

# Business Intelligence

by  
Sachin Gupta

# **BI (Business Intelligence)**

- The BI term refers to **Business Intelligence**.
- It is a data-driven decision support system (DSS),
- Helps
  - analyze the data
  - provide actionable information.
- It helps the business manager, corporate executives, and other users in making their decisions easily
- Business intelligence refers to the applications, technologies, and practices for the collection, analysis, integration, and presents the business information.
- The purpose of business intelligence is to support better decision making.

# BI Basics

## Business Intelligence

-  Improves decision making
-  Fast answers to any business query
-  Enables real-time analysis with quick navigation
-  Identify cross-Selling & Up-selling opportunities
-  Reduces the risk of bottlenecks
-  Helps you know your business better



# Importance of BI

- Business intelligence is used to improve all parts of a company by improving access to the firm's data and then using that data to increase profitability.
- Companies that practices BI can translate their collected data into **insights** their business processors.
- Then the insights can be used to create strategic business decisions that improve productivity and accelerate the growth.

# Benefits

- Driving new revenues.
- It increases operational efficiency.
- It optimizes internal business processes.
- It improves decision making.
- It is gaining a competitive advantage over business rivals.
- It is used in spotting business problems that need to be addressed.
- It can be used in assisting companies in the identification of market trends.

# Functionalities

- **Analysis:** Business Intelligence solution analyze the raw data with seasonal, trend analysis , what if scenarios, modeling using different data modeling techniques and similar functionalities.
- **Reports:** The solution provides very comprehensive functionalities for creating reports for better understanding of the business data. Business users can create, view, modify report and visualizations online and offline and with other office products.
- **Real time monitoring:** Business Intelligence solution provides tools to analyze the real time up to the second operational data to enable the business to take quick and informed decisions.
- **Dashboards:** The solution provides dashboards which helps the business to monitor, measure and manage business performance in a very quick manageable way.
- **Scorecards:** Business Intelligence solutions provides industry standard scorecards to enable the business to measure the key performance indicators.



# Functionalities

- **Collaborative Business Intelligence:** Business Intelligence Solution provides capabilities to collaboratively share the information with different stakeholders within the organization and outside.
- **Mobile Business Intelligence:** Business Intelligence solution provides functionalities to make the same information such as reports, dashboards, monitoring available in mobile device.
- **Advanced Analytics and visualization:** Some of the business intelligence solutions provide advanced analytics and visualisation capabilities enabling real time visualisation in different forms and methods and to augment additional functionalities.

# How business intelligence works



- Businesses and organizations have questions and goals. To answer these questions and track performance against these goals, they gather the necessary data, analyze it, and determine which actions to take to reach their goals.
- On the technical side, raw data is collected from business systems. Data is processed and then stored in data warehouses, the cloud, applications, and files. Once it's stored, users can access the data, starting the analysis process to answer business questions.
- BI platforms also offer data visualization tools, which convert data into charts or graphs, as well as presenting to any key stakeholders or decision-makers.

# BI Methods

- **Data mining:** Using databases, statistics, and **machine learning (ML)** to uncover trends in large datasets
- **Reporting:** Sharing data analysis to stakeholders so they can draw conclusions and make decisions
- **Performance metrics and benchmarking:** Comparing current performance data to historical data to track performance against goals, typically using customized dashboards
- **Descriptive analytics:** Using preliminary data analysis to find out what happened
- **Querying:** Asking the data-specific questions, BI pulling the answers from the data sets

- **Statistical analysis:** Taking the results from descriptive analytics and further exploring the data using statistics such as how this trend happened and why
- **Data visualization:** Turning data analysis into visual representations such as charts, graphs, and histograms to more easily consume data
- **Visual analysis:** Exploring data through visual storytelling to communicate insights on the fly and stay in the flow of analysis
- **Data preparation:** Compiling multiple data sources, identifying the dimensions and measurements, and preparing it for data analysis

## How BI, data analytics, and business analytics work together ?

- Business intelligence includes data analytics and **business analytics** but uses them only as parts of the whole process.
- BI helps users draw conclusions from data analysis.
- Data scientists dig into the specifics of data, using advanced statistics and **predictive analytics** to discover patterns and forecast future patterns.
- BI is designed to answer specific queries and provide at-a-glance analysis for decisions or planning. However, companies can use the processes of analytics to continually improve follow-up questions and iteration. B



Analysis complete ⓘ

Refresh

#### Dimensions

Abc Category

⊕ City

⊕ Country

Abc Customer Name

⊕ Manufacturer

⊕ Order Date

Abc Order ID

⊕ Postal Code

Abc Product Name

.III. Profit (bin)

Abc Region

Abc Segment

⊕ Ship Date

Abc Ship Mode

⊕ State

Abc Sub-Category

#### Measures

# Discount

=# Number of Records

# Profit

=# Profit Ratio

# Quantity

# Sales



Ask about fields in this data source

Or try one of these suggestions:

**Discount** at least 0%

by **Category**, top **Category** by most expensive **Profit**

sum of **Discount**

sum of **Discount** by **Order Date**'s year as a line chart

by **Category**

sort **Category** in alphabetical order

top **Category** by sum of **Number of Records**

# How to develop a business intelligence strategy

Here's how to create a BI strategy from the ground up:

1. Know your business strategy and goals.
2. Identify key stakeholders.
3. Choose a sponsor from your key stakeholders.
4. Choose your BI platform and tools.
5. Create a BI team.
6. Define your scope.
7. Prepare your data infrastructure.
8. Define your goals and roadmap.

# Categories of BI analysis

There are three major types of BI analysis, which cover many different needs and uses. These are predictive analytics, descriptive analytics, and prescriptive analytics.

- **Predictive analytics** takes historical and real-time data and models future outcomes for planning purposes.
- **Descriptive analytics** is the process of identifying trends and relationships in data using historical and current data.
- **prescriptive analytics** takes all the relevant data to answer the question, “what should my business do?”

## How to choose a BI platform

- Intuitive to use
- Variety of dashboard and visualization options
- Smart insights
- Alerts for good and bad metrics
- Built-in artificial intelligence (AI)
- Deployment flexibility
- Integration with other platforms and applications
- Data connectivity
- Embedding in business applications

## BI dashboards

The key features of a BI dashboard are:

- Interactivity
- Real-time data
- Customizable interface
- Standard templates
- Sharing ability

# BI and big data

- As the data atmosphere grows and the gathering, storage, and analysis of data becomes more complex, it's important to consider the relationship between BI and **big data**.
- Big data has become a bit of a buzzword in the industry lately, so what exactly is it? Well, data experts define it by “the four Vs”: Volume, velocity, value, and variety.
- These four define big data and set it apart. In particular, volume is what people normally point to as the main defining factor, as the amount of data is ever-increasing and relatively easy to store for long periods of time.
- As you can imagine, this is important for BI as businesses create more and more data by the year, and BI platforms have to keep up with the increasing demands made on them. A good platform will grow with increasing demands. But if not maintained, dashboards and data sources may fall behind as big data evolves.

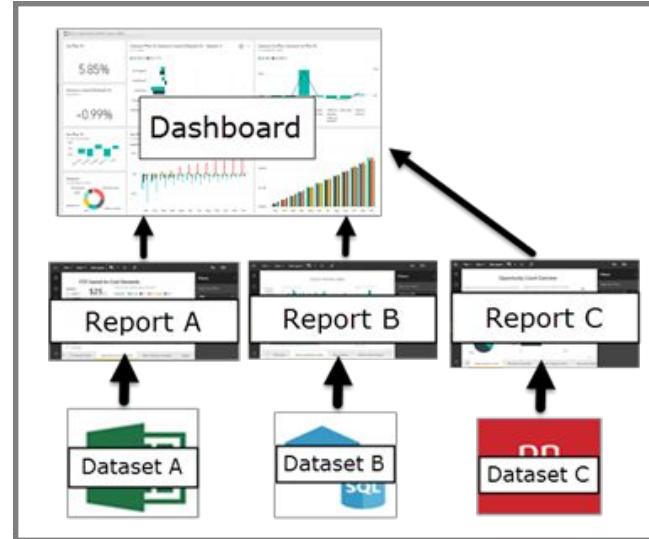
# The future role of business intelligence

- Business intelligence is continually evolving to keep pace with business needs and technology, so each year, we identify current trends to keep users up-to-date on innovations.
- Realize that artificial intelligence and machine learning will continue to grow, and businesses can integrate the insights from AI into a broader BI strategy
- As companies strive to be more data-driven, efforts to share data and collaborate will increase.
- Data visualization will be even more essential to work together across teams and departments.
- BI offers capabilities for near real-time sales tracking, allowing users to discover insights into customer behavior, forecast profits, and more.
- Diverse industries like retail, insurance, and oil have adopted BI, and more are joining each year. BI platforms adapt to new technology and the innovation of its users.

# Dashboard basics

The visualizations you see on the dashboard are called *tiles*. You *pin* tiles to a dashboard from reports. If you're new to Power BI

The visualizations on a dashboard originate from reports and each report is based on a dataset. One way to think of a dashboard is as an entryway to the underlying reports and datasets. Selecting a visualization takes you to the report (and dataset) that it's based on.



## **Advantages of dashboards**

- Dashboards are a wonderful way to monitor your business and see all of your most important metrics at a glance. The visualizations on a dashboard can come from one underlying dataset or many, and from one underlying report or many. A dashboard combines on-premises and cloud data, providing a consolidated view regardless of where the data lives.
- A dashboard isn't just a pretty picture. It's highly interactive and the tiles update as the underlying data changes.

# Dashboards versus reports

| Capability                                   | Dashboards  | Reports  |
|--|---|--|
| Pages  | One page  | One or more pages  |
| Data sources                                 | One or more reports and one or more datasets per dashboard.   | A single dataset per report  |
| Drilling down in visuals                     | Only if you pin an entire report page to a dashboard.   | Yes  |
| Available in Power BI Desktop                | No  | Yes. Can build and view reports in Power BI Desktop.                           |
| Filtering                                    | No. Can't filter or slice a dashboard. <i>Can</i> filter a dashboard tile in focus mode, but can't save the filter. | Yes. Many different ways to filter, highlight, and slice.                      |
| Feature content on colleagues' Home page     | Yes   | Yes  |
| Favorite                                     | Yes. Can set multiple dashboards as <i>favorites</i> .  | Yes. Can set multiple reports as <i>favorites</i> .                            |
| Natural language queries (Q&A)               | Yes   | Yes, provided you have edit permissions for the report and underlying dataset. |
| Set alerts                                   | Yes. Available for dashboard tiles in certain circumstances.  | No   |
| Subscribe                                    | Yes. Can subscribe to a dashboard.  | Yes. Can subscribe to a report page.   |
| Can see underlying dataset tables and fields | No. Can export data but can't see tables and fields in the dashboard itself.  | Yes  |

# Business Intelligence Requirements Gathering: What to do Before Building Dashboards



Too often people equate BI with reports or dashboards but those are the end results. There's a lot more at play before one reaches that stage (read: final stage) of business intelligence.

### **Why is business intelligence requirements gathering important?**

Each department of a business has different goals and objectives they wish to achieve with the help of business intelligence and analytics and therefore their requirements are different too. A well-documented BI Solution Requirements Report will result in the following:

- Address business, data, technical, and usage needs
- Uncover current and future needs
- Save time by providing the team with a constant point of reference during the BI software solution selection phase as well as in the implementation phase

# Business Assessments

## 1. Business Assessment

- A business assessment ensures your business is well equipped with what it needs to deliver on your goals. It provides a guideline to help you scale your business in a way that is smart and strategic. The secret to conducting useful business assessments that provide actionable insights is to analyze appropriate Key Performance Indicators (KPIs).
- KPIs keep business objectives at the forefront. They are the metrics that provide answers to the right questions or monitor the health of the areas of your business in consideration. With the right selection of KPIs you evaluate how a BI solution needs to address pain points experienced in the past, how should it deliver on the current needs and finally how will it scale to cover future wants.
- Since KPIs also play a pivotal role in keeping your business intelligence projects on track, we have created a [KPIs for BI Template](#) that will assist you in the exercise of choosing the right ones.

# Data Assessment

- The truthfulness of the analysis is directly proportional to the integrity of the data.
- A data assessment report can identify data quantity, complexity and active versus inactive data – critical for project scoping.
- It includes various quality checks which provide information on data relevancy, business data readiness, target readiness and gap analysis.
- Data Assessment must also confirm that the data is applicable to the KPIs.
- You may have the right KPIs defined in the Business Assessment Report but if there's a mismatch between your data and KPIs, you need to relook at your data.
- Finally, a thorough data assessment will also reduce unnecessary infrastructure costs and the complexity of your new system - ensuring better operational efficiencies.

# User Group Assessment

- Business intelligence is a process that sees many users participate at various stages starting with data transformation performed by technical users and ending with reports or dashboards consumed by business users.
- Each group has specific requirements from the BI solution.
- Profiling them helps in understanding their consumption and usage patterns which comes in handy while selecting and designing dashboards

# Security Assessment

- When there are multiple users and datasets involved in the project, it's crucial to know in advance
- how the interactivity will take place,
- how will users be authenticated,
- what is the back-end architecture, and where is the solution going to be integrated. Finally, if there are any security checkpoints already in place that needs to be considered, or if any extra security checks need to be put in place.

# Decision Making BI Concepts

## THE 2 SYSTEMS



### System 1 (Fast Thinking)

Continuously scans our environment.



Fast but error-prone



Works automatically & effortlessly via shortcuts, impulses and intuition.



### System 2 (Slow Thinking)

Used for specific problems, only if necessary



Takes effort to analyze, reason, solve complex problems, exercise self-control



Slow but reliable



## 1. Slow But Accurate

- You have a slow thinking system which you can think of as a spotlight of attention. You use this system when you really focus on and think about a problem step by step. This slow thinking system is very accurate. However, it can only process so much information at any one time and requires a lot of energy.
- Many of us don't use our slow thinking system for most decisions. Instead, we use our fast system, otherwise known as our intuition.

## 2. Fast But Error-Prone

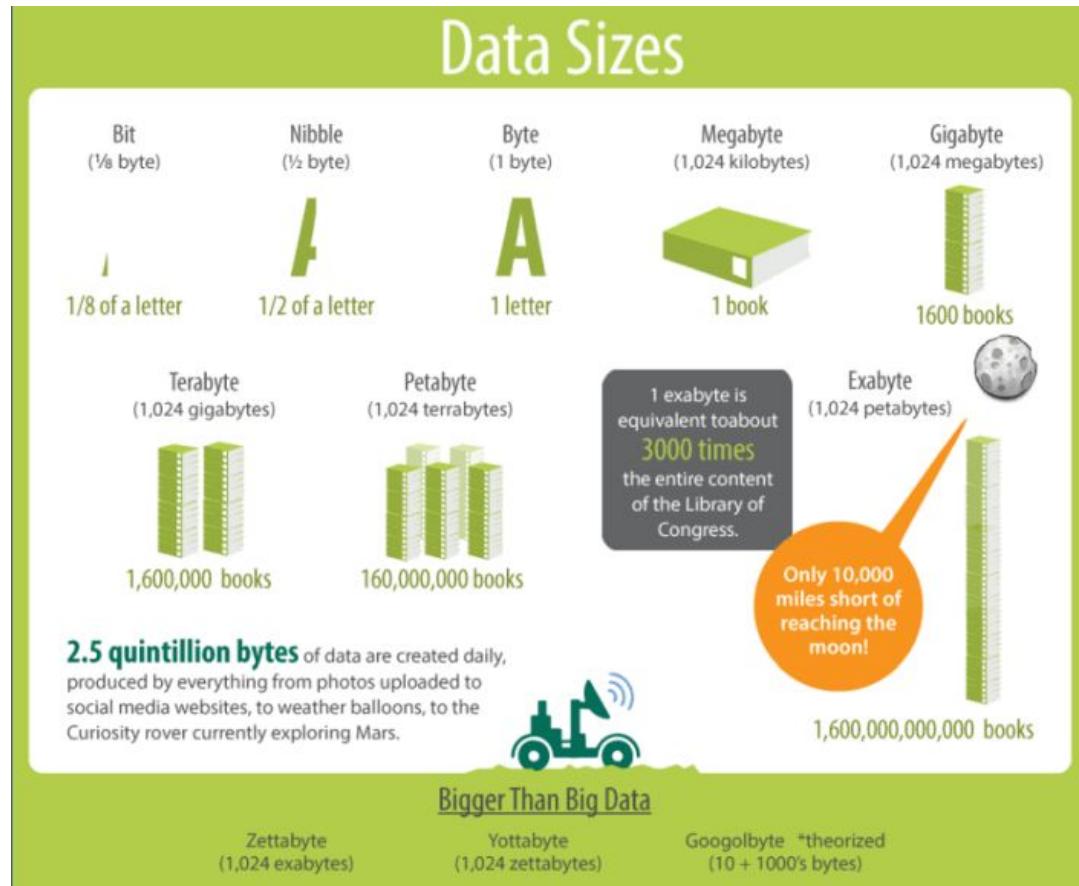
- Your fast thinking system can take in massive amounts of data at once. It can also make snap decisions very quickly, with pretty high accuracy. However, your fast thinking system has one big disadvantage. It is prone to [logical fallacies](#) and perceptual biases.
- If your slow thinking system is like a spotlight, your fast thinking system is like the area of a picture that isn't in focus. You don't have to burn much energy to use it, but you don't see things as clearly as you do with your slow thinking system.

Figure 1: A Comparison of System 1 and System 2 Thinking

| System 1<br>"Fast"  | System 2<br>"Slow"  |
|---|---|
| <b>DEFINING CHARACTERISTICS</b><br>Unconscious<br>Effortless<br>Automatic | <b>DEFINING CHARACTERISTICS</b><br>Deliberate and conscious<br>Effortful<br>Controlled mental process |
| WITHOUT self-awareness or control<br><br>"What you see is all there is."  | WITH self-awareness or control<br><br>Logical and skeptical   |
| <b>ROLE</b><br>Assesses the situation<br>Delivers updates                 | <b>ROLE</b><br>Seeks new/missing information<br>Makes decisions                                       |

# Why Are Business Intelligence Concepts And BI Solutions So Important Today?

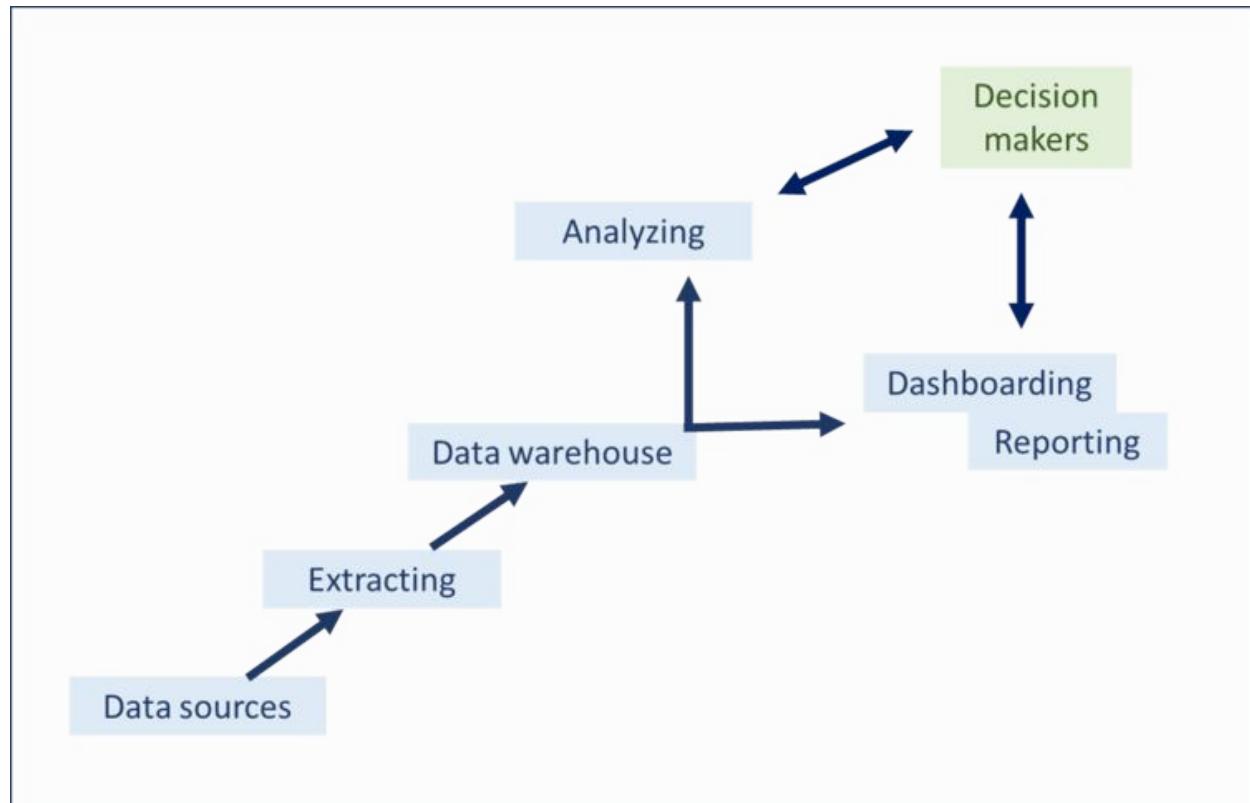
The fact is, without business intelligence, you risk the very real possibility of drowning in data. Just look at these numbers: according to [CloudTweaks](#), in 2015 there were 2.5 quintillion bytes of data produced daily. Do you know how much a quintillion is? It's a 2.5 followed by **18 zeros**. We have already entered the Zettabyte era, also mentioned as one of our [tech buzzwords](#) for 2019, and, for scale, in 2012, the entire Internet only contained  $\frac{1}{2}$  of one zettabyte in data.



1. They prevent you from drowning in data
2. They provide a wealth of insights
3. Benchmarking is more accurate
  - BI-based key performance indicators (KPIs),
4. You can predict your business future
5. They enable powerful data visualization

# The Concepts Of Business Intelligence?

- The data itself (raw data)
- The data warehouse
- Data access, analytics, and presentation
- Data dashboarding and reporting



# The raw data

- The first component of an integral BI solution is the data itself. This data could be anything like sales records for the year, the keywords implemented in your latest advertising program, salary and benefit tables, or profit and loss statements.
- A company's data is typically stored across a host of databases, depending on how each specific data set is collected (through CRMs, ERPs, flat files, APIs, etc.).
- As a result of this fragmentation, today's BI solutions are developed with various **data connectors** that let users consolidate all of their databases into one centralized data warehouse, allowing them to work on each insight conjointly and enhancing cross-database analysis.

# The Data Warehouse

- **Data storage warehouse** is the logistics platform that connects all of your different databases together and allows you to create relationships between them. This is an area that has seen great advances recently with the introduction of **cloud-based BI tools**.
- The legacy approach to a data warehouse was often a mishmash of different Excel sheets, old mainframe style databases that had to be accessed by technicians, paper-based records, and proprietary program databases.
- After realizing how difficult it was to make use of all of these scattered data sources, people began to integrate databases through the use of warehouses and systems. Modern systems are also superior to the legacy systems in that they often update in real time, as opposed to having to be manually updated – a process which often required the IT department.

# Data access, analytics, and presentation

- Once all of your data is connected and can ‘talk to each other,’ one of the next key business intelligence concepts is to make use of that data. This involves accessing the data, analyzing it for important trends, and presenting it in a way that is immediately understandable.
- These steps can often blend together, especially if you use [interactive dashboards](#) that let you zoom in and out of your data according to your business need. Data presentation has also come a long way since the Excel days. Now there are beautiful, intuitive [dashboard examples](#) that can give you the information you need at a glance.

# Data dashboarding and reporting

- Building on our previous point, the fourth and perhaps most pivotal component of an interactive dashboard is the ability to continuously track, monitor, and report your data.
- By having access to a flexible, customizable, data-driven **online dashboard**, you can set targets, identify patterns, spot trends, and uncover insights that foster growth and improvement. Through initiative functionality and seamless data visualization, it's also possible to share your discoveries with others within the organization in a way that's inclusive as well as digestible.
- Moreover, today's BI-based **dashboard reporting** is portable, meaning that it's possible to log in, analyze data, and share information wherever you may be in the world, 24/7, on a multitude of devices.
- It's clear that these 4 concepts of business intelligence are a recipe for data-driven success. Now let's look at how to apply them in a real-world context.

# Visual Analytics

- Visual analytics is the use of sophisticated tools and processes to analyze datasets using visual representations of the data.
- Visualizing the data in graphs, charts, and maps helps users identify patterns and thereby develop actionable insights.
- These insights help organizations make better, data-driven decisions.

# Benefits

- Sometimes confused with **data visualization**, visual analytics isn't simply a matter of representing data graphically. Modern, interactive visual analytics makes it easy to combine data from multiple sources and deeply analyze the data directly within the visualization itself. Plus, AI and machine learning algorithms can offer recommendations to help guide your exploration.
- Ultimately, visual analytics helps you turn massive data sets into business insights which can have a major positive impact on your organization.



Hospital Readmissions

Analyze  
SheetNarrate  
Storytelling

Hospital Readmissio...



No selections applied



Selections

## Hospital Readmission Reduction Program Dashboard



Total CMS 30 Day Readmission Penalties: \$936,453

Hospital

AMI Penalty

\$126,458

CHF Penalty

\$304,133

COPD Penalty

\$185,865

HAC Penalty

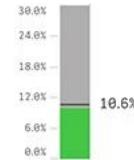
\$143,821

PNA Penalty

\$176,176

## Readmission Rates

All Causes



Readmission rate (10% target)

AMI



Readmission rate (10% target)

CHF



Readmission rate (20% target)

COPD



Readmission rate (15% target)

HAC



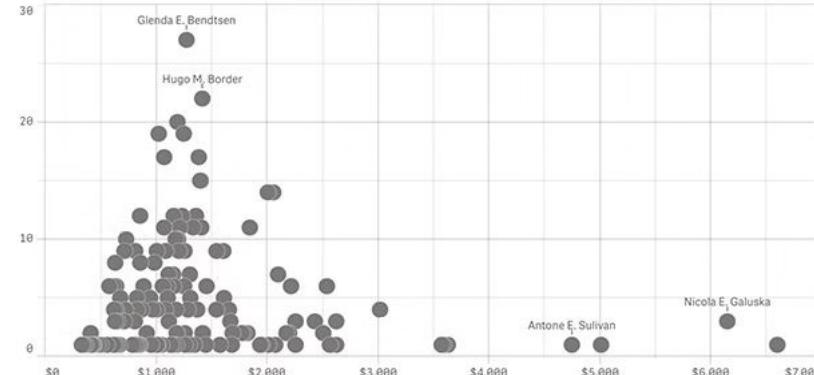
Readmission rate (10% target)

PNA



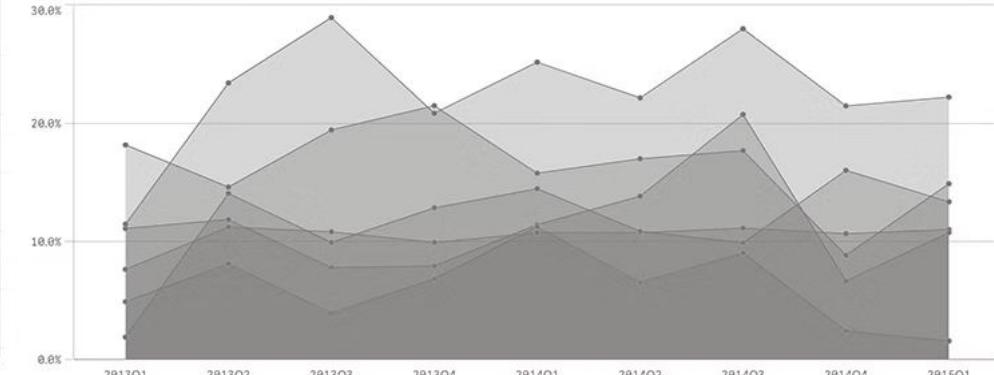
Readmission rate (10% target)

## Readmissions vs Average HRRP Penalties by Provider



## HRRP Readmission Rates Over Time

HRRP Conditions vs All Causes



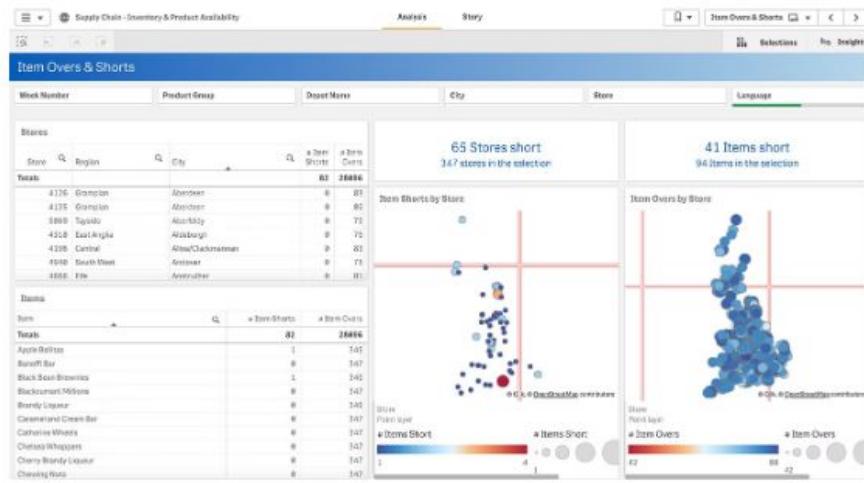
1. **Share findings and track progress:** Interactive reports and dashboards help users track, organize and share **key performance indicators** across an organization.
2. **Make faster decisions:** Users can understand data insights much more quickly by seeing and working with data sets when they are in a visual format.
3. **Explore data more easily:** **Self-service analytics** tools which allow users to interact with data in a visual context allows them to discover hidden relationships and patterns in the data without relying on help from IT.
4. **Promote data literacy:** Making data easier to work with and understand democratizes data analytics, getting more people across an organization involved.

# Examples



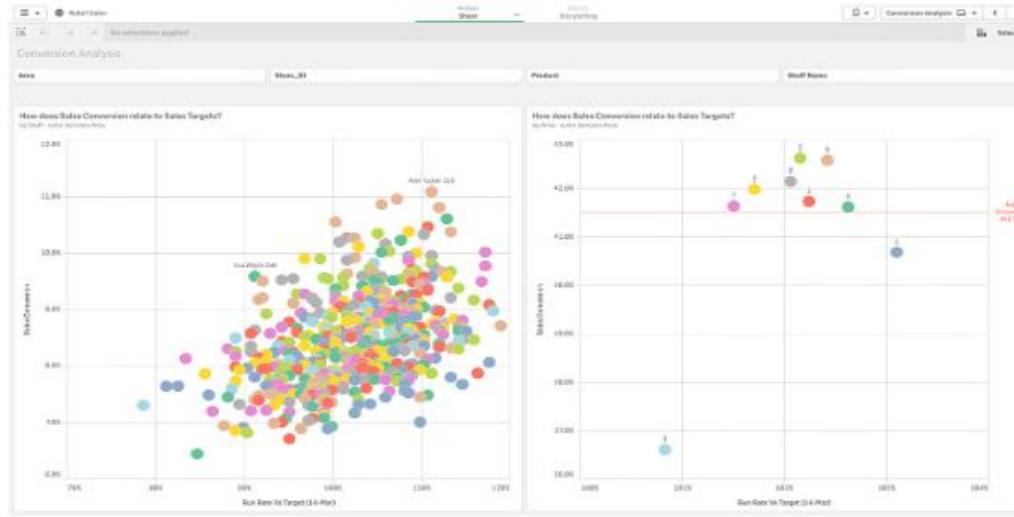
## Marketing

In this example, visual analytics helps a marketer boost ROI by letting them see and understand each phase of the customer life cycle. This is made possible by connecting data from the CRM, advertising tools, and web analytics



## Supply Chain

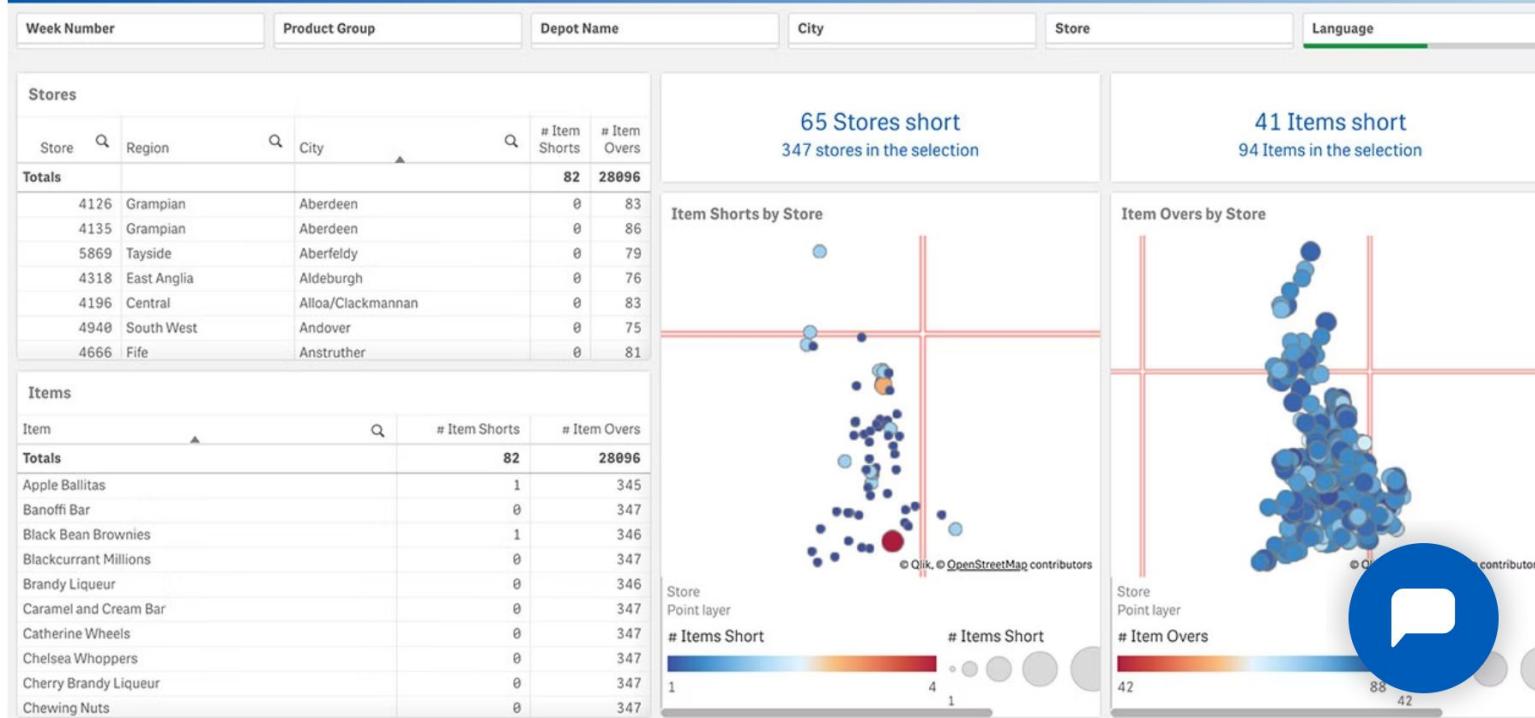
Big data visual analytics can help supply chain managers immediately find connections between complex, disparate data sources by surfacing KPIs and allowing interactive exploration.



## Sales

Sales managers benefit from the clean, organized presentation of sales data to increase revenue, improve forecasting, and identify key trends.

## Item Overs & Shorts



## Supply Chain

Big data visual analytics can help **supply chain** managers immediately find connections between complex, disparate data sources by surfacing KPIs and allowing interactive exploration.

