

SmartSell:
Site Performance Optimization
Overview

Table of Contents

Executive Summary	1
Introduction	2
Why SmartSell?	2
Key Concepts	3
What is Success?	3
Analysis: Models	4
Synthesis: Architecture and Content	5
SmartSell Components	5
Case Study	7
Introduction	7
SFP	7
Data Collection	7
Data Analysis Results	7
SFP Analysis and Recommendations	7
SiteEval	8
SiteEval Recommendations	9
DynModel (Dynamic Model)	10
I-Site Analysis	11
Methodology	11
Visualization Guide	13
I-Site Analysis of Client Rewards Site	15
Statistical Analysis	15
MAP Visualizations	16
3View Analysis	21
Conclusions	22

Executive Summary

SmartSell™ is a set of tools and an integrated process for designing measurably more effective websites, and for improving the effectiveness of existing websites. The tools and process leverage both the our "best practices" over many Fortune 1000 clients, and the plethora of data gathered by web servers. We have now refined and systemized these tools and processes and named them SmartSell™.

SmartSell's main focus is on providing an objective basis for design improvements: it provides both analytical components that identify problem areas, and evaluation components that can evaluate proposed solutions.

SmartSell also provides guidance in Search Engine Optimization (SEO) that's complementary to the well know techniques involving optimal use and placement of keywords. It's focus on prospect motivation and site effectiveness guides more effective choice of keywords, and refinement of the user visible content in search results - to get more clicks the site.

SmartSell's power derives from using models for both the analysis of website content and of many users' interactions with the site (via site logs). By using model of how the user makes a purchase decision, we can translate raw data back to sales-meaningful terms, and then use sales-focused techniques to drive a synthesis of changes to improve the site.

While SmartSell can provide maximum value in the case of the redesign of an existing site, where both site log data and the site itself can be incorporated into the analysis, it is also very effective in the design of completely new sites. SmartSell requires only that the client have some ideas of why customers might buy a product to capture and incorporate those ideas into an effective initial site design.

Introduction

Why SmartSell?

As uses of the Internet mature, and management demands real focus on the bottom line, there is a need for a site design process that guarantees results, and provides insight into those results before dollars are spent. We've responded by systemizing our "best practices" over six years of experience with FedEx, Sun, CNA Insurance, Cisco, etc. We've created technology and process that allows both predicable design of more effective websites and the estimation of effectiveness gains over an existing site.

Some examples of both the worldview and benefits that SmartSell™ provides:

Site One

SmartSell analysis was applied to a subsite selling a complex and unusual outsourcing service. Analysis of the site logs revealed that one component of the site was strongly — and unexpectedly — correlated with success. Bringing that component to be near the top of the site, along with many other design changes resulted in a 41% increase in the site's success rate.

Site Two

SmartSell static analysis of the site revealed that the product value and generality (in the sense of the range of businesses, which it could serve) was not being adequately presented and supported. This change, along with many others as part of a redesign resulted in 14% increase in sales inquiries via the site.

Site Three

After an initial SmartSell interview, and a review of some existing survey data, an internal advertising banner on a site was modified to reflect the perceived benefit of a new product, resulting in a click rate going for 2% to 5%, on an otherwise unchanged site.

In all the examples above, we've talked about the major contributor to site effectiveness improvement, but, in fact, each (other than the last) involved dozens of architectural and content changes. SmartSell is mostly about these architectural and content changes. It provides an objective basis for recommending and evaluating these design issues.

Key Concepts

What is Success?

To focus on improving website design, we must define what dimensions that improvement will be measured along — we must define "success"

Early websites were relatively small and focused on a small number of functions, usually aimed at providing general information about the company. Contemporary websites contain a profusion of information and opportunities for interaction. In well-designed sites, visitors will find distinct locales, or subsites, that provide particular kinds of information or offer particular kinds of interaction.

The definitions of "success" for a large, complex site may be difficult to pin down. This is because there are many constituent subsites and the definitions of success for individual subsites are different. Our research suggests that there are four main types of subsite, each with its own definition of success.

Information Subsites

These provide the prospect with information. An example is finding Specification Sheets for all the products offered in a particular product line.

Success is measured by successful and efficient retrieval of the needed information.

Value/Function Subsites

These enable the prospect to perform a task that delivers value. An example is a configurator that enables a prospect to specify requirements for a computer and receive a valid configuration for a computer that meets the requirements, together with a price quote.

Success is measured by successful and efficient (in terms of the visitor's time) delivery of the valuable result.

Persuasive Subsites

As the name suggests, these subsite are focused on changing the attitude of a prospect towards some action: signing up for a mailing list, inquiring about employment at a company, or — and most often — persuading the prospect to buy a product. In this last role, they acquaint the prospect with the virtues of a product and attempt to induce the prospect to initiate the next step in the sales process. The next step might be Registration (being added to the mailing list), a Request for Contact (actively seeking a call from a Sales Representative), or placing an order (initiating an actual on-line sales transaction).

Success is measured by the percentage of prospects that take one of the offered positive actions.

SmartSell can be applied to any of these, but is oriented especially to the analysis and design of persuasive subsites.

Advertising Banners and Subsites

Advertising fundamentally consists of two steps: attraction and conversion. Success in attraction is measured by "clickthrus" — the fraction of viewers of an ad that go to the next step. Success in conversion is measured by the fraction of clickthrus that are converted to either sales or the next step in the sales process.

