

ANALYZING MARKS IN VISUALIZATIONS BASED ON DATASET CHARACTERISTICS

CROSS-REFERENCE TO RELATED APPLICATION

5 This application is a Utility Patent application based on previously filed U.S. Provisional Patent Application No. 62/862,349 filed on June 17, 2019, the benefit of the filing date of which is hereby claimed under 35 U.S.C. §119(e) and which is further incorporated in entirety by reference.

TECHNICAL FIELD

10 The present invention relates generally to data visualization, and more particularly, but not exclusively to, automatically explaining the reasons for the appearance of values in a visualization.

BACKGROUND

 Organizations are generating and collecting an ever increasing amount of data. This data
15 may be associated with disparate parts of the organization, such as, consumer activity, manufacturing activity, customer service, server logs, or the like. For various reasons, it may be inconvenient for such organizations to effectively utilize their vast collections of data. In some cases the quantity of data may make it difficult to effectively utilize the collected data to improve business practices. In some cases, organizations employ various tools to generate
20 visualizations of the some or all of their data. Employing visualizations to represent this data may enable organizations to improve their understanding of critical business operations and help them monitor key performance indicators. However, in some cases, visualizations may include marks, signal, values, or the like, that may seem out of place or otherwise anomalous. In some cases, determining the source or otherwise analyzing these the source or cause of such marks
25 may require an inordinate level of understanding of the underlying data that was used to generate the visualizations. Disadvantageously, this may require organizations to direct skilled or specialized data analysts to review the visualization and data to determine an explanation for why the mark may have a given value. Also, in some cases, even if the user has the skills or technical background to perform their own analysis, the underlying data may be sensitive or
30 otherwise inaccessible to users that may be reviewing the visualizations. Thus, it is with respect to these considerations and others that the present invention has been made.

BRIEF DESCRIPTION OF THE DRAWINGS

Non-limiting and non-exhaustive embodiments of the present innovations are described with reference to the following drawings. In the drawings, like reference numerals refer to like parts throughout the various figures unless otherwise specified. For a better understanding of the described innovations, reference will be made to the following Detailed Description of Various Embodiments, which is to be read in association with the accompanying drawings, wherein:

FIGURE 1 illustrates a system environment in which various embodiments may be implemented;

FIGURE 2 illustrates a schematic embodiment of a client computer;

FIGURE 3 illustrates a schematic embodiment of a network computer;

FIGURE 4 illustrates a logical architecture of a system for analyzing marks in visualizations based on data characteristics in accordance with one or more of the various embodiments;

FIGURE 5 illustrates a logical representation of a portion of a visualization in accordance with one or more of the various embodiments;

FIGURE 6 illustrates a logical representation of a portion of a mark assessment system in accordance with one or more of the various embodiments;

FIGURE 7 illustrates a logical representation of a portion of a user interface for analyzing marks in visualizations based on data characteristics in accordance with one or more of the various embodiments;

FIGURE 8 illustrates an overview flowchart for a process for analyzing marks in visualizations based on data characteristics in accordance with one or more of the various embodiments;

FIGURE 9 illustrates a flowchart of a process for analyzing marks using assessment models in accordance with one or more of the various embodiments;

FIGURE 10 illustrates a flowchart of a process for adapting filtering or ranking of assessment models based on monitoring user interactions in accordance with one or more of the various embodiments;

FIGURE 11 illustrates a flowchart of a process for generating expected value distributions to for assessment models in accordance with one or more of the various embodiments; and

FIGURE 12 illustrates a flowchart of a process for comparing the results of assessment models in accordance with one or more of the various embodiments.

DETAILED DESCRIPTION OF VARIOUS EMBODIMENTS

Various embodiments now will be described more fully hereinafter with reference to the accompanying drawings, which form a part hereof, and which show, by way of illustration, specific exemplary embodiments by which the invention may be practiced. The embodiments may, however, be embodied in many different forms and should not be construed as limited to those 10 the embodiments set forth herein; rather, these embodiments are provided so that this disclosure will be thorough and complete, and will fully convey the scope of the embodiments to those skilled in the art. Among other things, the various embodiments may be methods, systems, media or devices. Accordingly, the various embodiments may take the form of an entirely 15 hardware embodiment, an entirely software embodiment or an embodiment combining software and hardware aspects. The following detailed description is, therefore, not to be taken in a limiting sense.

Throughout the specification and claims, the following terms take the meanings explicitly associated herein, unless the context clearly dictates otherwise. The phrase “in one 20 embodiment” as used herein does not necessarily refer to the same embodiment, though it may. Furthermore, the phrase “in another embodiment” as used herein does not necessarily refer to a different embodiment, although it may. Thus, as described below, various embodiments may be readily combined, without departing from the scope or spirit of the invention.

In addition, as used herein, the term “or” is an inclusive “or” operator, and is equivalent 25 to the term “and/or,” unless the context clearly dictates otherwise. The term “based on” is not exclusive and allows for being based on additional factors not described, unless the context clearly dictates otherwise. In addition, throughout the specification, the meaning of “a,” “an,” and “the” include plural references. The meaning of “in” includes “in” and “on.”

For example embodiments, the following terms are also used herein according to the 30 corresponding meaning, unless the context clearly dictates otherwise.

As used herein the term, “engine” refers to logic embodied in hardware or software instructions, which can be written in a programming language, such as C, C++, Objective-C, COBOL, Java™, PHP, Perl, JavaScript, Ruby, VBScript, Microsoft .NET™ languages such as C#, or the like. An engine may be compiled into executable programs or written in interpreted programming languages. Software engines may be callable from other engines or from themselves. Engines described herein refer to one or more logical modules that can be merged with other engines or applications, or can be divided into sub-engines. The engines can be stored in non-transitory computer-readable medium or computer storage device and be stored on and executed by one or more general purpose computers, thus creating a special purpose computer configured to provide the engine.

As used herein, the term “data source” refers to databases, applications, services, file systems, or the like, that store or provide information for an organization. Examples of data sources may include, RDBMS databases, graph databases, spreadsheets, file systems, document management systems, local or remote data streams, or the like. In some cases, data sources are organized around one or more tables or table-like structure. In other cases, data sources be organized as a graph or graph-like structure.

As used herein the term “data model” refers to one or more data structures that provide a representation of an underlying data source. In some cases, data models may provide views of a data source for particular applications. Data models may be considered views or interfaces to the underlying data source. In some cases, data models may map directly to a data source (e.g., practically a logical pass through). Also, in some cases, data models may be provided by a data source. In some circumstances, data models may be considered interfaces to data sources. Data models enable organizations to organize or present information from data sources in ways that may be more convenient, more meaningful (e.g., easier to reason about), safer, or the like.

As used herein the term “data object” refers to one or more entities or data structures that comprise data models. In some cases, data objects may be considered portions of the data model. Data objects may represent individual instances of items or classes or kinds of items.

As used herein the term “panel” refers to region within a graphical user interface (GUI) that has a defined geometry (e.g., x, y, z-order) within the GUI. Panels may be arranged to display information to users or to host one or more interactive controls. The geometry or styles

associated with panels may be defined using configuration information, including dynamic rules. Also, in some cases, users may be enabled to perform actions on one or more panels, such as, moving, showing, hiding, re-sizing, re-ordering, or the like.

As used herein the “visualization model” refers to one or more data structures that
5 represent one or more representations of a data model that may be suitable for use in a visualization that is displayed on one or more hardware displays. Visualization models may define styling or user interface features that may be made available to non-authoring user.

As used herein the term “display object” refers to one or more data structures that
comprise visualization models. In some cases, display objects may be considered portions of the
10 visualization model. Display objects may represent individual instances of items or entire classes or kinds of items that may be displayed in a visualization. In some embodiments, display objects may be considered or referred to as views because they provide a view of some portion of the data model.

As used herein, the term “mark” refers to a distinct or otherwise identifiable portion of a
15 visualization that may correspond to particular value or result in the visualization. For example, if a visualization includes a bar chart, one or more of the bars may be considered to be marks. Likewise, if a visualization includes a line plot, positions on the plot may be considered a marks.

As used herein, the term “mark-of-interest” refers to a mark in a visualization that has
been selected from among the other marks included in the visualization. In some cases, marks in
20 visualizations may incorporate one or more interactive features that may enable a user to select or identify one or more marks-of-interest from among the marks comprising a visualization. For example, a user may be enabled to select a mark-of-interest by right-clicking a mouse button while the mouse pointer may be hovering over a mark. In some cases, marks-of-interest may be selected via searching, filtering, or the like.

25 As used herein the term “configuration information” refers to information that may include rule based policies, pattern matching, scripts (e.g., computer readable instructions), or the like, that may be provided from various sources, including, configuration files, databases, user input, built-in defaults, or the like, or combination thereof.

The following briefly describes embodiments of the invention in order to provide a basic understanding of some aspects of the invention. This brief description is not intended as an extensive overview. It is not intended to identify key or critical elements, or to delineate or otherwise narrow the scope. Its purpose is merely to present some concepts in a simplified form
5 as a prelude to the more detailed description that is presented later.

Briefly stated, various embodiments are directed to managing visualizations of data using one or more processors that execute one or more instructions to perform as described herein.

In one or more of the various embodiments, one or more visualizations based on data from a data source may be provided such that the one or more visualizations include one or more
10 marks that may be associated with one or more values from the data source.

