

# PREDICTIVE DYNAMIX INC

*TURNING BUSINESS EXPERIENCE INTO BETTER DECISIONS*

## **Geospatial Data Mining for Market Intelligence**

**By Paul Duke, Predictive Dynamix, Inc.**

*Copyright © 2000-2001. All rights reserved.*

### **Overview**

Today, there is a huge amount of information readily available describing both consumer & business demographics related to practically every point on the map. In addition to this syndicated market data, most companies also have their own vast stores of strategic, market research, and operational data. There is so much data that information overload is virtually assured. More information is only better if you can effectively analyze it and assimilate it into actionable decisions.

Amid this information are the keys to making mission critical marketing decisions that cross industry boundaries, such as:

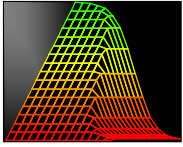
- Real Estate: Based on candidate locations, how should I allocate my capital budget between building new sites & rehabbing existing sites?
- Retail: How should I adjust inventory, merchandising, & pricing based on location in order to meet volume & revenue targets?
- Banking: Which of my savings account customers are the best candidates for brokerage accounts and what are the most effective incentives to attract them?
- Telecom: Which of my existing customers are most at risk to move to a competing carrier & which non-customers are the best prospects to switch to my service?
- Political Campaigning: Based on the latest polling research, what are the most significant profiles of swing voters and where are the most pivotal distributions?

Geospatial data mining describes the combination of two key market intelligence software tools: Geographical Information Systems (GIS) and Data Mining Systems. GIS and data mining are naturally synergistic technologies that can be synthesized to produce powerful market insight from a sea of disparate data.

### ***GIS: Putting Your Customers on the Map***

GIS technology has evolved far beyond the days of mainframe-plotted mapping systems. Today's business mapping software is very powerful and easy to use. This software allows you to integrate a virtually unlimited number of layers of spatial and relational databases.

Geographic map objects take the form of regions (zip codes, block groups, land use, etc.), lines (roads, transit routes, natural barriers, etc), and points (competitor site locations,



# PREDICTIVE DYNAMIX INC

*TURNING BUSINESS EXPERIENCE INTO BETTER DECISIONS*

individual customer locations, traffic counts, etc.). Attached to each map object can be a very detailed set of data describing it, such as, consumer & business demographics, site survey information, and historical customer purchasing patterns.

GIS provides operators that allow you to link tables based on proximity as well as join related databases key fields. With this capability you can address questions such as:

- What are the business population, residential population, and average annual income within 1 mile of a proposed site?
- Where do my best customers live?
- What competitive choices do my customers have in a given geographical area?

GIS can directly answer these and other similar questions. Demographic reporting is commonly used to generate a detailed characterization of potential store locations, broader marketing regions, or individual customers. Unfortunately, this detailed information, while qualitatively rich, often does not directly yield a decisive score that can be acted upon.

Luckily, this is the *raison d'être* for data mining technology.

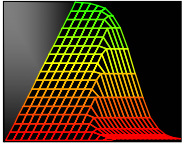
## ***Data Mining: Turning Business Experience into Better Decisions***

Data mining technology has evolved from the fields of statistics, artificial intelligence, and signal processing. Some of the underlying technologies include neural networks, clustering, genetic algorithms, fuzzy logic, decision tree, various regression methods, partial least squares, principal component analysis, and factor analysis.

The common goal for most of these techniques is to distill important trends and patterns from large amounts of historical cases in order to accurately forecast the outcome for future cases. For a given application, a “case” might represent a site, a customer, a block group, a transaction, a website hit, or some other unit of information to be evaluated.

