Modeling, Scoring & Analytics Overview

by Kevin F. Moore, Discovery Communications
Jim Doyle, USPS and American University

Discovery Communications Commerce Dept.

Discovery Communications – Online Store

Discovery Channel Commerce

True Multi Media Integration:

Media	Media Usage
Email	**Throughout the year**
Search	
– SEO	On Going
– Paid Search	Timed with campaigns
TV	**Remaining Times and Fall**
Direct Mail	**Fall Catalog – 3 drops**

The Process of Prospect Targeting

- Start with a Customer Profile

    - Demographics
        - Easy to identify verses costly and difficult data
        - At the register get the Zip Code
        - At the register get the Telephone number (cell phones ?)
        - Online – get the email and home address

        - What data will best allow you to target?
            - Income – Household over $ 125,000.
            - College Educated
            - Presence of Children

The Process of Prospect Targeting

- Purchase Histories - RFM
  - What did you buy ?
  - How recently did you buy it ?
  - How frequently did you purchase ?
  - What was the value of your purchase ?

RFM Scoring

- Select Universe
  - 12M, 24M, 48M, etc.?
  - All buyers, sub group of buyers?
- Download transaction records within that period
  - Address info (or unique customer identifier)
  - Order date
  - Order amount
  - Any other relevant info you might want to use later
- Profile the data to prep for scoring
  - Min & max order date by individual
  - Count of orders
  - Sum of order amount
  - Determine R, F & M breaks
- Run series of queries with if/then statements to start scoring.
  - If max order date between x and y, then z
  - If count of orders between x and y, then z
  - If order amount between x and y, then z
- "Zs" above should be defined as follows
  - Best = 3, Next Best = 2, Worst = 1

RFM Scoring

- Run a simple concatenation query to create a composite score for each individual
  - Recency is always first (R)
  - Followed by Frequency (F)
  - Followed by Monetary Value (M)
  - Examples: Customer A: 323, Customer B: 213, Etc.
- Sort file descending based on score
  - 333
  - 332
  - 331
  - 323
  - 322
  - ………
  - 111
- You now have a file that you can select against

RFM Scoring Sample Summary Report
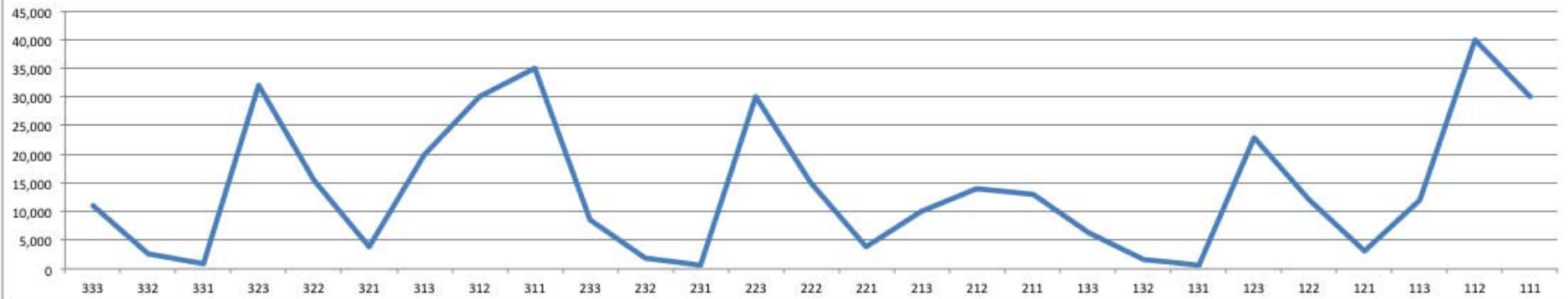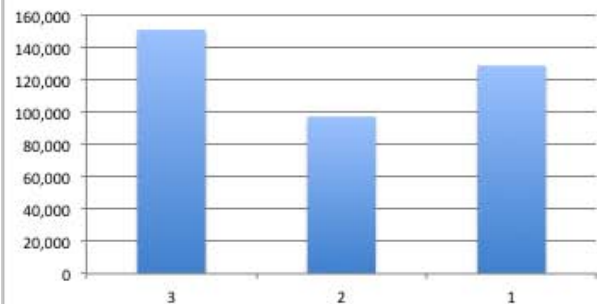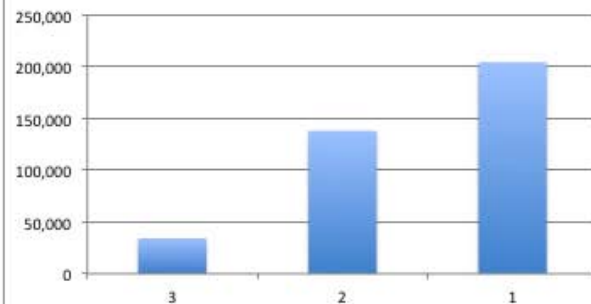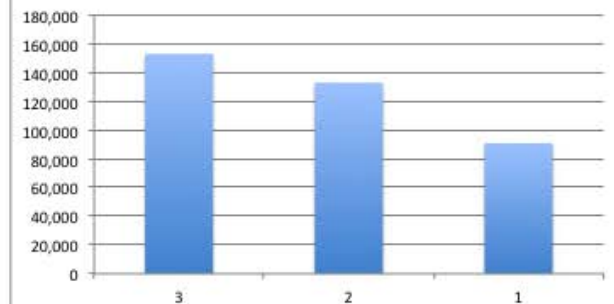
House File Modeling & Segmentation