In one or more of the various embodiments, a mark-of-interest may be determined from the one or more marks based on one or more characteristics of the one or more marks, the one or more visualizations, or the like. In one or more of the various embodiments, determining the mark-of-interest may include: automatically analyzing the one or more values associated with
15 the one or more marks; and determining the mark-of-interest based on the automatic analysis.

In one or more of the various embodiments, one or more assessment models may be employed to generate one or more assessment results for the one or more assessment models based on the mark-of-interest and the data from the data source such that each assessment result may be associated with one or more scores that may be based on a best fit to the data from the
20 data source and a value of the mark-of-interest. In one or more of the various embodiments, generating the one or more assessment results may include: generating one or more natural language narratives that describe one or more characteristics of the mark-of-interest based on its assessment and one or more narrative templates; and including the one or more natural language narratives in the one or more assessment results.

25 In one or more of the various embodiments, the one or more assessment results may be rank ordered based on their association with the one or more scores.

In one or more of the various embodiments, a report that includes the rank ordered list of the one or more assessment results may be provided. In one or more of the various embodiments, providing the report may include generating one or more other visualizations that

may be associated with the one or more assessment models, the one or more assessment results, the mark-of-interest, or the like.

In one or more of the various embodiments, feedback information that is associated with the report may be monitored. In one or more of the various embodiments, one or more other
5 scores based on the feedback information may be provided. In one or more of the various embodiments, the one or more other scores may be employed to determine one or more modifications to the one or more assessments results.

In one or more of the various embodiments, the one or more assessment results may be filtered based on the one or more scores such that each assessment result associated with a score
10 that may be less than a defined threshold value is excluded from the report.

In one or more of the various embodiments, data associated with the mark-of-interest that is included in the data source and omitted from the visualization may be evaluated. Accordingly, in some embodiments, a portion of the one or more assessment results may be generated based on the evaluation.

15 **Illustrated Operating Environment**

FIGURE 1 shows components of one embodiment of an environment in which embodiments of the invention may be practiced. Not all of the components may be required to practice the invention, and variations in the arrangement and type of the components may be made without departing from the spirit or scope of the invention. As shown, system 100 of
20 FIGURE 1 includes local area networks (LANs)/ wide area networks (WANs) – (network) 110, wireless network 108, client computers 102-105, visualization server computer 116, data source server computer 118, or the like.

At least one embodiment of client computers 102-105 is described in more detail below in conjunction with FIGURE 2. In one embodiment, at least some of client computers 102-105
25 may operate over one or more wired or wireless networks, such as networks 108, or 110. Generally, client computers 102-105 may include virtually any computer capable of communicating over a network to send and receive information, perform various online activities, offline actions, or the like. In one embodiment, one or more of client computers 102-105 may be configured to operate within a business or other entity to perform a variety of

services for the business or other entity. For example, client computers 102-105 may be configured to operate as a web server, firewall, client application, media player, mobile telephone, game console, desktop computer, or the like. However, client computers 102-105 are not constrained to these services and may also be employed, for example, as for end-user
5 computing in other embodiments. It should be recognized that more or less client computers (as shown in FIGURE 1) may be included within a system such as described herein, and embodiments are therefore not constrained by the number or type of client computers employed.

Computers that may operate as client computer 102 may include computers that typically connect using a wired or wireless communications medium such as personal computers,
10 multiprocessor systems, microprocessor-based or programmable electronic devices, network PCs, or the like. In some embodiments, client computers 102-105 may include virtually any portable computer capable of connecting to another computer and receiving information such as, laptop computer 103, mobile computer 104, tablet computers 105, or the like. However, portable computers are not so limited and may also include other portable computers such as cellular
15 telephones, display pagers, radio frequency (RF) devices, infrared (IR) devices, Personal Digital Assistants (PDAs), handheld computers, wearable computers, integrated devices combining one or more of the preceding computers, or the like. As such, client computers 102-105 typically range widely in terms of capabilities and features. Moreover, client computers 102-105 may access various computing applications, including a browser, or other web-based application.

20 A web-enabled client computer may include a browser application that is configured to send requests and receive responses over the web. The browser application may be configured to receive and display graphics, text, multimedia, and the like, employing virtually any web-based language. In one embodiment, the browser application is enabled to employ JavaScript, Hypertext Markup Language (HTML), eXtensible Markup Language (XML), JavaScript Object
25 Notation (JSON), Cascading Style Sheets (CSS), or the like, or combination thereof, to display and send a message. In one embodiment, a user of the client computer may employ the browser application to perform various activities over a network (online). However, another application may also be used to perform various online activities.

Client computers 102-105 also may include at least one other client application that is
30 configured to receive or send content between another computer. The client application may

include a capability to send or receive content, or the like. The client application may further provide information that identifies itself, including a type, capability, name, and the like. In one embodiment, client computers 102-105 may uniquely identify themselves through any of a variety of mechanisms, including an Internet Protocol (IP) address, a phone number, Mobile
5 Identification Number (MIN), an electronic serial number (ESN), a client certificate, or other device identifier. Such information may be provided in one or more network packets, or the like, sent between other client computers, visualization server computer 116, data source server computer 118, or other computers.

Client computers 102-105 may further be configured to include a client application that
10 enables an end-user to log into an end-user account that may be managed by another computer, such as visualization server computer 116, data source server computer 118, or the like. Such an end-user account, in one non-limiting example, may be configured to enable the end-user to manage one or more online activities, including in one non-limiting example, project
15 management, software development, system administration, configuration management, search activities, social networking activities, browse various websites, communicate with other users, or the like. Also, client computers may be arranged to enable users to display reports, interactive user-interfaces, or results provided by visualization server computer 116, data source server computer 118.

Wireless network 108 is configured to couple client computers 103-105 and its
20 components with network 110. Wireless network 108 may include any of a variety of wireless sub-networks that may further overlay stand-alone ad-hoc networks, and the like, to provide an infrastructure-oriented connection for client computers 103-105. Such sub-networks may include mesh networks, Wireless LAN (WLAN) networks, cellular networks, and the like. In one embodiment, the system may include more than one wireless network.

Wireless network 108 may further include an autonomous system of terminals, gateways,
25 routers, and the like connected by wireless radio links, and the like. These connectors may be configured to move freely and randomly and organize themselves arbitrarily, such that the topology of wireless network 108 may change rapidly.

Wireless network 108 may further employ a plurality of access technologies including
30 2nd (2G), 3rd (3G), 4th (4G) 5th (5G) generation radio access for cellular systems, WLAN,

Wireless Router (WR) mesh, and the like. Access technologies such as 2G, 3G, 4G, 5G, and future access networks may enable wide area coverage for mobile computers, such as client computers 103-105 with various degrees of mobility. In one non-limiting example, wireless network 108 may enable a radio connection through a radio network access such as Global
5 System for Mobil communication (GSM), General Packet Radio Services (GPRS), Enhanced Data GSM Environment (EDGE), code division multiple access (CDMA), time division multiple access (TDMA), Wideband Code Division Multiple Access (WCDMA), High Speed Downlink Packet Access (HSDPA), Long Term Evolution (LTE), and the like. In essence, wireless network 108 may include virtually any wireless communication mechanism by which
10 information may travel between client computers 103-105 and another computer, network, a cloud-based network, a cloud instance, or the like.

Network 110 is configured to couple network computers with other computers, including, visualization server computer 116, data source server computer 118, client computers 102, and client computers 103-105 through wireless network 108, or the like. Network 110 is
15 enabled to employ any form of computer readable media for communicating information from one electronic device to another. Also, network 110 can include the Internet in addition to local area networks (LANs), wide area networks (WANs), direct connections, such as through a universal serial bus (USB) port, Ethernet port, other forms of computer-readable media, or any combination thereof. On an interconnected set of LANs, including those based on differing
20 architectures and protocols, a router acts as a link between LANs, enabling messages to be sent from one to another. In addition, communication links within LANs typically include twisted wire pair or coaxial cable, while communication links between networks may utilize analog telephone lines, full or fractional dedicated digital lines including T1, T2, T3, and T4, or other carrier mechanisms including, for example, E-carriers, Integrated Services Digital Networks
25 (ISDNs), Digital Subscriber Lines (DSLs), wireless links including satellite links, or other communications links known to those skilled in the art. Moreover, communication links may further employ any of a variety of digital signaling technologies, including without limit, for example, DS-0, DS-1, DS-2, DS-3, DS-4, OC-3, OC-12, OC-48, or the like. Furthermore, remote computers and other related electronic devices could be remotely connected to either
30 LANs or WANs via a modem and temporary telephone link. In one embodiment, network 110 may be configured to transport information of an Internet Protocol (IP).

Additionally, communication media typically embodies computer readable instructions, data structures, program modules, or other transport mechanism and includes any information non-transitory delivery media or transitory delivery media. By way of example, communication media includes wired media such as twisted pair, coaxial cable, fiber optics, wave guides, and other wired media and wireless media such as acoustic, RF, infrared, and other wireless media.

Also, one embodiment of visualization server computer 116, data source server computer 118 are described in more detail below in conjunction with FIGURE 3. Although FIGURE 1 illustrates visualization server computer 116, data source server computer 118, or the like, each as a single computer, the innovations or embodiments are not so limited. For example, one or more functions of visualization server computer 116, data source server computer 118, or the like, may be distributed across one or more distinct network computers. Moreover, in one or more embodiments, visualization server computer 116, data source server computer 118 may be implemented using a plurality of network computers. Further, in one or more of the various embodiments, visualization server computer 116, data source server computer 118, or the like, may be implemented using one or more cloud instances in one or more cloud networks. Accordingly, these innovations and embodiments are not to be construed as being limited to a single environment, and other configurations, and other architectures are also envisaged.

