PREDICTING CUSTOMER CHURN

WHICH CUSTOMER
IS LIKELY TO LEAVE?

Cecily Rekart

Craig Lowenstein

Crystal Trainor

Ernest Chrappah

Maieka Hansard

AGENDA

1. Background
2. Why Predicting Customer Churn Matters
3. Data Modeling & Analysis
4. Recommendations
5. Questions and Answers

PREDICTING CUSTOMER CHURN

- **Background**
  - Competition is intense: 0% balance transfers

Aggressive Acquirers	Relationship Managers
Nextcard	American Express
GetSmart	Discover Card

  - High rates of customer defection: 20%-30%
  - Highly profitable
  - Cost $80 to acquire a customer that will generate $120 a year if he/she keeps the card

- **Firm**
  - Major credit card company with travel offices in UK, France, Britain, and Germany

- **Problem**
  - Which Current Customer Is Likely To Leave?

WHY PREDICTING CHURN MATTERS

- **0% offers cost the industry £1bn a year in lost revenue in UK.**
- **20% annual churn rate means half of your customer base will be gone in 4 years!**
- **Costs 7 times to acquire a new customer than to keep an existing customer**
- **Devise strategies for “rate tarts”**
- **Win-back high value customers**

DATA MODELING & ANALYSIS

Data Source: Professor P. Kannan, Harvey Sanders Associate Professor of Marketing. Director, Center for Excellence in Service, UMD.

Explanatory Analysis: Examined data by sorting, pivot tables, scatter plots, correlations, and histograms

Variable Reduction: Started out with 405 observations and 182 predictor variables. Utilized 11 predictors. Narrowed predictors through domain knowledge, elimination of null values, and creation of interaction variables.

Other activities: Logged variables to normalize data, recoded the data into format usable by analysis tool, looked at stepwise through logistic regression to find most relevant variables

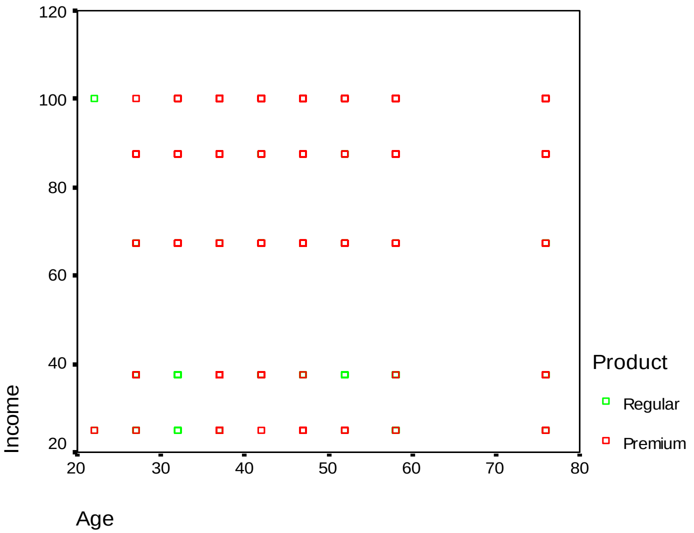
Model/Statistical technique Used: Discriminant Analysis because -

- Insufficient data to run classification trees
- Asymmetrical misclassification cost (7:1)
- Prior probabilities
- 30% more efficient than logistic regression given multivariate normal data