Larger or less predicable and proven audiences

Smaller or more predictable transactional audiences

Logistic Regression Modeling

RFM Scoring

Tightening Techniques

Product affinity segmentation

Retail Trade Areas

Acquisition Source (e.g. Google vs. Catalog Mailed)

Serial Season Buyer (x number of holiday purchases)
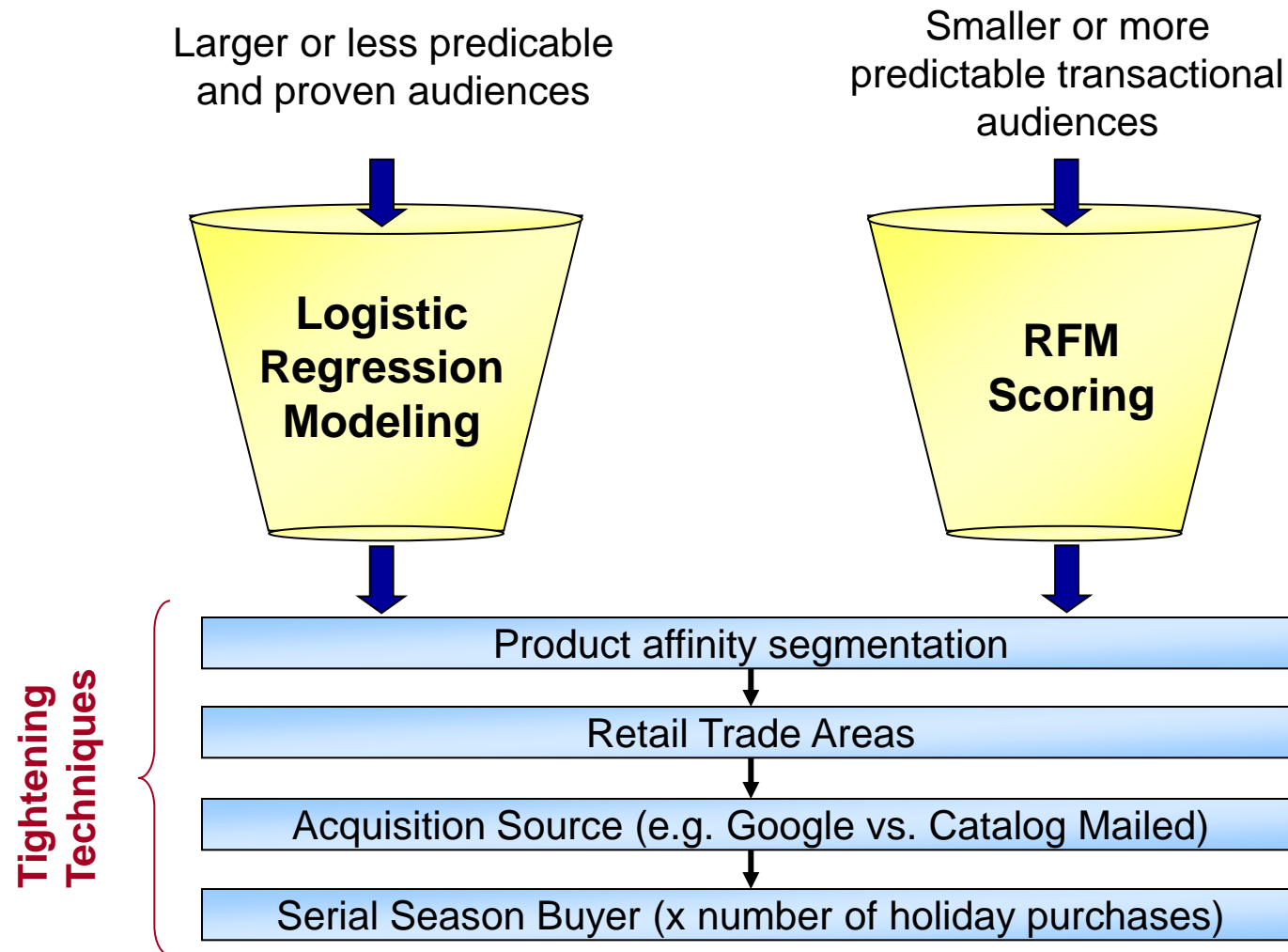
Catalog Multi-channel Planning & Analysis Process