SmartSell can drive both aspects: it can attract by making sure the ad messages are the most effective and it increases the conversion rate by guiding the development of more effective sales supporting content.

Analysis: Models

Introduction

SmartSell is based on the idea of using Models of decision processes to capture insights into effective site design: make the website optimally support the prospect's decision process, and you maximize the website's success.

The elements that make up SmartSell models are:

- Market Segments, which means some distinct group of customers;
- Factors, which are the context-specific considerations that go into making a decision: perceived price, company reputation, perceived quality, esthetics of product, etc.
- Decision Processes, the ways in which Factors are combined by a specific Market Segment to reach a decision
- Content, which is what the user can potentially perceive at a website: text, images, video, audio, etc.

The sections below explore these concepts in a bit more depth.

Market Segments

In SmartSell, Market Segments have a very specific meaning: each Segment is a group of prospects who have a significantly different Decision Process to make a purchase decision. The distinctions in Process often imply a corresponding difference in what information/positioning/arrangement would be the most likely to produce a sale with a given Market Segment.

Factors

Factors are the internal components that a prospect will consider in making a specific purchase decision. They can range from the easily measurable (price, delivery time, weight...) to the ultimately abstract (company image, perceived product quality, status value of product...).

Decision Processes

The Decision Process is how a given Market Segment combines Factors and Content to lead to a Purchase Decision. Decision Processes can differ both in the parameters associated with Factors and/or in the actual structure: a Factor used in a simple weighted way by one Market Segment may be used as a Threshold that must be exceeded by another Market Segment.

Content

Content influences the decision by changing the prospect's beliefs about Factors, and thereby influences the decision.

Synthesis: Architecture and Content

The architectural design of websites is where SmartSell provides its initial value: by using the SmartSell Models, the site architect can determine both the most important things to focus on, and the most effective implementation of that focus.

For example, imagine that prospects coming to a website are nervous that goods ordered from the site might be damaged in shipping, but are also have an unrealistically high impression of the products pricing. SmartSell can provide guidance on both the placement of content to deal with these sales objections, and guidance on the tenor of the content to most effectively surmount them.

SmartSell Components

SmartSell has a number of components, which we describe abstractly below.

Note that while both the I-Site and DynModel components require an actual site to exist, SFP and SiteEval require only that you have some ideas about why prospects might buy your products.

SFP: Model Acquisition

SFP is an interview process and supporting technology that acquires and organizes the Segments/Factors/Processes information from your knowledgeable sales or marketing personnel.

Its results are both the Models use on the other components, and quantified recommendations on both the relative importance and effective value of the various Factors, as well as an analysis of how valuable it will be to recognize different Market Segments on the website.

SiteEval: Model based Site Content/Structure Analysis

SiteEval captures a model of the effectiveness of the content and structure of an existing site, via a formal analysis and evaluation of the content, and flow implicit in the site structure.

It provides an indication of where you can the most benefit out of improving the site. It also allows both numeric and visualized comparisons of alternative site designs for effectiveness.

I-Site: Model Based Log Analysis

I-Site both translates site log data into Model-related terms, and provides statistical and visualization tools for analyzing all the data.

It's results are statistical analyses and visualized patterns that identify major contributors to both success and failure, Both model-based and model-free statistics and visualizations are provided, so that the viewpoint is not limited to the model.

DynModel

DynModel essentially "calibrates" the model with actual site log data, which allows prediction of the effects of many site changes.

The results allow various designs to be evaluated to predict effectiveness prior to implementation, assuming no significant change in Factors or Market Segments properties and no drastic site architecture changes.

3View: Evaluate each model element (Factors, Content, Decision Process) for accuracy/contribution

3View analyses how well each element of the model performs, relative to the actual data.

Its results are either explicit reports on how effective the element is, or how much it's effectiveness could be improved under various assumptions. Both provide strong guidance on what to change to get the most benefit.

Case Study

Introduction

The following sections represent a single analysis and recommendations to a client, which happens to involve a product with only a single market segment. The goal ("success") of the subsite is to sign up prospects for a "frequent buyer" program, in order to encourage more use of the basic product.

SFP

Data Collection

The raw data was obtained by interviewing two people familiar with customers and their motivations. Since there was strong agreement among these people, no further interviews were felt necessary.

Data Analysis Results

Factor Name	Raw Interest Percent	Site Interest Percent	Content Impact Percent
Facilities in Visited Areas	11%	17%	26%
Ease of Understanding	2%	17%	21%
Cost to Join	2%	14%	17%
Other Programs	8%	15%	12%
Hassle Accumulating	20%	1%	11%
Point Pricing	23%	17%	5%
Hassle-Collect	10%	1%	5%
Will I get Value?	24%	17%	4%

SFP Analysis and Recommendations

Many prospects are unaware that the client now has participating facilities in essentially all metropolitan areas — and that this is a distinct advantage over the competition.

- On any redesign, this point should be emphasized on the initial benefits page – it's a major advantage and a major concern to potential members. Note this lends itself to a graphics treatment: showing participating hotels scattered across a US map. It's also possible to show other, competing hotels (Hilton) to provide a visually striking contrast and clear advantage.
- Also note that these responses answer the "Other Programs" issue, by making it clear that the client's program is useful in many more localities than most competing programs.

