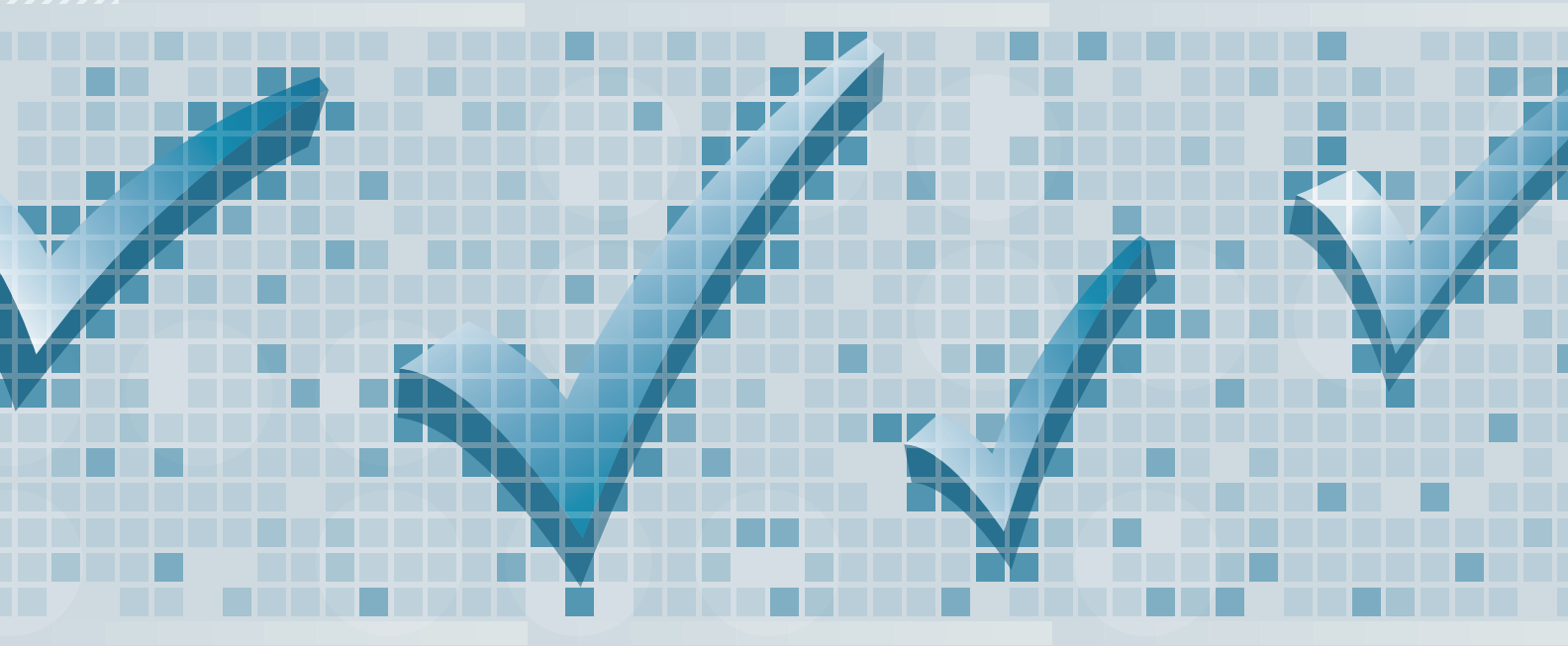
TDWI RESEARCH

TDWI CHECKLIST REPORT

SEVEN USE CASES FOR GEOSPATIAL ANALYTICS

By Fern Halper



Sponsored by

alteryx

Information Builders

tableau SOFTWARE

tdwi.org

tdwi

DECEMBER 2013

TDWI CHECKLIST REPORT

SEVEN USE CASES FOR GEOSPATIAL ANALYTICS

By Fern Halper

TABLE OF CONTENTS

555 S Renton Village Place, Ste. 700
Renton, WA 98057-3295

T 425.277.9126
F 425.687.2842
E info@tdwi.org

tdwi.org

FOREWORD

Increasingly, companies are looking to a variety of data types and new forms of analysis in order to remain competitive. Forward-looking companies are developing analytics ecosystems that make use of disparate kinds of data, including text data, social media data, machine data, and more. *Geospatial data*, sometimes referred to as location data or simply spatial data, is emerging as an important source of information both in traditional and in big data analytics.