Illustrative Client Computer

FIGURE 2 shows one embodiment of client computer 200 that may include many more or less components than those shown. Client computer 200 may represent, for example, one or more embodiment of mobile computers or client computers shown in FIGURE 1.

Client computer 200 may include processor 202 in communication with memory 204 via bus 228. Client computer 200 may also include power supply 230, network interface 232, audio interface 256, display 250, keypad 252, illuminator 254, video interface 242, input/output interface 238, haptic interface 264, global positioning systems (GPS) receiver 258, open air gesture interface 260, temperature interface 262, camera(s) 240, projector 246, pointing device interface 266, processor-readable stationary storage device 234, and processor-readable removable storage device 236. Client computer 200 may optionally communicate with a base station (not shown), or directly with another computer. And in one embodiment, although not

shown, a gyroscope may be employed within client computer 200 to measuring or maintaining an orientation of client computer 200.

Power supply 230 may provide power to client computer 200. A rechargeable or non-rechargeable battery may be used to provide power. The power may also be provided by an external power source, such as an AC adapter or a powered docking cradle that supplements or recharges the battery.

Network interface 232 includes circuitry for coupling client computer 200 to one or more networks, and is constructed for use with one or more communication protocols and technologies including, but not limited to, protocols and technologies that implement any portion of the OSI model for mobile communication (GSM), CDMA, time division multiple access (TDMA), UDP, TCP/IP, SMS, MMS, GPRS, WAP, UWB, WiMax, SIP/RTP, GPRS, EDGE, WCDMA, LTE, UMTS, OFDM, CDMA2000, EV-DO, HSDPA, or any of a variety of other wireless communication protocols. Network interface 232 is sometimes known as a transceiver, transceiving device, or network interface card (NIC).

Audio interface 256 may be arranged to produce and receive audio signals such as the sound of a human voice. For example, audio interface 256 may be coupled to a speaker and microphone (not shown) to enable telecommunication with others or generate an audio acknowledgment for some action. A microphone in audio interface 256 can also be used for input to or control of client computer 200, e.g., using voice recognition, detecting touch based on sound, and the like.

Display 250 may be a liquid crystal display (LCD), gas plasma, electronic ink, light emitting diode (LED), Organic LED (OLED) or any other type of light reflective or light transmissive display that can be used with a computer. Display 250 may also include a touch interface 244 arranged to receive input from an object such as a stylus or a digit from a human hand, and may use resistive, capacitive, surface acoustic wave (SAW), infrared, radar, or other technologies to sense touch or gestures.

Projector 246 may be a remote handheld projector or an integrated projector that is capable of projecting an image on a remote wall or any other reflective object such as a remote screen.

Video interface 242 may be arranged to capture video images, such as a still photo, a video segment, an infrared video, or the like. For example, video interface 242 may be coupled to a digital video camera, a web-camera, or the like. Video interface 242 may comprise a lens, an image sensor, and other electronics. Image sensors may include a complementary metal-oxide-semiconductor (CMOS) integrated circuit, charge-coupled device (CCD), or any other integrated circuit for sensing light.

Keypad 252 may comprise any input device arranged to receive input from a user. For example, keypad 252 may include a push button numeric dial, or a keyboard. Keypad 252 may also include command buttons that are associated with selecting and sending images.

Illuminator 254 may provide a status indication or provide light. Illuminator 254 may remain active for specific periods of time or in response to event messages. For example, when illuminator 254 is active, it may back-light the buttons on keypad 252 and stay on while the client computer is powered. Also, illuminator 254 may back-light these buttons in various patterns when particular actions are performed, such as dialing another client computer. Illuminator 254 may also cause light sources positioned within a transparent or translucent case of the client computer to illuminate in response to actions.

Further, client computer 200 may also comprise hardware security module (HSM) 268 for providing additional tamper resistant safeguards for generating, storing or using security/cryptographic information such as, keys, digital certificates, passwords, passphrases, two-factor authentication information, or the like. In some embodiments, hardware security module may be employed to support one or more standard public key infrastructures (PKI), and may be employed to generate, manage, or store keys pairs, or the like. In some embodiments, HSM 268 may be a stand-alone computer, in other cases, HSM 268 may be arranged as a hardware card that may be added to a client computer.

Client computer 200 may also comprise input/output interface 238 for communicating with external peripheral devices or other computers such as other client computers and network computers. The peripheral devices may include an audio headset, virtual reality headsets, display screen glasses, remote speaker system, remote speaker and microphone system, and the like. Input/output interface 238 can utilize one or more technologies, such as Universal Serial Bus (USB), Infrared, WiFi, WiMax, Bluetooth™, and the like.

Input/output interface 238 may also include one or more sensors for determining geolocation information (e.g., GPS), monitoring electrical power conditions (e.g., voltage sensors, current sensors, frequency sensors, and so on), monitoring weather (e.g., thermostats, barometers, anemometers, humidity detectors, precipitation scales, or the like), or the like.

- 5 Sensors may be one or more hardware sensors that collect or measure data that is external to client computer 200.

- 10 Haptic interface 264 may be arranged to provide tactile feedback to a user of the client computer. For example, the haptic interface 264 may be employed to vibrate client computer 200 in a particular way when another user of a computer is calling. Temperature interface 262 may be used to provide a temperature measurement input or a temperature changing output to a user of client computer 200. Open air gesture interface 260 may sense physical gestures of a user of client computer 200, for example, by using single or stereo video cameras, radar, a gyroscopic sensor inside a computer held or worn by the user, or the like. Camera 240 may be used to track physical eye movements of a user of client computer 200.

- 15 GPS transceiver 258 can determine the physical coordinates of client computer 200 on the surface of the Earth, which typically outputs a location as latitude and longitude values. GPS transceiver 258 can also employ other geo-positioning mechanisms, including, but not limited to, triangulation, assisted GPS (AGPS), Enhanced Observed Time Difference (E-OTD), Cell Identifier (CI), Service Area Identifier (SAI), Enhanced Timing Advance (ETA), Base Station Subsystem (BSS), or the like, to further determine the physical location of client computer 200 on the surface of the Earth. It is understood that under different conditions, GPS transceiver 258 can determine a physical location for client computer 200. In one or more embodiments, however, client computer 200 may, through other components, provide other information that may be employed to determine a physical location of the client computer, including for example, 20 a Media Access Control (MAC) address, IP address, and the like.

- 25 In at least one of the various embodiments, applications, such as, operating system 206, other client apps 224, web browser 226, or the like, may be arranged to employ geo-location information to select one or more localization features, such as, time zones, languages, currencies, calendar formatting, or the like. Localization features may be used in display objects, 30 data models, data objects, user-interfaces, reports, as well as internal processes or databases. In

at least one of the various embodiments, geo-location information used for selecting localization information may be provided by GPS 258. Also, in some embodiments, geolocation information may include information provided using one or more geolocation protocols over the networks, such as, wireless network 108 or network 111.

5 Human interface components can be peripheral devices that are physically separate from client computer 200, allowing for remote input or output to client computer 200. For example, information routed as described here through human interface components such as display 250 or keyboard 252 can instead be routed through network interface 232 to appropriate human interface components located remotely. Examples of human interface peripheral components
10 that may be remote include, but are not limited to, audio devices, pointing devices, keypads, displays, cameras, projectors, and the like. These peripheral components may communicate over a Pico Network such as Bluetooth™, Zigbee™ and the like. One non-limiting example of a client computer with such peripheral human interface components is a wearable computer, which might include a remote pico projector along with one or more cameras that remotely
15 communicate with a separately located client computer to sense a user's gestures toward portions of an image projected by the pico projector onto a reflected surface such as a wall or the user's hand.

A client computer may include web browser application 226 that is configured to receive and to send web pages, web-based messages, graphics, text, multimedia, and the like. The client
20 computer's browser application may employ virtually any programming language, including a wireless application protocol messages (WAP), and the like. In one or more embodiments, the browser application is enabled to employ Handheld Device Markup Language (HDML), Wireless Markup Language (WML), WMLScript, JavaScript, Standard Generalized Markup Language (SGML), HyperText Markup Language (HTML), eXtensible Markup Language
25 (XML), HTML5, and the like.

Memory 204 may include RAM, ROM, or other types of memory. Memory 204 illustrates an example of computer-readable storage media (devices) for storage of information such as computer-readable instructions, data structures, program modules or other data. Memory 204 may store BIOS 208 for controlling low-level operation of client computer 200. The
30 memory may also store operating system 206 for controlling the operation of client computer

200. It will be appreciated that this component may include a general-purpose operating system such as a version of UNIX, or LINUX™, or a specialized client computer communication operating system such as Windows Phone™, or the Symbian® operating system. The operating system may include, or interface with a Java virtual machine module that enables control of
5 hardware components or operating system operations via Java application programs.