1 Circulation Planning, Merge & Mailing

2 Matchback Process

3 Analysis

4 Reporting

Circulation Plan

Printer
Prints, addresses and mails catalogs to target audience.

Customers

House File Modeling
Logistic Regression, RFM Scoring, Retail Trade Area, Source Analysis, etc.

Mail File

Rental Affinity Lists

Merge / Purge
Receives files & performs merge/purge data hygiene, householding, radius mapping, etc. & Executes holdout instructions

Mktg Database
Matches back mail & holdout files and posts to promotion history. Non-matches (prospects) are held outside database and rerun against new purchasers

Mailed: Transaction File

Holdout Analysis
Compare rev per cust

Multi-Channel Revenue Report

Coop Prospects

Holdout File

Holdout: Transaction File

Non-Matched Prospect file

Holdout Methodology

Universe – Customer Database

Logistic Regression Modeling

Target Audience

Holdout Group

Representative Sample

Mailed Group

Nth'd sample not mailed.
Extraction criteria based on:
1) universe size
2) required sample size
3) projected response rates

Analyze:
Orders
Response Rate
ADT
Revenue

**Incremental revenue measured
by lift in performance of mailed
group vs. holdout group**

Analyze:
Orders
Response Rate
ADT
Revenue
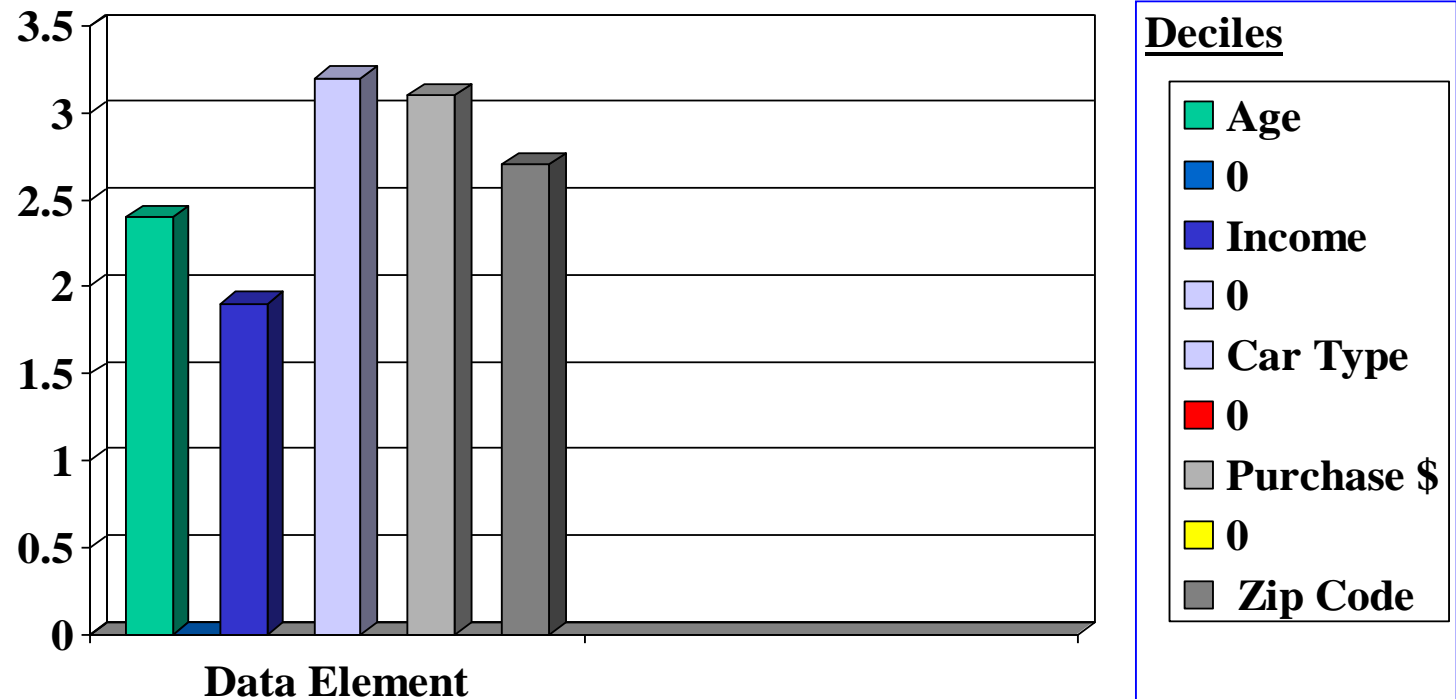
Modeling

- A profile of the attributes of the population of your database is descriptive of your customers.