Geospatial data and geographic information systems (GIS) software are being integrated with other analytics products to enable analytics that utilize location and geographic information. Such analytics are also moving past mapping to more sophisticated use cases such as advanced visualization and predictive analytics. According to a recent TDWI survey about analytics, for instance, the number of respondents who plan to use geospatial analytics will double between 2013 and 2016. Today, users want to better understand the value and use cases for this technology.

What is geospatial data and geospatial analytics? When people refer to geospatial data, they are often describing address-related data (a specific address, point of interest, ZIP code, and so on). This data can be matched to a specific latitude and longitude using a process known as geocoding. Geocodes for addresses and points of interest can also be integrated with other data sources to enhance analysis in dashboards, visualizations, and more advanced modeling. In addition, geospatial data sources include:

- **Global positioning system (GPS) data.** GPS data is gathered through satellite and ground-based radio navigation systems. A GPS-enabled smartphone can provide the location of a person, for instance.
- **Remote sensing data.** Remote sensing involves specialized instruments that capture data that can be converted into digital form. Satellites, scanners, and radar systems are all examples of these instruments.

Geospatial information can be extremely helpful in a variety of analytics ranging from marketing to operations management. This checklist introduces readers to the range of use cases where geospatial analytics is being used today to support analysis.

NUMBER ONE

TO REFINE SALES AND MARKETING.

A popular use case for geospatial analytics in marketing is market segmentation, where the goal is to divide customers into groups with common (unique) characteristics. These customers may have demographic (i.e., gender, income) and/or lifestyle data and even behavioral (i.e., buying patterns) features in common. The goal is to use these segments to help improve promotions and retention, and to obtain new customers.

For example, imagine a sporting goods company that uses demographic data such as age, sex, and income, together with purchase history, to develop its segments. By analyzing this data, the company has identified a target segment it calls "high-end soccer mom." This segment consists of customers who are female, aged 30 to 45, and buy high-end sporting gear for their young children and for themselves.

Adding geospatial data to the mix can help the company maximize its promotional activity and target new customers. The company might enrich its data set using geocoded data from ZIP codes associated with its current customers (e.g., from loyalty card information) to help visualize where customers in this segment live. Maps combining this data together with store locations help the firm visualize the distribution of its target segment around its stores to see whether these women are clustered geographically. If so, business analysts might create a new derived variable called distance, which is the distance between ZIP codes and various store locations.

The company can then look for other women (using third-party data) who meet the segment's profile and reside in clusters within a 15-mile radius of various stores. The retailer can then determine which stores would benefit from a targeted promotion to both current customers and prospects.

Taking this a step further, the company could also use location data from the GPS on the target segment's smartphones to push offers, perhaps on weekends, or even while the customers are in the store.

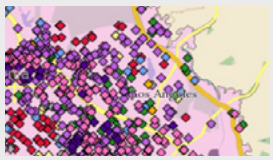


Figure 1

Click to enlarge

NUMBER TWO

TO UPGRADE ASSET MANAGEMENT.

Assets come in all shapes and sizes. Physical asset management refers to the systematic monitoring of equipment, buildings, and other physical property. A popular use case for geospatial analytics is to help companies manage assets. For example, a communications company might be interested in using geospatial data to help manage its network assets in order to deal more effectively with power outages in the event of a storm. These companies want to understand where network outages are occurring so they can determine which sites might be most vulnerable in the event of, say, a hurricane. The provider also needs to understand where its service assets (such as repair crews) are deployed to determine how to position them when maintenance is needed. Here, geospatial analytics can provide:

- **Network overview.** Geospatial analytics can provide insight into where the actual outages occur. It can determine where the biggest impacts are based on customer location and customer density as affected by the outages. Those areas with larger customer impact can be addressed first.
- **Service crew dispatch.** Geospatial analytics can help a provider evaluate which crews make sense to dispatch where based on the current location of the crews and the problems.