Memory 204 may further include one or more data storage 210, which can be utilized by client computer 200 to store, among other things, applications 220 or other data. For example, data storage 210 may also be employed to store information that describes various capabilities of client computer 200. The information may then be provided to another device or computer based
10 on any of a variety of methods, including being sent as part of a header during a communication, sent upon request, or the like. Data storage 210 may also be employed to store social networking information including address books, buddy lists, aliases, user profile information, or the like. Data storage 210 may further include program code, data, algorithms, and the like, for use by a processor, such as processor 202 to execute and perform actions. In one embodiment, at least
15 some of data storage 210 might also be stored on another component of client computer 200, including, but not limited to, non-transitory processor-readable removable storage device 236, processor-readable stationary storage device 234, or even external to the client computer.

Applications 220 may include computer executable instructions which, when executed by client computer 200, transmit, receive, or otherwise process instructions and data.
20 Applications 220 may include, for example, client visualization engine 222, other client applications 224, web browser 226, or the like. Client computers may be arranged to exchange communications one or more servers.

Other examples of application programs include calendars, search programs, email client applications, IM applications, SMS applications, Voice Over Internet Protocol (VOIP)
25 applications, contact managers, task managers, transcoders, database programs, word processing programs, security applications, spreadsheet programs, games, search programs, visualization applications, and so forth.

Additionally, in one or more embodiments (not shown in the figures), client computer 200 may include an embedded logic hardware device instead of a CPU, such as, an Application
30 Specific Integrated Circuit (ASIC), Field Programmable Gate Array (FPGA), Programmable

Array Logic (PAL), or the like, or combination thereof. The embedded logic hardware device may directly execute its embedded logic to perform actions. Also, in one or more embodiments (not shown in the figures), client computer 200 may include one or more hardware micro-controllers instead of CPUs. In one or more embodiments, the one or more micro-controllers
5 may directly execute their own embedded logic to perform actions and access its own internal memory and its own external Input and Output Interfaces (e.g., hardware pins or wireless transceivers) to perform actions, such as System On a Chip (SOC), or the like.

Illustrative Network Computer

FIGURE 3 shows one embodiment of network computer 300 that may be included in a
10 system implementing one or more of the various embodiments. Network computer 300 may include many more or less components than those shown in FIGURE 3. However, the components shown are sufficient to disclose an illustrative embodiment for practicing these innovations. Network computer 300 may represent, for example, one embodiment of at least one of visualization server computer 116, data source server computer 118, or the like, of FIGURE
15 1.

Network computers, such as, network computer 300 may include a processor 302 that may be in communication with a memory 304 via a bus 328. In some embodiments, processor 302 may be comprised of one or more hardware processors, or one or more processor cores. In some cases, one or more of the one or more processors may be specialized processors designed
20 to perform one or more specialized actions, such as, those described herein. Network computer 300 also includes a power supply 330, network interface 332, audio interface 356, display 350, keyboard 352, input/output interface 338, processor-readable stationary storage device 334, and processor-readable removable storage device 336. Power supply 330 provides power to network computer 300.

25 Network interface 332 includes circuitry for coupling network computer 300 to one or more networks, and is constructed for use with one or more communication protocols and technologies including, but not limited to, protocols and technologies that implement any portion of the Open Systems Interconnection model (OSI model), global system for mobile communication (GSM), code division multiple access (CDMA), time division multiple access
30 (TDMA), user datagram protocol (UDP), transmission control protocol/Internet protocol

(TCP/IP), Short Message Service (SMS), Multimedia Messaging Service (MMS), general packet radio service (GPRS), WAP, ultra-wide band (UWB), IEEE 802.16 Worldwide Interoperability for Microwave Access (WiMax), Session Initiation Protocol/Real-time Transport Protocol (SIP/RTP), or any of a variety of other wired and wireless communication protocols. Network
5 interface 332 is sometimes known as a transceiver, transceiving device, or network interface card (NIC). Network computer 300 may optionally communicate with a base station (not shown), or directly with another computer.

Audio interface 356 is arranged to produce and receive audio signals such as the sound of a human voice. For example, audio interface 356 may be coupled to a speaker and microphone
10 (not shown) to enable telecommunication with others or generate an audio acknowledgment for some action. A microphone in audio interface 356 can also be used for input to or control of network computer 300, for example, using voice recognition.

Display 350 may be a liquid crystal display (LCD), gas plasma, electronic ink, light emitting diode (LED), Organic LED (OLED) or any other type of light reflective or light
15 transmissive display that can be used with a computer. In some embodiments, display 350 may be a handheld projector or pico projector capable of projecting an image on a wall or other object.

Network computer 300 may also comprise input/output interface 338 for communicating with external devices or computers not shown in FIGURE 3. Input/output interface 338 can
20 utilize one or more wired or wireless communication technologies, such as USB™, Firewire™, WiFi, WiMax, Thunderbolt™, Infrared, Bluetooth™, Zigbee™, serial port, parallel port, and the like.

Also, input/output interface 338 may also include one or more sensors for determining geolocation information (e.g., GPS), monitoring electrical power conditions (e.g., voltage
25 sensors, current sensors, frequency sensors, and so on), monitoring weather (e.g., thermostats, barometers, anemometers, humidity detectors, precipitation scales, or the like), or the like. Sensors may be one or more hardware sensors that collect or measure data that is external to network computer 300. Human interface components can be physically separate from network computer 300, allowing for remote input or output to network computer 300. For example,
30 information routed as described here through human interface components such as display 350

or keyboard 352 can instead be routed through the network interface 332 to appropriate human interface components located elsewhere on the network. Human interface components include any component that allows the computer to take input from, or send output to, a human user of a computer. Accordingly, pointing devices such as mice, styluses, track balls, or the like, may
5 communicate through pointing device interface 358 to receive user input.

GPS transceiver 340 can determine the physical coordinates of network computer 300 on the surface of the Earth, which typically outputs a location as latitude and longitude values. GPS transceiver 340 can also employ other geo-positioning mechanisms, including, but not limited to, triangulation, assisted GPS (AGPS), Enhanced Observed Time Difference (E-OTD), Cell
10 Identifier (CI), Service Area Identifier (SAI), Enhanced Timing Advance (ETA), Base Station Subsystem (BSS), or the like, to further determine the physical location of network computer 300 on the surface of the Earth. It is understood that under different conditions, GPS transceiver 340 can determine a physical location for network computer 300. In one or more embodiments, however, network computer 300 may, through other components, provide other information that
15 may be employed to determine a physical location of the client computer, including for example, a Media Access Control (MAC) address, IP address, and the like.

In at least one of the various embodiments, applications, such as, operating system 306, assessment engine 322, visualization engine 324, modeling engine 326, other applications 329, or the like, may be arranged to employ geo-location information to select one or more
20 localization features, such as, time zones, languages, currencies, currency formatting, calendar formatting, or the like. Localization features may be used in user interfaces, dashboards, visualizations, reports, as well as internal processes or databases. In at least one of the various embodiments, geo-location information used for selecting localization information may be provided by GPS 340. Also, in some embodiments, geolocation information may include
25 information provided using one or more geolocation protocols over the networks, such as, wireless network 108 or network 111.

Memory 304 may include Random Access Memory (RAM), Read-Only Memory (ROM), or other types of memory. Memory 304 illustrates an example of computer-readable storage media (devices) for storage of information such as computer-readable instructions, data
30 structures, program modules or other data. Memory 304 stores a basic input/output system

(BIOS) 308 for controlling low-level operation of network computer 300. The memory also stores an operating system 306 for controlling the operation of network computer 300. It will be appreciated that this component may include a general-purpose operating system such as a version of UNIX, or LINUX™, or a specialized operating system such as Microsoft

5 Corporation's Windows® operating system, or the Apple Corporation's OSX® operating system. The operating system may include, or interface with one or more virtual machine modules, such as, a Java virtual machine module that enables control of hardware components or operating system operations via Java application programs. Likewise, other runtime environments may be included.

10 Memory 304 may further include one or more data storage 310, which can be utilized by network computer 300 to store, among other things, applications 320 or other data. For example, data storage 310 may also be employed to store information that describes various capabilities of network computer 300. The information may then be provided to another device or computer based on any of a variety of methods, including being sent as part of a header during a
15 communication, sent upon request, or the like. Data storage 310 may also be employed to store social networking information including address books, buddy lists, aliases, user profile information, or the like. Data storage 310 may further include program code, data, algorithms, and the like, for use by a processor, such as processor 302 to execute and perform actions such as those actions described below. In one embodiment, at least some of data storage 310 might
20 also be stored on another component of network computer 300, including, but not limited to, non-transitory media inside processor-readable removable storage device 336, processor-readable stationary storage device 334, or any other computer-readable storage device within network computer 300, or even external to network computer 300. Data storage 310 may include, for example, data models 314, data sources 316, visualization models 318, assessment
25 models 319, or the like.

Applications 320 may include computer executable instructions which, when executed by network computer 300, transmit, receive, or otherwise process messages (e.g., SMS, Multimedia Messaging Service (MMS), Instant Message (IM), email, or other messages), audio, video, and enable telecommunication with another user of another mobile computer. Other
30 examples of application programs include calendars, search programs, email client applications, IM applications, SMS applications, Voice Over Internet Protocol (VOIP) applications, contact

managers, task managers, transcoders, database programs, word processing programs, security applications, spreadsheet programs, games, search programs, and so forth. Applications 320 may include assessment engine 322, visualization engine 324, modeling engine 326, other applications 329, or the like, that may be arranged to perform actions for embodiments described

5 below. In one or more of the various embodiments, one or more of the applications may be implemented as modules or components of another application. Further, in one or more of the various embodiments, applications may be implemented as operating system extensions, modules, plugins, or the like.

