Unit 1

charts and graphs

What is a Graph?

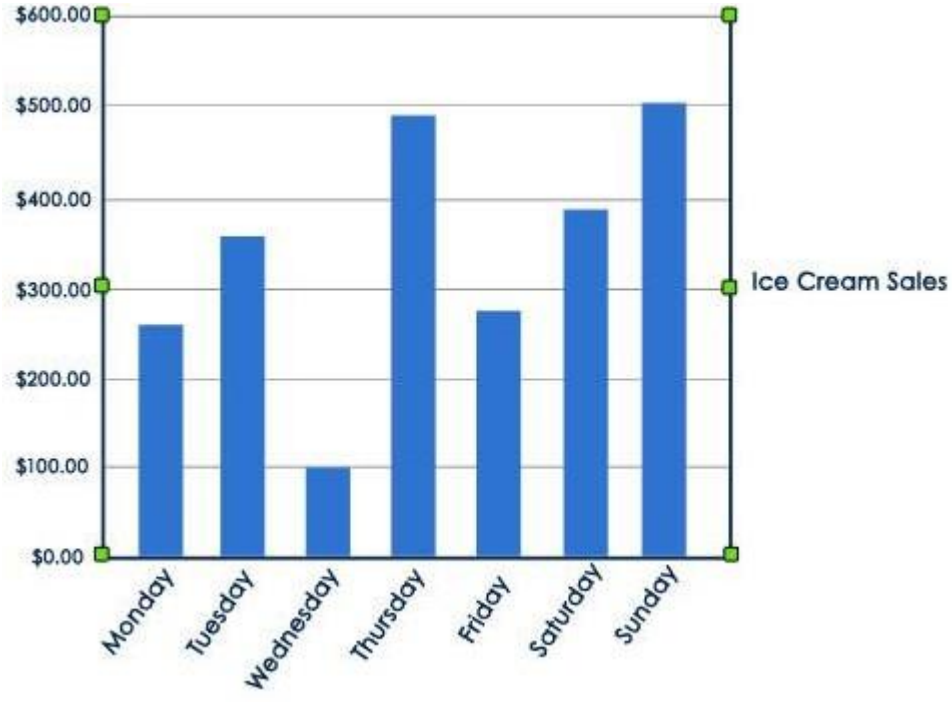
Graphs mainly focus on raw data and depict the trend overtime-related to such data. A two-dimensional graph shows the relationship between the data through a line, curve, etc., using the horizontal line along the bottom (called X-axis) and vertical line up the side (called Y-axis). As per the Advanced English Dictionary, “A Graph is a mathematical diagram that shows the relationship between two or more sets of numbers or measurements.” A graph allows the user to easily represent the values in the data through a visual representation. An example of a basic graph is shown below:

The above graph is a basic graph that allows the user to visualize that the data plotted on its Y- axes are on an increasing trend, which is shown in years on X-axes. There are two types of graphs: bar graphs and line graphs.

What is a Chart?

A chart is a type of representation of large sets of data, which makes the user understand the same better. Using the same helps predict existing data and forecast future data based on the present data pattern. A chart can take the form of a diagram, a picture, or a graph. We can transform datasets into a meaningful display of information using charts.

An example of a simple chart is shown below:



The above chart is a **simple Column Chart** depicting the company's sales of ice cream products on different days of the week. With just a glance of the same, the user can identify the week's highest and lowest sales day.

Charts can simplify data and categorize the same into easy-to-understand and analyze formats and find its excessive usage in a business where data is presented using different charts.

There are types of charts –

- Vertical bar charts
- Historical bar charts
- Stacked bar charts
- **Histograms**
- The **Pie Chart in excel**
- Line chart
- **Area Charts in Excel**

**Graphs vs Charts Comparative Table

Meaning	A graph is a chart used to show the mathematical relationship between varied data sets by plotting horizontal (X-axis) and vertical (Y-axis).	A chart represents information that can be in the form of a diagram, table, or graph. It comprises various methods for presenting large information.
Subset	All graphs are charts. It means that no matter which type of graph one uses to display the data, it will always be a type of chart subset.	All charts are not graphs. It means there can be other types of charts that are not graphs.
Data Analyzed	Graphs can be used for raw data and provide a visual representation of trends and changes in the data.	Ideal for those forms of data which can be easily structured or categorized into small subsets of simple and easily understandable figures.
Usage	Graphs find their usage more when analyzing raw data and exact numbers. As such shows, accurate numerical figures are plotted on its axes.	Charts find their excess use in business presentations and in showing survey results. Example pie charts are the most popular ones used in business presentations.
Trend Analysis	A graph is an ideal choice for those data which depicts some trend or relation	We can also use charts in those cases where the data shown does not depict any trend or relationship.

	between variables described on the graph.	
Common Types	Line Graph and bar graph.	Popular chart types are pie charts, histograms, and vertical and historical Bar Chart.

Types of graphs good for categorical data

Once the type of data, categorical or quantitative is identified, we can consider graphical representations of the data, which would be helpful for Maria to understand.

Frequency tables, pie charts, and bar charts are the most appropriate graphical displays for categorical variables. Below are a frequency table, a pie chart, and a bar graph for data concerning Mental Health Admission numbers.

Frequency Table

A table containing the counts of how often each category occurs.

Diagnosis	Count	Percent
Depression	40835	48.5%
Anxiety	29388	34.9%
OCD	5465	6.5%
Abuse	8513	10.1%
Total	84201	100.0%

Pie chart

Graphical representation for categorical data in which a circle is partitioned into “slices” on the basis of the proportions of each category.

Pie Chart of Diagnosis

Category

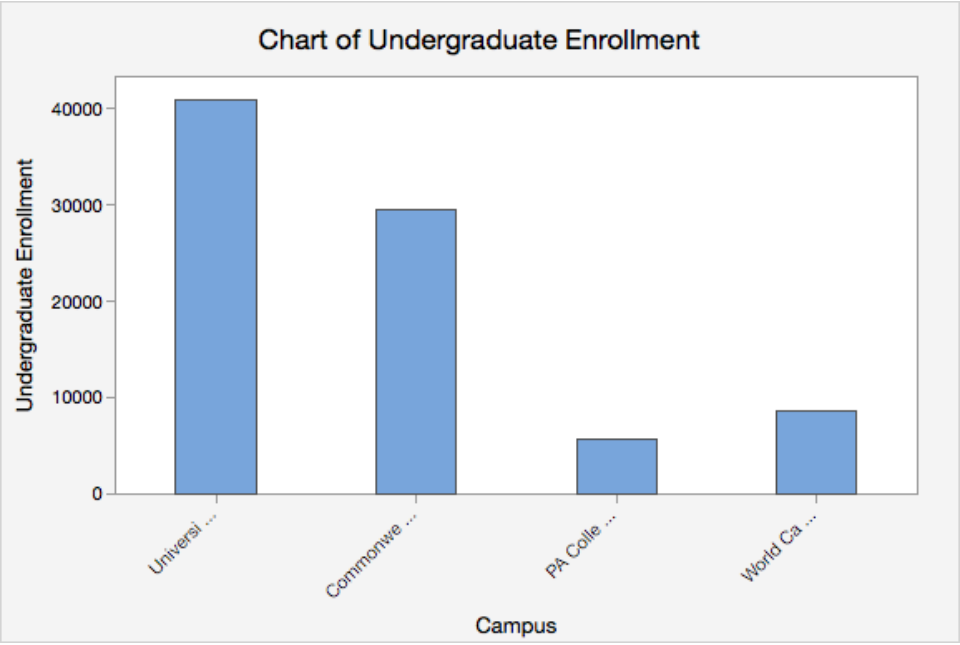
- Depression (48.5%)
- Anxiety (34.9%)
- OCD (6.5%)
- Abuse (10.1%)

Pitfalls

One of the pitfalls of a pie chart is that if the “slices” only represent percentages the reader does not know how many actual people fall in each category.

Bar chart

Graphical representation for categorical data in which vertical (or sometimes horizontal) bars are used to depict the number of experimental units in each category; bars are separated by space.



what types of charts good for time series data

TTime series line graphs and bar graphs

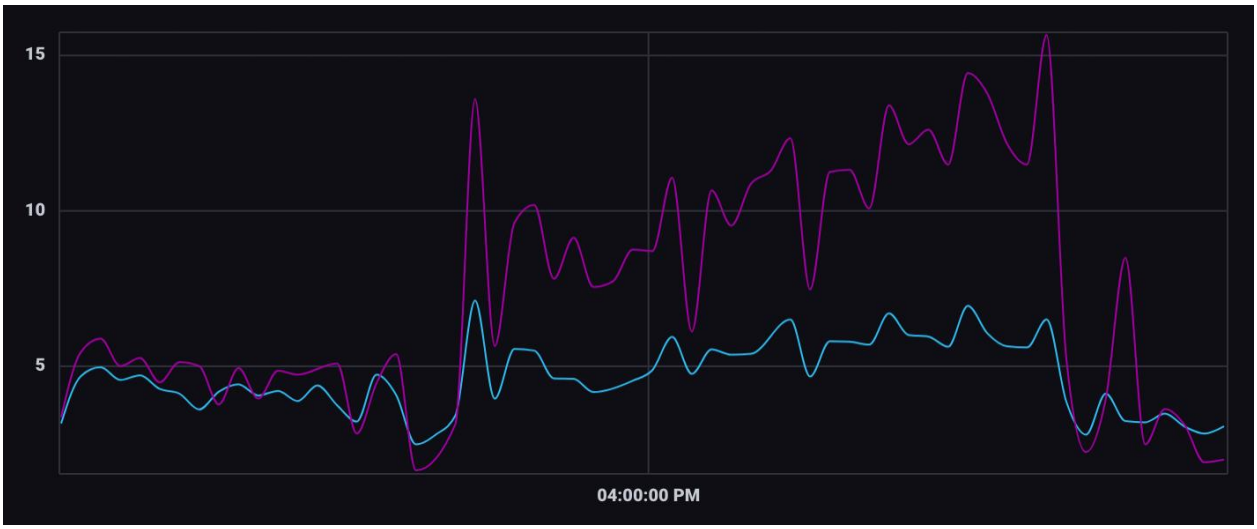
The Graph view in the InfluxDB 2.0 UI lets you select from multiple graph types such as line graphs and bar graphs (Coming).

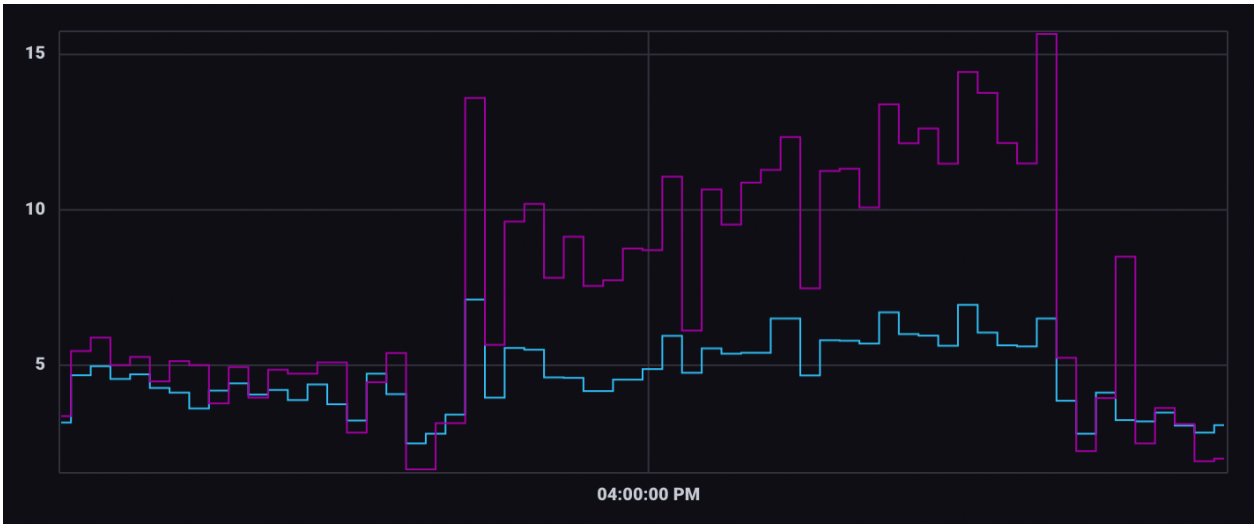
A line graph is the simplest way to represent time series data. It helps the viewer get a quick sense of how something has changed over time. A line graph uses points connected by lines (also called trend lines) to show how a dependent variable and independent variable changed:

- An independent variable, true to its name, remains unaffected by other parameters.
- The dependent variable depends on how the independent variable changes.

