Business intelligence and big data in hospitality and tourism: a systematic literature review

Article

Accepted Version

Tables


It is advisable to refer to the publisher’s version if you intend to cite from the work.

To link to this article DOI: http://dx.doi.org/10.1108/IJCHM-07-2017-0461

Publisher: Emerald

All outputs in CentAUR are protected by Intellectual Property Rights law, including copyright law. Copyright and IPR is retained by the creators or other copyright holders. Terms and conditions for use of this material are defined in the End User Agreement.

www.reading.ac.uk/centaur
Table 1: Major components and functional subdomains within the Business Intelligence umbrella

<table>
<thead>
<tr>
<th>Major Concept</th>
<th>Acronym</th>
<th>Short Explanation</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Decision Support System</td>
<td>DSS</td>
<td>Computer-based information system that supports decision making resulting in ranking, sorting, or choosing from among alternatives</td>
<td>Burstein and Holsapple, 2008; Sauter, 2011</td>
</tr>
<tr>
<td>Data Warehousing</td>
<td>DW</td>
<td>Central data repository system of integrated data from one or multiple sources that stores current and historical data in one single place and format</td>
<td>Kimball et al., 2008</td>
</tr>
<tr>
<td>Online analytical processing</td>
<td>OLAP</td>
<td>Provides multi-dimensional analytical queries encompassing data warehousing and reporting. Supports the operations of consolidation, drill-down, slicing and dicing</td>
<td>Kimball et al., 2008</td>
</tr>
<tr>
<td>Data Mining</td>
<td>DM</td>
<td>Discovers correlations and patterns in (usually large) data sets involving methods of machine learning, statistics and mathematical modelling</td>
<td>Larose, 2005; Rud, 2009</td>
</tr>
<tr>
<td>Business Intelligence</td>
<td>BI</td>
<td>Umbrella term comprising the domains of DW, OLAP and DM</td>
<td>Kimball and Ross, 2016</td>
</tr>
<tr>
<td>Descriptive Analytics</td>
<td>-</td>
<td>Uses data aggregation (e.g. sums, averages, percent, changes, etc.) and data mining to provide insight into the past to answer: “What has happened”?</td>
<td>Williams, 2016</td>
</tr>
<tr>
<td>Predictive Analytics</td>
<td>-</td>
<td>Uses statistical and DM models to forecast the future and answer: “What could happen?”</td>
<td>Đedić and Stanier, 2016</td>
</tr>
<tr>
<td>Prescriptive Analytics</td>
<td>-</td>
<td>Uses machine learning and computational modelling to advice on optimal outcomes and answers: “What should we do?”</td>
<td>Đedić and Stanier, 2016; Williams, 2016</td>
</tr>
<tr>
<td>Big Data</td>
<td>BD</td>
<td>Data sets that are so large or complex that traditional data processing application software is inadequate to deal with them. Includes challenges, as data extraction, storage, analytics, visualization, querying, updating and information privacy</td>
<td>Erl et al., 2015</td>
</tr>
<tr>
<td>Article (author &amp; title)</td>
<td>Research topic</td>
<td>Type of paper (conceptual/empirical)</td>
<td>Source(s) of data</td>
</tr>
<tr>
<td>-------------------------</td>
<td>----------------</td>
<td>-------------------------------------</td>
<td>------------------</td>
</tr>
<tr>
<td>Amadio and Procaccino (2016).</td>
<td>Text-based online review analysis using exploratory text mining techniques and visual analytics for SWOT analysis, applied to the hotel industry</td>
<td>Empirical</td>
<td>Online reviews from TripAdvisor</td>
</tr>
<tr>
<td>Arbelaitz et al. (2013).</td>
<td>Combined web usage and content mining to generate user navigation profiles and semantically enriched user interest profiles as input to website optimization and marketing</td>
<td>Empirical</td>
<td>Web page content and web server log files of Bidasoa Turismo website</td>
</tr>
<tr>
<td>Ashiabor et al. (2007).</td>
<td>Nested mixed logit models to estimate market share of automobile and commercial air transportation</td>