Furthermore, in one or more of the various embodiments, assessment engine 322,

10 visualization engine 324, modeling engine 326, other applications 329, or the like, may be operative in a cloud-based computing environment. In one or more of the various embodiments, these applications, and others, that comprise the management platform may be executing within virtual machines or virtual servers that may be managed in a cloud-based based computing environment. In one or more of the various embodiments, in this context the applications may

15 flow from one physical network computer within the cloud-based environment to another depending on performance and scaling considerations automatically managed by the cloud computing environment. Likewise, in one or more of the various embodiments, virtual machines or virtual servers dedicated to assessment engine 322, visualization engine 324, modeling engine 326, other applications 329, or the like, may be provisioned and de-commissioned automatically.

20 Also, in one or more of the various embodiments, assessment engine 322, visualization engine 324, modeling engine 326, other applications 329, or the like, may be located in virtual servers running in a cloud-based computing environment rather than being tied to one or more specific physical network computers.

Further, network computer 300 may also comprise hardware security module (HSM) 360

25 for providing additional tamper resistant safeguards for generating, storing or using security/cryptographic information such as, keys, digital certificates, passwords, passphrases, two-factor authentication information, or the like. In some embodiments, hardware security module may be employed to support one or more standard public key infrastructures (PKI), and may be employed to generate, manage, or store keys pairs, or the like. In some embodiments,

HSM 360 may be a stand-alone network computer, in other cases, HSM 360 may be arranged as a hardware card that may be installed in a network computer.

Additionally, in one or more embodiments (not shown in the figures), network computer 300 may include an embedded logic hardware device instead of a CPU, such as, an Application Specific Integrated Circuit (ASIC), Field Programmable Gate Array (FPGA), Programmable Array Logic (PAL), or the like, or combination thereof. The embedded logic hardware device may directly execute its embedded logic to perform actions. Also, in one or more embodiments (not shown in the figures), the network computer may include one or more hardware microcontrollers instead of a CPU. In one or more embodiments, the one or more microcontrollers may directly execute their own embedded logic to perform actions and access their own internal memory and their own external Input and Output Interfaces (e.g., hardware pins or wireless transceivers) to perform actions, such as System On a Chip (SOC), or the like.

Illustrative Logical System Architecture

FIGURE 4 illustrates a logical architecture of system 400 for analyzing marks in visualizations based on data characteristics in accordance with one or more of the various embodiments. In one or more of the various embodiments, system 400 may be comprised of various components, including, one or more modeling engines, such as, modeling engine 402; one or more visualization engines, such as, visualization engine 404; one or more visualizations, such as, visualization 406; one or more data sources, such as, data source 410; one or more visualization models, such as, visualization model 408; or one or more assessment engines, such as, assessment engine 412.

In one or more of the various embodiments, modeling engine 402 may be arranged to enable users to design one or more visualization models that may be provided to visualization engine 404. Accordingly, in one or more of the various embodiments, visualization engine 404 may be arranged to generate one or more visualizations based on the visualization models.

In one or more of the various embodiments, modeling engines may be arranged to access one or more data sources, such as, data source 410. In some embodiments, modeling engines may be arranged to include user interfaces that enable users to browse various data source information, data objects, or the like, to design visualization models that may be used to generate visualizations of the information stored in the data sources.

Accordingly, in some embodiments, visualization models may be designed to provide visualizations that include charts, plots, graphs, tables, graphics, styling, explanatory text, interactive elements, user interface features, or the like. In some embodiments, users may be provided a graphical user interface that enables them to interactively design visualization models
5 such that various elements or display objects in the visualization model may be associated with data from one or more data sources, such as, data source 410.

In one or more of the various embodiments, data sources, such as, data source 410 may include one or more of databases, data stores, file systems, or the like, that may be located locally or remotely. In some embodiments, data sources may be provided by another service
10 over a network. In some embodiments, there may be one or more components (not shown) that filter or otherwise provide management views or administrative access to the data in a data source.

In one or more of the various embodiments, visualization models may be stored in one or more data stores, such as, visualization model storage 408. In this example, for some
15 embodiments, visualization model storage 408 represents one or more databases, file systems, or the like, for storing, securing, or indexing visualization models.

In one or more of the various embodiments, visualization engines, such as, visualization engine 404 may be arranged to parse or otherwise interpret the visualization models and data from data sources to generate one or more visualizations that may be displayed to users.

20 In one or more of the various embodiments, assessment engines, such as, assessment engine 412 may be arranged to assess or otherwise evaluate marks in a visualization. Accordingly, in some embodiments, assessment engines may be arranged to automatically provide statistical explanations for the value of a specific data point (e.g., mark), even if the data for the explanation has been omitted from the visualization. In some embodiments, assessment
25 engines may be arranged to automatically pull in additional data from the data source associated with a visualization and evaluate its relevancy using statistical models, such as, assessment models. In one or more of the various embodiments, explanations about a mark may be conveyed to users as text strings and interactive visualizations, which be further explored.

In one or more of the various embodiments, assessment engines may enable users to
30 select one or more marks-of-interest from in visualization. Accordingly, in some embodiments,

visualization engines may be arranged to generate visualizations that include interactive user interfaces features that enable a user to select a mark-of-interest. For example, in one or more of the various embodiments, visualization engines may be arranged to include an assess-this-mark command in a right-click context menu. Thus, in some embodiments, users may right-click on a display object that represents the mark-of-interest to bring up a context menu and then select the assess-this-mark command from the context menu. In other embodiments, users may be enabled to search for marks-of-interest using names or labels associated with a mark.

FIGURE 5 illustrates a logical representation of a portion of visualization 500 in accordance with one or more of the various embodiments. As described above, visualization engines may be arranged to employ visualization models and data to generate visualizations, such as, visualization 500. In this example, visualization 500 represents a bar chart that shows sales revenue per day-of-week. One of ordinary skill in the art will appreciate that visualization models or visualization engines may be arranged to generate many different types of visualizations for various purposes depending on the design goals of users or organizations. Here, visualization 500 is presented as a non-limiting example to help provide clarity to the description of these innovations. One of ordinary skill in the art will appreciate that this example is at least sufficient to disclose the innovations herein and that visualization engines or visualization models may be arranged to generate many different visualizations for many different purposes in many domains.

In this example, visualization 500 includes mark 502 that represents the revenue earned on Sunday. Accordingly, in this example, mark 502 may appear to be an anomalous result given that it appears to be significantly lower than the other marks in visualization 500.

In this example, mark 502 may be determined to be a mark of interest because it may appear to anomalous compared to the other marks that may be associated with the revenue values for the other days-of-the-week. In some embodiments, users may be enabled to identify one or more marks-of-interest that seem interesting or anomalous. Also, in some embodiments, an assessment engine may be arranged to automatically identify one or more marks-of-interest based on automatically identifying marks that may be anomalies or statistical outliers.

In some embodiments, if a mark may be identified as a mark-of-interest, an assessment engine may be arranged to automatically perform one or more actions to analyze the mark-of-

interest to provide an explanation for the apparent discrepancies. In some cases, an analysis performed by the assessment engine may determine that the mark-of-interest may be within expectations rather than being an anomaly.

FIGURE 6 illustrates a logical representation of a portion of mark assessment system 600 in accordance with one or more of the various embodiments. In one or more of the various embodiments, system 600 may include one or more components, such as, assessment engine 602, assessment model 604, visualization model 606, data source 608, mark-of-interest 610, assessment result 612. In some embodiments, assessment results, such as, assessment result 612 may be arranged to include confidence score 614 or narrative 616.

In one or more of the various embodiments, assessment engine 602 may be arranged to assess mark-of-interest 610 based on assessment model 604. In one or more of the various embodiments, assessment models may be arranged to include one or more heuristics or machine-learning evaluators that may be executed to classify marks-of-interest.

As discussed herein, assessment engines may be arranged to employ one or more assessment models and provide one or more reports regarding how well a given assessment model matches (or classifies) a mark-of-interest. Accordingly, in this example, assessment result 612 includes a score, such as, confidence score 614 and natural language narrative 616.

In one or more of the various embodiments, assessment models may be arranged to provide a score that represents how well they explain the marks-of-interest. In some embodiments, assessment engines may be arranged to execute or apply assessment models to perform various evaluations of the mark-of-interests, visualization models, or data sources to classify the mark-of-interest. In some embodiments, confidence scores that represent how well the mark-of-interest fits the assessment model may be provided by the assessment model. For example, assessment model A may be arranged to execute ten tests or evaluate ten conditions that provide a score that includes ten points for each matched condition. Likewise, in some embodiments, assessment models may execute or apply one classifier that provides a confidence score.

In one or more of the various embodiments, different assessment models may employ different scoring criteria. Accordingly, in some embodiments, assessment engines may be arranged to weight or normalize confidence scores provided by different assessment models. In

some embodiments, the particular normalization rules or weight rules for normalizing or weighting assessment model confidence scores may be provided via configuration information.

Also, in one or more of the various embodiments, assessment models may be arranged to provide natural language narratives, such as, narrative 616. In some embodiments, natural language narratives may be employed in user interfaces or reports that may be provided to a user to explain the assessment of the marks-of-interest with respect a given assessment model. In some embodiments, narratives may be based on templates that enable labels, units, values, or the like, that may be associated with mark-of-interest or visualization model to be included in the user interfaces or report information.