For temporal visualizations, time is always the independent variable, which is plotted on the horizontal axis. Then the dependent variable is plotted on the vertical axis.

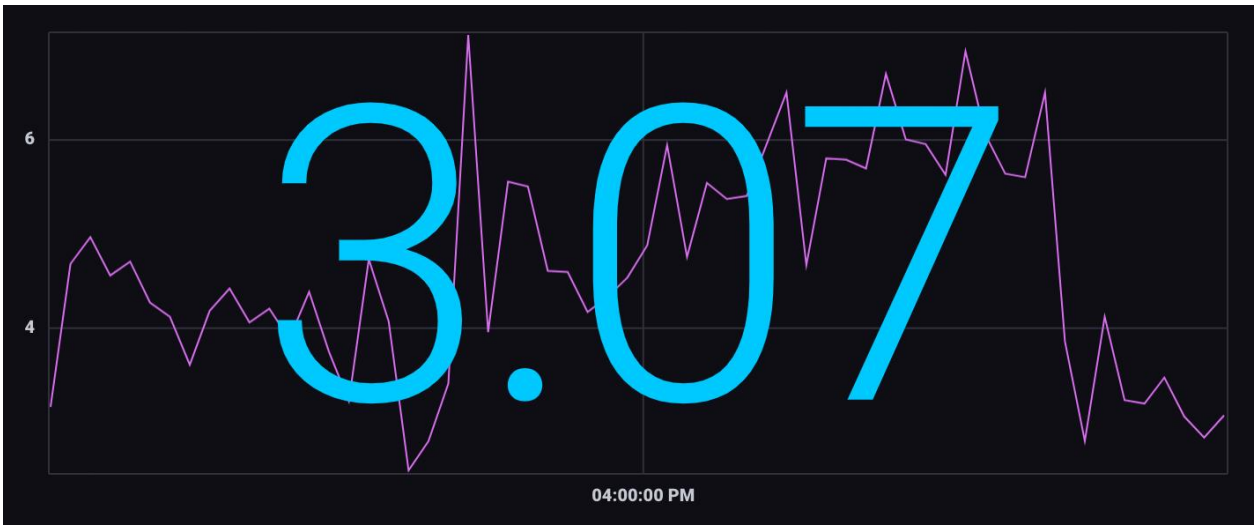
While the above graph is an example of a line graph with linear interpolation (interpolation is the estimation of a value within two known values in a sequence of values), the below two graphs depict smooth interpolation and step interpolation.





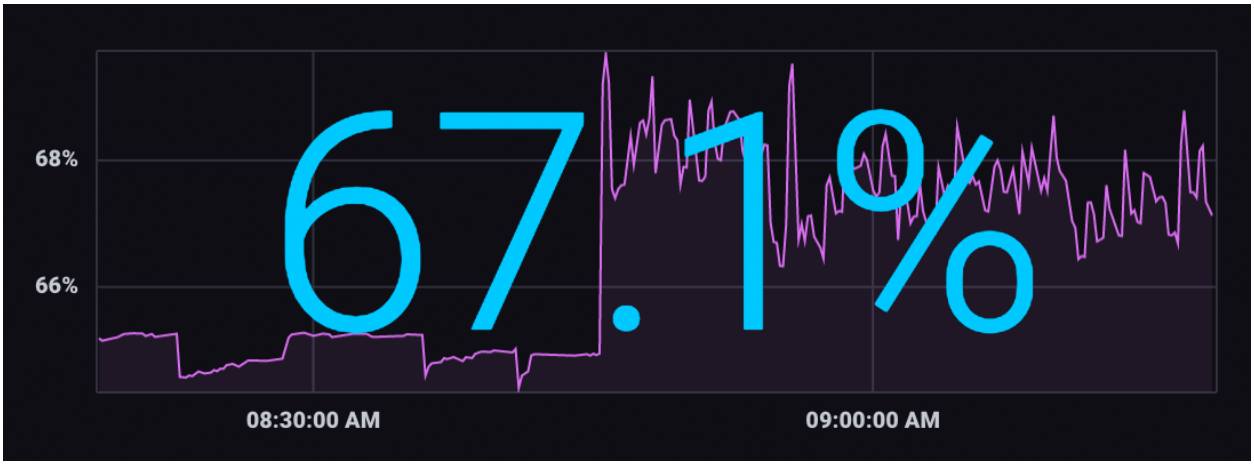
Graph + Single Stat visualization for time series data

The Graph + Single Stat view displays the specified time series in a line graph and overlays the single most recent value as a large numeric value. The Single Stat visualization displays a single numeric data point. It uses the latest point in the first table (or series) returned by the query.



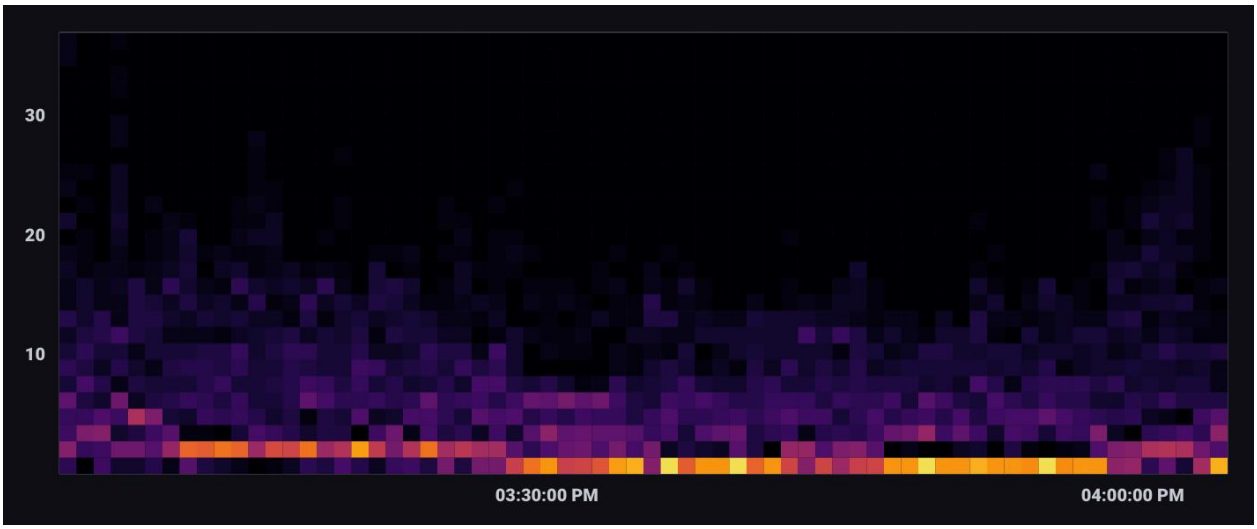
The primary use case for the Graph + Single Stat visualization is to show the current or latest value as well as historical values.

The following example shows the current percentage of memory used as well as memory usage over time.

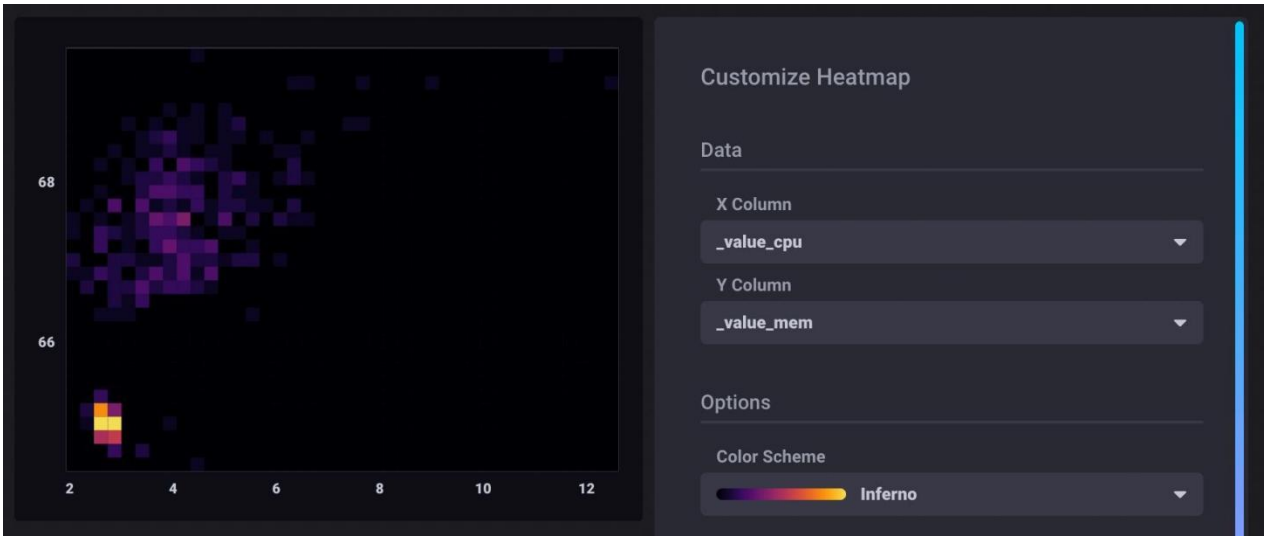


Heatmap

A Heatmap displays the distribution of data on an x and y axes where color represents different concentrations of data points. Heatmaps divide data points into “bins” – segments of the visualization with upper and lower bounds for both X and Y axes. The Bin Size option determines the bounds for each bin. The total number of points that fall within a bin determine its value and color. Warmer or brighter colors represent higher bin values or density of points within the bin.

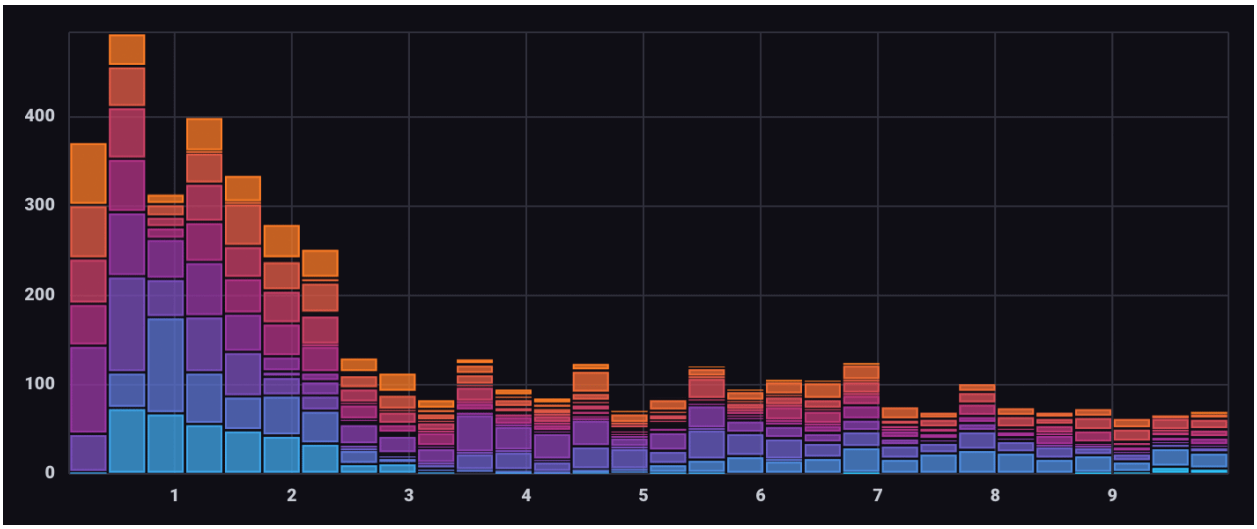


A heatmap, as shown below, can be used to visualize correlation.

