Quick Access to Relevant Business Insights

Adopting visual data discovery, business organizations improve their ability to find the information they need when they need it and do so more productively than other companies. According to a study conducted recently, business managers in organizations that use visual data discovery tools are 28 percent more likely to find timely information than those who rely solely on managed reporting and dashboards. Moreover, 48 percent of business intelligence users at companies that use visual data discovery are able to find the information they need without the help of IT staff all or most of the time.

Determine patterns in business operations

Data visualization enables users to see interesting and previously unknown patterns – like, for example, being able to picture the relationship between business and operations – and then related performance measures. In fact, with data visua-lization, it is easier to see how day-to-day job impacts the overall business performance, and find if any operational changes caused an increase/decrease in business performance.

Today, the amount of customer and market information that organizations are able to gather is overwhelming. However, to ensure organizations actually derive key insights from all this data, things need to be simplified. Data visualization tools allow decision makers to quickly visually identify any changes in customer behavior or market conditions and make the necessary adjustments. The identified business pattern helps the decision makers to easily identify the reason for growth or dip in the organization's performance and address it.

Rapid Identification of Latest Trends

In this age, the volume of data that companies are able to gather about customers and market conditions can provide business leaders with insights into new revenue and business opportunities, presuming they can spot the opportunities in the mountain of data. Using data visualization, decision makers are able to grasp shifts in customer behaviors and market conditions across multiple data sets much more quickly.

Accurate Customer Sentiment Analysis

Using data visualization, companies can attain a deeper dive into customer sentiment and other data, which reveals emerging opportunities for them to launch new services to their customers. These useful insights enable the enterprises to act on new business opportunities for staying ahead of their rivals.

Direct Interaction with Data

Data Visualization also helps the companies to manipulate and interact with their data in a direct manner. One of the greatest strengths of data visualization is how it brings actionable insights to the surface. Unlike one-dimensional tables and charts that can only be viewed, data visualization tools enable users to interact with data.

Geo-Spatial Visualization

Another scope of data visualization that has emerged in the business world lately is the geo-spatial visualization. The popularity of geo-spatial visualization has occurred due to a lot of websites providing web-services, attracting visitors' interest. This type of business needs to take advantage of location specific information, which is already present in the system in the form of customer's zip code, providing better daily analysis experience. This type of visualization adds a new dimension to the figures and helps in better understanding of the matter.

Predictive Sales Analysis

With the help of real-time data-visualization, sales executives can carry out advanced predictive analytics for their sales figures, viewing up-to-date sales figures and see why certain products are underperforming and the reasons that sales are lagging. For example, discounts offered by competitors may be one of those reasons.

Drill-Down Sales Analysis

Using heat map data-visualization, business executives can illustrate which product groups are performing well or underperforming, and drill down into the data to determine the factors that are shaping sales. For example, the data might reveal that pet-care products are underperforming, but that higher-income customers represent the majority of sales. These insights could be used to target promotions to this customer segment to increase conversion rates and revenue growth for this category.

Easy Comprehension of Data

Utilizing data-visualization, companies may approach huge data and makes it easily comprehensible, be it the field of entertainment, current affairs, financial issues or political affairs. It also builds in them a deep insight, prompting them to take a good decision and an immediate business action if needed.

Customized Data-Visualization

A yet another vital advantage of Data-visualization is that it not only provides graphical representation of data but also allows changing the form, omitting what is not required, and filter more to get further details. This is a great eye catcher and attracts business executive's attention better and provides better communication. Additionally, it provides a great advantage over traditional methods of presenting data.

1.3 DATA VISUALIZATION TOOLS

There is more number of commercial and non-commercial data visualization tools available in the market. Some of the popular data visualization tools in use are Tableau, Qlikview, Sisense, Looker, Google Data Studio, Zoho Analytics, Fusioncharts, Highcharts, Datawrapper, Klipfolio, Kibana, Chartio, Plotly, Infogram, Visme, Geckoboard, AnyChart, D3.js, Microsoft PowerBI, IBM Watson Analytics and SAP Analytics Cloud.

The features, advantages and disadvantages few important tools in the market are elaborated below.

I.3. I Google Data Studio

The data visualization tool, **Google Data Studio** which can be accessed freely using the Google Account. Google Data Studio is a data visualization tool used to create effective data reports from data sources like Google Analytics, Google Sheets, Google Ads, Google Search Console, YouTube, and MySQL. Google Data Studio has templates to set up reports and dash boards quickly and easily.

The youngest tool and a part of Google's analytics solutions - Google Data Studio. Being relatively new to the field, it strives to take its position among many competitors via ease of usage, simple yet beautiful design, innovative problem-solving and straightforward, habitual ways to share dashboards (just as sharing documents). It is a fully web-based solution, and there is no desktop version.

Google aspires to hit the right spot on the market with not just going for a single BI tool, but also promoting all their other tools for working with data by conveniently combining them into the Google Analytics Solutions data toolkit, a software suite for analyzing data and facilitating data-driven solutions.

Google Data Studio allows for the transformation of raw data to present it in interactive visualizations that will be compiled into dashboards. In addition, the tool is perfectly accommodated for use with Google specific data sources. It provides easy access to the data through the convenient facility of data connectors.

Finally, one of the best parts concerns the collaboration techniques that are used in Google Data Studio, bringing the team of developers to work together on a single problem. Data Studio, allows others to view and edit the dashboard in the same way as in Google Docs.

Unique Features

- Connectors to Google Data Sources
- Transformation tools for working with raw data

- Decent library of built-in visual types
- Great teamwork capabilities

Sharing of reports is straightforward and functions analogously to Google Drive. The control of access levels also works similarly,provides facility to send invitations to access a report or a folder of reports via either email, or a shareable link, and choose to either grant permission to view only or allow editing.

Google Data Studio's overall capabilities are still limited. Compared to other tools, it is in short of creating interactivity of reports, customizing visuals and calculating functions.

Advantages

- Developed as part of the Google Analytics suite, fully integrated with other relevant Google products
- Simple in all major aspects, easy to use
- Great collaboration capabilities

Disadvantages

- Not flexible as its competitors
- Less ability to add custom visuals, only to modify the existing ones to some extent
- Interactivity is not supported
- No functionality for mixing and blending data
- The data must be ready for visualization, only minor changes can be done to it

I.3.2 Tableau

It is a popular and market-leading data visualization tool used to visualize and analyze the data in an easily digestible format. It is an extremely powerful tool that focuses on business intelligence and analysis, utilized by thousands of companies worldwide. It allows working with live data-sets and spends more time on data analysis. It has a very large customer base across industries due to its simplicity and ability to produces interactive visualization. It is particularly well suited to handling the huge and very fast changing data-sets which are used in big data operations, including artificial intelligence and machine learning applications.

Tableau has different licensing plans similar to other BI solutions. Tableau offers three distinct products with drastically different prices. The three products are the Tableau desktop, Tableau online, and Tableau server that come at different price points.

It supports integration with large number of advanced variety of data sources with many systematic types, such as data systems organized in file formats (CSV, JSON, XML, MS Excel, etc.), relational and non-relational data systems (PostgreSQL, MySQL, SQL Server, MongoDB, etc.), cloud systems (AWS, Oracle Cloud, Google BigQuery, Microsoft Azure).

The core distinction from competitors is that Tableau has a special feature of Data Blending. Another unique feature is the ability for collaboration in real time that makes it a valuable investment for commercial and non-commercial organizations alike. There are several ways to share the reports in Tableau, by publishing them to a Tableau server, via email Tableau Reader capability, by publishing Tableau workbook openly and giving access to anyone who has a link. This magnitude of options enables great flexibility and removes many restrictions.

Tableau offers a broad variety of visualization capabilities with distinct features, enabling smart ways of data discovery and deep insight. The rich library of visualizations types includes *word clouds* and *bubble charts* that provide high levels of comprehension unique to Tableau. Tree diagram and Treemap provide contextual information to the visuals. The latter is usually utilized for the depiction of parts categorical data, focusing attention on the most relevant pieces of the information.

Tableau dashboards are amazingly flexible. Its central features allow the remarkable ability to layout the dashboard in the desired way with any overlaps, which comes really handy in the screen space ergonomics.

Tableau is easy to apprehend as a working tool, its learning curve is pretty gentle, as it strives to provide all of its powers to any kind of users, even those who haven't been previously exposed to technical details of visualization workflows. This objective is accomplished by using intuitive interface, everything is always no more than two clicks away, robust filters and drill-downs are easy to find and use, operations are well documented and labeled.

Unique Features

- Complimentary sharing ability (with certain limitations)
- Support for connection to 30+ data source types
- Mixing data sources
- Support for cubes

Advantages

- Intuitive and attractive user interface
- Seamless integration with big data platforms, from Hadoop to Google BigQuery
- Provides an extensive roaster of native data connections allowing easy integration with data from many resources

8

- Responsiveness- supported by mobile platforms
- Powerful community collaboration
- Constant development new updates are regularly released and are easy to install

Disadvantages

- Initial data preparation is required (structured data)
- Although great for analytical purposes, Tableau and other BI tools cannot replace financial reporting applications
- No concept of versioning with Tableau server

1.3.3 Qlikview

It is another data visualization tool which is a major player in the market and Tableau's biggest competitor. The key advantage of the tool is highly customizable and wide range of features. In addition to its data visualization capabilities, Qlikview offers powerful business intelligence analytics and enterprise reporting capabilities and clean and clutter-free user interface.QlikView is rated as one of the most expensive platforms in the BI field.

QlikView is a solution that focuses on the user as the receiver of data. It allows users to explore and discover your data in a workflow similar to the way developers work when processing data. To sustain flexibility in its approach to data exploration and visualization, this software strives to maintain the association between data. This facilitates the discovery of your data by the end-user seeking a certain piece of data with awareness of retrieval of any relevant items, in spite of any circumstances, even if the origins of the items applicable to the search are incredibly disjoint.

QlikView is incredibly flexible. It allows setting and tweaking every little aspect of each object and customizing the look and feel of any visualizations and dashboards. With such great deal of flexibility, there also comes an incorporated ETL (Extract, Transform, Load) Engine that enables you to conduct the ordinary data cleansing operations. However, it may turn out to be costly.

Unique Features

- Uniqueness and flexibility
- Rich set of features for creating advanced dashboards
- Manipulate data associations automatically
- Allows faster queries and quicker data exploration by keeping data in-memory

Advantages

- Attractive user interface
- · Easy to set up filtering for any kind of visuals
- Fast rendering of both graphs and tables
- Ability to mail reports in the convenient form of PDF

Disadvantages

- Unintentional combining of some data aspects while filtering
- No ability to union bookmark results together
- Complications in using it as an enterprise tool

1.3.4 Power BI

Power BI is the software solution, developed and supported by Microsoft, for business intelligence and analytics needs. At the core of Power BI is an online service with various options for interaction, also featuring several outlets for connection to data provided by thirdparty software and services.

Power BI provides a simple web-based interface with lotof useful features varying from customizable visualization to certainly limited controls of data sources. The desktop application expands the available functionality to an even larger extent with the addition of tools for data cleansing and normalization. Another way to work and make data-driven decisions on the go is through the mobile app, which is available for multiple platforms. It is also amazingly simple to share insights by publishing your work to Power BI service and forming lively dashboards from a combination of reports which makes the data communication centralized and easy to follow up for all the participants. Power BI is concise and minimalistic, yet powerful and robust. However, like any other software, it also has its ups and downs which must be carefully considered.

First of all, as it is a Microsoft product, it follows a philosophy, principles, and architecture similar to other major Microsoft products. It also exposes a familiar interface for the Windows users.Power BI was created and designed with the aim to build upon the functionalities of MS Excel, upgrade it to the next level, extend its operability even further to unlock new use cases, cover more platforms, and reach out to the cloud.

As a Microsoft product, Power BI has connections to some other software from the Microsoft's toolbelt but goes much farther than that by utilizing a whole suite of novel business analytics tools. Thus, Power BI is not just related to other products; it is tightly integrated with the main Microsoft tools including MS Excel, Azure Cloud Service, and SQL Server.

Unique Features

- Power BI has a free basic version, giving users a chance to explore it first
- It supports plenty of ways to incorporate or import your data (streaming data, cloud services, excel spreadsheets, and third-party connections)
- It has interactive dashboards with real-time feed of data
- Simple API for integrating Power BI with your applications
- Different ways to share reports and dashboards
- Multiplatform support (Web, Desktop, Mobile)

Advantages

- More affordable than other tools and offers a free version
- Integration with other Microsoft products, Azure, Excel, SQL Server
- The built-in library of visuals is impressive and robust
- Ability to connect almost any kind of data source

Disadvantages

- Unnecessary complexity
- Performance issues when streaming and importing big data sets
- Lack of data preparation and cleaning tools

I.4 FEATURES

1.4.1 Common features of Data Visualization Tool

Data Visualization tool helps enterprises, organizations and companies to display data in the structured and ordered format, which is not only easy to interpret but meaningful and receptive tomaking decisions. It identifies patterns, limits noise and insignificant values from the data to produce actionable insights.So, in order to make most out of the Data Visualization, companies need to select the right tool with variety of features and capabilities. Some of the features that a data visualization tool should have are discussed below. The following are a variety of features and capabilities that experts recommend organizations consider when adopting visualization tools for big data:

Clear and Customizable Dashboard

The dashboard could easily be called as the important feature of a data visualization tool. Just like one look at the car's dashboard which gives all the vital information that is in need such as speed, indicator, light, seat belt, fuel etc. Similarly, a visualization dashboard should be able to present all the key information in a glance. A sample dashboard prepared using Microsoft Power BI visualization tool is shown in figure 1.1.



