SELECTING CITIES FOR A/B EXPERIMENT

RELEVANT JMP PLATFORMS AND STATISTICAL TECHNIQUES

Distribution : Histograms, Boxplots, and Summary Statistics

Graph Builder : Bar Charts

Multivariate : Correlation, Matrix Plot, Parallel Coordinate Plot

Clustering : Heirarchical Clustering

PROBLEM STATEMENT

The marketing analytics team within a national investment brokerage firm has been asked to assist the client services division in designing an experiment to test a new strategy to help the company gain share of the Mid-Net-Worth household (MNW) client segment. This new offering consists of providing this client segment with access to services and tools at no cost which are currently only being offered to the company's High-Net-Worth client segment (e.g., free quarterly meetings with a financial advisor, online planning tools, etc.).

The plan is to run a pilot of the new program in 7-10 test markets (i.e., large U.S. cities) over the course of a year and evaluate its potential impact on acquisition and retention of this client segment. The primary question the client service division has is how should it be determined which cities will serve as the test markets (i.e., receive the new offering) and which cities will each of those test markets be compared to in order to quantify the impact of the program.

JMP® STATISTICAL DISCOVERY

Clearly the amount of client acquision and retention in a city at the end of the year will be based upon not only the impact of whatever program that has been offered, but also a large number of factors and characteristics of that city (many of them unknown). In other words, if the results in a city that receive the new offering are compared to the results of another city that didn't receive the offering, how can those factors and characertitics of each city be controlled for so that it is reasonable to attribute any differeces observed in client acquisition and retention to the offering?

20 cities are candidates to be used in the experiment. Ideally 10 will be selected to receive the new offering (i.e., test group) and 10 others will be paired with each to continue with the current offering (i.e., control group). These 10 pairs are to be selected such that it is reasonable to assume that client acquisition and retention would be similar for each if they did in fact receive the same offering. The client services division is OK if only 7 pairs of cities are found, as they want to make sure that each city that does receive the new offering in the pilot program does have a corresponding city to match up with to serve as the control.

To describe each city hopefully identifying the key factors that would impact the client acquisition and retention outcome, 9 characteristics for each city was gathered.

DATA SET

Selecting_Cities_for_A-B_Experiment.jmp

City	Metropolitan Statistical Area (MSA). A city and its surrounding area.
Population	Population (in thousands)
PCPI (Adj)	Average Personal Income adusted for the cost of living of each city.
MNW Ret	The % share of the Mid-Net-Worth Retiree segment the company has in that city
MNW Mid	The % share of the Mid-Net-Worth Midlife segment the company has in that city
MNW Young	The % share of the Mid-Net-Worth Young segment the company has in that city
Brand Sent	The sentiment that people in each city has towards the this particular companies brand. Score 1-100 with high number representing a more positive sentiment.
FS Sent	The sentiment that people in each city has towards the financial services industry in general. Score 1-100 with high number representing a more positive sentiment.
Growth Rate	A score representing rate of growth (or decline) company has been experiencing the past 5 years in each city. Growth rate is a composite score comprised of many variables such as acquisition, retention, asset inflows and balances, customer engagement, among others.

2

Marketing Eff	A score representing how effective this company's marketing and advertising efforts has been in each city during the past 5 years. It's a composite score made up of many different variables related to the success of various marketing/advertising campaigns in those cities. Positive number means the effectiveness of marketing/advertising has been increasing while a negative value indicates it has been declining.

EXCERCISES

1. As a means to get familiar with the data, create histograms, boxplots, and summary statistics for each of the variables. *Hint: Use Distribution platform.*

2. Are there any cities that stand out as being much different than the others relative to any of the variables? If so, plan to peform the cluster analyses in Exercise 5 with and without those cities to see if these cities have a strong influence on the results.

3. As a means to examine the % market share the company has for the three different client segments across the 20 cities, create the following visualizations: a) Bar Chart showing the % market of the three segments for each city (*Hint: use Graph Builder*); b) Scatterplot Matrix and Parallel Coordinate Plot for the three client segment variables (*Hint: use Multivariate Platform*). Use these graphs to identify cities that have very high/low % share in these three client segnments. Are there any cities that have a relatively large or small % share across all three segments? Are there cites that have a high/low % share in one or more client segments but opposite in another? Provide a few bullet point description of some of these findings in a way that would be useful for management of client services and marketing to understand.

4. Calculate the correlations between the 9 city characteristic variables and create a Scatterplot Matrix (note: *there are 36 pairs*). Identify some of the variables that are most correlated? Provide a brief intepretation of what some of these correlations mean in a way that would be useful for management of client services and marketing to understand.

5. Perform a Hierarchical Cluster Analysis. *Hint: Use Clustering platform.* Add a Constellation Plot.

   a. Examining the Dendogram and Constellation Plot, identify at least 7 pairs of cities that are "similar" to each other relative to the 9 characteristics.

   b. Perform the cluster analysis a second time excluding any cities that you identified in #2 as being very different from the rest. Find at least 7 pairs of cities. Compare to the solution in 5a. If there are differences, which set would you recommend be used in the experiment?

   c. For each pair, how will you decide which city receives the new service offering and which one does not?

JMP® STATISTICAL DISCOVERY

6. The plan is to analyze 3 outcome variables from the experiment: 'New Client Acquisitions' , 'Retention Rate' , and 'Amount of Deposits'. Since the cities have different population sizes, what are some idea on how to modify the resulting outcome data so that analyses comparing the results across the cities are on the same scale?

7. Are there any other characteristics of these cities that you think would be helpful to include in the cluster analysis in order to improve upon matching the cities together? *Hint: These are factors that would potentially impact the outcome variables.*

jmp STATISTICAL DISCOVERY