Actionable Analytics

DMA of Northern California

May 20, 2009

Jim Porzak

THE
generations
network

What We'll Cover

☞ Introduction

☞ Test to Tune

☞ Segment for Understanding

☞ Predict for profit

THE
generations
network

My Background & Biases

☞ DM projects over last few years

- 24 HourFitness
- LA Times, Chicago Sun Times, Chicago Tribune
- Fleetwood Motor Homes
- Master Foods (My M&M's)
- Philosophy
- And some to special to mention...

☞ Ancestry.com

Test to Tune - Outline

- Why do we test?

- Control Group – Why? How Big?

- Simple A/B testing

- Testing more with A/BCD... testing

- Multi-variable Testing?

- Testing at Home

Why Do We Test?

☞ Direct Marketing is an experimental art form

- No real learnings without experiments!
- Harrah's Gary Loveman:
  "Three ways to get fired at Harrah's:
  1. Steal,
  2. Harass Women, or
  3. Institute a program without first running an experiment
- Sara Lee's Barry Beracha:
  "In God we trust, all others bring data."

Control Groups

☞ No control, no test. No test, no learnings!

☞ OK, so we'll have a control. How big should it be?

- CFO: zero

- Geek: 50%

- DM'er: Just big enough to show significance.

Gary Loveman's 3rd reason sometimes quoted as "Failing to use a control group."

How Big is a "Just Big Enough" Control?

- Types of Significance
  - Statistics
  - Business
- Answer questions in this order:
  - Total N?
    - Or, N per week?
  - What is current, or expected, baseline response rate?
  - What is minimum lift to have meaningful business impact?
  - What are time and/or budget constraints?
- Cop-out: Use rule of 100:
  - Ensure at least 100 responses in each group.

THE
generations
network

Control Group Size – Example 1

☞ Current marketing piece (A)

- Sent to 100,000 prospects
- $0.50 cost, fully loaded
- Converts at 1%, with $75 average profit

☞ Test piece (B)

- Try a glossier piece
- $0.80 cost (when, & if, eventually rolled out)
- Assuming same average profit:
  - What will conversion rate need to be to equal performance of A?
  - What control size control group needed to verify this performance with 95% confidence?

Control Group Size – Example 2

🖎 Currently our subscribers renew at 85%

- Value of a renewal is $100
- Each month 100k subscribers are up for renewal

🖎 In addition to everything we already do to encourage retention, next month we will test a special informational mailing costing $0.50 per piece.

🖎 Goal in to increase retention rate to 86%.

🖎 On theory, "marketing can do no harm," we have budget to mail to all 100k subscribers up for renewal, but...

🖎 What size hold-out should we have to be able to verify the expected 1 point lift with 95% confidence?

THE
generations
network

Evaluating Results of Test

☞ Basic stats

- Don't cheat!
  - Use 2-sided always
  - 90% confidence *minimum*

☞ Use prop.test function in R

```
> prop.test(Subs, Regs)
2-sample test for equality of proportions with continuity correction
data: Subs out of Regs
X-squared = 7.3232, df = 1, p-value = 0.006807
alternative hypothesis: two.sided
95 percent confidence interval:
0.001459448 0.008540552
sample estimates:
prop 1 prop 2
0.049 0.044
```
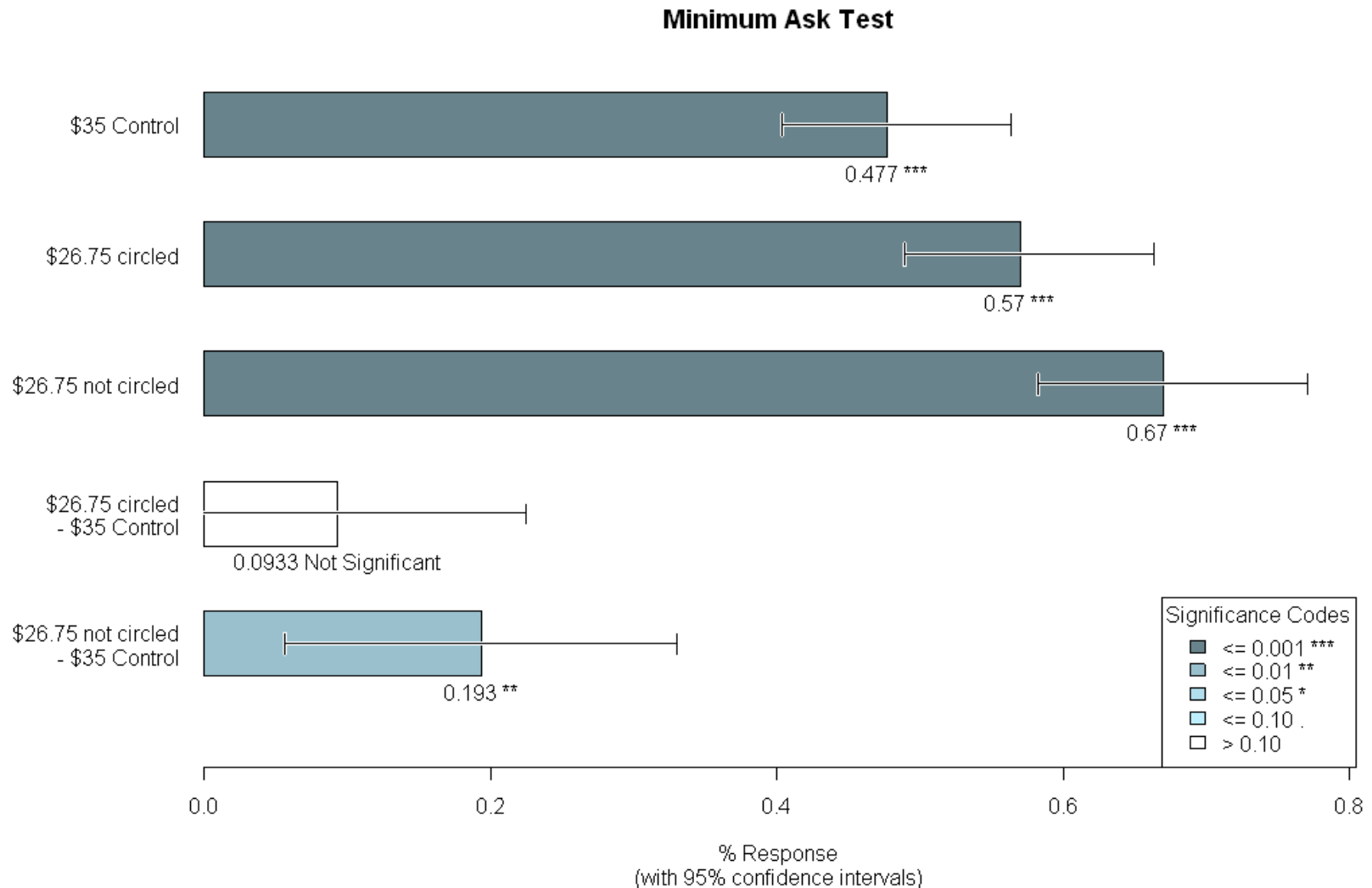
THE
generations
network

A/BC Test Example

Minimum Ask Test

$35 Control — 0.477 ***

$26.75 circled — 0.57 ***

$26.75 not circled — 0.67 ***

$26.75 circled - $35 Control — 0.0933 Not Significant

$26.75 not circled - $35 Control — 0.193 **

Significance Codes
- <= 0.001 ***
- <= 0.01 **
- <= 0.05 *
- <= 0.10 .
- > 0.10

% Response
(with 95% confidence intervals)

Data from Mal Warwick's Newsletter November, 2007

Segment for Understanding

- Making Customers Personable
- Purpose(s) of Segmentation
- Real World Segments
- RFM Segments
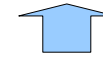- Using Cluster Analytics to Segment
- Clustering Resources

"We've broken your list into eighty-four subgroups. Our work here is done."