Fig. 1.1 Sample Dashboard prepared using Microsoft Power BI visualization tool

A good data visualization dashboard should be a few things simultaneously. For starters, it should look great. It needs to be clear, with pops of color amid adequate whitespace. Too white is boring and too much color is overwhelming. The dashboard should strike a balance.

The dashboard should be able to accurately summarize all the data that matters. The top Key Process Indicators (KPI) you are trying to track, the vital trends you are monitoring or any other dataset that is pivotal to your business should be clearly visualized on the dashboard in a way that you have the general overview within a few seconds of launching the dashboard. Facts presented on the dashboard should be clear and decipherable in a glance.

Another very important quality a dashboard must possess is customizability. At any given time, your company may be tracking dozens of different datasets. You should have the power to customize which datasets appear prominently on the dashboard. Different teams have different priorities and hence, the data visualization tool must allow complete customizability.

Embeddability

The ability to seamlessly integrate the visual reports into any other applications in use is important to really utilize the power of data visualization. For the teams to work efficiently, collaborate better and share across different platforms, the data visualization software must have the feature to allow using the various media like graphs and charts into different mobile or web applications. The quality and richness of the visual reports should not diminish when transporting it onto a different application. The reports should still remain interactive and allow further investigation of the data.

Not all departments need to analyze all the high-level data your tool collects. Most of them would only want one part of that data to seamlessly integrate with their specific applications. They need immediate actionable insights that can help them increase the efficiency of their tasks and campaigns. A good data visualization tool must allow for easy embeddability.

Performance

If visualization tools for big data distract workers from the flow of their work, they're less likely to be used. A few seconds delay may not be significant for some use cases but may discourage users tasked with evaluating hundreds of decisions throughout the day. Features that help improve performance include prompts, data optimization settings and dynamic loading options.

Another performance-related feature to consider is the tool's ability to run computations on GPUs. As data sets have grown, rendering large amounts of data with traditional architectures has become harder. GPUs used with direct memory access can help crunch large volumes of data faster and more efficiently. This makes it easier to build high-definition visualizations on the server side that simply get served by the application via a web application.

Interactive Reporting

The visual reports generated by a data visualization tool must be extremely interactive, allowing easy investigation into trends and insights. Interactive data visualization helps identify trends and tell a story through data. This includes capabilities for filtering, slicing and dicing, and drilling up and down at speeds that make it possible for users to investigate huge volumes of data and get answers to their questions immediately.

Data analysts and decision makers need to be able to collate data from various sources and combine datasets to produce insightful reports. The tool should allow for the reports to be viewed in various different formats and different parts can be highlighted at different times. Industry specific KPIs need to be customized to provide tailored insights. To enable all of this, the business intelligence and data visualization tool needs to be highly interactive.

Data Collection and Sharing

Importing raw data to the visualization tool and then exporting the visual reports in various different forms is something that needs to be managed by the company in the way it likes. Some datasets can be fed into the tool in their rawest form while others will need to be aggregated first because they are too large. Sometimes, data can be taken from just one source while sometime it needs to be collected from different sources and visualized by the tool. Also the tool must provide the facility to share the reports to team members and other stakeholders. The reports must be exportable to other applications.

Geo-tagging and Location Intelligence

The globalization of business demands location intelligence in the data visualization tool. Where is the data coming from? Which states or regions are more actively using the services and which areas need more work? The ability to layer sets of data chronologically and spatially is important for businesses that need to track locationbased KPIs.

Collaboration

Real-time collaboration capabilities in visualization tools for big data allow employees to have more meaningful conversations about their discoveries. This includes the ability for employees to collaborate in real time on current data, rather than requiring them to send static files and screenshots to one another.

Streaming data support

Enterprises are now faced with wrangling massive volumes of complex, streaming data from a variety of different sources. Many visualization tools use legacy back ends based on structured batch data analysis. This makes it difficult to analyze extreme data in real time. Support for streaming data can allow more visualization use cases involving data from social media, internet of things devices and mobile applications.

Artificial Intelligence Integration

Visualization tools for big data are starting to experiment with machine learning, deep learning and natural language processing to make it easier to analyze, explore, predict and prescribe actions.

1.4.2 Salient features of popular data visualization tools

Tableau - Tableau has for a long time now, been hailed as one of the best data visualization tools out there. Their clientele includes giants like LinkedIn, Deloitte, Lufthansa and PepsiCo. Some of its best features are,

- Customizable dashboards that is embeddable with applications like Salesforce, SharePoint and Jive etc.
- Real-time interactive dashboards for filter on demand and click to dig deeper
- Plenty of data connections with live and in-memory data
- Secure collaboration
- Mobile optimized

Qlikview- Qlikview is probably the strongest competitor to Tableau. It was in fact chosen as the Gartner Magic Quadrant Leader 2019 and boasts clients like Conde Nast, Subaru and Global retail Bank. Some of its best features include,

- Embedded Analytics
- Advanced Analytics Integration with third party engines like Python
- Customizable Dashboard
- Predictive Analysis
- Shared file management

Sisense - Sisense is more than a traditional analytics tool. It is scalable and can handle all sorts of data. With high profile clients including NASA, NASDAQ, Samsung and Comcast, Sisense is definitely one of the best. Top futures include,

- Customizable dashboard with sharing, drag and drop and built-in chart widgets
- In-memory columnar database can crunch terabytes of data on a single server
- Extremely fast implementation
- Advanced machine learning and AI
- Instant insights that update in real-time
- Interactive and automatic scheduled reporting

Domo - Domo is not just a data visualization tool but a complete business management platform that handles your analytics and reporting from just one platform, with clients ranging from eBay to National Geographic and Sage. Its features include,

- Hundreds of data connectors including Facebook, Salesforce etc.
- Workbench functionality imports on-premise data into Domo easily
- Easily cleans, combines and transforms data in multiple ways
- Easy data sharing with custom tools
- Mobile optimized with automatic alerts
- Automatic schedule reporting and customizable dashboard

Microsoft Power BI -Coming from Microsoft gives Power BI a familiarity that makes it easy for new entrants to adopt and explore. To add to this ease of adoption, Power BI offers a free basic version and is open-source. With clients like Adobe, HP and Toshiba, it offers features like,

- Interactive dashboard with real-time data feed and easy sharing
- Customized reports that can be created from scratch
- Easy data capture and sharing with Datasets
- Explore data by asking questions in natural language
- Cloud based and easy to implement

Klipfolio - With over 500 data sources it can connect to including Google Analytics, Twitter and Moz, Klipfolio is a great choice indeed. Top features include,

- Widespread data sourcing
- Financial forecasting
- Customizable dashboard with built in templates
- Real-time accuracy

Plotly -One of the most colorful, yet elegant BI solutions out there, Plotly helps create interactive graphs for easy comprehension. Some of its top features are,

- 2D and 3D charts with designer input and customizability
- Integration with analytics oriented languages like Python, R and Matlab
- User friendly with inbuilt APIs

Chartio - Chartio is a BI and data visualization tool for all businesses big and small. Some features include,

- Real-time analytics with live changes
- Comparative analytics
- Easy set-up
- Multiple chart formats

Geckoboard - With over 80 pre-built services for real-time analysis, Geckoboard make data visualization easy for anyone. Some of its best features are,

- Custom dashboards with pre-built widgets
- Rich integrations with APIs for Facebook, twitter, Salesforceetc
- Pull and push data integrations
- Customizable style sheets, schema and widgets

Datawrapper - Datawrapper's simple, clear and easy to use interface has quickly made it a top choice among non-technical clients like media organizations such as Fortune, Mother Jones and The Times. Some of its best features are,

- Easy to use with no coding or design skills required
- Fast and interactive charts
- Styled to your branding

To summarize, choosing the right data visualization tool is a big decision not only because they are fairly expensive, but also because they play a huge role in shaping the business strategy. A tool that can present the most clear, interactive and accurate visual reports can help business people to take better decisions, make better plans and track

16

the KPI's better. So depending on what features matter most to the business, choose a tool that will give just the representations that are in need.

1.5 DATA ACCESS FROM DATA SOURCES

Data visualization tools provide excellent facility to connect to variety of data sources and fetch data to prepare reports. For example, Google Data Studio is a data visualization tool used to create effective data reports from data sources like Google Analytics, Google Sheets, Google Ads, Google Search Console, YouTube, and MySQL. A sample screenshot of Data Studio is given in fig 1.2 to understand the data sources supported by it.

CWNED BY ME SH	ARED WITH ME TRASH		Q Search	
	Earlier	Owner	Last Modulieit	
REPORTS	[Sample] Google Analytics Data	Google Data Budio	Dec 13, 2016	j.
DATA SOURCES	[Sample] YouTube Data	Google Date Studio	Nev 29, 2016	3
	[Sample] Firebase Analytics Data (Us	Google Data Mudio	Nov 4, 2016	3
Video totorials Learn by watching!	[Sample] Firebase Analytics Data (Ev	Google Data Illutio	Nov 4, 2010	1
	[Sample] AdWords Data	Google Data Studio	Sep 22, 2016	â
	[Sample] Rio Olympics Data	Roogle Date Studia	Sep 22, 2016	1
	[Sample] World Population Data 2005	Google Data Studia	Aug 5, 2016	3

Fig. 1.2 Screenshot of Google Data Studio Data Sources

The tool, Tableau can connect to all the popular data sources which are widely used. Tableau's native connectors shown in figure 1.3can connect to the following types of data sources.

- File Systems such as CSV, Excel, etc.
- **Relational Systems** such as Oracle, Sql Server, DB2, etc.
- Cloud Systems such as Windows Azure, Google BigQuery, etc.
- Other Sources using ODBC

Connect Live –In Tool Tableau, the Connect Live feature is used for real-time data analysis. In this case, Tableau connects to real-time data source and keeps reading the data. Thus, the result of the analysis is up to the second, and the latest changes are reflected in the result. However, on the downside, it burdens the source system as it has to keep sending the data to Tableau.



Fig. 1.3 Sample screenshot of Tableau's native connect

In-Memory - Tableau can also process data in-memory by caching them in memory and not being connected to the source anymore while analyzing the data. Of course, there will be a limit to the amount of data cached depending on the availability of memory. **Combine Data Sources -** Tableau can connect to different data sources at the same time. For example, in a single workbook, connecting to a flat file and a relational source by defining multiple connections is possible. This is used in data blending, which is a very unique feature in Tableau.

REVIEW QUESTIONS

- 1. Write the benefits of Data Visualization tools.
- 2. What is Geo-Spatial Visualization?

- 3. Differentiate predictive sales analysis and drill-down sales analysis.
- 4. List out popular data visualization tools and write the unique features of any one tool.
- 5. Pen down the advantages and disadvantages of Google Data Studio.
- 6. What is Tableau?
- 7. Describe the unique features of Tableau, Qlik View and Power BI.
- 8. Elaborate the Geo-tagging and Location Intelligence.
- 9. Differentiate Plotly and chartio.
- 10. Explain Klipfolio.
- 11. Explain the accessign of Google Data Studion using free google account.
- 12. Elucidate the salient features of Tableau tool.
- 13. Describe the tool Qlik view.
- 14. Brief the Power BI online service.
- 15. Explain the clear and customizabe dashboard.
- 16. Elaborate the steps to fetch data and prepare reports using data visualization tools.
- 17. Write the steps to create a report in Google Data Studio
- 18. Describe the data access from data sources with an example.
- 19. Compare the unique features of Google Data Studio, Tableau and qlik view.
- 20. Create a sample report for world population data using Google Data Studio.

CHAPTER II

2.1 DATA TRANSFORMATION

Data comes in many forms such as text, numerical, images and videos. For example, a customer details form where few fields are not filled and left empty. Such data are known as missing data. In most of the cases, data may be missing data, unstructured data, or data that lacks regular structure. In data visualization, before processing the data, there is a need of cleaning data to make it fit to process further.

Data cleansing has a long history in databases and is a key step known as *extract, transform, load* (ETL), commonly used in data warehouses shown in figure 2.1, where data is extracted from one or more sources; transformed into its proper format and structure, including cleansing of the data; and finally loaded into a final target location, such as a single database or file which can be used for business analytics &data visualization.