<td>Empirical</td>
<td>American Travel Survey</td>
</tr>
<tr>
<td>Authors and Year</td>
<td>Description</td>
<td>Data Source</td>
<td>Data Type</td>
</tr>
<tr>
<td>------------------</td>
<td>-------------</td>
<td>-------------</td>
<td>-----------</td>
</tr>
<tr>
<td>Chen and Tsai (2016).</td>
<td>Data mining framework based on rough set theory to improve location selection decisions: A case study of a restaurant chain</td>
<td>Data mining framework based on rough set theory (RST) to support location selection decisions</td>
<td>Empirical</td>
</tr>
<tr>
<td>Chiang (2015).</td>
<td>Applying data mining with a new model on customer relationship management systems: A case of airline industry in Taiwan</td>
<td>Mining high-value family travelers for CRM systems of online airlines and travel agencies to identify decision rules for discovering market segments</td>
<td>Empirical</td>
</tr>
<tr>
<td>Dursun and Caber (2016).</td>
<td>Using data mining techniques for profiling profitable hotel customers: An application of RFM analysis</td>
<td>Profiling hotel customers by recency, frequency and monetary (RFM) indicators</td>
<td>Empirical</td>
</tr>
<tr>
<td>Fuchs et al. (2013).</td>
<td>A knowledge destination framework for tourism sustainability: A business intelligence application from Sweden</td>
<td>A Destination Management Information system focusing on Online-Analytical Processing (OLAP) to measure proportion of tourists with smallest ecological footprint</td>
<td>Empirical</td>
</tr>
<tr>
<td>Fuchs et al. (2014).</td>
<td>Big data analytics for knowledge generation in tourism</td>
<td>BI-based knowledge infrastructure implemented at the Swedish mountain</td>
<td>Empirical</td>
</tr>
<tr>
<td>Study Title</td>
<td>Description</td>
<td>Methodology</td>
<td>Data Management</td>
</tr>
<tr>
<td>---------------------------------------------------------------------------</td>
<td>-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------</td>
<td>------------------------------------------------------------------------------------------------------------------------------------------------</td>
<td>-------------------------------------------------------------------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>destinations - A case from Sweden</td>
<td>destination, Are and examples of use by tourism managers</td>
<td>user-generated content)</td>
<td>automatically (periodically)</td>
</tr>
<tr>
<td>Holland et al. (2016). The role and impact of comparison websites on the consumer search process in the US and German airline markets</td>
<td>Examines how consumers search for airline tickets based on a comparative analysis of the US and German markets</td>
<td>Empirical</td>
<td>Structured and unstructured (periodically)</td>
</tr>
<tr>
<td>Höpken et al. (2015). Business intelligence for cross-process knowledge extraction at tourism destinations</td>
<td>A novel approach for BI-based cross-process knowledge extraction and decision support for tourism destinations</td>
<td>Empirical and conceptual</td>
<td>Structured and unstructured (periodically)</td>
</tr>
<tr>
<td>Kisilevich et al. (2013). A GIS-based decision support system for hotel room rate estimation and temporal price prediction: The hotel brokers’ context</td>
<td>A tool that assists travel intermediaries to acquire missing strategic information about hotels to leverage profitable deals. The GIS-based DSS</td>
<td>Empirical</td>
<td>Structured and unstructured (periodically) and manually (one-time)</td>
</tr>
</tbody>
</table>
estimates room rates using hotel and location characteristics hotel facilities, room amenities, hotel categories. Dynamic: room prices for one night customers received during their search, date of search, date of order additive regression with isotonic regression; Locally Weighted Learning with Linear Regression; LibSVM nu-SVR; Multilayer Perceptron (ANN) Graphs to visualize price estimation results