Preview – RFM Based Segments
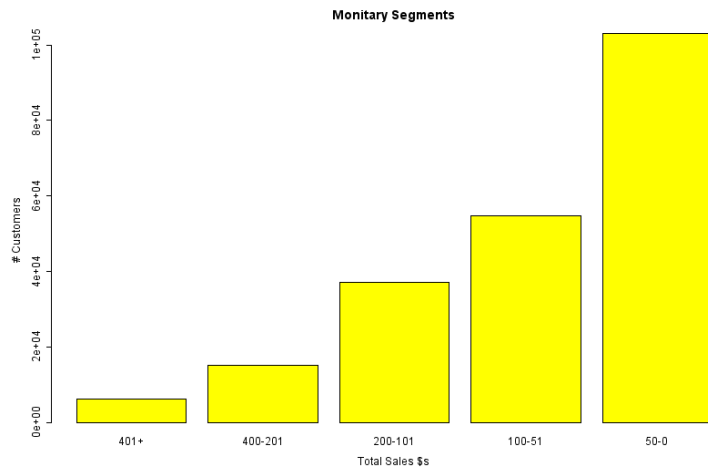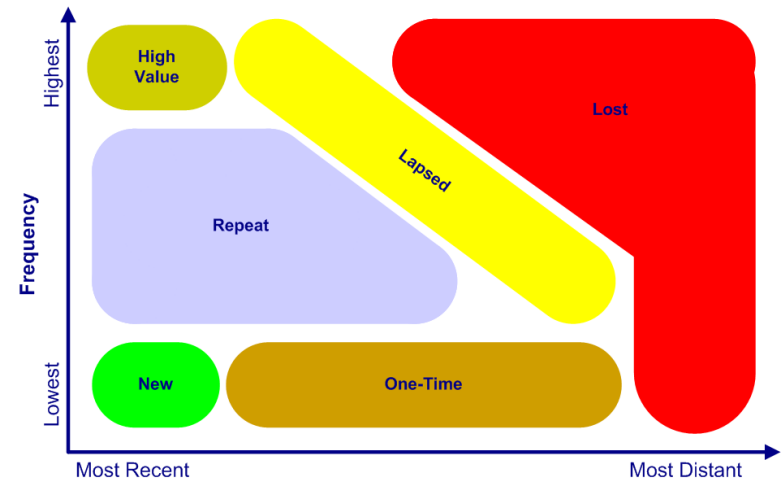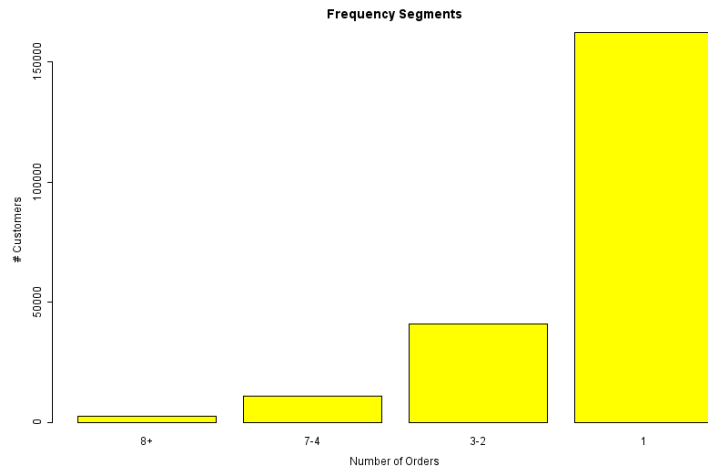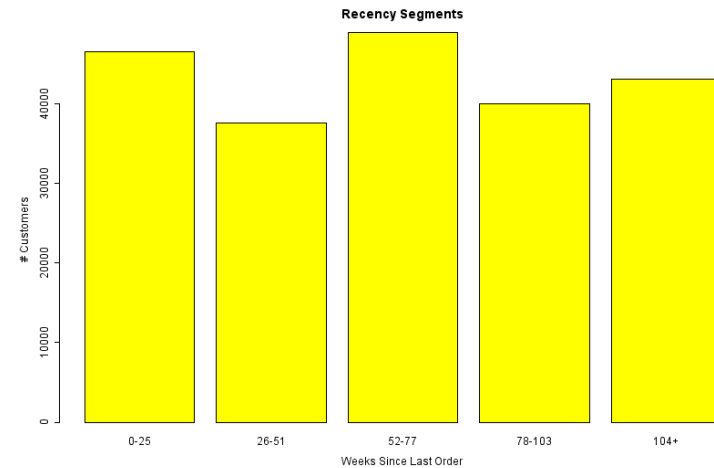
Cluster Based Segment Example

☞ Attitudinal Segmentation

☞ Marketing Challenge

- Our client offers free download of software with high perceived value, but

- First asks user to fill out a simple survey

- Challenge is to come up with a "few" segments that will be used by segment to:
  - Prioritize contact strategy
  - Craft marketing messages based on profile

Data Set

- Surveys from 20k respondents
- All within same time frame (a number of weeks)
- All requested the software download

Survey Description

- 35 check boxes or radio buttons
  - None required. Coded as binary responses
- Arranged in 5 sections
  - License: W and/or X
  - Role: one of D, SA, ITM, ITA, Str, Oth (radio buttons)
  - System: any of S, T, A, B, C, D, O (check boxes)
  - Interest: any of M, O Pl, Pr, Sup, 64, Con, Per, DT, Z, Oth. (check boxes)
  - Application: any of Web, Inf, Col, Db, J2, Top, Dev, Per, Other (check boxes)