Fig. 2.1 ETL Process

2.1.1 Extraction, Transformation and Load (ETL)

Extraction

The first step of the ETL process is extraction. In this step, data from various source systems is extracted which can be in various formats like relational databases, No SQL, XML and flat files into the staging area. It is important to extract the data from various source systems and store it into the staging area first and not directly into the data warehouse because the extracted data is in various formats and can be corrupted

also. Hence loading it directly into the data warehouse may damage it and rollback will be much more difficult. Therefore, this is one of the most important steps of ETL process.

Transformation

The second step of the ETL process is transformation. In this step, a set of rules or functions are applied on the extracted data to convert it into a single standard format. It may involve following processes/tasks:

- Filtering loading only certain attributes into the data warehouse.
- Cleaning filling up the NULL values with some default values, mapping U.S.A, United States and America into USA, etc.
- Joining joining multiple attributes into one.
- Splitting splitting a single attribute into multiple attributes.
- Sorting sorting tuples on the basis of some attribute (generally key-attribute).

Loading

The third and final step of the ETL process is loading. In this step, the transformed data is finally loaded into the data warehouse. Sometimes the data is updated by loading into the data warehouse very frequently and sometimes it is done after longer but regular intervals. The rate and period of loading solely depends on the requirements and varies from system to system.

2.1.2 Messy data

Data sets large and small are rarely ready to use. As figure 2.1 shows, simple commaseparated value (CSV) data set has a variety of issues, including invalid fields, missing and additional values, and other issues. It cannot be used directly for further processing.



Figure 2.1 Simple messy data set example

22

This example is a simple one, but anyone who has worked with a public data set will understand these issues and the need to preprocess data to make it useful. Data sets that have such obvious errors make the results of the processed data somewhat questionable. The observations with errors result in incomplete data or invalid observations that can lead to incorrect results. Cleansing data is therefore a key step in the data processing pipeline.

Data may also come from multiple sources. Although each source may be valid in isolation, bringing the data together may require processing for consistency and uniformity. For example, one data set may have a different unit of measure for a given field than another, requiring that they be normalized. One key factor for data validity, then, is the format in which the data is represented.

2.1.3 Data formats and schemas

Data sets can be in many forms, but the majority is stored as delimited text files. As shown in fig 2.1, these data sets delimit their fields by using a character, commonly a comma, but in other cases through white space (space, tab, etc.). These raw data sets are particularly prone to error because they lack any information that indicates their structure and so require data scientists to interpret the data set manually.

So-called "self-describing formats" can greatly improve our ability to maintain data correctly. These formats include XML and JSON. These data formats allow the data to be embedded within metadata to make it fully self-describing within a single file. They also permit complex data formats that are more difficult to describe with simple flat text files (such as variant arrays of data or relationships within the data). Figure 2.2 shows the representation of temperature data by using the JSON format. Here the data is labeled, and the labels are predefined such that the ingest tool understands what to expect.

```
[
required: [Sensor, Temperature, DateTime],
properties: [
Sensor: { type: string },
Temperature: [ type: real },
DateTime: [ type: string ]
}
[Sensor : A1,
Temperature : 45.2,
DateTime : 2015-11-25T19:17:53
],
[Sensor : A2,
Temperature : 47.9,
DateTime : 2015-11-25T19:19:06"
}
```

Figure 2.2 JSON format to self-describe a data set

2.1.4 Data blending or fusion

Data blending is the process by which a data set is constructed from two or more independent data sets. Blending data may not be a one-time process; instead, it can be performed on demand based on the machine learning use case.For example, users in a marketing department might blend data from a CRM system and a spreadsheet with product profitability information. They could then quickly see which products not only make the most money, but also attract the most customer purchasing interest.

Blending data has all the problems namely, the need to cleanse more than one data source. Fusing multiple data sets has additional problems, however, in the representation of the data in each source (such as one data set that uses Celsius and another that uses Kelvin, as illustrated in figure 2.3). The data may not be consistent across sources and may require transforming and reordering data fields so that the fused data can be properly used.



Blended dataset

Figure 2.3 Blending and transforming two data sets

Some data blending tools may not preserve all the data detail when combining datasets. For example, data visualization software may do blending by simply aggregating data. In this case, users will get rapid views and summary information from the combined data. However, more in-depth data exploration may not be possible. Users may not be able to ask ad-hoc questions, which in turn could limit creativity and innovation.

2.1.5 Methods for data cleansing

Data cleansing begins with *data parsing*, which means taking each observation from its data file and extracting each independent element. The parsing can be easily identified if the records are similar such as same number of elements, similar types, etc.

Schema, a higher-level representation of the data observations, we can typecheck the observation to ensure that it matches the schema and the user's expectation for later data analysis. For example, we can ensure that a number is contained at a given field location instead of a string given that you intend to perform numerical operations on it. A schema informs whether the proper number of fields is represented for each observation.

Some data-cleansing applications permit the construction of rules with functions that permit more complex transformations of data. For example, interrogating fields to create or modify other fields based on their contents. The rules can also validate the consistency of an observation to remove invalid data or to transform data for greater accuracy. For example, modifying the U.S. ZIP code from five digits to the enhanced nine digits. We can also identify duplicates, although there are applications for duplicate observations in a data set so that duplicate elimination isn't always required or necessary. When a data set is syntactically correct, we can apply methods to ensure that the data is semantically correct.

2.1.6 Data profiling

When the data is clean, the next step is to profile the data as a secondary step in the cleansing process. *Profiling* is an analysis of the data to ensure that the data is consistent. Through profiling, we can dig into the data to see the distribution of the individual fields to look for outliers and other data that doesn't match the general data set. For example, Figure 2.4 illustrates this process.



Figure 2.4 Data set errors made visible through data profiling

In line 1, given that the real values represent physical measurements, a zero value may indicate an issue with this observation. In line 3, you see that the range of the measurement is obviously not in the same range as other measurements of this field (and its type differs). Finally, in line 5, notice that the class name is misspelled. In some cases, these issues can be detected automatically through profiling. We could indicate that all measurements should be greater than 0 to catch the first issue. Through

statistical analysis, we could identify the second outlier measurement. The final issue could be identified by capturing the unique class names and through their frequencies understand that this particular class name is an outlier (likely an occurrence of one). We can validate time-series data in the context of flow to ensure that the data is processed in the correct order given timestamps.

2.1.7 Open source data-cleansing tools

There are many open source data-cleansing tools in the market. One interesting example is called *Drake*, which performs data cleansing for text-based data by using a workflow approach that automatically handles dependencies in the available data and the commands to cleanse them. It supports multiple input and output files and has a similar operation as the make utility in the context of managing dependencies.

The Data Cleaner tool is a framework and data-profiling engine that exposes an API and allows user-defined extensions for data cleansing. Data Cleaner supports multiple input and output formats, with the ability to create rules for data quality over the data.

2.2 BAR CHART

Bar charts involve rectangular blocks of varying heights, and the height of the block corresponds to the value of the quantity being represented. The vertical axis shows the values – for example, the total number of each type of object counted and the horizontal axis shows the categories. In case of counting the different types of vehicles in a parking lot, the individual blocks could represent cars, vans, motorcycles and jeeps, and their heights could represent the count of each vehicle

In other words, a bar chart uses horizontal or vertical bars to show comparisons among categories. The longer the bar, the greater the value it represents. In the bar chart, an axis of the chart shows the specific categories (dimensions) which is being compared and the other axis represents a discrete value (metric).

The bars can represent pretty much anything that can fit into categories, though, or even the values of the same quantity at different points in time. The height of the bar could also represent a wide range of things, including counts, total revenues, percentages, frequencies or values in any unit of measurement (e.g., heights, speeds or masses). Bar graphs are incredibly versatile, so anybody dealing with data will undoubtedly use them often.

2.2.1 Types of bar chart

Types of bar chart supported by various visualization tools are,

- Vertical bar chart / Column chart
- Stacked vertical bar chart / Stacked column chart

- 100% Stacked vertical bar chart / 100% Stacked column chart
- Waterfall chart
- Horizontal bar chart
- Stacked horizontal bar chart
- 100% Stacked horizontal bar chart

Bar charts, also known as column charts, use vertical or horizontal bars to represent data along both an x-axis and a y-axis visually. Each bar represents one value. When the bars are stacked next to one another, the viewer can compare the different bars, or values, at a glance.

For example, a bar chart might show how smartphone use has changed over time. Along the vertical axis, or axis Y, the maker of the graph would plot a quantitative or numerical scale such as smartphone users by the millions. On the horizontal axis, or axis X, the graph maker might plot a category, such as years from 2009 to 2019. In this way, viewers can easily see how many millions of people started using smartphones during each of those years and whether that number steadily increased or decreased over time.

2.2.1 Bar chart example

The bar charts below show 2 different views of Google Analytics web traffic data. The base dimension for both charts is *Medium*. The left hand chart uses stacked bars to show several metrics (*Sessions, User* and *Exits*) for medium. The right hand chart uses a second dimension, *Country ISO Code*, to breakdown each medium according to its country of origin. Instead of stacked bars, this chart use grouped bars: each bar is a data series corresponding to one of the countries. Two dimensional charts can only plot a single metric (*Sessions* again in this example).



Figure 2.5 Bar chart example

2.2.4 Bar chartAdvantages

- Its simplicity makes a bar graph a good choice to represent data across to large groups of people.
- It is their ability to represent data that shows changes over time, which helps people visualize trends.
- Along with more complicated types of graph, the simple bar graph can present many different types of data clearly and concisely.
- That historical context can lead to a greater understanding of the data and why it is important. For example, a bar graph could be useful for people who want to show how McDonald's preferences have changed over time.
- The bar chart is the one that gets the key pieces of information across in the most readable and digestible format, without sacrificing accuracy.
- It is in widespread use everywhere from textbooks to newspapers, most audiences understand how to read a bar graph and can grasp the information the graph conveys.
- Other types of graphs, such as those with compressed scales, matrix graphs or MTF charts, are difficult to read for someone who isn't already familiar with that type of data visualization.

2.2.5 Steps to create a Bar chart in Google Data Studio

- Type datastudio.google.com in the url address bar in the browser and signin using your google account
- Data Studio overview page will be shown.
- Click the '+' box with title Start a new report as shown below.

🛫 Data Studio Home				:: 0	9	0
Start a new	report		ALL TOMPLATES	c.		
i.		Colorana Bar Bar - D Colorana				
ALL CHARTER ME INVESTIGATE	riet Habi	and trees	langto tay	l facili		4
	(mark)	See.	Law spectra in			
Carl Street B	El Downsteines DavidouarDe (2015)	weap free,	11114	÷		
D DITANDUCIN	Bitty Opublication	1044934	22226.004	E		
di totuntera	Provinces of Marci	and a	Last speed by two			
	GA Budence Overvew with Itels Control	Amp) sets to re-	Aur 12 3311	1		
🗂 Product Districts 🧉	Province Moderate	1000	And an interaction			
10 Angert Sellers 3	Alt Requiring Overview with Data Control	Surgital and South State	Apr 16, 2211	E.		
P. Connet in Sale (Carey of eCommerce specifi	Ajacimat	Air SIGHT	1		
	Copy of elignments speed	Add terrain	Apr 16 2019	F	Y	+
O New Texture of	Communit speed	Dime Solution)	Air 1C2014	E		~

Figure 2.6 Data studio starting page

A new Untitled Report page will be opened. Click File → New report



Figure 2.7 A blank report page in Data studio

- From the right side pane select data source. (for eg. [Sample] World Population Data 2..)
- Click "add a chart" in the toolbar. Data Studio makes it easy to compare chart types with some handy illustrations as shown below.



Figure 2.8 Add a chart menu item in Data studio

• Choose anyone chart type under "Bar". It may be column chart / horizontal chart, etc., once it appears on the report page, the right-hand pane will change with Data and Style options.



Figure 2.9 Stacked column bar chart for world population data

- By default, the dimension is "Year", Data Studio will automatically select a metric (eg. 'population', what's displayed on the Y axis). Metric and dimension can be changed. For instance, female%. Internet users% etc.
- In the style tab, how many bars and how many country details can be changed.
- To see the finished product, click "View" in the top corner. This transitions you from Editor to Viewer mode. To edit the report Click "Edit."
- To finish up, we need to give the report a name. Double-click the title (right now it's "Untitled Report") to change it as 'World Population Barchart'



Figure 2.10 Naming the report in data studio

2.2.6 Advanced options for Bar chart in Google Data Studio

There are many advanced option in the data visualization tool, Google Data Studio to effectively present the bar chart. Those options are,

- Drilling Down gives viewers a way to reveal additional levels of detail within a chart
- Breakdown dimension displays the metric data broken down according to the selected dimension.
- Date range dimension used as the basis for limiting the date range of the chart. For example, this is the dimension used if you set a date range property for the chart, or if a viewer of the report uses a date range control to limit the time frame.