Based upon both general impressions and competitive experiences, many travelers believe such programs to be complex, and are therefore skeptical about receiving real benefits:

- It's therefore important to emphasize the relatively simplicity and unrestrictive nature of the client's program in the initial persuasive material.
- Another idea would be to indicate what percentage (hopefully large) of members take at least annual advantage of the client's reward program.

Since some programs have a cost-to-join, there is an initial resistance to pursuing such reward programs.

- While currently emphasized on the page just prior to sign-up, the free nature of the program should be moved to and emphasized on the initial page, and reiterated at key points for additional impact.

SiteEval

With SiteEval, we use the SFP derived Models, plus a static analysis of the effectiveness of the site content, to analyze the effectiveness the site by examining the impact of its content and structure on the Factors affecting the prospect's purchase decision.

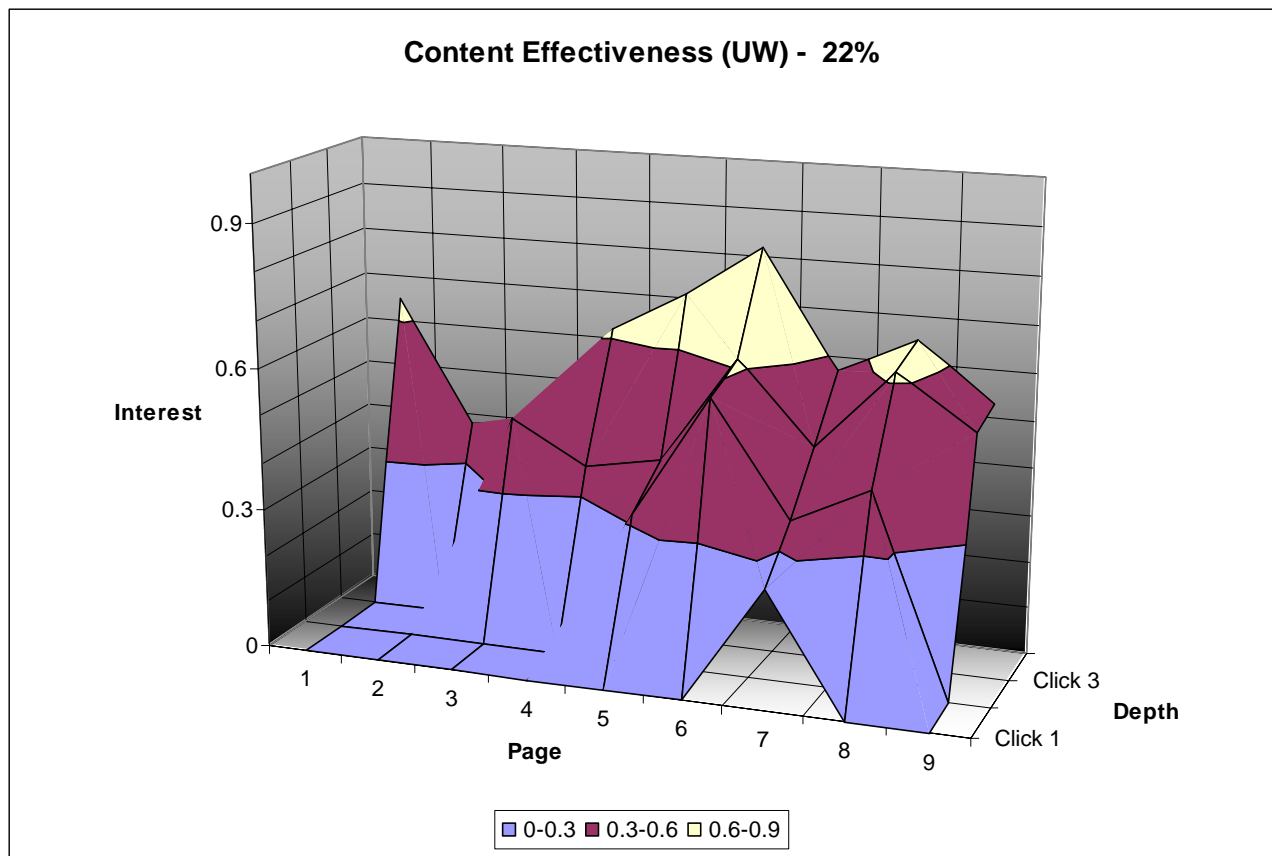
Besides providing an indication of what the optimum site performance could be, this static analysis allows us to evaluate alternative site designs to determine how much more (or less) effective they might be, relative to the original design.

