Training Session I

DATA MANAGER | INSIGHT | PREDICT

http://support.lityxiq.com/

Training Session I Agenda

1 ½ - 2 hrs

Introduction to Predictive Modeling

Quick Q&A session based on prior review of the deck "Introduction to Predictive Modeling"

- What Are Predictive Models
- Overview of Model Building
- Why Build Models
- Examples of Models
- Typical Data Used to Build Models
- Classifying Models and Algorithms
- Deciles*

- Model Development Process
  - Business Understanding
  - Data Understanding
  - Data Preparation
  - Modeling
  - Performance Evaluation
  - Deployment

**Open separate excel sheet called 'Deciles Discussion'*

Predictive Modeling in Action

Short walk-through of some example Lityx case studies. Any of:

- Predicting Retail Customer Behavior
- Cost Benefit Comparison
- Predictive Analytics for Email Campaigns
- Cancer Detection with Sensor Data
- Optimizing Casino Patron Loyalty
- Optimizing Print Channel Acquisition

Data Manager

A solution embedded into LityxIQ used to import, clean, manage, and manipulate data.

- Additional Tools:
  - Notification Settings
  - External Connections
  - File Manager
- Create New Dataset (RAW)
  - Go over everything in Source Definition
- Create New Dataset (VIEW)
  - Incoming Data
  - Joins
  - New Fields
  - Filter
  - Drop Fields
  - Aggregate & Sort

- Data Manager Links
  - Manage Dataset Libraries
- Console & Refresh
- Help

- Selected Dataset
  - Browse Data
  - Summary Statistics
  - Export
  - View Data Dictionary
  - Create View
  - Copy Definition
  - Deactivate

Insight

A solution embedded into LityxIQ used to create graphs and tables from LityxIQ datasets.

- Insight Links
  - Manage Chart Libraries
  - Dashboards & Galleries
- New Chart | Edit Chart Box
  - Setup
  - Series | Category Series | Analysis Metrics
- New Table | Edit Table Box
  - Rows and Columns
  - Metrics

Predict

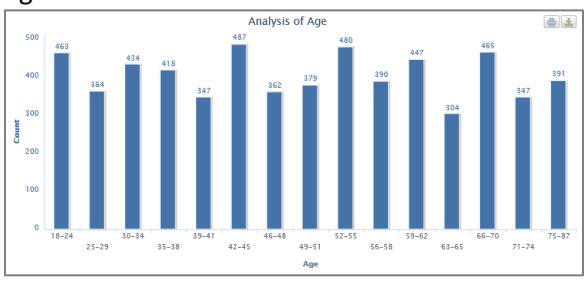
A solution embedded into LityxIQ used to create predictive models and score files with those models

- Predict Links
  - Manage Model Libraries
  - Scoring Catalogs
  - Scoring Jobs
  - Predictive Models
- Create New Model
  - Model Types; *Review Help Page on Model Types*
  - Walk through Model Build Settings
- Selected Model
  - Edit Model Build Settings
  - Execute Model; *Discuss use of Schedule*
  - Performance Analysis; *Output the results in a table*
  - Approvals & Implementation
  - Advance Major Model Version
  - Copy Model
  - Deactivate
- Scoring Catalog and Scoring Job
  - Create New Scoring Catalog
  - Create New Scoring Job
  - Export Data
  - Join Scores with Modeling Dataset and Graph

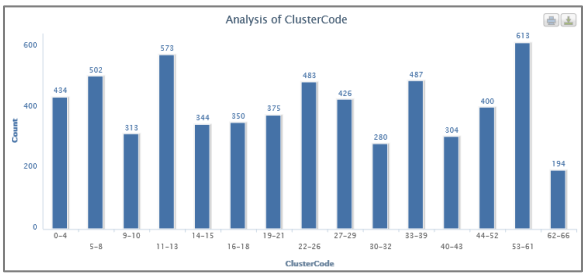
Analysis Scenario

Sample file from a marketer with information from an acquisition effort. Some of the individuals marketed were former customers and some had never been customers before.