# Marketing

## Dashboard

### Q ADVOCACY SEGMENT (100%)

Loyals  
Planning to Leave  
Switchers  
Trapped

TOTAL SALES  
**\$294,920k**

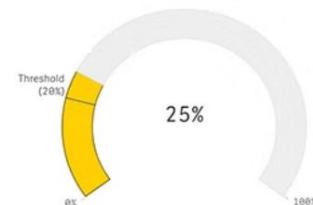
### NET PROMOTER SCORE



### OVERALL SATISFACTION



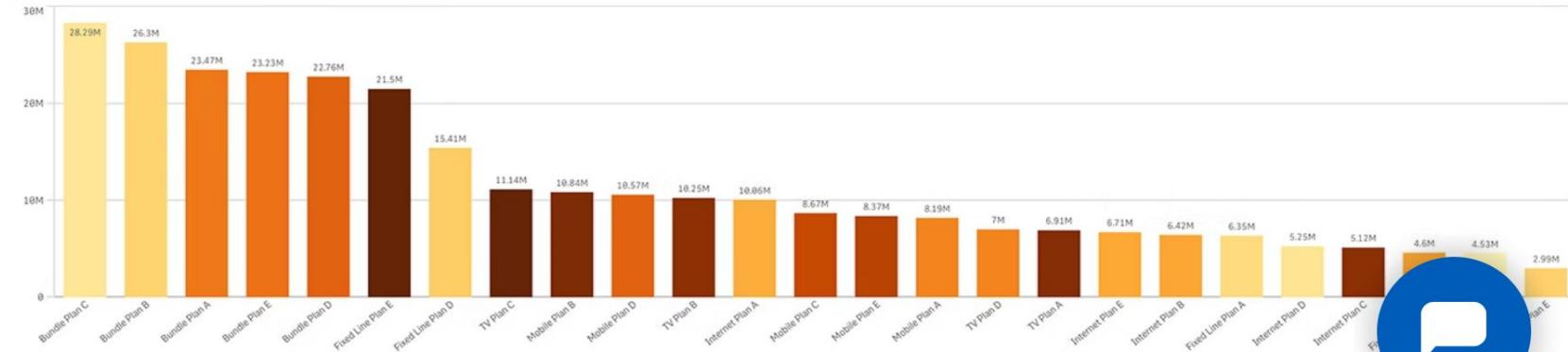
### CUSTOMER EFFORT



GROSS MARGIN  
**15%**

### How Does Revenue Compare by Product?

Amount in \$



Stronger Colour indicates better average NPS



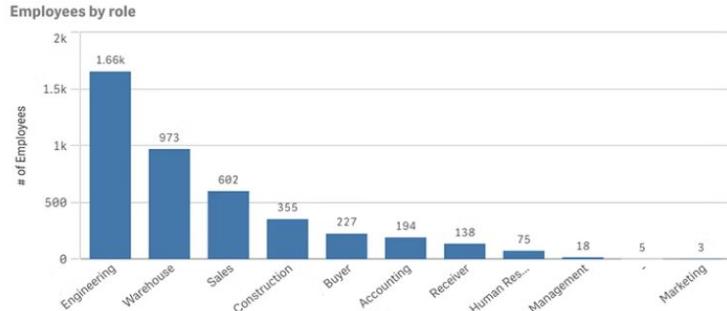
# HR

## Dashboard

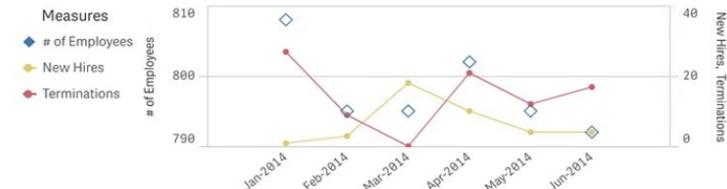
# of Employees  
**4.25k**

Salaries  
\$379.3M Avg Compensation  
\$79,382.31

New Hires  
**40** Terminations  
86



### New Hires and Terminations



Women Employees Ratio  
**36.3%**

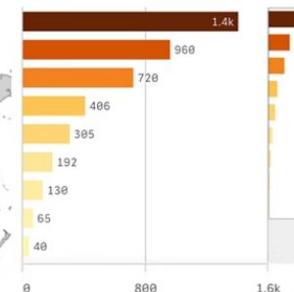
Completed Training ratio  
**4.6%**

Employee Satisfaction Ratio  
**48.7%**

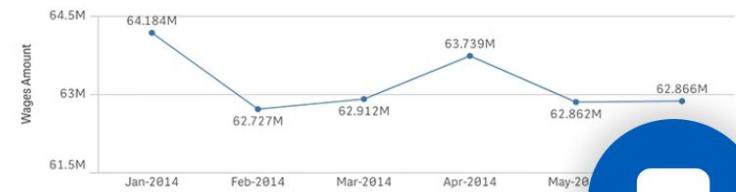
### Number of Employees by location



### Employees by country



### Compensation



# Finance

In this example, a loan manager at a consumer bank can explore how different regions, product and loan officers perform over time and drive the biggest impact on revenue and margins.



Current Balance (millions)  
**14,014**

Current # of Customers  
**537,829**

Current Margin %  
**5.31%**

Wtd Avg Credit Score (max Month)  
**767**



© Qlik

Original Credit Score

Original Deal Size

...

Loan Category

Fico Score Bucket

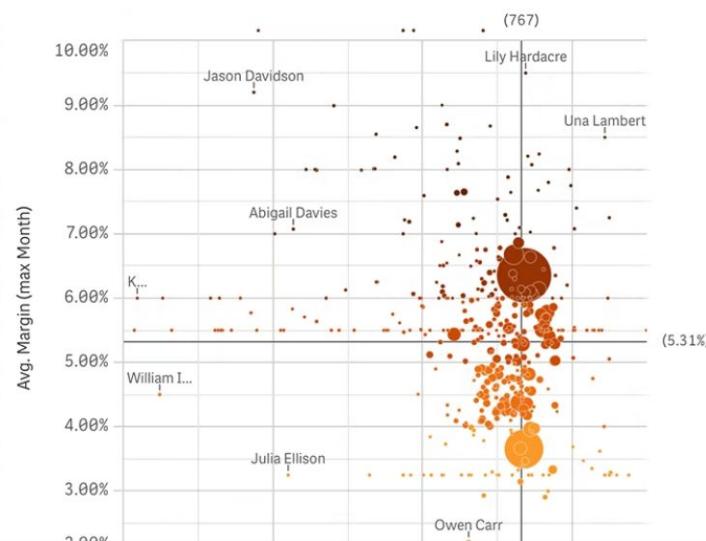
Loan Status

Delinquency Status

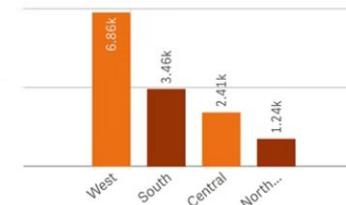
Loan APR

Term

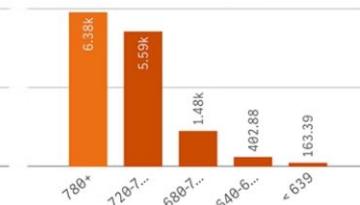
Margin v. Credit Score (100.0% of Total Selected)



Balance by Product



FICO Scores Distribution



Region  Loan Product Ca...

Loan Officer  CustomerID

Values

|               | Balance                | Avg. Margin (max Month) |
|---------------|------------------------|-------------------------|
| <b>Totals</b> | <b>\$13,980,057.28</b> | <b>5.31%</b>            |
| • Northeast   | \$1,241,455.04         | 6.02%                   |
| • Central     | \$2,414,386.79         | 4.91%                   |
| • South       | \$3,462,936.53         | 6.06%                   |
| • West        | \$6,861,278.92         | 4.95%                   |

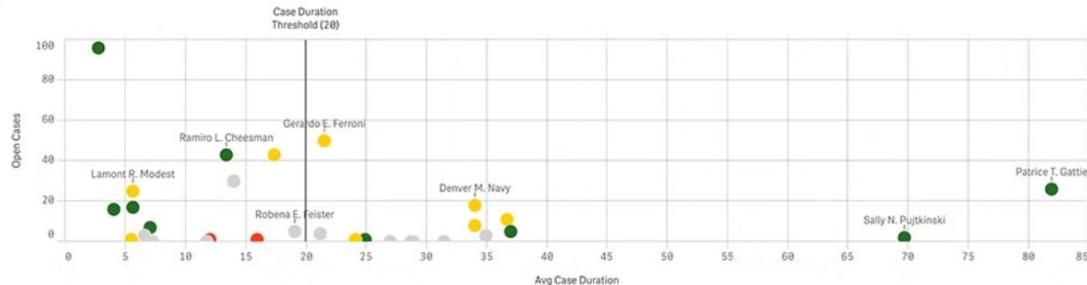


# IT

## High Priority Cases

72

Open Cases vs Duration  
\* color represents the priority with the most open cases



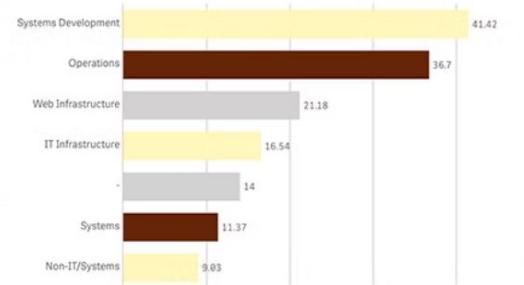
## Medium Priority Cases

142

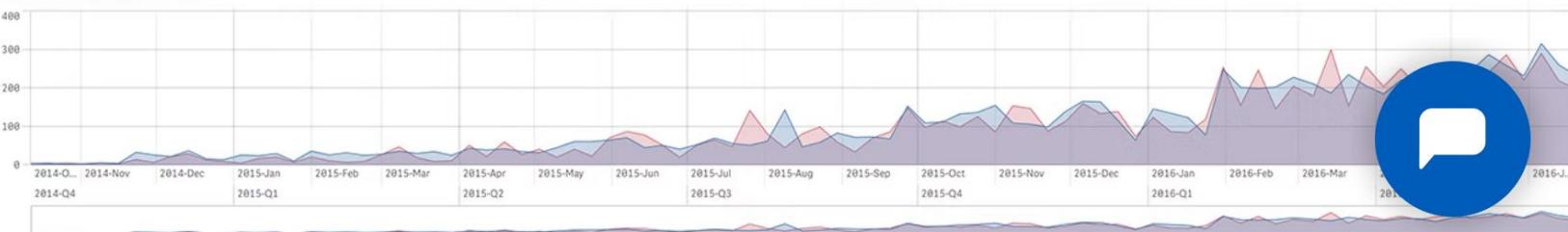
## Low Priority Cases

203

Avg Case Resolution Time (Days)  
\* color represents the priority with the most open cases



## Open & Resolved Cases Over Time



# Best Practices

## Define goals

Before you begin, be sure to define specific goals for your visual data analysis work. What specific questions are you trying to answer?

## Integrate and manage the data

Your source data needs to be transformed into clean, business-ready information. You'll need to combine and replicate data from a variety of sources and then bring it into standardized formats stored in a repository such as a [data lake](#) or [data warehouse](#).

## Simplify visualizations

Choose the right visual technique to present your story in the simplest way possible. Learn more about [visualizing data](#).

## Get Inspired

See the ten most compelling and interesting [data visualization examples](#) from recent years.

# Data Visualization

- Every day your business generates more data on sales revenue, marketing performance, customer interactions, inventory levels, production metrics, staffing levels, costs, and other KPIs. But with so much data to sift through, it can be difficult for people to see the story it tells.
- Data visualization helps you turn all that granular data into easily understood, visually compelling—and useful—business information. By tapping into external data sources, today's data visualization tools don't simply let you see your KPIs more clearly, they unify data and apply AI-driven analytics to reveal relationships between your KPIs, the market, and the world.

# Benefits

- Hidden within your data lie important insights that can help drive business forward. But the challenge is that you can't always connect the dots by looking at raw numbers alone. When you look at your data presented in a visual format, patterns, connections, and other "a-ha" insights emerge that would otherwise remain out of sight.
- Data visualization brings data to life, making you the master storyteller of the insights hidden within your numbers. Through live data dashboards, interactive reports, charts, graphs, and other visual representations, data visualization helps users develop powerful business insight quickly and effectively.

Measures: People with at least one dose

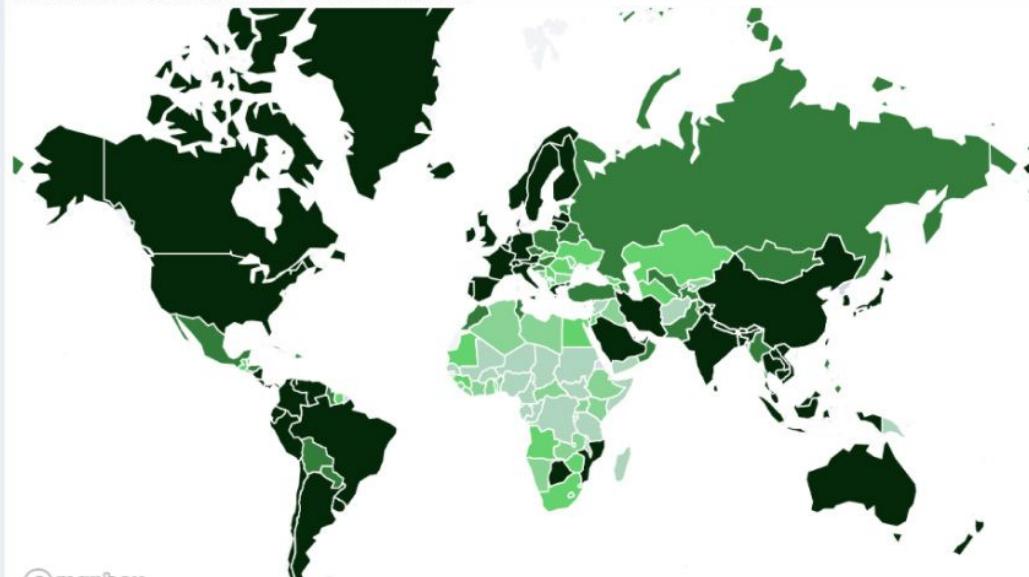
People fully vaccinated

View: By Location

By Date

This chart shows the percentage of people who have initiated COVID-19 vaccinations (given at least one dose). [Learn more](#)

#### PEOPLE INITIATING VACCINATION BY COUNTRY/REGION



© Mapbox © OpenStreetMap [Improve this map](#)

No data

>0 to 17.5%

>17.5% to 35%

>35% to 52.5% >52.5% to 70%

>70%

| Country/Region       | % of People Initiating Vaccination |
|----------------------|------------------------------------|
| Brunei               | 100.0%                             |
| Gibraltar            | 100.0%                             |
| Malta                | 100.0%                             |
| Samoa                | 100.0%                             |
| Tonga                | 100.0%                             |
| United Arab Emirates | 100.0%                             |
| Tokelau              | 96.8%                              |
| Portugal             | 95.3%                              |
| Cuba                 | 94.2%                              |
| Macao SAR            | 94.2%                              |
| Chile                | 94.1%                              |
| Cayman Islands       | 93.0%                              |
| Qatar                | 91.4%                              |
| Argentina            | 91.1%                              |
| Iceland              | 90.8%                              |
| Cambodia             | 90.0%                              |
| Bhutan               | 89.8%                              |
| China (mainland)     | 89.7%                              |
| Hong Kong SAR        | 89.7%                              |
| Peru                 | 89.2%                              |
| New Zealand          | 89.0%                              |
| Taiwan               | 88.7%                              |
| Nicaragua            | 88.6%                              |
| Spain                | 88.2%                              |

## Drive better decision making with data visualization

- **See the big picture.** There's a clear picture of performance buried within the transaction, interaction, process, and behavioral data stored in your systems. Data visualization allows you to recognize the broader context and higher-level scenario within it. As a result, you'll notice trends and spot patterns you wouldn't be able to see if you were looking at numbers on their own.
- **Identify the significance.** Bringing visual clarity to the story told within your data helps you identify insights that lead to better decision making, planning, strategies, and actions. How is your business performing, what needs to be modified, and where should you focus your resources? The ability to understand the significance of your data drives more effective operations and decisions.

- **Make informed decisions.** With concrete numbers and tangible insights, you can be confident your decisions are backed by data. Having clear insight into performance metrics empowers you with the knowledge and arms you with the tools to make the right decisions at the right time.
- **Track trends over time.** Once you've established a baseline, trends will begin to emerge. Track progress, spot trends, and begin using your insights to drive informed, strategic decisions. As you build your trends, shifts in patterns indicate if things drift off track, allowing you to immediately address any sign of lowered performance.

## Evaluating data visualization tools

- Today's data visualization tools run the gamut from free versions you can access with a browser to feature-rich platforms that integrate with a wide variety of mainstream business applications.
- But with so many data visualization tools, how do you choose the right solution for your specific needs? In the end, you want a tool that gives you the insights you need to help you focus on what matters most.
- Here are some key questions you'll want to keep in mind as you look for a data visualization solution with the right balance of power, flexibility, and ease-of-use for your needs:

# Key Questions Before Choose

## **What are your goals and needs?**

Take a moment to list out what types of data you have and what kinds of information and insights you want to gather.

## **What features do you require?**

Top data visualization solutions include features like business intelligence, analytics, enterprise reporting capabilities, giving you flexibility to meet your needs both today and tomorrow.

## **Where will your data come from?**

Consider how you'll connect or import your data and determine if the tool offers a seamless fit with your systems.

## **Where will you need to see your data?**

Find a tool that allows you to take your data and dashboards with you on any device, for quick insight and collaboration anywhere you are.

## **Where would you like to publish your visualization?**

Think about where your final presentation will be published. On a website, in a blog post, or in an online news article? Find a tool that allows flexibility to post and share your data visualization wherever your story should be told—one that lets you create data stories with amazing interactive visualizations and easily publish them to your blog or website.

# DATA Binding

Data binding is the process that couples two data sources together and synchronizes them. With data binding, a change to an element in a data set automatically updates in the bound data set.

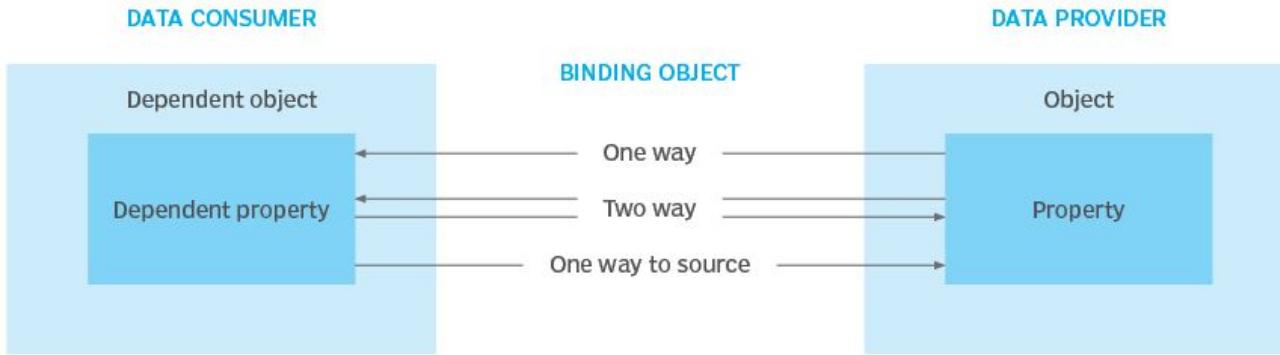
Data binding may be used for many reasons, such as to link an application's user interface (UI) and the data it displays, for data entry, reporting and in text box elements. It also lets internet users manipulate the representation of data in the elements of a web page without needing to use a complicated programming or scripting processes.

- Data and data objects of different logic functions can be bound together in data binding. Data types with different languages can also be bound, such as data binding Extensible Markup Language (XML) and UI.
- Each data change in one data set automatically reflects in the other bound data set. In binding syntax, the data source is the data provider, and the other data set is the data consumer. The binding forms the link between the data provider and the data consumer, enabling the connection between visual element data and a data source.

# Types

- **One-way binding** is a relatively simple type of data binding. Changes to the data provider update automatically in the data consumer data set, but not the other way around.
- **Two-way binding** is where changes to either the data provider or the data consumer automatically updates the other.
- **One-way-to-source binding** is the reverse of one-way binding. Changes to the data consumer automatically update the data provider but not the other way around.
- **One-time binding** is where changes to the data provider do not automatically update the data consumer. This approach is useful when only a snapshot of the data is needed, and the data is static.

# How types of data binding work



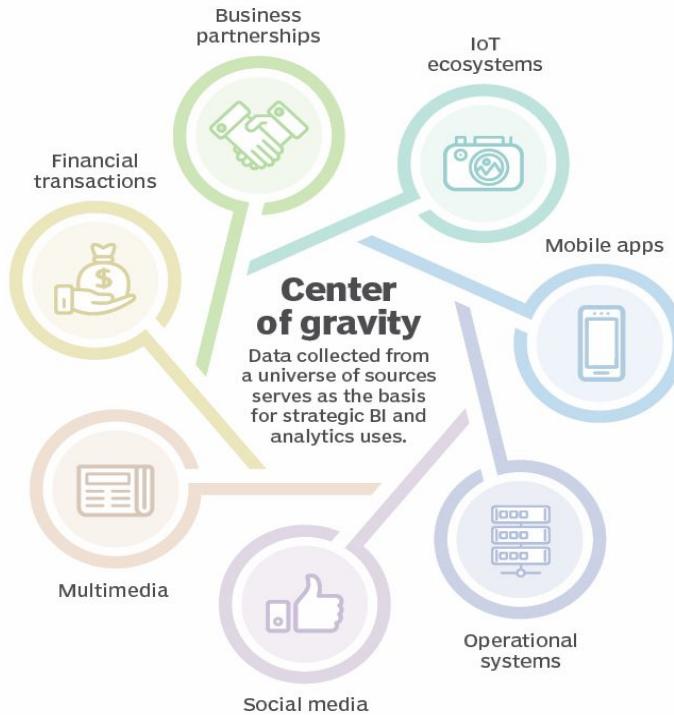
The following examples show how data binding can be used:

- **Reporting.** Binding is a common way to compile reports that display data from a data source to a screen or printer.
- **Data entry.** Data binding is also a common way to enter large amounts of data and keep it updated and synchronized to a data source.
- **Lookup tables.** Lookup tables are information tables that are typically a part of larger data displays. Data binding and controls are used to display and change data.
- **Master-detail formats.** This is a model for communication protocols where one device or process controls another. These formats may have two tables of data bound together.

# Data Collection (SQL)

- Data collection is the process of gathering data for use in business decision-making, strategic planning, research and other purposes.
- It's a crucial part of data analytics applications and research projects
- Effective data collection provides the information that's needed to answer questions, analyze business performance or other outcomes, and predict future trends, actions and scenarios.

- In businesses, data collection happens on multiple levels.
- IT systems regularly collect data on customers, employees, sales and other aspects of business operations when transactions are processed and data is entered.
- Companies also conduct surveys and track social media to get feedback from customers.
- Data scientists, other analysts and business users then collect relevant data to analyze from internal systems, plus external data sources if needed.



# Methods Data Collection

Data can be collected from one or more sources as needed to provide the information that's being sought. For example,

- To analyze sales and the effectiveness of its marketing campaigns,
- A retailer might collect customer data from transaction records, website visits, mobile applications, its loyalty program and an online survey.

# Methods

- automated data collection functions built into business applications, websites and mobile apps;
- sensors that collect operational data from industrial equipment, vehicles and other machinery;
- collection of data from information services providers and other external data sources;
- tracking social media, discussion forums, reviews sites, blogs and other online channels;
- surveys, questionnaires and forms, done online, in person or by phone, email or regular mail;
- focus groups and one-on-one interviews; and
- direct observation of participants in a research study.

# EX



# Challenges in Data Collection

- **Data quality issues.** Raw data typically includes errors, inconsistencies and other issues. Ideally, data collection measures are designed to avoid or minimize such problems. That isn't foolproof in most cases, though. As a result, collected data usually needs to be put through data profiling to identify issues and data cleansing to fix them.
- **Finding relevant data.** With a wide range of systems to navigate, gathering data to analyze can be a complicated task for data scientists and other users in an organization. The use of data curation techniques helps make it easier to find and access data. For example, that might include creating a data catalog and searchable indexes.

- **Deciding what data to collect.** This is a fundamental issue both for upfront collection of raw data and when users gather data for analytics applications. Collecting data that isn't needed adds time, cost and complexity to the process. But leaving out useful data can limit a data set's business value and affect analytics results.
- **Dealing with big data.** Big data environments typically include a combination of structured, unstructured and semistructured data, in large volumes. That makes the initial data collection and processing stages more complex. In addition, data scientists often need to filter sets of raw data stored in a data lake for specific analytics applications.
- **Low response and other research issues.** In research studies, a lack of responses or willing participants raises questions about the validity of the data that's collected. Other research challenges include training people to collect the data and creating sufficient quality assurance procedures to ensure that the data is accurate.

## **key steps in the data collection process**

1. Identify a business or research issue that needs to be addressed and set goals for the project.
2. Gather data requirements to answer the business question or deliver the research information.
3. Identify the data sets that can provide the desired information.
4. Set a plan for collecting the data, including the collection methods that will be used.
5. Collect the available data and begin working to prepare it for analysis.

# Data collection considerations and best practices

- There are two primary types of data that can be collected: quantitative data and qualitative data. The former is numerical -- for example, prices, amounts, statistics and percentages. Qualitative data is descriptive in nature -- e.g., color, smell, appearance and opinion.
- Organizations also make use of secondary data from external sources to help drive business decisions. For example, manufacturers and retailers might use U.S. census data to aid in planning their marketing strategies and campaigns. Companies might also use government health statistics and outside healthcare studies to analyze and optimize their medical insurance plans.

The European Union's General Data Protection Regulation ([GDPR](#)) and other privacy laws enacted in recent years make data privacy and security bigger considerations when collecting data, particularly if it contains personal information about customers. An organization's [data governance](#) program should include policies to ensure that data collection practices comply with laws such as GDPR.

Other data collection best practices include the following:

- Make sure you collect the right data to meet business or research needs.
- Ensure that the data is accurate, either as it's collected or as part of the data preparation process.
- Don't waste time and resources collecting irrelevant data

# Decision Making

- Visual Functionality & Presentation
- Integration
- Customization
- Mobility
- Training Requirements
- Pricing (If licensed)
- IT Support (If licensed, If open source how is the community)

# User Segmentation

# Market Segmentation

---

## **Types of Market Segmentation:**

- Demographic Segmentation
  - Behavioral Segmentation
  - Geographic Segmentation
  - Psychographic Segmentation
- 





---

## Demographic Segmentation

- Age
  - Education Level
  - Sex
  - Income
  - Marital Status
  - Race
  - Family Size
  - Nationality Religion
  - Occupation
-

---

## **Behavioral Segmentation**

- Online shopping habits
  - Actions taken on a website
  - Benefits sought
  - Loyalty
- 



---

**Geographic Segmentation,  
splitting up your market based on  
their location, is a basic but highly  
useful segmentation strategy**

---



# EXCEL FORMULA

A formula in Excel is used to do mathematical calculations. Formulas always start with the equal sign (=) typed in the cell, followed by your calculation.

Formulas can be used for calculations such as:

- $=1+1$
- $=2*2$
- $=4/2=2$

Step by step:

1. Select C1 and type (=)
2. Right click A1
3. Type (+)
4. Right click A2
5. Press enter
- 6.

SUM ▾

X ✓ fx =(A1+A2)

|   | A | B | C   | D | E |
|---|---|---|---|---|---|
| 1 | 2 |   | = <span style="border: 1px solid green; padding: 2px;">(A1+A2)</span> |   |   |
| 2 | 4 |   |   |   |   |
| 3 |   |   |   |   |   |
| 4 |   |   |   |   |   |
| 5 |   |   |   |   |   |

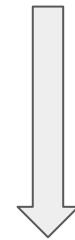
C1 ▾ fx =A1+A2

|    | A | B | C | D |
|----|---|---|---|---|
| 1  | 2 |   | 6 |   |
| 2  | 4 |   |   |   |
| 3  |   |   |   |   |
| 4  |   |   |   |   |
| 5  |   |   |   |   |
| 6  |   |   |   |   |
| 7  |   |   |   |   |
| 8  |   |   |   |   |
| 9  |   |   |   |   |
| 10 |   |   |   |   |

Lets change from addition to multiplication, by replacing the (+) with a (\*). It should now be =A1\*A2, press enter to see what happens.

| C1 | A | B | C      | D |
|----|---|---|--------|---|
| 1  | 2 |   |        |   |
| 2  | 4 |   | =A1*A2 |   |
| 3  |   |   |        |   |
| 4  |   |   |        |   |
| 5  |   |   |        |   |
| 6  |   |   |        |   |
| 7  |   |   |        |   |
| 8  |   |   |        |   |
| 9  |   |   |        |   |
| 10 |   |   |        |   |

| Trainers | Pokeball | Great Balls | Ultra Balls | Total     |
|----------|----------|-------------|-------------|-----------|
| sachin   | 2        | 3           | 1           | =B8+C8+D8 |
| ankita   | 5        | 5           | 2           |           |
| rahul    | 10       | 2           | 3           |           |



DRAG

# Sample

|   | A                    | B           | C   | D     | E        |
|---|----------------------|-------------|-----|-------|----------|
| 1 | Home supplies budget |             |     |       |          |
| 2 | S/N                  | Item        | Qty | Price | Subtotal |
| 4 | 1                    | Mangoes     | 9   | 600   |          |
| 5 | 2                    | Oranges     | 3   | 1200  |          |
| 6 | 3                    | Tomatoes    | 1   | 2500  |          |
| 7 | 4                    | Cooking Oil | 5   | 6500  |          |
| 8 | 5                    | Tonic water | 13  | 3900  |          |

FORMULA  
GOES HERE  
 $C4*D4$

## SUM()

The SUM() function, as the name suggests, gives the total of the selected range of cell values. It performs the mathematical operation which is addition. Here's an example of it below:

A screenshot of Microsoft Excel showing a table with four columns (A, B, C, D) and five rows (1, 2, 3, 4, 5). The first row contains column headers: 'Qty' in A1, 'Price per Unit' in B1, and 'Total Sales' in C1. Rows 2, 3, and 4 show data points: Row 2 has Qty 10, Price per Unit 30, and Total Sales 300; Row 3 has Qty 11, Price per Unit 35, and Total Sales 385; Row 4 has Qty 12, Price per Unit 40, and Total Sales 480. Row 5 is a summary row with 'Total' in C5 and 1165 in D5. The formula bar at the top shows '=SUM(C2:C4)'.

|   | A   | B              | C           | D |
|---|-----|----------------|-------------|---|
| 1 | Qty | Price per Unit | Total Sales |   |
| 2 | 10  | 30             | 300         |   |
| 3 | 11  | 35             | 385         |   |
| 4 | 12  | 40             | 480         |   |
| 5 |     | Total          | 1165        |   |

# Average()

The AVERAGE() function focuses on calculating the average of the selected range of cell values. As seen from the below example, to find the avg of the total sales, you have to simply type in “AVERAGE(C10, C11, C12)”.

| Average |                |                       |
|---------|----------------|-----------------------|
| QTY     | Price per Unit | Total Sales           |
| 10      | 30             | 300                   |
| 11      | 35             | 385                   |
| 12      | 40             | 480                   |
|         | AVERAGE        | =AVERAGE(C10,C11,C12) |

AVERAGE(number1, [number2], [number3], [number4], ...)

# count()

The function **COUNT()** counts the total number of cells in a range that contains a number. It does not include the cell, which is blank, and the ones that hold data in any other format apart from numeric.

|   | QTY     | Price per Unit | Total Sales         |
|---|---------|----------------|---------------------|
| 5 | 10      | 30             | 300                 |
| 6 | 11      | 35             | 385                 |
| 7 | 12      | 40             | 480                 |
| 8 | AVERAGE |                | =COUNT(C17,C18,C19) |

# Counta()

Count numerical, text, value etc However, COUNTA() does not count any blank cells.

| QTY   | Price per Unit | Total Sales                   |
|-------|----------------|-------------------------------|
| 10    | 30             | 300                           |
| 11    | 35             | 385                           |
| 12    | 40             | 480                           |
| COUNT |                | =COUNTA(C24:C27)              |
|       |                | COUNTA(value1, [value2], ...) |

Moving ahead, let's now understand how the subtotal function works. The SUBTOTAL() function returns the subtotal in a database. Depending on what you want, you can select either average, count, sum, min, max, min, and others. Let's have a look at two such examples.

The screenshot shows a Microsoft Excel interface. The formula bar at the top displays the formula `=SUBTOTAL(1,A2:A4)`. Below the formula bar is a table with five rows and five columns. The columns are labeled A, B, C, D, and E. The first row contains column headers: "Qty" in A, "Price per Unit" in B, and "Total Sales" in C. The second row contains values: 10 in A, 30 in B, and 300 in C. The third row contains values: 11 in A, 35 in B, and 385 in C. The fourth row contains values: 12 in A, 40 in B, and 480 in C. The fifth row is a subtotal row, with "Subtotal" in A and 11 in C. The cells in row 5 are highlighted with green backgrounds, while the other cells are white with black borders.

|   | A   | B              | C           | D | E |
|---|-----|----------------|-------------|---|---|
| 1 | Qty | Price per Unit | Total Sales |   |   |
| 2 | 10  | 30             | 300         |   |   |
| 3 | 11  | 35             | 385         |   |   |
| 4 | 12  | 40             | 480         |   |   |
| 5 |     | Subtotal       | 11          |   |   |

|  |                           |                       |   |   |   |  |
|--|---------------------------|-----------------------|---|---|---|--|
| 29   |                           |                       |   |   |   |  |
| 30   | subtotal                  |                       |   |   |   |  |
| 31   |                           |                       |   |   |   |  |
| 32   | <b>QTY</b>                | <b>Price per Unit</b> | <b>Total Sales</b>  |   |   |  |
| 33   | 10                        | 30                    | 300   |   |   |  |
| 34   | 11                        | 35                    | 385   |   |   |  |
| 35   | 12                        | 40                    | 480   |   |   |  |
| 36   | subtotal                  | =SUBTOTAL(            | (... 1 - AVERAGE<br>(... 2 - COUNT<br>(... 3 - COUNTA<br>(... 4 - MAX<br>(... 5 - MIN<br>(... 6 - PRODUCT<br>(... 7 - STDEV.S<br>(... 8 - STDEV.P<br>(... 9 - SUM<br>(... 10 - VAR.S<br>(... 11 - VAR.P<br>(... 101 - AVERAGE |   |   |  |
| 37   |                           | SUBTOTAL(function_n   |   |   |   |  |
| 38   |                           |                       |   |   |   |  |
| 39   |                           |                       |   |   |   |  |
| 40   |                           |                       |   |   |   |  |
| 41   |                           |                       |   |   |   |  |
| 42   |                           |                       |   |   |   |  |
| 43   |                           |                       |   |   |   |  |
| 44   |                           |                       |   |   |   |  |
|  |                           | SUM                   |   |   |   |  |
|  |                           | Formulas              |   |   |   |  |
|  |                           | +                     |   |   |   |  |
| Edit   | Accessibility: Good to go |                       |   |   |   |  |
|  28°C |                           |                       |   |   |   |  |
| Cloudy   |                           |                       |   |  |  |  |

# Modulus

The MOD() function works on returning the remainder when a particular number is divided by a divisor. Let's now have a look at the examples below for better understanding.

- In the first example, we have divided 10 by 3. The remainder is calculated using the function “=MOD(A2,3)”. The result is stored in B2. We can also directly type “=MOD(10,3)” as it will give the same answer.

A screenshot of a Microsoft Excel spreadsheet. The formula bar shows the formula =MOD(A2,3). The spreadsheet contains the following data:

|   | A       | B | C | D | E |
|---|---------|---|---|---|---|
| 1 | Modulus |   |   |   |   |
| 2 | 10      | 1 |   |   |   |
| 3 | 12      | 0 |   |   |   |
| 4 | 45      | 3 |   |   |   |

# Modulus

- Similarly, here, we have divided 12 by 4. The remainder is 0 is, which is stored in B3.

