



# Home Ownership

---

Mamak Ashtari  
Alexander Basilia  
Chien-Ting Chen  
Ashish Markanday  
Santosh



# Objective

- Objective: What variables distinguish homeowners from non - homeowners
- Research questions:
  - What variables characterize homeowners and non-owners
  - Is education important?
  - What about geography?



# Initial Hypothesis

- We expected to see:
  - More married people to be homeowners than single people.
  - positive relationship between level of income and homeownership
  - the number of homeowners to be more among older people.



# Data Source

- Consumer Expenditure Survey
- Last three months of 2000
- With the exception of Age all data is categorical
- After conversion into dummy variables 56 variables
- 5106 records



# Cleaning Up the Data Set

- We reduced the size of independent variables to eight:
  - region, urban, age, having or not children, marital status, income and race.
  - made smaller bins of data.
  - added price to the data set
- 5099 records remain



# Structure of Data

Count of ID	Region				
House Ownership	Midwest	Northeast	South	West	Grand Total
No	27.51%	24.04%	37.07%	11.38%	100.00%
Yes	29.94%	22.01%	38.68%	9.37%	100.00%
Grand Total	28.87%	22.91%	37.97%	10.26%	100.00%

Count of ID	Area Type		
House Ownership	Rural	Urban	Grand Total
No	1.87%	98.13%	100.00%
Yes	1.90%	98.10%	100.00%
Grand Total	1.88%	98.12%	100.00%



# Structure of Data

House Ownership	No Income	Below 20k	Between 20 and 50	Above 50k	Grand Total
No	13.16%	37.96%	34.09%	14.80%	100.00%
Yes	8.42%	17.62%	44.44%	29.52%	100.00%
Grand Total	10.51%	26.59%	39.87%	23.02%	100.00%

Count of ID	Race				Grand Total
House Ownership	Asian	Black	Spanish	White	Grand Total
No	4.49%	18.40%	0.09%	77.02%	100.00%
Yes	3.65%	11.79%	0.07%	84.49%	100.00%
Grand Total	4.02%	14.71%	0.08%	81.19%	100.00%



# Conclusions

- Social-economic variables explain about 65% of homeownership or non-ownership.
- Income, marital status, children, race, region.
- Education does not contribute to home ownership based on decision tree.
- People in the South and Midwest are more likely to own homes than people in west and northeast.
- Age is not important.



# Analysis using Logistic Regression Model

- Key characteristics to House Owners
  - **Income Level:** The income levels are critical factors to decide whether residents own house or not. The higher the income, the more likely residents own house
  - **Children:** Having children is a big motivation that residents consider to own houses
  - **Race:** Compared with Asian, white people are more likely to own their house, Spanish are less likely to own their house
- Other Factors Common to House Owners
  - **Marriage status:** Married couple tends to own their house
  - **Education:** Higher education has a positive influence on owning a house.
  - **Geography Area:** People living in the Midwest and the South are more likely to own their house than those in the west.
- Irrelevant Factors to House Ownership
  - **Urban:** There is no indication/tendency about whether people living in Urban area or in Rural area to own a house or not. From managerial point of view, estate agent should ignore this factor while considering targeting potential clients
  - **Age:** The age of residents has little influence on the house ownership