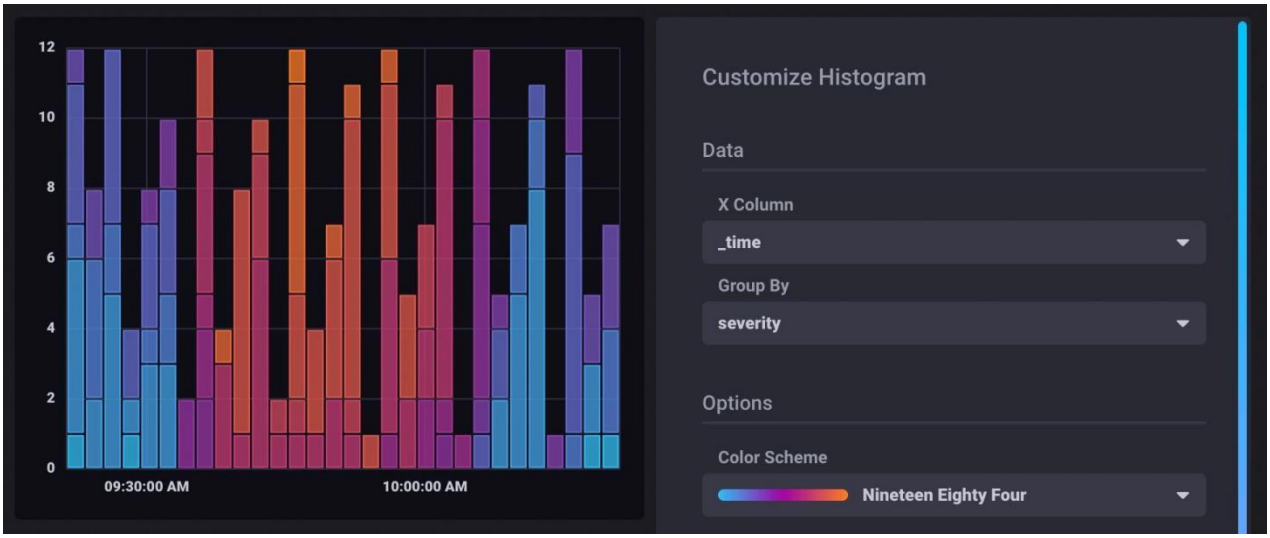


Histogram

A histogram is a way to view the distribution of data. The y-axis is dedicated to count, and the x-axis is divided into bins. The Histogram visualization is a bar graph that displays the number of data points that fall within “bins” – segments of the X axis with upper and lower bounds.



For example, the below histogram shows error counts by severity over time.



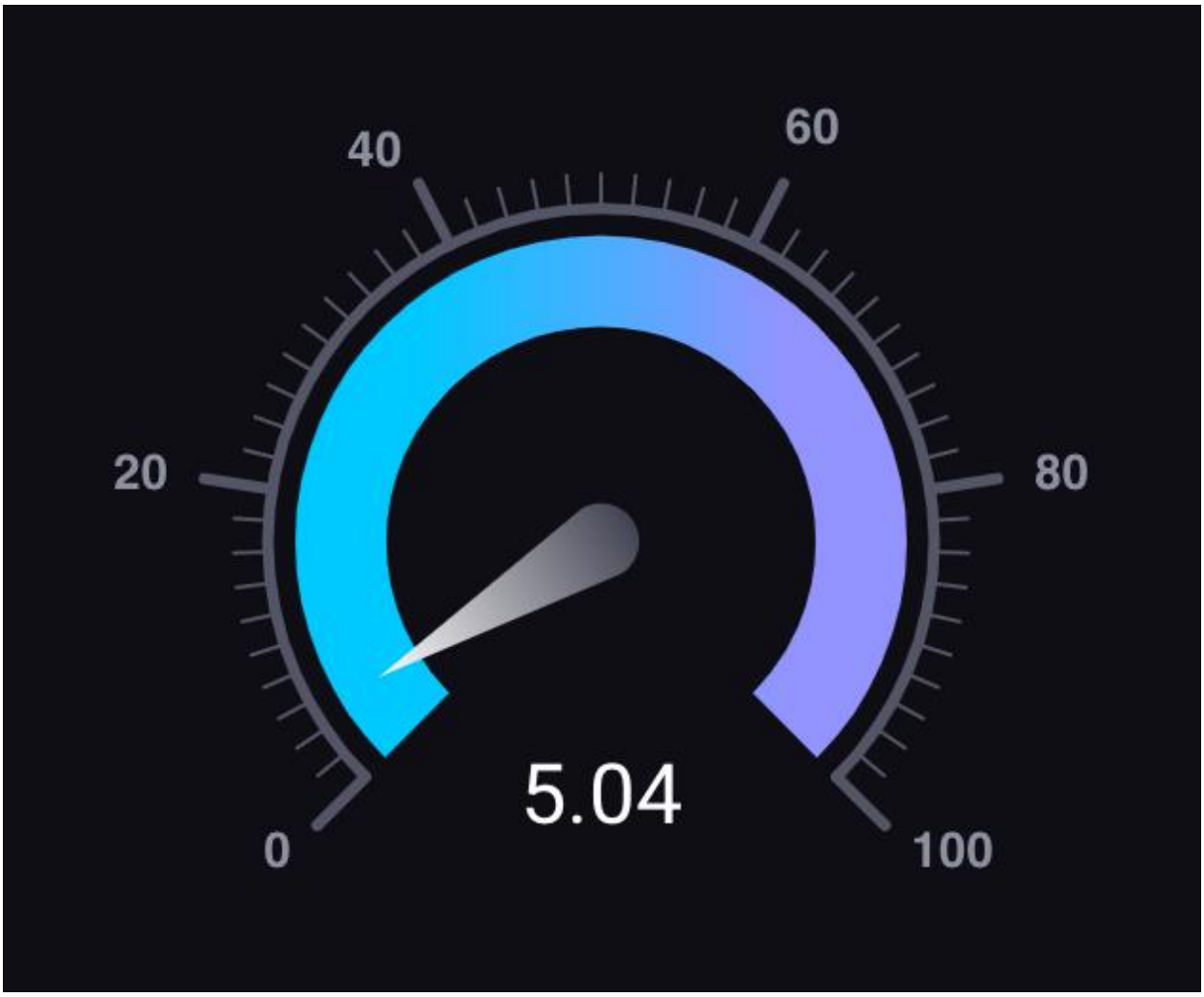
Single Stat

The Single Stat view displays the most recent value of the specified time series as a numerical value. It uses the latest point in the first table (or series) returned by the query.

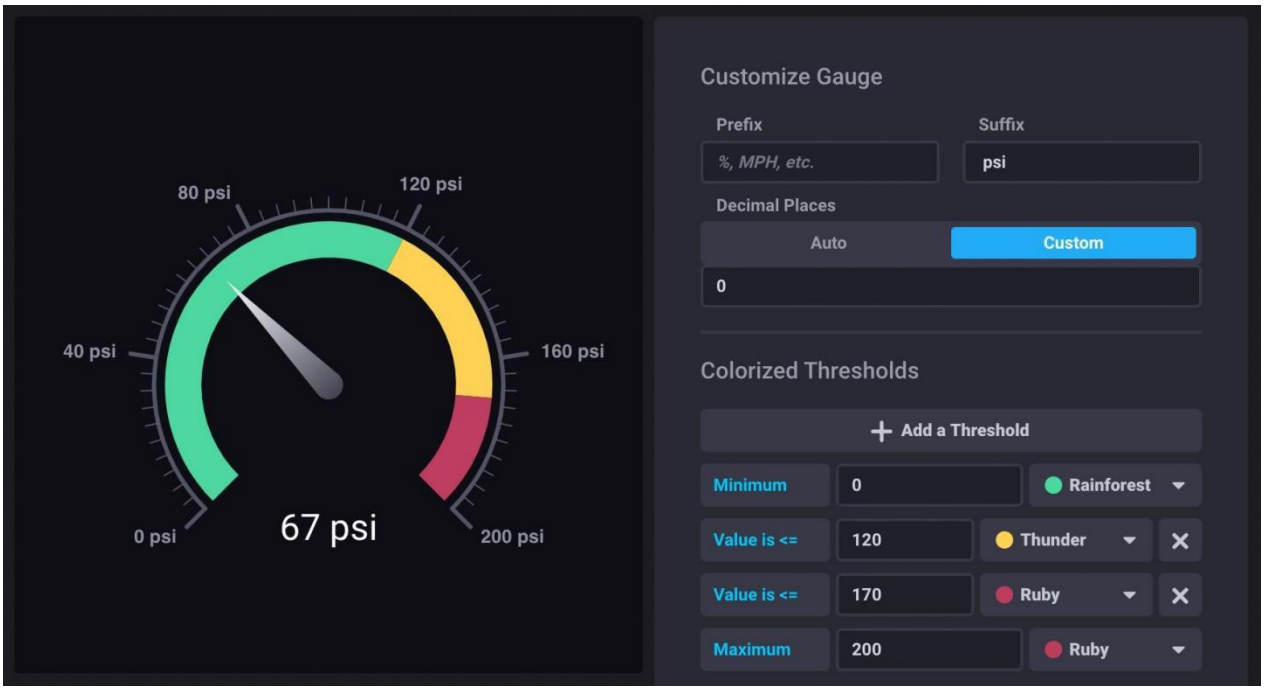
The following visualization example shows the current memory usage as a percentage.

Gauge

The Gauge view displays the single most recent value for a time series in a gauge view. Gauge visualizations are useful for showing the current value of a metric and displaying where it falls within a spectrum.

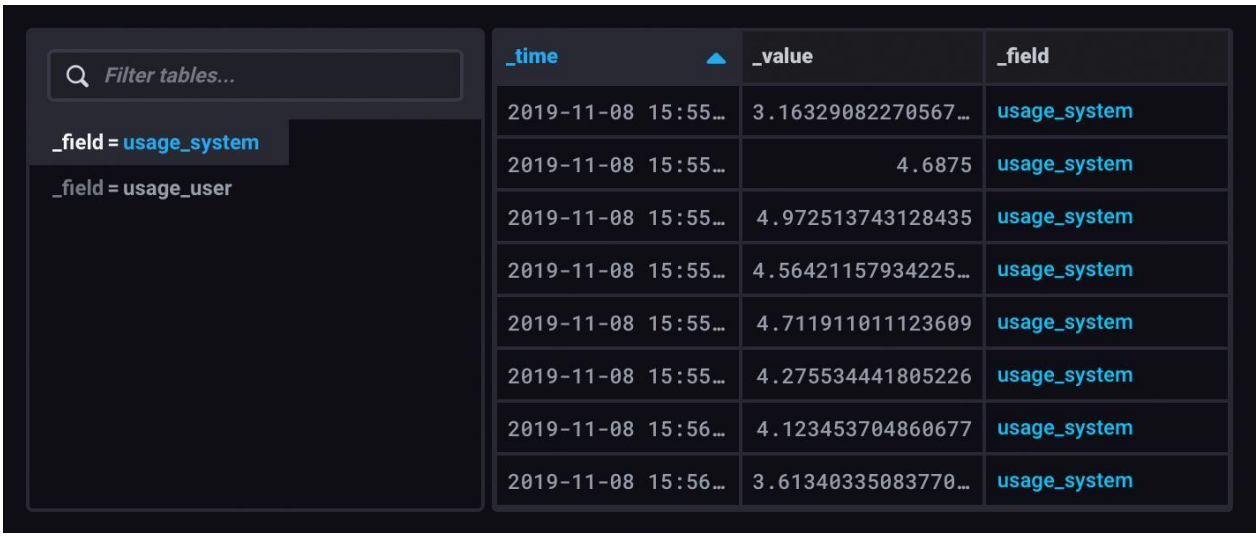


The following gauge visualization displays the pressure of steam pipes in a facility.