Taking this a step further, the communications company can also use more advanced analytics employing location data to determine where a failure might occur. This could involve using historical data (such as weather data, previous asset performance, number affected, time to repair, and coded geographic land forms) as input for a model. Asset managers can look at the characteristics of previous asset issues to predict future asset problems. For instance, the analyst could determine that network problems occur in low-lying areas with a limestone foundation when storms occur at high tide. In this way, managers can better plan where to position service crews in advance of a storm in areas that meet these conditions and decrease the time needed to begin repairs.

NUMBER THREE

TO AUGMENT SITUATIONAL AWARENESS AND INTELLIGENCE.

Geospatial data has been used by law enforcement agencies to predict and prevent crime in several ways:

- **Mapping.** Police or other agencies can utilize data from multiple sources and visualize it on layered maps. This data might include incident reports and crime types (i.e., robbery, homicide, narcotics, assault), date and time of the crime, as well as other data about the area from demographic sources (e.g., census data about education, population density, and housing). Geospatial data such as the location of an incident as well as other points of interest in the area (vacant lots, bars, restaurants) can be used to help visualize where high rates of crime are occurring. This can reveal patterns. For example, the police department might want to see where the highest rates of drug-related crimes occur in their city. Layering that specific type of crime on a map that also includes population density and when incidents occurred might indicate that drugs are being sold right after school within a two-block radius of certain fast food restaurants where teenagers congregate after classes are over.
- **Alerts.** Geospatial analysis can also be used for issuing alerts. For instance, an ankle bracelet with a GPS tracking device can be used to notify law enforcement if a person crosses designated boundaries.
- **Predictions.** Some police departments are using data from multiple sources, including geospatial data, as input to a predictive model. For instance, the target outcome of interest might be a certain kind of robbery. Predictors might include historical data consisting of time of event, location of event, geographic features (such as parking lots, empty lots, and parks), or population density. Data with known outcomes can be fed into a predictive model (such as a decision tree) to predict the probability of a robbery occurring in another area with similar characteristics.

Figure 2

Click to enlarge

NUMBER FOUR

TO HONE RISK ANALYSIS.

The idea behind risk analysis is to determine how often certain events might occur and what their impact could be in order to mitigate exposure. Financial institutions assess risk associated with certain investments. Insurance companies assess risk to set premiums. Insurance companies in the property and casualty space, for instance, model risk in order to spread that risk while maintaining reasonable premiums that ensure solid financial returns. A good example of this are the risks associated with disaster claims from floods or hurricanes. Geospatial data can be very useful here.

On a simple level, an insurance company can use geospatial analytics to see if it is overexposed in a certain area—say, where the risk of flooding from storm surge is high. The company can map its policyholders and overlay that map with hazard maps of historical flooding to see potential impact. This kind of analysis can be quite detailed. For instance, a given geography might have sub-geographies that may be at lower risk than the larger area.

Risk models can be made more sophisticated by incorporating location-based information to enrich a data set that is already used for risk modeling. For instance, insurance companies might use predictive models to calculate the loss for a group of policies related to a possible weather event. Location-enriched data might include location features and characteristics such as soil type, proximity to rivers, home characteristics, and weather data such as rainfall amounts, storm intensity, or historical river floods. This location-based information can be used as part of the model to predict where the probability of payout will be high. The insurance company can then adjust premiums accordingly.

On the flip side of premium pricing, geospatial data can also help insurance companies become more proactive in claims management. For example, an insurance company can use a geospatially enhanced map to determine where a large percentage of its claims are coming from after an event and then send adjusters and other personnel to those areas first. This helps to make claims management more effective and efficient.

Figure 3

Click to enlarge

NUMBER FIVE

TO ENHANCE TRANSPORTATION AND LOGISTICS PLANNING.