In one or more of the various embodiments, assessment models may be designed or tailored to evaluate one or more statistical features of data associated with a mark-of-interest. Accordingly, in one or more of the various embodiments, assessment engines may be arranged to apply one or more assessment models to assess if the data associated with a mark-of-interest have the one or more of the statistical features targeted by an assessment model. In some embodiments, assessment models may be arranged to provide the confidence score as a form of a self-grade that represents how close the data associated with the mark-of-interest matches the statistical features the assessment model may be designed to match or otherwise evaluate.

In one or more of the various embodiments, one or more assessment models may focus on general, well-known, or commonplace statistical features that may be expected to be associated with marks-of-interest.

Also, in one or more of the various embodiments, one or more assessment models may be customized or directed to particular problem domains or business domains. For example, assessment models directed to financial information may be arranged differently than assessment models directed to employee information. Likewise, for example, assessment models directed to the automobile industry may be arranged differently than assessment models directed to the cruise (ship) industry. Further, in one or more of the various embodiments, one or more assessment models may be customized for particular data sources or visualization models for a particular organization or user. Accordingly, in one or more of the various embodiments, assessment models may be stored in data store that enables them to be configured independently from each other.

In one or more of the various embodiments, assessment engines may be arranged to generate or maintain profiles for one or more assessment models. In some embodiments, profiles may be arranged to track information that may be used for adapting assessment model results to particular organizations, users, problem domains, or the like.

- 5 Accordingly, in one or more of the various embodiments, assessment engines may be arranged to employ user activity information or user feedback to automatically build assessment model profiles that may be employed to modify or customize assessment reports. For example, if users of an organization consistently report mismatches between the marks-of-interest and assessment results, the assessment engine may be arranged to introduce weighting rules that
- 10 increase or decrease the effective confidence scores used for ranking assessment results for the organization based on the user feedback information.

- In one or more of the various embodiments, if a user selects a mark-of-interest in a visualization, the assessment engine may determine one or more assessment models and apply them to the mark-of-interest and its associated visualization model or data source. Accordingly,
- 15 non-limiting examples of assessment models are discussed below. For brevity and clarity this discussion is limited to a few examples, however, one of ordinary skill in the art will appreciate that other assessment models that may incorporate other or additional assessment strategies are contemplated.

- NUMBER OF ROWS: An assessment model may be arranged to determine if there may
- 20 be an unusually high or low number of data values or data records rolled up into the value represented by the mark-of-interest as compared to other marks in the visualization. For example, in some cases, if a value for a mark-of-interest seems high compared other marks it might be explained because the number of underlying records contributing the mark-of-interest may be greater than other marks, such as, Chinese GDP being high because there are many
- 25 people in China.

- MEASURE OF ROWS: An assessment model may be enabled to determine if the average value in the rows aggregated to the mark-of-interest is high as compared to other marks in the visualization. For example, in some cases, if a value for a mark-of-interest seems high compared other marks it might be explained because the underlying records contributing the
- 30 mark-of-interest may have above average values compared to the underlying records for other

marks, such as, the US GDP is high because people in the US earn more money than people in other countries

UN-VISUALIZED DIMENSION: An assessment model may be arranged to evaluate if the data records rolled up into the mark-of-interest are different from other rows in some dimension that is not included in visualization seen by the users. For example, store X has a high average order value because it sells diamonds, and most other stores do not sell diamonds. And, other stores that sell diamonds also have a high average order values.

VISUALIZED DIMENSION: An assessment model may be arranged to evaluate the if data records rolled up into the mark-of-interest are different from other data records in some dimension that is included in the visualization. For example, store X in 2015 had high total sales, but this is part of the overall trend in sales by store rather than geographic region being relevant to difference between the mark-of-interest and other marks.

ONE OUTLIER: An assessment model may be arranged to determine if there may be a single outlier in the data records associated with the mark-of-interest that may lead to mark-of-interest looking different than other marks in the visualization. Accordingly, in some embodiments, this may suggest that the one data record may be worth investigation further. For example, region X had high sales on 2018-04-28 because outlier order Y occurred on that day.

MULTIPLE OUTLIERS: An assessment model may be arranged to determine if there may be multiple outliers that may collectively cause the mark-of-interest to appear different than other marks in the visualization. This result may suggest a missing latent factor. For example, region X had high sales on 2018-04-28 because orders Y1, Y2, and Y3 were all on that day in that region. Such a result may be worth further investigation.

INTERMEDIATE AGGREGATE: An assessment model may be arranged to determine if a sub-set of data records associated with the mark-of-interest value produce values that may be inconsistent or improbable across the entire set of data records or as compared to other sub-sets of data records. For example, an assessment model may report that region X had high sales on 2018-04-28 because store Y is in region X, and it had high sales that day, but there were no individual orders at store Y that dominate the sales.

FIGURE 7 illustrates a logical representation of a portion of user interface 700 for analyzing marks in visualizations based on data characteristics in accordance with one or more of the various embodiments. In some embodiments, user interface 700 may be arranged to include one or more panels, such as, panel 702, panel 704, panel 706, panel 708, explanations
5 710, or the like.

in this example, for one or more of the various embodiments, panel 702 may be arranged to display a rank order list of the explanations for the mark-of-interest that was analyzed by the assessment engine. In some embodiments, each explanation item may include a natural language narrative that may provide a potential explanation of the values associated with the mark-of-
10 interest as determined by the assessment engine. Likewise, in some embodiments, if the assessment engine may be unable to determine a potential explanation, panel 702 may be arranged to display a narrative to that effect.

In one or more of the various embodiments, the text for the explanation narratives may be based on templates such that the field names or labels that may be associated with the mark or
15 the explanation for the mark may be included in the explanation narrative.

In one or more of the various embodiments, panel 704 may be arranged to include additional explanations or visualizations that may be related to the mark that was analyzed by the assessment engine. In some embodiments, panels, such as, panel 706 may include additional explanation narrative related to the mark or the one or more visualizations that may help explain
20 the mark to users. Also, in some embodiments, panels, such as, panel 708 may include visualizations that help explain the mark. For example, panel 708 may include one or more alternate visualizations that help explain the reasons why a mark appeared anomalous.

In one or more of the various embodiments, assessment engines, visualization engines, or the like, may be arranged to use specific visualization models or explanation text/narrative
25 templates for different assessment models. In some embodiments, the visualization models or explanation/narrative templates for an assessment model may be defined in configuration information. Accordingly, the explanations/narrative text or visualizations included in user interface 700 may be tailored for a particular organization, locale, or the like.

In one or more of the various embodiments, assessment engines or visualization engines
30 may be arranged to determine some or all of the content or styling for user interface 700 via

configuration information. Accordingly, in some embodiments, the narrative explanation text or templates, panel layout, visualizations, or the like, may be configured differently depending on the needs of the organization, users, or the like.

Generalized Operations

5 FIGURES 8-12 represent generalized operations for analyzing marks in visualizations based on data characteristics in accordance with one or more of the various embodiments. In one or more of the various embodiments, processes 800, 900, 1000, 1100, and 1200 described in conjunction with FIGURES 8-12 may be implemented by or executed by one or more processors on a single network computer (or network monitoring computer), such as network computer 300
10 of FIGURE 3. In other embodiments, these processes, or portions thereof, may be implemented by or executed on a plurality of network computers, such as network computer 300 of FIGURE 3. In yet other embodiments, these processes, or portions thereof, may be implemented by or executed on one or more virtualized computers, such as, those in a cloud-based environment. However, embodiments are not so limited and various combinations of network computers,
15 client computers, or the like may be utilized. Further, in one or more of the various embodiments, the processes described in conjunction with FIGURES 8-12 may be used for analyzing marks in visualizations based on data characteristics in accordance with at least one of the various embodiments or architectures such as those described in conjunction with FIGURES 4-7. Further, in one or more of the various embodiments, some or all of the actions performed by
20 processes 800, 900, 1000, 1100, and 1200 may be executed in part by assessment engine 322, visualization engine 324, modeling engine 326 one or more processors of one or more network computers.

FIGURE 8 illustrates an overview flowchart for process 800 for analyzing marks in visualizations based on data characteristics in accordance with one or more of the various
25 embodiments. After a start block, at block 802, in one or more of the various embodiments, a visualization engine may be arranged to generate one or more visualizations based on one or more visualization models or data sources. As described above, a visualization system may be arranged to include, one or more modeling engines, one or more data sources, one or more visualization engines, or the like, that may be arranged to generate visualizations based on one or
30 more visualization models and data provided by the one or more data sources.

At block 804, in one or more of the various embodiments, an assessment engine may be arranged to determine one or more marks of interest in the one or more visualizations. In some embodiments, the marks of interest may be selected by a user.

In one or more of the various embodiments, assessment engines may be arranged to
5 automatically identify one or more marks that may appear inconsistent with other marks in visualization. In some embodiments, assessment engines may be arranged to execute actions to perform one or more heuristic evaluations to determine if one or more marks may be identified as marks-of-interest. In some embodiments, assessment engines may be arranged to highlight or otherwise indicate in the visualization user interface which marks may be marks-of-interest. The
10 particular styling for indicating a potential mark-of-interest may vary depending on the organization, as well as, configuration information.

In some embodiments, marks-of-interest automatically determined by the assessment engine may be considered potential marks-of-interest. Accordingly, in some embodiments, users may be enabled to confirm if a potential mark-of-interest should be assessed by the assessment
15 engine.