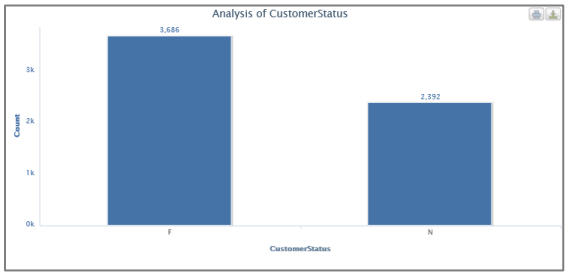
Age



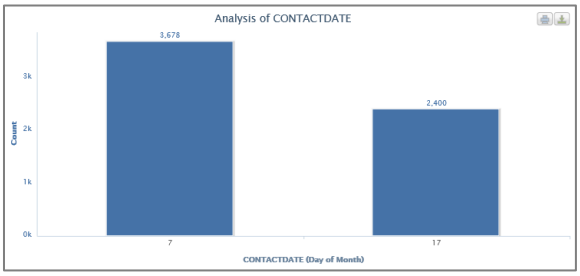
Cluster Code



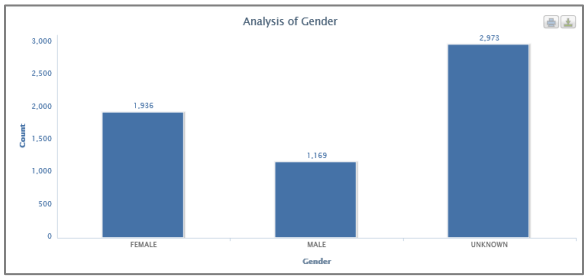
Customer Status (Former | Never)



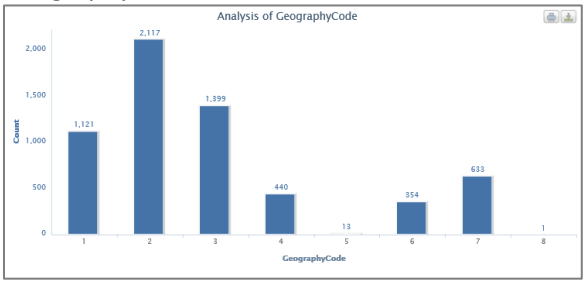
Contact Date (3/7 & 3/17)



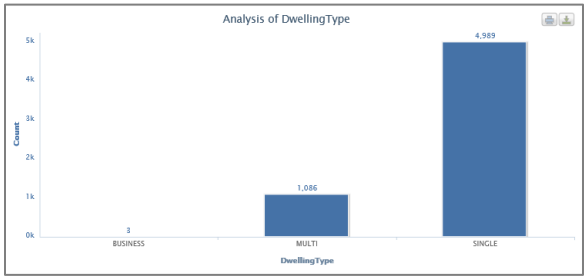
Gender



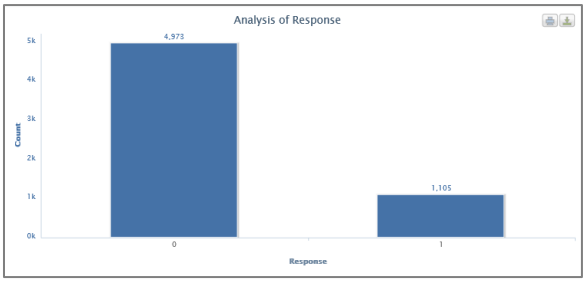
Geography Code



Dwelling Type



Response (18%)



EXERCISE

Data Manager

Read As Needed (links):

- What is Data Manager?
- LityxIQ Datasets
- Uploading A File
- Create a New Dataset
- Using the Console Window
- Creating a Dataset View
- Defining a Dataset View
- Defining Joins for a View

Exercise:

1. Change Project to 'Training Area' and within this project go to Libraries and create a dataset library called 'Training Session I_*your initials*'
2. Select this new dataset library. Choose Create New Dataset then Raw Data Source and call it 'Campaign_*your initials*' by reading in 'DemoResponse.csv' from the Public folder. How many records and variables were read in?
3. Open the Console Window to see that everything ran correctly.
4. QC the Campaign dataset using 'Browse Data' and 'Summary Statistics'. What is the mean age on the file?
5. Create another new dataset called 'Zip Demos_*your initials*' by reading in 'GeoTraits.csv' from the Public folder. How many records and variables were read in?
6. Open the Console Window to see that everything ran correctly.
7. QC the Zip Demos_*your initials* dataset using 'Browse Data' and 'Summary Statistics'. What is the max '# of 'HH' on the file?
8. Create a Dataset View called 'Campaign Data with Demos_*your initials*' and join 'Zip Code Demographics_*your initials*' to 'Campaign_*your initials*' joining on the field 'zip'. Keep from the zip data # of HH, # of Persons, % AA/B, % Asian /Pac Isl, % HH/L, % Pop 50+, % Pop 60+, % Pop 75+, % Spanish Speaking, % White, Avg HH Inc (x100), Avg Home Val (x100), Med HH Inc (x100), State, Zip. You should have 15 selected. What is the max '# of HH' on the new view?

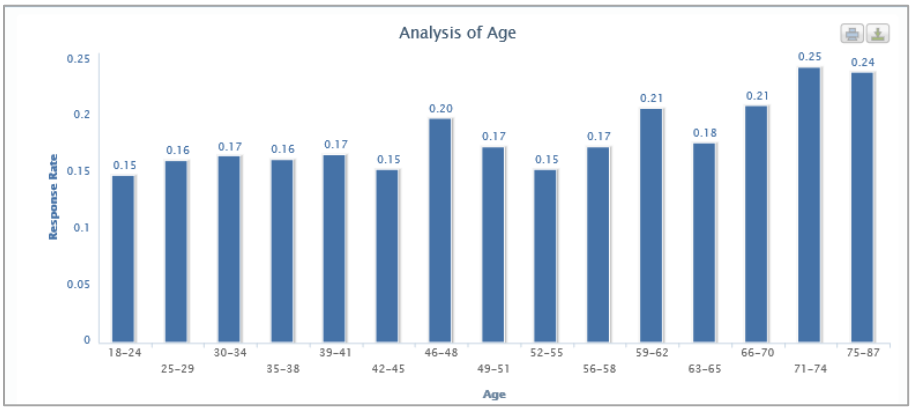
Insight

Read As Needed (links):

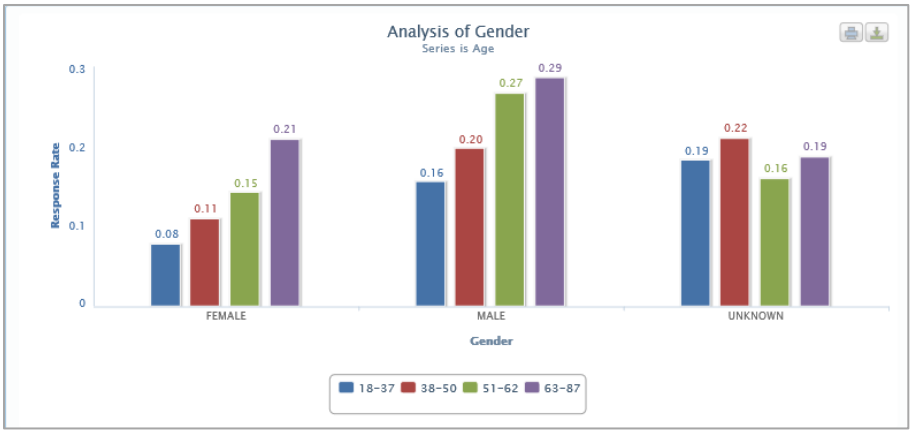
- Approaches for Generating Insights
- Navigating the Main Insight Links
- Creating a New Chart
- Defining a Chart
- Computing Summary Metrics
- Change Display Settings