2.3 Pie chart

Pie charts are extensively used in presentations and offices. Pie Charts help show proportions and percentages between categories, by dividing a circle into proportional segments. Each arc length represents a proportion of each category, while the full circle represents the total sum of all the data, equal to 100%.Pie Charts are ideal for giving the reader a quick idea of the proportional distribution of the data. Comparing a given category (one slice) within the total of a single pie chart, then it can often be more effective. However the major disadvantages to pie charts are:

- They cannot show more than a few values, because as the number of values shown increases, the size of each segment/slice becomes smaller. This makes them unsuitable for large amounts of data.
- They take up more space than their alternatives, like stacked bar chart for example. Mainly due to their size and for the usual need for a legend.
- They are not great for making accurate comparisons between groups of pie charts. This being that it is harder to distinguish the size of items via area when it is for length.

An example of bar chart and pie chart is shown in fig 2.11 which visualizes of the sales from a fictitious fruit stand.



Figure 2.11 Bar chart and Pie Chart

In the bar chart, it's easy to see the relative sales values of each fruit. In the pie chart, legends on the right to be referred. To know the sales of apple, in the bar chart, we have to do some mental math summing all the non-apple sales. But with the pie chart it's immediately obvious, even without looking at the value in the legend, that apples do in fact make up more than half of sales.

If the goal for the visualization is to convey the sales amount of each product, the bar chart is the better selection. But if the point of the visualization isn't to know the precise value of each of the products, but instead to bring home the point that apples are more than half of your business, then the pie chart is a more powerful visualization.

2.3.1 BAR CHART VS PIE CHART

Bar charts and pie charts are very common data visualization tools, but it is important to use them correctly to ensure you convey clear and concise information. Most data visualization tools make it very easy to plot the data as either a bar chart or a pie chart. According to the goal of visualization i.e. what to visualize, we have to select the chart. Because of the pitfalls of pie charts, bar charts tend to be the better choice. However, there are times when the pie chart is actually the better, more powerful visualization. Additionally, there are many instances where it's easy for the human eye to tell the slice values approximately enough for the purpose of the visualization. In that case the choice between the bar chart and the pie chart is purely a matter of preference. As a rule of thumb:

- Use either a bar chart or a pie chart when comparing parts of a whole, the categories are few (up to four), and it's easy for the human eye to estimate their value when presented as slices (as when the values are close to 25%, 50% or 75%)
- Use a pie chart when there is a specific and clear point related to the share of the total that we are trying to get across and the individual values of each slice is not important
- Use a bar chart otherwise

2.3.2 Steps to create a Pie chart in Google Data Studio

- Type datastudio.google.com in the url address bar in the browser and sign-in using your google account
- Data Studio overview page will be shown.
- Click the '+' box with title Start a new report as shown below.

🚦 Data Studio 🛛 Home					1	6
Start a ne	w report		ALL TEMPLATES	C		
-	+	Enderson Britten B				
ALL CHRISTIAN INCOME IN	Source Resident		Sector Sector), Secon		4
🖬 mene (Devectories Davidouards (30.0)	and fail	122794	1		
C DEALERS (eller Opsträumete	Apartola	TTTE MA	E		
🖬 tanunis 💳 😦	Providence, C. Mitania	(mark)	Last union () a re-			
	GA Audience Overvew with Itels Control	Anappi take traine	Aut 12.3370			
🗂 Product Diservices 🧉	Promising the bags	and a	Sale generation.			
III Angert Sellers 1	🚮 Add Augustitus Derregen with Data Control	Security Days States	-lar 14, 2214	÷.		
21 Connet to Date +	Capy of sConvenience speciel	(Add Enter)	Air 2010	1		
	Copy of all ormination speeds	Add Second	Apr 16 2019	Ē		Ŧ
O New Perchant	aCommunic speed	Directory)	Air XC2014	E		-

Figure 2.12 Data studio starting page

• A new Untitled Report page will be opened. Click File→New report



Figure 2.12A A blank report page in Data studio

- From the right side pane select data source. (for eg. [Sample] World Population Data 2..)
- Click "add a chart" in the toolbar. Data Studio makes it easy to compare chart types with some handy illustrations as shown below.



Figure 2.13 Add a chart menu item in Data studio

• Choose anyone chart type under "Pie". There are 2 types: Pie chart and Donut chart. Select any one. Once it appears on the report page, the right-hand pane will change with Data and Style options.



Figure 2.14 Pie chart for world population data



Figure 2.15 Donut chart for world population data with country as dimension and population as metric

- By default, the dimension is "Year", Data Studio will automatically select a metric (eg. 'population', what's displayed on the Y axis). Metric and dimension can be changed. For instance, female%. Internet users % etc.
- In the style tab, how many bars and how many country details can be changed.
- To see the finished product, click "View" in the top corner. This transitions you from Editor to Viewer mode. To edit the report Click "Edit."
- To finish up, we need to give the report a name. Double-click the title (right now it's "Untitled Report") to change it as 'World Population pie chart'

2.4 DATA TABLES

Data tables display the data in a grid of rows and columns. Each column represents a dimension or metric, while each row is one record of the data. Tables automatically summarize the data. Each row in the table displays the summary for each unique combination of the dimensions included in the table definition. Each metric in the table is summarized according to the aggregation type for that metric (sum, average, count, etc.). For example, in Google Data Studio, table can have up to 10 dimensions and 20 metrics.

A data table which presents sales data for a fictional pet store is shown in Table 2.1. The store sells items for dogs, cats, and birds, with several products in each category.

Date	ltem	Category	Qty Sold
10/1/2016	Happy Cat Catnip	Cat	1
10/1/2016	Healthy Dog Food	Dog	3
10/1/2016	Pretty Bird Seed	Bird	5
10/2/2016	Pretty Bird Seed	Bird	3
10/2/2016	Happy Cat Catnip	Cat	2
10/3/2016	Playful Puppy Toy	Dog	6
10/5/2016	Pretty Bird Seed	Bird	7

Table 2.2 shows just the category dimension and quantity metric for table 2.1. It has aggregated the quantities sold per category. Since there are only 3 categories in the data set, the table shows just 3 rows.

Category	Qty Sold
Bird	28
Dog	27
Cat	12

Table 2.3 contains 6 rows, 1 for each item. The quantity sold metric is now aggregated per item.

Category	ltem	Qty Sold
Bird	Pretty Bird Bird Seed	20
Dog	Healthy Dog Dog Food	17
Dog	Playful Puppy Toy	10
Bird	Parrot Perch	8
Cat	Happy Cat Catnip	4
Cat	Hungry Kitty Cat Food	3

2.5 SCATTER CHART

Scatter charts can be used to look for relationships between variables. These charts show the data as points or circles on a graph using X (left to right) and Y (top to bottom) axes. Scatter charts can include a trend line that shows how the variables in the chart are related. They tend to be more frequently used in scientific fields. Though infrequent, there are use cases for scatter charts in the business world as well.

For example, to manage bus fleet, we have to understand the relationship between miles driven and cost per mile. Thescatterplot may look something like in figure 2.16.



Figure 2.16 Sample scatter chart

To focus primarily on those cases where cost per mile isabove average, a slightly modified scatter chart designed as given in figure 2.17.



Figure 2.17 Scatter chart with conditions

From the figure 2.17, cost per mile ishigher than average when less than about 1,700 miles or more than about 3,300 miles observations can be made.

2.5.1 Steps to create a Scatter /Bubble chart in Google Data Studio

- Type datastudio.google.com in the url address bar in the browser and sign-in using your google account
- Data Studio overview page will be shown.
- Click the '+' box with title Start a new report as shown below.

🗧 Data Studio 🛛 Home					0	0
Start a ne	w report		ALL YEMPLATE	• C		
Ref		Constanting Barness (P)				
ALL CHARTENAL INVESTIG	Singa Anayko Sa Tirint Hukun		lange ten	Q Seria	_	4
	1 miles	See.	Last space if by real			
Carl House I	Dovertances Davidouardia (36.5)	way hay	4.11194	÷		
D DEALDREN !	(Birg Orathinanda	494 (1934)	22216.004	E.		
di terune - :	President C 4000	Dest.	Last speed by two			
	GA Audience Overvew with Itels Control	Anapie take to be	Aur 12,3319	1		
🗂 Product Districts 📫	Province Shi hand	Dare .	Linuinitaria	_		
80 Annort Sellers 18	Sill Requirement Overveew with Data Control	Service States States	Apr: 14, 2211	Ē.		
P. Connet in Easte 🛛 🤞	Carey of eCommenta spece	April Control	AirTEIT	1		
	Copy of eCommunic speed	Add Named	April 2019	E		(H)
O New Testhered						-

Figure 2.18 Data studio starting page

A new Untitled Report page will be opened. Click File → New report



Figure 2.19 A blank report page in Data studio

- From the right side pane select data source. (for eg. [Sample] World Population Data 2..)
- Click "add a chart" in the toolbar. Data Studio makes it easy to compare chart types with some handy illustrations as shown below.



Figure 2.20 Add a chart menu item in Data studio

• Choose anyone chart type under "Scatter". There are 2 types: Scatter chart and Bubble chart. Select any one. Once it appears on the report page, the right-hand pane will change with Data and Style options.



Figure 2.21 Scatter chart for world population data


Figure 2.22 Bubble chart for world population data with country as dimension and population as metric

- By default, the dimension is "Year", Data Studio will automatically select a metric (eg. 'population', what's displayed on the Y axis). Metric and dimension can be changed. For instance, female%, Internet users % etc.
- In the style tab, how many bars and how many country details can be changed.
- To see the finished product, click "View" in the top corner. This transitions you from Editor to Viewer mode. To edit the report Click "Edit."
- To finish up, we need to give the report a name. Double-click the title (right now it's "Untitled Report") to change it as 'World Population Scatter Chart".

REVIEW QUESTIONS

- 1. Explain Transformation process in ETL.
- 2. Elaborate messy data
- 3. What is self -describing format?
- 4. Describe data blending.
- 5. Brief the data parsing method in data cleaning
- 6. Elucidate data profiling with an example.
- 7. List out the advanced options for Barchart in Google Data Studio.
- 8. Write the advantages of barchart.
- 9. Explain Data tables with an example.
- 10. Elaborate Scatter chat with an eample.

- 11. Describe the ETL process with a neat sketch.
- 12. Discuss the methods in data cleansing with necessary diagram.
- 13. List out the types of barchat and write the procedute to draw a bar chart for Google Analytics Web Traffic Data.
- 14. How to create a bar chart in google data studio.
- 15. Differentiate Barchart and Pie chart.
- 16. Elaborate the steps to create a pie chart in Google Data Studio.
- 17. How to create a Scatter or bubble chart in Google Data Studio.
- 18. Write the steps to create a time series chart in Google Data Studio.

CHAPTER III

3.1 TIME SERIES CHART

3.1.1 Introduction

Time series forecasting is a critical requirement for many organizations. The starting point of forecasting is a time series visualization, which provides the flexibility to reflect on historical data and analyze trends and seasonal components. It also helps to compare multiple dimensions over time, spot trends and identify seasonal patterns in the data. A few examples include stock market analysis, population trend analysis using a census, or sales and profit trends over time.

Time series analysis is a statistical technique used to record and analyze data points over a period of time, such as daily, monthly, yearly, etc. A time series chart is the graphical representation of the time series data across the interval period. A time series chart, also called as times series graph or time series plot, is a data visualization tool. Each point on the chart corresponds to both a time and a quantity that is being measured. A sample time series chart shown in fig 3.1.





Generally, the horizontal axis of the chart or graph is used to plot increments of time and the vertical axis pinpoints values of the variable that is being measured. When the values are connected in chronological order by a straight line that creates a series of peaks and valleys, a time series chart may also be referred to as a fever chart.

3.1.2 Working with Tableau – An Introduction

Tableau, a popular data visualization tool can be used to create various charts. The opening screen of the Tableau Desktop version is shown in fig 3.2.

File Data Server Help					
*					
Connect	Open				Discover
To a File				Open a Workbook	
Microsoft Excel					Training
Text file					Catting Started
JSON file					detting started
Microsoft Access					Connecting to Data
PDF file					Visual Analytics
Spatial file					Understanding Tableau
Statistical file					More training videos
More					-
To a Server					Viz of the Week
Tableau Server					Boeing Commercial Market Outlook
Microsoft SQL Server					2018-2037
MySQL					
Oracle					Les Resources
Amazon Redshift					Get Tableau Prep
More >	Sample Workbook	(S		More Samples	Blog - Practice your data viz skills and build your portfolio with these Tableau Commu
Saved Data Sources		C.	1.11		Register for Tableau Conference
Sample - EU Superstore	4120 4000 0000 0				Forums
Sample - Superstore	Superstore	Regional	World Indicators		

Fig 3.2 Snapshot of Tableau Opening Screen

The Tableau workspace is a collection of worksheets, menu bar, toolbar, marks card, shelves and a lot of other elements as shown in fig 3.3. Sheets can be worksheets, dashboards, or stories. The image below highlights the major components of the workspace. However, more familiarity will be achieved once when the users work with actual data.