At block 806, in one or more of the various embodiments, the assessment engine may be arranged to analyze the one or more marks of interest based one or more assessment models. In one or more of the various embodiments, assessment engines may be arranged to execute one or more assessment models to assess the marks-of-interest to determine explanations that may
20 provide insight into which how the characteristics of underlying data/data-source contributed to the marks-of-interest.

At block 808, in one or more of the various embodiments, the assessment engine may be arranged to generate one or more assessment reports that may include one or more visualizations or one or more explanation narratives associated with the one or more marks. As described
25 above, in some embodiments, assessment reports may include interactive user interfaces that enable users to review the assessment results that may include explanation narratives that may describe one or more characteristics of the marks-of-interest.

Next, in one or more of the various embodiments, control may be returned to a calling process.

FIGURE 9 illustrates a flowchart of process 900 for analyzing marks using assessment models in accordance with one or more of the various embodiments. After a start block, at block 902, in one or more of the various embodiments, an assessment engine may be one or more assessment models and one or more marks of interest. As described above, in some

5 embodiments, one or more assessment models may be defined for assessing marks. Accordingly, in some embodiments, the assessment engine may be arranged to obtain one or more assessment models from a data store, or the like. In some embodiments, assessment engines may be arranged to determine the particular assessment models based on rules, conditions, or the like, that may be provided via configuration information.

10 At block 904, in one or more of the various embodiments, the assessment engine may be arranged to evaluate the one or more assessment models based on the marks, visualization models, data sources, or the like. In one or more of the various embodiments, assessment models may include or be associated rules, conditions, computer readable instructions, or the like, that an assessment engine may execute or apply to assessment a mark.

15 At decision block 906, in one or more of the various embodiments, if there may be more assessment models to evaluate, control may loop back to block 904; otherwise, control may flow to block 908. In some embodiments, assessment engines may be arranged to end the assessment early by omitting or skipping one or more assessment models. For example, in some

20 embodiments, assessment engines may be configured stop evaluating additional assessment models if a previously evaluated assessment model may be considered to provide a sufficient explanation for the mark-of-interest.

 Accordingly, in one or more of the various embodiments, assessment engines may be arranged to associate a confidence score with assessment results provided by an assessment model. In some embodiments, assessment models may be arranged to provide a confidence score

25 that indicates how well the assessment model matches or explains the mark.

 Also, in one or more of the various embodiments, assessment engines may be arranged to maintain another set of confidence scores that may be combined with the confidence score provided by the assessment model. Thus, in some embodiments, the overall confidence score of an assessment of a mark may be a combination of a score provided by the assessment model and

30 a score provided by the assessment engine.

For example, in some embodiments, the confidence score provided by the assessment engine may be generated based on other factors or metrics that may be unavailable or irrelevant to an assessment model. In some embodiments, such factors or metrics may be based on user feedback, organization requirements (e.g., configuration information), metrics associated with
5 organization or community activity, or the like.

At block 908, in one or more of the various embodiments, the assessment engine may be arranged to filter and rank the results provided by the assessment models. In one or more of the various embodiments, assessment models may be arranged to provide one or more confidence scores that represents a self-assessment of how well an explanation provided by the model
10 matches the mark.

Also, in one or more of the various embodiments, assessment engines may be arranged to modify confidence scores provided by assessment models to provide a final confidence score.

Accordingly, in one or more of the various embodiments, assessment engines may be arranged to eliminate one or more explanations based on filtering out one or more assessment
15 models based on rules, conditions, or the like. For example, in some embodiments, an assessment engine may be configured automatically exclude explanations that are associated with confidence scores that fall below a defined threshold value.

Also, in one or more of the various embodiments, other metrics or features of one or more of the explanations, visualization models, data sources, data, users, organizations, or the
20 like, may be incorporated into conditions, rules, or the like, that may be executed to filter out one or more assessments or explanations. For example, in some embodiments, filters may include conditions associated with one or more user roles. Accordingly, for example, a filter may be configured to be more inclusive for some user roles or less inclusive for other user roles. Thus, in this example, the number of explanations included in an assessment report may be increased
25 for data scientists or reduced for regular users.

In some embodiments, other metrics or features used for filtering may be based on various sources, including, organization preferences, user preferences, localization, user input, other configuration information, or the like.

Similarly, in one or more of the various embodiments, assessment engines may be arranged to rank order the one or more explanations based on various conditions, rules, metrics, or features. For example, in some embodiments, explanations may be rank ordered based on confidence scores, or the like. Further, for example, in some embodiments, one or more
5 explanations may be rank ordered based on user preferences, organization preferences, or the like.

Similar to filters, in some embodiments, the various conditions, rules, metrics, or features used for rank ordering explanations may be based on various sources, including, organization preferences, user preferences, localization, user input, other configuration information, or the
10 like.

At block 910, in one or more of the various embodiments, the assessment engine may be arranged to generate an assessment report. In one or more of the various embodiments, assessment reports may include a variety of information, including the rank ordered list of explanations, descriptions of one or more characteristics of the mark-of-interest that was
15 assessed, visualizations that show the mark-of-interest in different contexts or views, or the like.

Next, in one or more of the various embodiments, control may be returned to a calling process.

FIGURE 10 illustrates a flowchart of process 1000 for adapting filtering or ranking of assessment models based on monitoring user interactions in accordance with one or more of the
20 various embodiments. After a start block, at block 1002, in one or more of the various embodiments, an assessment engine may be arranged to generate an assessment report based on an automatic evaluation of one or more marks or interest. As described above, assessment engines may be arranged to provide information for generating interactive reports that include one or more user interface features that enable users to interact with the explanations.

Also, in one or more of the various embodiments, interactive assessment reports may include one or more user interface components that may enable users to provide feedback information to grade the explanations. For example, in some embodiments, an interactive assessment report may enable users to evaluate the explanations or visualizations using various grading systems, such as, helpful/not-helpful, rating scales, (e.g., 1-100, stars, letter grades, or
30 the like), natural language narratives, or the like.

At block 1004, in one or more of the various embodiments, the assessment engine may be arranged to monitor one or more user interactions or user feedback that may be associated with the assessment report. As mentioned above, interactive assessment reports may be arranged to include components or features for directly collecting user feedback. For example, in some
5 embodiments, users may be enabled to grade the usefulness or correctness of one or more explanations. Accordingly, in one or more of the various embodiments, assessment engines may monitor or record the feedback associated with various portions of assessment reports. In some embodiments, assessment engines may be arranged to enabled users to provide an overall grade for entire assessment reports.

10 Also, in one or more of the various embodiments, assessment engines may be arranged to passively monitor how one or more users interact with assessment reports. In some embodiments, monitoring various user interaction metrics may provide additional insights into the performance or quality of the assessment that may be provided by various assessment
15 models. For example, in some embodiments, if the monitored interactions shows that users are doing things like dismissing one or more explanations, re-sorting the explanation lists, spending a lot of time reviewing lower ranked explanations, clicking on help screens, running several mark assessments over short period, or the like, it may indicate that the explanation narratives or the assessment models may not be explaining marks satisfactorily. In contrast, for example,
20 monitored information that shows that users are accepting the explanations without performing additional confirmation actions or reviews, it may indicate that the assessment reports may be satisfactory.

At block 1006, in one or more of the various embodiments, the assessment engine may be arranged to provide of modify confidence scores based on the monitored interactions or feedback. In one or more of the various embodiments, assessment engines may be arranged to
25 execute one or more formulas to generate a confidence scores that indicate how well assessment models may be performing.

Accordingly, in one or more of the various embodiments, assessment engines may be arranged to increase or decrease assessment model confidence scores based on user feedback or observed user activity. In some embodiments, the impact of various types of feedback or
30 observed activity may vary depending on the type of feedback or observed activity.

In one or more of the various embodiments, assessment engines may be arranged to provide or modify confidence scores based on one or more rules or actions that may be defined in configuration information.

At decision block 1008, in one or more of the various embodiments, if the confidence scores for one or more assessment models fall below a defined threshold, control may flow block 1010; otherwise, control may flow to block 1012.

At block 1010, in one or more of the various embodiments, the assessment engine may be arranged to discard or re-train one or more low scoring assessment models. In some embodiments, assessment engines may be arranged to exclude one or more low scoring assessment models from use for particular organizations. Also, in some embodiments, assessment engines may be arranged to globally exclude one or more assessment models from use by some or all other organizations based on confidence scores that may be associated with one or more assessment models.

In one or more of the various embodiments, rather than discarding poor scoring assessment models, assessment engines may be arranged to indicate that one or more poor scoring assessment models may require re-training or re-certification before they may be employed for assessing marks. Accordingly, in some embodiments, assessment engines may be arranged to generate notifications or include re-training or re-certification information in assessment reports.

At block 1012, in one or more of the various embodiments, the assessment engine may be arranged to employ the confidence scores for filtering or ranking subsequently generated mark assessments or explanations. In one or more of the various embodiments, assessment engines may be arranged to enable organizations or users to define threshold values for including or excluding assessment models or assessment model results from assessment reports. Next, in one or more of the various embodiments, control may be returned to a calling process.

FIGURE 11 illustrates a flowchart of process 1100 for generating expected value distributions to for assessment models in accordance with one or more of the various embodiments. After a start block, at block 1102, in one or more of the various embodiments, an assessment engine may be provided an assessment model and a mark-of-interest. As described above, assessment engines may be arranged to evaluate an ensemble of two or more assessment

models to determine which one may explain the appearance of the mark-of-interest.