- A database Model can predict which names are most likely to respond.

Modeling

Modeling Steps

- Determine which data elements are likely to be predictive of response

- Assign Data to all records

- Random Nth Select test promotions to list containing different segments

- Analyze response data – Runs separate reports

- Test model projections against customer files

Zip Penetration / Trade Area Scoring

- Zip Penetration / Trade Area scoring utilized to target & optimize geographically
- Create primary, secondary and tertiary zip code markets.
- Tag customers within each market
- Mainly utilized as an optimization technique
- Can supplement by utilizing desktop mapping software like MicroSoft MapPoint

Sample Data

Markets	% of Zips	% of Orders	% of Sales	% of Population	Orders/Pop
Primary	28%	83%	83%	71%	0.30%
Secondary	72%	17%	17%	29%	0.15%

Modeling Works

- Summit Racing can predict which names are like to respond so accurately that they can have rented list out perform catalog requests.

- Credit Card Companies
- Charities

Contact Information

- Jim Doyle

  jdoyle@american.edu

  james.m.doyle@usps.gov

GeoSelector

The following slides are screen grabs from www.geoselector.com

Zip Code 10304 Staten Island New York

Mapping for the Entire Zip Code

City:

State: [▼] Zip: 10304

Find Address

Interactive Help

Your count is complete, and
is 9,929 records. It is now
displayed in the selection
window below. You may
click on Select
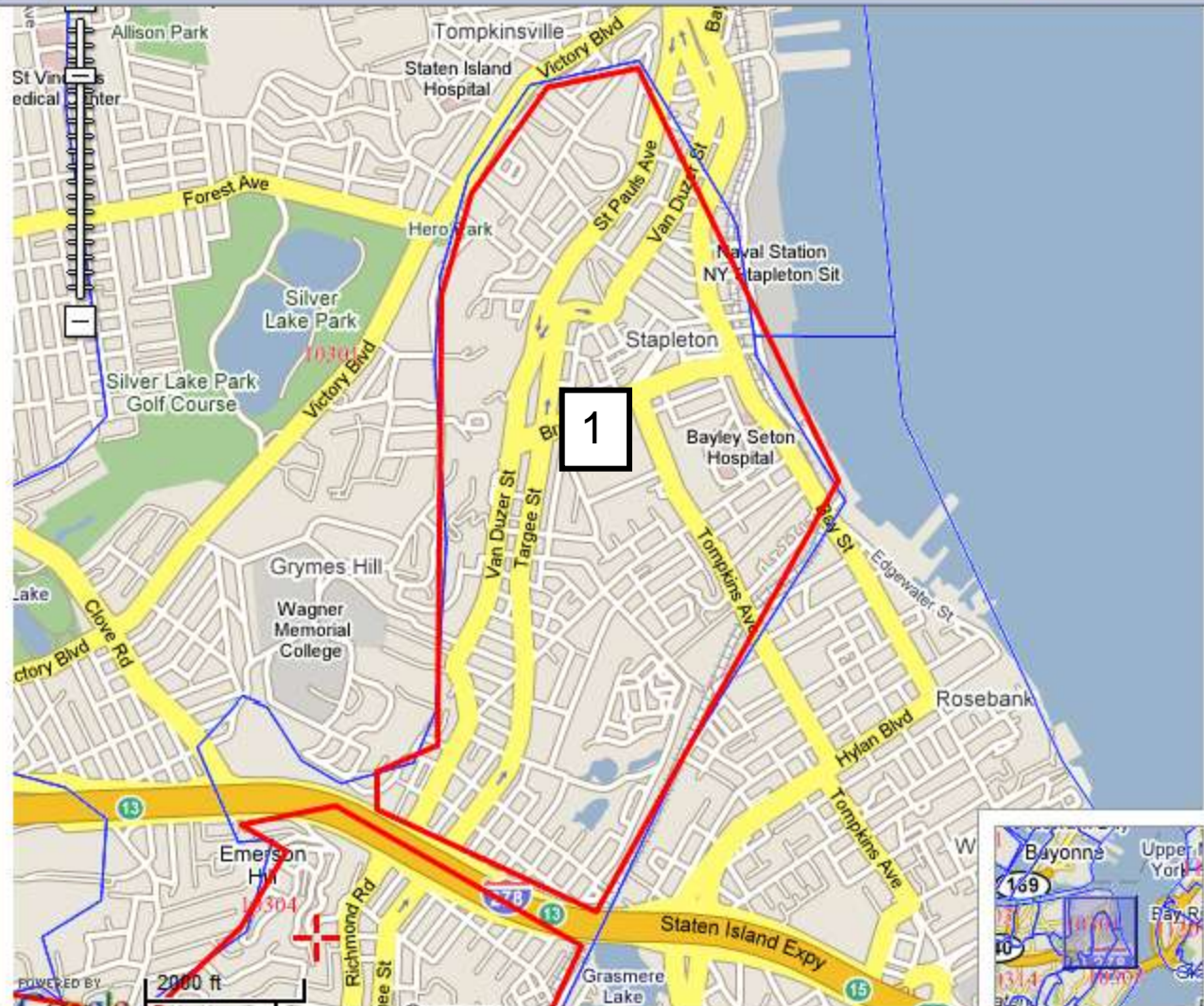Demographics, or Continue
Order to proceed

