown and community downtown (1=Downtown)
 - Combined strip development with regional, community, and neighborhood commercial to make one commercial variable (2=Commercial)
- Created several dummy variables

Exploration

Scatter Plot

0=Washington DC, 1=Baltimore

MSA



0=Suburban, 1=Urban



= A



= B



= C



= D



= E

Analysis



To perform the analysis we:

- Ran Discriminant Analysis and Classification Tree with all predictors for entire group
- Compared coefficients and error rates across different banks
- Split the groups and repeated process until we had a reasonable, parsimonious model
- Compared what the model tells us with what we expected from exploration

Analysis-Discriminant Analysis

Before Split

Error Report			
Class	# Cases	# Errors	% Error
A	266	135	50.75
B	219	65	29.68
C	157	46	29.30
D	227	58	25.55
E	193	101	52.33
Overall	1062	405	38.14

After Split
Non-Metro

Error Report			
Class	# Cases	# Errors	% Error
A	13	1	7.69
B	2	2	100.00
C	157	49	31.21
D	227	47	20.70
E	1	0	0.00
Overall	400	99	24.75

After Split
Metro

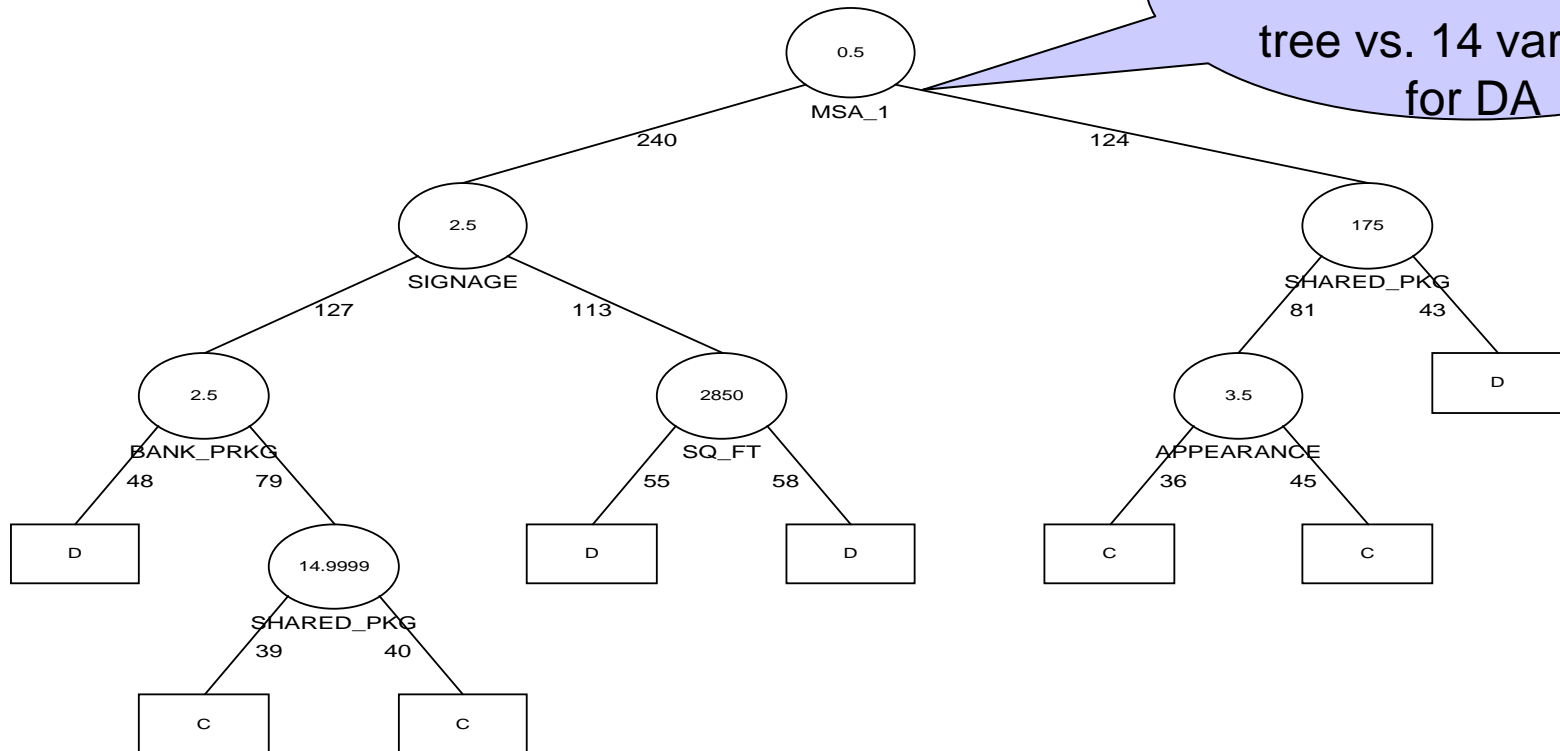
Error Report			
Class	# Cases	# Errors	% Error
A	253	96	37.94
B	217	66	30.41
E	192	95	49.48
Overall	662	257	38.82

Analysis-Classification Tree for Non Metro

Error Report			
Class	# Cases	# Errors	% Error
C	148	43	29.05
D	216	55	25.46
Overall	364	98	26.92

Similar Error Rate for Class Tree

Parsimony—only 4 variables for class tree vs. 14 variables for DA





Conclusions: Profiles

Urban areas

- **A Bank:** Less free standing, less store front, less inside mall with more ATMs
- **B Bank:** Less ATMs, smaller size with less signage
- **E Bank:** More office building, more free standing, more store front, more inside mall, less drive-up tellers with more signage

Suburban areas

- **D Bank:** in Baltimore with more shared parking, or in DC with more signs or less bank parking
- **C Bank:** in Baltimore with less shared parking, or in DC with less signs and more bank parking

Questions