<table>
<thead>
<tr>
<th>Köseoğlu et al. (2016). Competitive intelligence practices in hotels</th>
<th>Assessment of awareness and knowledge about competitive intelligence efforts in the hotel industry</th>
<th>Empirical</th>
<th>23 hoteliers’ knowledge and awareness about competitive intelligence</th>
<th>Unstructured</th>
<th>In-depth interview</th>
<th>N/A</th>
<th>N/A</th>
</tr>
</thead>
<tbody>
<tr>
<td>Kwok and Yu (2016). Taxonomy of Facebook messages in business-to-consumer communications: What really works?</td>
<td>Combines machine learning and human intelligence to analyze Facebook messages initiated by hospitality companies</td>
<td>Empirical</td>
<td>2,654 Facebook messages initiated by 26 hospitality companies</td>
<td>Unstructured and structured</td>
<td>Automatically/ manually</td>
<td>Machine learning (support vector machines)</td>
<td>Taxonomy of Facebook message types</td>
</tr>
<tr>
<td>Li et al. (2015). Identifying emerging hotel preferences using Emerging Pattern Mining technique</td>
<td>Identification of emergent hotel features by extracting frequent keywords from online reviews</td>
<td>Empirical</td>
<td>118,000 online reviews from TripAdvisor</td>
<td>Unstructured</td>
<td>Automatically (one-time)</td>
<td>Unsupervised feature extraction by frequent keywords, emerging pattern mining (EPM)</td>
<td>N/A</td>
</tr>
<tr>
<td>Reference</td>
<td>Title</td>
<td>Methodology</td>
<td>Data Type</td>
<td>Data Source</td>
<td>Pre-processing</td>
<td>Analysis</td>
<td></td>
</tr>
<tr>
<td>-----------</td>
<td>-------</td>
<td>-------------</td>
<td>-----------</td>
<td>-------------</td>
<td>----------------</td>
<td>----------</td>
<td></td>
</tr>
<tr>
<td>Marine-Roig and Anton Clavé (2015).</td>
<td>Tourism analytics with massive user-generated content: A case study of Barcelona</td>
<td>Empirical</td>
<td>Heterogeneous</td>
<td>Unstructured</td>
<td>Data was extracted automatically through Offline Explorer Enterprise</td>
<td>Tables created through word count</td>
<td></td>
</tr>
<tr>
<td>Pope et al. (2009).</td>
<td>Conceptual framework for collecting online airline pricing data: Challenges, opportunities, and preliminary results</td>
<td>Conceptual</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td></td>
</tr>
<tr>
<td>Ritchie and Ritchie (2002).</td>
<td>A framework for an industry supported destination marketing information system</td>
<td>Empirical</td>
<td>Industry stakeholders’ knowledge needs and current use of research &amp; intelligence (Inter-)National Travel Survey</td>
<td>Semi-structured interview</td>
<td>N/A</td>
<td>N/A</td>
<td></td>
</tr>
<tr>
<td>Rossetti et al. (2016).</td>
<td>Analyzing user reviews in tourism with topic models</td>
<td>Empirical</td>
<td>Yelp Data set Challenge; TripAdvisor Dataset</td>
<td>Structured and unstructured</td>
<td>Yelp is existing data set; TripAdvisor automatically collected by crawler</td>
<td>K-Nearest Neighbor User Based (KNN-UB), K-Nearest Neighbor Item Based</td>
<td>Illustrative examples for selected Topics related to multi-criteria dimensions</td>
</tr>
<tr>
<td>Authors</td>
<td>Title</td>
<td>Methodology</td>
<td>Dataset</td>
<td>Analysis Approach</td>
<td>Model/Algorithm</td>
<td></td>
<td></td>
</tr>
<tr>
<td>---------</td>
<td>-------</td>
<td>-------------</td>
<td>---------</td>
<td>-------------------</td>
<td>-----------------</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sánchez-Franco et al. (2016)</td>
<td>Online Customer Service Reviews in Urban Hotels: A Data Mining Approach</td>
<td>Extraction of features from hotel reviews and analysis of their relationship with guests’ hotel rating in the online travel agencies environment</td>
<td>Empirical</td>
<td>19,318 hotel reviews from 2014 to 2015 from booking.com</td>
<td>Structured and unstructured</td>
<td>Automatically (one-time)</td>