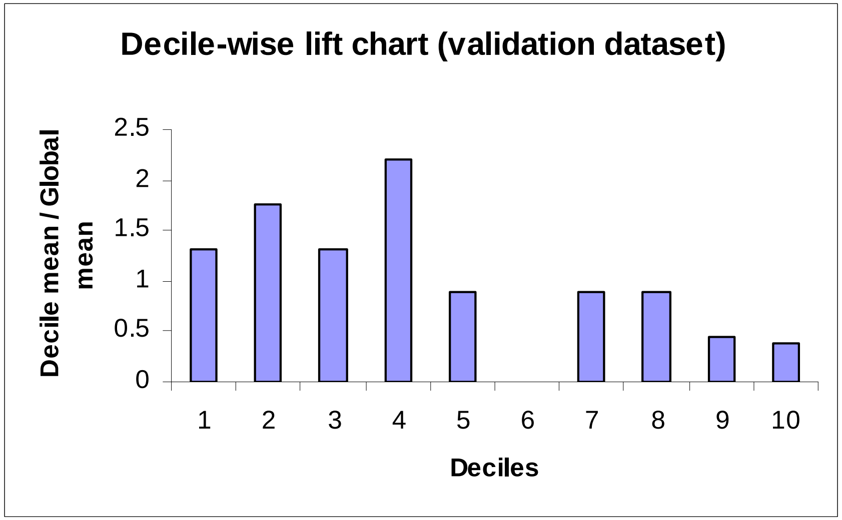
VARIABLES IN OUR MODEL

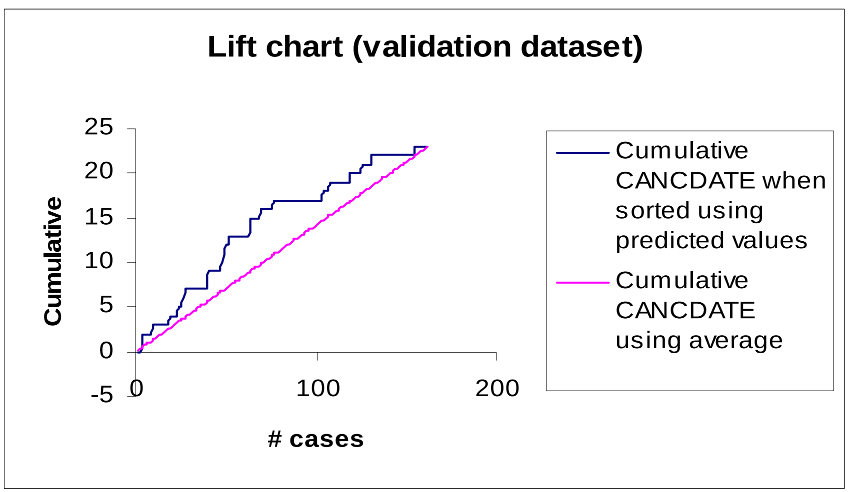
- Income
- Age
- Product
- Amount Spent
- # Transactions
- Marital Status
- Statement Ease
- Marketing Quality
- Anniversary Date
- Loyalty
- Gender

Variables	Classification Function	
	1	0
Constant	-146.1769714	-145.9470673
Log(Age)	56.04314804	55.95493698
MARITAL STATUS	-6.37859583	-6.39721918
Log(INCOME)	21.88240623	21.65524483
PRODUCT	9.39768505	9.12808704
Log (ANNDATE)	-7.50112486	-7.17833376
TOTAL_SPEND24	-0.03029263	-0.02248576