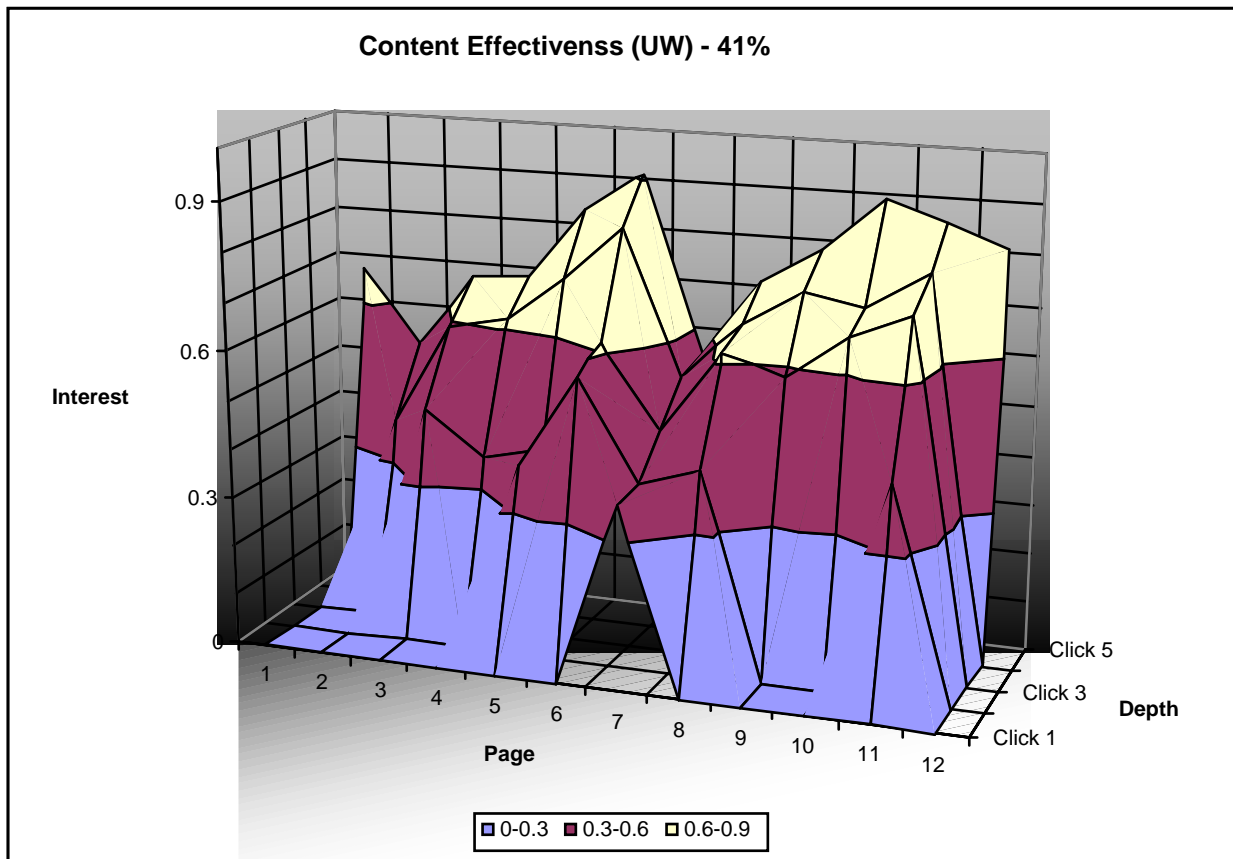
The graphs below show such a comparison of an original and an improved site design. The improvement is visualized by the increased size of the white areas at the top of the peaks, and is reduced to the single numbers (22% and 41%) at the top of the graphs. It's important to note that this does not predict a doubling of actual sales — it would indicate only some increase in sales, whose actual amount would depend on the distribution of visitors to the site, among other things.

SiteEval Recommendations

A cursory examination of the SiteEval data revealed that the site content at the top levels was very ineffective at persuading prospects to sign up. This was both because it diluted the marketing message with support for actual use of the product, and because the marketing messages present were only unclear "teasers" to link to the actual marketing content. The marketing content linked to was, in fact, quite effective. The solution was to restructure the site to bring the key marketing messages to the top, and link them to in-depth support.

As can be seen from the graphs below, this design roughly doubles the surface effectiveness of the site.





DynModel (Dynamic Model)

DynModel connects the actual behavior analyzed by I-Site with the Prospect Models and Site Models produced by SFP and SiteEval to allow us to predict the actual effectiveness of a new site design.

Essentially fitting the actual site's session data (I-Site data) to the two models, and then modifying the SiteEval model to reflect the design changes, and using the changed model to drive DynModel does this. Because DynModel is largely about producing a runnable model, it doesn't produce a lot of analytical output — but does produce the crucial prediction: sales increases.

The results of building a DynModel based on the existing Decision Models, SiteEval models and actual log data:

Success Behavior Explained:	47%
Failure Behavior Explained:	17%
Actual Sales/Predicted Sales (from Actual)	93%

The Success prediction is typical. The Failure rate is lower than usual, perhaps due to the relatively low success rate for the site, or it may indicate many visitors are there by mistake. It's suggested that referrer logs be kept and examined to determine the source of visitors to resolve this issue.

[The above numbers (especially Success Behavior Explained) serve to validate that our Models are sufficiently valid for use. If they weren't, we'd look for additional Factors, market segments, or errors in the SiteEval analysis — then refine the model to get improved validity.]

Results of testing revised design with DynModel:

Design change evaluation for Model Two, using Log data 7/2/98 to 8/7/98

Success Rate Change:	+27%
Sensitivity Analysis:	+5%, -12% for 10% noise, so expected range of Success Rate change is +15% to +32%

I-Site Analysis

Methodology

A team skilled in Web design, mathematical modeling and marketing analysis conducts an I-Site analysis.