<td>Pathfinder network scaling, principal component analysis (PCA), linear mixed-effects regression</td>
</tr>
<tr>
<td>Snavely et al. (2008)</td>
<td>Modeling the world from internet photo collections</td>
<td>Presents algorithms and results as a step towards 3D modeling of the world’s well-photographed sites, cities, and landscapes from Internet imagery</td>
<td>Empirical</td>
<td>Flickr</td>
<td>Large sets of image data</td>
<td>Automatically downloaded from Flickr</td>
<td>Keypoint detection (SIFT keypoint detector) and matching (by approximate nearest neighbors (ANN) kd-tree); Structure for motion (to recover camera parameters); geo-registration (by digital elevation maps)</td>
</tr>
<tr>
<td>Solnet et al. (2016)</td>
<td>An untapped gold mine? Exploring the potential of market basket analysis to grow hotel revenue</td>
<td>Market Basket Analysis to identify and predict the purchasing behavior of customers based on their expenditure patterns in order to determine the most attractive additional products and services</td>
<td>Empirical</td>
<td>56,906 guest sales records from a luxury hotel group in Australia from 2009 to 2014</td>
<td>Structured</td>
<td>Automatically (one-time)</td>
<td>Multivariate logit model</td>
</tr>
<tr>
<td>Sun et al. (2016)</td>
<td>Chinese Customers’ Evaluation of Travel Website Quality: A Decision-Tree Analysis</td>
<td>Identification of critical attributes that influence quality levels of a customer’s travel</td>
<td>Empirical</td>
<td>Survey data from 25 individuals</td>
<td>Structured</td>
<td>Manually</td>
<td>Attention-interest-desire-action (AIDA) model, C4.5 decision tree</td>
</tr>
<tr>
<td>Reference</td>
<td>Title</td>
<td>Description</td>
<td>Methodology</td>
<td>N/A</td>
<td>Exploration</td>
<td>Analysis</td>
<td>Tools</td>
</tr>
<tr>
<td>-----------</td>
<td>-------</td>
<td>-------------</td>
<td>-------------</td>
<td>-----</td>
<td>-------------</td>
<td>---------</td>
<td>-------</td>
</tr>
<tr>
<td>Tseng and Won (2016). Integrating multiple recommendation schemes for designing sales force support system: A travel agency example</td>
<td>Proposes a design of sales force support (SFS) system with business intelligence methodologies</td>
<td>Empirical</td>
<td>N/A</td>
<td>N/A</td>
<td>Explorative data analysis (EDA) and data mining (e.g. sequential pattern discovery)</td>
<td>Dashboards, On-Line Analytical Processing (OLAP)</td>
<td></td>
</tr>
<tr>
<td>Wu et al. (2010). Data mining for hotel occupancy rate: An independent component analysis approach</td>
<td>Identification of major factors determining the hotel occupancy rate and incorporation of these factors to decompose hotel occupancy rates and examine the effect of each factor on the hotel occupancy rate</td>
<td>Empirical</td>
<td>Monthly hotel occupancy rate time series for each district of Hong Kong from January 1996 to May 2009</td>
<td>Structured</td>
<td>Manually</td>
<td>Independent component analysis (ICA)</td>
<td>N/A</td>
</tr>
<tr>
<td>Zhang and Huang (2015). Mining tourist motive for marketing development via twice-learning</td>
<td>Application of twice-learning framework to predict tourists’ travel motives from tourists’ external and internal features, useful for targeted marketing strategy development</td>
<td>Empirical</td>
<td>On-site surveys in Nanjing, China, from October to November 2012 with 121 responses</td>
<td>Structured</td>
<td>Manually</td>
<td>Twice-learning framework, neural networks, C4.5 decision tree, Naïve Bayes</td>
<td>N/A</td>
</tr>
<tr>
<td>Zhu et al. (2016). Get into the spirit of a location by mining user-generated travelogues</td>
<td>Location information extraction from user-generated travelogues, examining contents and structures of travelogues, as well as their interplay</td>
<td>Empirical</td>
<td>80,384 travelogues related to tourist destinations in the United States</td>
<td>Unstructured</td>