Clustering Method

- flexclust package by Fritz Leisch
  - See his 2006 paper (on his personal page):
    A Toolbox for K-Centroids Cluster Analysis
  - This is (mostly) an optional response type survey
    - 1 = "yes" is significant
    - 0 is just absence not really a "no"
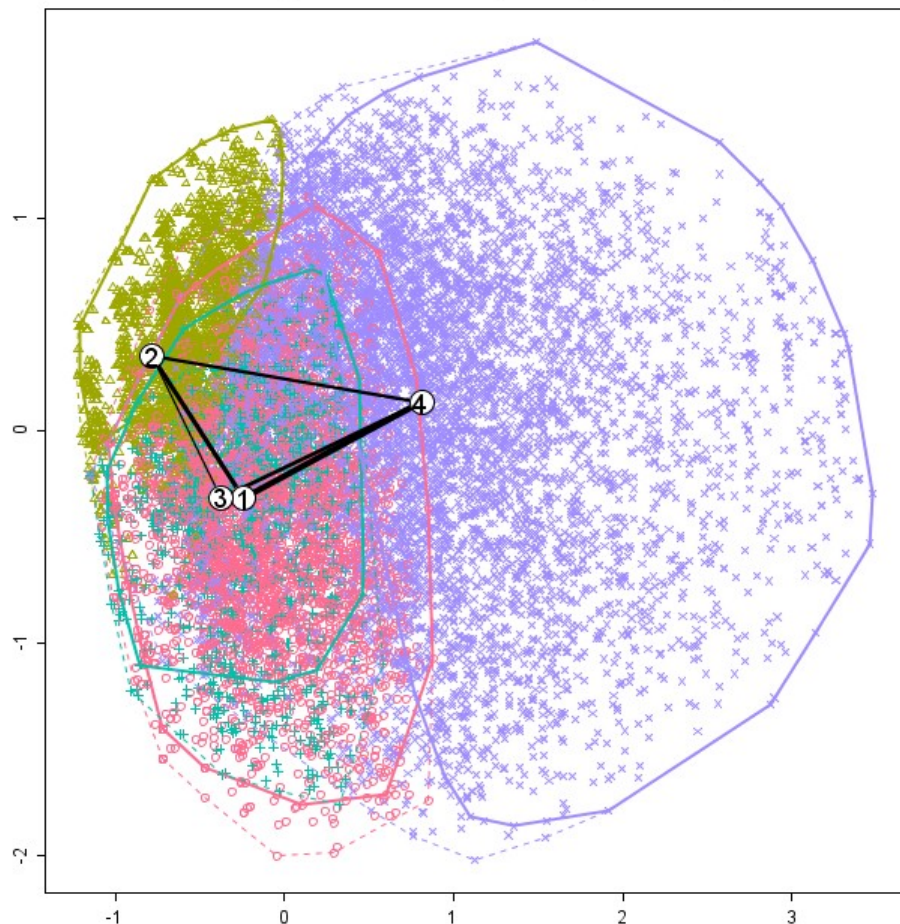 - Respondents checking Role_SA have much more in common than those not checking Role_SA
- Following Fritz's argument we use the expectation based Jaccard distance measure.
- We pick # clusters desired
  - Try 3 to 8 clusters. Repeat 4 times with a random start.

Plots (k=4, seed = 1)
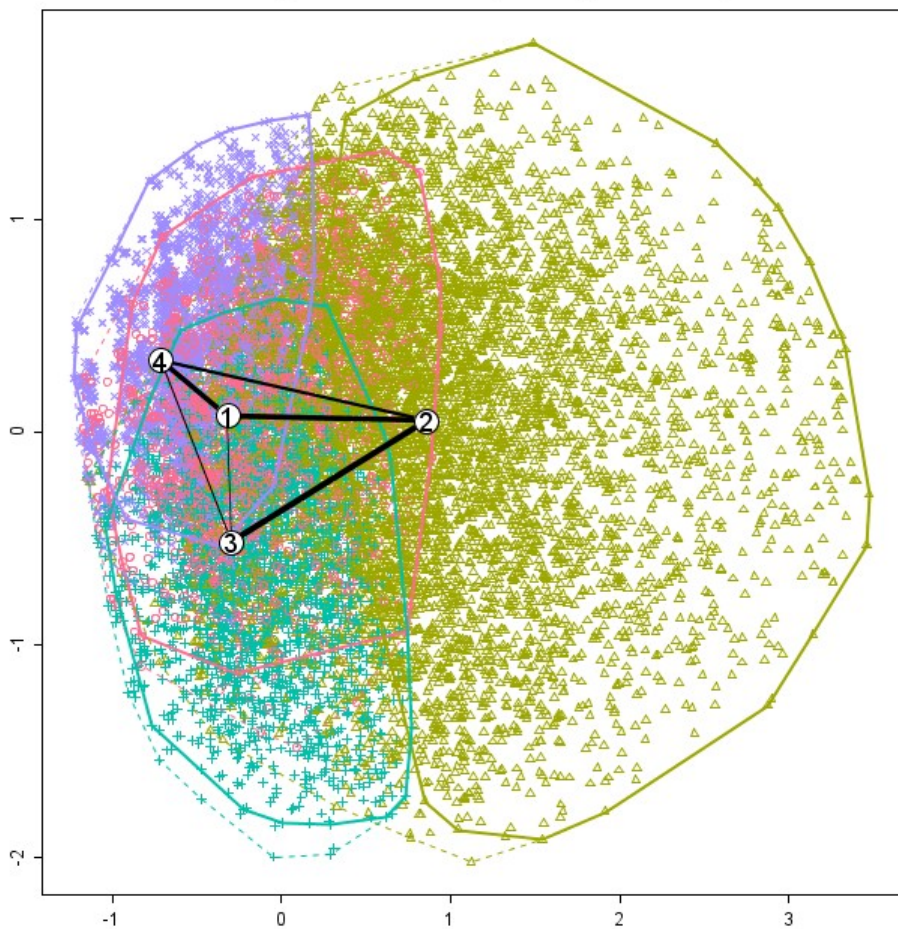

kcca ejaccard - 4 clusters (20k sample, seed = 1)

Av Dist = 0.73465, k = 5
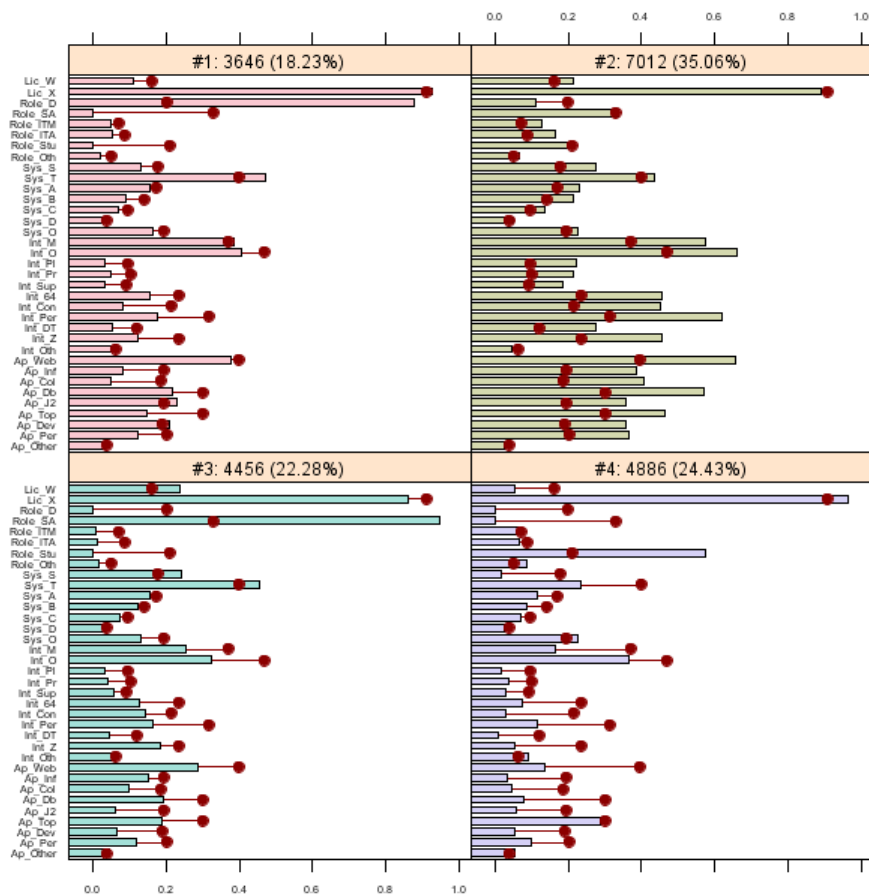
kcca ejaccard - 4 clusters (20k sample, seed = 1)