The screenshot shows a Microsoft Excel spreadsheet. The formula bar at the top displays the formula `=MOD(A3,4)`. The table below has columns labeled A, B, C, D, and E. Row 1 contains the header "Modulus". Row 2 contains values 10 and 1. Row 3 contains values 12 and 0, where 0 is highlighted with a green selection border. Row 4 contains values 45 and 3. The cell B3, which contains the value 0, is also highlighted with a green selection border. The status bar at the bottom of the screen shows the text "1 cell selected".

| A | B       | C | D | E |
|---|---------|---|---|---|
| 1 | Modulus |   |   |   |
| 2 | 10      | 1 |   |   |
| 3 | 12      | 0 |   |   |
| 4 | 45      | 3 |   |   |

# Power

The function “Power()” returns the result of a number raised to a certain power. Let’s have a look at the examples shown below:

A screenshot of Microsoft Excel showing a table and the formula bar. The formula bar shows the formula =POWER(A2,3). The table has columns A and B. Row 1 contains the header "Power". Row 2 contains values 10 and 1000. Row 3 contains values 4 and 256. Row 4 is empty.

|   | A     | B    | C | D | E |
|---|-------|------|---|---|---|
| 1 | Power |      |   |   |   |
| 2 | 10    | 1000 |   |   |   |
| 3 | 4     | 256  |   |   |   |
| 4 |       |      |   |   |   |

# Ceiling

we have the ceiling function. The CEILING() function rounds a number up to its nearest multiple of significance

| B2 |         |    | X | ✓ | fx | =CEILING(A2,5) |
|----|---------|----|---|---|----|----------------|
| 1  | A       | B  | C | D | E  |                |
| 2  | Ceiling |    |   |   |    |                |
|    | 35.316  | 40 |   |   |    |                |

The nearest highest multiple of 5 for 35.316 is 40.

# Floor

Contrary to the Ceiling function, the floor function rounds a number down to the nearest multiple of significance

The nearest lowest multiple of 5 for 35.316 is 35.

| B2 | A      | B  | C | D | E |  |
|----|--------|----|---|---|---|--|
| 1  | Floor  |    |   |   |   |  |
| 2  | 35.316 | 35 |   |   |   |  |

# Concatenate

This function merges or joins several text strings into one text string. Given below are the different ways to perform this function.

In this example, we have operated with the syntax =CONCATENATE(A25, " ", B25)

| A26                            |          | X ✓ fx =CONCATENATE(A25, " ", B25) |
|--------------------------------|----------|------------------------------------|
| 23                             | A        | B                                  |
| Concatenate - combines strings |          |                                    |
| 25                             | Hello    | World                              |
| 26                             |          | Hello World                        |
| 27                             | Excel is | fun to learn                       |
| 28                             |          | Excel is fun to learn              |

# LEN

The function LEN() returns the total number of characters in a string. So, it will count the overall characters, including spaces and special characters. Given below is an example of the Len function.

A screenshot of a Microsoft Excel spreadsheet. The formula bar at the top shows the cell reference B7 and the formula =LEN(A7). The main area displays two rows of data. Row 6 contains the header "length" in cell A6. Row 7 contains the word "World" in cell A7 and the value 5 in cell B7. Row 8 contains the word "Microsoft" in cell A8 and the value 9 in cell B8. The columns are labeled A and B.

|   | A         | B |
|---|-----------|---|
| 6 | length    |   |
| 7 | World     | 5 |
| 8 | Microsoft | 9 |

# REPLACE

As the name suggests, the REPLACE() function works on replacing the part of a text string with a different text string.

The syntax is “=REPLACE(old\_text, start\_num, num\_chars, new\_text)”. Here, start\_num refers to the index position you want to start replacing the characters with. Next, num\_chars indicate the number of characters you want to replace.

Let's have a look at the ways we can use this function.

- Here, we are replacing A101 with B101 by typing “=REPLACE(A15,1,1,"B")”.

A screenshot of Microsoft Excel demonstrating the use of the REPLACE function. The formula bar shows the formula =REPLACE(A15,1,1,"B"). The spreadsheet has two columns, A and B. Row 14 contains the word "Replace". Row 15 shows the result of the formula: the value "A101" in column A is replaced by "B101" in column B. Rows 16 and 17 show other data: "A102" in A16 and "A2102" in B16; "Adam" in A17 and "Saam" in B17.

|    | A       | B     |
|----|---------|-------|
| 14 | Replace |       |
| 15 | A101    | B101  |
| 16 | A102    | A2102 |
| 17 | Adam    | Saam  |

# REPLACE

- Next, we are replacing A102 with A2102 by typing “=REPLACE(A16,1,1, "A2")”.

| Replace |       |
|---------|-------|
| A       | B     |
| 14      |       |
| 15 A101 | B101  |
| 16 A102 | A2102 |
| 17 Adam | Saam  |

- Finally, we are replacing Adam with Saam by typing “=REPLACE(A17,1,2, "Sa")”.

| Replace |       |
|---------|-------|
| A       | B     |
| 14      |       |
| 15 A101 | B101  |
| 16 A102 | A2102 |
| 17 Adam | Saam  |

# SUBSTITUTE

The **SUBSTITUTE()** function replaces the existing text with a new text in a text string.

The syntax is “=SUBSTITUTE(text, old\_text, new\_text, [instance\_num])”.

Here, [instance\_num] refers to the index position of the present texts more than once.

Given below are a few examples of this function:

- Here, we are substituting “I like” with “He likes” by typing “=SUBSTITUTE(A20, “I like”, “He likes”)”.

|                   | A                           | B                           |
|-------------------|-----------------------------|-----------------------------|
| <b>Substitute</b> |                             |                             |
| 19                |                             |                             |
| 20                | I like Excel                | He likes Excel              |
| 21                | MS Excel 2010, MS Word 2010 | MS Excel 2010, MS Word 2016 |
| 22                | MS Excel 2010, MS Word 2010 | MS Excel 2016, MS Word 2016 |

- Next, we are substituting the second 2010 that occurs in the original text in cell A21 with 2016 by typing “=SUBSTITUTE(A21,2010, 2016,2)”.

The screenshot shows a Microsoft Excel interface. The formula bar at the top contains the formula `=SUBSTITUTE(A21,2010, 2016,2)`. Below the formula bar is a small toolbar with a 'Substitute' button. The main area displays a table with four rows and two columns. The first row is a header with column labels 'A' and 'B'. Rows 20 and 21 are visible, while row 19 is partially visible above them. Row 22 is the next row down. The data in the table is as follows:

| A  | B                           |
|----|-----------------------------|
| 19 | Substitute                  |
| 20 | I like Excel                |
| 21 | MS Excel 2010, MS Word 2010 |
| 22 | MS Excel 2010, MS Word 2010 |

- Now, we are replacing both the 2010s in the original text with 2016 by typing “=SUBSTITUTE(A22,2010,2016)”.

The screenshot shows a Microsoft Excel interface. The formula bar at the top contains the formula `=SUBSTITUTE(A22,2010,2016)`. Below the formula bar is a table with two columns, A and B. The table has four rows, labeled 19, 20, 21, and 22. Row 19 is a header row with the title "Substitute". Rows 20, 21, and 22 contain data. In column A, the values are "I like Excel", "MS Excel 2010, MS Word 2010", and "MS Excel 2010, MS Word 2010" respectively. In column B, the values are "He likes Excel", "MS Excel 2010, MS Word 2016", and "MS Excel 2016, MS Word 2016" respectively. The cell containing the formula in the formula bar is highlighted with a green border.

|            | A                           | B                           |
|------------|-----------------------------|-----------------------------|
| Substitute |                             |                             |
| 20         | I like Excel                | He likes Excel              |
| 21         | MS Excel 2010, MS Word 2010 | MS Excel 2010, MS Word 2016 |
| 22         | MS Excel 2010, MS Word 2010 | MS Excel 2016, MS Word 2016 |

# LEFT()

The LEFT() function gives the number of characters from the start of a text string

- In the example below, we use the function left to obtain the leftmost word on the sentence in cell A5.

A screenshot of a Microsoft Excel spreadsheet. The formula bar at the top shows the formula =LEFT(A5,5). The spreadsheet has four columns labeled A, B, C, and D. Row 4 contains empty cells. Row 5 contains the text "Excel is fun to learn". Cell A5 is highlighted in orange and contains the text "Excel". Cell B5 is also highlighted in orange and contains the text "Excel". Cells C5 and D5 are empty. The formula bar also shows the cell reference B5.

|   | B5                    | ⋮     | X      | ✓        | fx | =LEFT(A5,5) |
|---|-----------------------|-------|--------|----------|----|-------------|
| 4 | A                     | B     | C      | D        |    |             |
| 5 | Excel is fun to learn | Excel | is fun | to learn |    |             |
| 6 |                       |       |        |          |    |             |

# Right ( )

the right() function returns the number of characters from the end of a text string.

| D5 | A                     | B     | C      | D        |
|----|-----------------------|-------|--------|----------|
| 4  |                       |       |        |          |
| 5  | Excel is fun to learn | Excel | is fun | to learn |
| 6  |                       |       |        |          |

# Mid ( )

MID() function returns the characters from the middle of a text string, given a starting position and length.

| C5 | A                           | B | C      | D        |
|----|-----------------------------|---|--------|----------|
| 4  |                             |   |        |          |
| 5  | Excel is fun to learn Excel |   | is fun | to learn |
| 6  |                             |   |        |          |

# UPPER

The UPPER() function converts any text string to uppercase

|        |        |            |   |  |
|--------|--------|------------|---|--|
|        |        | =UPPER(E1) |   |  |
| D      | E      | F          | G |  |
| amrita | AMRITA |            |   |  |

# Lower ( )

LOWER() function converts any text string to lowercase

|        |            |   |
|--------|------------|---|
| x      | =LOWER(E1) |   |
| D      | E          | F |
| SACHIN | sachin     |   |

(Ctrl) ▾

# PROPER ( )

The PROPER() function converts any text string to proper case, i.e., the first letter in each word will be in uppercase, and all the other will be in lowercase.

|   |   |                       |   |
|---|---|-----------------------|---|
|   |   | =PROPER(E1)           |   |
| D | E | sachiN dATA aRchitect | F |
|   |   | Sachin Data Architect |   |

# Now ( )

The NOW() function in Excel gives the current system date and time.

|   |       | Get & Transform  |   | Connections |  |  |
|---|-------|------------------|---|-------------|--|--|
|   |       | =NOW()           |   |             |  |  |
| D | E     | F                | G |             |  |  |
|   | Now() | 05-08-2022 23:44 |   |             |  |  |
|   |       |                  |   |             |  |  |

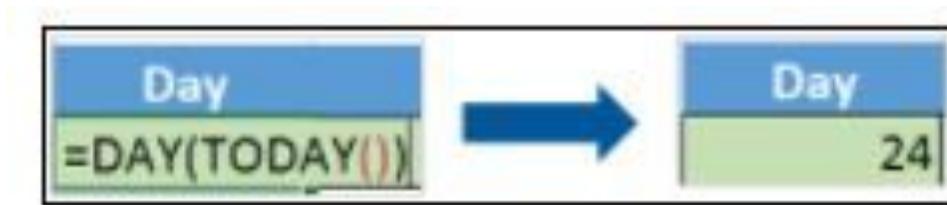
# TODAY ( )

The TODAY() function in Excel provides the current system date.



# Day ( )

The function DAY() is used to return the day of the month. It will be a number between 1 to 31. 1 is the first day of the month, 31 is the last day of the month.



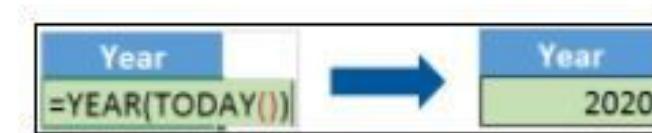
# Month ( )

The MONTH() function returns the month, a number from 1 to 12, where 1 is January and 12 is December.



# Year ( )

The YEAR() function, as the name suggests, returns the year from a date value.



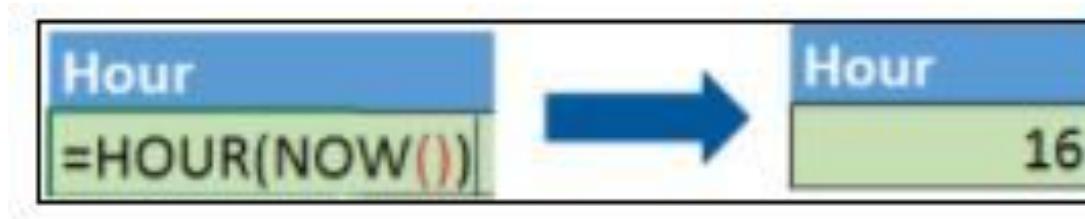
# Time ( )

The TIME() function converts hours, minutes, seconds given as numbers to an Excel serial number, formatted with a time format.



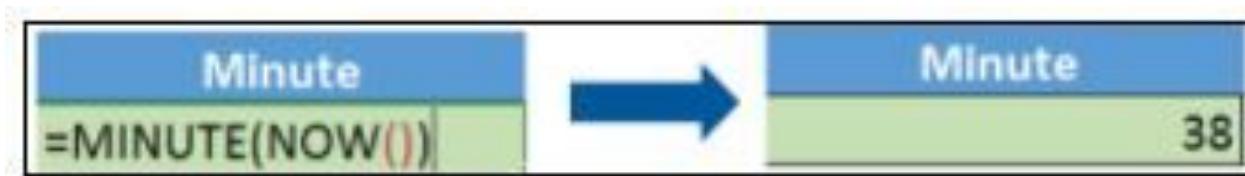
# Hour ( )

The HOUR() function generates the hour from a time value as a number from 0 to 23. Here, 0 means 12 AM and 23 is 11 PM



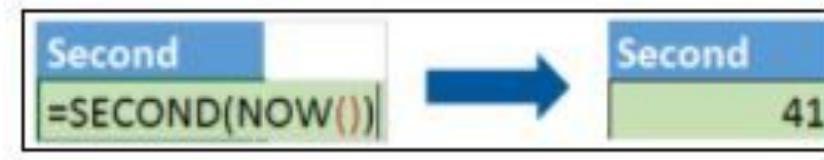
# Minute ( )

The function MINUTE(), returns the minute from a time value as a number from 0 to 59.



# Second ( )

The SECOND() function returns the second from a time value as a number from 0 to 59.



# Datedif ( )

The DATEDIF() function provides the difference between two dates in terms of years, months, or days.

Below is an example of a DATEDIF function where we calculate the current age of a person based on two given dates, the date of birth and today's date.

| Calculate Age | Datedif               |
|---------------|-----------------------|
| DOB           | 30-12-1994            |
| Today         | 24-08-2020            |
| Age           | =DATEDIF(B12,B13,"y") |

→

| Calculate Age | Datedif    |
|---------------|------------|
| DOB           | 30-12-1994 |
| Today         | 24-08-2020 |
| Age           | 25         |

# lookup

the LOOKUP function searches for value in the *lookup\_range* and returns the value in the *result\_range* that is in the same position.

The syntax for the LOOKUP function in Microsoft Excel is:

```
LOOKUP( value, lookup_range, [result_range] )
```

A screenshot of Microsoft Excel version 12.0 (Office 2007) showing a table of fruit orders. The table has columns for Order ID, Product, Unit Price, and Quantity. Row 5 contains the formula =LOOKUP(10251, A1:A6, B1:B6), which looks up the value 10251 in the range A1:A6 and returns the corresponding value from the range B1:B6, which is "Pears".

|   | A        | B       | C          | D        | E | F     | G |
|---|----------|---------|------------|----------|---|-------|---|
| 1 | Order ID | Product | Unit Price | Quantity |   | Pears |   |
| 2 | 10247    | Apples  | \$14.00    | 12       |   |       |   |
| 3 | 10249    | Oranges | \$9.80     | 10       |   |       |   |
| 4 | 10250    | Bananas | \$34.80    | 5        |   |       |   |
| 5 | 10251    | Pears   | \$18.60    | 9        |   |       |   |
| 6 | 10252    | Grapes  | \$42.30    | 40       |   |       |   |
| 7 |          |         |            |          |   |       |   |

# VLOOKUP ( )

In its simplest form, the VLOOKUP function says: =VLOOKUP(**What you want to look up, where you want to look for it, the column number in the range containing the value to return, return an Approximate or Exact match – indicated as 1/TRUE, or 0/FALSE**).

VLOOKUP supports approximate and exact matching

```
=VLOOKUP (lookup_value, table_array, column_index_num, [range_lookup])
```

# HLOOKUP ( )

**Use HLOOKUP when your comparison values are located in a row across the top of a table of data, and you want to look down a specified number of rows.** Use VLOOKUP when your comparison values are located in a column to the left of the data you want to find. The H in HLOOKUP stands for "Horizontal."

HLOOKUP([value], [range], [row number], [false or true])

# Dropdown

# Create a Drop-down List

To create a drop-down list in Excel, execute the following steps.

1. On the second sheet, type the items you want to appear in the drop-down list.

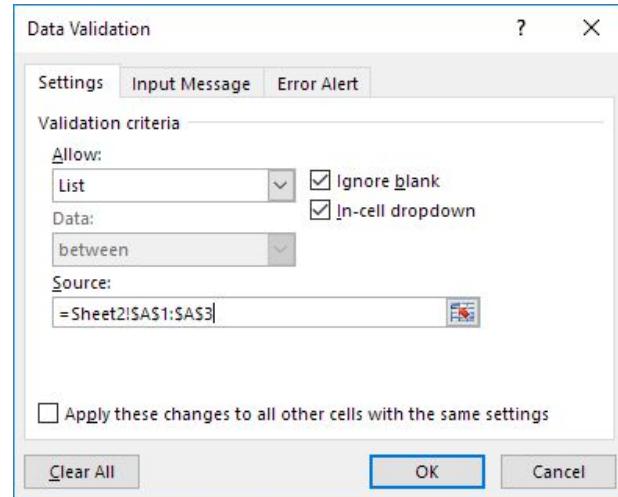
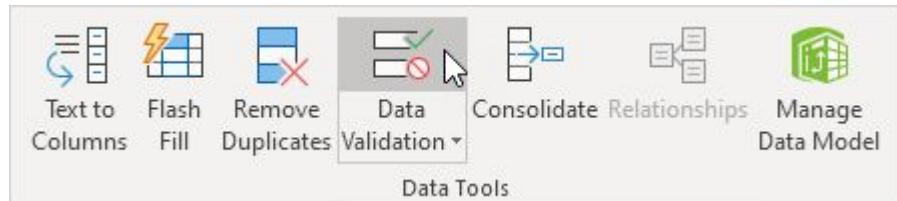
. On the first sheet, select cell B1.

The 'Data Validation' dialog box appears.

4. In the Allow box, click List.

5. Click in the Source box and select the range A1:A3 on Sheet2.

|   | A              | B        | C | D | E | F | G | H | I |
|---|----------------|----------|---|---|---|---|---|---|---|
| 1 | Favorite Food: |          |   |   |   |   |   |   |   |
| 2 |                | Pizza    |   |   |   |   |   |   |   |
| 3 |                | Pancakes |   |   |   |   |   |   |   |
| 4 |                | Chinese  |   |   |   |   |   |   |   |



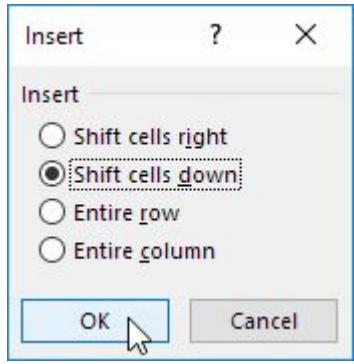
# Add/Remove Items

You can add or remove items from a drop-down list in Excel without opening the 'Data Validation' dialog box and changing the range reference. This saves time.

1. To add an item to a drop-down list, go to the items and select an item.

|   | A        | B | C | D | E | F | G | H | I |
|---|----------|---|---|---|---|---|---|---|---|
| 1 | Pizza    |   |   |   |   |   |   |   |   |
| 2 | Pancakes |   |   |   |   |   |   |   |   |
| 3 | Chinese  |   |   |   |   |   |   |   |   |
| 4 |          |   |   |   |   |   |   |   |   |





| A |          | B | C | D | E | F | G | H | I |
|---|----------|---|---|---|---|---|---|---|---|
| 1 | Pizza    |   |   |   |   |   |   |   |   |
| 2 |          |   |   |   |   |   |   |   |   |
| 3 | Pancakes |   |   |   |   |   |   |   |   |
| 4 | Chinese  |   |   |   |   |   |   |   |   |
| 5 |          |   |   |   |   |   |   |   |   |

The screenshot shows a Microsoft Excel spreadsheet. The ribbon at the top is visible with the Home tab selected. The formula bar shows cell A2. The main workspace displays a list of food items: Pizza, Pancakes, and Chinese. The 'Sheet2' tab is currently active. The 'Insert' dialog box is open on the left, showing options for inserting rows or columns. The 'Shift cells down' option is selected. The 'OK' button in the dialog box is being clicked, as indicated by the mouse cursor icon.

4. Type a new item.

|   | A        | B | C | D | E | F | G | H | I |
|---|----------|---|---|---|---|---|---|---|---|
| 1 | Pizza    |   |   |   |   |   |   |   |   |
| 2 | Soup     |   |   |   |   |   |   |   |   |
| 3 | Pancakes |   |   |   |   |   |   |   |   |
| 4 | Chinese  |   |   |   |   |   |   |   |   |
| 5 |          |   |   |   |   |   |   |   |   |

|   | A              | B        | C | D | E | F | G | H | I |
|---|----------------|----------|---|---|---|---|---|---|---|
| 1 | Favorite Food: |          |   |   |   |   |   |   |   |
| 2 |                | Pizza    |   |   |   |   |   |   |   |
| 3 |                | Soup     |   |   |   |   |   |   |   |
| 4 |                | Pancakes |   |   |   |   |   |   |   |
| 5 |                | Chinese  |   |   |   |   |   |   |   |

# Unique

Let's say we want to find a list of distinct values from the Item column (C) in the table below:

```
=UNIQUE(C7:C21)
```

The screenshot shows a Microsoft Excel spreadsheet with a table of data and a formula bar. The formula bar displays the formula `=UNIQUE(C7:C21)`. A red dashed arrow points from the formula bar down to cell C25, which contains the word "UNIQUE". Another red dashed arrow points from cell C25 down to cell C26, which contains the value "Stamps". The table below has columns labeled Department, Item, Quantity, Price, and Total \$. The total value for the table is 716.60.

|    | B            | C        | D        | E     | F        |
|----|--------------|----------|----------|-------|----------|
| 6  | Department   | Item     | Quantity | Price | Total \$ |
| 7  | Support      | Stamps   | 50       | 0.60  | 30.00    |
| 8  | Support      | Paper    | 10       | 4.99  | 49.90    |
| 9  | Support      | Pens     | 15       | 0.95  | 14.25    |
| 10 | Support      | Staples  | 1        | 1.10  | 1.10     |
| 11 | Support      | Scissors | 1        | 4.70  | 4.70     |
| 12 | Front Office | Stamps   | 250      | 0.60  | 150.00   |
| 13 | Front Office | Paper    | 30       | 4.99  | 149.70   |
| 14 | Front Office | Pens     | 50       | 0.95  | 47.50    |
| 15 | Front Office | Staples  | 10       | 1.10  | 11.00    |
| 16 | Front Office | Scissors | 0        | 4.70  | -        |
| 17 | Sales        | Stamps   | 200      | 0.60  | 120.00   |
| 18 | Sales        | Paper    | 20       | 4.99  | 99.80    |
| 19 | Sales        | Pens     | 25       | 0.95  | 23.75    |
| 20 | Sales        | Staples  | 5        | 1.10  | 5.50     |
| 21 | Sales        | Scissors | 2        | 4.70  | 9.40     |
| 22 | Total        |          |          |       | 716.60   |
| 23 |              |          |          |       |          |
| 24 |              |          |          |       |          |
| 25 | UNIQUE       |          |          |       |          |
| 26 | Stamps       |          |          |       |          |
| 27 | Paper        |          |          |       |          |
| 28 | Pens         |          |          |       |          |
| 29 | Staples      |          |          |       |          |

# IF ( )

The **IF()** function checks a given condition and returns a particular value if it is TRUE. It will return another value if the condition is FALSE.

In the below example, we want to check if the value in cell A2 is greater than 5. If it's greater than 5, the function will return "Yes 4 is greater", else it will return "No".

In this case, it will return 'No' since 4 is not greater than 5.

|   | A  | B                                   | C | D |
|---|----|-------------------------------------|---|---|
| 1 | IF |                                     |   |   |
| 2 |    | =IF(A2>5, "Yes 4 is greater", "No") |   |   |

# IFERROR

'IFERROR' is another function that is popularly used. This function returns a value if an expression evaluates to an error, or else it will return the value of the expression.

Suppose you want to divide 10 by 0. This is an invalid expression, as you can't divide a number by zero. It will result in an error.

|  | =IFERROR(E1/E2,H1) |    |             |   |             |   |
|--|--------------------|----|-------------|---|-------------|---|
|  | D                  | E  | F           | G | H           | I |
|  |                    | 10 |             |   | cant divide |   |
|  |                    | 0  | cant divide |   |             |   |

# countif() , count occurrence with condition

- =COUNTIF(A2:A5,"London")
- =COUNTIF(A2:A5,A4)

| Trainers | Pokeball | Great Balls | Ultra Balls | Total |
|----------|----------|-------------|-------------|-------|
| sachin   | 2        | 3           | 1           | 6     |
| ankita   | 5        | 5           | 2           | 12    |
| rahul    | 10       | 2           | 3           | 15    |

=COUNTIF(B8:D10,2)  
COUNTIF(range, criteria)

The function COUNTIF() is used to count the total number of cells within a range that meet the given condition.

# sumif()

The SUMIF() function adds the cells specified by a given condition or criteria.

=SUMIF(C5:C456,"Blue",F5:F456)

Sumif

=SUMIF(G:G, "India",E:E)



Sumif

207615

# Analysis ToolPak in Excel

## Excel Options

- General
- Formulas
- Proofing
- Save
- Language
- Accessibility
- Advanced
- Customize Ribbon
- Quick Access Toolbar
- Add-ins**
- Trust Center

View and manage Microsoft Office Add-ins.

### Add-ins

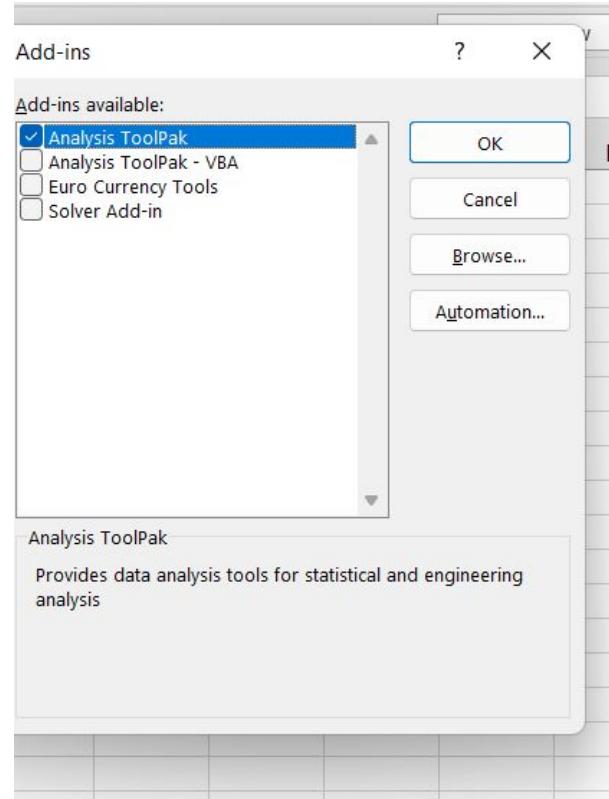
| Name                                | Location                                       | Type               |
|-------------------------------------|--|--------------------|
| <b>Active Application Add-ins</b>   |  |                    |
| No Active Application Add-ins       |  |                    |
| <b>Inactive Application Add-ins</b> |  |                    |
| Analysis ToolPak                    | C:\...t\Office16\Library\Analysis\ANALYS32.XLL | Excel Add-in       |
| Analysis ToolPak - VBA              | C:\...fice16\Library\Analysis\ATPVBAEN.XLAM    | Excel Add-in       |
| Date (XML)                          | C:\...es\Microsoft Shared\Smart Tag\MOFLDLL    | Action             |
| Euro Currency Tools                 | C:\...e\root\Office16\Library\EUROTOOL.XLAM    | Excel Add-in       |
| Microsoft Actions Pane 3            |  | XML Expansion Pack |
| Microsoft Power Map for Excel       | C:\...ap Excel Add-in\EXCELPLUGINSHELL.DLL     | COM Add-in         |
| Solver Add-in                       | C:\...t\Office16\Library\SOLVER\SOLVER.XLAM    | Excel Add-in       |
| <b>Document Related Add-ins</b>     |  |                    |
| No Document Related Add-ins         |  |                    |
| <b>Disabled Application Add-ins</b> |  |                    |
| No Disabled Application Add-ins     |  |                    |

Add-in: Analysis ToolPak  
Publisher: Microsoft Office  
Compatibility: No compatibility information available  
Location: C:\Program Files\Microsoft Office\root\Office16\Library\Analysis\ANALYS32.XLL

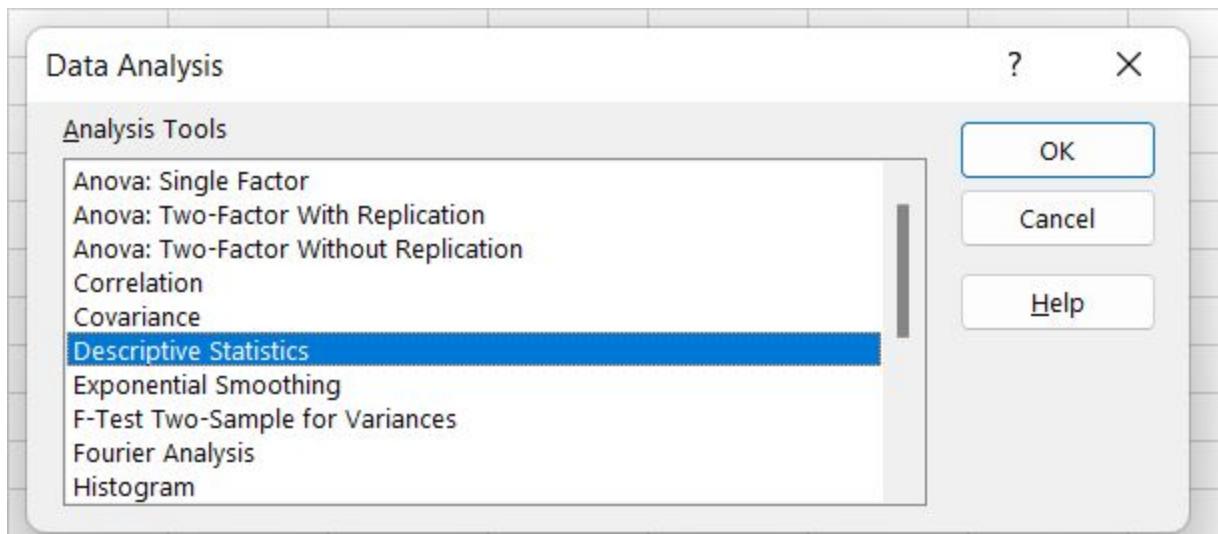
Description: Provides data analysis tools for statistical and engineering analysis

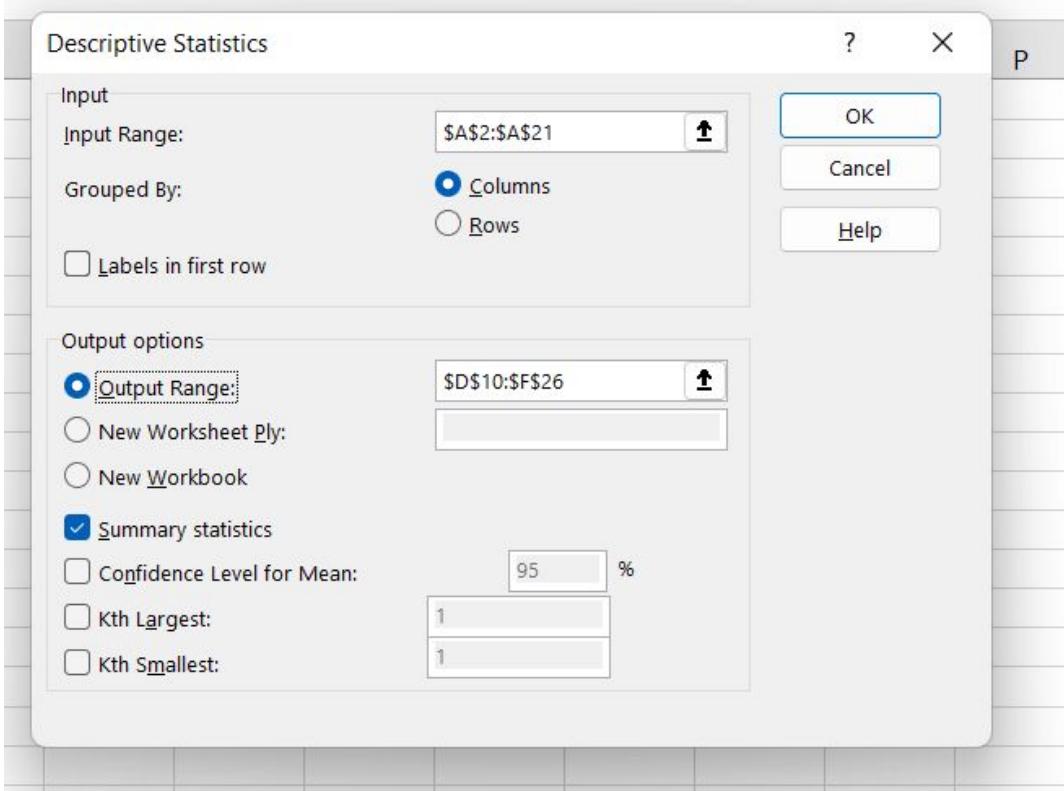
Manage: Excel Add-ins Go...

OK Cancel



# Descriptive Statistics in Excel



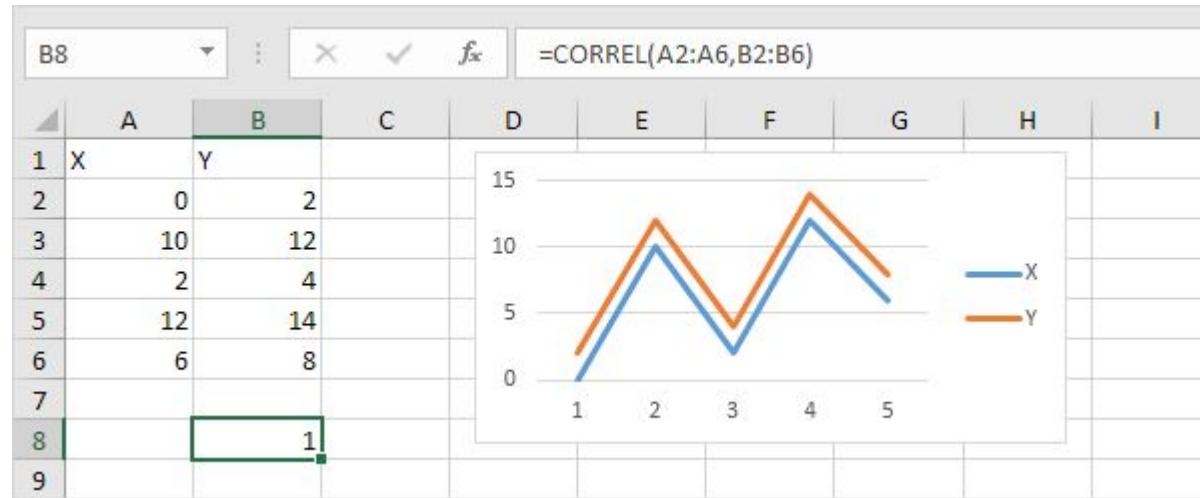


| Column1    |              |
|------------|--------------|
| Mean       | 47.25        |
| Standard E | 7.879311685  |
| Median     | 55           |
| Mode       | 87           |
| Standard D | 35.23735309  |
| Sample Va  | 1241.671053  |
| Kurtosis   | -1.715300569 |
| Skewness   | -0.003594223 |
| Range      | 96           |
| Minimum    | 4            |
| Maximum    | 100          |
| Sum        | 945          |
| Count      | 20           |
|            | 0            |

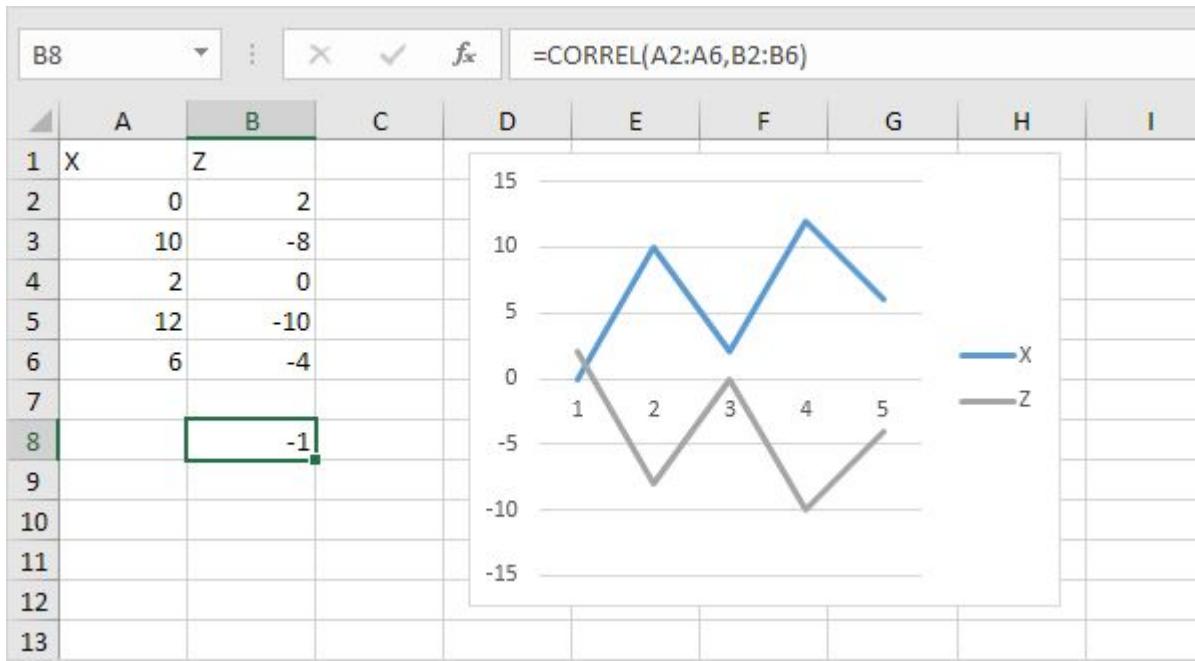
# Correlation in Excel

The correlation coefficient (a value between -1 and +1) tells you how strongly two variables are related to each other. We can use the CORREL function or the Analysis Toolpak add-in in Excel to find the correlation coefficient between two variables.