Fig 3.3 Tableau Workspace

To begin working with Tableau, data source to be connected. Tableau is compatible with a lot of data sources. The data sources supported by Tableau appear on the left side of the opening screen. Some commonly used data sources are excel, text file, relational database or even on a server. One can also connect to a cloud database source such as Google Analytics, Amazon Redshift, etc.

One of the sample dataset is superstore data set that comes pre-loaded with Tableau. The data is that of a superstore. It contains information about products, sales, profits, etc. The aim of data analysts is to analyze the data and find critical areas of improvement within this fictitious company.

Tableau provides convenient options for building time series charts. The built-in date and time functions allow the user to use the drag-and-drop option to create and analyze time trends, drill down with a click, and easily perform trend analysis comparisons. Dimensions are qualitative data, such as a name or date. By default, Tableau automatically classifies data that contains qualitative or categorical information as a dimension, for example, any field with text or date values. These fields generally appear as column headers for rows of data, such as Customer Name or Order Date, and also define the level of granularity that shows in the view.

Measures are quantitative numerical data. By default, Tableau treats any field containing this kind of data as a measure, for example, sales transactions or profit. Data that is classified as a measure can be aggregated based on a given dimension, for example, total sales (Measure) by region (Dimension). Aggregation is the row-level data rolled up to a higher category, such as the sum of sales or total profit. Tableau automatically sorts the fields in Measures and Dimensions.

3.1.3 Procedure to Create Time Series Chart in Tableau

- Step 1: Start Tableau
- Step 2: Get into Tableau workspace
- Step 3: Connecting to a Data Source -Sample-Superstore data set
- **Step 4:** Go to the worksheet. Click on the tab sheet1 at the bottom left of the tableau workspace as shown in fig 3.4.



Fig 3.4 Worksheet in Tableau

- **Step 5:** In the worksheet, from Dimension under the Data pane, drag the Order Date to the Column shelf. (On dragging the Order Date to the columns shelf, a column for each year of Orders is created in the dataset. An 'Abc' indicator is visible under each column which implies that text or numerical or text data can be dragged here. On the other hand, if the Sales pulled here, a cross-tab would be created which would show the total Sales for each year.)
- **Step 6:** Similarly, from the Measures tab, drag the Sales field onto the Rows shelf. (Tableau populates a chart with sales aggregated as a sum. Total aggregated sales for each year by order date is displayed. Tableau always populates a line chart as shown in fig 3.5 for a view that includes time-field which in this example is Order Date.)

Pages		III Coxumns	(TAR Unter One i	
			≡ Roya	(SUM(Saire)
Filters			Sheet 1	Order Date
Maries			10008	1
in Aut	analle -	(3)		/
tt Color	D tite	(U) Libel	Hodik -	
 Detai	Colta	er' Pati	BDDK	/
			7995	
			0000	
			8 szek	
			4004	
			andic -	
			25dk)	
			100H)	
			0k. 200	8 2017 2018 2018

Fig 3.5 Line Chart for Aggregated Sales for Each Year

Step 7: In the chart above, the display is in years. To further drill down to quarter and month levels, click on the plus icon on the order date in the Columns shelf. This will generate the following output, which now displays the data broken down to the month and quarter level as shown in fig 3.6.



Fig 3.6 Line Chart for Aggregated Sales for Each Month and Quarter Level

The above chart is useful, but it is displayed in a discrete format. It will be more beneficial if the data is displayed in continuous form.

Step 8: To convert the chart into a continuous format time series chart, the first step is to roll up the YEAR (Order Date) back to year level, and then the second step is to right-click on it and select the Year and Continuous options as shown in fig 3.7.



Fig 3.7 Converting Discrete to Continuous Line Chart

Step 9: Drill down to quarter and Month level as in step 7 by changing the Columns shelf from YEAR (Order Date) to MONTH (Order Date). This will generate a monthly time series chart. From an analytics perspective, this chart shown in fig 3.8 is more insightful as it allows us to see the sales fluctuations across months and years. This is also useful for decomposing the seasonality and trend components of the time series data.



Fig 3.8 Time-Series Chart For Sales Analysis for Each Month

Step 10: Change the Path Property by going into the Marks shelf and clicking on the Path option. There are three options for the type of line graph for the view, and selecting the second option will produce the chart as shown in fig 3.9. The output is like the previous chart, but the trend shifts are more pronounced now.



Fig 3.9 Time-Series Chart in Different Path Property

Step 11: Adding Categories to Time Series

A time series chart with two variables, Sales and time can be further improved by adding more variables to a chart. For instance, it could be useful to visualize sales by segment across time. This can be done easily in two ways. First, simply drag the Segment field to the Color pane in the Marks shelf. The second method is to move the category to the Rows shelf to show it separately as shown in fig 3.10.



Fig 3.10 Time Series Chart With Sales with Different Category

3.2 SCORECARDS

3.2.1 Introduction

Scorecards offer organizations a snapshot of their current performance when compared to their goals. They are useful tools for organizations which need to manage performance and make strategic decisions better based on the distance between current performance and the goal. As such, scorecards present a more static view of an organization at a point in time rather than a dynamic hub to monitor success.

Scorecards are most commonly used to track KPIs, as they focus on both the current status of the metric being tracked and the target value. However, scorecards aren't live, so data is not updated in real-time. Instead, scorecards serve to monitor strategic goals relative to KPIs and to make decisions on a larger scale.

These decisions can include tracking the progress of a set strategy, measuring the efficiency of particular teams or departments towards meeting goals or even identifying problems and how they can be resolved. Scorecards are generally periodic measures, usually updated at set intervals such as weekly or monthly. For example, a scorecard can summarize total sales, average bounce rate, count of ad impressions, maximum hold time, minimum failure rate, etc. Scorecards in Google Data Studio appear as numbers, and, optionally, the name of the metric being summarized. The format of the displayed number depends on how the metric is configured in the data source. For example, the data source for a fictional pet store contains the following metrics. The *Qty Sold* metric is simply a number coming from the data set. The *Avg Qty Sold* metric is a duplicate of the *Qty Sold* field, with the Average aggregation type. *Total Items*, and *Unique Items* are calculated fields as given in table below as well as Scorecards for these metrics are shown in fig 3.11.

Name	Calculation	Aggregation Type	Qty Sold	Total Items
Qty Sold	none	Sum	02	19
Avg Qty Sold	none	Average		
Total Items	COUNT(Items)	Auto	Avg Qty Sold	Unique Items
Unique Items	COUNT_ DISTINCT(Items)	Auto	3.26	0

Fig 3.11 Sample scorecard

3.2.2 Scorecard Vs Dashboard

- Dashboards offer a broad way to track strategic goals and measure a company's overall efficiency. Scorecards, on the other hand, provide a quick and concise way to measure KPIs and give a clear indication of how well organizations are working to achieve their targets.
- Dashboards provide dynamic data i.e., Data is constantly updated, giving organizations an opportunity to track their operational performance in real time. but scorecard provides static data
- Dashboards are used daily in organizations as they offer a more operational view of success than scorecards' focus on strategic goals.
- Scorecards are ideal for a concise view of a specific area. It is used to determine how well marketing KPIs are being met, and illuminates how close or far they are from their goals. This can be useful to identify areas for improvement or ways to make specific tasks more efficient. Dashboards are advantageous to provide a bird's-eye view of the organization's operations.
- Scorecards and dashboards are not necessarily mutually exclusive. Choose carefully according to the need of the business.

Dashboards and scorecards don't have to be separate entities. Scorecards can
also be included in dashboards, offering an individual location to view multiple
KPIs and their accompanying progress.

3.3 BULLET CHART

3.3.1 Introduction

In data visualization, there are situations where a single value data to be compared to target value and also indicate if it is good, bad or excellent, and all this in a limited space. A bullet chart is very useful for such kind of situations. A *bullet graph* is a variation of a bar *graph* developed by Stephen Few. The *bullet chart* serves as a replacement for dashboard gauges and meters. A sample bullet chart in fig 3.12 representing the performance of sales representatives against their target as well as color coded to indicate their performance is ok or good or excellent.



Fig 3.12 Sample bullet chart

3.3.2 Components of Bullet Chart

The Bullet graph consists of 5 primary components as in figure 3.13.



Fig 3.13 Components of Bullet Chart

- Text label: X axis and Y axis unit of measurement.
- Quantitative Scale: Measures the value of metric (eg. sales value) on a linear axis.
- The Performance Measure: The bar that displays the primary performance measure (eg: Sales).
- Comparative Measure: Target value
- **Qualitative Scale:** The background color fill that encodes qualitative ranges like bad, satisfactory and good.

3.3.3 Types of Bullet chart

A bullet graph can be either horizontal or vertical depending on the alignment of the quantitative scale as shown in fig 3.14. The choice of vertical or horizontal alignment depends on the available space for the visualization.



Fig 3.14 Sample Horizontal and Vertical bullet chart

When a set of bullet graphs are displayed together and they include some measures that are considered good when they are high (eg: revenue) and others that are considered good when they are low (eg: expenses)

While plotting single measures, it is considered that target as the point which a measure should reach or exceed. While this works for measures like revenue and profit, for other measures like expenses, the situation is reversed. Expenses have to stay below target to be optimum.

In the chart above, the background fill uses the darkest color for poor performance and the lightest color for best performance. It works for revenue, new customers and avg. order size but not for expenses. Usually while color coding the qualitative scale, distinct intensities from dark to light of a single hue are used. Darker color intensities for the poor states and the lighter color intensities for the favorable states are used.

3.3.4 Advantages and Disadvantage of Bullet Chart

The advantages of using bullet graphs in place of gauges on the dashboard are:

- It can be oriented horizontally or vertically, depending upon the real estate available
- It can display multiple measures
- Information is presented in an easier to digest format

The disadvantages of bullet graphs are:

- Too much is going on in a single graph, and can quickly confuse a person new to data visualization
- It is difficult for the reader to understand how big the variation is
- The background, with many shades of one color, is distracting to the reader

3.3.5 Procedure to create bullet chart in Tableau

The procedure will enable the user to find the size of profits for the respective sales figures in each Sub-Category for the Sample - Superstore data source.

- **Step 1:** Open Tableau Desktop and connect to the Sample Superstore data source.
- Step 2: Navigate to a new worksheet.
- **Step 3:** From the Data pane, under Dimensions, drag and drop Sub-Category to the Columns shelf
- **Step 4:** Drag and drop the Measures Profit and Sales to the Rows shelf and observe the chart appears as in fig 3.15 which shows the two measures as two separate categories of bar charts, each representing the values for sub-categories.



Fig 3.15 Two categories of bar chart

Step 5: Drag the sales measure to the Marks card. Using Show Me, choose the bullet graph option. Observe the shown bullet graph as given in fig 3.16.



Fig 3.16 Bullet graph

3.4 AREA CHART

3.4.1 Introduction

An area chart or area graph is basically a line graph with the area below the lined filled with colors or textures. Like line graphs, area charts are used to represent the development of quantitative values over a time period. Area charts often used to show overall trends over time rather than specific values. For example, for the quarterly sales data for five years, to do comparison of sales column or bar chart is used. But to show the trend of how the sales values have changed over the years, time series chart or an area chart is useful.

Area charts are commonly used to showcase data that depicts a time-series relationship. A sample area chart is shown in fig 3.17. The area chart can be visualized in two ways:

One with data plots overlapping each other



Another with data plots stacked on top of each other

Fig 3.17 Sample area chart

3.4.2 Line chart Vs Area Chart

Both line chart and area chart can be used interchangeably. A sample line chart and area chart for the same data is shown in fig 3.18. But there are few differences, which are discussed below.



Fig 3.18 Line chart Vs Area chart

- A line chart would be good for showing net change in population over time, while an area chart would be good for showing the total population over time. The filled area below the line can help to indicate that it is a physically countable amount.
- Line charts and area charts are very closely related. They are both good for time series data. They both show continuity across a dataset. They are both good for seeing trends rather than individual values.
- Line charts are good at showing multiple different series and comparing them against each other. They can support up to about seven lines in a static version; interactive versions can go higher. Area charts are not as good at comparing between that many areas because they have problems with occlusion.
- Use line charts to compare several data series, or for individual intangible values like rates. Use area charts for multiple data series with part to whole relationships, or for individual series representing a physically countable set, or cumulative series of values.

3.4.3 Procedure to create an area chart in Tableau

- **Step 1:** Open Tableau Desktop and connect to the Sample Superstore data source.
- Step 2: Navigate to a new worksheet.
- Step 3: From the Data pane, under Dimensions, drag Order Date to the Columns shelf
- **Step 4:** On the Columns shelf, right-click YEAR(Order Date) and select Month as shown in fig 3.19.