Exercise:

1. Create a new chart called 'Analysis of Age on RR%_*your initials*'
2. Use the dataset call 'Campaign_*your initials*'
3. X-Axis is Age and Number of Buckets set to 15
4. Analysis metrics create a summary called Response Rate applying Summary Function 'Average' on Variable 'Response'; Add data labels.
5. Your chart should look like this:



6. Create a new chart called 'Analysis of Age & Gender on RR%_*your initials*'
7. Use the dataset called 'Campaign_*your initials*'
8. X-Axis is Gender
9. Use Category Series 'Age'; Variable Settings, Number of Buckets 4; Display Settings add Data Labels. Analysis metrics create a summary called Response Rate applying Summary Function 'Average' on Variable 'Response';
10. Your chart should look like this:



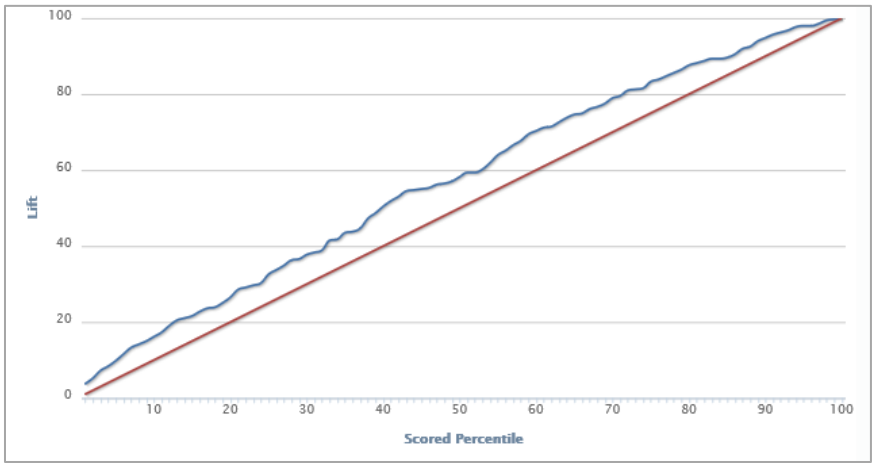
Predict

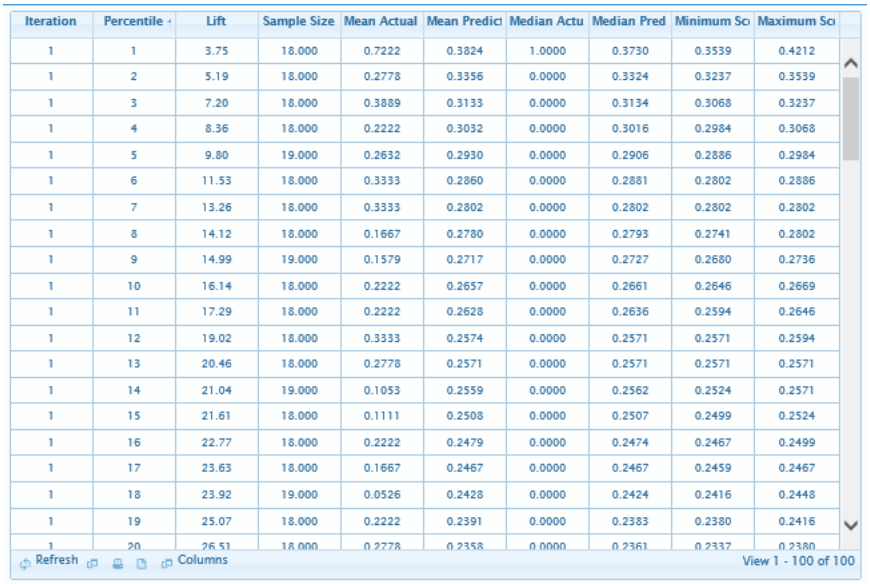
Read As Needed:

- Building Models for Novice Users
- What Can You Do in Predict
- Creating a New Predictive Model
- Types of Models
- Define Settings for Building a Model
- Running a Model

Exercise:

1. Create a new response model called 'Response Model 1_*your initials*'. Use the dataset called 'Campaign Data with Demos_*your initials*'; Response indictor field choose 'Response'. Remove from predictors 'AddressID' and 'CustomerID'.
2. Examine the model performance and show a graph that looks like the one below. What is the Lift of the model? Bring up a chart and table like the ones below. What predictors came into your model?





Iteration	Percentile	Lift	Sample Size	Mean Actual	Mean Predict	Median Actu	Median Pred	Minimum Sc	Maximum Sc
1	1	3.75	18.000	0.7222	0.3824	1.0000	0.3730	0.3539	0.4212
1	2	5.19	18.000	0.2778	0.3356	0.0000	0.3324	0.3237	0.3539
1	3	7.20	18.000	0.3889	0.3133	0.0000	0.3134	0.3068	0.3237
1	4	8.36	18.000	0.2222	0.3032	0.0000	0.3016	0.2984	0.3068
1	5	9.80	19.000	0.2632	0.2930	0.0000	0.2906	0.2886	0.2984
1	6	11.53	18.000	0.3333	0.2860	0.0000	0.2881	0.2802	0.2886
1	7	13.26	18.000	0.3333	0.2802	0.0000	0.2802	0.2802	0.2802
1	8	14.12	18.000	0.1667	0.2780	0.0000	0.2793	0.2741	0.2802
1	9	14.99	19.000	0.1579	0.2717	0.0000	0.2727	0.2680	0.2736
1	10	16.14	18.000	0.2222	0.2657	0.0000	0.2661	0.2646	0.2669
1	11	17.29	18.000	0.2222	0.2628	0.0000	0.2636	0.2594	0.2646
1	12	19.02	18.000	0.3333	0.2574	0.0000	0.2571	0.2571	0.2594
1	13	20.46	18.000	0.2778	0.2571	0.0000	0.2571	0.2571	0.2571
1	14	21.04	19.000	0.1053	0.2559	0.0000	0.2562	0.2524	0.2571
1	15	21.61	18.000	0.1111	0.2508	0.0000	0.2507	0.2499	0.2524
1	16	22.77	18.000	0.2222	0.2479	0.0000	0.2474	0.2467	0.2499
1	17	23.63	18.000	0.1667	0.2467	0.0000	0.2467	0.2459	0.2467
1	18	23.92	19.000	0.0526	0.2428	0.0000	0.2424	0.2416	0.2448
1	19	25.07	18.000	0.2222	0.2391	0.0000	0.2383	0.2380	0.2416
1	20	26.51	18.000	0.2778	0.2358	0.0000	0.2361	0.2337	0.2380

Refresh Columns View 1 - 100 of 100

3. Put 'Response Model 1_*your initials*' into production so you can score a file with it.
4. Create a Scoring Catalog called 'Scoring Catalog_*your initials*' using 'Campaign_*your initials*' as the template dataset and CustomerID as the primary key.
5. Create a New Scoring Job called 'Scoring Job 1_*your initials*' selecting your response model, the dataset to score 'Campaign_*your initials*' in your scoring catalog with any grouping method or number of groups you prefer.
6. Run your scoring job now. Open the console to see when it finishes.
7. Go to Scoring Catalogs, select your catalog and browse your scores and deciles.

EXTRA (FUN)

8. Create a new dataset view called 'Campaign with Scores_*your initials*' joining 'Campaign_*your initials*' with 'Scoring Catalog_*your initials*' *on 'CustomerID'*
9. Go into Insight and create a new chart using 'Campaign with Scores_*your initials*' to graph your model deciles showing response rate.