Transportation charges are often a large part of logistics costs. Fuel, maintenance, and driver time add to the cost. Transportation companies (often third-party providers) typically build complex models using linear programming techniques to try to maximize fleet utilization for optimum efficiency. Factors considered might include historical data of routes taken, trucks used, orders loaded on each truck, shift times, and delivery data. These companies often perform alternate scenario modeling to determine which routes make the most sense.

Currently, this software is being combined with GPS truck tracking software and data to monitor the actual versus planned time for a delivery. For example, some algorithms might run on distance from point A to B. The shortest distance might be used. However, the shortest distance might not mean the shortest time, especially at certain times of the day. Traffic, construction, and other delays might occur. Using past data of actual time to get from point A to point B, based on GPS data, logistics companies can re-plan and improve their routes. Some even do this in real time.

Geospatial analytics can be used in other scenarios. For example, a service company planning daily routes for its service technicians to arrive on time for specific customer appointments also benefits from geospatial data. Accurate street data and customer locations can be used to compute driving times for maximum efficiency. In addition, these companies are installing GPS devices in their trucks to help infield technicians plan routes in case of a problem. These devices can also transmit data back to the home base for monitoring service activity (e.g., tracking the length of a stop, the route taken, and driving behavior) to help optimize operations and meet response time targets.

Figure 4

Click to enlarge

NUMBER SIX

TO SHARPEN STRATEGIC LOCATION DETERMINATION.

Companies face numerous considerations when they are deciding where to locate a new store, restaurant, healthcare facility, or other commercial property. This often becomes a market optimization exercise where the planners need to understand the interplay of various factors that affect the choice of a certain location. Geospatial analysis can help to enrich this exercise. For instance, restaurant chains can use geospatial analysis to determine where a new restaurant should be located, including:

- **Target market.** Analytics applied to demographic data can determine where a specific target group is located. This data can be visualized using geocoded information to analyze potential target customer densities.
- **Competitors in an area.** This involves analyzing other restaurants in the chain to determine opportunities for a new restaurant. Visualization together with analytics can help determine where chain restaurants are in relation to the new location as well as where competitor restaurants are located. This distance/time measure could feed models that estimate cannibalization of the existing restaurant customer base.
- **Operational considerations.** These include understanding the costs associated with supplying new locations as well as other factors such as availability of ample parking in the area.
- **Change.** This refers to using geographic information to look at new plans for housing, commercial buildings, or roads that might impact the number of potential customers coming to a location.

Geospatial analytics can also help in situating assets other than property. Communications companies are using geospatial data to determine where to locate cell phone towers. A company might send drivers out with devices used to measure signal strength from particular towers and the range of that signal. This signal strength is plotted on a map to see where coverage is lacking. Based on population density and other factors such as roads and where customers are located, the communications company can decide where to install new cell towers.

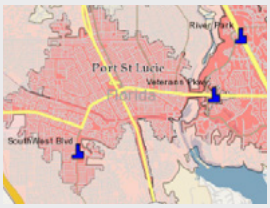


Figure 5

Click to enlarge

NUMBER SEVEN

TO IMPROVE FRAUD DETECTION AND PREVENTION.

Fraud is a multi-billion-dollar problem in the insurance industry, affecting premium costs for consumers and profits for the insurance industry. Insurance companies have begun to deploy sophisticated techniques such as predictive analytics in claims analysis. For example, insurance companies are using historical claims data with known fraud outcomes to predict the probability that new claims are fraudulent. Flagged claims are sent to a special investigation unit for follow-up.

Geospatial data can help enterprises detect and prevent fraud by providing location information about where fraud might occur. Automobile claims fraud is a good example. An automobile insurance company might be performing an analysis to identify repair shops that inflate repair estimates for automobile body work. Data might include the claimants' addresses, which can be geocoded, as well as the location of repair shops and the average claim estimate for a particular kind of problem. Using this geocoded data together with other claims information, the analyst can map areas where estimates are higher than the average for a bumper repair (for example). On top of this could be layered the claimants' home locations. The claims analyst might find that people are traveling a long distance to get to a certain repair shop. This shop might be providing estimates that are also above average. This suggests an inflated estimate, which indicates potential fraud.