DISCRIMINANT ANALYSIS RESULTS

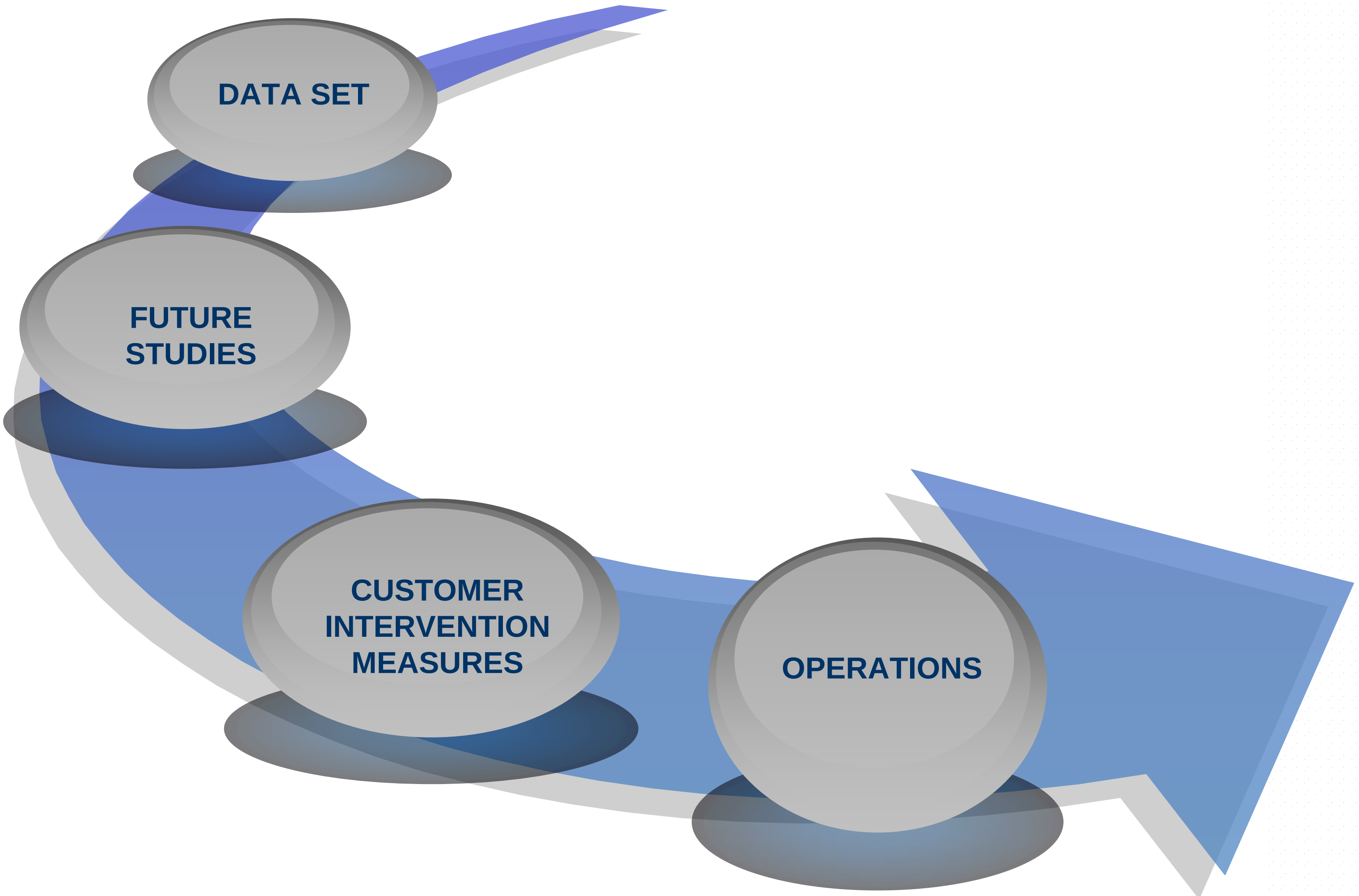




	Predicted Class	
Actual Class	1	0
1	22	1
0	127	12

Error Report			
Class	# Cases	# Errors	% Error
1	23	1	4.35
0	139	127	91.37
Overall	162	128	79.01

RECOMMENDATIONS



WHICH CUSTOMER IS LIKELY TO LEAVE?

- **Age: 40's**
- **Shorter Tenure**
- **Regular Product**
- **Lower Income**
- **Spend Less**
- **Single**

	Stay	Leave
Premium	86.14%	13.86%
Standard	80.30%	19.70%

Variables	Classification Function	
	1	0
Constant	-146.1769714	-145.9470673
Log(Age)	56.04314804	55.95493698
MARITAL STATUS	-6.37859583	-6.39721918
Log(INCOME)	21.88240623	21.65524483
PRODUCT	9.39768505	9.12808704
Log (ANNDATE)	-7.50112486	-7.17833376
TOTAL_SPEND24	-0.03029263	-0.02248576

RECOMMENDATIONS

- **Data Set –**
  - **Collect more ratio data**
  - **Track carrying balance, payoff trends**
- **Conduct future studies – *Why* are customers leaving?**
- **Operationalize the Model**
- **Intervention Measures**
  - **Incentives**

PREDICTING CUSTOMER CHURN

Thank You!