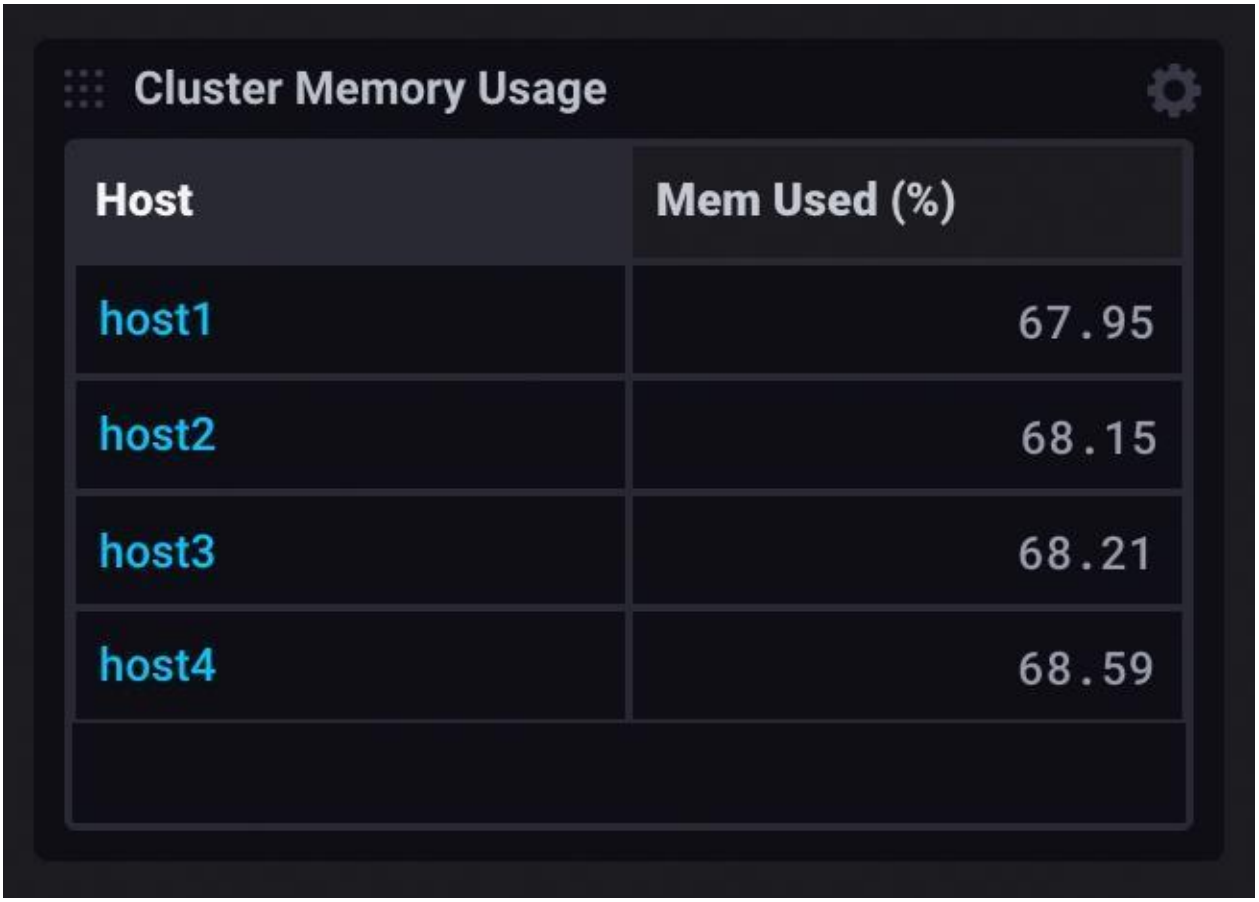


Table

The Table visualization option displays the results of queries in a tabular view, which is sometimes easier to analyze than graph views of data.

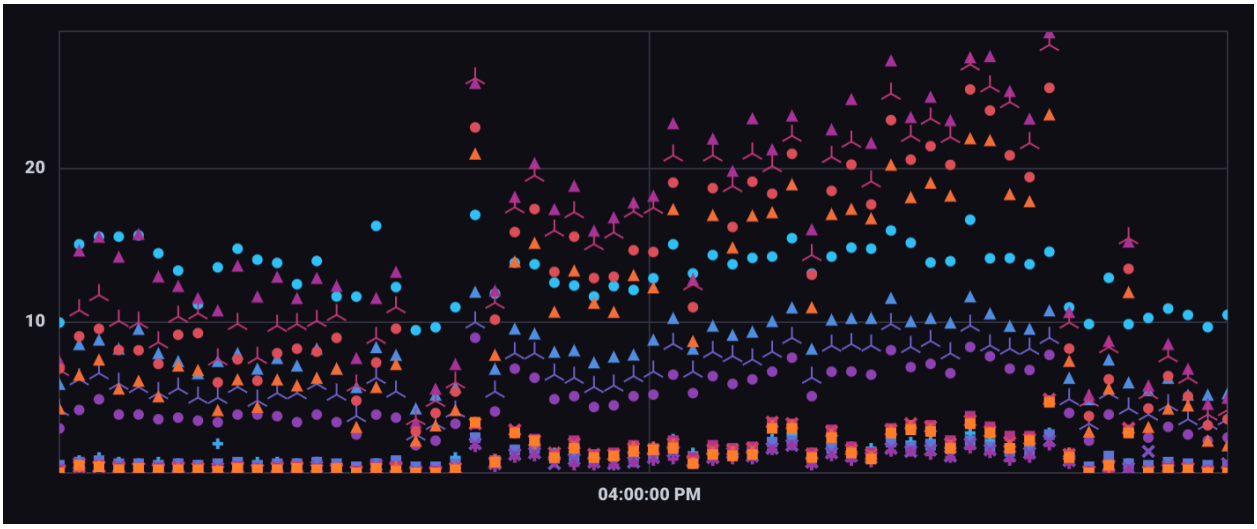


The table visualization renders queried data in structured, easy-to-read tables. Columns and rows match those in the query output. Tables are helpful when displaying many human-readable metrics in a dashboard such as cluster statistics or log messages.



Scatter

The Scatter view uses a scatter plot to display time series data. A scatter plot can have anything on the horizontal axis, in any transformation, and points are not connected or ordered.



The scatter visualization maps each data point to X and Y coordinates.

The following example explores possible correlation between CPU and Memory usage. In the Scatter visualization controls, points are differentiated based on their group keys.

Types of charts or graphs good for Continuous Data

Continuous data

Continuous variables can take on any numeric value, and it can be meaningfully divided into smaller increments, including fractional and decimal values. There are an infinite number of possible values between any two values. Typically, you measure continuous variables on a scale. For example, when you measure height, weight, and temperature, you have continuous data.

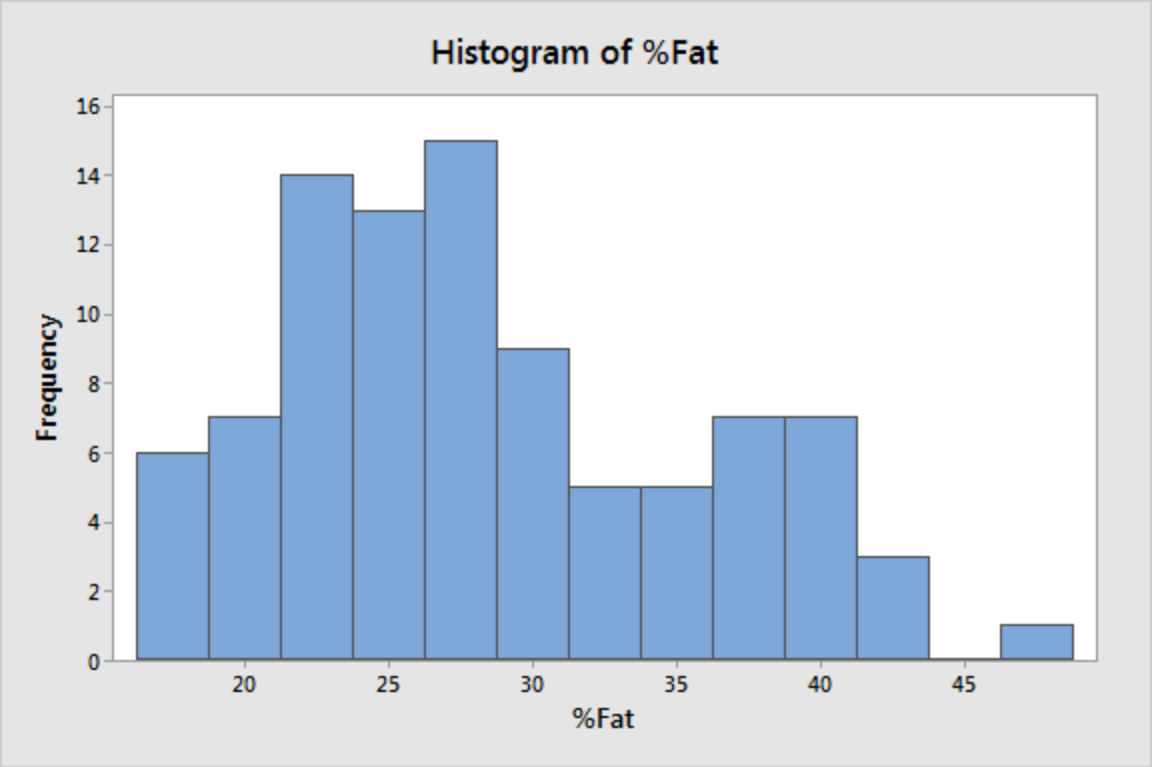
With continuous variables, you can assess measures of central tendency and variability, such as the mean, median, distribution, range, and standard deviation. For example, the mean height in the U.S. is 5 feet 9 inches for men and 5 feet 4 inches for women.

Related posts: Measure of Central Tendency and Measures of Variability

How to graph continuous data

%Fat
23.9
28.8
32.4
25.8
22.5

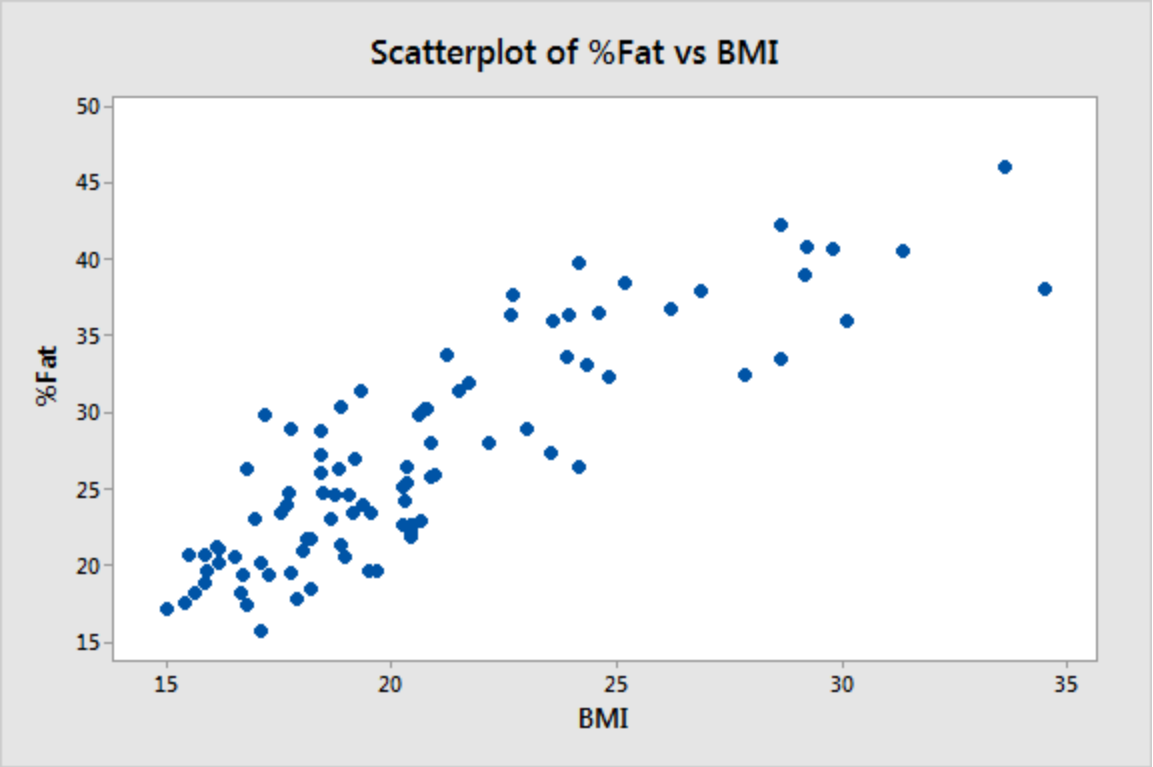
Histograms are a standard way to graph continuous variables because they show the distribution of the values. The histogram below helps you determine whether the distribution of body fat percentage values for adolescent girls are symmetric or skewed; understand the range of values; and, identify where the most common values fall.