🖽 MONTH(Order 👻	
Filter	
Show Filter	
Show Highlighter	
Sort	
Format	
✓ Show Header	
✓ Include in Tooltip	
Show Missing Value	15
Year	2015
Quarter	Q2
✓ Month	May
Day	8
More	•
Year	2015

Fig 3.19 Drill down to Month Sales

- Step 5: From the Data pane, under Measures, drag Quantity to the Rows shelf.
- **Step 6:** From the Date pane, under Dimensions, drag Ship Mode to Color on the Marks card.
- **Step 7:** On the Marks card, click the Mark Type drop-down and select Area as given in fig 3.20 and observe the chart visualization changes to display ship mode details for monthly order data.



Fig 3.20 Selecting Area Chart in Mark Type and visualization

- **Step 8:** Add formatting to an area chart (if required), using Format menu. Choose the part of the view that is to be formatted, such as Font, Borders, or Filters.
- **Step 9:** Add a highlight action using the highlight button in the toolbar if required.

REVIEW QUESTIONS

- 1. Create the Procedure to Create Time Series Chart in Tableau
- 2. Differentiate between Scorecard Vs Dashboard.
- 3. What are the Components of Bullet Chart?
- 4. Explain the advantages and disadvantages of Bullet Chart.
- 5. Discuss the uses of area chart.
- 6. Demonstrate the procedure to create bullet chart in Tableau
- 7. Compare and contrast Line chart Vs Area Chart.
- 8. Write the procedure to create an area chart in Tableau
- 9. Explain the uses of times series chart with suitable example.
- 10. How to Connect a Data Source in tableau? Explain.
- 11. Create a table with your own data set and draw the time-series chart for sales analysis for each Month.

4.1 HEAT MAP

4.1.1 Introduction

Heat maps originated in 2D displays of the values in a data matrix. Larger values were represented by small dark gray or black squares (pixels) and smaller values by lighter squares. Software designer Cormac Kinney trademarked the term heat map to describe a 2D display depicting financial market information.

Heat maps visualize data through variations in colouring. When applied to a tabular format, Heat maps are useful for cross-examining multivariate data, through placing variables in the rows and columns and colouring the cells within the table. Heat maps are good for showing variance across multiple variables, revealing any patterns, displaying whether any variables are similar to each other, and for detecting if any correlations exist in-between them.

The cells in the data matrix, either contain categorical data (eg., Male or Female) or numerical data (eg., 10, 50). Categorical data is colour-coded, while numerical data requires a colour scale that blends from one colour to another, in order to represent the difference in high and low values. A selection of solid colours can be used to represent multiple value ranges (0-10, 11-20, 21-30, etc) or a gradient scale for a single range (for example 0 - 100) by blending two or more colours together. A sample heat map is shown in fig 4.1.



Fig. 4.1 Sample Heat Map Chart

A density heat map is used to analyze the areas in a plot where data points are dense or scattered. Heat maps are specifically used where there is a huge data set with overlapping data values. This helps the analyst to see the areas with greater density and discover data trends. A sample heat map shown in fig 4.2 visualizes the visitor's behaviour on a web page and provides clue on where the most important content to be placed.



Fig 4.2 Heat map showing visitors behaviour on a web page

4.1.2 Uses of Heat Map

- Heat maps are good for showing variance across multiple variables, revealing any patterns, displaying whether any variables are similar to each other, and for detecting if any correlations exist in-between them.
- Heat maps can also be used to show the changes in data over time
- Heat maps are used to indicate the weight of each point in the geographical range. It is usually displayed in a special highlight.

4.1.3 Procedure to Create Heat Map in Tableau

The procedure provides steps to create a density heat map using sample dataset pertaining to sales in an electronics store. The following requirements are to be considered before creating heat map, which are,

- Rows: At least one measure or dimension
- Columns: At least one continuous measure

- Mark type: Density
- Marks card: At least one dimension
- Step 1: Start Tableau.
- Step 2: Connect to Data Source Electronic Store Sales.
- **Step 3:** Add Measure Profit to the Columns section as shown in fig 4.3.



Fig 4.3 Add to the columns section

Step 4: Select the aggregation type as AVG as shown in fig 4.4, that is, an average of the field values. Also, make sure that the measure is continuous type.



Fig 4.4 Select average measure

Step 5: Next Add Measure Sales to Rows Section and again select **Average** of the field values. Observe the chart as shown in fig 4.5, an empty plot with two axes appears on the canvas.



Fig 4.5 An empty plot of Chart

Step 6: Add Dimension Field State into Detail card present in Marks section. This will add a group circle representing different states on the plot showing average sales and average profit for each state as shown in fig 4.6.



Fig 4.6 Adding a Dimension

Step 7: To convert this plot into a density heat map, select the Shape as Density. This will change the shape of data points from circles to density spots. That is, the color scheme of data points will follow a density gradient as shown in fig 4.7. The regions with most data points or dense regions will be in red/orange whereas, the areas with lesser or scattered data points appear in greenishblue shades. The color schemes can be modified for heat maps.

Step 8: Right-click on Color card and set the intensity, opacity and other border effects for the heat map as given in fig 4.7. Select the color scheme from a long list of available options.



Fig 4.7 Color scheme for heat map

Step 9: Using size in Mark type, increasing or decreasing the size of density spots in heat map is shown in fig 4.8.



Fig 4.8 Increase or Decrease the density spots

4.2 GEO MAP

4.2.1 Introduction -Need of map visualization

There are many reasons to put the data on a map. The user can choose map, if the data is location data or looking for a report relating to a location. Map visualization can be used when there is a need to find answers to spatial questions such as,

- Which state has the most farmers markets?
- Where are the regions in the India with the high obesity rates?
- Which metro station is the busiest for each metro line in Chennai?

Geo Maps represent a family of geospatial visualizations that change according to data type.

- If variables are coordinates (longitude, latitude) or city-specific, the user can get a dot map or bubble map.
- If variables are shapes or areas (e.g. states, countries, continents), the user can get a Choropleth map, Choropleth Maps are geospatial visualizations that use shape files, or polygons of geographic areas as given in fig 4.9. Choropleth maps allow each area to represent quantitative data using fill saturation. In other words, typically, the *higher* the quantity, the *darker* the area.



Fig 4.9 Choropleth map (Polygon Shape)

4.2.2 Uses of Geo map

Combining geospatial information with data over time creates a greater scope of understanding. Some benefits of using maps in data visualization include:

- A greater ability to more easily understand the distribution of the organization's presence across the city, state, or country
- The ability to compare the activity across several locations at a glance
- More intuitive decision making for company leaders
- Contextualizing the data in the real world

4.2.3 Types of maps

Many types of map visualization can be made using different visualization tools, some of them are,

- Proportional symbol maps
- Choropleth maps (filled maps)
- Point distribution maps
- Heat maps (density maps)
- Flow maps (path maps)
- Spider maps (origin-destination maps)

Proportional symbol maps are great for showing quantitative data for individual locations. For example, earthquakes around the world and size them by magnitude using symbol map. Choropleth maps known as filled maps, different from heap maps. Choropleth maps are great for showing ratio data. Geo maps are used to show the measure details of different regions using color code. More value shown by darker color and less value are shown by light color.

4.2.4 Procedure to create Geo map / Symbol map in Tableau

Creating map involves many tasks such as

- Connecting to data source
- Joining geographic data
- Formatting geographic data
- Creating location hierarchies
- Building and presenting a basic map view
- Applying key mapping features
- Step 1: Start Tableau.
- Step 2: Connect to data source Sample Superstore Data.
- Step 3: (Joining) On the left side of the Data Source page, under Sheets, doubleclick Orders. Next Under Sheets, double-click People. Tableau creates an inner-join between the two spreadsheets, using the Region column from both spreadsheets as the joining field. To edit this join, click the join icon (the two circles), if required edit the join in the Join dialog box that opens as shown in fig 4.10.



Fig 4.10 Join Dialog Box

Step 4: (Formatting) Depending on the type of map to create, assign certain data types, data roles, and geographic roles to the fields (or columns). On the Data Source page, click the data type icon (the globe) for Postal Code and select String as shown in fig 4.11.

Orders State	 Orders Postal Code 	Abc Orders Region	Abc Orders Product ID
Kentucky	42420	South	FUR-B00001798
Kentucky	42420	South	FUR-CH-10000454
California	90036	West	OFF-LA-10000240
Florida	33311	South	FUR-TA-10000577
Florida	33311	South	OFF-ST-10000760
California	90032	West	FUR-FU-10001487
California	90032	West	OFF-AR-10002833
California	90032	West	TEC-PH-10002275
California	90032	West	OFF-BI-10003910
California	90032	West	OFF-AP-10002892

Fig 4.11 Formatting data type

Step 5: On the Data Source page, click Sheet 1. Observe the worksheet as in fig 4.12 with generated Latitude and Longitude fields in Measures. It indicates that the data is ready with geographic roles to prepare map.

Analytics 4	Esam	AR, Coturne		
B Ddes+(Tarol+ Son		I Rows		
Category Category City City City	Plans,	Sheet 1	Quan Para Mare	
Clusterner (D)	Marks			
Customer Rame	Container an			
Order ID Postal Code				
Another Service Another Service	A Q Boot Santa			
		trou Seut Mer	September	

Fig 4.12 Preparing Geographic roles

Step 7: In the Data pane, under Dimensions, select a field, such as Row ID, and drag it down to the Measures section. The field is added to the Measures

section and changes from blue to green. A Dimension field is converted to a Measure.

- **Step 8:** Create a geographic hierarchy. Creating a geographic hierarchy allows the user to quickly drill into the levels of geographic detail the data contains, in the specified order.
 - **Step 8.1:** In the Data pane, right-click the geographic field, Country, and then select Hierarchy > Create Hierarchy.
 - **Step 8.2:** In the Create Hierarchy dialog box that opens, give the hierarchy a name, such as Mapping Items, and then click OK. At the bottom of the Dimensions section, the Mapping Items hierarchy is created with the Country field.
 - **Step 8.3:** In the Data pane, drag the State field to the hierarchy and place it below the Country field.
 - **Step 8.4:** Repeat step 8.3 for the City and Postal Code fields.
- **Step 9:** Build a basic map. In the Data pane, double-click Country. The Country field is added to Detail on the Marks card, and Latitude (generated) and Longitude (generated) are added to the Columns and Rows shelves. A map view with one data point is created. Since a geographic role is assigned to Country, Tableau creates a map view. Double-click any other field, such as a dimension or measure, Tableau adds that field to the Rows or Columns shelf, or the Marks card, depending on what the user already have in the view. Geographic fields are always placed on Detail on the Marks card, however.



Step 10: On the Marks card, click the + icon on the Country field. The State field is added to Detail on the Marks card and the map updates to include a data point for every state in the data source as in fig 4.13.



Fig 4.13 Add Details on the Marks Card

Step 11: From Measures, drag Sales to Color on the Marks card. Each state is colored by sum of sales. Since Sales is a measure, a qualitative color palette is used. Place a dimension on color, then a categorical color palette is used as in fig 4.14.



Fig 4.14 Categorical color palette for Geographical map

Step 12: From Measures, drag Sales to Label on the Marks card. Each state is labeled with sum of sales. The numbers need a little bit of formatting, In the Data pane, right-click Sales and select Default Properties > Number Format. In the Default Number Format dialog box that opens, select Number (Custom), and then do the following as in fig 4.15,

Sheet 1	e Default	Number Format [Sales]
1294 144 15 15 16 17 16 17 16 17 18 18 18 18 18 18 18 18 18 18	Automatic Number (Standard) Number (Custom) Currency (Standard) Currency (Custom) Scientific Percentage Custom	Number (Custom) Decimal places. 0 0 Negative values: -1234 Units: Thousands (K) Prefix / Suffix:
	Clear	Cancel

Fig 4.15 Numbers are mapped in Geographical areas

Step 12: From the Show Me Pane, select Geo map / Symbol map for different visualization. Customize the background map using map menu item if required.

4.3 SYMBOL MAP

4.3.1 Introduction

Symbol Maps are simply maps that have a mark displayed over a given Longitude and Latitude. Using the "Marks" card in Tableau the user can quickly build up a powerful visual that informs users about their data in relation to its location. These types of maps are called proportional symbol maps. Proportional symbol maps are great for showing quantitative values for individual locations. They can show one or two quantitative values per location (one value encoded with size, and, if necessary, another encoded with color). For example, the user can plot earthquakes recorded from 1981 to 2014 around the world, and size them by magnitude as shown in fig 4.16. The user can also color the data points by magnitude for additional visual detail.



Fig 4.16 Visualization of Earthquake using Symbol Map

4.3.2 Procedure to create Symbol Map

Step 1: Start Tableau

Step 2: Connect to Earthquake data source as given in table below.