<td>Manually</td>
<td>Gazetteer-based location detection, semantic correlation detection by natural language</td>
<td>N/A</td>
</tr>
<tr>
<td>parsing techniques, location concept network by PLSA</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Table 3. Big Data works in Hospitality and Tourism (selected works; in “type of data and size” an asterisk indicates large quantities of data, > 100 000 records)

<table>
<thead>
<tr>
<th>Article (author and title)</th>
<th>Research topic</th>
<th>Type of paper (conceptual/ empirical)</th>
<th>Sources of data</th>
<th>Type of data and size</th>
<th>Data collection methods</th>
<th>Data analysis techniques</th>
<th>Data reporting and visualization</th>
</tr>
</thead>
<tbody>
<tr>
<td>Buhalis and Foerste (2015). SoCoMo marketing for travel and tourism: Empowering co-creation of value</td>
<td>Proposes social context mobile (SoCoMo) marketing as a new framework that enables marketers to increase value for all stakeholders at the destination</td>
<td>Conceptual</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>Carter (2016). Where are the enslaved?: TripAdvisor and the narrative landscapes of southern plantation museums</td>
<td>Explores what visitors learn about the history of the enslaved on two tours (Laura and Oak Valley) and how they participate in the narrative construction of the plantation</td>
<td>Empirical</td>
<td>TripAdvisor visitor reviews (Laura and the Oak Alley museums, USA)</td>
<td>Unstructured</td>
<td>Web (manual) scraping</td>
<td>Word frequency and words associations in reviews</td>
<td>Standard tables</td>
</tr>
<tr>
<td>Dolnicar and Ring (2014). Tourism marketing research: Past, present and future</td>
<td>Critical review of tourism marketing research</td>
<td>Literature review</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>Fuchs et al. (2014). Big data analytics for knowledge generation in tourism destinations - A case from Sweden</td>
<td>Presents a knowledge infrastructure implemented at the Swedish mountain tourism destination, Åre and examples of use by tourism managers</td>
<td>Empirical</td>
<td>Web search, booking and feedback data (e.g., survey-based, user-generated content)</td>
<td>Structured and unstructured</td>
<td>Data Warehouse (DW) including Facts and Dimensions Tables</td>
<td>On-Line Analytical Processing (OLAP); Support Vector Machines (SVM), Naïve Bayes (NB) and</td>
<td>Html-based web application</td>
</tr>
<tr>
<td>Reference</td>
<td>Title</td>
<td>Methodology</td>
<td>Data Source</td>
<td>Data Type</td>
<td>Natural Language Processing</td>
<td>Results</td>
<td></td>
</tr>
<tr>
<td>-----------</td>
<td>-------</td>
<td>-------------</td>
<td>-------------</td>
<td>-----------</td>
<td>----------------------------</td>
<td>---------</td>
<td></td>
</tr>
<tr>
<td>García-Pablos et al. (2016).</td>
<td>Automatic analysis of textual hotel reviews</td>
<td>Empirical</td>
<td>Online reviews from Zoover and HolidayCheck</td>
<td>Unstructured</td>
<td>Web crawler</td>
<td>Natural Language Processing: Named Entity Recognition, Sentiment Analysis and Opinion Mining</td>
<td>Standard tables</td>
</tr>
<tr>
<td>Gretzel et al. (2015).</td>
<td>Smart tourism: Foundations and developments</td>
<td>Conceptual</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>Gunter and Önder (2016).</td>
<td>Forecasting city arrivals with google analytics</td>
<td>Empirical</td>
<td>Google analytics variables collected on a monthly basis over the period August 2008-October 2014</td>
<td>Structured</td>
<td>Simple access to Google analytics</td>
<td>VAR model class: BVAR, FAVAR, BFAVAR.</td>
<td>Basic tables of descriptive statistics</td>
</tr>
<tr>
<td>He et al. (2016). Travel-package recommendations leveraging social influence of different relationship types</td>
<td>Develops a probabilistic topic model leveraging individual travel history and social influence of co travelers to capture personal interests and propose a recommendation method to utilize the proposed model</td>