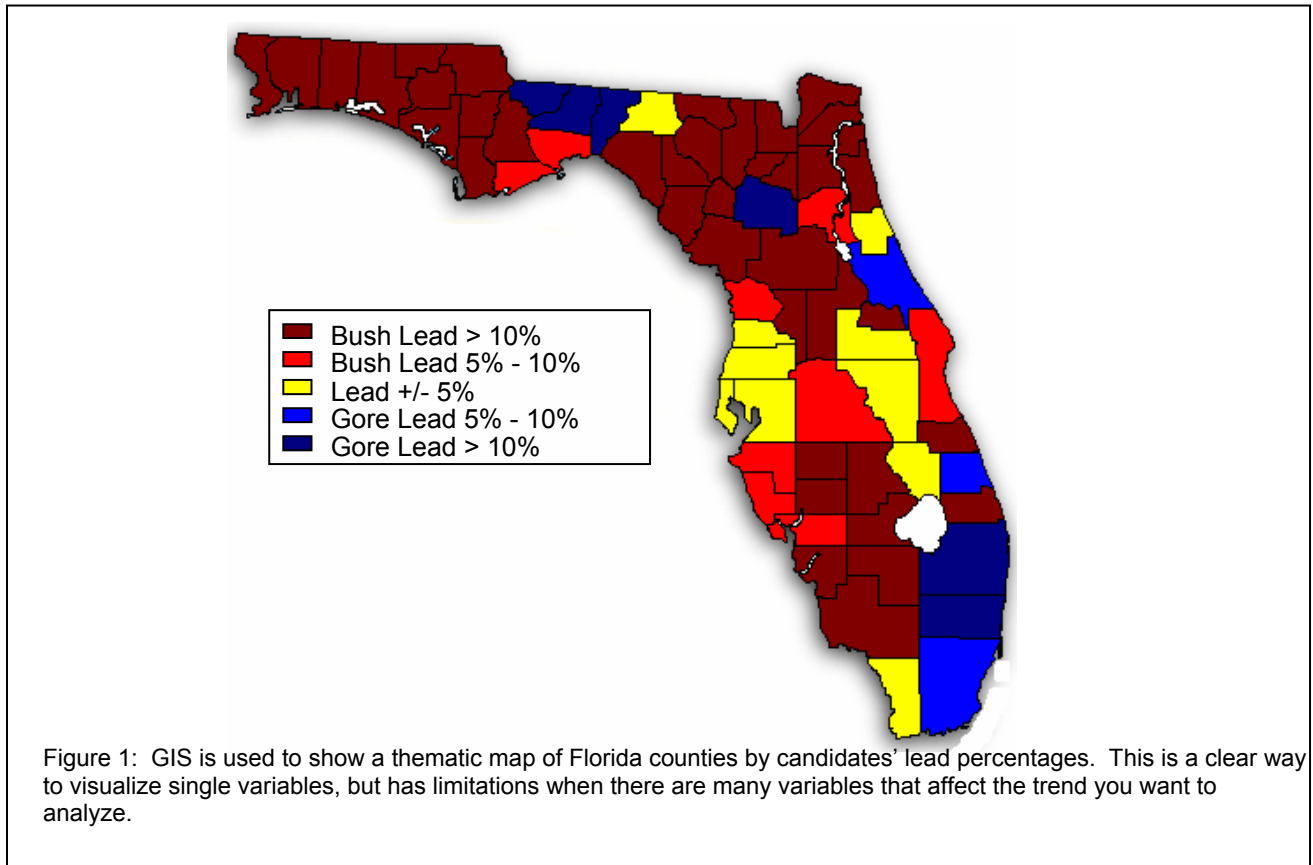
## **Geospatial Data Mining: Technology Synergy**

The strength of GIS is in providing a rich data infrastructure for combining disparate data in meaningful ways by using spatial proximity. Also, through the use of thematic map coloring, geographic visualization of individual variables can be very effective for identifying competitive hot zones, merchandising opportunities, etc.



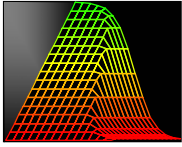
# PREDICTIVE DYNAMIX INC

*TURNING BUSINESS EXPERIENCE INTO BETTER DECISIONS*



The next logical step to take GIS analysis beyond demographic reporting to true market intelligence is to incorporate the ability to analyze and condense a large number of variables into a single forecast or score. This is the strength of predictive data mining technology and the reason why there is such a hand-in-glove fit between GIS & data mining.

Depending upon the specific application, GIS can combine historical customer or retail store sales data with syndicated demographic, business, traffic, and market research data. This dataset is then ideal for building predictive models to score new locations or customers for sales potential, cross-selling, targeted marketing, customer churn, and other similar applications.