Continue Order

#1: Count=9,929 Selection 4
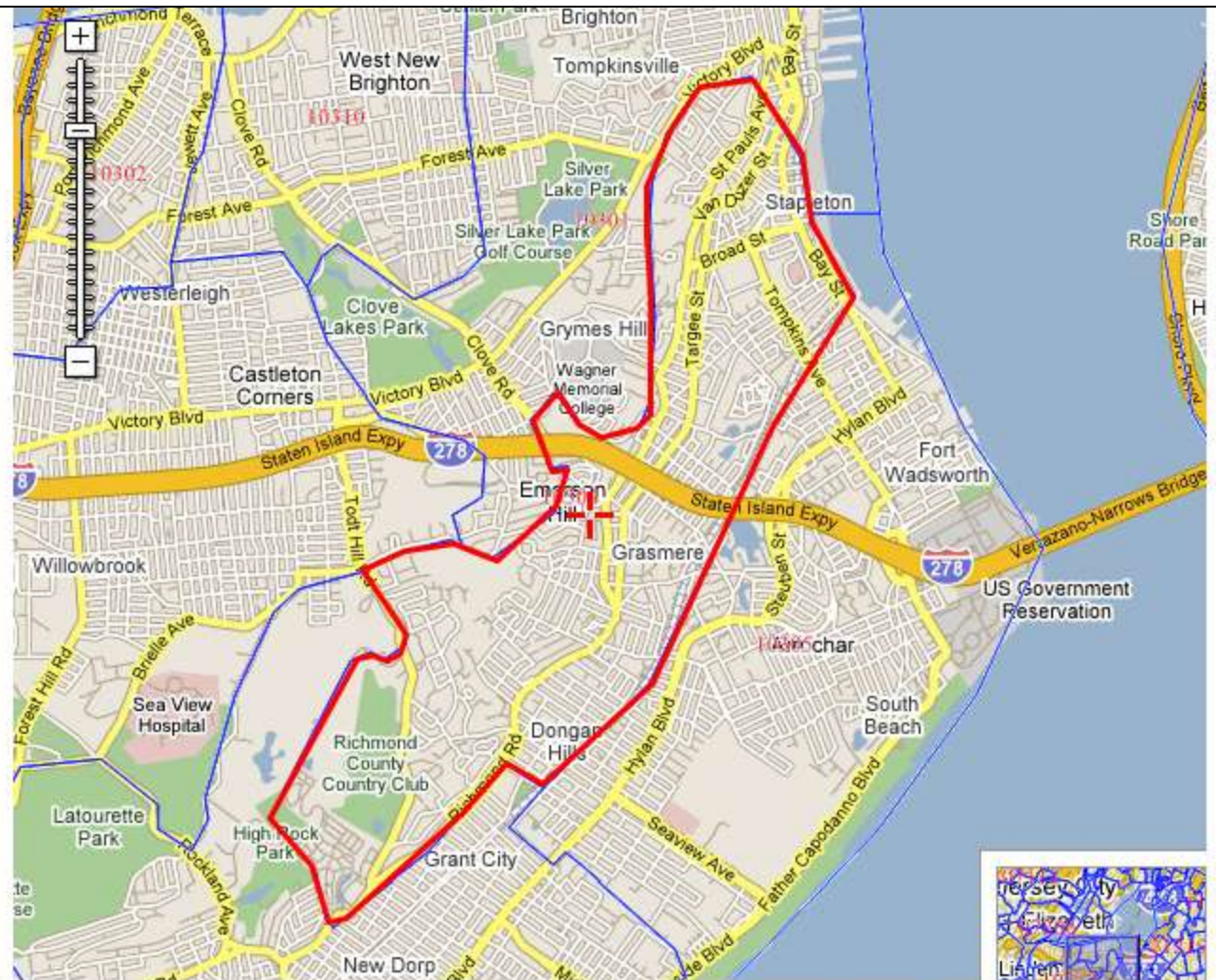
Internet

The Huntington Learning Center location