<td>Empirical</td>
<td>Structured travel records on travel packages</td>
<td>Structured</td>
<td>Access to a private company database</td>
<td>Biggs sampling</td>
<td>Basic tables of descriptive statistics</td>
</tr>
<tr>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td>Jackson (2016). Prediction, explanation and big(ger) data: A middle way to measuring and modelling the perceived success of a volunteer tourism sustainability campaign based on 'nudging'</td>
<td>Uses ‘automatic linear modelling’ that can cope with big data and presents the results as visualizations</td>
<td>Empirical</td>
<td>Structured (responses from questionnaire)</td>
<td>Structured</td>
<td>Survey</td>
<td>Automatic linear modelling and preparation through IBM SPSS</td>
<td>Basic tables of descriptive statistics and graphs stemming from automatic linear modelling (IBM SPSS)</td>
</tr>
<tr>
<td>Kong and Song (2016). A study on customer feedback of tourism service using social big data</td>
<td>Design of an analysis model for the top Korean travel agency to help the company improve customer satisfaction and service quality</td>
<td>Empirical</td>
<td>Internal sources (emails, counselling data, bulletin information, after use comments/evaluations) and external sources (Twitter, Facebook, OnlineNews, Blog, Community)</td>
<td>Mostly unstructured (e.g., data from emails, social media networks) and several structured (e.g., bulleting info)</td>
<td>BuzzMonitoring (Types and proportion of keywords from the extracted data are digitized to analyze incidents and phenomena)</td>
<td>BuzzMonitoring including the following modules: NLP, data clustering, text summarization, sentiment analysis, structure data joiner.</td>
<td>No table nor graphs stemming from the Buzz Monitoring</td>
</tr>
<tr>
<td>Study</td>
<td>Title</td>
<td>Methods</td>
<td>Data Sources</td>
<td>Analysis Tools</td>
<td>Results</td>
<td></td>
<td></td>
</tr>
<tr>
<td>-------</td>
<td>-------</td>
<td>---------</td>
<td>--------------</td>
<td>----------------</td>
<td>---------</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Law et al. (2011).</td>
<td>Identifying changes and trends in Hong Kong outbound tourism</td>
<td>Trends in Hong Kong outbound tourism in terms of Future trip intentions Travel destinations Motivation to travel.</td>
<td>Empirical</td>
<td>Tourism behavior survey data</td>
<td>Historical domestic Surveys</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mariani et al. (2016).</td>
<td>Facebook as a destination marketing tool: Evidence from Italian regional destination management organizations</td>
<td>Explores how Italian regional Destination Management Organisations (DMOs) strategically employ Facebook to promote and market their destinations, and improves on the current metrics for capturing user engagement</td>
<td>Empirical</td>
<td>Overall number of Facebook posts posted on the official Italian regional DMOs’ Facebook pages</td>
<td>Data parser and analyzer calculating per post statistics</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Marine-Roig and Anton Clavé (2015).</td>
<td>Tourism analytics with massive user-generated content: A case study of Barcelona</td>
<td>Studying the online image of Barcelona as transmitted via social media through the analysis of more than 100,000 relevant travel blogs and online travel reviews (OTRs) written in English</td>
<td>Empirical</td>
<td>Heterogeneous including the travel blogs, webpages, travelogues and travel reviews about Barcelona (250,000 pages)</td>
<td>Data was extracted through Offline Explorer Enterprise (OEE).</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mi et al. (2014).</td>
<td>A new method for evaluating tour online review based on grey 2-tuple linguistic</td>
<td>Evaluation of online reviews</td>
<td>Empirical</td>