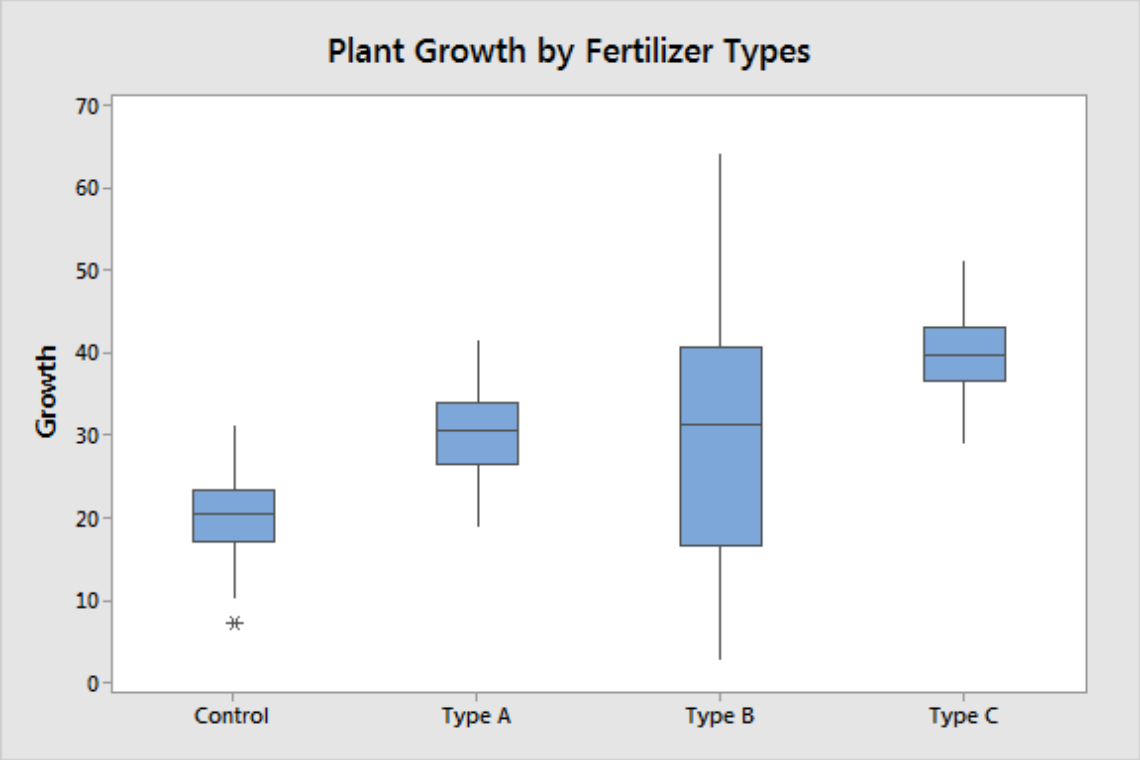
Dot plots provide the same types of information as histograms. For more information, read my Guide to Dot Plots.

Related post: Using Histograms to Understand Your Data

When you have two continuous variables, you can graph them using a scatterplot. The scatterplot shows how the body fat percentage tends to rise as BMI increases. Use correlation to assess the strength of this relationship or regression analysis to derive the equation for the line that provides the best fit for these data. For more information, read my Guide to Scatterplots.



When you have continuous variables that are divided into groups, you can use a boxplot to display the central tendency and spread of each group. Fertilizer Type C is associated with the highest plant growth while Type B produces the greatest variability.



Please notice how with continuous variables you can assess the wide variety of properties that I illustrate above. You'll see a contrast when we get to qualitative variables.

Related posts: Graphing Continuous Data by Groups: Boxplots vs. Individual Value Plots and Time Series Plots

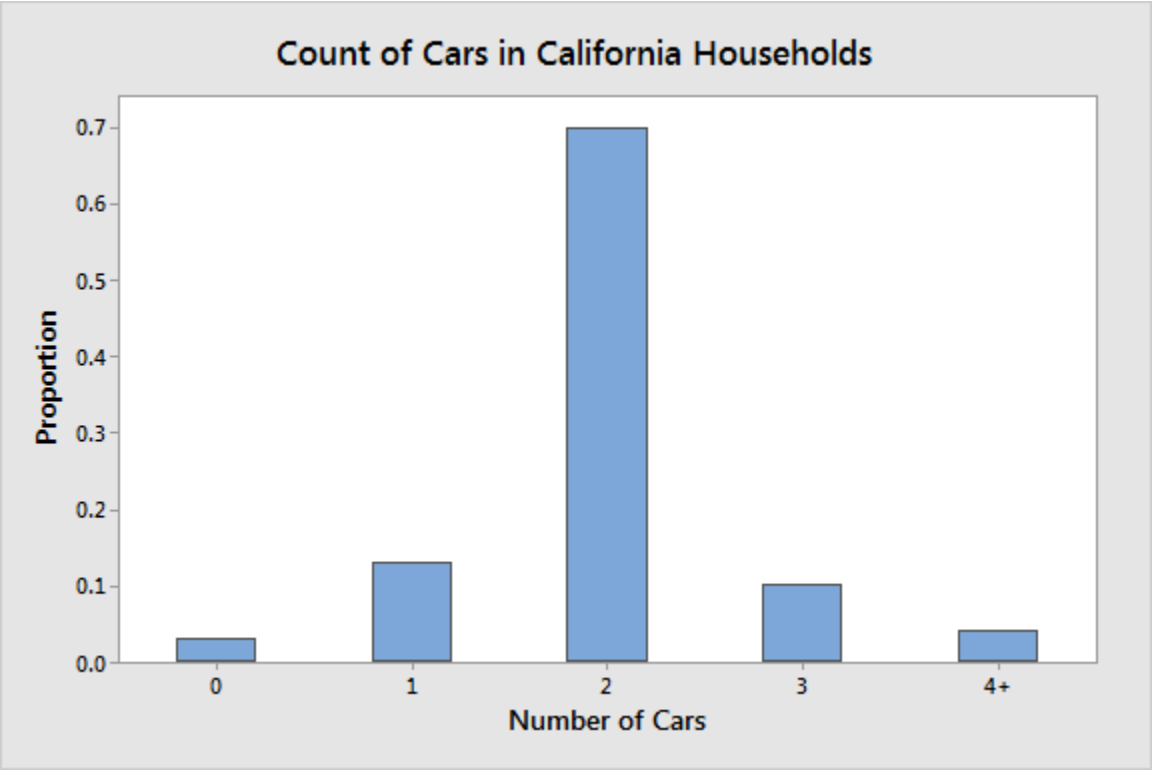
Discrete data

Car Count
2
1
2
2
3

Discrete quantitative data are a count of the presence of a characteristic, result, item, or activity. These measures cannot be meaningfully divided into smaller increments. For example, a single household can have 1 or 2 cars, but it cannot have 1.6. There are a finite number of possible values that you can record for an observation.

With discrete variables, you can calculate and assess a rate of occurrence or a summary of the count, such as the mean, sum, and standard deviation. For example, U.S. households had an average of 2.11 vehicles in 2014.

Bar charts are a standard way to graph discrete variables. Each bar represents a distinct value, and the height represents its proportion in the entire sample.



See how I used a line plot to graph the count of coronavirus cases by country.

Related posts: Guide to Bar Charts and Guide to Line Charts

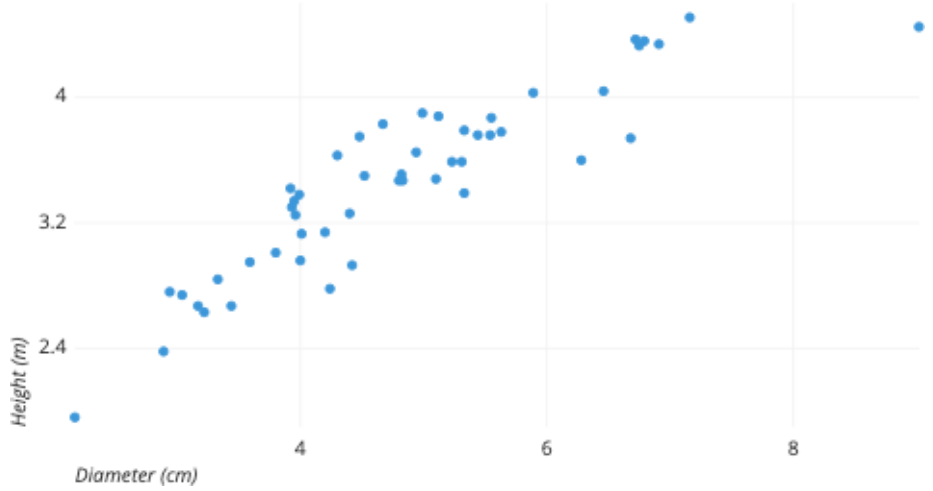
Qualitative Data: Categorical, Binary, and Ordinal

When you record information that categorizes your observations, you are collecting qualitative data. There are three types of qualitative variables—categorical, binary, and ordinal. With these data types, you're often interested in the proportions of each category. Consequently, bar charts and pie charts are conventional methods for graphing qualitative variables because they are useful for displaying the relative percentage of each group out of the entire sample.

As I mentioned in the section about continuous variables, notice how we learn much less from qualitative data. I highlight this aspect in the section about binary variables. In cases where you have a choice about recording a characteristic as a continuous or qualitative variable, the best practice is to record the continuous data because you can learn so much more.

Types of charts good for bivariate distribution:

A scatter plot (aka scatter chart, scatter graph) uses dots to represent values for two different numeric variables. The position of each dot on the horizontal and vertical axis indicates values for an individual data point. Scatter plots are used to observe relationships between variables.



The example scatter plot above shows the diameters and heights for a sample of fictional trees. Each dot represents a single tree; each point's

horizontal position indicates that tree's diameter (in centimeters) and the vertical position indicates that tree's height (in meters). From the plot, we can see a generally tight positive correlation between a tree's diameter and its height. We can also observe an outlier point, a tree that has a much larger diameter than the others. This tree appears fairly short for its girth, which might warrant further investigation.

Unit II

BUSINESS INTELLIGENCE

Business intelligence (BI) refers to the procedural and technical infrastructure that collects, stores, and analyzes the data produced by a company's activities.

BI is a broad term that encompasses data mining, process analysis, performance benchmarking, and descriptive analytics. BI parses all the data generated by a business and presents easy-to-digest reports, performance measures, and trendsthat inform management decisions.

KEY TAKEAWAYS

- BI represents the technical infrastructure that collects, stores, and analyzes company data.
- BI parses data and produces reports and information that help managers to make better decisions.
- Software companies produce BI solutions for companies that wish to make better use of their data.
- BI tools and software come in a wide variety of forms such as spreadsheets, reporting/query software, data visualization software, data mining tools, and online analytical processing (OLAP).
- Self-service BI is an approach to analytics that allows individuals without a technical background to access and explore data.

Types of BI Tools and Software

BI tools and software come in a wide variety of forms. Let's take a quick look at somecommon types of BI solutions.

- **Spreadsheets:** Spreadsheets like Microsoft Excel and Google Docs are some of the most widely used BI tools.
- **Reporting software:** Reporting software is used to report, organize, filter, and display data.

- **Data visualization software:** Data visualization software translates datasets into easy-to-read, visually appealing graphical representations to quickly gain insights.
- **Data mining tools:** Data mining tools "mine" large amounts of data for patterns using things like artificial intelligence, machine learning, and statistics.

- **Online analytical processing (OLAP):** OLAP tools allow users to analyze datasets from a wide variety of angles based on different business perspectives.

Benefits of Business Intelligence

There are many reasons why companies adopt BI. Many use it to support functionsas diverse as hiring, compliance, production, and marketing. BI is a core business value; it is difficult to find a business area that does not benefit from better information to work with.

Some of the many benefits companies can experience after adopting BI into their business models include faster, more accurate reporting and analysis, improved dataquality, better employee satisfaction, reduced costs, and increased revenues, and the ability to make better business decisions.

BI was derived to help businesses avoid the problem of "garbage in and garbage out," resulting from inaccurate or insufficient data analysis.

If, for example, you are in charge of production schedules for several beverage factories and sales are showing strong month-over-month growth in a particular region, you can approve extra shifts in near real-time to ensure your factories canmeet demand.