- A correlation coefficient of +1 indicates a perfect positive correlation. As variable X increases, variable Y increases. As variable X decreases, variable Y decreases.



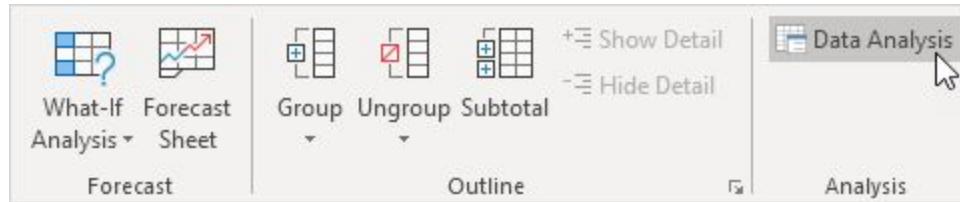
A correlation coefficient of -1 indicates a perfect negative correlation. As variable X increases, variable Z decreases. As variable X decreases, variable Z increases.



- A correlation coefficient near 0 indicates no correlation.

To use the Analysis Toolpak add-in in Excel to quickly generate correlation coefficients between multiple variables, execute the following steps.

1. On the Data tab, in the Analysis group, click Data Analysis.



Note: can't find the Data Analysis button? Click here to load the [Analysis ToolPak add-in](#).

2. Select Correlation and click OK.

## Correlation

?

X

### Input

Input Range:

\$A\$1:\$C\$6



OK

Grouped By:

Columns

Rows

Cancel

Help

Labels in first row

### Output options

Output Range:

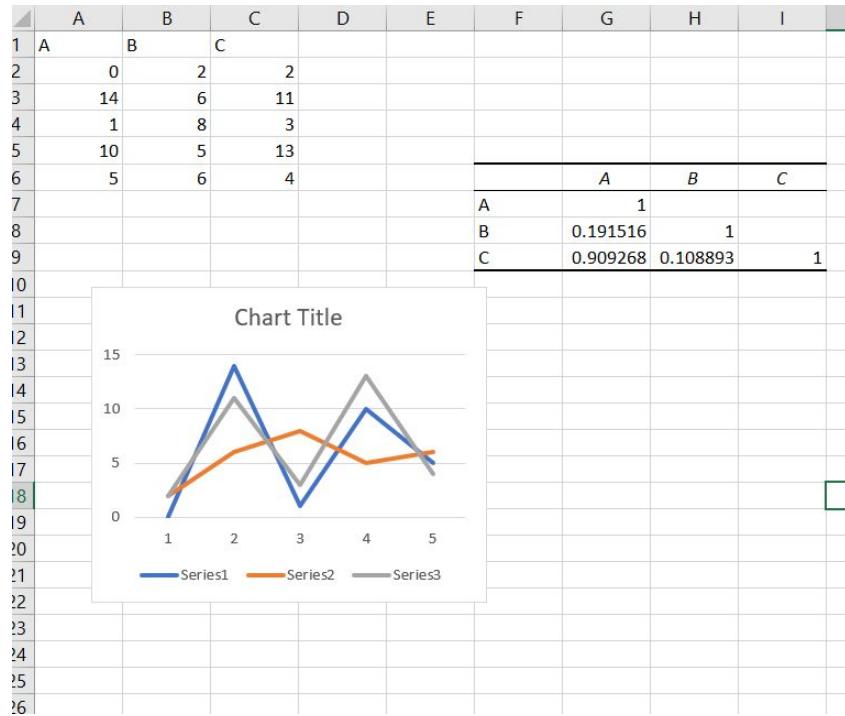
\$F\$6



New Worksheet Ply:

New Workbook

Conclusion: variables A and C are positively correlated (0.91). Variables A and B are not correlated (0.19). Variables B and C are also not correlated (0.11) . You can verify these conclusions by looking at the graph.



# Pivot Table

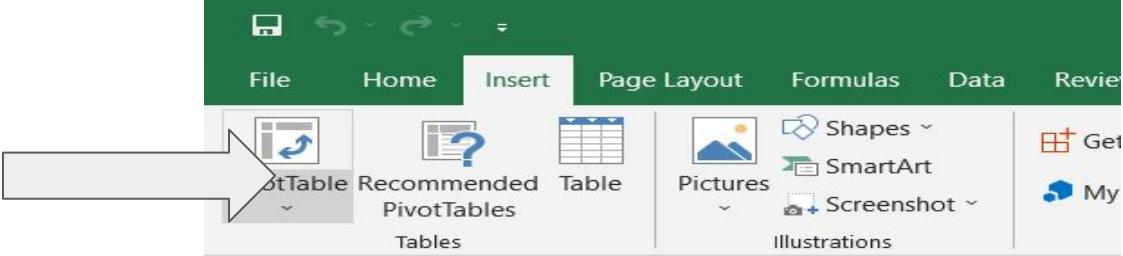
PivotTable is an extremely powerful tool that you can use to slice and dice data. In this tutorial, you will learn these PivotTable features in detail along with examples. By the time you complete this tutorial, you will have sufficient knowledge on PivotTable features that can get you started with exploring, analyzing, and reporting data based on the requirements.

# Features

- ◆ Creating a PivotTable is extremely simple and fast
- ◆ Enabling churning of data instantly by simple dragging of fields, sorting and filtering and different calculations on the data.
- ◆ Arriving at the suitable representation for your data as you gain insights into it.
- ◆ Ability to create reports on the fly.
- ◆ Producing multiple reports from the same PivotTable in a matter of seconds.
- ◆ Providing interactive reports to synchronize with the audience.



# Pivot Table



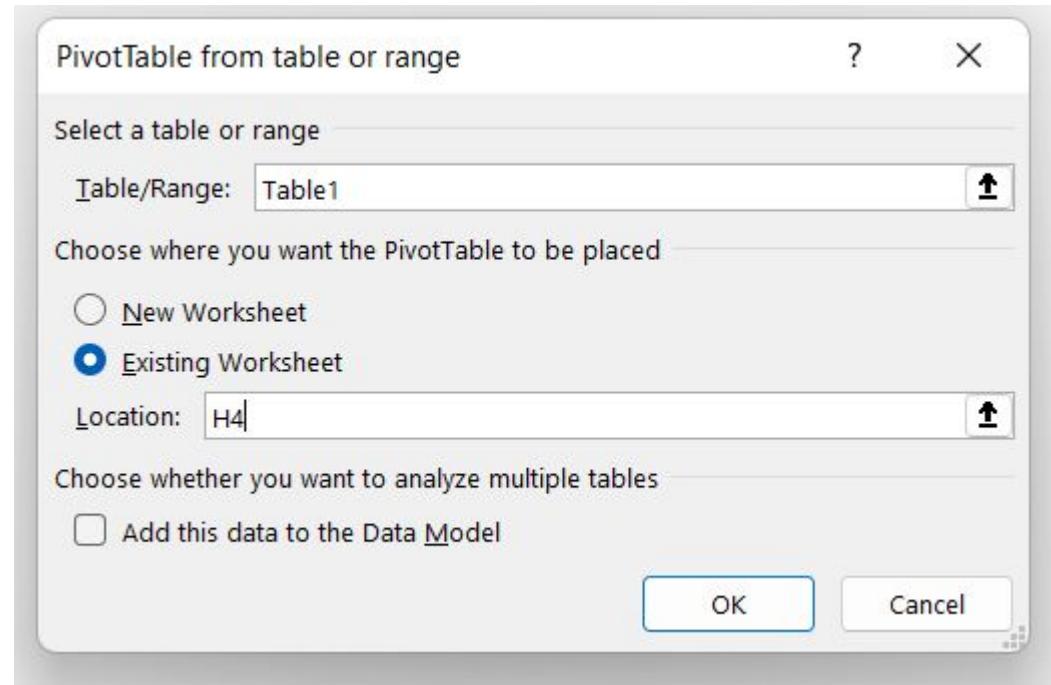
The screenshot shows the Microsoft Excel ribbon with the 'Insert' tab selected. A callout arrow points from the text in step 1 to the 'PivotTable' icon in the 'Tables' group. Below the ribbon, a sample sales data table is displayed in a worksheet. The table has columns for Date, Color, Region, Units, and Sales.

| Date      | Color | Region | Units | Sales    |
|-----------|-------|--------|-------|----------|
| 3-Jan-16  | Red   | West   | 1     | \$11.00  |
| 13-Jan-16 | Blue  | South  | 8     | \$96.00  |
| 21-Jan-16 | Green | West   | 2     | \$26.00  |
| 30-Jan-16 | Blue  | North  | 7     | \$84.00  |
| 7-Feb-16  | Green | North  | 8     | \$104.00 |
| 13-Feb-16 | Red   | South  | 2     | \$22.00  |
| 21-Feb-16 | Blue  | East   | 5     | \$60.00  |
| 1-Mar-16  | Green | West   | 2     | \$26.00  |
| 13-Mar-16 | Blue  | East   | 8     | \$96.00  |
| 23-Mar-16 | Blue  | North  | 7     | \$84.00  |
| 28-Mar-16 | Green | West   | 2     | \$26.00  |
| 3-Apr-16  | Blue  | South  | 8     | \$96.00  |

1. To start off, select *any cell in the data* and click Pivot Table on the Insert tab of the ribbon:

The [sample data](#) contains 452 records with 5 fields of information: Date, Color, Units, Sales, and Region. This data is perfect for a pivot table.

2. Override the default location and enter H4 to place the pivot table on the current worksheet:



Excel will display the Create Pivot Table window. Notice the data range is already filled in. The default location for a new pivot table is New Worksheet.

3. Click OK, and Excel builds an empty pivot table starting in cell H4.

H4 : X ✓ f<sub>x</sub>

A B C D E F G H I J K

1  
2 Sample sales data      New empty Pivot Table  
3  
4 Date Color Region Units Sales  
5 3-Jan-16 Red West 1 \$11.00  
6 13-Jan-16 Blue South 8 \$96.00  
7 21-Jan-16 Green West 2 \$26.00  
8 30-Jan-16 Blue North 7 \$84.00  
9 7-Feb-16 Green North 8 \$104.00  
10 13-Feb-16 Red South 2 \$22.00  
11 21-Feb-16 Blue East 5 \$60.00  
12 1-Mar-16 Green West 2 \$26.00  
13 13-Mar-16 Blue East 8 \$96.00  
14 23-Mar-16 Blue North 7 \$84.00  
15 28-Mar-16 Green West 2 \$26.00  
16 3-Apr-16 Blue South 8 \$96.00  
17 12-Apr-16 Green South 1 \$13.00  
18 16-Apr-16 Red East 8 \$88.00

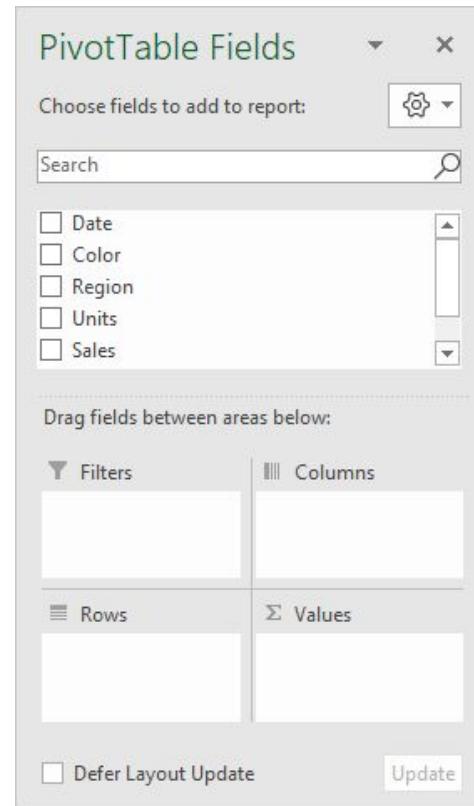
PivotTable2

To build a report, choose fields from the PivotTable Field List

The image shows a Microsoft Excel spreadsheet with data from row 4 to 18. The columns are labeled Date, Color, Region, Units, and Sales. A handwritten note 'New empty Pivot Table' is written above the empty PivotTable2 window. A yellow arrow points from the text to the PivotTable2 window. The PivotTable2 window contains the text 'To build a report, choose fields from the PivotTable Field List'. Below the window are icons for a grid, a checkmark, and a blue H-shaped icon.

Excel also displays the PivotTable Fields pane, which is empty at this point. Note all five fields are listed, but unused:

To build a pivot table, drag fields into one the Columns, Rows, or Values area. The Filters area is used to apply global filters to a pivot table.



## Add fields

1. Drag the Sales field to the Values area.

Excel calculates a grand total, 26356. This is the sum of all sales values in the entire data set:

The screenshot shows a Microsoft Excel spreadsheet titled "Sample sales data". The table has columns for Date, Color, Region, Units, and Sales. A formula bar at the top shows "Sum of Sales" with the value "26356". A yellow arrow points from the text "Grand total of ALL data" to the cell containing the value "26356".

|    | A                 | B     | C      | D     | E        | F | G | H            | I |
|----|-------------------|-------|--------|-------|----------|---|---|--------------|---|
| 1  | Sample sales data |       |        |       |          |   |   |              |   |
| 2  | Date              | Color | Region | Units | Sales    |   |   | Sum of Sales |   |
| 3  | 3-Jan-16          | Red   | West   | 1     | \$11.00  |   |   |              |   |
| 4  | 13-Jan-16         | Blue  | South  | 8     | \$96.00  |   |   |              |   |
| 5  | 21-Jan-16         | Green | West   | 2     | \$26.00  |   |   |              |   |
| 6  | 30-Jan-16         | Blue  | North  | 7     | \$84.00  |   |   |              |   |
| 7  | 7-Feb-16          | Green | North  | 8     | \$104.00 |   |   |              |   |
| 8  | 13-Feb-16         | Red   | South  | 2     | \$22.00  |   |   |              |   |
| 9  | 21-Feb-16         | Blue  | East   | 5     | \$60.00  |   |   |              |   |
| 10 | 1-Mar-16          | Green | West   | 2     | \$26.00  |   |   |              |   |
| 11 |                   |       |        |       |          |   |   |              |   |
| 12 |                   |       |        |       |          |   |   |              |   |

Grand total of  
ALL data

2. Drag the Color field to the Rows area.

Excel breaks out sales by Color. You can see Blue is the top seller, while Silver comes in last:

The screenshot shows a Microsoft Excel interface with a pivot table and a pivot chart. The pivot table is titled "Sample sales data" and contains the following data:

|    | Date      | Color | Region | Units | Sales    |
|----|-----------|-------|--------|-------|----------|
| 1  | 3-Jan-16  | Red   | West   | 1     | \$11.00  |
| 2  | 13-Jan-16 | Blue  | South  | 8     | \$96.00  |
| 3  | 21-Jan-16 | Green | West   | 2     | \$26.00  |
| 4  | 30-Jan-16 | Blue  | North  | 7     | \$84.00  |
| 5  | 7-Feb-16  | Green | North  | 8     | \$104.00 |
| 6  | 13-Feb-16 | Red   | South  | 2     | \$22.00  |
| 7  | 21-Feb-16 | Blue  | East   | 5     | \$60.00  |
| 8  | 1-Mar-16  | Green | West   | 2     | \$26.00  |
| 9  | 13-Mar-16 | Blue  | East   | 8     | \$96.00  |
| 10 | 23-Mar-16 | Blue  | North  | 7     | \$84.00  |
| 11 | 28-Mar-16 | Green | West   | 2     | \$26.00  |
| 12 | 3-Apr-16  | Blue  | South  | 8     | \$96.00  |
| 13 |           |       |        |       |          |
| 14 |           |       |        |       |          |
| 15 |           |       |        |       |          |
| 16 |           |       |        |       |          |

To the right of the pivot table, a pivot chart is displayed, showing a 3D bar chart of Sales by Color. The chart has four bars: Blue (approx. 74), Green (approx. 64), Red (approx. 55), and Silver (approx. 69). The chart title is "Sales by Color".

Below the pivot table, a pivot chart is shown with the following data:

| Color              | Sum of Sales |
|--------------------|--------------|
| Blue               | 7464         |
| Green              | 6414         |
| Red                | 5508         |
| Silver             | 6970         |
| <b>Grand Total</b> | <b>26356</b> |

At the bottom right of the slide, there is a logo for "EXCELJET" with a small orange icon.

You can see Color is a Row field, and Sales is a

PivotTable Fields ▼ X

Choose fields to add to report: ⚙️ ▾

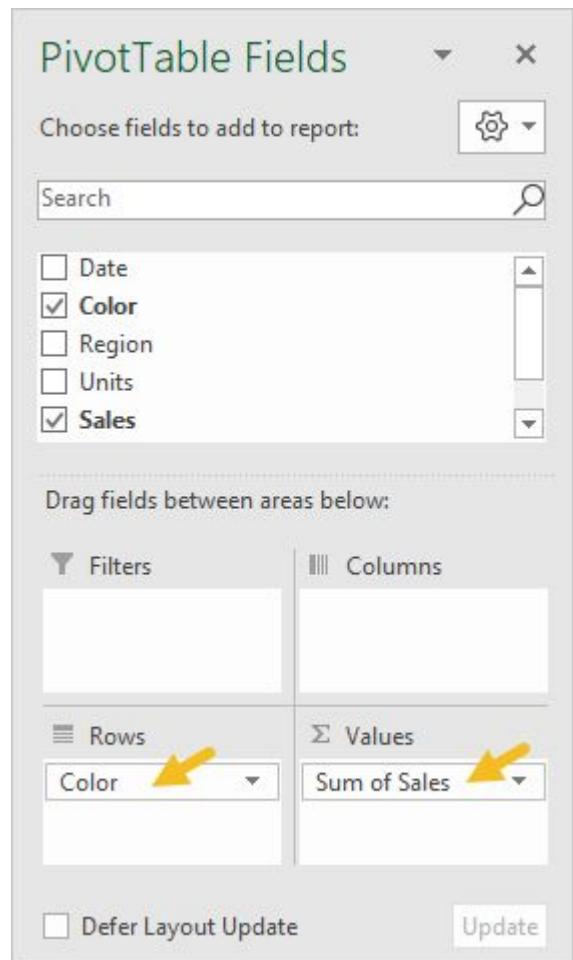
Search 🔍

Date  
 Color  
 Region  
 Units  
 Sales

Drag fields between areas below:

| Filters                                     | Columns  |
|---|--|
|   |  |
| Rows  | Values   |
| Color <span style="color: yellow;">↗</span> | Sum of Sales <span style="color: yellow;">↗</span> |
|   |  |

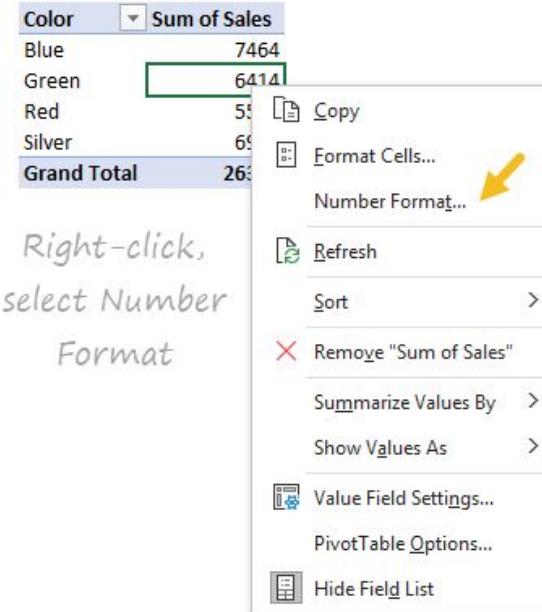
Defer Layout Update Update



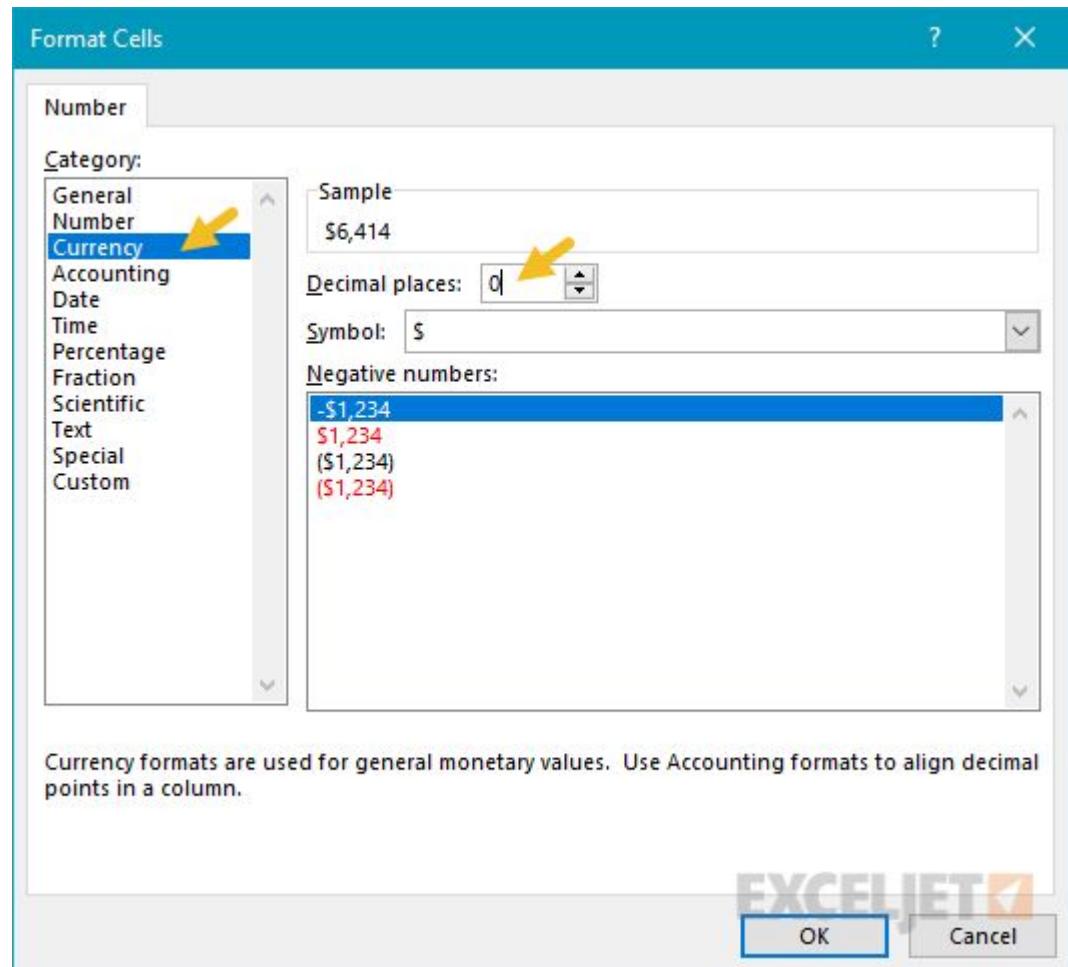
## Number formatting

Pivot Tables can apply and maintain number formatting automatically to numeric fields. This is a big time-saver when data changes frequently.

1. Right-click any Sales number and choose Number Format:



2. Apply Currency formatting with zero decimal places, then click OK:



- In the resulting pivot table, all sales values have Currency format applied
- Currency format will continue to be applied to Sales values, even when the pivot table is reconfigured, or new data is added.

Sample sales data

| Date      | Color | Region | Units | Sales    | Color              | Sum of Sales |
|-----------|-------|--------|-------|----------|--------------------|--------------|
| 3-Jan-16  | Red   | West   | 1     | \$11.00  | Blue               | \$7,464      |
| 13-Jan-16 | Blue  | South  | 8     | \$96.00  | Green              | \$6,414      |
| 21-Jan-16 | Green | West   | 2     | \$26.00  | Red                | \$5,508      |
| 30-Jan-16 | Blue  | North  | 7     | \$84.00  | Silver             | \$6,970      |
| 7-Feb-16  | Green | North  | 8     | \$104.00 | <b>Grand Total</b> |              |
| 13-Feb-16 | Red   | South  | 2     | \$22.00  | <b>\$26,356</b>    |              |
| 21-Feb-16 | Blue  | East   | 5     | \$60.00  |                    |              |
| 1-Mar-16  | Green | West   | 2     | \$26.00  |                    |              |
| 13-Mar-16 | Blue  | East   | 8     | \$96.00  |                    |              |
| 23-Mar-16 | Blue  | North  | 7     | \$84.00  |                    |              |
| 28-Mar-16 | Green | West   | 2     | \$26.00  |                    |              |
| 3-Apr-16  | Blue  | South  | 8     | \$96.00  |                    |              |

# Sorting by value

1. Right-click any Sales value and choose Sort > Largest to Smallest.

A screenshot of a Microsoft Excel PivotTable context menu. The menu is open over a cell containing the value '\$6,414'. The menu items include:

- Copy
- Format Cells...
- Number Format...
- Refresh
- Sort** (highlighted)
- Remove "Sum of Sales"
- Summarize Values By >
- Show Values As >
- Value Field Settings...
- PivotTable Options...
- Hide Field List

Below the 'Sort' item, there are two options:

- $Z \downarrow$  Sort Smallest to Largest
- $Z \downarrow A \downarrow$  Sort Largest to Smallest (highlighted with a yellow arrow)

Handwritten text on the right side of the menu says: "Right-click, select Sort".

| Color              | Sum of Sales |
|--------------------|--------------|
| Blue               | \$7,464      |
| Green              | \$6,414      |
| Red                | \$           |
| Silver             | \$           |
| <b>Grand Total</b> | \$2          |

Excel now lists top-selling colors first. This sort order will be maintained when data changes, or when the pivot table is reconfigured.

I5 : 7464

|    | A | B                 | C     | D      | E     | F        | G | H                          | I | J |
|----|---|-------------------|-------|--------|-------|----------|---|----------------------------|---|---|
| 1  |   |                   |       |        |       |          |   |                            |   |   |
| 2  |   | Sample sales data |       |        |       |          |   | Sorted largest to smallest |   |   |
| 3  |   | Date              | Color | Region | Units | Sales    |   |                            |   |   |
| 4  |   | 3-Jan-16          | Red   | West   | 1     | \$11.00  |   |                            |   |   |
| 5  |   | 13-Jan-16         | Blue  | South  | 8     | \$96.00  |   |                            |   |   |
| 6  |   | 21-Jan-16         | Green | West   | 2     | \$26.00  |   |                            |   |   |
| 7  |   | 30-Jan-16         | Blue  | North  | 7     | \$84.00  |   |                            |   |   |
| 8  |   | 7-Feb-16          | Green | North  | 8     | \$104.00 |   |                            |   |   |
| 9  |   | 13-Feb-16         | Red   | South  | 2     | \$22.00  |   |                            |   |   |
| 10 |   | 21-Feb-16         | Blue  | East   | 5     | \$60.00  |   |                            |   |   |
| 11 |   | 1-Mar-16          | Green | West   | 2     | \$26.00  |   |                            |   |   |
| 12 |   | 13-Mar-16         | Blue  | East   | 8     | \$96.00  |   |                            |   |   |
| 13 |   | 23-Mar-16         | Blue  | North  | 7     | \$84.00  |   |                            |   |   |
| 14 |   | 28-Mar-16         | Green | West   | 2     | \$26.00  |   |                            |   |   |
| 15 |   | 3-Apr-16          | Blue  | South  | 8     | \$96.00  |   |                            |   |   |
| 16 |   |                   |       |        |       |          |   |                            |   |   |

Color Sum of Sales

|                    |                 |
|--------------------|-----------------|
| Blue               | \$7,464         |
| Silver             | \$6,970         |
| Green              | \$6,414         |
| Red                | \$5,508         |
| <b>Grand Total</b> | <b>\$26,356</b> |



# Refresh data

Pivot table data needs to be "refreshed" in order to bring in updates. To reinforce how this works, we'll make a big change to the source data and watch it flow into the pivot table.

1. Select cell F5 and change \$11.00 to \$2000.
2. Right-click anywhere in the pivot table and select "Refresh".

Sample sales data

| Date      | Color | Region | Units | Sales      |
|-----------|-------|--------|-------|------------|
| 3-Jan-16  | Red   | West   | 1     | \$2,000.00 |
| 13-Jan-16 | Blue  | South  | 8     | \$96.00    |
| 21-Jan-16 | Green | West   | 2     | \$26.00    |
| 30-Jan-16 | Blue  | North  | 7     | \$84.00    |
| 7-Feb-16  | Green | North  | 8     | \$104.00   |
| 13-Feb-16 | Red   | South  | 2     | \$22.00    |
| 21-Feb-16 | Blue  | East   | 5     | \$60.00    |
| 1-Mar-16  | Green | West   | 2     | \$26.00    |
| 13-Mar-16 | Blue  | East   | 8     | \$96.00    |
| 23-Mar-16 | Blue  | North  | 7     | \$84.00    |
| 28-Mar-16 | Green | West   | 2     | \$26.00    |
| 3-Apr-16  | Blue  | South  | 8     | \$96.00    |
| 12-Apr-16 | Green | South  | 1     | \$13.00    |
| 16-Apr-16 | Red   | East   | 8     | \$88.00    |
| 23-Apr-16 | Red   | West   | 6     | \$66.00    |
| 30-Apr-16 | Green | South  | 5     | \$65.00    |

Right-click, select "Refresh"

Color Sum of Sales

|              |         |
|--------------|---------|
| Blue         | \$7,464 |
| Silver       | \$6,970 |
| Green        |         |
| Red          |         |
| <b>Grand</b> |         |

EXCELJET

Notice "Red" is now the top selling color, and automatically moves to the top:

|  | A         | B     | C      | D     | E          | F | G | H | I | J |
|--|-----------|-------|--------|-------|------------|---|---|---|---|---|
| 1  |           |       |        |       |            |   |   |   |   |   |
| 2  |           |       |        |       |            |   |   |   |   |   |
| Sample sales data                          |           |       |        |       |            |   |   |   |   |   |
| Date Color Region Units Sales Sum of Sales |           |       |        |       |            |   |   |   |   |   |
| 4  | Date      | Color | Region | Units | Sales      |   |   |   |   |   |
| 5  | 3-Jan-16  | Red   | West   | 1     | \$2,000.00 |   |   |   |   |   |
| 6  | 13-Jan-16 | Blue  | South  | 8     | \$96.00    |   |   |   |   |   |
| 7  | 21-Jan-16 | Green | West   | 2     | \$26.00    |   |   |   |   |   |
| 8  | 30-Jan-16 | Blue  | North  | 7     | \$84.00    |   |   |   |   |   |
| 9  | 7-Feb-16  | Green | North  | 8     | \$104.00   |   |   |   |   |   |
| 10   | 13-Feb-16 | Red   | South  | 2     | \$22.00    |   |   |   |   |   |
| 11   | 21-Feb-16 | Blue  | East   | 5     | \$60.00    |   |   |   |   |   |
| 12   | 1-Mar-16  | Green | West   | 2     | \$26.00    |   |   |   |   |   |
| 13   | 13-Mar-16 | Blue  | East   | 8     | \$96.00    |   |   |   |   |   |
| 14   | 23-Mar-16 | Blue  | North  | 7     | \$84.00    |   |   |   |   |   |
| 15   | 28-Mar-16 | Green | West   | 2     | \$26.00    |   |   |   |   |   |
| 16   | 3-Apr-16  | Blue  | South  | 8     | \$96.00    |   |   |   |   |   |
| 17   | 12-Apr-16 | Green | South  | 1     | \$13.00    |   |   |   |   |   |
| 18   | 16-Apr-16 | Red   | East   | 8     | \$88.00    |   |   |   |   |   |

| Color              | Sum of Sales    |
|--------------------|-----------------|
| Red                | \$7,497         |
| Blue               | \$7,464         |
| Silver             | \$6,970         |
| Green              | \$6,414         |
| <b>Grand Total</b> | <b>\$28,345</b> |

*Red moves to the top as the best selling color*

3. Change F5 back to \$11.00 and refresh the pivot again.

*Note: changing F5 to \$2000 is not realistic, but it's a good way to force a change you can easily see in the pivot table. Try changing an existing color to something new, like "Gold" or "Black". When you refresh, you'll see the new color appear. You can use undo to go back to original data and pivot.*

# Second value field

You can add more than one field as a Value field.

1. Drag Units to the Value area to see Sales and Units together:

The screenshot shows a Microsoft Excel spreadsheet with a pivot table. The top ribbon has tabs for Home, Insert, Page Layout, Formulas, Data, etc., with 'Data' selected. The formula bar shows 'I5' and '7464'. The pivot table is titled 'Sample sales data' and contains the following data:

| Date      | Color | Region | Units | Sales    |
|-----------|-------|--------|-------|----------|
| 3-Jan-16  | Red   | West   | 1     | \$11.00  |
| 13-Jan-16 | Blue  | South  | 8     | \$96.00  |
| 21-Jan-16 | Green | West   | 2     | \$26.00  |
| 30-Jan-16 | Blue  | North  | 7     | \$84.00  |
| 7-Feb-16  | Green | North  | 8     | \$104.00 |
| 13-Feb-16 | Red   | South  | 2     | \$22.00  |
| 21-Feb-16 | Blue  | East   | 5     | \$60.00  |
| 1-Mar-16  | Green | West   | 2     | \$26.00  |
| 13-Mar-16 | Blue  | East   | 8     | \$96.00  |
| 23-Mar-16 | Blue  | North  | 7     | \$84.00  |
| 28-Mar-16 | Green | West   | 2     | \$26.00  |
| 3-Apr-16  | Blue  | South  | 8     | \$96.00  |
| 12-Apr-16 | Green | South  | 1     | \$13.00  |
| 16-Apr-16 | Red   | East   | 8     | \$88.00  |

To the right of the pivot table, there is a separate table showing the sum of sales and units by color:

| Color              | Sum of Sales    | Sum of Units |
|--------------------|-----------------|--------------|
| Blue               | \$7,464         | 608          |
| Silver             | \$6,970         | 473          |
| Green              | \$6,414         | 481          |
| Red                | \$5,508         | 486          |
| <b>Grand Total</b> | <b>\$26,356</b> | <b>2048</b>  |

A handwritten note 'Units added as a Value field' is written next to the second column of the pivot table.

# Percent of total

There are different ways to display values. One option is to show values as a percent of total. If you want to display the same field in different ways, add the field twice.