#1: 5551 (27.75%)
#2: 4577 (22.89%)
#3: 2535 (12.68%)
#4: 7337 (36.69%)
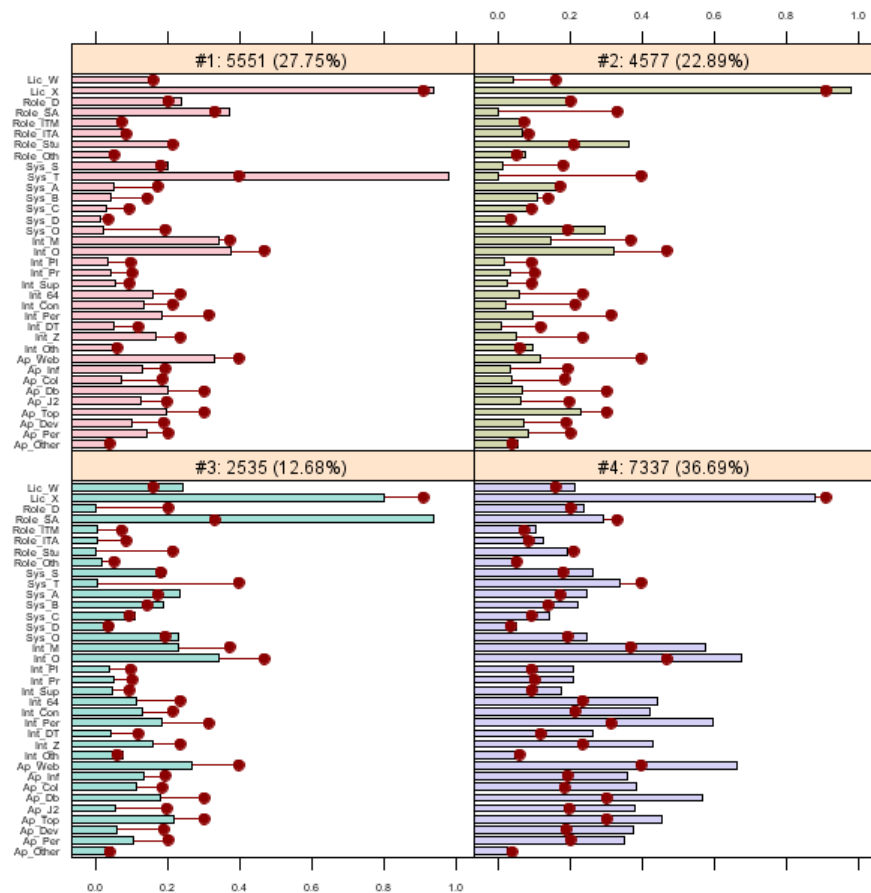
Plots (k=4, seed = 2)

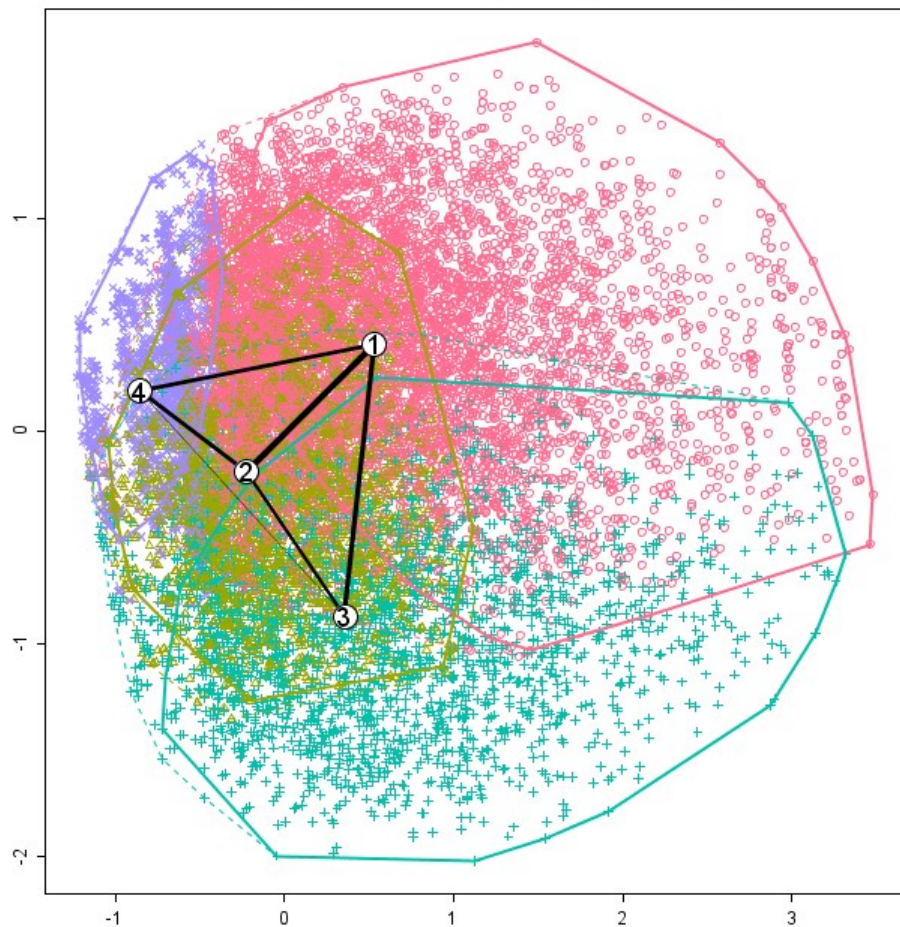

kcca ejaccard - 4 clusters (20k sample, seed = 2)