Because of the fine-grained nature of the study, each analysis is typically focused on a particular subsite of up to approximately 50 pages. All that is needed for the basic analysis is http server log data covering at least several hundred traversals of the chosen subsite. A more detailed analysis can be performed if further data relating to these sessions is available, such as additional logs or visitor registration data.

A crucial element of I-Site analyses is that they are focused on discovering the “reasons” behind successes and failures across a statistically significantly sample of sessions.

Success is defined as getting to some specific page(s) on the subsite that indicate whether the user did what the site owner desired for that subsite: ordered product, on an ordering subsite; or retrieved some class of useful content on an informational subsite.

Failure is defined as exiting the session without visiting any of the pages defined as success.

The steps of an I-Site analysis include the following:

Site Inspection: A browser inspection of the site under study, including a detailed and systematic traversal of the chosen subsite.

High-Level Log Analysis: A preliminary analysis of the basic page hit statistics for the overall site using a conventional server log analysis package.

Data Reduction: The server log data is passed through a set of special programs to obtain session-oriented data views used in the rest of the study.

I-Site Statistical/Visualization Analyses: The data are subjected to Regression and Cluster Analysis procedures to find correlations (positive and negative) between various session characteristics (such as pages visited) with positive outcomes.

This process leads to performing some relevant subset of the following specific analyses:

Entry Opportunity Analysis: The data are viewed using multi-axis visualization software to examine which paths are followed from the overall site home page to the chosen subsite.

The specific goal is to identify patterns of access to the subsite, especially ones that are unanticipated or patterns of very high use.

Opening Move Analysis: A multi-axis visualization study identifying the paths taken during the first ten moves from the subsite entry point.

The primary goals are to identify patterns of access indicating unanticipated market segments, use of the site in an unexpected way, or to detect early exits due to poor site design.

Exit Analysis: A multi-axis visualization study of the last ten moves before exiting the subsite.

The interest here is in patterns that indicate unanticipated reasons for people exiting. Since exiting can occur after both success and failures, patterns can also provide insight into why a subsite works, by noting the non-trivial patterns in the last ten pages visited.

Page-Hit Analysis: A multi-axis visualization of pages hit before and after positive outcomes.

This presents a finer-grain view of content effectiveness than regression or cluster analysis because sequencing effects can be seen, providing information about the site structure effectiveness, as opposed to just the single page effectiveness that regression provides.

Page-Time Analysis: A multi-axis visualization of time-on page data for before and after positive outcomes.

Since page time is a better measure of overall page value/interest than simple hit counts, this provides more insight into the effectiveness of content.

Demographic Analysis: Multi-axis visualization study of differences between paths taken by prospects with known demographic differences. In its simplest form, this includes a breakdown by domain names. If more data, such as user registration data are available, more detailed breakdowns can be obtained.

This can detect both unusual patterns of access by demographic segment, as well as unanticipated needs, if unusual and inexplicable patterns of access are seen that are distinct for a demographic segment.

Content Topic/Factor Correlation Analysis: This identifies pages that strongly affect the frequency of successful outcomes, for better or worse. The text and imagery of these crucial pages is analyzed in terms of the key topics that appear to influence prospect interest, and the key decision-making factors embedded in marketing messages — price, perceived quality, product positioning, etc. — that appear to most affect the prospect's decision-making process.

Where the population of prospects breaks down into distinct segments (via additional demographic data) of differing behavior, this analysis examines how responses to different topics and factors differ among different segments.

The analysis provides a way to clearly focus on pages associated with success and failure. This analysis usually provides a few surprises. Pages strongly correlated with failure may mean they are having a negative impact and need to be revised. Pages with a strong positive correlation with success for which there's no obvious explanation imply we don't adequately understand the prospect's decision-making process.

Content Topic/Factor Cluster Analysis: Using traditional cluster analysis, content can be analyzed to locate groups of points that are close in some selected dimensional mappings of the data.

Such clusters could indicate market segmentation, differing task goals in site use, or undetected demographics groups.

Visualization Guide

While the outputs from I-Site require quite a bit of interpretation, these are some immediate insights to be gained from the various imaging tools:

Correlation Outcomes Charts

This chart indicates how often each page of the subsite was associated with a positive outcome for a session.

The chart is sorted and colored by correlation (graphed on the vertical axis), so that green bars on the right indicate pages associated with success, and red bars on the left indicate pages that made success less likely when visited.

Cluster Charts

Cluster charts are simply 3D plots of three variables associated with a set of sessions, with the points being colored based on their closeness on the graph. The most common is to plot time-on-page value for three different pages of interest that are indicative of some market segment. If clearly distinct patterns (clusters) of time-per-page are apparent, and there are not obvious tasks to associate with those clusters, then it's likely new market segments or new task types have been discovered.

Since there are a very large number of such charts that could be generated, the ones examined are usually found through MAP, when interesting multi-variable patterns are detected.

Multi-Axis Plots (MAP)

MAP's are simple to explain, but difficult to interpret in the static view on a printed page.

They consist of a series of vertical axes on which the various variables (for instance, time-on-the-page, where each axis is a different page) associated with each session are plotted, with those points representing a session then being connected by line segments between the axes. The result is a "spaghetti" of lines, which are then examined for patterns by dynamically selecting subsets of the lines, and looking for visual patterns.