Earth- quake Date Time	ID	Magnitude	Magnitude^10	Latitude	Longitude
1/1/73	centennial19730101114235	6.00000	17,488,747.04	-35.570	-15.427
1/2/73	pde19730102005320300_66	5.50000	25,329,516.21	-9.854	117.427
1/3/73	pde19730103022942800_33	4.80000	6,492,506.21	1.548	126.305
1/4/73	pde19730104003142000_33	4.50000	3,405,062.89	41.305	-29.272
1/5/73	pde19730105003948200_36	4.70000	5,259,913.22	0.683	-80.018
1/6/73	pde19730106061852300_83	4.90000	7,979,226.63	-22.354	-69.310

Basic map building blocks:

Columns shelf:	Longitude (continuous measure, longitude geographic role assigned)
Rows shelf:	Latitude (continuous measure, latitude geographic role assigned)
Detail:	One or more dimensions
Size:	Measure (aggregated)
Mark type:	Automatic

- **Step 3:** Open a new worksheet. In the Data pane, under Measures, double-click Latitude and Longitude. Latitude is added to the Rows shelf, and Longitude is added to the Columns shelf. A map view with one data point is created.
- **Step 4:** From Dimensions, drag ID to Detail on the Marks card. If a warning dialog appears, click Add all members. A lower level of detail is added to the view as shown in Fig 4.17.



Fig 4.17 Add members in a map

Step 5: From Measures, drag Magnitude^10 to Size on the Marks card. The Magnitude^10 field is used to encode size, instead of the Magnitude field. This is because Magnitude^10 contains a wider range of values, so the differences between values can be seen visually. A proportional symbol map appears as in fig 4.18. The larger data points represent earthquakes with larger magnitudes, and the smaller data points represent earthquakes with smaller magnitudes.



Fig 4.18 Symbol map appears as small and larger data points as per the proportion of the Earthquake Magnitude

- Step 6: From Measures, drag Magnitude to Color on the Marks card.
- **Step 7:** On the Marks card, click Color > Edit Colors.
- Step 8: In the Edit Colors dialog box, do the following:
 - Click the color drop-down and select the Orange-Blue Diverging palette from the list.
 - Select Stepped Color, and then enter 8. This creates eight colors: four shades of orange, and four shades of blue.
 - Select Reversed. This reverses the palette so that orange represents a higher magnitude than blue.
 - Click Advanced, select Center, and then enter 7. This shifts the color palette and ensures that any earthquake over 7.0 magnitudes will appear orange in color, and any earthquake under 7.0 magnitudes will appear blue in color.
 - Click OK.
- **Step 9:** On the Marks card, click Color again, and then do the following: For Opacity, enter 70%. Under Effects, click the Border drop-down menu and select a dark blue border color. The map view updates with new colors. The dark orange

data points represent earthquakes with higher magnitudes, while the dark blue data points represent earthquakes with lower magnitudes. The opacity of the marks is at 70% allows to see where the data points overlap as in fig 4.19.



Fig 4.19 Enhance the visualization by using boundary colors and opacity

Step 10: On the Marks card, right-click the ID field and select Sort.

Step 11: In the Sort dialog box, do the following: For Sort Order, select Descending. For Sort By, select Field, and then click the drop-down and select Magnitude. Click OK. This sorts the data points in the view so that the larger magnitudes appear on top. Observe the completed proportional symbol map as in fig 4.20.



Fig 4.20 Proportional Symbol Map

4.4 FILLED MAP

Filled maps in Tableau are similar to symbol maps, but they include many more data points. While a symbol map draws a symbol at the intersection of each latitude and longitude pair, filled maps draw a polygon around the entire border. A filled map colored by region in Tableau is shown in fig 4.21.



Fig 4.21 Sample Filled Map

Filled maps are one of the easier chart types to create in Tableau using Show Me. To create a filled map in Tableau, simply click a geographic dimension (identified by a globe icon) from the Dimensions Shelf and choose 'filled maps' under Show Me. To create a filled map manually, double-click on the geographic dimension for State in the Sales – Superstore data set that comes with Tableau.



Fig 4.22 Default symbol map for each state

By default, Tableau generates a symbol map, placing a circle at the intersection of Longitude and Latitude for each state as given in fig 4.22. First, Longitude is on the Columns Shelf, which can also be thought of as the X-axis. Conversely, Latitude is on the Rows Shelf, or the Y-axis. On the Marks Shelf, State is the most granular level of detail in the view.

In order to change this from a symbol map to a filled map, change the mark type from 'Automatic' to 'Filled Map'. By selecting this special mark type in Tableau, the single circles on each state have been converted to nice, smooth polygons that trace the entire border of each state. Then encode the filled map by color by placing a field on the Color Marks Card. To color the territories by a measure such as Sales or Profit, color the marks by a dimension. The filled map view will appear as in fig 4.23.



Fig 4.23 Sales or Profit based colored Filled map

4.5 EDITING LOCATION IN MAP

In Geo map or symbol map, sometimes Tableau will not recognize one or more of the location names in the geological data. When this happens those values are marked as **unknown** in the lower right corner of the map view. This may happen if there is a location or abbreviation that Tableau does not recognize or if the location is ambiguous and could exist in multiple places. When this happens, additional information can be added to the view to define locations, or unknown location names can be edited to map to known locations.

4.5.1 Add more fields to view

If the data set includes ambiguous locations (for example, "Aberdeen," which could appear in multiple states or countries), adding another geographic field, like State
or Country, to the view defines the correct location for that data. If there is a hierarchy in the data pane, Tableau will automatically use the appropriate levels of the hierarchy to solve location ambiguities.

4.5.2 Edit locations in the Special Values menu

Sometimes, unknown locations can be prompted because there's a misspelling or another issue. When this happens, we can correct the unknown or ambiguous locations in the data using the **Special Values** menu ad given below.

- In the bottom right corner of the view, click the special values indicator that lists the number of unknown locations in your map. This opens the Special Values menu.
- In the Special Values dialog box, select **Edit Locations**. If the special values indicator is not visible, select **Map** > **Edit Locations** as shown in fig 4.24.



Fig 4.24 Edit Locations Menu Item

4.5.3 Edit ambiguous locationsv

Both unrecognized and ambiguous locations can be listed in the Edit Locations dialog box. Some ambiguous locations can be fixed by mentioning the **Country/Region** and/or **State/Province** for each city from a specific field in the data, either by hard-coding a value if the data only spans one state or country, or by telling Tableau which fields to refer to for that information. This option is in the Geographic Roles section of the Edit Locations dialog box. For example, if we have several cities that are ambiguous, we can specify a State/Province to fix them as shown in fig 4.25.

eographic roles							
Country/Region:	United States						
State/Province:	None None G Fixed: Alabama			•			
City:					A 242 issues		
latch values to loca				*			
📥 City	From field: State or Province			O None		2	
	Your Data		Matchi	Fixed:	Alabama		
Aberdeen		Ambiguous	From field:	State or Province			
Albany			Ambiguous				
Alexandria			Ambiguous				

Fig 4.25 Ambiguous regions solved by selecting state or province

If there are ambiguous or unknown locations in the Country/Region or State/Province tabs in the Edit Locations menu, fixing those first may resolve some ambiguous locations in smaller geographic roles, such as cities.

4.5.4 Edit unknown locations

In the Edit Locations dialog box, click on one of the **Unrecognized** cells to match a known location to the unknown data. When we click on an unrecognized cell, a search box appears. As we begin typing in the search box, Tableau generates a list of possible locations. Select a location from the list as shown in fig 4.26.

Country/Region: United States State/Province: State or Province City: City		•								
		•								
		A 3/15045								
ch values to i	ocations									
1 City	State/Province									
Your Data		Matching Location								
L.A. NYC SEA		Los Enter a Latitude and Longitude Exist Los Angeles								
								Aberdeen		Los Antos
								Abilene Abington Acton Addison		Los Banos
Los Gatos										
Palos Hills Paweba Palos Verder										
San Carlos Park										
Adrian		(Adrian)								
Telephone and the	and the second	a lat								

Fig 4.26 Solving unknown locations

Alternatively, we can enter latitude and longitude to manually map a value to a point location on the map. To do this, begin typing into the unrecognized cell and select **Enter Latitude and Longitude** from the drop-down menu as given in fig 4.27. When we type a latitude and longitude, enter the values in decimal format. For example, Addis Ababa is Latitude: 9.033140, Longitude: 38.750080.

Bamako Baraki Baraki Barnsley Beni Mellal Berrechid Dinzhou Cainta Cainta Calamba Calamba		INVOCUTING	Location
Baraki Enter a Latitude and Longitude. Barnsley Unrecognized Unrecogni	mako	1	
Barnsley Unrecognized Beni Melial Unrecognized Berrechid Unrecognized Binzhou Unrecognized Cainta Unrecognized Calamba Unrecognized	raki	Enter a Latitude and L	ongitude.
Beni Mellal Unrecognized Latitude and Longitude Berrechid Unrecognized Latitude: Binzhou Unrecognized Longitude: Cainta Unrecognized Longitude: Calamba Unrecognized Cancel	imsley	Unrecognized	Entert still, do and I are it do
Berrechid Unrecognized Latitude: 190.0 to 90.0 Binzhou Unrecognized Lonptude: -180.0 to 190.0 Cainta Unrecognized Calcer	rri Mellal	Unrecognized	Enter Latitude and Longitude
Binzhou Unrecognized Longtude: -180.0 to 190.0 Cainta Unrecognized Calcel Unrecognized	rrechid	Unrecognized	Latitude: 90.0 to 90.0
Cainta Unrecognized Carcel	nzhou	Unrecognized	Longitude: -180.0 to 190.0
Calamba Unrecognized	linta	Unrecognized	1000 Million 1 2000
Calescent Unrecondend	lamba	Unrecognized	Cancel
		Harmonialand	9(I)

Fig 4.27 Latitude and Longitude of a location

Summary

In Data Visualization Tools, map charts are very effective in visualizing the geographical data for providing real world scenario to the business people.

REVIEW QUESTIONS

- 1. Discuss about the Uses of heat map.
- 2. How to create tableau heat map?
- 3. Why we need map visualization?
- 4. List out the uses of Geo map.
- 5. When should you use a map to represent your data?
- 6. What is latitude and longitude on a map?
- 7. Tableau is revolutionizing data analysis and has truly made geographic analysis accessible to everyone. Comment your views.
- 8. Elucidate the salient features of categorical color palette for geographical map.
- 9. Discuss about the different types of maps in geo map.

- 10. Explain about join dialog box in geo map.
- 11. Explain the tasks involved in Geo map /symbol map in tableau
- 12. Elaborate the procedure to create a symbol Map in tableau.
- 13. Create a density heat map using your sample dataset pertaining to sales in an electronics store.
- 14. Compare and contrast heap map and symbol map in tableau.
- 15. Create a table for Earthquake data source and visualize using the symbol map.

CHAPTER V

5.1 DASHBOARD

5.1.1 Introduction

Today, the use of dashboards forms an important part of decision making. In information technology, a dashboard is an easy to read, often single page, real-time user interface, showing a graphical presentation of the current status (snapshot) and historical trends of an organization's or department's key performance indicators to enable instantaneous and informed decisions to be made at a glance.

5.1.2 Dashboard – Definition

Stephen Few has defined a dashboard as "A visual display of the most important information needed to achieve one or more objectives which fits entirely on a single computer screen so it can be monitored at a glance".

In the present terms, a dashboard can be defined as a data visualization tool that displays the current status of metrics and key performance indicators (KPIs) simplifying complex data sets to provide users with at a glance awareness of current performance. Dashboards consolidate and arrange numbers and metrics on a single screen. They can be tailored for a specific role and display metrics of a department or an organization on the whole. Dashboards can be static for a one-time view, or dynamic showing the consolidated results of the data changes behind the screen. They can also be made interactive to display the various segments of large data on a single screen. A sample dashboard is shown in fig 5.1.



Fig 5.1 Sample Dashboard

5.1.3 Key Metrics for Dashboard

The core of the dashboard lies in the key metrics required for monitoring. Thus, based on whether the dashboard is for an organization on the whole or for a department such as sales, finance, human resources, production, etc. the key metrics that are required for display vary.

Further, the key metrics for a dashboard also depend on the role of the recipients (audience), for example, Executive (CEO, CIO, etc.), Operations Manager, Sales Head, Sales Manager, etc. This is due to the fact that the primary goal of a dashboard is to enable data visualization for decision making.

The success of a dashboard often depends on the metrics that were chosen for monitoring. For example, Key Performance Indicators, Balanced Scorecards and Sales Performance Figures could be the content appropriate in business dashboards.

5.1.4 Benefits of Dashboard

Dashboards allow managers to monitor the contribution of the various departments in the organization. To monitor the organization's overall performance, dashboards allow you to capture and report specific data points from each of the departments in the organization, providing a snapshot of current performance and a comparison with earlier performance.