1. Remove the Units from the Values area
2. Add the Sales field (again) to the Values area.
3. Right-click the second instance and choose "% of grand total":

*Changing calculation to show percent of total*

| Color              | Sum of Sales    | Sum of Sales2 |
|--------------------|-----------------|---------------|
| Blue               | \$7,464         | 7464          |
| Silver             | \$6,970         | 69            |
| Green              | \$6,414         | 64            |
| Red                | \$5,508         | 55            |
| <b>Grand Total</b> | <b>\$26,356</b> | <b>263</b>    |

A context menu is open over the cell containing "7464". The menu options are:

- Copy
- Format Cells...
- Number Format...
- Refresh
- Sort >
- Remove "Sum of Sales2"
- Summarize Values By >
- Show Values As > (This option is highlighted with a yellow arrow.)
  - No Calculation
  - % of Grand Total (This option is checked with a yellow checkmark and has a yellow arrow pointing to it.)
  - % of Column Total
  - % of Row Total
  - % Of...
  - % of Parent Row Total
- Value Field Settings...
- PivotTable Options...
- Hide Field List

**EXCELJET**

The result is a breakdown by color along with a percent of total:

The screenshot shows a Microsoft Excel interface with a sample sales dataset and a pivot table analysis.

**Sample sales data:**

| Date      | Color | Region | Units | Sales    |
|-----------|-------|--------|-------|----------|
| 3-Jan-16  | Red   | West   | 1     | \$11.00  |
| 13-Jan-16 | Blue  | South  | 8     | \$96.00  |
| 21-Jan-16 | Green | West   | 2     | \$26.00  |
| 30-Jan-16 | Blue  | North  | 7     | \$84.00  |
| 7-Feb-16  | Green | North  | 8     | \$104.00 |
| 13-Feb-16 | Red   | South  | 2     | \$22.00  |
| 21-Feb-16 | Blue  | East   | 5     | \$60.00  |
| 1-Mar-16  | Green | West   | 2     | \$26.00  |
| 13-Mar-16 | Blue  | East   | 8     | \$96.00  |
| 23-Mar-16 | Blue  | North  | 7     | \$84.00  |
| 28-Mar-16 | Green | West   | 2     | \$26.00  |
| 3-Apr-16  | Blue  | South  | 8     | \$96.00  |
| 12-Apr-16 | Green | South  | 1     | \$13.00  |
| 16-Apr-16 | Red   | East   | 8     | \$88.00  |

**Pivot Table Analysis:**

| Color              | Sum of Sales    | Sum of Sales2 |
|--------------------|-----------------|---------------|
| Blue               | \$7,464         | 28.3%         |
| Silver             | \$6,970         | 26.4%         |
| Green              | \$6,414         | 24.3%         |
| Red                | \$5,508         | 20.9%         |
| <b>Grand Total</b> | <b>\$26,356</b> | <b>100.0%</b> |

*Second Sales field displays % of total*

**EXCELJET**

Note: the number format for percentage has also been adjusted to show 1 decimal.

# Group by date

Screenshot of the PivotTable Fields ribbon. Under "Choose fields to add to report", "Color" is selected and highlighted in green. Other fields listed are Date, Region, Units, and Sales. Under "Drag fields between areas below:", "Color" is in the Rows area, and "Sales" is in the Columns area. The "Σ Values" dropdown shows "Sum of Sales" and "Sum of Sales2". At the bottom, there are checkboxes for "Defer Layout Update" and "Update".

Pivot tables have a special feature to group dates into units like years, months, and quarters. This grouping can be customized.

1. Remove the second Sales field (Sales2).
2. Drag the Date field to the Columns area.
3. Right-click a date in the header area and choose "Group":

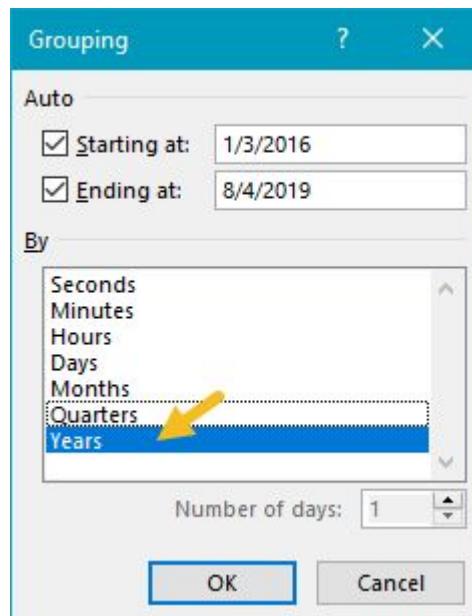
Screenshot of a PivotTable showing sales data grouped by year. The header row includes "Sum of Sales", "Years", "Quarters", "Date", and "Grand Total". A context menu is open over the "Date" column header, listing options: Copy, Format Cells..., Refresh, Sort, Filter, Subtotal "Years", Expand/Collapse, Group... (highlighted with a yellow arrow), Ungroup..., Move, Remove "Years", Show Details, Field Settings..., PivotTable Options..., and Hide Field List. The PivotTable data is as follows:

| Color       | 2016    | 2017    | 2018     | 2019 | 2020 | Grand Total |
|-------------|---------|---------|----------|------|------|-------------|
| Blue        | \$1,764 | \$2,000 | \$7,464  |      |      | \$7,464     |
| Silver      | \$1,001 | \$2,000 | \$6,970  |      |      | \$6,970     |
| Green       | \$682   | \$2,000 | \$6,414  |      |      | \$6,414     |
| Red         | \$3,447 | \$2,000 | \$5,508  |      |      | \$5,508     |
| Grand Total | \$3,447 | \$2,000 | \$26,356 |      |      | \$26,356    |

*Right-click a date, select Group*

**EXCELJET**

4. When the Group window appears, group by Years only (deselect Months and Quarters):



We now have a pivot table that groups sales by color and year:

G H I J K L M N

Date grouped by Years only

| Color              | Sum of Sales Date |                |                |                |  | Grand Total     |
|--------------------|-------------------|----------------|----------------|----------------|--|-----------------|
|                    | 2016              | 2017           | 2018           | 2019           |  |                 |
| Blue               | \$1,764           | \$2,136        | \$2,147        | \$1,417        |  | \$7,464         |
| Silver             |                   |                | \$3,565        | \$3,405        |  | \$6,970         |
| Green              | \$1,001           | \$2,431        | \$1,330        | \$1,652        |  | \$6,414         |
| Red                | \$682             | \$2,024        | \$1,986        | \$816          |  | \$5,508         |
| <b>Grand Total</b> | <b>\$3,447</b>    | <b>\$6,591</b> | <b>\$9,028</b> | <b>\$7,290</b> |  | <b>\$26,356</b> |

Notice there are no sales of Silver in 2016 and 2017. We can guess that Silver was introduced as a new color in 2018. Pivot tables often reveal patterns in data that are difficult to see otherwise.

PivotTable Fields

Choose fields to add to report:

Search

Date  
 Color  
 Region  
 Units  
 Sales

Drag fields between areas below:

| Filters | Columns |
|---------|---------|
|         | Date    |

| Rows  | Values       |
|-------|--------------|
| Color | Sum of Sales |

Defer Layout Update

## Two-way Pivot

Pivot tables can plot data in various two-dimensional arrangements.

1. Drag the Date field out of the columns area
2. Drag Region into the Columns area.

Excel builds a two-way pivot table that breaks down sales by color and region:

| G | H | I | J | K | L | M | N |
|---|---|---|---|---|---|---|---|
|---|---|---|---|---|---|---|---|

Two-way Pivot table, Color by Region

| Sum of Sales Region ▾ |  | East           | North          | South          | West           | Grand Total     |
|-----------------------|--|----------------|----------------|----------------|----------------|-----------------|
| Color                 |  |                |                |                |                |                 |
| Blue                  |  | \$1,121        | \$2,358        | \$1,710        | \$2,275        | \$7,464         |
| Silver                |  | \$697          | \$2,645        | \$1,149        | \$2,479        | \$6,970         |
| Green                 |  | \$1,198        | \$1,678        | \$1,059        | \$2,479        | \$6,414         |
| Red                   |  | \$714          | \$2,028        | \$1,126        | \$1,640        | \$5,508         |
| <b>Grand Total</b>    |  | <b>\$3,730</b> | <b>\$8,709</b> | <b>\$5,044</b> | <b>\$8,873</b> | <b>\$26,356</b> |

3. Swap Region and Color (i.e. drag Region to the Rows area and Color to the Columns area).

Excel builds another two-dimensional pivot table:

| G | H | I | J | K | L | M | N |
|---|---|---|---|---|---|---|---|
|---|---|---|---|---|---|---|---|

Two-way Pivot table, Region by Color

| Sum of Sales       |  | Color          |                |                |                |                 |
|--------------------|--|----------------|----------------|----------------|----------------|-----------------|
| Region             |  | Blue           | Silver         | Green          | Red            | Grand Total     |
| East               |  | \$1,121        | \$697          | \$1,198        | \$714          | \$3,730         |
| North              |  | \$2,358        | \$2,645        | \$1,678        | \$2,028        | \$8,709         |
| South              |  | \$1,710        | \$1,149        | \$1,059        | \$1,126        | \$5,044         |
| West               |  | \$2,275        | \$2,479        | \$2,479        | \$1,640        | \$8,873         |
| <b>Grand Total</b> |  | <b>\$7,464</b> | <b>\$6,970</b> | <b>\$6,414</b> | <b>\$5,508</b> | <b>\$26,356</b> |

Again notice total sales (\$26,356) is the same in *all pivot tables above*. Each table presents a different view of the *same data*, so they all sum to the *same total*.

The above example shows how quickly you can build different pivot tables from the same data. You can create [many other kinds of pivot tables](#), using all kinds of data.

## Key Pivot Table benefits

**Simplicity.** Basic pivot tables are very simple to set up and customize. There is no need to learn complicated formulas.

**Speed.** You can create a good-looking, useful report with a pivot table in minutes. Even if you are very good with formulas, [pivot tables are faster to set up and require much less effort.](#)

**Flexibility.** Unlike formulas, pivot tables don't lock you into a particular view of your data. You can quickly rearrange the pivot table to suit your needs. You can even [clone a pivot table](#) and build a separate view.

**Accuracy.** As long as a pivot table is set up correctly, you can rest assured results are accurate. In fact, a pivot table will often highlight problems in the data faster than any other tool.

**Formatting.** A Pivot table can apply automatically apply consistent number and style formatting, even as data changes.

**Updates.** Pivot tables are designed for on-going updates. If you base a pivot table on an Excel Table, the table resize as needed with new data. [All you need to do is click Refresh](#), and your pivot table will show you the latest.

**Filtering.** Pivot tables contain several tools for filtering data. Need to look at North America and Asia, but exclude Europe? A pivot table makes it simple.

**Charts.** Once you have a pivot table, you can easily [create a pivot chart](#).

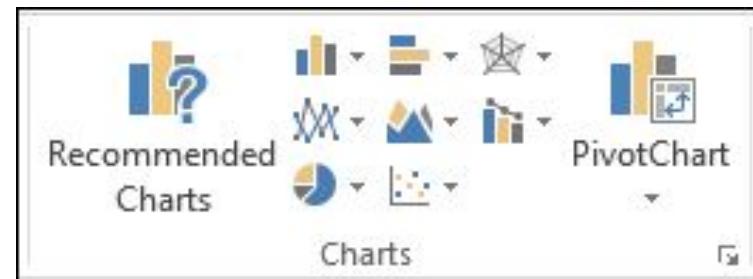
# Charts

In Microsoft Excel, charts are used to make a graphical representation of any set of data. A chart is a visual representation of data, in which the data is represented by symbols such as bars in a bar chart or lines in a line chart.

# Charts Group

The Charts group is formatted in such a way that –

- Types of charts are displayed.
- The subgroups are clubbed together.
- It helps you find a chart suitable to your data with the button Recommended Charts.

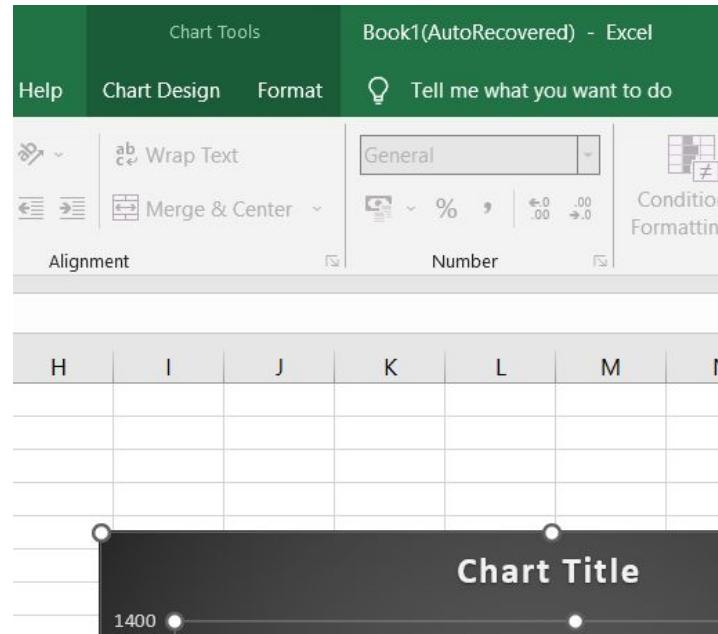


# CHART TOOLS

## Chart Tools

When you click on a chart, a new tab Chart Tools is displayed on the ribbon. There are two tabs under CHART TOOLS –

- DESIGN
- FORMAT



# Recommended Charts

The Recommended Charts command on the Insert tab helps you to create a chart that is just right for your data.

A screenshot of Microsoft Excel demonstrating the Recommended Charts feature. The ribbon at the top shows the 'INSERT' tab selected. Below the ribbon, a PivotTable is visible with data for book genres from 2008 to 2012. The 'CHARTS' icon in the 'CHARTS' group on the ribbon is highlighted with a red arrow. A second red arrow points to the text 'Recommended Charts' which is overlaid on the screenshot. The chart area below the table is currently empty.

| Genre       | 2008     | 2009     | 2010     | 2011     | 2012     |
|-------------|----------|----------|----------|----------|----------|
| Classics    | \$88,580 | \$49,225 | \$36,338 | \$50,017 | \$26,134 |
| Mystery     | \$78,970 | \$62,262 | \$48,640 | \$43,985 | \$73,428 |
| Romance     | \$34,236 | \$31,390 | \$79,022 | \$71,009 | \$81,474 |
| Sci-Fi & Fa | \$16,730 | \$19,730 | \$12,109 | \$11,355 | \$17,686 |
| Young Adr   | \$35,358 | \$42,685 | \$20,893 | \$16,065 | \$21,388 |

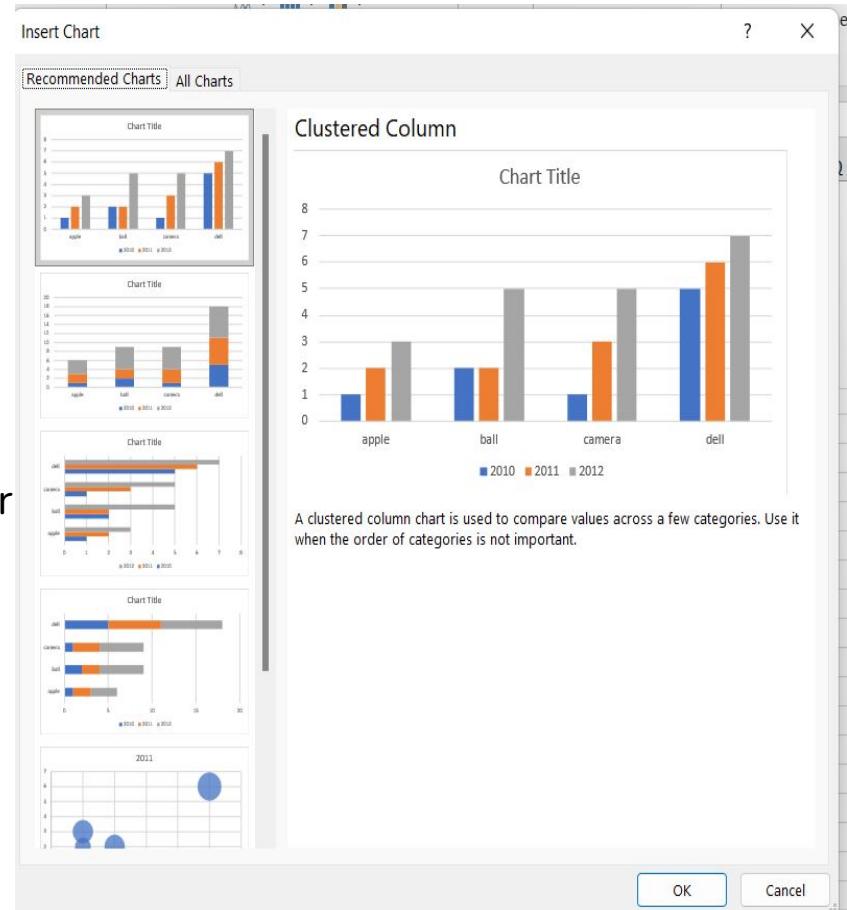
## Recommended Charts

## To use Recommended charts –

## Step 1 – Select the data.

## Step 2 – Click Recommended Charts.

A window displaying the charts that suit your data will be displayed.



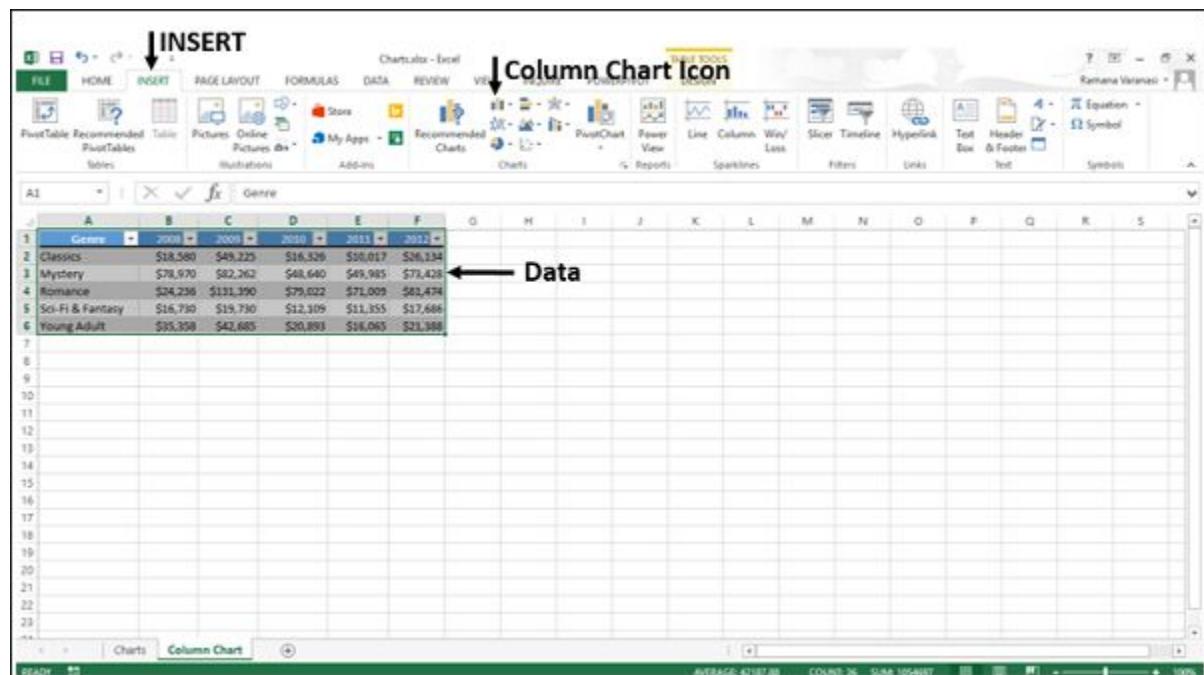
# Creating Charts with Insert Chart

To create charts using the Insert Chart tab, follow the steps given below.

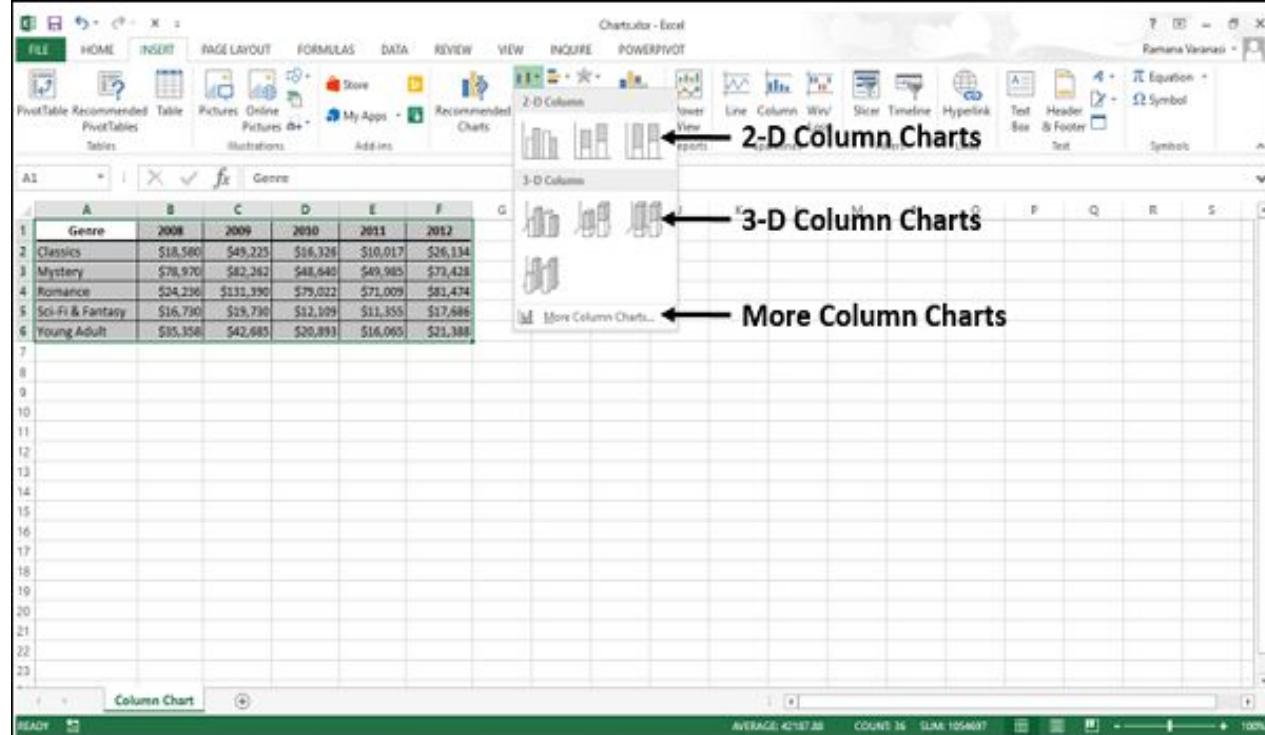
Step 1 – Select the data.

Step 2 – Click the Insert tab on the Ribbon.

Step 3 – Click the Insert Column Chart on the Ribbon.

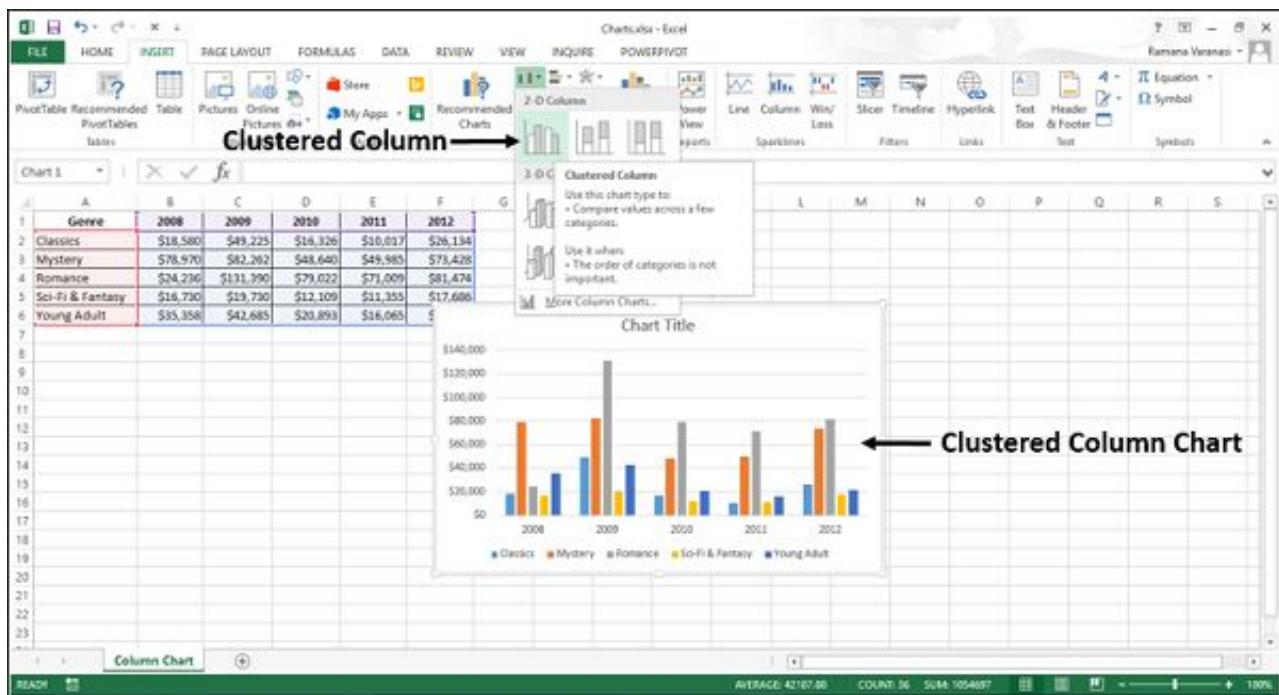


The 2-D column, 3-D Column chart options are displayed. Further, More Column Charts... option is also displayed.



Step 4 – Move through the Column Chart options to see the previews.

Step 5 – Click Clustered Column. The chart will be displayed in your worksheet



## Step 6 – Give a meaningful title to the chart by editing Chart Title.

The screenshot shows a Microsoft Excel spreadsheet titled "Chart.xlsx - Excel (Trial)". The ribbon at the top has tabs for FILE, HOME, INSERT, PAGE LAYOUT, FORMULAS, DATA, REVIEW, VIEW, INQUIRE, POWERPIVOT, DESIGN, and FORMAT. The DESIGN tab is selected. The chart area contains a grouped bar chart with five categories on the x-axis: Classics, Mystery, Romance, Sci-Fi & Fantasy, and Young Adult. Each category has five bars representing years from 2008 to 2012. The y-axis ranges from \$0 to \$140,000. A legend at the bottom identifies the years by color: 2008 (blue), 2009 (orange), 2010 (green), 2011 (yellow), and 2012 (red). The chart is set against a white background with light gray gridlines. To the right of the chart, the "CHART ELEMENTS" pane is open, showing checkboxes for various chart components. The "Chart Title" checkbox is checked, indicated by a green checkmark and a black arrow pointing to it. The data table on the left shows sales figures for each genre per year:

| Genre            | 2008     | 2009      | 2010     | 2011     | 2012     |
|------------------|----------|-----------|----------|----------|----------|
| Classics         | \$18,580 | \$49,225  | \$16,326 | \$30,017 | \$26,134 |
| Mystery          | \$78,970 | \$82,262  | \$48,640 | \$49,985 | \$73,428 |
| Romance          | \$24,236 | \$131,390 | \$79,022 | \$71,009 | \$81,474 |
| Sci-Fi & Fantasy | \$16,730 | \$19,730  | \$12,109 | \$11,355 | \$17,686 |
| Young Adult      | \$35,358 | \$42,685  | \$20,893 | \$16,063 | \$21,388 |

# Creating Charts with Quick Analysis

The screenshot shows a Microsoft Excel spreadsheet with a data table and a 'Quick Analysis' ribbon.

**Data Table:**

| item list | 2010 | 2011 | 2012 |
|-----------|------|------|------|
| apple     | 1    | 2    | 3    |
| ball      | 2    | 2    | 5    |
| camera    | 1    | 3    | 5    |
| dell      | 5    | 96   | 7    |

**Quick Analysis Ribbon:**

- Formatting
- Charts** (selected)
- Totals
- Tables
- Sparklines

**Chart Preview:**

A preview chart displays two stacked bars. The top bar represents the 'dell' item, and the bottom bar represents the 'camera' item. Both bars have four segments corresponding to the years 2010, 2011, 2012, and an unlabeled segment at the top.

**Chart Options:**

- Clustered Column
- Stacked Column
- Clustered Bar
- Stacked Bar
- Bubble
- More Charts

**Text at the Bottom:**

Recommended Charts help you visualize data.

Chaturax - Excel

FILE HOME INSERT PAGE LAYOUT FORMULAS DATA REVIEW VIEW POWERPIVOT DESIGN TABLE TOOLS

Cut Copy Format Painter Paste

Font Alignment Number Styles Cells

General Conditional Formatting Table Styles Insert Delete Format

AutoSum Filter Sort & Find & Filter Select Clear

Genre

|             | 2008     | 2009      | 2010     | 2011     | 2012     |
|-------------|----------|-----------|----------|----------|----------|
| Classics    | \$18,580 | \$49,225  | \$16,328 | \$10,017 | \$26,134 |
| Mystery     | \$78,970 | \$82,262  | \$48,640 | \$49,981 | \$71,428 |
| Romance     | \$24,298 | \$131,390 | \$79,022 | \$71,009 | \$81,474 |
| Sci-Fi & Fa | \$16,730 | \$19,730  | \$12,109 | \$11,355 | \$17,486 |
| Young Adt   | \$35,358 | \$42,683  | \$20,893 | \$16,065 | \$21,388 |

CHARTS

CHARTS TOTALS TABLES SPARKLINES

Clustered Clustered Stacked Stacked Clustered More

Recommended Charts

Sheet1 Sheet1 (2)

AVERAGE: 42187.08 COUNT: 36 SUM: 1054007

READY 100%

# Totals

| item list | 2010 | 2011 | 2012 |  |
|-----------|------|------|------|--|
| apple     | 1    | 2    | 3    |  |
| ball      | 2    | 2    | 5    |  |
| camera    | 1    | 3    | 5    |  |
| dell      | 5    | 96   | 7    |  |

A screenshot of Microsoft Excel showing a table of item counts across three years. A green selection box highlights the first four rows of the table. A context menu is open over the last row ('dell') with the 'Totals' tab selected. The 'Sum' option is highlighted, showing a preview of a bar chart with a total value of 108.

The context menu includes the following tabs: Formatting, Charts, Totals, Tables, and Sparklines. Below the tabs, there are six options: Sum, Average, Count, % Total, Running Total, and another Sum option. A note at the bottom states: "Formulas automatically calculate totals for you."

Formatting    Charts    Totals    Tables    Sparklines

Sum    Average    Count    % Total    Running Total    Sum

Formulas automatically calculate totals for you.

# Excel Charts Type

- Column Chart
- Line Chart
- Pie Chart
- Doughnut Chart
- Bar Chart
- Area Chart
- XY (Scatter) Chart
- Bubble Chart
- Stock Chart
- Surface Chart
- Radar Chart
- Combo Chart

# Column Chart

A Column Chart typically displays the categories along the horizontal (category) axis and values along the vertical (value) axis. To create a column chart, arrange the data in columns or rows on the worksheet.

A column chart has the following sub-types –

- Clustered Column.
- Stacked Column.
- 100% Stacked Column.
- 3-D Clustered Column.
- 3-D Stacked Column.
- 3-D 100% Stacked Column.
- 3-D Column.

Step 1 – Arrange the data in columns or rows on the worksheet.

Step 2 – Select the data.

Step 3 – On the INSERT tab, in the Charts group, click the Column chart icon on the Ribbon.

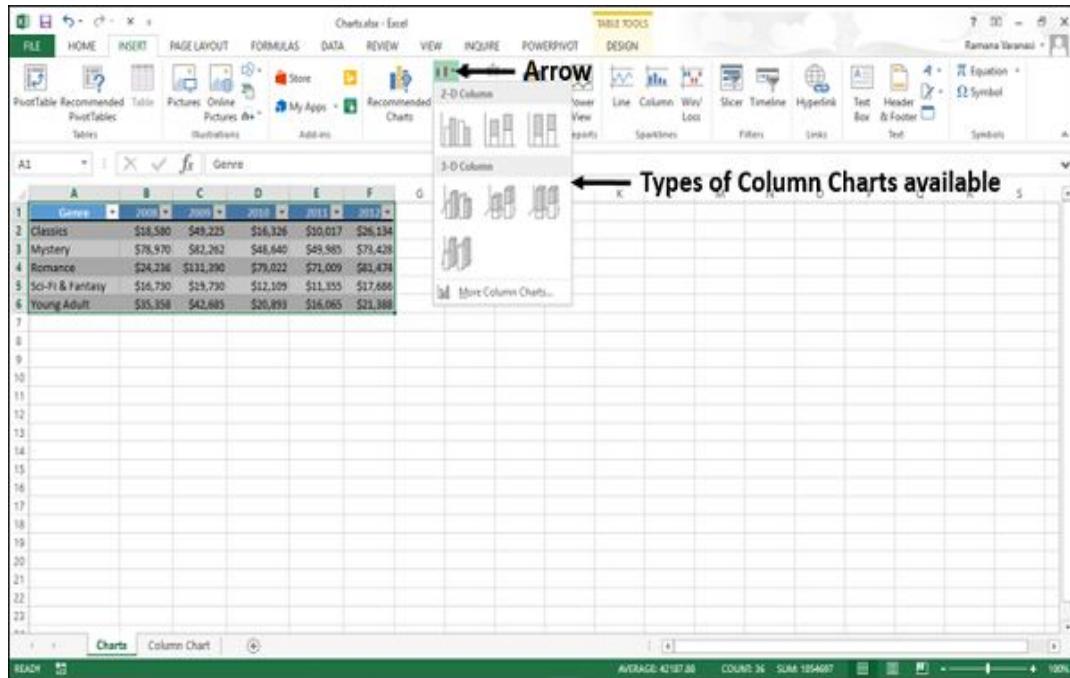
The screenshot shows a Microsoft Excel spreadsheet with data arranged in a table. The table has columns for 'Genre' and years from 2000 to 2012. The 'Genre' column is selected. The 'INSERT' tab is selected on the ribbon, and the 'Column Chart' icon is highlighted. An arrow points to the 'Data' range A1:F6. The ribbon also shows other tabs like HOME, PAGE LAYOUT, FORMULAS, DATA, REVIEW, and VBA.

| Genre            | 2000     | 2009      | 2010     | 2011     | 2012     |
|------------------|----------|-----------|----------|----------|----------|
| Classics         | \$18,580 | \$49,225  | \$16,526 | \$10,017 | \$26,134 |
| Mystery          | \$78,970 | \$82,362  | \$48,640 | \$49,985 | \$71,428 |
| Romance          | \$24,236 | \$131,390 | \$79,032 | \$71,009 | \$81,474 |
| Sci-Fi & Fantasy | \$16,730 | \$19,730  | \$12,109 | \$11,355 | \$17,686 |
| Young Adult      | \$35,358 | \$42,685  | \$28,893 | \$16,065 | \$23,388 |

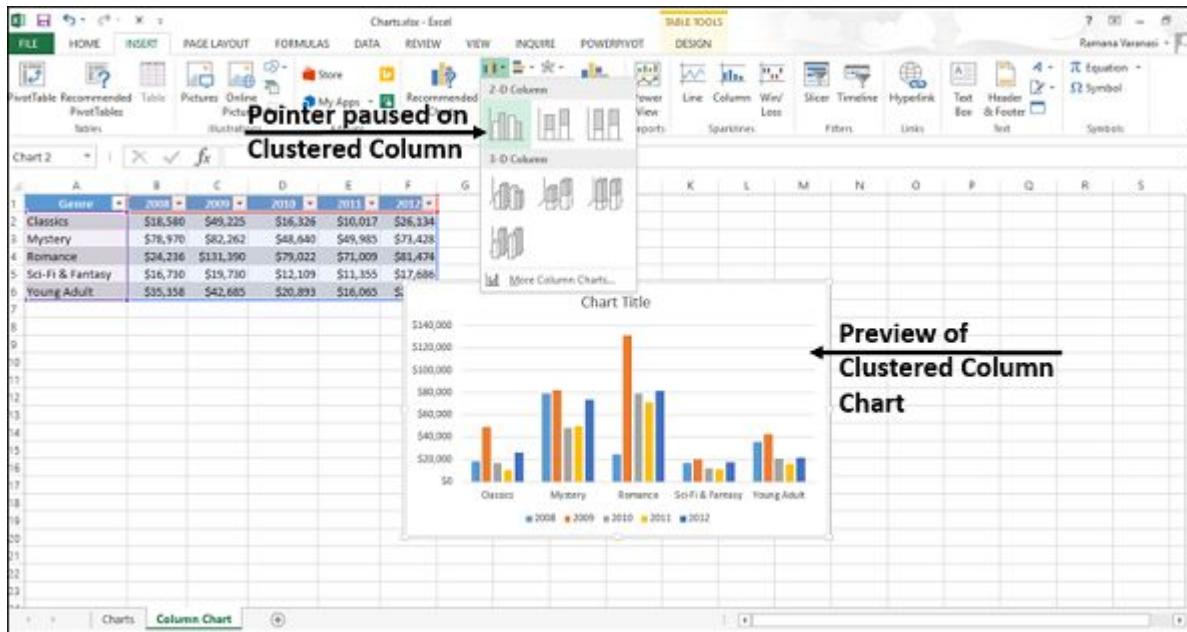
You will see the different options available for Column Charts.