What is generally discovered are both simple and complex correlations — both the same ones discovered by the "Correlations Outcomes" and more subtle ones that shows as patterned bundles of lines. It is also possible to discover multi-dimensional clusters that would not be caught by conventional cluster analysis.

I-Site Analysis of Client Rewards Site

Statistical Analysis

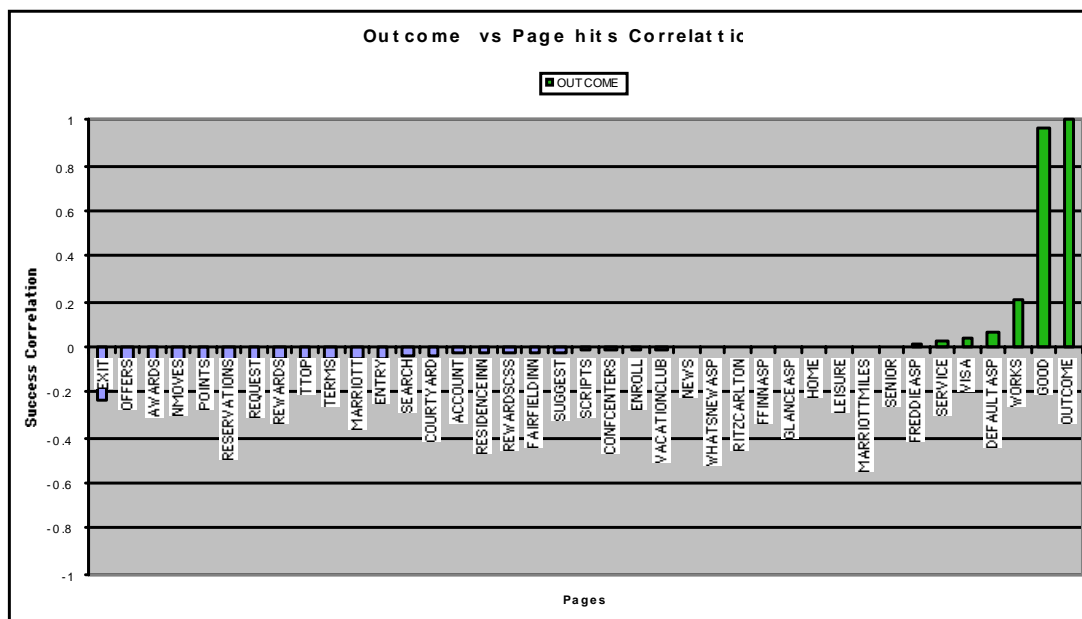
Exploring Correlation with Success

In the graphs below, you can see which pages are associated with success, considering both time on page and simple hit counts.

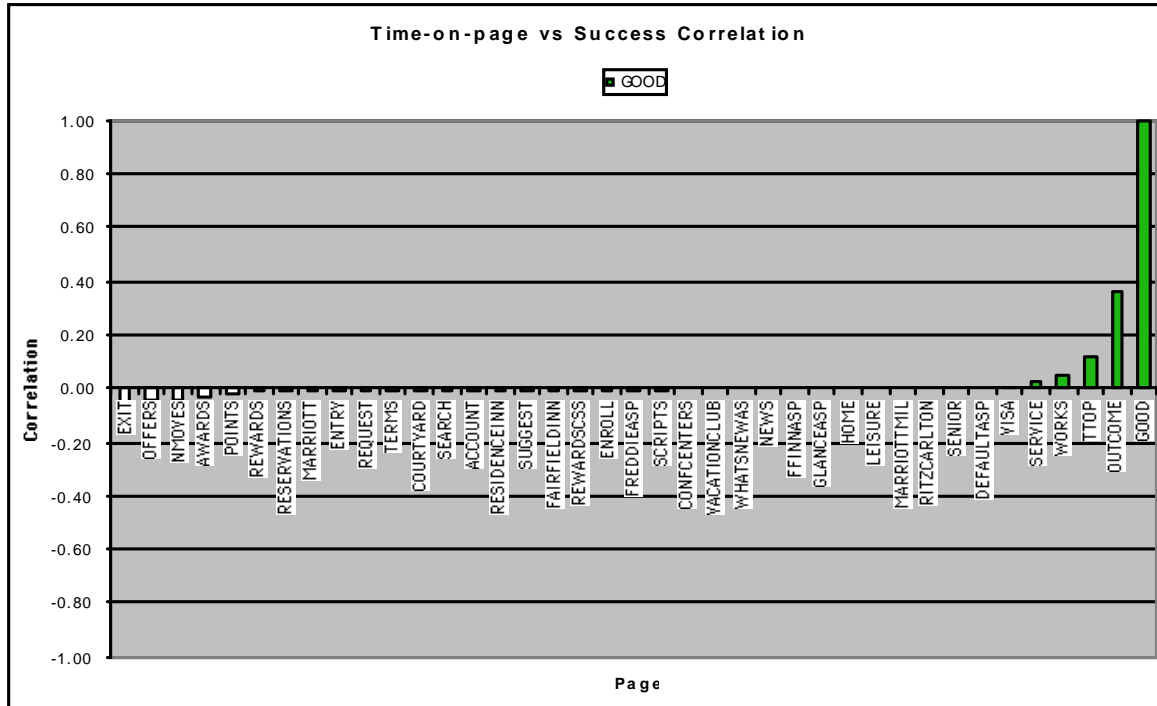
This analysis generated two items of interest:

- The Works page was the only one with a substantial positive correlation with success, probably because it answers all the basic questions succinctly, and as noted elsewhere, enrollment is not a tough sell.
- There were no substantial negative correlations with success, which is unusual in our experience. This both reinforces, the easy sell aspect of enrollment, and indicates there are no pages working against encouraging enrollment.
 - ⇒ The first conclusion above strongly suggests a small, direct, simplified pitch will lead to an increase in Enrollments through the site. It also suggests that simply putting a prominent link to the Enrollment page on the Works page will lead to increased enrollments.