# PREDICTIVE DYNAMIX INC

*TURNING BUSINESS EXPERIENCE INTO BETTER DECISIONS*

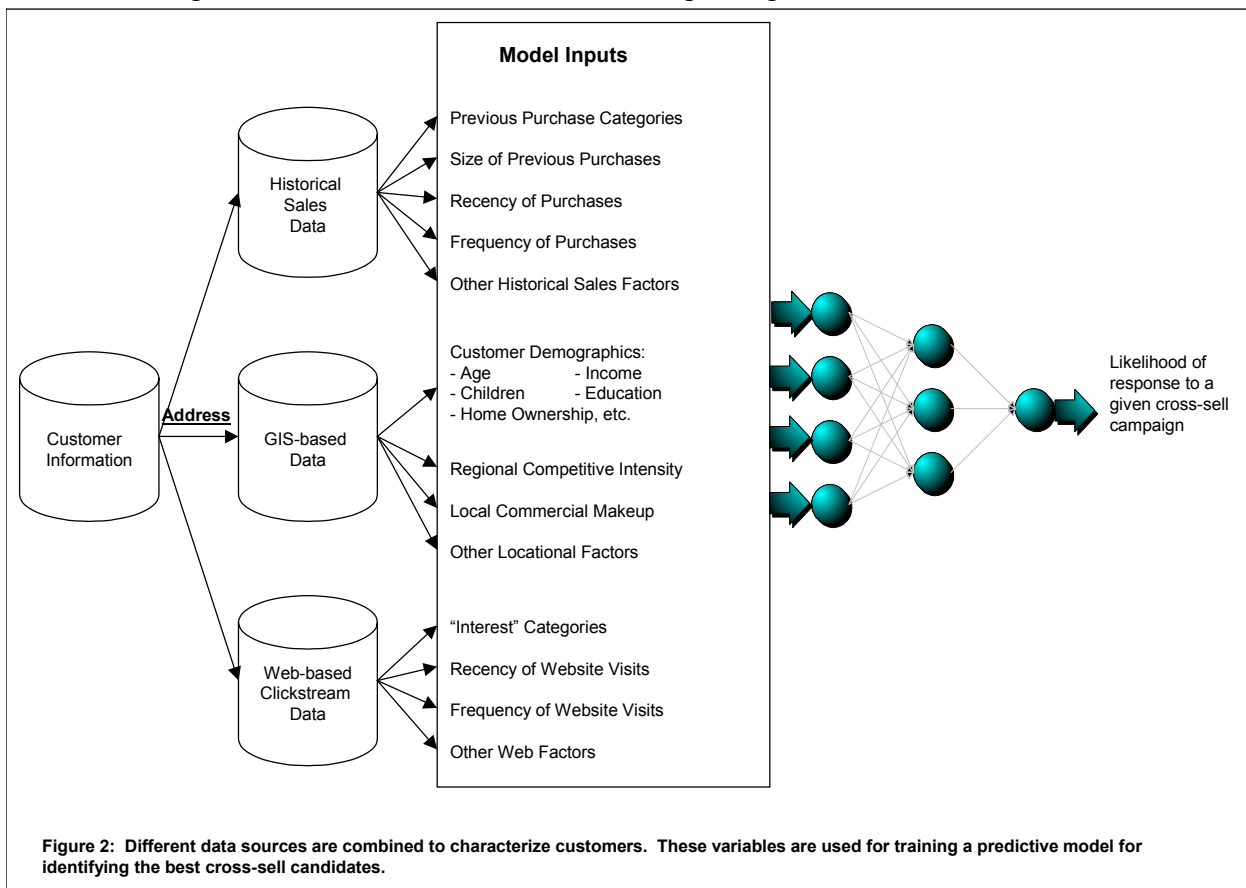
## ***Customer Relationship Management: Knowing Your Target Audience***

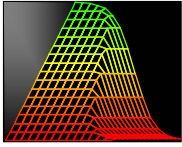
Often the goals of customer relationship management (CRM) systems read like the holy grail of marketing. Some CRM application examples include:

- Identify the most profitable new customers for direct mail marketing
- Identify the most profitable existing customers for cross selling
- Reduce churn of existing customer-base

With the prevalence of the Internet, eCRM systems are integrating click-stream data with historical sales & demographic data to address real-time, on-line marketing issues such as personalized content, intelligent cross selling, price management, and banner ad management.

At the heart of these applications are a variety of data sources, often linked geographically. The underlying intelligence comes from predictive models that can accurately forecast future customer response based on historical customer response patterns.





# PREDICTIVE DYNAMIX INC

*TURNING BUSINESS EXPERIENCE INTO BETTER DECISIONS*

## ***Retail Site Selection: If You Build It, Will They Come?***

Conventional wisdom in the site selection field is that the three most important factors in site performance are: Location, Location, & Location. This is undoubtedly true. However, in order to build an application around this logic you must be able to quantify, in detail, the factors that make a location good or poor.

Characterizing a location is where GIS technology shines. By incorporating block group demographic data, traffic data, business demographics, land use information, and other syndicated data, you can readily obtain information that will take you a long way in describing a location.

It is also critical to have detailed site information related to both your stores & your competitors' stores. This site data is used to characterize factors such as saturation & competitive intensity in a given area. Having detailed site data also allows you to characterize brand strength and other important operational and facility factors.