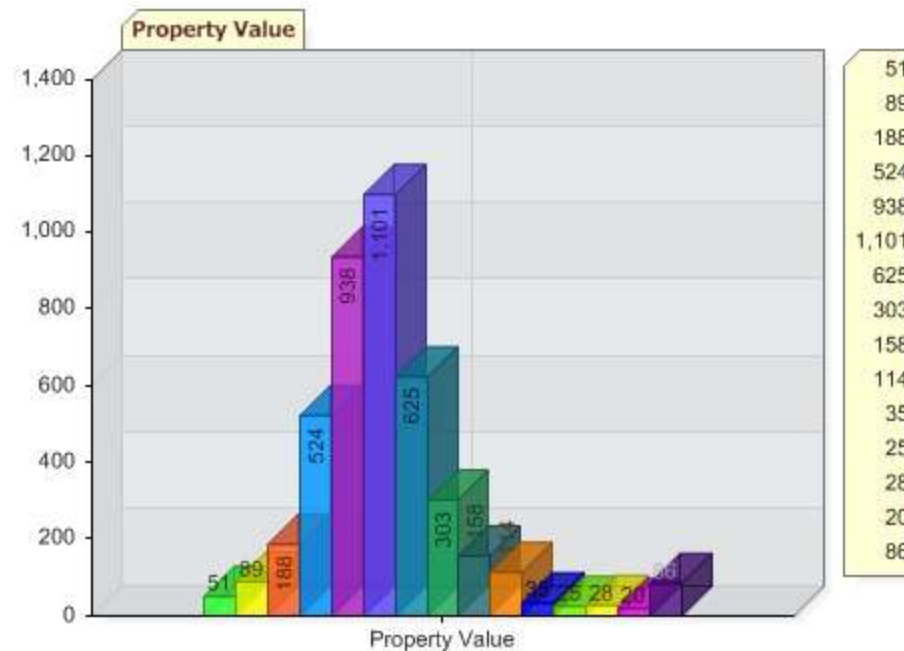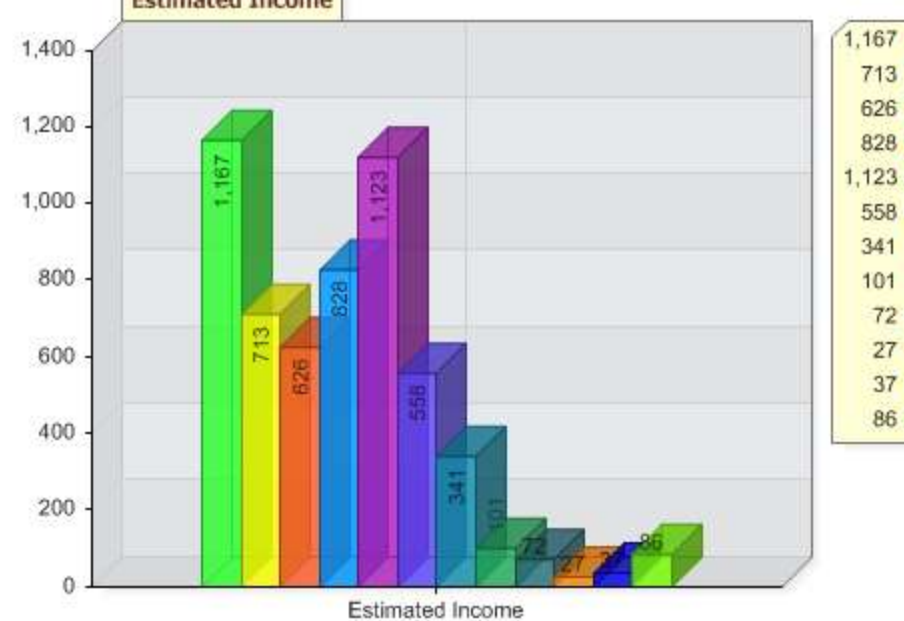in relation to the zip code.