Correlation of Hit Count with Success



Correlation of Time-on-page with Success



MAP Visualizations

In this section, we use the MAP technology to discover and analyze patterns beyond the reach of traditional statistics.

Exploring Success

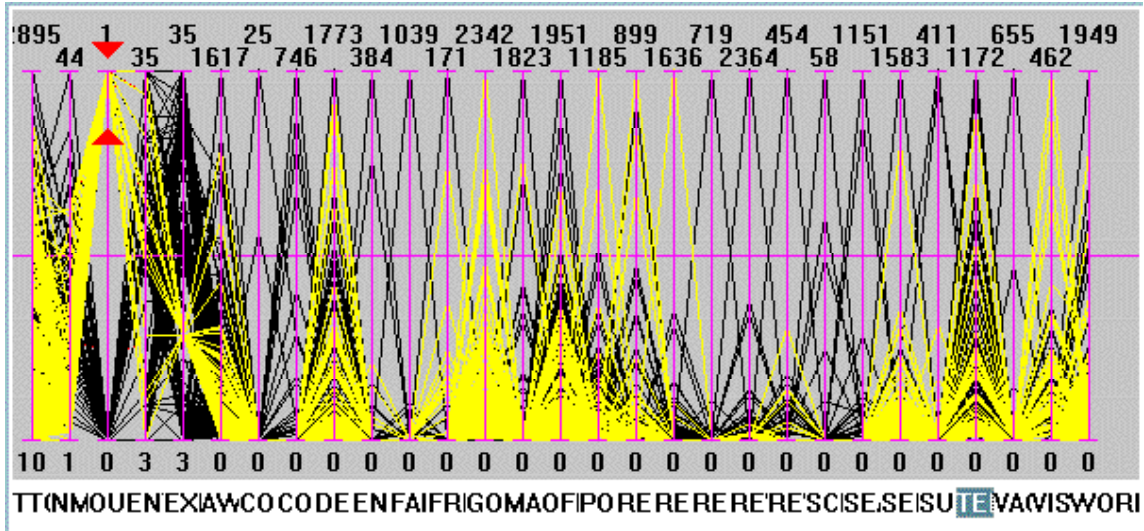
The graph below shows the Time on Page for each sampled successful session. The lines shown in yellow are those sessions for which three or fewer pages were viewed.

This analysis generated several interesting items:

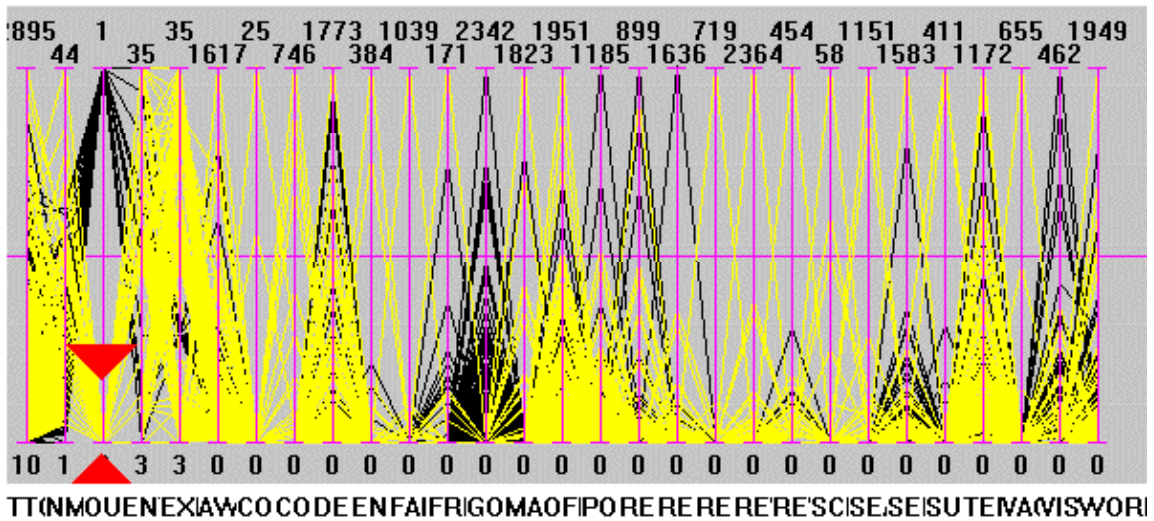
- For successful Enrollments, 47% occur after visiting only three pages (the corresponding percentage requires 15 pages for non-enrollees) and 87% after only seven pages, indicating that each prospect examines relatively little content before enrolling.

This suggests that prospects are quite motivated to enroll, with little resistance, when they enter the site.