Common applications of business intelligence:

Here are four key business intelligence applications that can help improve your organization's operations:

1. Sales Intelligence

A key application of BI focuses on where your business meets the customer. Customernegotiation is a crucial skill that every organization's sales department should foster.
Sometimes it can be hard to move leads along the pipeline and convince potential clientsto buy your product or service. Through the applications of business analytics and intelligence, this process is becoming smoother and more predictable.

Business intelligence collects data on specific KPIs like customer demographics, conversion rates, sales metrics, etc. Then it organizes this data into structured visualizations like graphs, pie charts and scattergrams. Users can identify trends from thisdata that provide insights into customer behavior and business operations. Knowing thecustomer means you can better serve them!

2. Visualization

Business intelligence software utilizes a range of data analytic tools that are designed toanalyze and manage data related to your business operations. This data, presented in theform of visualizations, allows the organization to monitor logistics, sales, productivity andmuch more. Some business intelligence platforms offer custom reporting abilities whereusers can specify their parameters. Others offer out-of-the-box reporting templates thatalready include industry-standard metrics.

By presenting the data in intuitive visuals and easy to comprehend formats, business intelligence systems enable even the least experienced employee to draw insights fromdata.

3. Reporting

A crucial business application of BI is reporting. As we've covered, business intelligence tools collect and study unstructured sets of data in addition to organizing and using themto generate a range of different types of reports. These can include staffing, expenses, sales, customer services, and other processes.

Reporting and data analysis are similar, but they vary significantly in purpose, delivery, tasks and value. Reporting is the process of organizing data in summaries with the intention of monitoring business performance. Analysis is the process of exploring datato extract insights that can be applied to improve business practices.

4. Performance management

With BI applications, organizations can monitor goal progress based on pre-defined or customizable timeframes. The data-driven goals may include project completion deadlines, target delivery time, or sales goals. For example, if you'd like to reach a certainsales goal, your BI system can analyze previous months of data and suggest a reasonable goal to aim for based on past performance.

Users can also monitor goal fulfillment and use progress data to gauge the overallproductivity of an organization. information is always readily accessible. This savesbusinesses time and money

What is a business intelligence (BI) tool?

Business intelligence tools collect, process, and analyze large amounts of structured and unstructured data from both internal and external systems. Data sources might include documents, images, email, videos, journals, books, social media posts, files, and more. BI tools find this information through queries, which

can present the data in user-friendly formats such as reports, dashboards, charts,and graphs.

The tools can perform functions such as data mining, data visualization, performance management, analytics, reporting, text mining, predictive analytics,and much more. As a result, employees can harness this information to make better decisions based on predictions, market trends, and key performance indicators (KPIs).

1. Microsoft Power BI

One of the most popular BI tools is Power BI, offered by leading software giant Microsoft. This tool is downloadable software, so you can choose to run analytics either on the cloud or in a reporting server. Syncing with sources such as Facebook,Oracle, and more, generate reports and dashboards in minutes with this interactivetool. It comes with built-in AI capabilities, Excel integration, and data connectors, and offers end-to-end data encryption and real-time access monitoring.

2. Tableau

Tableau is known for its user-friendly data visualization capabilities, but it can do more than make pretty charts. Their offering includes live visual analytics, an interface that allows users to drag and drop buttons to spot trends in data quickly.The tool supports data sources such as Microsoft Excel, Box, PDF files, Google Analytics, and more. Its versatility extends to being able to connect with most databases.

3. QlikSense

QlikSense is a BI tool that emphasizes a self-service approach, meaning that it supports a wide range of analytics use cases, from guided apps and dashboards tocustom and embedded analytics. It offers a user-friendly interface optimized for touchscreens, sophisticated AI, and high-performance cloud platforms. Its associative exploration capability, Search & Conversational Analytics, allows userstо ask questions and uncover actionable insights, which helps increase data literacy for those new to using BI tools.

4. Dundas BI

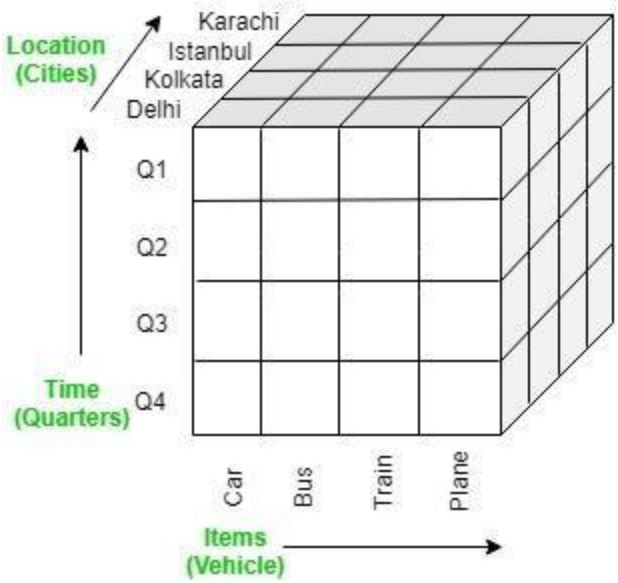
Dundas BI is a browser-based BI tool that's been around for 25 years. Like Tableau,Dundas BI features a drag-and-drop function that allows users to analyze data ontheir own, without involving their IT team. The tool is known for its simplicity andflexibility through interactive dashboards, reports, and visual analytics. Since its inception as a data visualization tool in 1992, it has evolved into an end-to-end analytics platform that is able to compete with the new BI tools available today.

5. Sisense

Sisense is a user-friendly BI tool that focuses on being simplified and streamlined. With this tool, you can export data from sources like Google Analytics, Salesforce,and more. Its in-chip technology allows for faster data processing compared to other tools. Key features include the ability to embed white-label analytics, meaning a company can fully customize the services to its needs. Like others, it hasa drag-and-drop feature. Sisense allows you to share reports and dashboards withyour team members as well as externally.

What is OLAP (Online Analytical Processing)?

OLAP stands for ***Online Analytical Processing*** Server. It is a software technology that allows users to analyze information from multiple database systems at the same time. It is based on multidimensional data model and allows the user to query on multi-dimensional data (eg. Delhi -> 2018 -> Sales data). OLAP databases are divided into one or more cubes and these cubes are known as *Hyper-cubes*.

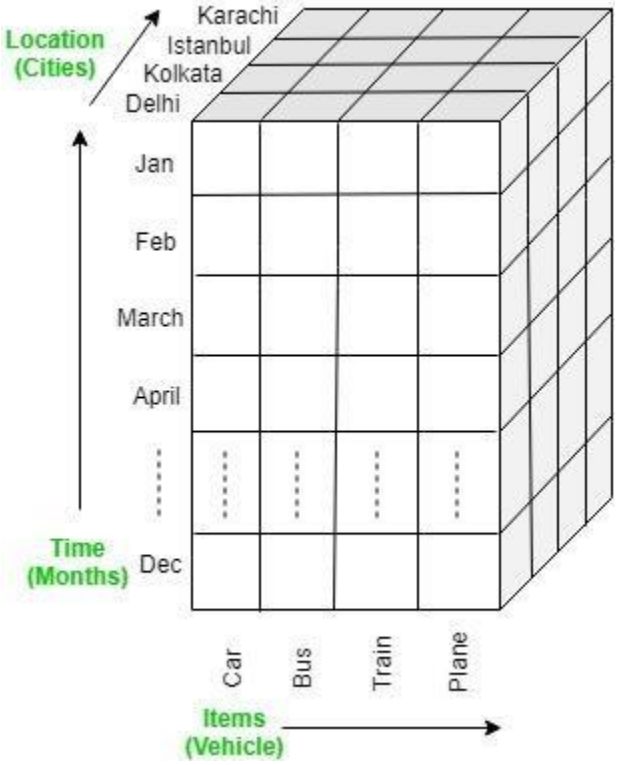


OLAP operations:

There are five basic analytical operations that can be performed on an OLAP cube:

1. **Drill down:** In drill-down operation, the less detailed data is converted into highly detailed data. It can be done by:
    - Moving down in the concept hierarchy
    - Adding a new dimension

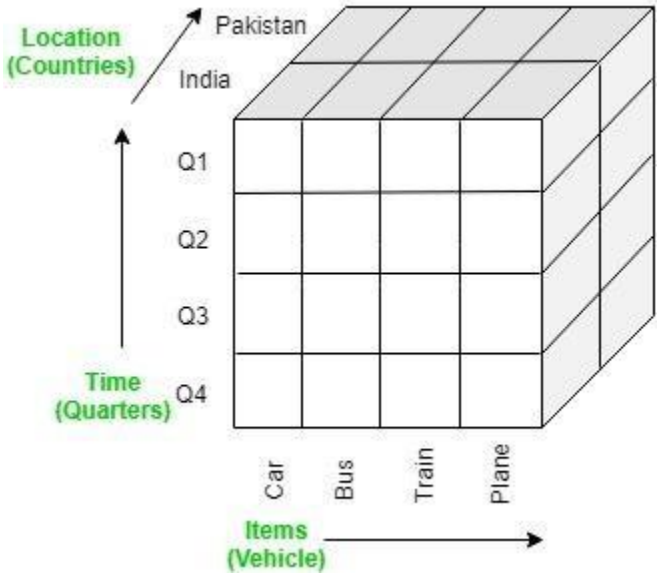
- In the cube given in overview section, the drill down operation is performed by moving down in the concept hierarchy of *Time* dimension (Quarter -> Month).



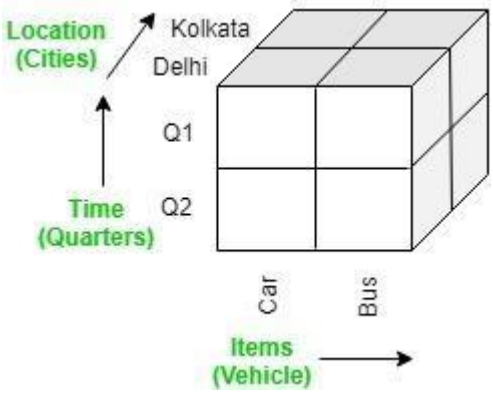
2. **Roll up:** It is just opposite of the drill-down operation. It performs aggregationon the OLAP cube. It can be done by:
    - Climbing up in the concept hierarchy
    - Reducing the dimensions

- In the cube given in the overview section, the roll-up operation is performedby climbing up in the concept hierarchy of *Location* dimension (City ->

Country).

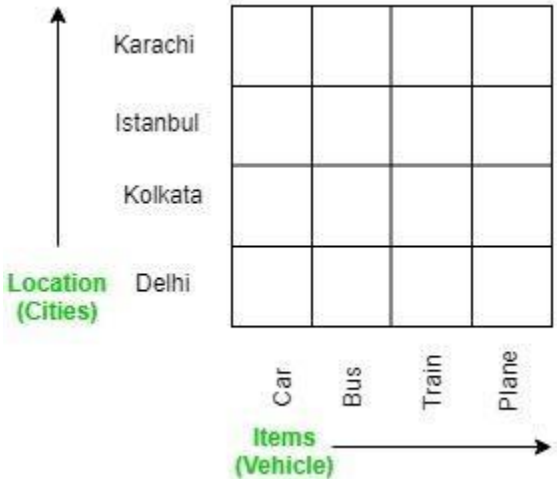


3. **Dice:** It selects a sub-cube from the OLAP cube by selecting two or more dimensions. In the cube given in the overview section, a sub-cube is selectedby selecting following dimensions with criteria:
    - Location = “Delhi” or “Kolkata”
    - Time = “Q1” or “Q2”
    - Item = “Car” or “Bus”

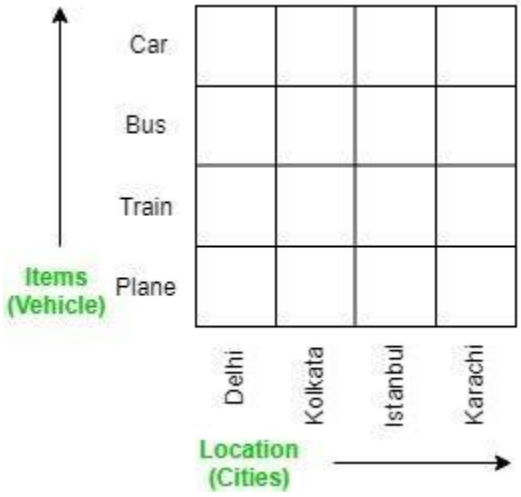


4. **Slice:** It selects a single dimension from the OLAP cube which results in anew sub-cube creation. In the cube given in the overview section, Slice is

performed on the dimension Time = “Q1”.