Section One Map

Section One - Demographic Report

The top personality for your selection is African-American Neighborhoods K0

Young, working-class minority city dwellers and single-parent families with low educational levels living in older homes and low-rise apartments

For a detailed look at your top personality, and to browse other personalities click here

Estimated Income

Estimated Income
1,167
713
626
828
1,123
558
341
101
72
27
37
86

Property Value

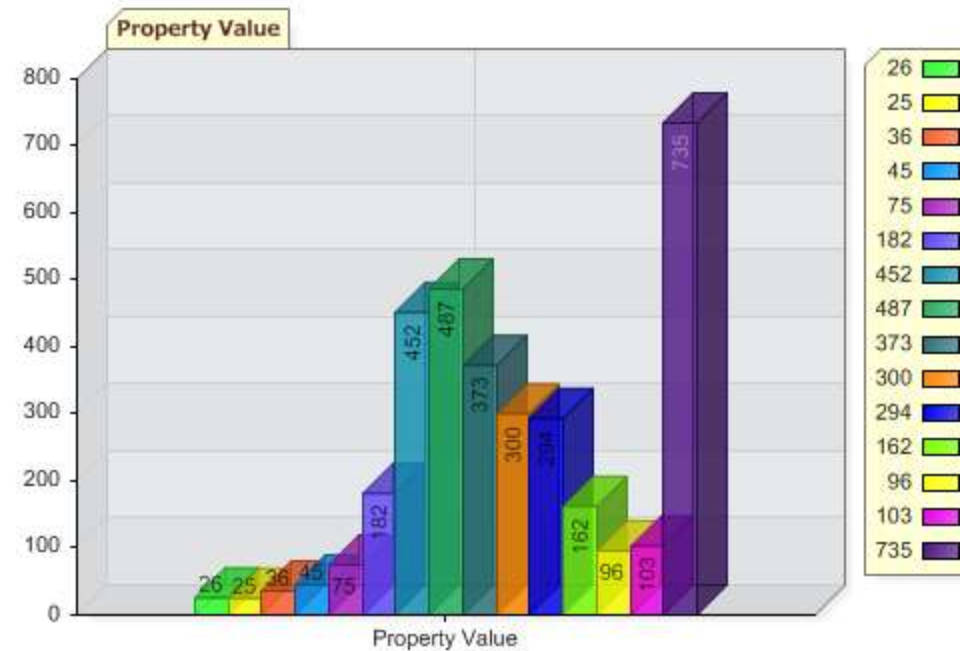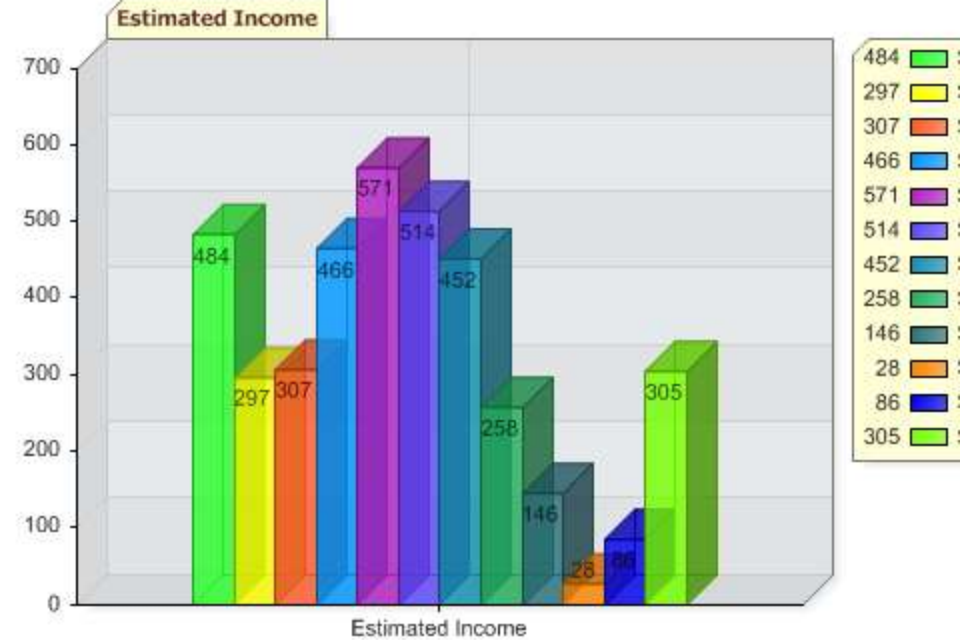| Property Value |
|---|
| 51 |
| 89 |
| 188 |
| 524 |
| 938 |
| 1,101 |
| 625 |
| 303 |
| 158 |
| 114 |
| 35 |
| 25 |
| 28 |
| 20 |
| 86 |

The top personality for your selection is Urban Commuter Families B03

Upscale, college educated Baby Boomer families and couples living in comfortable, single detached homes in city neighborhoods on the metropolitan fringe