<td>Reviews from tourism website</td>
<td>Grey 2-tuple linguistic evaluation (expert evaluations)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Authors</td>
<td>Title</td>
<td>Methodology</td>
<td>Data Source</td>
<td>Analysis Method</td>
<td>Technology</td>
<td>Visualization</td>
<td></td>
</tr>
<tr>
<td>---------</td>
<td>-------</td>
<td>-------------</td>
<td>-------------</td>
<td>-----------------</td>
<td>------------</td>
<td>---------------</td>
<td></td>
</tr>
<tr>
<td>Mocanu et al. (2013).</td>
<td>The twitter of babel: Mapping world languages through microblogging platforms</td>
<td>Survey on worldwide linguistic indicators and trends through</td>
<td>Empirical</td>
<td>Large-scale dataset of geotagged tweets</td>
<td>Twitter API</td>
<td>Maps + charts and tables</td>
<td></td>
</tr>
<tr>
<td>Noguchi et al. (2016).</td>
<td>Advanced, high-performance big data technology and trial validation</td>
<td>Presents technology for analyzing data and location data</td>
<td>Conceptual, application design, case study</td>
<td>Smartphone application</td>
<td></td>
<td>Maps + charts and tables</td>
<td></td>
</tr>
<tr>
<td>Orellana et al. (2012).</td>
<td>Exploring visitor movement patterns in natural recreational areas</td>
<td>Explores the properties of the collective movement of visitors in recreational natural areas</td>
<td>Empirical</td>
<td>Global Positioning System tracking data</td>
<td></td>
<td>GIS + standard tables</td>
<td></td>
</tr>
<tr>
<td>Paldino et al. (2015).</td>
<td>Urban magnetism through the lens of geo-tagged photography</td>
<td>Tastes of individuals, and what attracts them to live in a particular city or spend vacation there.</td>
<td>Empirical</td>
<td>Geo-tagged photos</td>
<td>Flickr API</td>
<td>Maps + charts and tables</td>
<td></td>
</tr>
<tr>
<td>Study (2016)</td>
<td>Methodology</td>
<td>Data Source</td>
<td>Analysis Techniques</td>
<td>Results</td>
<td>Visualization</td>
<td></td>
<td></td>
</tr>
<tr>
<td>-------------</td>
<td>-------------</td>
<td>-------------</td>
<td>---------------------</td>
<td>---------</td>
<td>---------------</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Park et al.</td>
<td>Analysis of social media data on cruise tourism</td>
<td>Tweets containing search words</td>
<td>Twitter API and Web scraping</td>
<td>Word frequency, content analysis, and network analysis</td>
<td>Charts, tables, network diagrams</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Raun et al.</td>
<td>Measure space-time tracking data to analyze, monitor and compare destinations based on data describing actual visits</td>
<td>Anonymized roaming data of the foreign mobile phone call detail records</td>
<td>From telcom operator</td>
<td>Statistical analyses, ArcGIS for spatial analyses, binary logistic regression</td>
<td>Charts, tables, maps</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Shi et al.</td>
<td>Apply semantic web and big data techniques to help collect data, and implement platform and questionnaire design to construct stakeholder balance model for tourism intangible cultural heritage</td>
<td>Questionnaire + User Generated Content (reviews)</td>
<td>Structural equation model, path analysis</td>
<td>Structural equation model, path analysis</td>
<td>Charts, tables, maps</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Su et al.</td>
<td>Characterize geographical preferences of international tourists and quantify local influential factors of tourists’ destination preferences across time and space and origins</td>
<td>Metadata online geotagged photos</td>
<td>Flickr API</td>
<td>Statistical and spatial analyses</td>
<td>Charts, tables, maps</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Authors</td>
<td>Title</td>
<td>Methodology</td>
<td>Data Source</td>
<td>Analysis Methods</td>
<td>Visualizations</td>
<td></td>
<td></td>
</tr>
<tr>
<td>------------------</td>