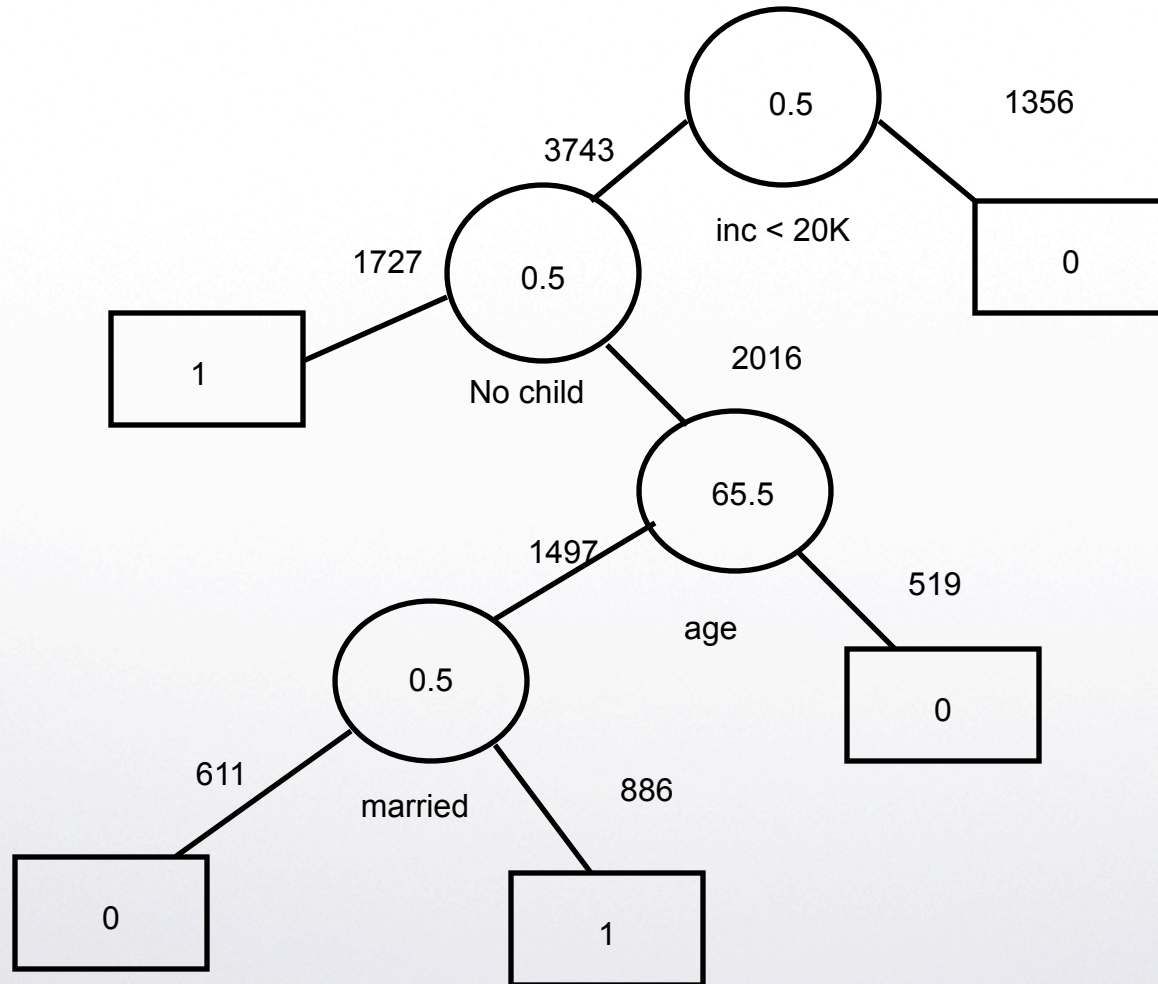
# Linear Discriminant Analysis

Variables	Classification Function		Differences
	1	0	
Constant	-58.827007	-58.28051	
northeast	9.3300238	9.2325001	0.0975237
midwest	6.3864851	6.1066895	0.2797957
south	7.6124458	7.2924681	0.3199778
Urban	56.865795	56.950943	-0.0851478
age_ref	0.1433503	0.1454592	-0.002109
no_Child	3.8544216	4.396318	-0.5418963
married	7.8073816	7.4710493	0.3363323
inc_120k	8.7256985	8.9146719	-0.1889734
inc_20t50k	8.9554653	8.2564783	0.698987
inc_50kp	8.4978781	7.4585118	1.0393662
white	22.041281	21.702395	0.3416862
black	20.587074	20.754723	-0.1676483
spanish	-1.3209963	-1.0675249	-0.2534714
High_Scl	10.421017	10.574903	-0.1538858
College	11.143661	11.188146	-0.0444851

- LDA once more empathizes the importance of characteristics
- It also states that people without children, or with income less than 20k are less likely to own homes
- The use of dummy variables violates LDA's assumption, therefore it is not as useful as logistic regression



# Classification Tree





# Cluster Analysis of Home Owners

- This was an exploratory exercise to better understand the structure of home ownership
- Based on within-cluster distances we conclude that there are four natural clusters in the data
  - Cluster 1: Young, married, educated couples with children living in the south
  - Cluster 2: Older, low-income, single individuals living mostly in the mid-west
  - Cluster 3: High income, college educated married whites without children
  - Cluster 4: Medium income, high school educated white and hispanic families with children living mainly in the North-East



# Clusters and Dominant Characteristics

<b>CLUSTERS</b>	<b>REGION</b>	<b>MARITAL STATUS</b>	<b>RACE</b>	<b>INCOME</b>	<b>EDUCATION</b>
<b>YOUNG PARENTS</b>	South	Married with Children	Mix, but mostly white	Medium	College Educated
<b>RETIRED</b>	Mid-West	Single	Mix, but mostly white	Low Income	Mostly high school
<b>WEALTHY COUPLES</b>	Mostly North-East	Married without children	White	Wealthy	College Educated
<b>LOWER STRATA</b>	North-East	Married with children	Hispanic and Whites	Low Income	High School



## Going Forward

- We believe that the data does not fully capture local pricing variations
  - Supplied regional pricing data with limited explanatory power in our analyses
- It would also be interesting to investigate if there are any cultural aspects impinging on home ownership
  - Since home owners seem to be concentrated mostly in the mid-west and south



# Q & A