Av Dist = 0.72916, k = 5
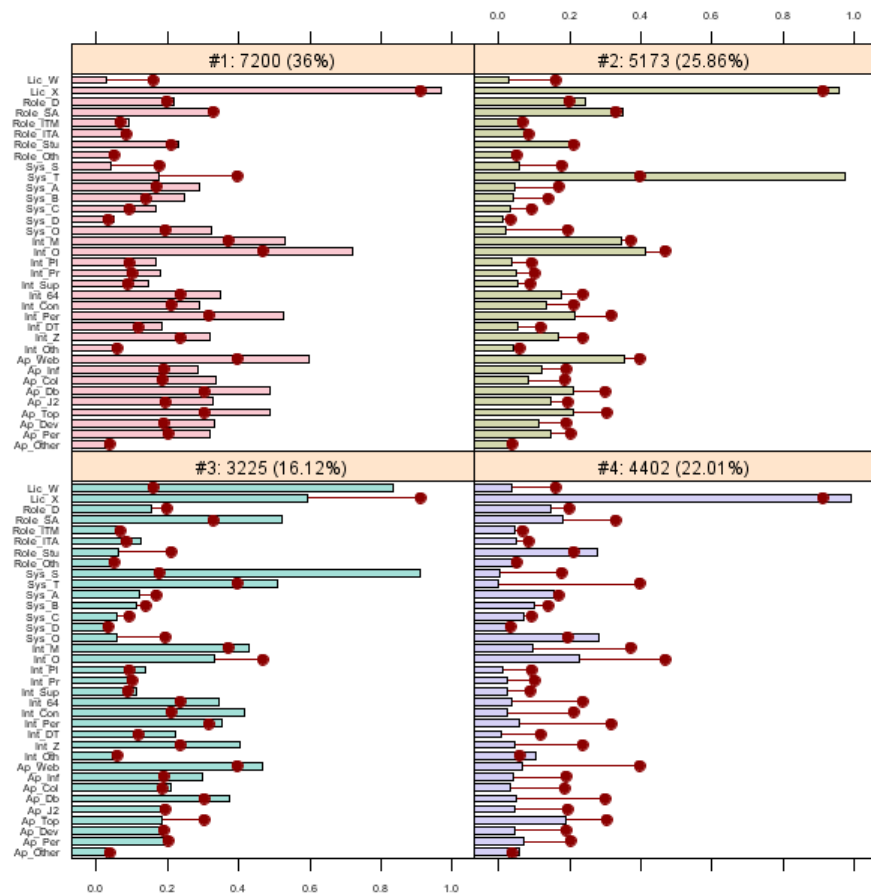
kcca ejaccard - 4 clusters (20k sample, seed = 2)

Plots (k=4, seed = 3)

kcca ejaccard - 4 clusters (20k sample, seed = 3)

Av Dist = 0.73046, k = 5

Are any of these any good?

☞ If so, which?

☞ How to decide?

☞ Quoting Fritz (pg 15): *The actual choice … has been made manually by comparing various solutions and selecting the one which made most sense from the practitioners point of view. This may seem unsatisfying because the decision is subjective, but cluster analysis here is used as a tool for exploratory data analysis and offers simplified views of a complex data set.*

☞ Our Criteria

- Choice of k, must have mostly ~ stable solutions, and
- Cluster profiles must be interpretable. IOW, what is the story you can tell about each cluster? Can the marketers relate to the stories?
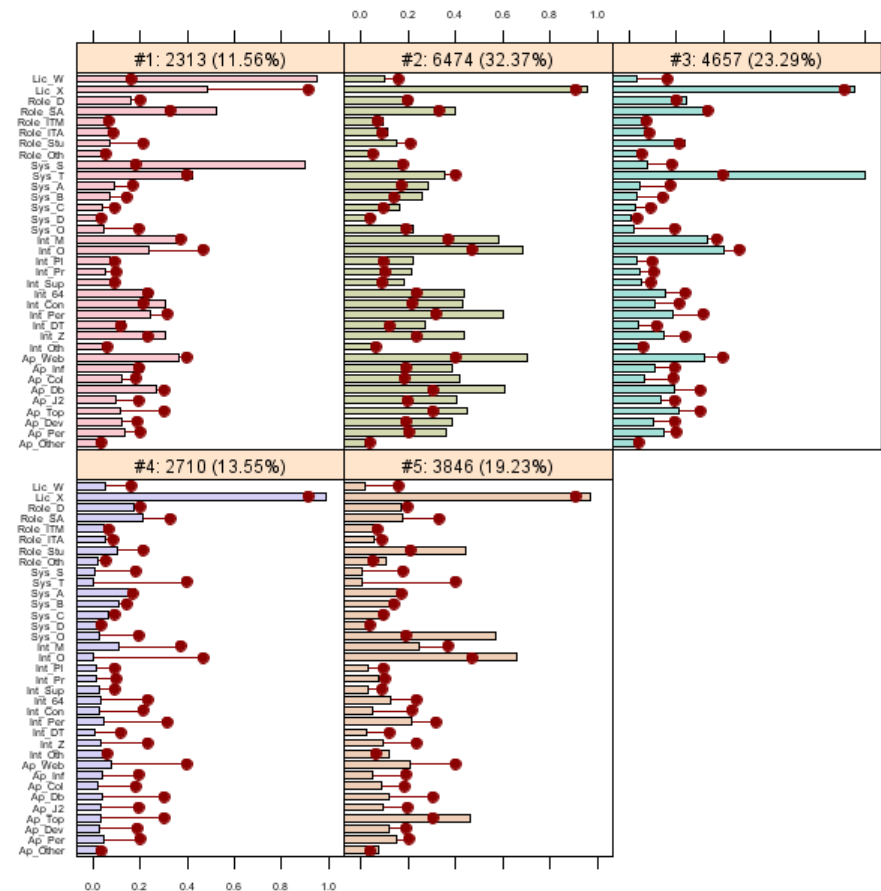
THE
generations
network

kcca ejaccard - 5 clusters (20k sample, seed = 9)
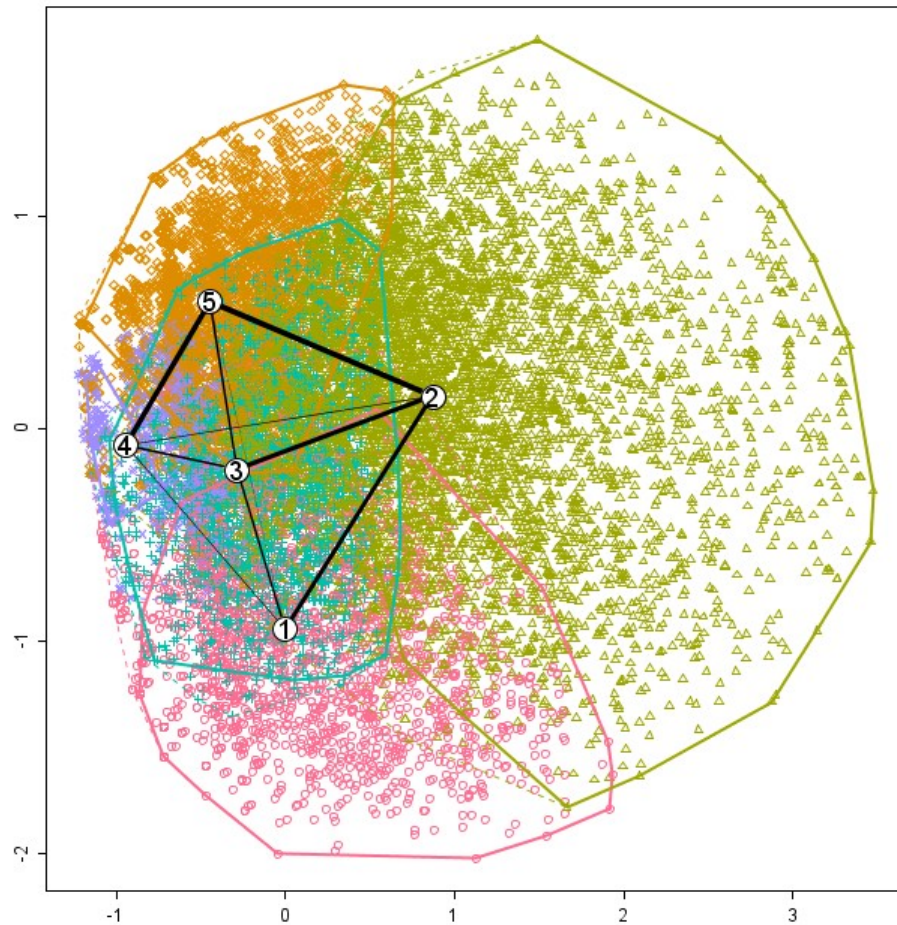
Av Dist = 0.71212, k = 5


kcca ejaccard - 5 clusters (20k sample, seed = 9)

And the stories are...