For a detailed look at your top personality, and to browse other personalities click here

Done

Estimated Income

Estimated Income

Property Value

Property Value

Section Three Map

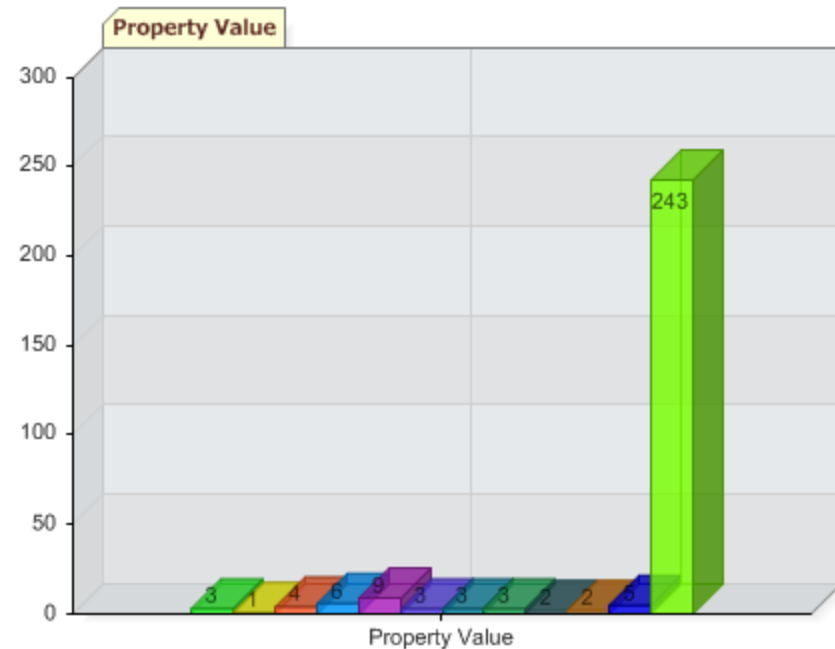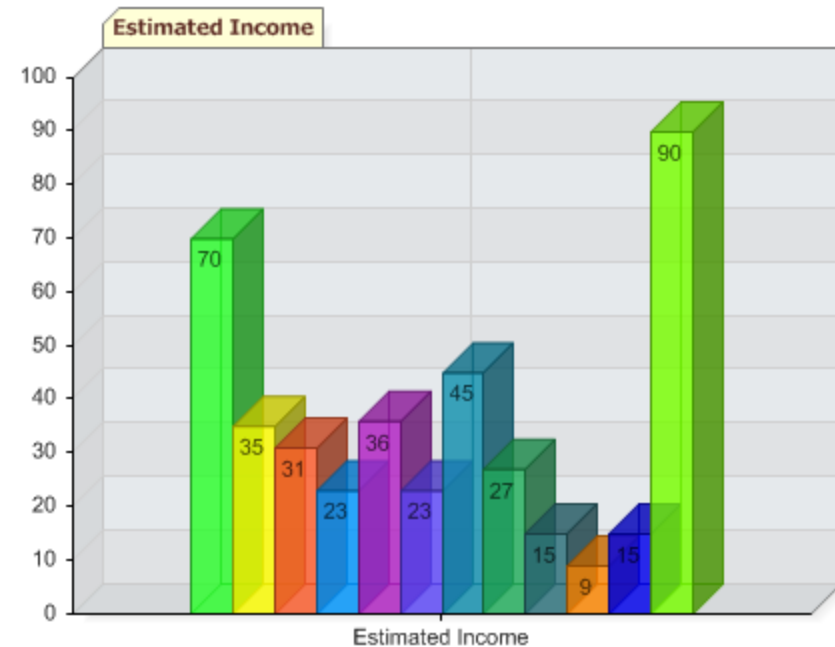**Total Households Analyzed:** 419

Section Three Demographics

The top personality for your selection is America's Wealthiest A01

The nation's most wealthy households are financially secure with expensive tastes, living the most luxurious of lifestyles in the most affluent and exclusive communities

For a detailed look at your top personality, and to browse other personalities click here

Your top 10 GeoSelector Personality Types

Estimated Income

70 35 31 23 36 23 45 27 15 9 15 90

Estimated Income

Property Value

3 1 4 6 9 3 3 3 2 2 3 243

Property Value

This Huntington Learning Center location
is within a few miles of some of best demographic households in the county.
However, excluding all of zip code 10304, these families would be missed.

In the January mailing, sections 2 and 3 were included. Section one was omitted.

Westport for Feb 5th Mailing

Contact Information

- Jim Doyle

  jdoyle@american.edu

  james.m.doyle@usps.gov