A Column Chart has the following sub-types –

- 2-D Column Charts
  - Clustered Column
  - Stacked Column
  - 100% Stacked Column
- 3-D Column Charts
  - 3-D Clustered Column
  - 3-D Stacked Column
  - 3-D 100% Stacked Column
  - 3-D Column



Step 4 – Point your mouse on each of the icons. You will see a preview of the chart type.



Step 5 – Double-click the chart type that suits your data.

## Clustered Column and 3-D Clustered Column

These chart types are useful to compare the values across a few categories, when the order of the categories is not important.

Remember that –

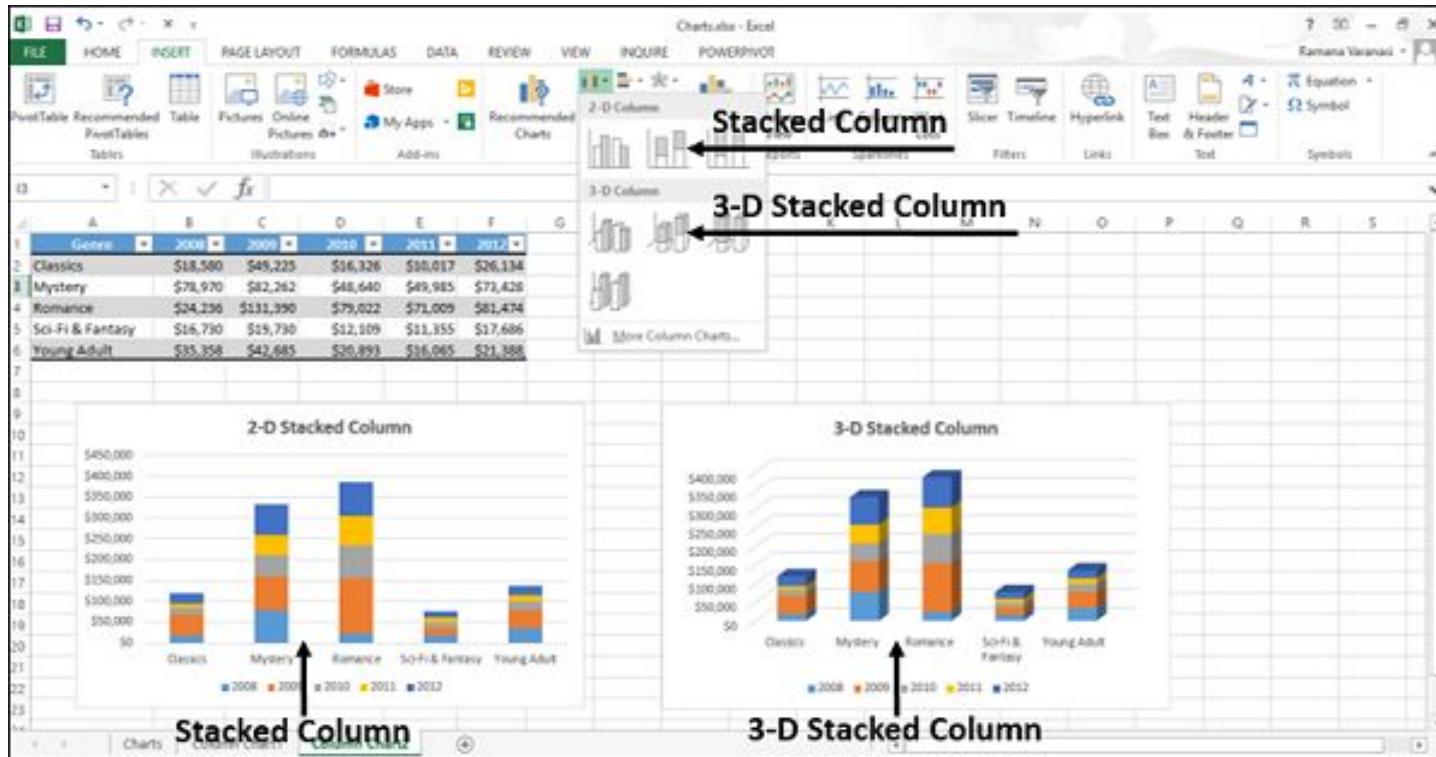
- A Clustered Column chart shows values in 2-D rectangular columns.
- A 3-D Clustered Column chart shows Columns in 3-D perspective, but it does not use a third value axis (depth axis).



You can use Clustered Column charts when you have multiple data series with categories that represent –

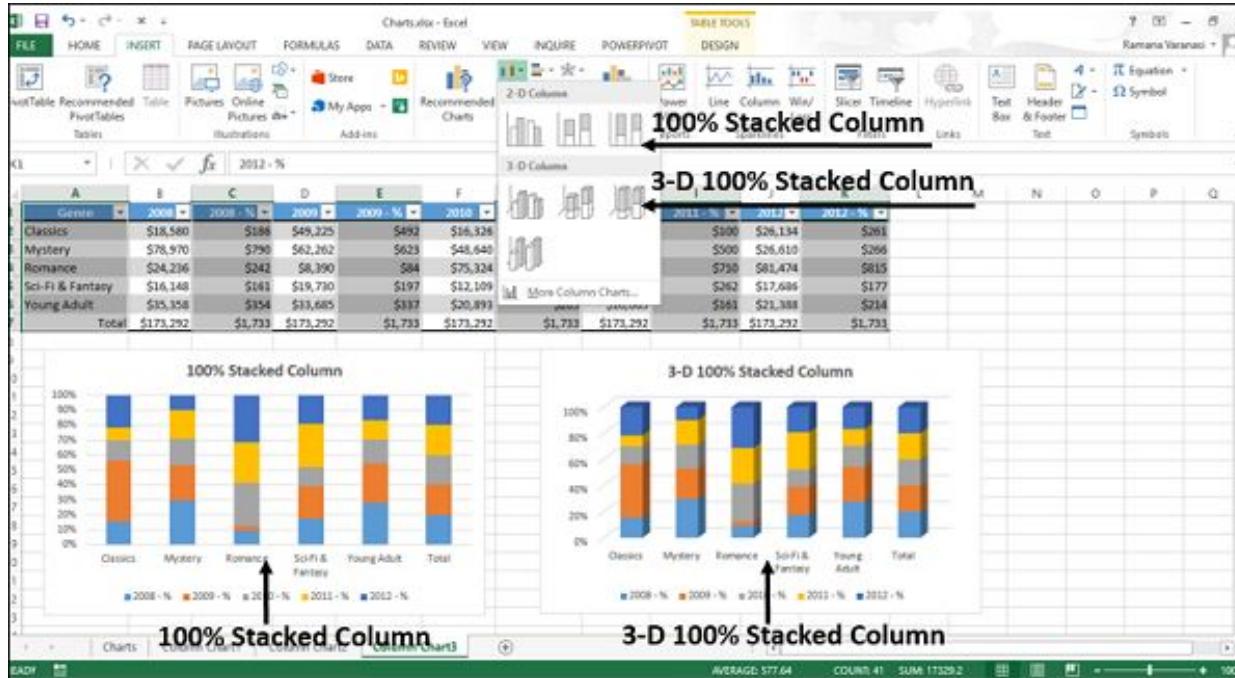
- Ranges of values (e.g. item counts).
- Specific scale arrangements (e.g. a Likert scale with entries like Strongly agree, Agree, Neutral, Disagree, Strongly disagree).
- Names that are not in any specific order (e.g. item names, geographic names, or the names of people).

# Stacked Column and 3-D Stacked Column

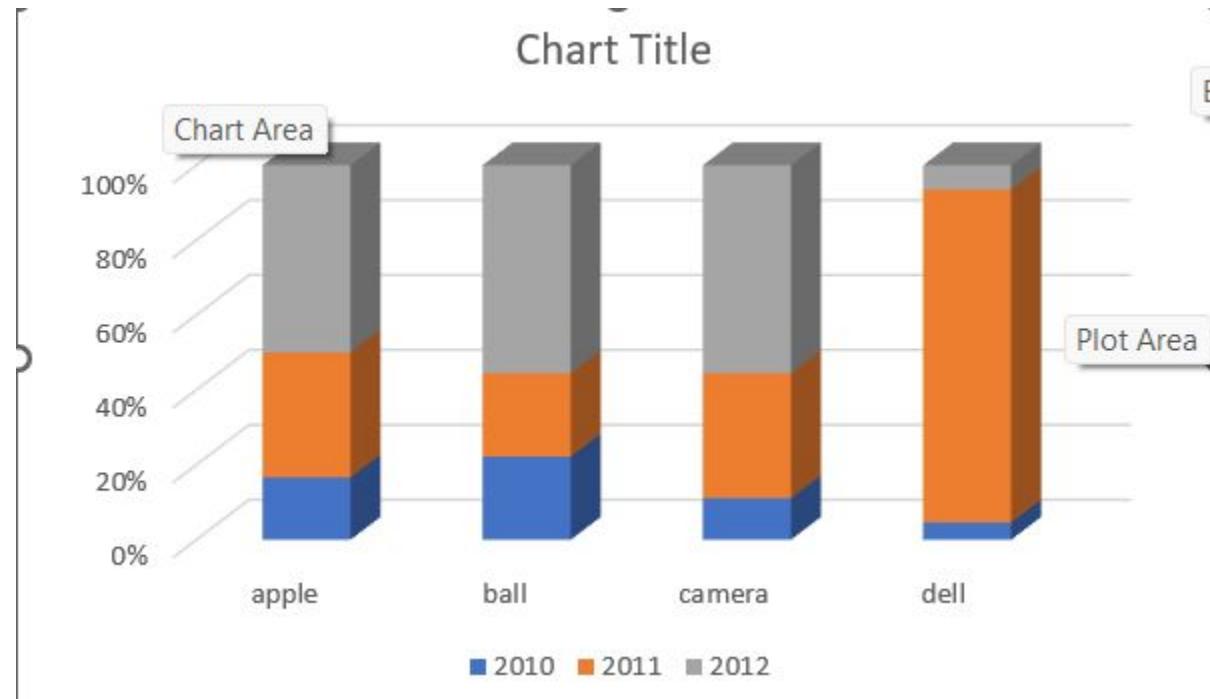


A Stacked Column chart displays values in 2-D vertical stacked rectangles. A 3-D Stacked Column chart displays the data by using a 3-D perspective, but it does not use a third value axis (depth axis).

# 100% Stacked Column and 3-D 100% Stacked Column

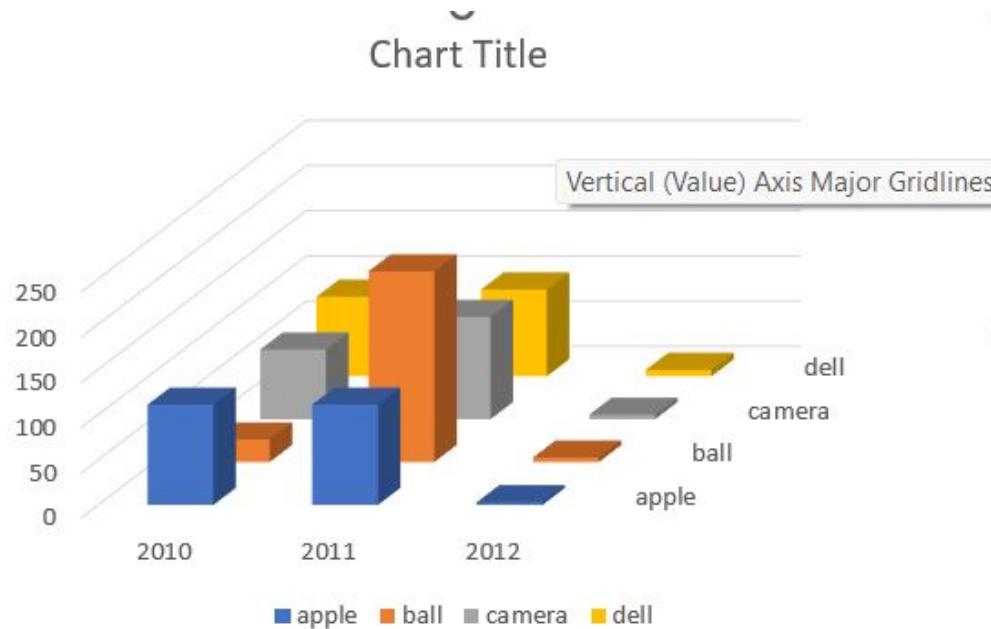


A 100% Stacked Column chart shows values in 2-D columns that are stacked to represent 100%. A 3-D 100% Stacked Column chart shows the columns using a 3-D perspective, but it does use a third value axis (depth axis).



# 3-D Column

3-D Column charts use three axes that you can modify (a horizontal axis, a vertical axis, and a depth axis), and they compare data points along the horizontal and the depth axes.



## Line Chart

Line charts can show continuous data over time on an evenly scaled Axis. Therefore, they are ideal for showing trends in data at equal intervals, such as days, months, quarters or years.

In a Line chart –

- Category data is distributed evenly along the horizontal axis.
- Value data is distributed evenly along the vertical axis.

Follow the steps given below to insert a Line chart in your worksheet.

Step 1 – Arrange the data in columns or rows on the worksheet.

Step 2 – Select the data.

Step 3 – On the INSERT tab, in the Charts group, click the Line chart icon on the Ribbon.

Step 4 – Point your mouse on each of the icons. A preview of that line type will be shown on the worksheet.

Step 5 – Double-click the chart type that suits your data.

- 2-D Line charts
  - Line
  - 100% Stacked Line
  - Line with Markers
  - Stacked Line with Markers
  - 100% Stacked Line with Markers
- 3-D Line charts
  - 3-D Line

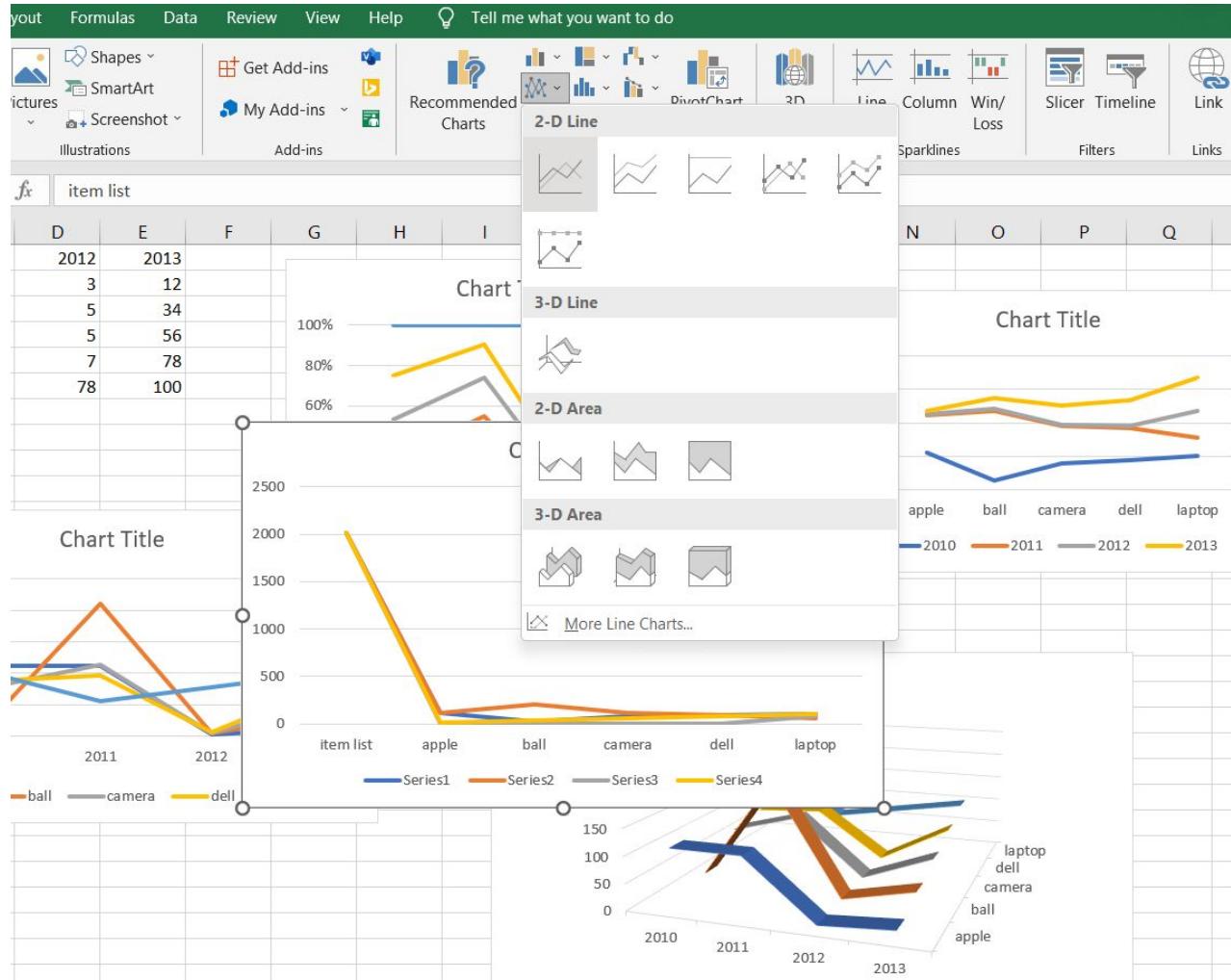
**Arrow**

**Types of Line Charts available**

The screenshot shows a Microsoft Excel spreadsheet titled "Chart.xlsx - Excel (Trial)". The data is organized into a table with columns for genre (Genre) and years (2008, 2009, 2010, 2011, 2012). The "Line" chart type dropdown menu is open, displaying various chart styles. Arrows point from the text labels to the corresponding chart icons in the menu.

| Genre            | 2008     | 2009      | 2010     | 2011     | 2012     |
|------------------|----------|-----------|----------|----------|----------|
| Classics         | \$18,500 | \$49,225  | \$16,326 | \$10,017 | \$26,334 |
| Mystery          | \$78,570 | \$82,262  | \$48,640 | \$49,985 | \$71,428 |
| Romance          | \$24,236 | \$131,390 | \$79,622 | \$71,009 | \$81,474 |
| Sci-Fi & Fantasy | \$16,730 | \$19,730  | \$12,109 | \$11,355 | \$17,686 |
| Young Adult      | \$33,358 | \$42,685  | \$20,893 | \$16,063 | \$21,398 |

Line Chart1



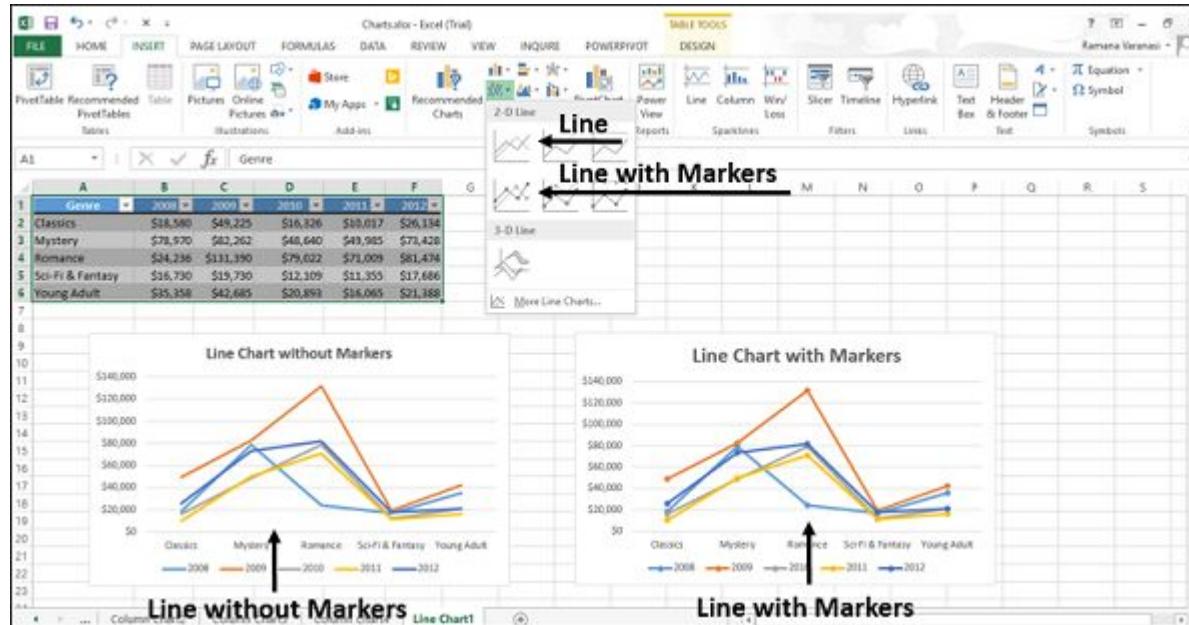
# Line and Line with Markers

Line charts indicate individual data values. Line charts work best when you have multiple data series in your chart.

Line charts can show trends over –

- Time (days, months, quarters or years), or
- Evenly spaced Categories.

# Line and Line with Markers



You can use a Line chart without markers when –

- The order of categories is important.
- There are many categories or if the values are approximate.

You can use a Line chart with Markers when –

- The order of categories is important.
- There are only a few categories.

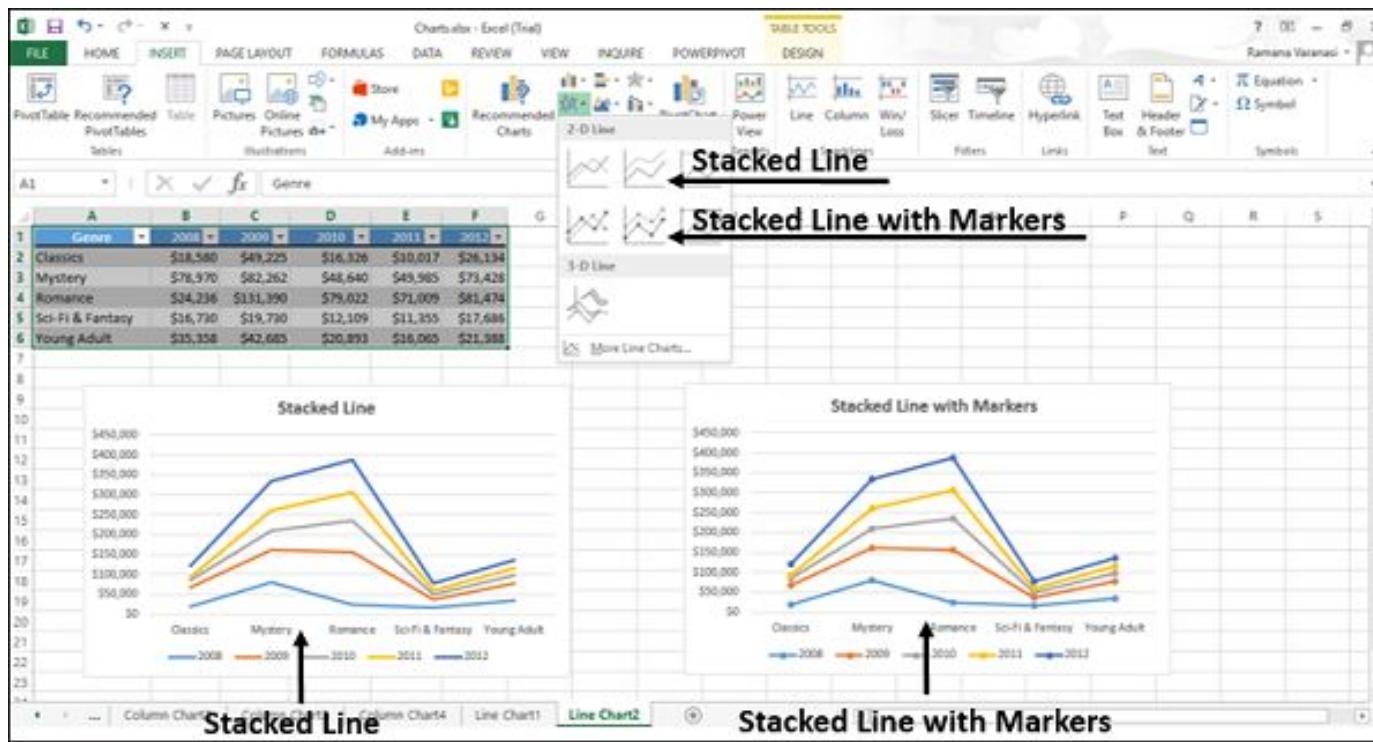
# Stacked Line and Stacked Line with Markers

Stacked Line charts indicate individual data values. Stacked Line Charts can show the trend of the contribution of each value over –

- Time, or
- Evenly spaced Categories.

Stacked Line charts can be with or without markers.

# Stacked Line and Stacked Line with Markers



## Stacked Line and Stacked Line with Markers

You can use a stacked line chart without markers when there are many categories or if the values are approximate. You can use a stacked line chart with markers when there are only a few categories.

Reading Stacked Line charts can be difficult as –

- They sum the data, which might not be the result you want.
- It might not be easy to see that the lines are stacked.

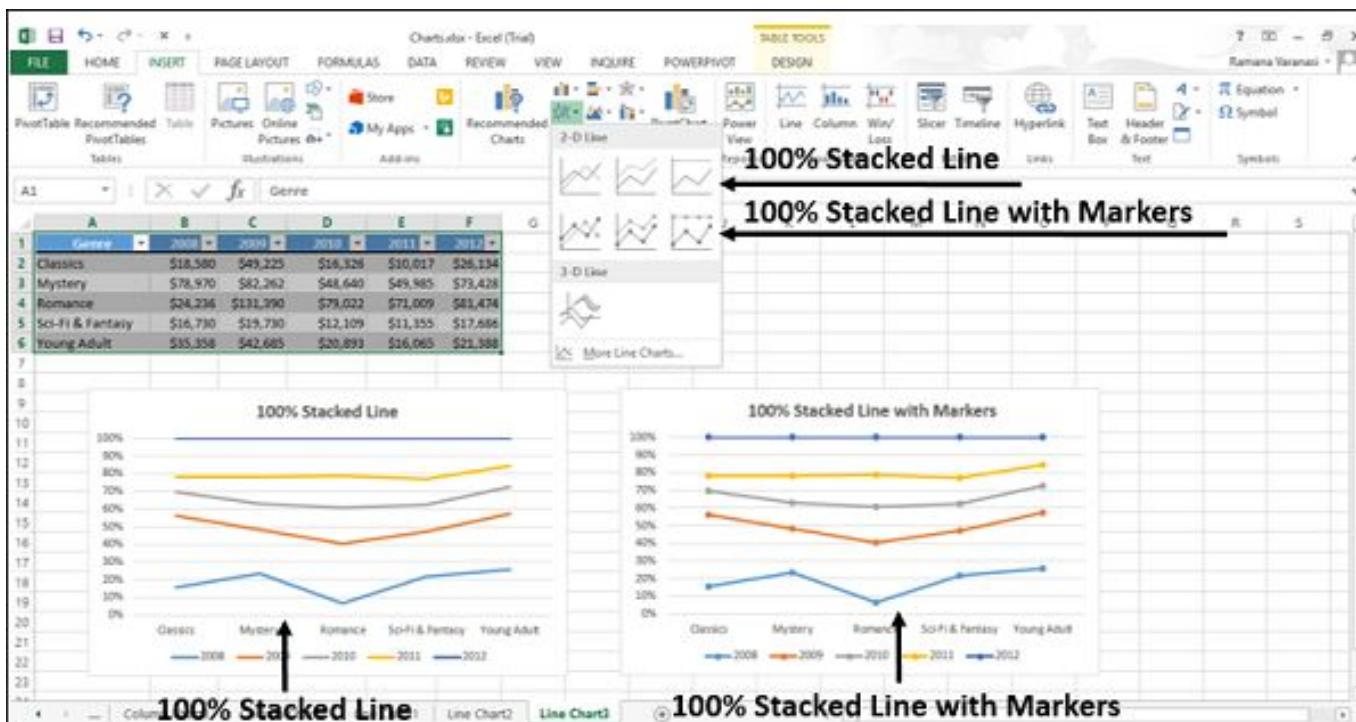
To overcome the difficulties, you can use a Stacked Area chart instead.

## 100% Stacked Line and 100% Stacked Line with Markers

100% Stacked Line charts indicate individual data values. 100% Stacked Line charts can show the trend of the percentage of each value over –

- Time, or
- Evenly spaced Categories

100% Stacked Line chart can be with or without Markers.



## 3-D Line

3-D Line charts show each row or column of data as a 3-D Ribbon. 3-D Line charts can show trends over –

- Time (days, months, quarters or years), or
- Categories.

A 3-D Line chart has horizontal, vertical, and depth axes that you can change. The third axis can show some lines in front of others.

# Pie Chart

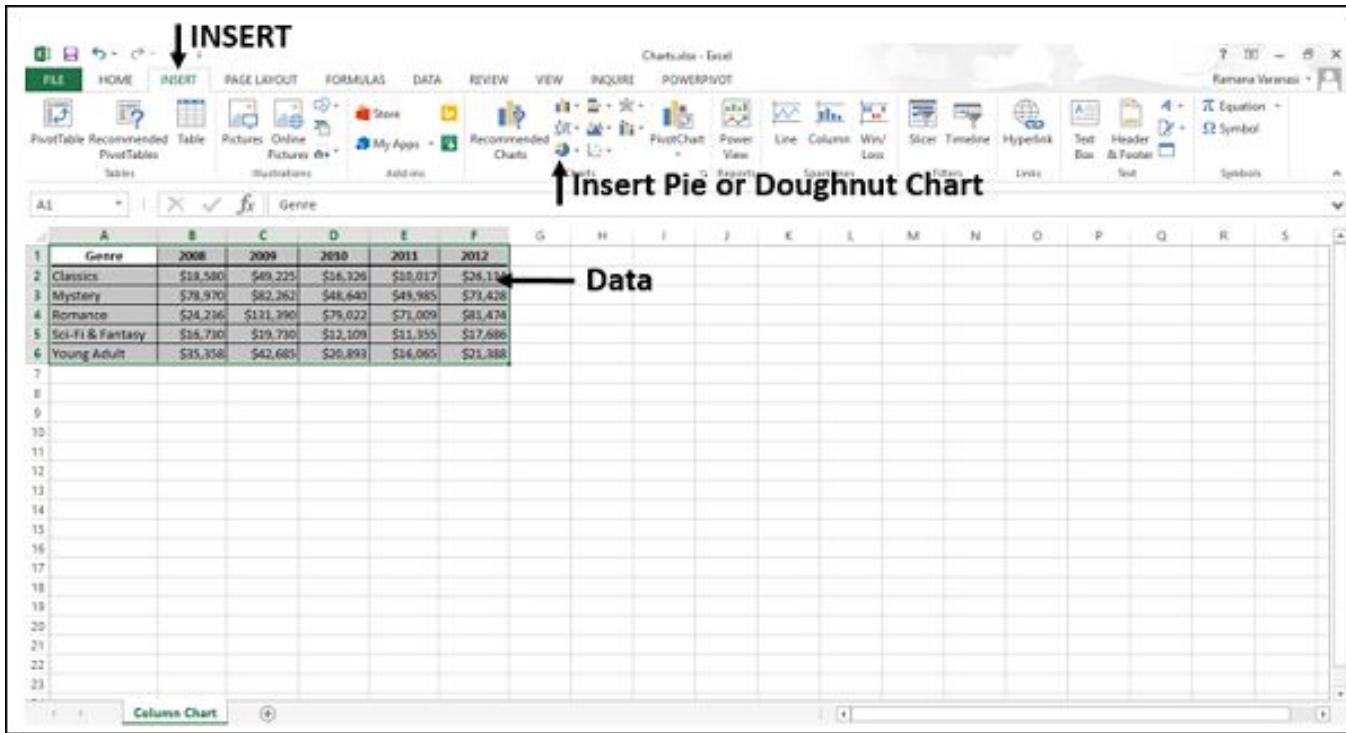
Pie charts show the size of the items in one data series, proportional to the sum of the items. The data points in a Pie chart are shown as a percentage of the whole Pie.

Follow the steps given below to insert a pie chart in your worksheet.

Step 1 – Arrange the data in columns or rows on the worksheet.

Step 2 – Select the data.

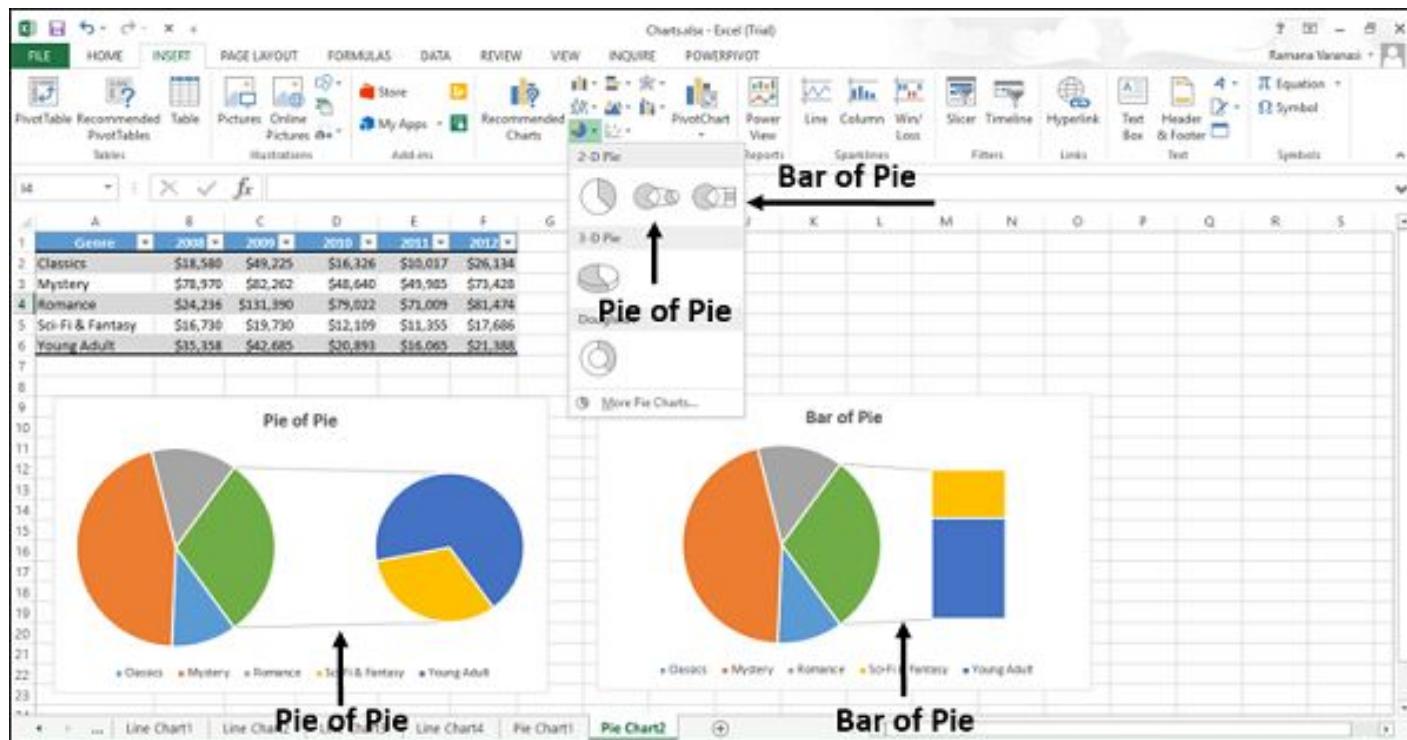
Step 3 – On the INSERT tab, in the Charts group, click the Pie chart icon on the Ribbon.