☞ Based on knowing a bit more about the client than I can share with you.

☞ #1: An "S" loyalist, high % SA's

☞ #2: Favors name brands, high responders

☞ #3: A "T" loyalist, broad but reduced responses

☞ #4: Favors name brands, but otherwise low resp.

☞ #5: Student, gray box, open source, desktop.

THE generations network

Cluster Based Targeting

☞ Once cluster model is built…

- New survey takers can be assigned to cluster
  - Could be real time
  - Marketed to appropriately
- Tied back to other metrics, eg a purchase model:

Probability of Purchase by Persona Cluster
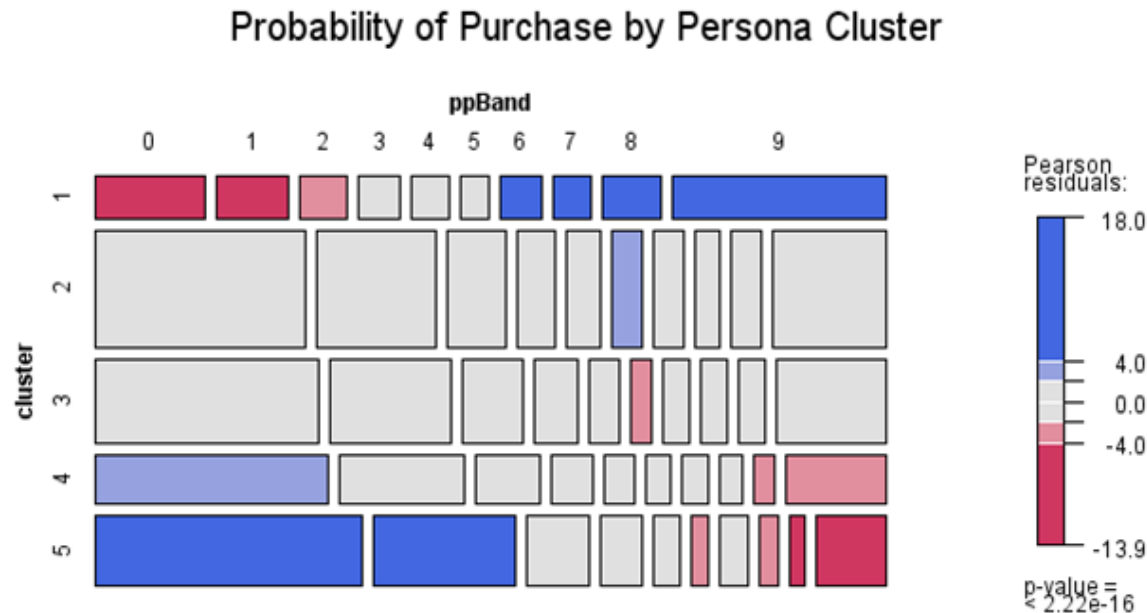
Predict for Profit

☞ Why do we Predict?

☞ What's a Prediction Model?

☞ Classical Response Models

☞ Propensity Models

THE
generations
network

How & Why do we Predict?

Five ways to cut costs using predictive modeling:

☞ Response Modeling

Don't contact those who won't respond.

☞ Response-uplift Modeling

Don't contact those who would have purchased anyway.

☞ Churn Modeling

Don't waste retention dollars on those who will stay anyway.

☞ Churn-uplift Modeling

Don't offend those who would otherwise stay.

☞ Risk Modeling

Don't acquire "loss customers."

After Eric Siegel, Predictive Analytics World 2009

THE
generations
network

An At-Risk Model

☞ Sample Data from a sports club

- Challenge – predict "at-risk" members based on membership usage data & simple demographics

- Training & Test data sets ~ 2k records each

  - MembID (identifier)
  - Status = M or C
    (Member or Cancel)
  - Gender
  - Age
  - MembDays
  - NumUses1st30d
  - NumUsesLast30d
  - TotalUses
  - FirstCkInDay

  - LastCkInDay
  - DaysSinceLastUse
  - TotalPaid
  - MonthlyAmt
  - MilesToClub
  - NumExtras1st30d
  - NumExtrasLast30d
  - TotalExtras
  - DaysSinceLastExtra

THE generations network

randomForest run

```
> Members.rf <- randomForest(Status ~ ., data = Members, importance = TRUE, proximity =
TRUE)                                                    ## 30 sec

> Members.rf
```

```
Call:
 randomForest(x = Members[-1], y = Members$Status, ntree = 500,
             mtry = 3, importance = TRUE, proximity = TRUE, data = Members)

               Type of random forest: classification
                     Number of trees: 500
No. of variables tried at each split: 3

       OOB estimate of  error rate: 21.4%
Confusion matrix:
    C    M class.error
C 546 263   0.3250927
M 147 960   0.1327913
```

☞Rather good results. Only ~20% overall error rate.
- 33% false positive
- 13% false negative