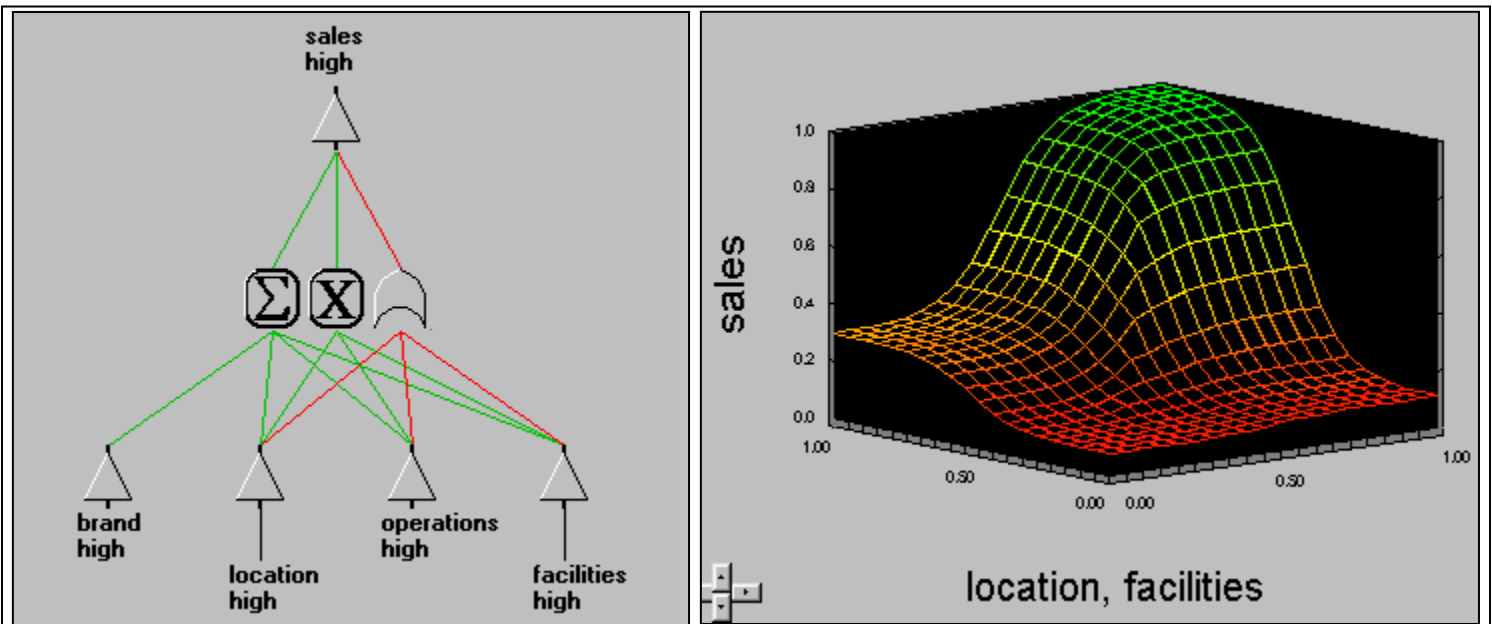
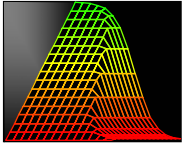


Figure 3: Using GIS is critical to characterizing a site's location which is invariably the most important factor in scoring a site. Location factors would include traffic, population, competition, etc. The decision surface shows how a model might generate a site score as a function of location & facilities.



# PREDICTIVE DYNAMIX INC

*TURNING BUSINESS EXPERIENCE INTO BETTER DECISIONS*

## Conclusion

This article has described how two business intelligence technologies, both important in their own right, have a unique synergy. GIS provides a spatial data infrastructure and the capability to generate insightful information that can be achieved in no other way. Data mining then provides the means to couple that information with operational data to produce actionable business decisions.

The two applications described above (CRM & Site Selection) show that this approach is directly relevant to some of the most important strategic & tactical business decisions faced by corporations today. These key tools facilitate applications easily valuing in the billions of dollars on a global basis.

Beyond marketing, there are many engineering, environmental, and other applications that clearly benefit from geospatial data mining. In many respects the science & technical communities are significantly ahead of the business community because they were the pioneers & early adopters of these technologies. However, with the advent of applications like CRM, value chain management, sales force automation, and others, the business community is quickly gaining ground.

Certainly many adept businesses are already applying approaches similar to those described. Unfortunately, the current processes tend to be very disjointed, often requiring the manual movement of data between various software applications. This tends to limit the use of the technology to those who have skills in database management, GIS, data analysis & predictive modeling, programming, and, most importantly, the specific business application. Given this, look for more tools in the near future that will integrate the necessary technologies in a manner that will streamline the process. This will allow companies to focus on the business opportunities instead of the barriers to implementation.

## About the Author

***Paul Duke has been developing commercial predictive modeling applications since the 1980's. His significant industry experience includes applications in retail network optimization, sales forecasting, pricing, securities trading, analytical CRM, homeland security, and diverse engineering & science areas.***

*This article was published in PC AI (March/April 2001) & The Data Administration Newsletter/TDAN.com (April 2001).*