Geospatial data can also be used in other industries to detect fraud:

- **Healthcare:** Analyzing claims forms to identify physicians who might be inflating claims or falsifying reports
- **Government agencies:** Studying where recipients are defrauding social programs such as the supplemental nutrition assistance (food stamps) program
- **Credit card companies:** Inspecting transaction data, including a combination of geographic location of a transaction, transaction amount, date/time, and merchandise category, might be used to detect fraud

Figure 6

Click to enlarge

ABOUT OUR SPONSORS

alteryx

www.alteryx.com

Alteryx provides indispensable analytic solutions for enterprise companies making critical decisions about how to expand and grow. Our product, Alteryx Strategic Analytics, is a desktop-to-cloud agile BI and analytics solution designed for data artisans and business leaders that brings together the market knowledge, location insight, and business intelligence today's organizations require. For more than a decade, Alteryx has enabled strategic planning executives to identify and seize market opportunities, outsmart their competitors, and drive more revenue. Customers such as Experian Marketing Services and McDonald's rely on Alteryx daily for their most important decisions. Headquartered in Irvine, California, and with offices in Boulder and Silicon Valley, Alteryx empowers 250+ customers and 200,000+ users worldwide. Get inspired today at www.alteryx.com or call 888.836.4274.

Information Builders

www.informationbuilders.com

Information Builders' "3i" strategy incorporates technology for business intelligence, integration, and data integrity. Although power user tooling is an important part of their approach to BI, superior scalability and an emphasis on InfoApps™ for non-technical users enables the WebFOCUS BI platform to support the largest BI deployments—some going outside the firewall, reaching millions of users—while mashing up GIS, search, and predictive analytics in easy-to-consume ways. Data quality and integration software, combined with top-rated customer service and satisfaction, ensure that clients turn data into business value.

tableau SOFTWARE

www.tableausoftware.com

Tableau Software helps people see and understand data. Tableau helps anyone quickly analyze, visualize, and share information. More than 15,000 customer accounts get rapid results with Tableau in the office and on the go. And tens of thousands of people use Tableau Public to share data in their blogs and websites. See how Tableau can help you by downloading the free trial at www.tableausoftware.com/trial.

ABOUT THE AUTHOR

Fern Halper, Ph.D., TDWI Research Director for Advanced Analytics, has more than 20 years of experience in data analysis, business analysis, and strategy development. She is a recognized expert in the field of analytics and was among the first analysts to write about text analytics. She has authored hundreds of reports and articles on data mining and information technology and is frequently quoted on the subject in the press. Prior to joining TDWI, Fern was a partner at Hurwitz & Associates, an industry analyst firm where she focused her research primarily on analytics and cloud computing. She also spent eight years at Bell Laboratories leading the development of innovative approaches and systems to analyze marketing and operational data for AT&T's marketing and business teams. Fern has taught at both Bentley College and Colgate University. She is the co-author of numerous "Dummies" publications, including *Cloud Computing for Dummies*, *Hybrid Cloud for Dummies*, and *Big Data for Dummies* (Wiley). Fern received her BA from Colgate University and her Ph.D. from Texas A&M University. You can reach her at fhalper@tdwi.org.

ABOUT THE TDWI CHECKLIST REPORT SERIES

TDWI Checklist Reports provide an overview of success factors for a specific project in business intelligence, data warehousing, or a related data management discipline. Companies may use this overview to get organized before beginning a project or to identify goals and areas of improvement for current projects.

ABOUT TDWI RESEARCH

TDWI Research provides research and advice for business intelligence and data warehousing professionals worldwide. TDWI Research focuses exclusively on BI/DW issues and teams up with industry thought leaders and practitioners to deliver both broad and deep understanding of the business and technical challenges surrounding the deployment and use of business intelligence and data warehousing solutions. TDWI Research offers in-depth research reports, commentary, inquiry services, and topical conferences as well as strategic planning services to user and vendor organizations.