THE
generations
network

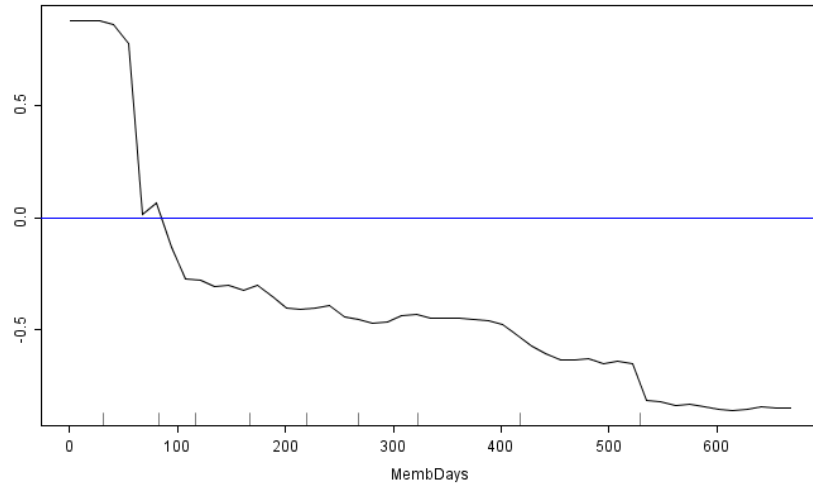# randomForest Variable Importance

Variable Importance Plot

> `varImpPlot(Members.rf)`
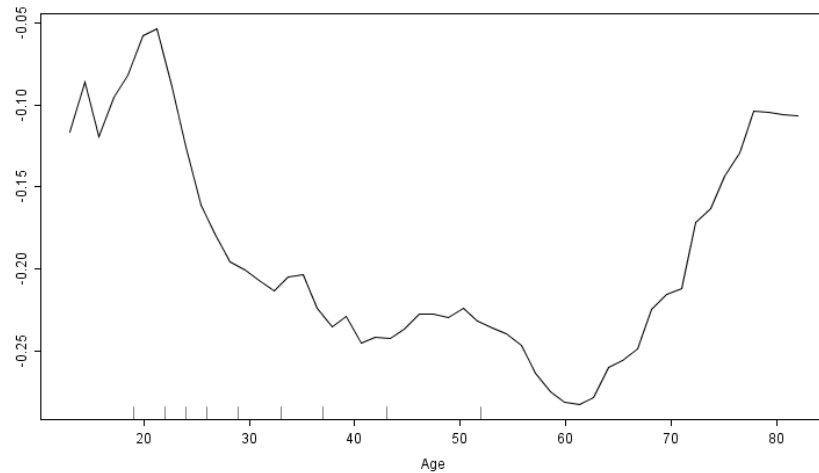

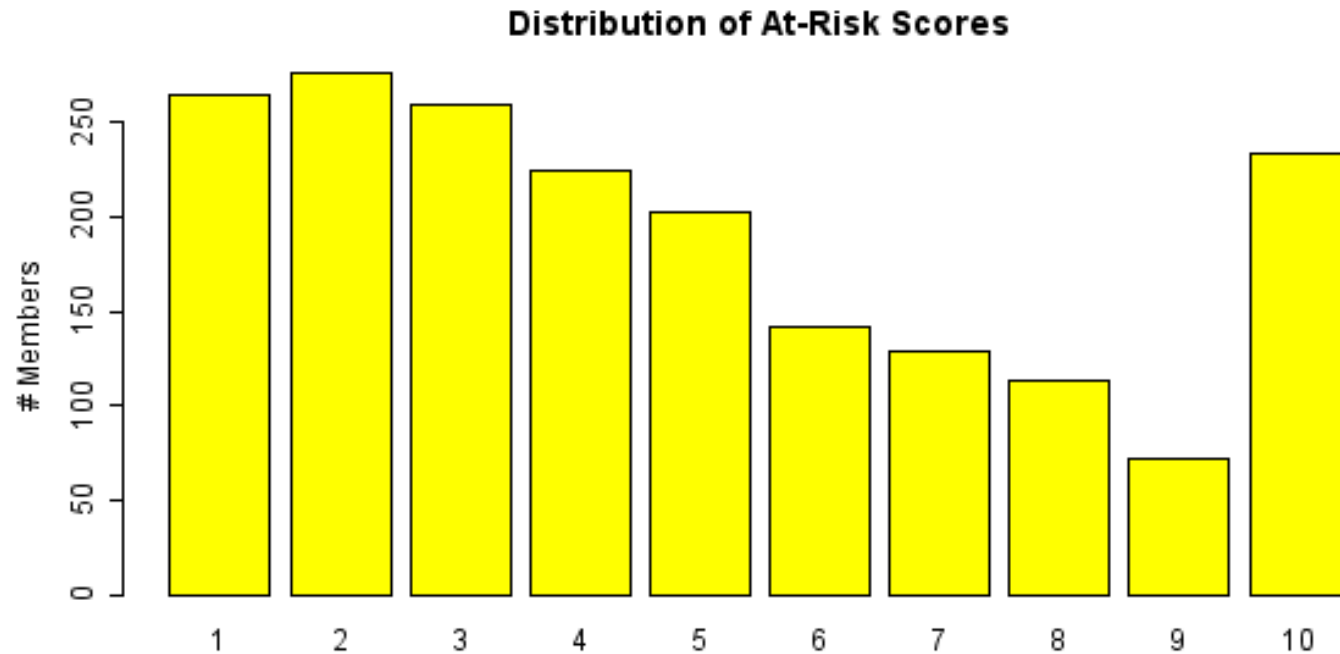Members.rf
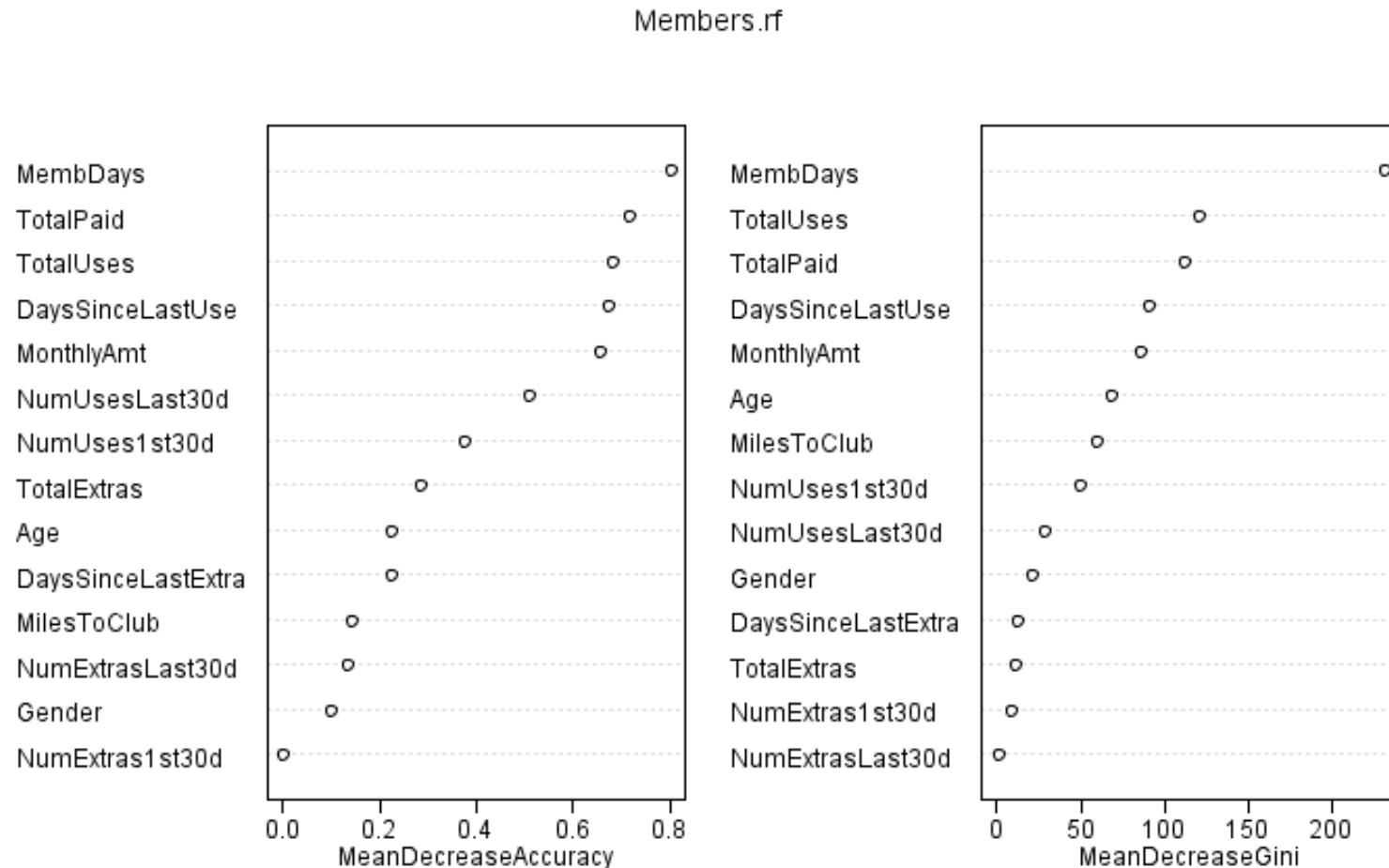
# How Does Outcome Depend on Variables?
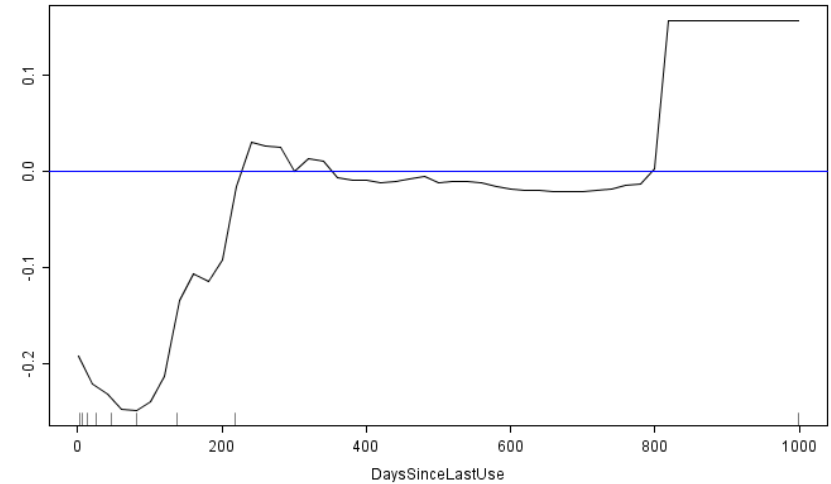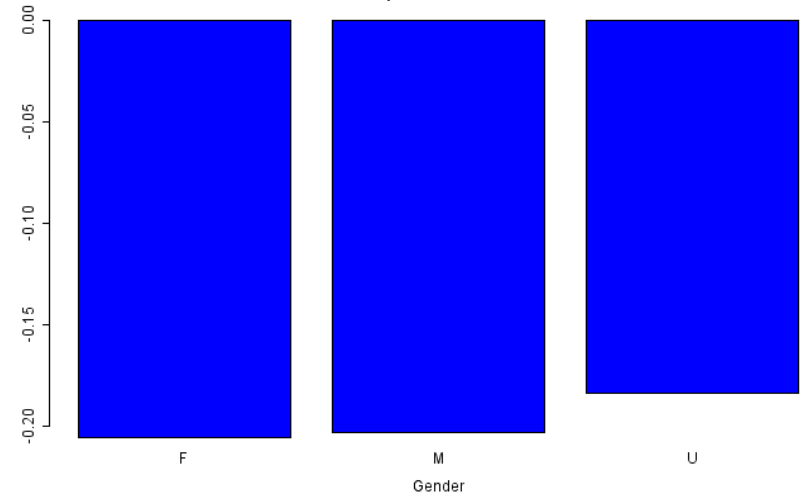
The At-Risk Score