The screenshot shows a Microsoft Excel spreadsheet with a table of book sales data. The table has columns for Genre (Classics, Mystery, Romance, Sci-Fi & Fantasy, Young Adult) and years (2008-2012). The data shows a general decline in sales over time for most genres. Two pie charts are displayed below the table: a 2-D pie chart on the left and a 3-D pie chart on the right. Both charts use the same color scheme: orange for Classics, blue for Mystery, grey for Romance, yellow for Sci-Fi & Fantasy, and green for Young Adult. The legends for both charts are located directly beneath their respective charts.

| Genre            | 2008     | 2009      | 2010     | 2011     | 2012     |
|------------------|----------|-----------|----------|----------|----------|
| Classics         | \$18,580 | \$49,225  | \$16,326 | \$10,617 | \$26,334 |
| Mystery          | \$78,570 | \$42,262  | \$46,640 | \$49,985 | \$73,428 |
| Romance          | \$24,336 | \$131,390 | \$79,022 | \$71,009 | \$81,474 |
| Sci-Fi & Fantasy | \$16,730 | \$19,730  | \$12,109 | \$11,355 | \$17,686 |
| Young Adult      | \$35,308 | \$42,683  | \$20,893 | \$16,065 | \$21,388 |

Pie of Pie or Bar of Pie charts show Pie charts with smaller values pulled out into a secondary Pie or Stacked Bar chart, which makes them easier to distinguish.



# Doughnut Chart

Doughnut charts show the size of items in a data series, proportional to the sum of the items. The doughnut chart is similar to a pie chart, but it can contain more than one data series.

Step 1 – Arrange the data in columns or rows on the worksheet.

Step 2 – Select the data.

Step 3 – On the INSERT tab, in the Charts group, click the Pie chart icon on the Ribbon. It is used to insert a Doughnut chart also.

Step 4 – Point your mouse on the Doughnut icon. A preview of that chart type will be shown on the worksheet.

# Doughnut charts

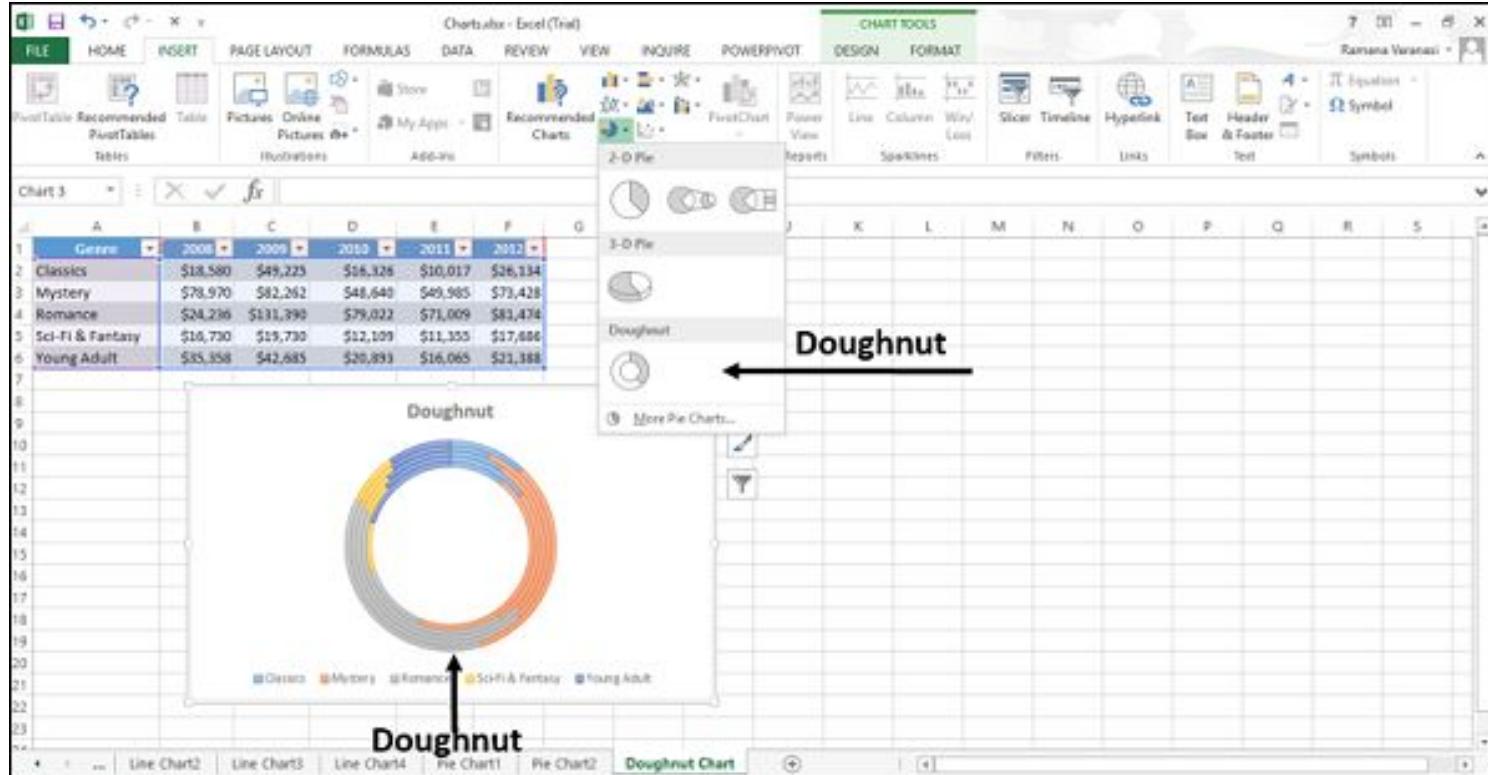
The screenshot shows a Microsoft Excel window titled "Chart.xlsx - Excel (Trial)". The ribbon is visible at the top with tabs: FILE, HOME, INSERT, PAGE LAYOUT, FORMULAS, DATA, REVIEW, VIEW, DESIGN, POWERPIVOT, and CHART TOOLS. The CHART TOOLS tab is selected, showing sub-tabs: PICTURES, FORMATTING, and DESIGN. The DESIGN tab has its own ribbon with options: 2-D Pie, 3-D Pie, Doughnut, and More Pie Charts... A black arrow points from the text "Arrow" to the 2-D Pie icon.

**Types of Doughnut Charts available**

| Genre            | 2008     | 2009      | 2010     | 2011     | 2012     |
|------------------|----------|-----------|----------|----------|----------|
| Classics         | \$18,580 | \$49,225  | \$16,326 | \$10,017 | \$26,134 |
| Mystery          | \$18,970 | \$82,262  | \$48,640 | \$49,585 | \$73,428 |
| Romance          | \$24,236 | \$131,190 | \$79,022 | \$71,009 | \$81,474 |
| Sci-Fi & Fantasy | \$16,750 | \$19,730  | \$12,109 | \$11,355 | \$17,686 |
| Young Adult      | \$15,358 | \$42,685  | \$20,893 | \$16,065 | \$21,388 |

Consider using a Doughnut chart when –

- You have more than one data series.
- None of the values in your data are negative.
- Almost none of the values in your data series are zero values.
- You have no more than seven categories, all of which represent parts of the whole pie.
- Doughnut Charts show data in rings, where each ring represents a data series. If percentages are shown in data labels, each ring will total to 100%.



# Bar Chart

Bar charts illustrate the comparisons among individual items. A Bar chart typically displays categories along the vertical (category) axis and values along the horizontal (value) axis.

Follow the steps given below to use a Bar chart.

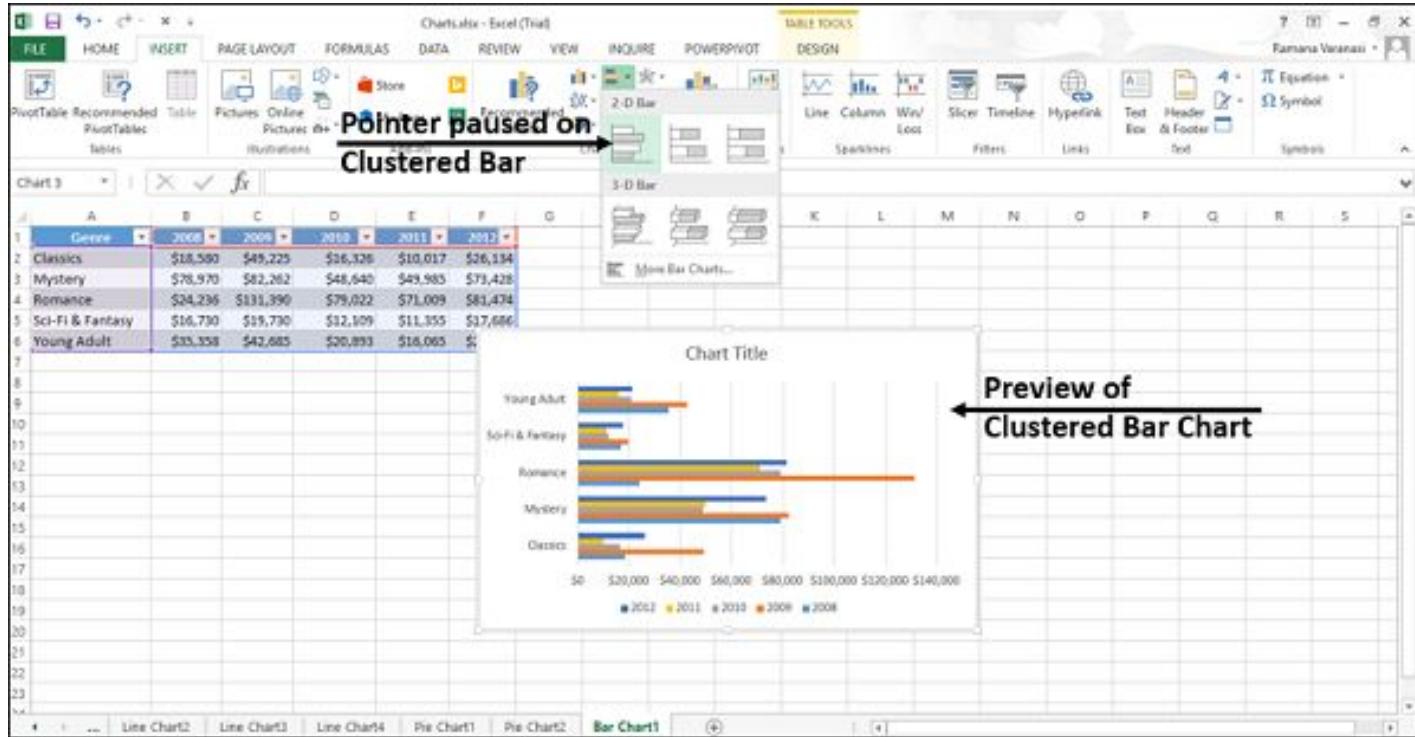
Step 1 – Arrange the data in columns or rows on the worksheet.

Step 2 – Select the data.

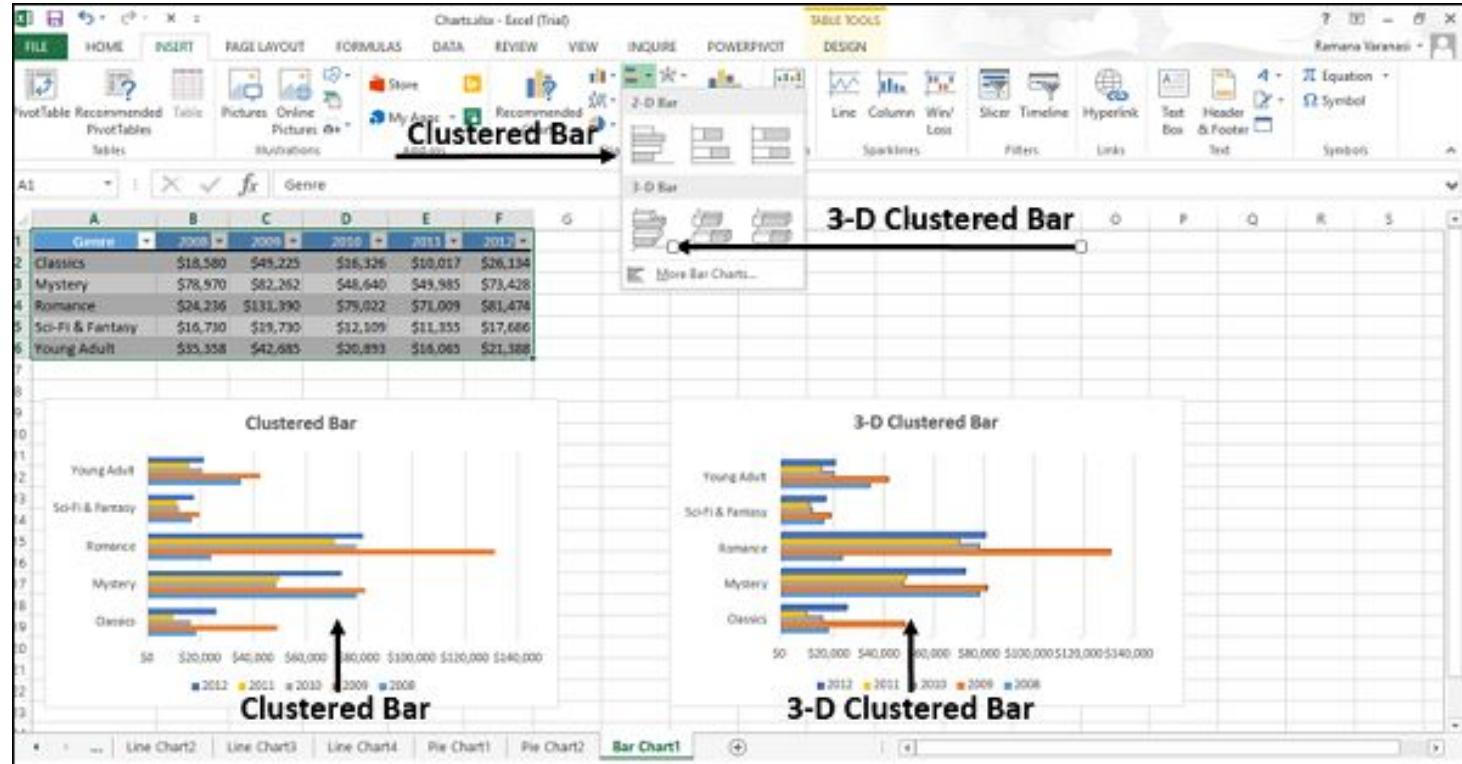
Step 3 – On the INSERT tab, in the Charts group, click the Bar chart icon on the Ribbon.

Step 4 – Point the mouse on each of the icons. A preview of that chart type will be shown on the worksheet.

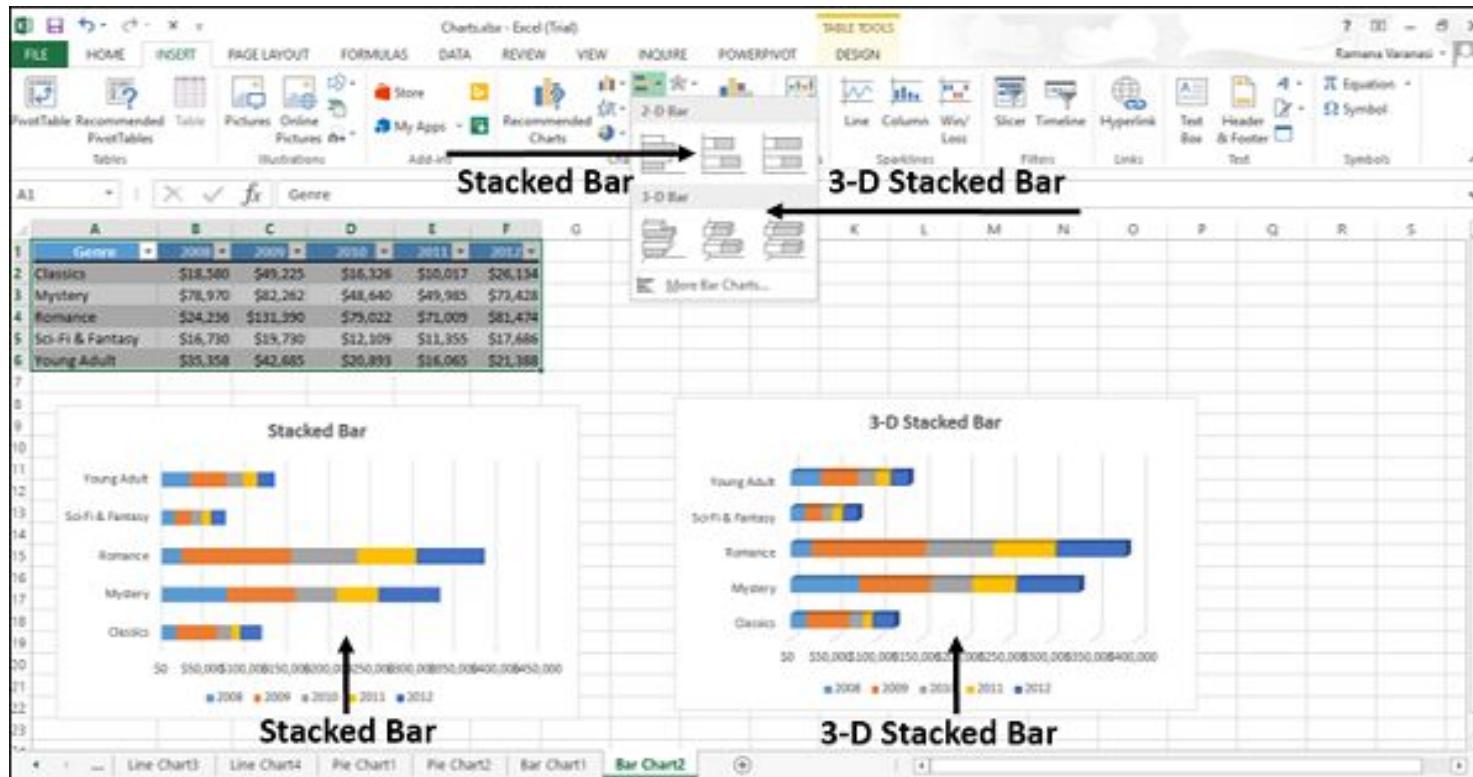
Step 5 – Double-click the chart type that suits your data.



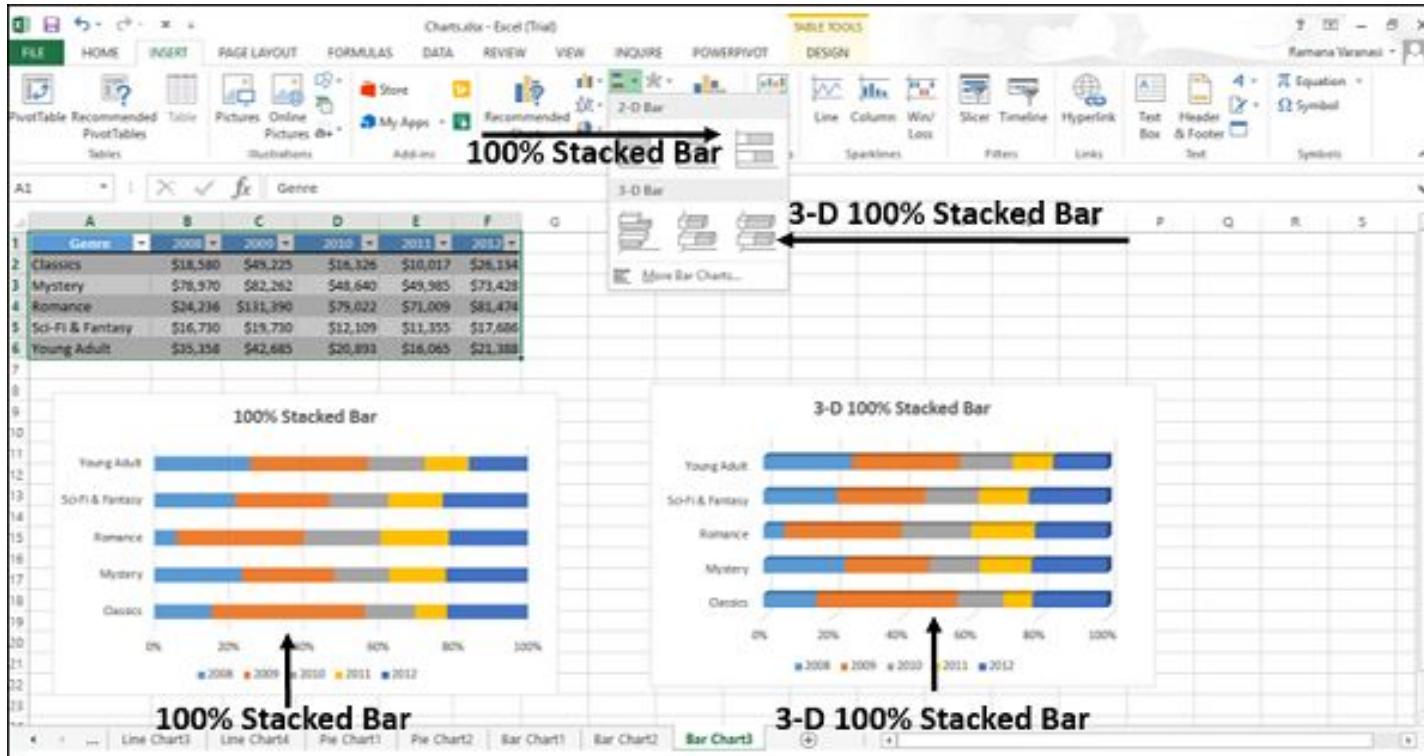
# Clustered Bar and 3-D Clustered Bar



## Stacked Bar and 3-D Stacked Bar



## 100% Stacked Bar and 3-D 100% Stacked Bar



# Area Chart

rea charts can be used to plot change over time (years, months and days) or categories and draw attention to the total value across a trend. By showing the sum of the plotted values, an Area chart also shows the relationship of parts to a whole.

You can use Area charts to highlight the magnitude of change over time.

Step 1 – Arrange the data in columns or rows on the worksheet.

Step 2 – Select the data.

Step 3 – On the INSERT tab, in the Charts group, click the Area chart icon on the Ribbon.

Step 4 – Point your mouse on each of the icons. A preview of that chart type will be shown on the worksheet.

Step 5 – Double-Click the chart type that suits your data. In this chapter, you will understand when each of the Area Chart Types is useful.

**Arrow**

**Types of Area Charts available**

The screenshot shows a Microsoft Excel window titled 'Chart.xlsx - Excel (Trial)'. The ribbon is visible at the top with tabs: FILE, HOME, INSERT, PAGE LAYOUT, FORMULAS, DATA, REVIEW, VIEW, INSURE, POWERPWT, and TABLE TOOLS. The 'TABLE TOOLS' tab is selected, indicated by a yellow background. The 'DESIGN' tab is active under 'TABLE TOOLS'. On the far right of the ribbon, there is a search bar with the text 'Rama Venanai' and a small icon. Below the ribbon, the 'Insert' tab is selected, showing icons for PivotTable, Recommended PivotTables, Table, Pictures, Online Pictures, My Apps, Recommended Charts, and Chart. A dropdown menu for 'Area' charts is open, showing two sections: '2-D Area' and '3-D Area', each with three preview icons. An arrow points from the text 'Types of Area Charts available' to the open dropdown menu. The main workspace contains a table with data for book genres and years, and a chart area below it.

| Genre            | 2008     | 2009      | 2010     | 2011     | 2012     |
|------------------|----------|-----------|----------|----------|----------|
| Classics         | \$18,580 | \$49,225  | \$16,326 | \$50,017 | \$26,134 |
| Mystery          | \$78,970 | \$82,262  | \$48,640 | \$49,385 | \$73,428 |
| Romance          | \$24,236 | \$131,390 | \$79,022 | \$71,009 | \$81,474 |
| Sci-Fi & Fantasy | \$16,730 | \$19,730  | \$12,109 | \$11,355 | \$17,686 |
| Young Adult      | \$25,358 | \$42,685  | \$20,893 | \$26,065 | \$21,388 |

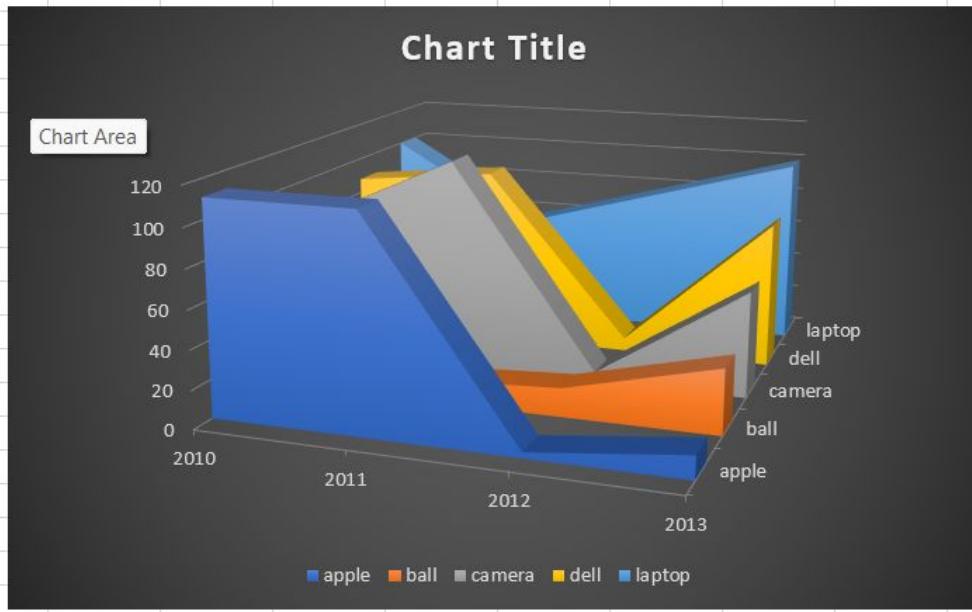
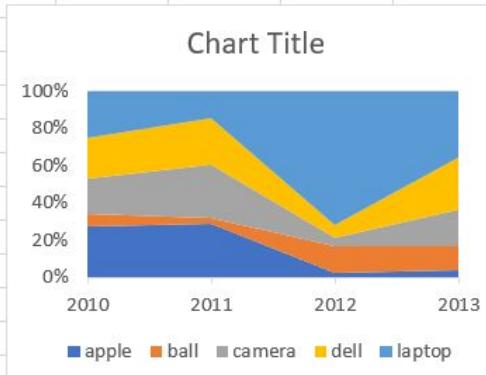
These chart types are useful to show the trend of values over time or other category data.

An Area chart shows the values in 2-D format. A 3-D Area chart shows values in 3-D format. 3-D Area charts use three axes (horizontal, vertical, and depth) that you can change.

You can use Area charts –

- When the category order is important.
- To highlight the magnitude of change over time.

|        |     |     |    |     |
|--------|-----|-----|----|-----|
| camera | 77  | 113 | 5  | 56  |
| ball   | 88  | 96  | 7  | 78  |
| laptop | 100 | 56  | 78 | 100 |



## Excel Charts - Scatter (X Y) Chart

Scatter (X Y) charts are typically used for showing and comparing numeric values, like scientific, statistical, and engineering data.

A Scatter Chart has two value axes –

- Horizontal (x) value axis
- Vertical (y) value axis

It combines x and y values into single data points and shows them in irregular intervals, or clusters.

Consider using a Scatter chart when –

- You want to change the scale of the horizontal axis.
- You want to make that axis a logarithmic scale.
- Values for horizontal axis are not evenly spaced.
- There are many data points on the horizontal axis.
- You want to adjust the independent axis scales of a scatter chart to reveal more information about the data that includes pairs or grouped sets of values.
- You want to show similarities between large sets of data instead of the differences between the data points.
- You want to compare many data points regardless of the time.
  - The more data that you include in a Scatter chart, the better the comparisons.

Step 1 – Arrange the data in columns or rows on the worksheet.

Step 2 – Place the x values in one row or column, and then enter the corresponding y values in the adjacent rows or columns.

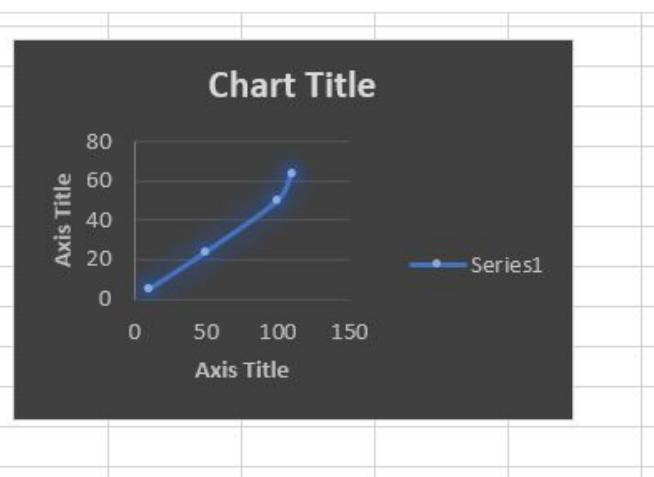
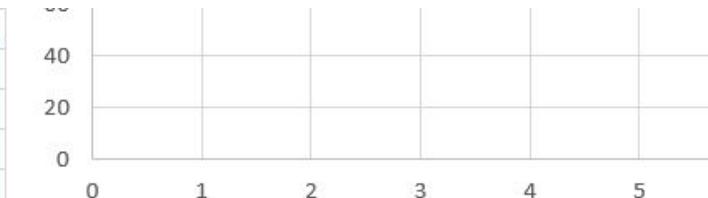
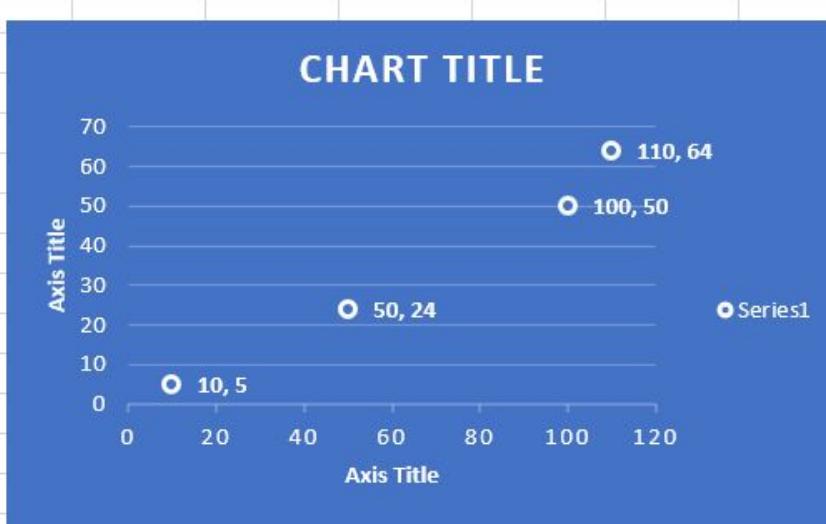
Step 3 – Select the data.

Step 4 – On the INSERT tab, in the Charts group, click the Scatter chart icon on the Ribbon.

You will see the different types of available Scatter charts.

A Scatter chart has the following sub-types –

- Scatter
- Scatter with Smooth Lines and Markers
- Scatter with Smooth Lines
- Scatter with Straight Lines and Markers
- Scatter with Straight Lines

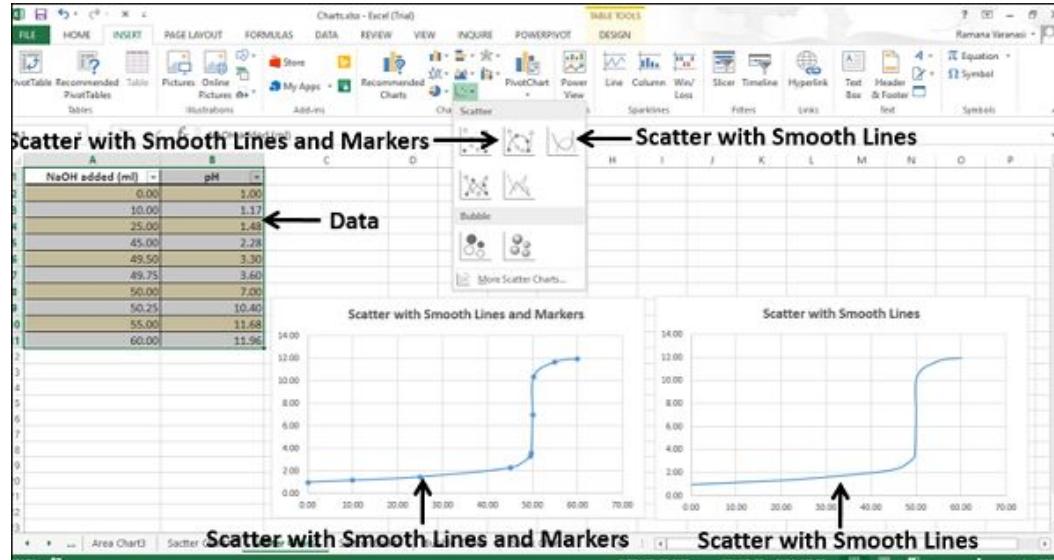


# Types of Scatter Charts

The following section explains the different options available to display a Scatter chart.

## Scatter with smooth lines and markers and scatter with smooth lines.

Scatter with Smooth Lines and Markers and Scatter with Smooth Lines display a smooth curve that connects the data points. Scatter with Smooth Lines and Markers and Scatter with Smooth Lines are useful to compare at least two sets of values or pairs of data.

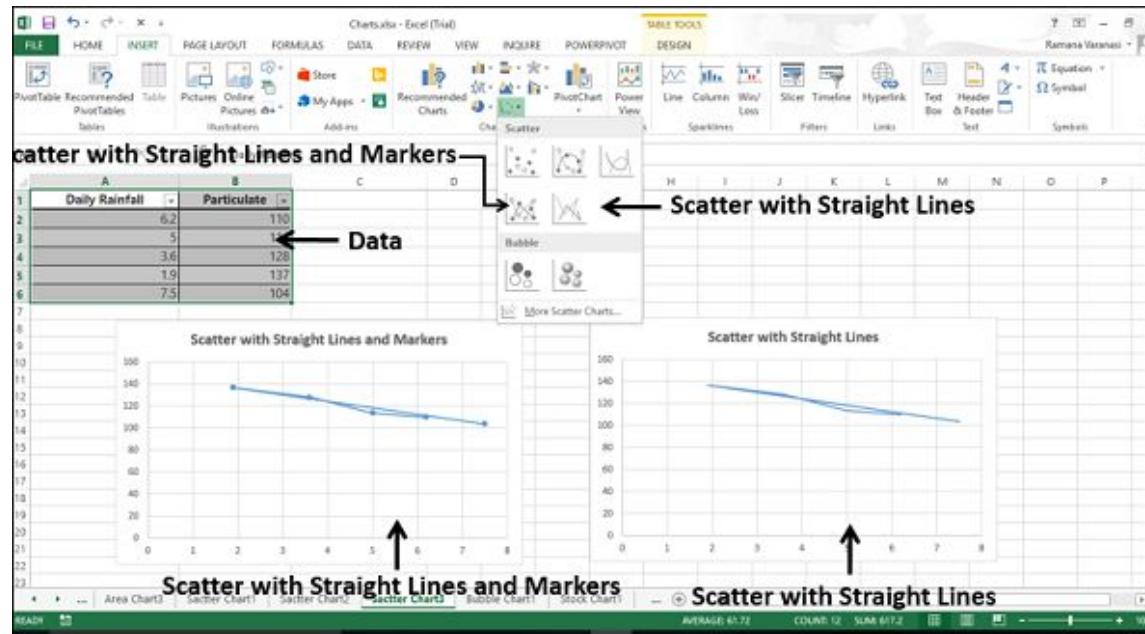


Use Scatter with Smooth Lines and Markers and Scatter with Smooth Lines charts when the data represents a set of x, y pairs based on a formula.

- Use Scatter with Smooth Lines and Markers when there are a few data points.
- Use Scatter with Smooth Lines when there are many data points.