Success



Failures



Exploring Failure and Other Tasks

This graph focuses time per page for non-Enrollments. The pattern that proved interesting was the roughly similar densities of the Points, Terms, and Awards pages (boxed).

Exploring the graph below suggests strongly that users come to the site with one of a specific set of tasks in mind:

- Enroll, if some simple objection can be overcome
- Check current offers for any of interest
- Check the expected benefits (points) for a specific trip, presumably to make the decision to stay at a participating facility.
 - ⇒ This last task from the analysis suggests that adding a capability focused on providing easy access to the benefits and options for a specific trip would generate interest — and bookings.

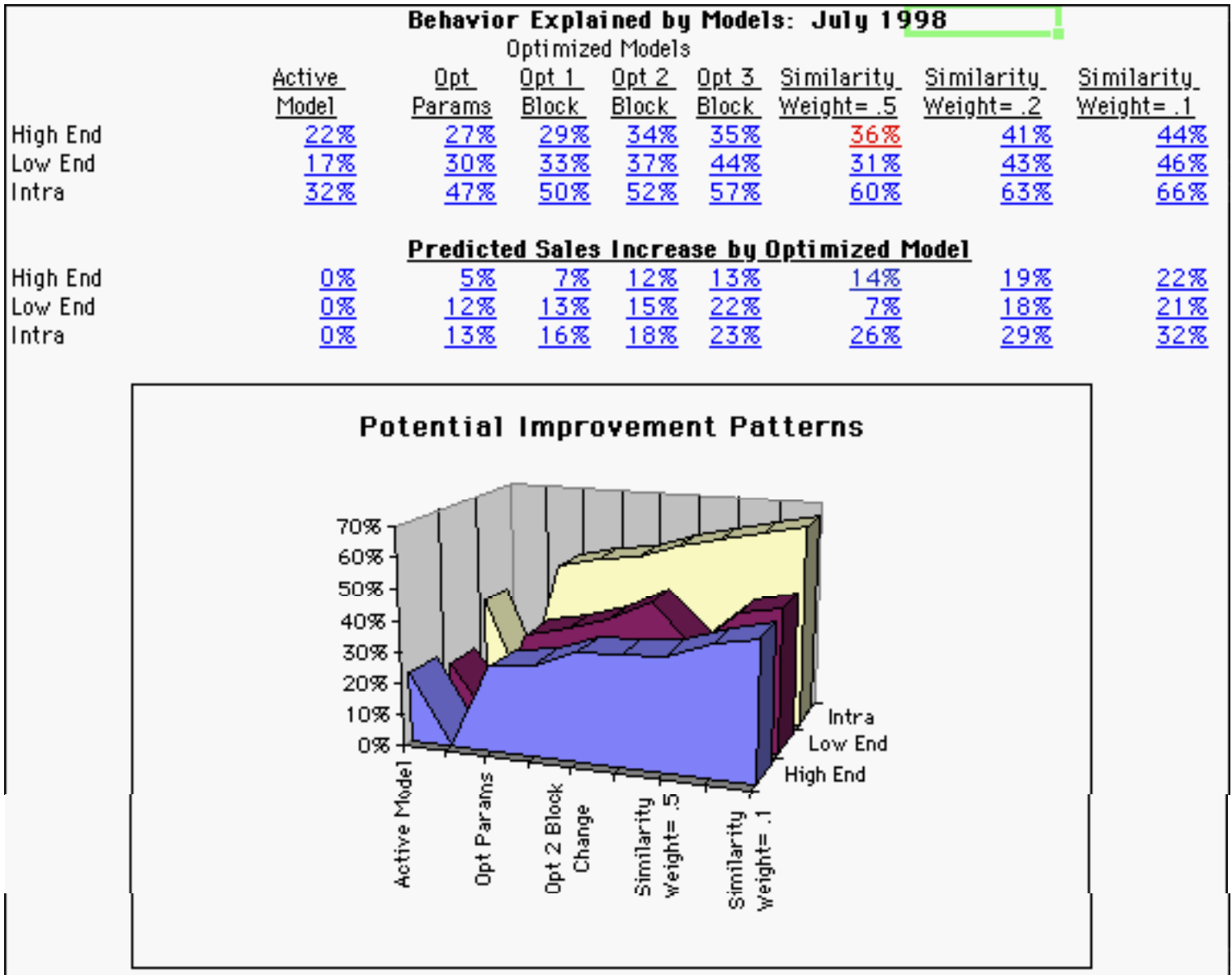
Specifically proposed would be the following capabilities:

- Provide a very specialized search/calculation, where a user can enter a destination city (metro area), a date and number of days to stay
- The output would consist of the number of points to be gained at each of the local facilities, including — and indicating — any special offers in effect, and a list of other services that could generate additional points, if used.
- The above would be formatted to allow one-click access directly to participating facility reservations, with appropriate fields already filled in.

The net effect would be to increase bookings, and add major value and competitive distinction to the client's rewards site.

3View Analysis

In this case, only the Decision Process analysis is shown, and it indicates that relatively little can be gained by improving the Process portion of the model, which is inconsistent with the high fraction of behavior explained, both in this analysis and in the DynModel portion of the analysis.



Conclusions

Many of the analysis techniques result from the same conclusions and recommendations. While redundant, it is very reassuring, so we generally leave the redundancy in the final reports. In fact, if only one technique supports a given conclusion, it is generally suspect.