☞ Need a score? Count the trees.

```
AtRiskScore <- floor(9.99999 * Members.rf$votes[, 1]) + 1
barplot(table(AtRiskScore), col = "yellow",
        ylab = "# Members", main = "Distribution of At-Risk Scores")
```

Distribution of At-Risk Scores

☞ Another randomForest run 4k in each group.

- Validated against entire population

Paid Support Model Validation

% Paid	% X86
20	3.40
40	7.01
60	13.11
80	29.83
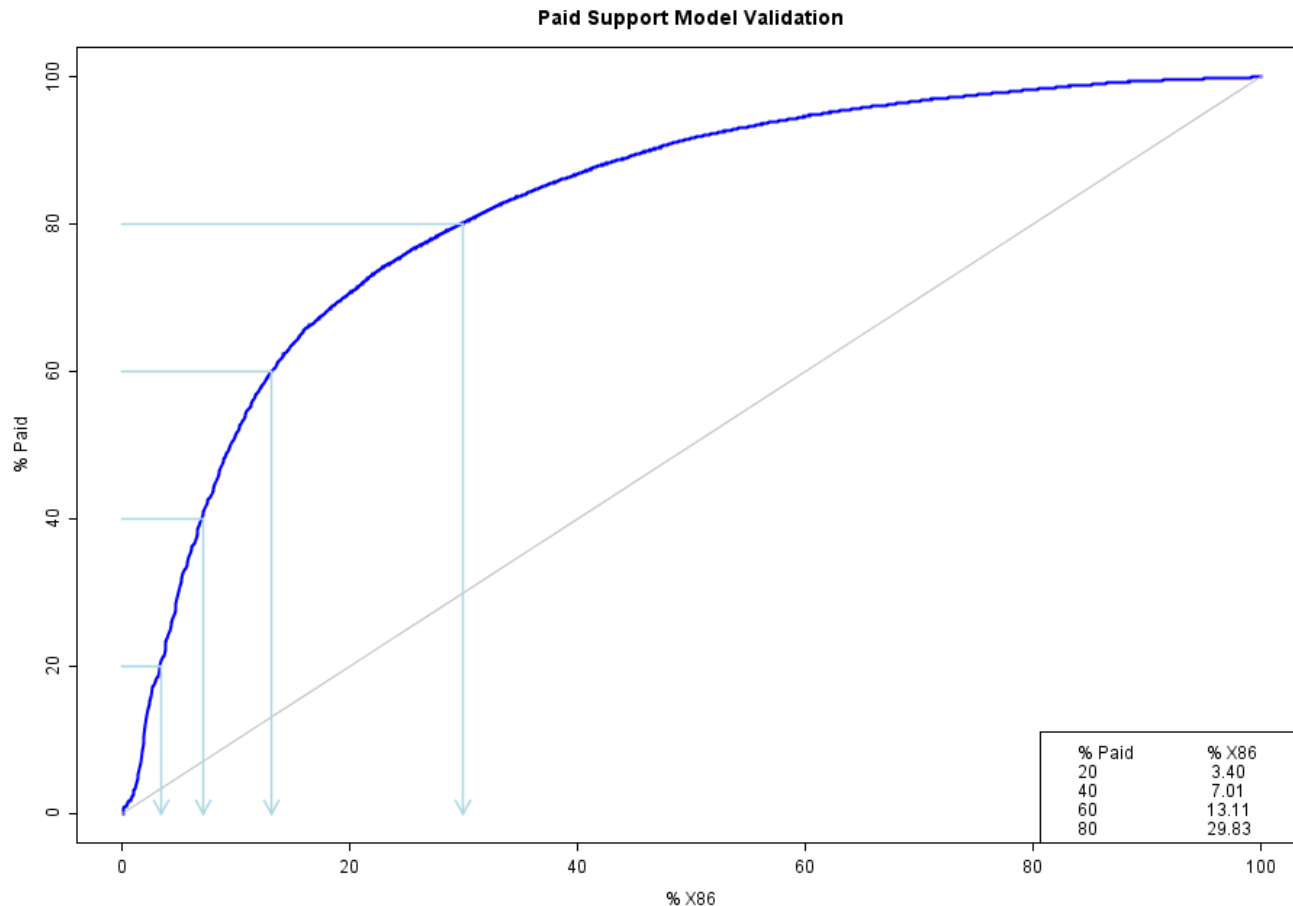
Questions? Comments?

☞ Now would be the time!

THE
generations
network