5. **Pivot:** It is also known as *rotation* operation as it rotates the current view to get a new view of the representation. In the sub-cube obtained after the sliceoperation, performing pivot operation gives a new view of it.



6.

Types of OLAP systems

OLAP (online analytical processing) systems typically fall into one of threetypes:

Multidimensional OLAP (MOLAP) is OLAP that indexes directly into a multidimensional database.

Relational OLAP (ROLAP) is OLAP that performs dynamic multidimensionalanalysis of data stored in a relational database.

Hybrid OLAP (HOLAP) is a combination of ROLAP and MOLAP. HOLAP was developed to combine the greater data capacity of ROLAP with the superior processing capability of MOLAP.

Trends

In technical analysis, trends are identified by trend lines or price action that highlight when the price is making higher swing highs and higher swing lows for an uptrend, or lower swing lows and lower swing highs for a downtrend. The three basic types of trends are up, down, and sideways.

An uptrend is marked by an overall increase in price. Nothing moves straight up for long, so there will always be oscillations, but the overall direction needs to be higher.

A downtrend occurs when the price of an asset moves lower over a period of time. While the price may move intermittently higher or lower, downtrends are characterized by lower peaks and lower troughs over time.

Trends may be discovered in the short, medium, and long term. Generally, investors take positions in assets that will be profitable as long as the current trend continues. Taking positions that profit only if a trend reverses is riskier. Analysts use trendlines and channels, which are essentially boundaries for price fluctuations, in an attempt to spot and define trends. Upward trends are characterized by an asset price hitting a series of higher highs and higher lows, while downward trends are marked by lower highs and lower lows. Most traders trade in the direction of the trend. Traders who go opposite the trend are called contrarian investors.

Alerts

A Alert is a series of data that repeats in a recognizable way. It can be identified in the history of the asset being evaluated or other assets with similar characteristics. Alerts often include the study of sale volume, as well as price. Alerts can occur within a downward or upward trend, or they can mark the beginning of a new trend.

Alerts are the distinctive formations created by the movements of security prices on a chart. A pattern is identified by a line that connects common price points, such as closing prices or highs or lows, during a specific period of time. Chartists seek to identify Alerts as a way to anticipate the future direction of a security's price.

There are bottoming, topping, and continuation Alerts. A "follow-through day" pattern is an example of a pattern used by some analysts to identify market bottoms. The "head-and-shoulders" topping pattern is popular among day and swing traders, while continuation Alerts include the "cup-and-handle," "flat base," and "three weeks tight."

"The trend is your friend" is a common catchphrase among technical analysts. A trend can often be found by establishing a line chart. A trend line is the line formed between a high and a low. If

that line is going up, the trend is up. If the trend line is sloping downward, the trend is down. Trend lines are the foundation for most chart Alerts.

Self Service Reporting.

The relevance of data with the success of an enterprise goes hand-in-hand. What do we know about Self Service Reporting?
Well, the data is compelling. However, it is self-service reporting that offers all the access we require to the power. Comprehensive analysis is the prerequisite for IT leaders who face challenges with data accessibility that comes with the changing infrastructure management. With Self-Service Business Intelligence, you can profess about the future of your organization and gain access to thorough analysis, reports, and dashboards.
Business Intelligence (BI) solutions require individual users to be able to access reports and dashboards. The need to hire a data-scientist for every statistical move to be taken within an organization is increasingly becoming redundant. Providing users access to data to help them stay informed with the changing trends is what self service reporting is all about. It is irrelevant as to which domain or analytical background the user belongs.
Every company requires a self service reporting tool for the kind of benefits it offers. Not only to the end user, but for IT and the company itself.

Benefits of Self Service Reporting tool

- It improves flexibility and agility within the business departments by providing user independence.
- IT departments are not overloaded in data requests or drowned in information. Their primary and straightforward tasks are removed or reduced, which reduces their overall workload.
- IT department can focus more on value-added tasks and help the organization with a higher degree of expertise from their department.
- Tackle business problems at the right time with self service BI's flexibility. The flexibility it offers to the users with the information is precious. With, ad hoc reporting, reports, and dashboards, once can take decisions time specifically.
- Efficiency is the key and the most promising attribute of self-service analytics tools; with it, there is minimal dependence on resources (external). It eliminates the rapid translation process and replaces it with business needs.
- Make quicker decisions based on data and compete with your competition in the forefront.
- IT is relieved of the tedious changing requests that come from various departments.
- Quick insights and data discovery are the primary mottoes of self-service business intelligence. It brings all the power to the users. Well, in the past enterprise space, traditional BI was rigid, slow, time-consuming and a burden for IT teams. That said, let's weigh in the pros and cons that are making self-service reporting emerge and unleash the power of data and analytics.

Cons of Traditional BI	Pros of Self-Service Reporting
The technology is way past its prime; it is around 20 years old. Traditional BI is a design with architectural limitations, i.e., memory, hardware, CPU, etc. The age-old use of filters, disk space access, aggregation, etc. requires a lot of technologies	Advanced technology for any operations, any tasks are performed on the go, user independence. Reports and dashboards replace any raw data existing and give practical and time-saving information.

Cons of Traditional BI	Pros of Self-Service Reporting
to support the architecture.	
It has only a pre-defined view on data, before the users view the data, it gets aggregated.	Decreases longer turnaround times and improves productivity with predictive analysis.

With the dreadful technology dying with age, business intelligence (BI) solutions are paving the way to transitioning and replacing the traditional BI. Users require advanced analytics and search-based analytics. The liberation from conventional systems and access to easy-to-use tools helps the users to gain access to data and flexibility in performing operations.

Can they two co-exist?

Many organizations use both the business intelligence tools, traditional and self-service BI. That is to say; it is a common practice for businesses to perform their daily business operations, for functional reporting with conventional BI.
Traditional BI is still great for a few organizations for dash boarding and compliance reporting. Organizations that have been using traditional BI for a long time still use it to seek answers about what happened back then.
With the changing dimensions, what users require is individual power. That being said, traditional BI is not designed to provide the users what it wants most. Which is the reason for self-service to be the preferred one?

Characteristics of a good Report:

the reports have some objective and purpose behind it. That objective and purpose can only be achieved if a report has the following qualities and characteristics:

1. It should be factual: Every report should be based on facts, verified information and valid proofs.
2. Clear and Easily understandable: Explained below
3. Free from errors and duplication
4. Should facilitate the decision makers in making the right decision:
5. Result focused and result oriented
6. Well organized and structured
7. Ethical reporting style

Reader-Friendly

Readers are various stakeholders who receive reports generated by M&E. If reports are reader-friendly, they are likely to be read, remembered and acted upon. Following decisions need to be made by CSOs to make their reports reader-friendly:

- What do they need to know?
- When do they need to know?
- How do they like to know?

Easy, Simple Language

M&E reports are meant to inform not impress. Using easy, simple language, be it Urdu or English makes the report friendly on reader. To do this, here are some useful tips:

- Write only what is necessary
- Avoid repetition and redundancy
- Give interesting and relevant information
- Avoid preaching or lecturing
- Compose short and correct sentences

Purposeful Presentation

Each report has some objective(s) to meet. The "objective" comes from analyzing the needs of the reader. A CSO is working for a project that has several donors, and is channeled through an agency that needs to be informed about some specific things going on in the field. CSOs reports are the main pathways or channels of information to the people who decide to fund this and other such projects. Similarly, field reports are the admin vehicles for the management of the CSOs to make decision regarding the project itself. A good report presents facts and arguments in a manner that supports the purpose of the report.

Organized and Well-Structured

Each CSO comes up with a format of internal reporting to suit its requirements. Reporting to donors is done on their prescribed formats. The M&E system should be able to generate information that can be organized using different formats. In the annex, this manual provides some useful formats that can be customized by a CSO.

Result-Focused

In general, all readers are interested in the RESULTS. Therefore, one over-riding principle that CSOs should aim for in all report writing is to report on the results of their activities. This requires some analysis on their part that goes beyond a mere description of their activities. Result-focused means that description of activities is liked with the project objectives. This aspect must be addressed especially in the project progress reports. According to Phil Bartle, "A good progress report is not merely a descriptive activity report, but must analyze the results of those reported activities. The analysis should answer the question, "How far have the project objectives been reached?"

Timely Prepared and Dispatched

M&E generate "Information Products", a customized set of information according to needs to a defined group of users. M&E's information products are time-bound for both internal and external stakeholders. Reports, in suitable formats, need to be timely produced and made available to the readers. It is useful to develop an Information Product Matrix (IPM) like the one described below:

Straightforward

A good report is straight forward, honest description. It contains no lies, no deception, no fluff. It

is neat, readable and to-the-point. It is well spaced, has titles and subtitles and is free of language errors.

Features of Power BI Layout

Layouts Are Custom

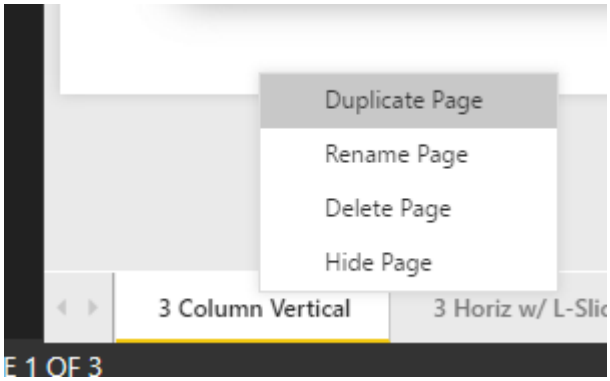
Each layout is a totally custom setup. Each visual has been modified in a way to make the presentation of the data look the best. This means in some cases the visuals fonts, labels, label precision, x or y axis has been turned off or modified in some way.

Since the layout has been highly stylized, you may not recognize the visual or feel like something is missing, for example the y-axis. This was intentional and done to tell the data story.

You can modify all the properties by clicking on the visual and selecting the **Paint Roller** to modify the properties.

Duplicating a Page

Once you have found a layout that you like, you might want to reuse that page layout multiple times. To reuse a layout, simply **Right Click** on the report layout tab that you like and select **Duplicate Page**.



Duplicate a Page

Delete Unwanted Page

Most layout files will contain multiple report pages. Some layout documents have multiple layouts per PBIT file. Thus, there will be situations where you want to delete un-needed pages. This is a simple **Right Click** on the report tab that you wish to remove and selecting **Delete Page**.

The Layout files are created with many different report tabs, so you can have incredible flexibility. For example, you may want a main page with two (2) visuals, and slicers, but later in the report you need a three (3) column layout. By including many different designs within one PBIT Layout file you can choose what works best with your data.

Visuals Layout in Selection Pane

In the selection window there will be multiple items which represent each item on the report page. Each item in the **Selection** pane will attempt to be laid out with applicable names and in

order of left to right typically. There will be some exceptions to this rule, but items will be laid out as intuitive as possible.

*Note: If you don't see the selection window you can turn this on by clicking on the **View** ribbon, and then checking the box titled **Selection Pane**.*

Report Pages Background

Each Layout Report page will come with a background image. If it is desired this background image can be removed. To remove the background image, make sure no visuals are selected. This can be achieved by pressing the **ESC** key. Then in the **Visualizations** pane expand the **Page Background**. To remove the background image, click the little **X** next to the image name, found under the Transparency setting. By default, the **Image Fit** is set to **Fit**.

Caution: If a background image is removed, it can't be replaced. So, once it is gone, you won't be able to re-apply a background image.

Changing Visuals

Now, not all layouts will be able to meet your report needs. Maybe you want to add a custom visual, or maybe change a visual from a map to a bar chart. Even though this is a predefined layout you can easily change the visual. To change the visual first select the visual you want to change, and then on the **Visualizations** pane click a new visual type. It is that easy.

Moving Visuals

All the visuals on each report page are locked by default and therefore you cannot move them with the mouse. This is a setting that has been turned on. The setting is called **Lock Objects** and it can be changed by navigating to the **View Ribbon** and then **Unchecking** the **Lock Objects** item. By unchecking this feature, you will be able to move the visuals around the page.