<td>-----------------------------------------------------------------------</td>
<td>-------------</td>
<td>----------------------------------------------------------------------------</td>
<td>---------------------------------------------------------------------------------</td>
<td>----------------------------------</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sun et al. (2016)</td>
<td>Integration of Internet of Things (IoT) and big data analytics for smart connected communities</td>
<td>Conceptual + case study</td>
<td>Design of an IoT system personal sensors, open data, and participatory sensing to enhance the services in the area of tourism and cultural heritage with a Context-Aware Recommendation System</td>
<td>N/A</td>
<td>N/A</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Supak et al. (2015)</td>
<td>Geospatial analytics for federally managed tourism destinations and their demand markets</td>
<td>Empirical</td>
<td>Examine the general geospatial demand for overnight recreation on federal lands prior to the 2008 recession and the specific geospatial demand for national park regions</td>
<td>National Recreation Reservation Service reservations database Structured (*)</td>
<td>DB access Statistical and spatial analyses Charts, tables, maps</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tang et al. (2016)</td>
<td>Study related to spatial network of tourist flows and its structure in the urban areas</td>
<td>Empirical</td>
<td>Geotagged microblog posts Structured (*)</td>
<td>API from Sina Microblog Statistical, spatial and network analyses</td>
<td>Charts, tables, maps network diagrams</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Reference</td>
<td>Title</td>
<td>Research Objectives</td>
<td>Methodology</td>
<td>Data Sources</td>
<td>Analysis Method</td>
<td>Presentation Format</td>
<td></td>
</tr>
<tr>
<td>-----------</td>
<td>-------</td>
<td>---------------------</td>
<td>-------------</td>
<td>--------------</td>
<td>-----------------</td>
<td>---------------------</td>
<td></td>
</tr>
<tr>
<td>Wang et al. (2015). Revenue management: Progress, challenges, and research prospects</td>
<td>Discuss evolution and future developments of revenue management and use of big data analytics</td>
<td>Conceptual</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td></td>
</tr>
<tr>
<td>Wood et al. (2013). Using social media to quantify nature-based tourism and recreation</td>
<td>Online posted photos are used to estimate visitation rates and travelers’ origins. Compare to empirical data showing that crowd-sourced information can serve as reliable proxy for empirical visitation rates</td>
<td>Empirical</td>
<td>Empirical datasets that quantified visitation to 836 sites in 31 countries around the world + Flickr metadata</td>
<td>Dataset + Flickr API</td>
<td>Statistical and spatial analyse</td>
<td>Charts, tables, maps</td>
<td></td>
</tr>
<tr>
<td>Yang et al. (2014). Predicting hotel demand using destination marketing organization’s web traffic data</td>
<td>Demonstrate the value of website traffic data in predicting demand for hotel rooms at a destination, and potentially future revenue and performance</td>
<td>Empirical</td>
<td>Website traffic data and local hotel room demand data</td>
<td>Google analytics + standard data</td>
<td>Statistical and time series forecasts</td>
<td>Charts, tables</td>
<td></td>
</tr>
</tbody>
</table>
Yang et al. (2016). The big data analysis of land use evolution and its ecological security responses in silver beach of China by the clustering of spatial patterns

<table>
<thead>
<tr>
<th>Method</th>
<th>Data Sources</th>
<th>Results</th>
</tr>
</thead>
<tbody>
<tr>
<td>Empirical</td>
<td>Landsat satellite high-definition images</td>
<td>Pictures + metadata</td>
</tr>
</tbody>
</table>

Use remote-sensing images to analyze the land use evolution and to evaluate its ecological security.