Benefits of dashboards include the following,

- Customizable, Interactive, Predictive and great story teller of data
- Visual presentation of performance measures.
- Ability to identify and correct negative trends.
- Measurement of efficiencies /inefficiencies.
- Ability to generate detailed reports showing new trends.
- Ability to make more informed decisions based on collected data.
- Alignment of strategies and organizational goals.
- Instant visibility of all systems in total.
- Quick identification of data outliers and correlations.
- Time saving with the comprehensive data visualization than running multiple reports.

5.1.5 Types of Dashboards

Dashboards can be categorized based on their utility as follows,

- Strategic Dashboards
- Analytical Dashboards

- Operational Dashboards
- Informational Dashboards

Strategic Dashboards

Strategic dashboards support managers at any level in an organization for decision making. They provide the snapshot of data, displaying the health and opportunities of the business, focusing on the high level measures of performance and forecasts.

- Strategic dashboards require to have periodic and static snapshots of data (e.g. daily, weekly, monthly, quarterly and annually). They need not be constantly changing from one moment to the next and require an update at the specified intervals of time.
- They portray only the high level data not necessarily giving the details.
- They can be interactive to facilitate comparisons and different views in case of large data sets at the click of a button. But, it is not necessary to provide more interactive features in these dashboards.

A screenshot given in fig 5.2 shows an example of an executive dashboard which displaying goals and progress.



Fig 5.2 Executive Dashboard

Analytical Dashboards

Analytical dashboards include more context, comparisons and history. They focus on the various facets of data required for analysis. Analytical dashboards typically support interactions with the data, such as drilling down into the underlying details and hence should be interactive. Examples of analytical dashboards include Finance Management dashboard and Sales Management dashboard as shown in fig 5.3.

-www	~	Arg. Visit Duration • Arg. Visit Duration sector accord	1-1		• Guar Carrier	ion Rate ion Rate	M
May 24 Art 3 Art 17 Visits by Traffic Type	e	Visits and Pageverse	try Mable	54 U 0	Goal Convers	an 7 a	a la
		Mobile	Vinim	Pageviews	16.24%		No. 1
		film .	567	1,720	100 may 10.245		
		Yes	:57	1427	Goal Complet	sona	4
		Vote and Avg. Visit D	utation by Cour	try/Tentury 0	102	5,000	Amon
56.69% organic 214 Vels		Country/Territory	Visits	Avg. Visit Duration	Goal Value E182.00		1.5
23.88% referral		United Kingdom	298	00.02.52	internet address	N (0.000 (0.0	an a
= 19.44% Other		United States	110	00.00 59	Gital Complet	torns and Goat I	Comienta 14
122 Visite		bridia	43	00.00.39	Contration Sectore in the		122/221
		Canada	21	00:01:04	Source	Completions	Conversion
Pagennews by Traffic Type	fraffic Type @	Fignoe	17	00-00-01	through a	ET.	16.62%
					(Concert)	18	16 (17%)
		Visits and Pages / Vis	s by Country/Te	rationy @	1.00	10	28.47%
		Country/Territory	Vision	Pages / Visit	fact and the second	ĩ	*****
		United Wingdom	.2116	3.48	Access to an		10112

Fig 5.3 Analytical Dashboard for Finance and Sales Management

Operational Dashboards

Operational dashboards are for constant monitoring of operations. They are often designed differently from strategic or analytical dashboards and focus on monitoring of activities and events that are constantly changing and might require attention and response at a moment's notice. Thus, operational dashboards require live and up to date data available at all times and hence should be dynamic. An example of an operation dashboard could be a support-system dashboard, displaying live data on service tickets that require an immediate action from the supervisor on high-priority tickets as given in fig 5.4.



Fig 5.4 Sample Operational Dashboard

Informational Dashboards

Informational dashboards are just for displaying figures, facts and/or statistics. They can be either static or dynamic with live data but not interactive. For example, flights arrival/departure information dashboard in an airport as in fig 5.5.

Arrivals - London Heathrow Airport			LHR 5:32 PM GMT+1		
тионт		ORIGIN	GATE	ADRIVES 🕈	STATUS
BRITISHIRIR	1335	NEWGRISTLE		5:17 PH	ARRIVED
BRITISHAIR	861	PRAGUE		617 PH	RREDED
BRITISHAIR	323	PERLS		6:18 PM	ABRINED
JET HERMANS	118	mineral		6:21 PH	BRRIVED
RER L'INGLIS	176	DUBLIN		6.22 PH	RREIVED
BRITISH RIR.	485	AMSTERDAM		8231911	ARRIVED
BRITISHAIR	10	BANGKOK		8:25 PH	BRRIVED
ATSCHOOL	0.00	DOT THE		0.00 00	2000111000

Fig 5.5 Informational Dashboard for Airport Flight Arrival/Departure

5.1.6 Customized Dashboard

Depends on the data domain the dashboards are customized and categorized as,

- Business Dashboard
- Executive Dashboard
- KPI Dashboard
- Project Dashboard

Business Dashboard

In a way, virtually any dashboard used by a business falls into this category. The term "business dashboard" specifically refers to reporting tools that fulfill these purposes:

- Tracking important business metrics
- Monitoring business intelligence initiatives
- Reporting data to stakeholders

An effective business dashboard should focus on top-level data related to the overall success of the business. In most cases, every metric in the dashboard should

support the business' most important metric: the bottom line. The goal of a business dashboard is to not only communicate data about the business' success, though; it facilitates understanding ad alignment between departments, holds each team accountable for their goals and progress, and helps users identify areas that need immediate action.

Executive Dashboard

An executive dashboard gathers and holds information that top-level stakeholders need to run a company, business, or organization. Executive dashboards function much like business dashboards, except the information in them should cater specifically to the needs and expectations of executives. Executives only have so much bandwidth to gather and understand information, which means they need access to the information they need, when they need it. Some key benefits include:

- Performance Management: An overview of how departments are meeting their goals
- Scorecards: Insight regarding specific employee performance and goals
- Visibility: Access to high-level goals and metrics related to the overall success of the organization
- **Time Management:** Cohesive reports and drilldowns in one place, accessible on any device

KPI Dashboard

KPIs (Key Performance Indicators) are the heart and soul of the organization's performance. They are the stepping stones that will guide your business to long-term success, so tracking and comparing them in one place is vital. KPIs should be measurable, tangible metrics that let each employee, team, and department understand how their performance influences the success of the organization – and the KPI dashboard is where these metrics are stored. A successful KPI dashboard should:

- Set tangible goals and targets for each department
- Facilitate accountability within each department
- Provide real-time updates on goals and progress

Project Dashboard

Much like KPI dashboards, project dashboards track tangible goals; however, the goal of a project dashboard is not about hitting a sales quota or increasing marketing revenue by a certain margin. Instead, project dashboards track specific metrics related to the progress and complete of a project. This means, project dashboards involve more scheduling metrics as,

- When does the project need to be completed?
- Does each team member have the bandwidth to complete his or her portion of the project?
- What is the project budget? Is the project on pace to accommodate it?

These are the questions most – if not all – project managers ask themselves on a daily basis. By having these metrics in one place, project managers can avoid unnecessary time logging into multiple data sources and comparing information to get a simple progress report. With the help of a project management dashboard, they can simply open the dashboard and see exactly where the project stands, make accommodations or changes as necessary, and provide an accurate assessment of when the project will be complete.

5.2 CREATING A DASHBOARD

5.2.1 Creating a Dashboard using Google Data Studio

The steps involved in creating the dashboard are the following:

- Step 1: Login & Create a Blank Report
- Step 2: Add the First Data Source
- Step 3: Add the Second Data Source
- Step 4: Adding Elements
 - Line Graphs / Time Series
 - Pie Charts
 - Scorecards
 - Tables
- Step 5: Adding Date Range Filter

5.2.2 Creating a dashboard in Tableau

- Step 1: Start Tableau.
- **Step 2:** Create a dashboard by clicking click the New Dashboard icon at the bottom of the workbook as shown in Fig 5.6.



Fig 5.6 Dashboard Icon

Step 3 : From the **Sheets** list which were created already are shown at left, drag views to the dashboard at right as given in fig 5.7.



Fig 5.7 Sheets list

- **Step 4:** To replace a sheet, select it in the dashboard at right. In the Sheets list at left, hover over the replacement sheet, and click the Swap Sheets button.
- **Step 5:** Adding objects and set their options: The user can add dashboard objects that add visual appeal and interactivity. The various objects are,
 - Horizontal and Vertical objects provide layout containers that allows to group related objects together and fine-tunes when the dashboard resizes during user interaction.
 - Web Page objects display target pages in the context of the dashboard.
 - Blank object helps to adjust spacing between dashboard items.
 - Text objects serve can provide headers, explanations, and other information.
 - Button objects helps to navigate from one dashboard to another, or to other sheets or stories. For the button style, choose an image or text to indicate the button's destination to the users.
 - Image objects add to the visual flavor of a dashboard, and can be linked to specific target URLs.
 - Extension objects helps to add unique features to dashboards or integrate them with applications outside Tableau.

To add an object in dashboard, Select an item under Objects on the left, and drag it to the dashboard sheet on the right as given in fig 5.8.



Fig 5.8 Add an object in Dashboard

To set options for objects, Click the object container to select it. Then click the arrow in the upper corner to open the shortcut menu. (The menu options vary depending on the object.) Detailed options for navigation buttons: Button objects have several unique options that helps to visually indicate the navigation destination as given in fig 5.9.



Fig 5.9 Navigation Button

A navigation button using text for the button style: In the upper corner of a Button object, click the object menu, and choose **Edit Button** as in fig 5.10.



Fig 5.10 Edit button

- From the Navigate to menu, choose a sheet outside the current dashboard.
- Choose image or text for **Button Style**, specify the image or text you want to appear, and then set related formatting options.
- For **Tooltip text**, add explanatory text that appears when viewers hover over the button. This text is optional and typically best used with image buttons.

To add interactivity in the dashboard,

• In the upper corner of sheet, enable the **Use as Filter** option to use selected marks in the sheet as filters for other sheets in the dashboard as shown in fig 5.11.



Fig 5.11 Filter Options in Dashboard

Actions often have unique behavior when the source or destination is a dashboard. Because a dashboard can contain multiple views, a single filter or highlight action can have broad impact. Dashboards can also contain web page objects, which you can target with interactive URL actions. Use a single view to filter other views in a dashboard. Imagine the created dashboard that contains three views about profitability: a map, a bar chart, and a table of customer names. A filter action can be used to make one of the views in the dashboard, such as the map, the "master." When the users select a region in the map, the data in the other views is filtered so that it relates to just that region.

- On the dashboard, select the view that the user wants to use as a filter.
- On the view's shortcut menu, choose **Use as Filter**. Perform the same action by clicking the Use as Filter icon as shown in fig 5.12.



Fig 5.12 Shortcut menu for 'use as Filter'

Use filter actions to filter the data on a dashboard when the data comes from multiple data sources.

5.3 FORMATTING A DASHBOARD

There is a need to resize and reorganize elements in a dashboard to provide better appearance. There are many formatting options for dashboard like, size, layout, rename, background, layout container, border and shading.

5.3.1 Dashboard Size

There are different dashboard size options to control overall dashboard size which are,

Fixed size (default): The dashboard remains the same size, regardless of the size of the window used to display it. If the dashboard is larger than the window, it becomes scrollable. You can pick from a preset size, such as Desktop Browser (the default), Small Blog, and iPad. Fixed size dashboards let you specify the exact location and position of objects, which can be useful if there are floating objects. Select this setting if you know the precise size at which your dashboard will be displayed. Published dashboards that use a fixed size can load faster because they're more likely to use a cached version on the server. (Dashboards with variable sizes need to be freshly rendered for every browser request.)

Range: The dashboard scales between minimum and maximum sizes that you specify. If the window used to display the dashboard is smaller than the minimum size, scroll bars are displayed. If it's larger than the maximum size, white space is displayed. Use this setting when we are designing for two different display sizes that need the same content and have similar shapes such as small- and medium-sized browser windows. Range also works well for mobile dashboards with vertical layouts, where the width may change to account for different mobile device widths, but the height is fixed to allow for vertical scrolling.

Automatic: The dashboard automatically resizes to fill the window used to display it. Use this setting for Tableau to take care of any resizing. For best results, use a tiled dashboard layout. In Tableau Desktop, we can create dashboard layouts for different device types to create unique layouts optimized for desktop computers, tablets, and phones. In addition to adapting to different screen sizes, each device layout can contain different items.

5.3.2 Steps to set overall dashboard size

 Under Size on the Dashboard pane, select the dashboards dimensions (such as Desktop Browser) or sizing behavior (for example, Automatic) as shown in fig 5.13.



Fig 5.13 Dashboard size setting

5.3.3 Group items using layout containers

Layout containers used to group related dashboard items together so that user can quickly position them. When we change the size and placement of items inside a container, other container items automatically adjust.

Layout container types

A horizontal layout container resizes the width of the views and objects it contains and a vertical layout container adjusts height as shown in fig 5.14.