Adding Custom Report Themes

he layouts have been made using the default colors for Power BI Desktop. However, once you have connected your data you can easily add your own color theme. If you need help making color themes using JSON you can use this **free Theme Generator** to help add custom colors to your report. For the full specification on the Reports Themes visit the **official Microsoft documentation page**..

ou may have seen various people use the terms "reporting" and "analysis" as though they were interchangeable terms or almost synonyms. While both of these areas of web analytics draw upon the same collected web data, reporting and analysis are very different in terms of their **purpose**, **tasks**, **outputs**, **delivery**, and **value**. Without a clear distinction of the differences, an organization may sell itself short in one area (typically analysis) and not achieve the full benefits of its web analytics investment. Although I'm primarily focusing on web analytics, companies can run into the same challenge with other analytics tools (e.g., ad serving, email, search, social, etc.).

Most companies have analytics solutions in place to derive greater value for their organizations. In other words, the ultimate goal for reporting and analysis is to increase sales and reduce costs (i.e., add value). Both reporting and analysis play roles in influencing and driving the actions which lead to greater value in organizations.

For the purposes of this blog post, I'm not going delve deeply into what happens before or after the reporting and analysis stages, but I do recognize that both areas are critical and challenging steps in the overall data-driven decision-making process. It's important to remember that reporting and analysis only have the opportunity of being valuable **if they are acted upon**.

Purpose

Before covering the differing roles of reporting and analysis, let's start with some high-level definitions of these two key areas of analytics.****

Reporting: *The process of organizing data into informational summaries in order to monitor how different areas of a business are performing.*****

Analysis: *The process of exploring data and reports in order to extract meaningful insights, which can be used to better understand and improve business performance.*

__Reporting translates raw data into **information**. Analysis transforms data and information into **insights**. Reporting helps companies to monitor their online business and be alerted to when data falls outside of expected ranges. Good reporting should **raise questions** about the business from its end users. The goal of analysis is to **answer questions** by interpreting the data at a deeper level and providing actionable recommendations. Through the process of performing analysis you may raise additional questions, but the goal is to identify answers, or at least potential answers that can be tested. In summary, reporting shows you ***what is happening*** while analysis focuses on explaining ***why it is happening*** and ***what you can do about it.***

Tasks

Companies can sometimes confuse "analytics" with "analysis". For example, a firm may be focused on the general area of analytics (strategy, implementation, reporting, etc.) but not necessarily on the specific aspect of analysis. It's almost like some organizations run out of gas after the initial set-up-related activities and don't make it to the analysis stage. In addition, sometimes the lines between reporting and analysis can blur – what feels like analysis is really just another flavor of reporting.

One way to distinguish whether your organization is emphasizing reporting or analysis is by identifying the primary tasks that are being performed by your analytics team. If most of the team's time is spent on activities such as *building, configuring, consolidating, organizing, formatting,* and *summarizing* – that's reporting. Analysis focuses on different tasks such as *questioning, examining, interpreting, comparing*, and *confirming* (I've left out testing as I view optimization efforts as part of the action stage). Reporting and analysis tasks can be intertwined, but your analytics team should still evaluate where it is spending the majority of its time. In most cases, I've seen analytics teams spending most of their time on reporting tasks.

Outputs

When you look at reporting and analysis deliverables, on the surface they may look similar – lots of charts, graphs, trend lines, tables, stats, etc. However, there are some subtle differences. One of the main differences between reporting and analysis is the overall approach. Reporting follows a **push approach**, where reports are pushed to users who are then expected to extract meaningful insights and take appropriate actions for themselves (i.e., self-serve). I've identified three main types of reporting: *canned reports*, *dashboards*, and *alerts*.

1. **Canned reports:** These are the out-of-the-box and custom reports that you can access within the analytics tool or which can also be delivered on a recurring basis to a group of end users. Canned reports are fairly static with fixed metrics and dimensions. In general, some canned reports are more valuable than others, and a report's value may depend on how relevant it is to an individual's role (e.g., SEO specialist vs. web producer).
2. **Dashboards:** These custom-made reports combine different KPIs and reports to provide a comprehensive, high-level view of business performance for specific audiences. Dashboards may include data from various data sources and are also usually fairly static.
3. **Alerts:** These conditional reports are triggered when data falls outside of expected ranges or some other pre-defined criteria is met. Once people are notified of what happened, they can take appropriate action as necessary.

In contrast, analysis follows a **pull approach**, where particular data is pulled by an analyst in order to answer specific business questions. A basic, informal analysis can occur whenever someone simply performs some kind of mental assessment of a report and makes a decision to act or not act based on the data. In the case of analysis with actual deliverables, there are two main types: *ad hoc responses* and *analysis presentations*.

1. **Ad hoc responses:** Analysts receive requests to answer a variety of business questions, which may be spurred by questions raised by the reporting. Typically, these urgent requests are time sensitive and demand a quick turnaround. The analytics team may have to juggle multiple requests at the same time. As a result, the analyses cannot go as deep or wide as the analysts may like, and the deliverable is a short and concise report, which may or may not include any specific recommendations.
2. **Analysis presentations:** Some business questions are more complex in nature and require more time to perform a comprehensive, deep-dive analysis. These analysis projects result in a more formal deliverable, which includes two key sections: *key findings* and *recommendations*. The key findings section highlights the most meaningful and actionable insights gleaned from the analyses performed. The recommendations section provides guidance on what actions to take based on the analysis findings.

When you compare the two sets of reporting and analysis deliverables, the different purposes (information vs. insights) reveal the true colors of the outputs. Reporting pushes information to the organization, and analysis pulls insights from the reports and data. There may be other hybrid outputs such as annotated dashboards (analysis sprinkles on a reporting donut), which may appear to span the two areas. You should be able to determine whether a deliverable is primarily focused on reporting or analysis by its purpose (information/insights) and approach (push/pull).

Another key difference between reporting and analysis is **context**. Reporting provides **no or**

limited context about what's happening in the data. In some cases, the end users already possess the necessary context to understand and interpret the data correctly. However, in other situations, the audience may not have the required background knowledge. Context is critical to good analysis. In order to tell a meaningful story with the data to drive specific actions, context becomes an essential component of the storyline.

Although they both leverage various forms of data visualization in their deliverables, analysis is different from reporting because it **emphasizes data points that are significant, unique, or special – and explain why they are important to the business**. Reporting may sometimes automatically highlight key changes in the data, but it's not going explain why these changes are (or aren't) important. Reporting isn't going to answer the "so what?" question on its own.

If you've ever had the pleasure of being a new parent, I would compare canned reporting, dashboards, and alerts to a six-month-old infant. It cries – often loudly – when something is wrong, but it can't tell you what is exactly wrong. The parent has to scramble to figure out what's going on (hungry, dirty diaper, no pacifier, teething, tired, ear infection, new Baby Einstein DVD, etc.). Continuing the parenting metaphor, reporting is also not going to tell you how to stop the crying.

The **recommendations** component is a key differentiator between analysis and reporting as it provides specific guidance on what actions to take based on the key insights found in the data. Even analysis outputs such as ad hoc responses may not drive action if they fail to include recommendations. Once a recommendation has been made, **follow-up** is another potent outcome of analysis because recommendations demand decisions to be made (go/no go/explore further). Decisions precede action. Action precedes value.

Delivery

As mentioned, reporting is more of a push model, where people can access reports through an analytics tool, Excel spreadsheet, widget, or have them scheduled for delivery into their mailbox, mobile device, FTP site, etc. Because of the demands of having to provide periodic reports (daily, weekly, monthly, etc.) to multiple individuals and groups, **automation** becomes a key focus in building and delivering reports. In other words, once the report is built, how can it be automated for regular delivery? Most of the analysts who I've talked to don't like manually building and refreshing reports on a regular basis. It's a job for robots or computers, not human beings who are still paying off their student loans for 4-6 years of higher education.

On the other hand, analysis is all about human beings using their superior reasoning and analytical skills to extract key insights from the data and form actionable recommendations for their organizations. Although analysis can be "submitted" to decision makers, it is more effectively presented **person-to-person**. In their book "Competing on Analytics", Thomas Davenport and Jeanne Harris emphasize the importance of **trust and credibility between the analyst and decision maker**. Decision makers typically don't have the time or ability to perform analyses themselves. With a "close, trusting relationship" in place, the executives will frame their needs correctly, the analysts will ask the right questions, and the executives will be more likely to take action on analysis they trust.

Value

When it comes to comparing the different roles of reporting and analysis, it's important to understand the relationship between reporting and analysis in driving value. I like to think of the data-driven stages (data > reporting > analysis > decision > action > value) as a series of dominoes. If you remove a domino, it can be more difficult or impossible to achieve the desired value.

In the "Path to Value" diagram above, it all starts with having the **right data that is complete and accurate**. It doesn't matter how advanced your reporting or analysis is if you don't have good, reliable data. If we skip the "reporting" domino, some seasoned analysts might argue that they don't need reports to do analysis (i.e., just give me the raw files and a database). On an individual basis that might be true for some people, but it doesn't work at the organizational level if you're striving to democratize your data.

Most companies have abundant reporting but may be missing the "analysis" domino. Reporting will rarely initiate action on its own as analysis is required to **help bridge the gap between data and action**. Having analysis doesn't guarantee that good decisions will be made, that people will actually act on the recommendations, that the business will take the right actions, or that teams will be able to execute effectively on those right actions. However, it is a necessary step closer to action and the potential value that can be realized through successful web analytics.

Core Tools for Reporting and Dashboard layout Design:

the most commonly cited among the websites that ranked the best BI tools. These five BI tools are not ranked in any particular order.

1. Microsoft Power BI

One of the most popular BI tools is Power BI, offered by leading software giant Microsoft. This tool is downloadable software, so you can choose to run analytics either on the cloud or in a reporting server. Syncing with sources such as Facebook, Oracle, and more, generate reports and dashboards in minutes with this interactive tool. It comes with built-in AI capabilities, Excel integration, and data connectors, and offers end-to-end data encryption and real-time access monitoring.

Learn Power BI with Coursera: In just two hours, you can learn the basics of Power BI Desktop with this guided project. You'll load and transform data to create interactive reports and dashboards.

2. Tableau

Tableau is known for its user-friendly data visualization capabilities, but it can do more than make pretty charts. Their offering includes live visual analytics, an interface that allows users to drag and drop buttons to spot trends in data quickly. The tool supports data sources such as Microsoft Excel, Box, PDF files, Google Analytics, and more. Its versatility extends to being able to connect with most databases.

Learn Tableau with Coursera: There are several options for learning how to use Tableau.

- Data Visualization with Tableau specialization from the University of California Davis
- Use Tableau for Your Data Science Workflow specialization from the University of California Irvine
- Data Visualization and Communication class with Tableau from Duke University
- Guided Project on Visualizing Citibike Trips with Tableau

3. QlikSense

QlikSense is a BI tool that emphasizes a self-service approach, meaning that it supports a wide range of analytics use cases, from guided apps and dashboards to custom and embedded analytics.

It offers a user-friendly interface optimized for touchscreens, sophisticated AI, and high-performance cloud platforms. Its associative exploration capability, Search & Conversational Analytics, allows users to ask questions and uncover actionable insights, which helps increase data literacy for those new to using BI tools.

4. Dundas BI

Dundas BI is a browser-based BI tool that's been around for 25 years. Like Tableau, Dundas BI features a drag-and-drop function that allows users to analyze data on their own, without involving their IT team. The tool is known for its simplicity and flexibility through interactive dashboards, reports, and visual analytics. Since its inception as a data visualization tool in 1992, it has evolved into an end-to-end analytics platform that is able to compete with the new BI tools available today.

5. Sisense

Sisense is a user-friendly BI tool that focuses on being simplified and streamlined. With this tool, you can export data from sources like Google Analytics, Salesforce, and more. Its in-chip technology allows for faster data processing compared to other tools. Key features include the ability to embed white-label analytics, meaning a company can fully customize the services to its needs. Like others, it has a drag-and-drop feature. Sisense allows you to share reports and dashboards with your team members as well as externally.

Other popular BI tools include: Zoho Analytics, Oracle BI, SAS Visual Analytics, Domo, Datapine, Yellowfin BI, Looker, SAP Business Objects, Clear Analytics, Board, MicroStrategy, IBM Cognos Analytics, Tibco Spotfire, BIRT, Intercom, Google Data Studio, and HubSpot.