Accordingly, in some embodiments, assessment engines may be arranged to evaluate each assessment model in the ensemble of models to rank them based on how well the mark-of-interest may be explained.

5 At block 1104, in one or more of the various embodiments, the assessment engine may be incrementally provided some or all of the data that may be directly displayed in the visualization. In some embodiments, the visualization that may be displayed to the user may be comprised of one or more portions of the data source. For example, in some embodiments, a visualization may have been generated based on a portion of the columns or fields that may be
10 available in a data source. Accordingly, in some embodiments, columns or fields that may be directly included in the visualization may be determined.

 In one or more of the various embodiments, assessment engines may be arranged to incrementally add data included in the visualization to an evaluation data set. For example, in some embodiments, if the displayed visualization may be based on five columns from the data
15 source, assessment engines may be arranged to evaluate assessment models based on two of the visualized columns, then three of the visualized columns, and so on.

 In some embodiments, the assessment engine may be arranged to provide the data portions (e.g., columns or fields) in various sequences, such as, the order they appear in the visualization, the order they appear in the data source, random order, one or more sorted orders,
20 or the like. Accordingly, in some embodiments, assessment engines may be arranged to determine the order data portions may be added to a data set based on rules, or the like, included in configuration information.

 At block 1106, in one or more of the various embodiments, the assessment engine may be arranged to employ the one or more assessment models and the current data set to provide
25 probability distributions for the mark-of-interest. Accordingly, in some embodiments, assessment models may employ data fields or columns in the current data set to attempt to predict a value for the mark-of-interest based on the current data set.

 Accordingly, in one or more of the various embodiments, models may be employed to generate probability distributions that may predict a value for the mark-of-interest each time
30 additional data may be added to the evaluation data set.

At decision block 1108, in one or more of the various embodiments, if the displayed data may be arranged to exhausted, control may flow to block 1110; otherwise, control may loop back block 1104. For example, if the visualization associated with the mark-of-interest included five columns, after each of the assessment models in the ensemble have generated predicted value distributions for the mark-of-interest for each combination of data portions associated with the visualization. Note, assessment engines may be arranged to employ configuration information to determine the number, order, or type of combinations of the data portions used to predict mark-of-interest values. Also, in some embodiments, while most examples described herein involve assessment models that produce probability distributions other types of models or predictors are anticipated, including models that produce a single value and an error range. Also, in one or more of the various embodiments, generally it is anticipated that assessment engines may be arranged to employ various predictive models as long as they may predict a value for the mark-of-interest given a data set selected from the data source.

At block 1110, in one or more of the various embodiments, the assessment engines may be arranged to incrementally modify the model data set to data not included in the visualization. In some embodiments, data sources may include additional data (e.g., columns or fields) that were not included in the visualization. In some cases, the information in other columns or fields may explain the appearance of the mark in the visualization. Accordingly, similar to as described for block 1104, these data portions may be incrementally added to the assessment data set.

At block 1112, in one or more of the various embodiments, the assessment engine may be arranged to employ the assessment models to generate one or more probability distributions based on the incrementally modified data set. Accordingly, similar to as described for block 1106, assessment engines may be arranged to employ the assessment data set and assessment models to predict values for the mark-of-interest.

At decision block 1114, in one or more of the various embodiments, if the remainder of the data portions associated with the mark-of-interest have been exhausted, control may be returned to a calling process; otherwise, control may loop back to block 1110. See, the description for decision block 1108 for additional details.

FIGURE 12 illustrates a flowchart of process 1200 for comparing the results of assessment models in accordance with one or more of the various embodiments. After a start

block, at block 1202, in one or more of the various embodiments, assessment engines may be arranged to generate various expected value distributions as described above.

At block 1204, in one or more of the various embodiments, the assessment engines may be arranged to generate model performance scores based on the amount of information provided by a given models best predictions. In some embodiments, the amount information provided by different models may be compared by computing the inherent randomness (e.g., entropy) of the probability distributions produced by a given model. Accordingly, in one or more of the various embodiments, assessment engines may be arranged to employ entropy to compare disparate models. In some embodiments, the particular mechanism for computing entropy for a given model and its results may vary depending on the model types. However, one of ordinary skill in the art may be expected to appreciate the assessment engines may be arranged to execute one or more conventional or well-known information theory computations to estimate to produce an entropy measurement for a given probability distribution. In some embodiments, assessment engines may be arranged to employ one or more rules, instructions, or the like, defined in configuration information to determine entropy values for a given probability distribution or assessment model. Accordingly, in some embodiments, assessment engines may be arranged to employ one or more unconventional or custom methods for computing entropy.

In some embodiments, the amount of information provided by an assessment model prediction may be represented or measured using conventional or well-known information theory units, such as, bits, nats, dits, or the like, that may enable the amount information included in a prediction to be expressed. Accordingly, in some embodiments, the amount of information provided by a different models or distributions may be compared.

At block 1206, in one or more of the various embodiments, the assessment engines may be arranged to generate model complexity scores for each assessment model based on the complexity associated with the provided predictions. Information theory provides various mechanisms for measuring complexity of predictive models that may be employed to fairly compare assessment models. In some embodiments, the complexity may be considered a measurement of how difficult it is for a model to produce a given result. Accordingly, in some embodiments, complexity may be measured based on a variety of factors including, the number of inputs, number of parameters, or the like.

Also, in one or more of the various embodiments, a complexity score may reflect how complex the proposed explanation for the mark-of-interest. For example, in some embodiments, a proposal that suggests a mark-of-interest depends strongly on one other data portion (e.g., column) may be considered less complex than a proposal that depends on two or more other data portions.

In one or more of the various embodiments, assessment engines may be arranged to employ rules, instructions, or the like that may be provided by configuration information to compute a measure of complexity for assessment models. Generally, in some embodiments, these may include well-known or convention information theoretic methods as well as custom or proprietary methods.

At block 1208, in one or more of the various embodiments, the assessment engines may be arranged to provide an overall model score for each assessment model based on associated information scores and complexity scores. In one or more of the various embodiments, the overall score may be arranged to select the assessment model that provided the most information with the least complexity. For example, in some embodiments, information scores (see, block 1204) and complexity scores (see, block 1206) may be normalized or mapped to a 0-100 scale and summed to provide an overall score for an assessment model. However, in some embodiments, assessment engines may be arranged to employ rules, instructions, or the like, provided via configuration information to compute overall scores based on information scores or complexity scores.

At block 1210, in one or more of the various embodiments, assessment engines may be arranged to modify one or more overall model scores based on one or more defined weights, heuristics, user preferences, or the like. For example, as described above, user feedback may be employed to weight the overall score of one or more assessment models. Likewise, in one or more of the various embodiments, one or more assessment models may be observed to produce erroneous results in some conditions, over fitting in scarce data applications, or the like. In one or more of the various embodiments, assessment engines may be arranged to modify overall score for certain assessment models in particular circumstances. Accordingly, in one or more of the various embodiments, assessment engines may be arranged to employ one or more rules,

instructions, or the like, provided via configuration information to determine if overall model scores should be modified.

At block 1212, in one or more of the various embodiments, the assessment engines may be arranged to rank order the assessment models based on the modified overall scores.

5 Next, in one or more of the various embodiments, control may be returned to a calling process.

It will be understood that each block in each flowchart illustration, and combinations of blocks in each flowchart illustration, can be implemented by computer program instructions. These program instructions may be provided to a processor to produce a machine, such that the
10 instructions, which execute on the processor, create means for implementing the actions specified in each flowchart block or blocks. The computer program instructions may be executed by a processor to cause a series of operational steps to be performed by the processor to produce a computer-implemented process such that the instructions, which execute on the processor, provide steps for implementing the actions specified in each flowchart block or blocks. The
15 computer program instructions may also cause at least some of the operational steps shown in the blocks of each flowchart to be performed in parallel. Moreover, some of the steps may also be performed across more than one processor, such as might arise in a multi-processor computer system. In addition, one or more blocks or combinations of blocks in each flowchart illustration may also be performed concurrently with other blocks or combinations of blocks, or even in a
20 different sequence than illustrated without departing from the scope or spirit of the invention.

Accordingly, each block in each flowchart illustration supports combinations of means for performing the specified actions, combinations of steps for performing the specified actions and program instruction means for performing the specified actions. It will also be understood that each block in each flowchart illustration, and combinations of blocks in each flowchart
25 illustration, can be implemented by special purpose hardware-based systems, which perform the specified actions or steps, or combinations of special purpose hardware and computer instructions. The foregoing example should not be construed as limiting or exhaustive, but rather, an illustrative use case to show an implementation of at least one of the various embodiments of the invention.

- Further, in one or more embodiments (not shown in the figures), the logic in the illustrative flowcharts may be executed using an embedded logic hardware device instead of a CPU, such as, an Application Specific Integrated Circuit (ASIC), Field Programmable Gate Array (FPGA), Programmable Array Logic (PAL), or the like, or combination thereof. The
- 5 embedded logic hardware device may directly execute its embedded logic to perform actions. In one or more embodiments, a microcontroller may be arranged to directly execute its own embedded logic to perform actions and access its own internal memory and its own external Input and Output Interfaces (e.g., hardware pins or wireless transceivers) to perform actions, such as System On a Chip (SOC), or the like.