# Scatter with Straight Lines and Markers and Scatter with Straight Lines

Scatter with Straight Lines and Markers and Scatter with Straight Lines connects the data points with straight lines. Scatter with Straight Lines and Markers and Scatter with Straight Lines are useful to compare at least two sets of values or pairs of data.

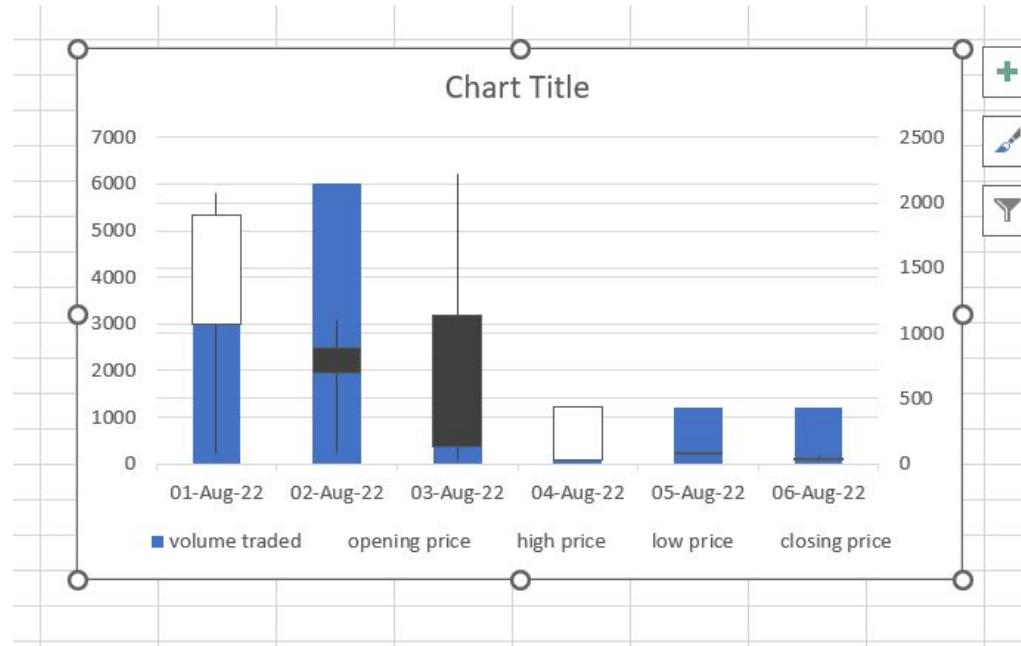


Use Scatter with Straight Lines and Markers and Scatter with Straight Lines charts when the data represents separate measurements.

- Use Scatter with Straight Lines and Markers when there are a few data points.
- Use Scatter with Straight Lines when there are many data points.

A Stock chart has the following sub-types –

- High-Low-Close
- Open-High-Low-Close
- Volume-High-Low-Close
- Volume-Open-High-Low-Close



# Surface Chart

Surface charts are useful when you want to find the optimum combinations between two sets of data. As in a topographic map, the colors and patterns indicate the areas that are in the same range of values.

To create a Surface chart, ensure that both the categories and the data series are numeric values.

Step 1 – Arrange the data in columns or rows on the worksheet.

Step 2 – Select the data.

Step 3 – On the INSERT tab, in the Charts group, click the Stock, Surface or Radar Chart icon on the Ribbon.

ChartTools - Excel (Trial)

FILE HOME INSERT PAGE LAYOUT FORMULAS DATA REVIEW VIEW INQUIRE POWERPIVOT DESIGN

PivotTable Recommended PivotTables Tables Pictures Online Pictures Store My Apps Recommended Charts Charts

All X ✓ fx Genre

|   | A                | B        | C         | D        | E        | F        | G | H |
|---|------------------|----------|-----------|----------|----------|----------|---|---|
| 1 | Genre            | 2008     | 2009      | 2010     | 2011     | 2012     |   |   |
| 2 | Classics         | \$18,580 | \$49,225  | \$18,326 | \$20,017 | \$26,134 |   |   |
| 3 | Mystery          | \$78,970 | \$82,262  | \$48,640 | \$45,985 | \$73,429 |   |   |
| 4 | Romance          | \$24,236 | \$131,390 | \$79,022 | \$73,009 | \$81,428 |   |   |
| 5 | Sci-Fi & Fantasy | \$16,730 | \$19,790  | \$12,109 | \$11,355 | \$17,686 |   |   |
| 6 | Young Adult      | \$15,358 | \$42,685  | \$20,891 | \$16,065 | \$21,388 |   |   |

Stock Win/Loss Slicer Timeline Hyperlink Text Box Header & Footer Symbol

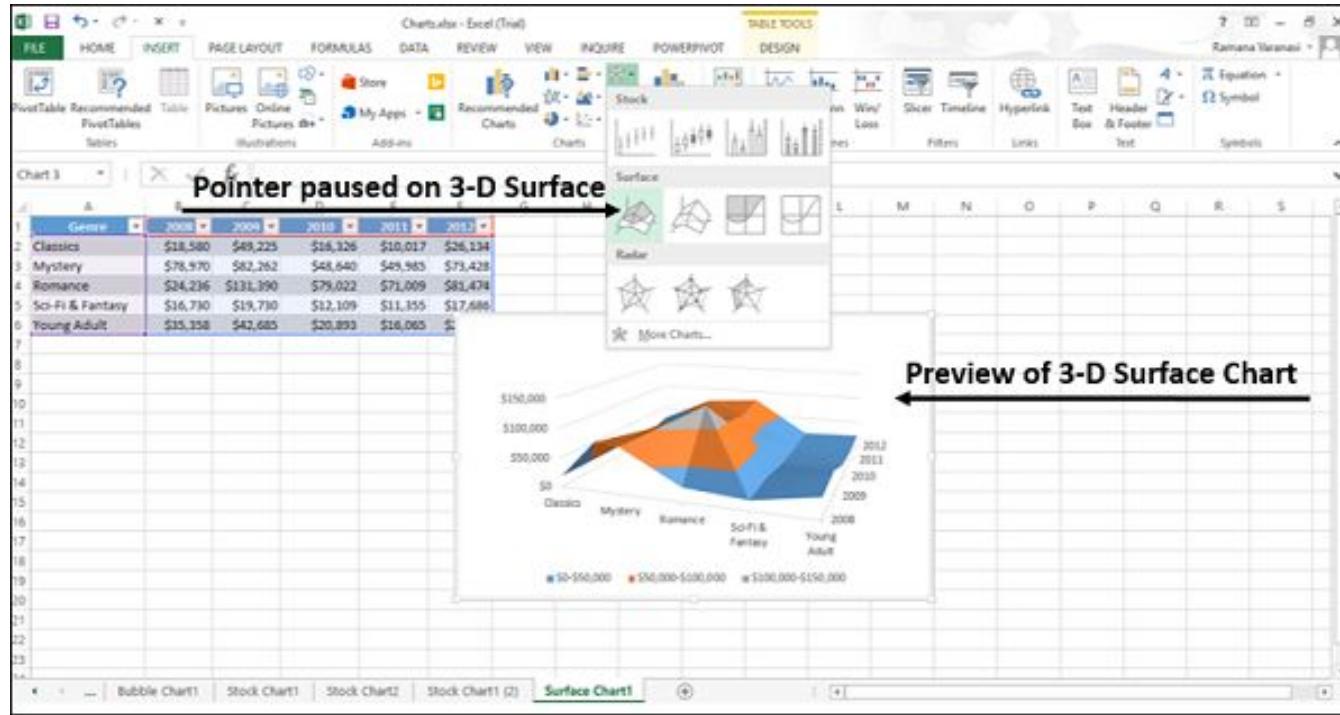
Surface Filters Lines Symbols

Radar More Charts...

**Arrow**

**Types of Surface Charts available**

The screenshot shows a Microsoft Excel spreadsheet titled "ChartTools - Excel (Trial)". The "Chart Tools" ribbon tab is selected, indicated by a yellow bar above the ribbon. The "DESIGN" tab is currently active. On the left, there is a PivotTable with data for book genres from 2008 to 2012. The "Surface" chart type is highlighted in the ribbon's chart section. A large black arrow points to the "Stock" chart icon in the "Surface" group. Another black arrow points to the "More Charts..." button at the bottom right of the ribbon.



# RADAR

Radar charts compare the aggregate values of several data series.

To insert a Radar chart in your worksheet, follow the steps given below.

A Radar chart has the following sub-types –

- Radar
- Radar with Markers
- Filled Radar

# Sparklines

Sparklines are tiny charts placed in single cells, each representing a row of data in your selection. They provide a quick way to see trends.

Sparklines have the following types –

- Line Sparkline
- Column Sparkline
- Win/Loss Sparkline

Quick Analysis Button

| Student          | Exam 1 | Exam 2 | Exam 3 | Exam 4 |
|------------------|--------|--------|--------|--------|
| Kreiger, Doris   | 87     | 90     | 79     | 96     |
| Oliviera, Manuel | 92     | 94     | 94     | 97     |
| Kodieda, Adam    | 96     | 95     | 95     | 80     |
| Langes, Michael  | 85     | 87     | 87     | 88     |
| Taylor, Maurice  | 81     | 88     | 82     | 85     |

# Data Analytics Life Cycle



# Data Discovery

- This is the initial phase to set your project's objectives and find ways to achieve a complete data analytics lifecycle. Start with defining your business domain and ensure you have enough resources (time, technology, data, and people) to achieve your goals.
- The biggest challenge in this phase is to accumulate enough information. You need to draft an analytic plan, which requires some serious leg work.

## Accumulate resources

First, you have to analyze the models you have intended to develop. Then determine how much domain knowledge you need to acquire for fulfilling those models.

The next important thing to do is assess whether you have enough skills and resources to bring your projects to fruition

## Frame the issue

Problems are most likely to occur while meeting your client's expectations. Therefore, you need to identify the issues related to the project and explain them to your clients. This process is called "framing." You have to prepare a problem statement explaining the current situation and challenges that can occur in the future. You also need to define the project's objective, including the success and failure criteria for the project.

## Formulate initial hypothesis

Once you gather all the clients' requirements, you have to develop initial hypotheses after exploring the initial data.

**Data discovery, also known as data mining, can be used in the research field to discover and extract patterns in large data sets** and help spot common data structures that can be brought together for more profound, insightful information.

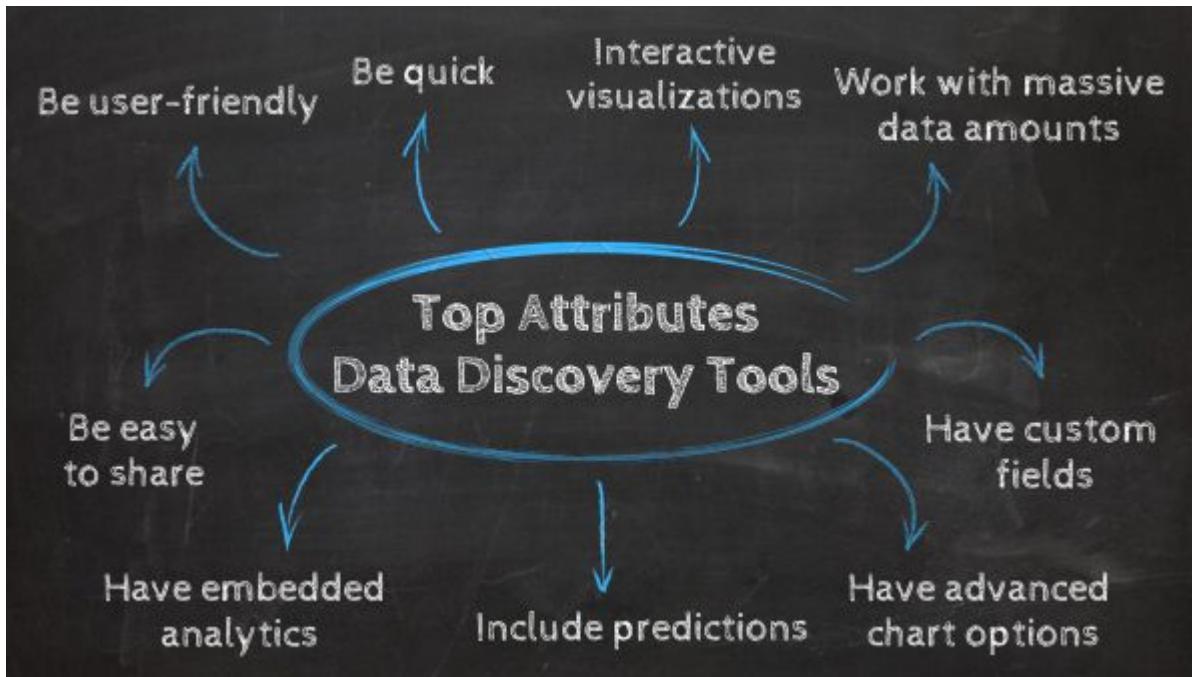
# Data Discovery



## Benefits

- ✓ Enables data visualization
- ✓ As part of business intelligence
- ✓ Aid behavioral analysis
- ✓ Understand the full life cycle
- ✓ Predictive analytics

# Attributes



## How does data discovery help?

- **Discover new opportunities** It helps uncover new insights for methods of business value creation.
- **Replicate success** Can drive similar high-value business outcomes where data was the catalyst of a modern business' operations success story.
- **Secure data** Can apply data protection to lower the risk of its exposure and prevent abuse, theft, and leaks.
- **Achieve compliance** Businesses can keep track of their data and its security to make sure they understand how safe it is to ensure they are compliant with industry standards.
- **Adopt the cloud** In cases where a move to the cloud (or further expansion) is needed, and data discovery tools gather all the digital assets in an ecosystem. This helps ensure that not a single piece of data is overlooked.

## What makes for a good data discovery tool?

- **Ease of use** A great tool is always easy to set up and start using. Likewise, a data discovery tool needs to be simple enough for non-technical users to create the dashboards and insights they need straight out of the box.
- **Deep discovery capabilities** Should also be able to track data regardless of its location – be it in the cloud or on-premises; as long as it belongs to the organization, the data should be displayed in its dashboards.
- **Ability to process big data** Most companies need the help of data discovery tools because they have large amounts of data. Therefore, a good data discovery tool should be able to find data, process it, and present it with ease and in the shortest time possible.

- **Recognition of data types** The tool should also identify data types in whatever format they may be stored in, and even if it has been corrupted or is missing attributes.
- **Display data in insightful dashboards** The reports and dashboards created from discovered data should help with easy and informed decision-making.
- **Collaboration features** It is rare that only one user creates a dashboard and then uses it Therefore, a good data discovery tool should allow dashboards and reports to be easily shared among stakeholders.
- **The price** Cost-effectiveness and a positive return on investments (ROI) will always be at the fore of any product

# **The best data discovery tools**

**1. Informatica**

**2. Qlik Sense**

**3. Tableau**

**4. SyncSpider**

**5. Nightfall**

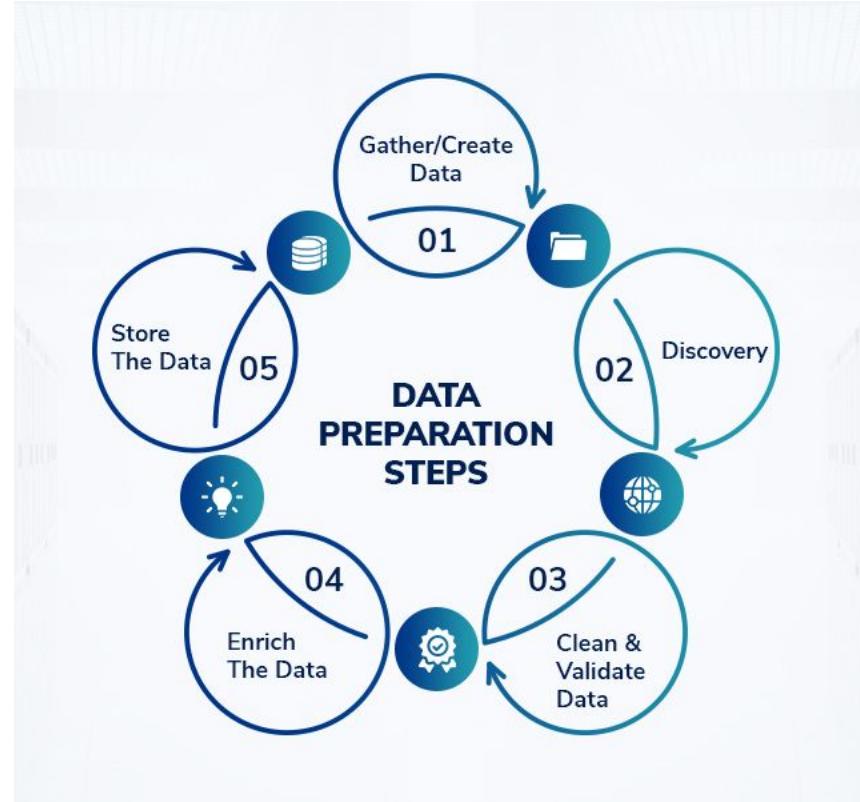
**6. Atlan**

**7. PowerBI**

**8. Tibco**

**9. Grow**

# Data Preparation



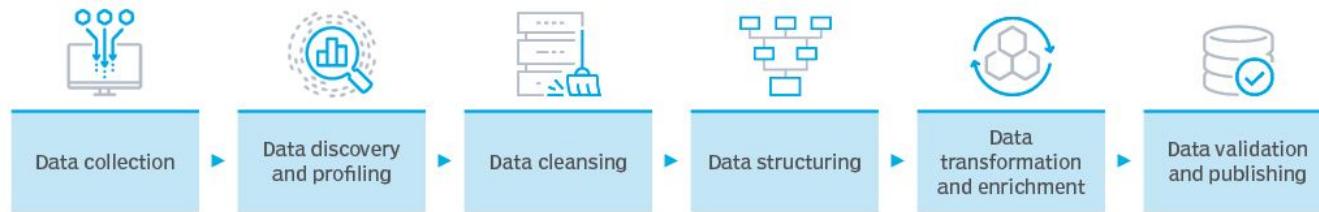
Data preparation is the process of getting raw data ready for analysis and processing. This can mean restructuring the data at hand, merging sets for a more complete view, and even making corrections to data that isn't recorded properly. While this sort of work is highly time-consuming, it is essential for any job that involves working with large amounts of complex data.

# Steps in the data preparation process

- **Data collection.** Relevant data is gathered from operational systems, data warehouses, data lakes and other data sources. During this step, data scientists, members of the BI team, other data professionals and end users who collect data should confirm that it's a good fit for the objectives of the planned analytics applications.
- **Data discovery and profiling.** The next step is to explore the collected data to better understand what it contains and what needs to be done to prepare it for the intended uses. To help with that, data profiling identifies patterns, relationships and other attributes in the data, as well as inconsistencies, anomalies, missing values and other issues so they can be addressed.
- **Data cleansing.** Next, the identified data errors and issues are corrected to create complete and accurate data sets. For example, as part of cleansing data sets, faulty data is removed or fixed, missing values are filled in and inconsistent entries are harmonized.

- **Data structuring.** At this point, the data needs to be modeled and organized to meet the analytics requirements. For example, data stored in comma-separated values (CSV) files or other file formats has to be converted into tables to make it accessible to BI and analytics tools.
- **Data transformation and enrichment.** In addition to being structured, the data typically must be transformed into a unified and usable format. For example, data transformation may involve creating new fields or columns that aggregate values from existing ones. Data enrichment further enhances and optimizes data sets as needed, through measures such as augmenting and adding data.
- **Data validation and publishing.** In this last step, automated routines are run against the data to validate its consistency, completeness and accuracy. The prepared data is then stored in a data warehouse, a data lake or another repository and either used directly by whoever prepared it or made available for other users to access.

# Key data preparation steps

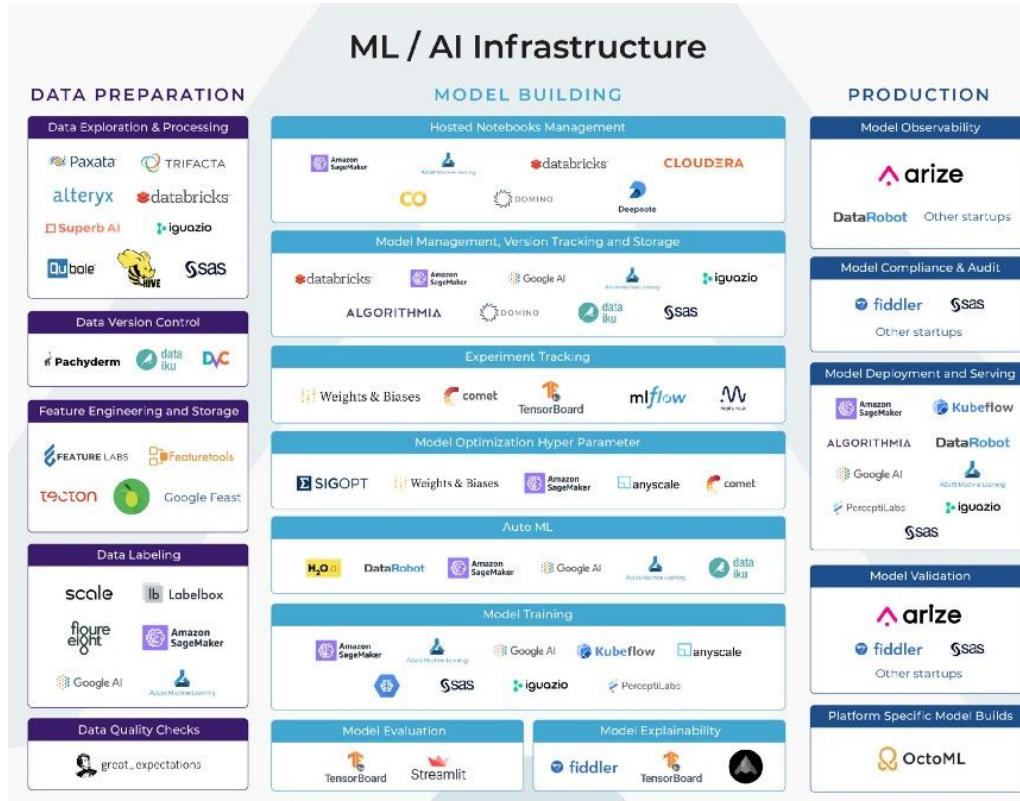


ICONS FROM LEFT: PRIVANIKAGUPTA/GETTY IMAGES, TIM\_UKRI/GETTY IMAGES, ENIS AKSOY/GETTY IMAGES, FINGERMEDIUM/GETTY IMAGES, ENOTMAKS/GETTY IMAGES, BROWNDOGSTUDIOS/GETTY IMAGES

# Benefits of Data Preparation And The Cloud

- **Fix Errors Quickly:** Fixing errors before processing data is much faster than doing it after the fact.
- **High-Quality Data:** With errors being fixed so quickly, your data will always see a quality increase after preparation.
- **More Usable Data:** Higher quality data will be easier to read and make use of, making this process well worth it.
- **Easy Collaboration:** Storing all of your data on the cloud will make it easier for the whole team to access, aiding collaboration.
- **Future Proof:** Unlike having your own servers, cloud options can scale with your business, securing your future without forcing you to constantly upgrade.

# Tools



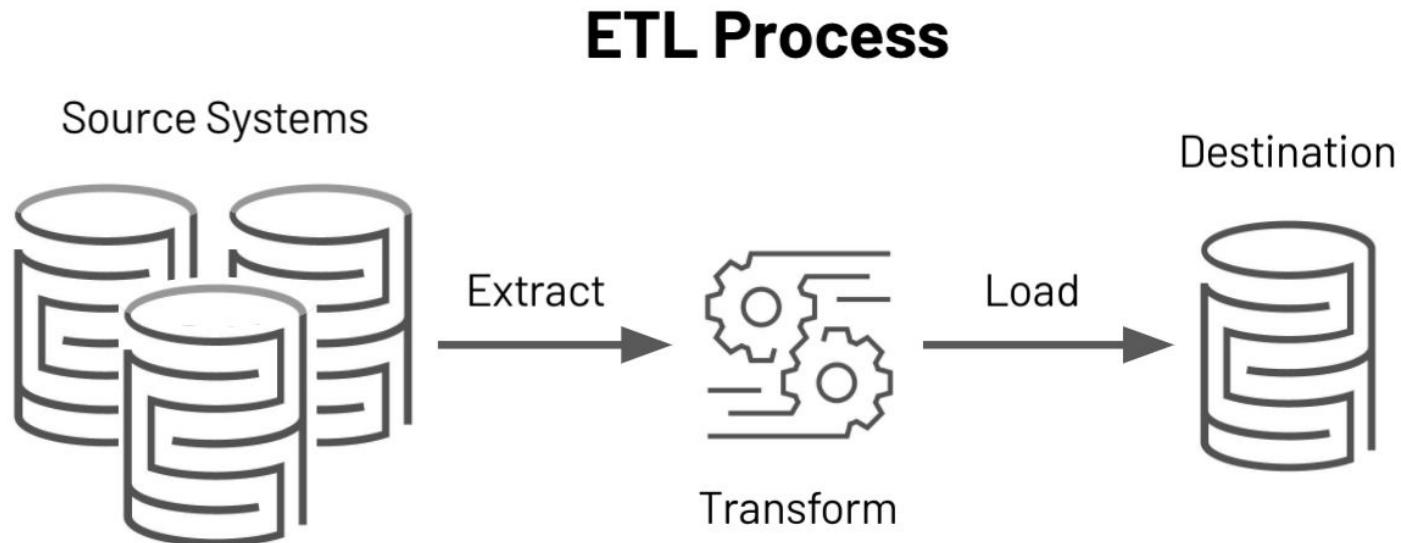
## Model Planning

This is a phase where you have to analyze the quality of data and find a suitable model for your project.

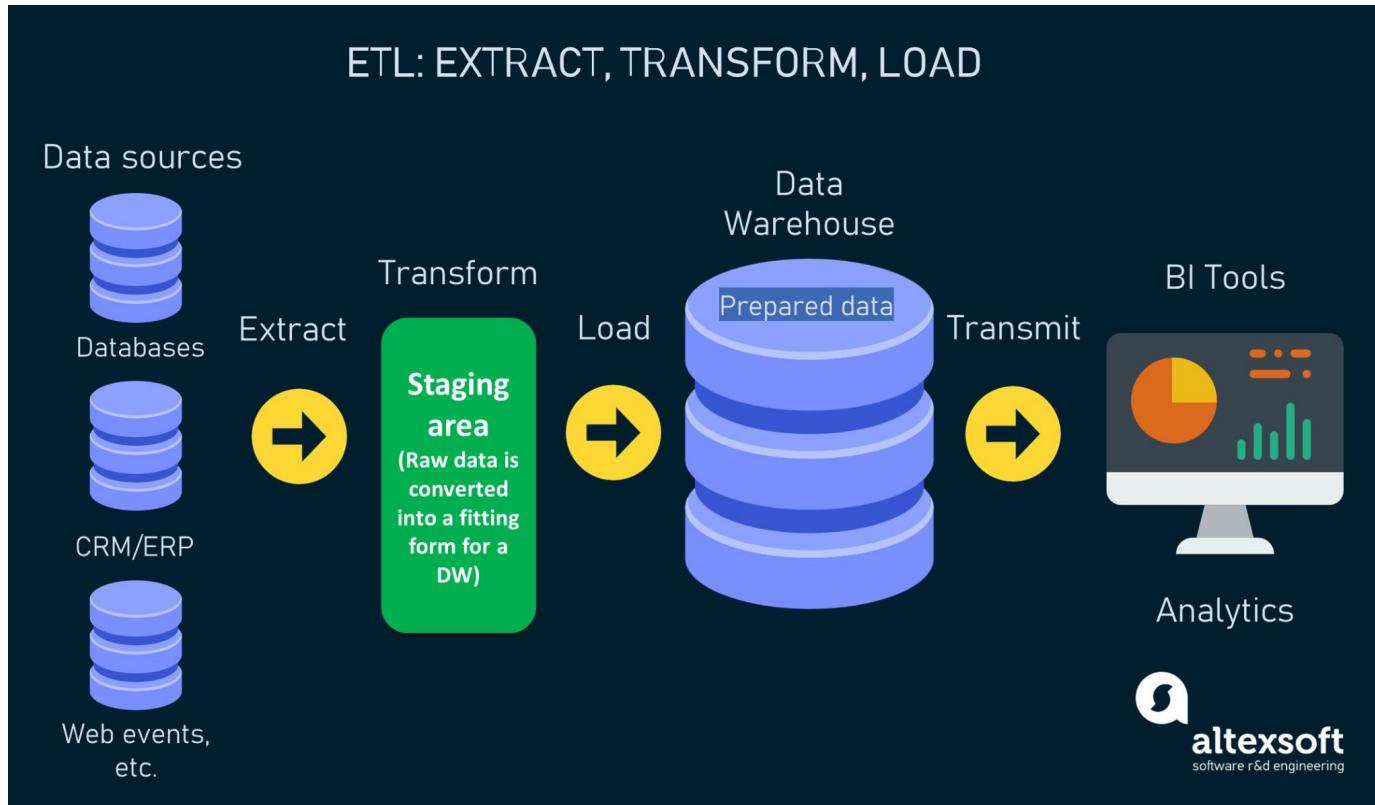
An analytics sandbox is a part of data lake architecture that allows you to store and process large amounts of data. It can efficiently process a large range of data such as big data, transactional data, social media data, web data, and many more. It is an environment that allows your analysts to schedule and process data assets using the data tools of their choice. The best part of the analytics sandbox is its agility. It empowers analysts to process data in real-time and get essential information within a short duration.

Data are loaded in the sandbox in three ways:

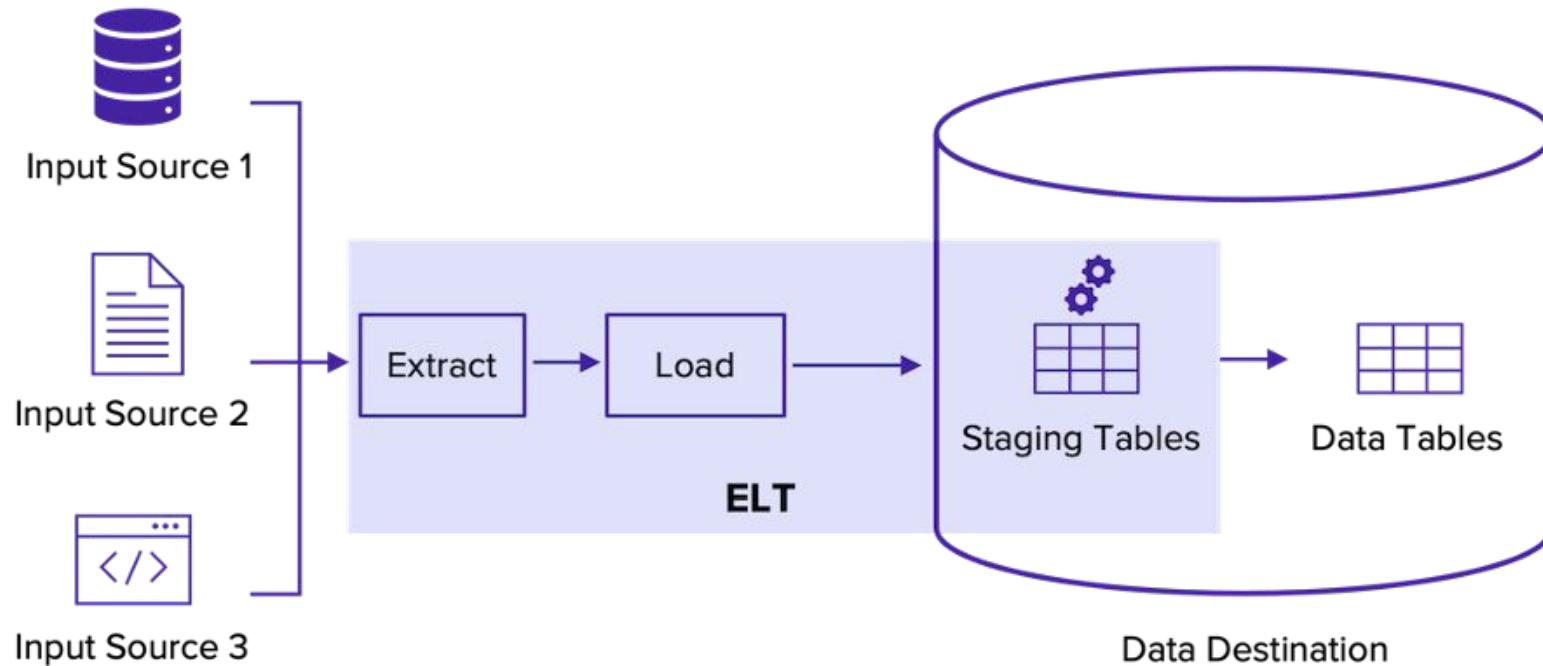
ETL



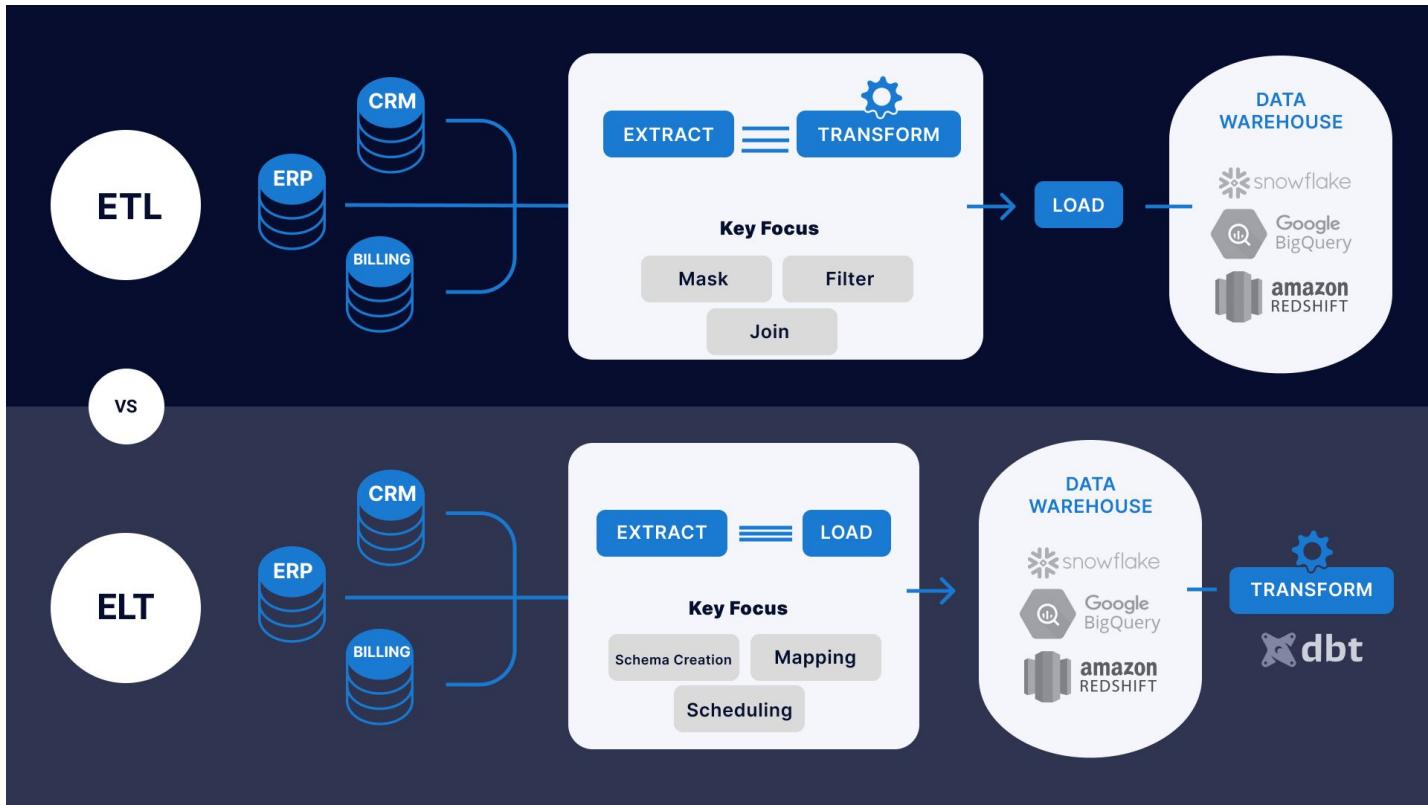
# ETL



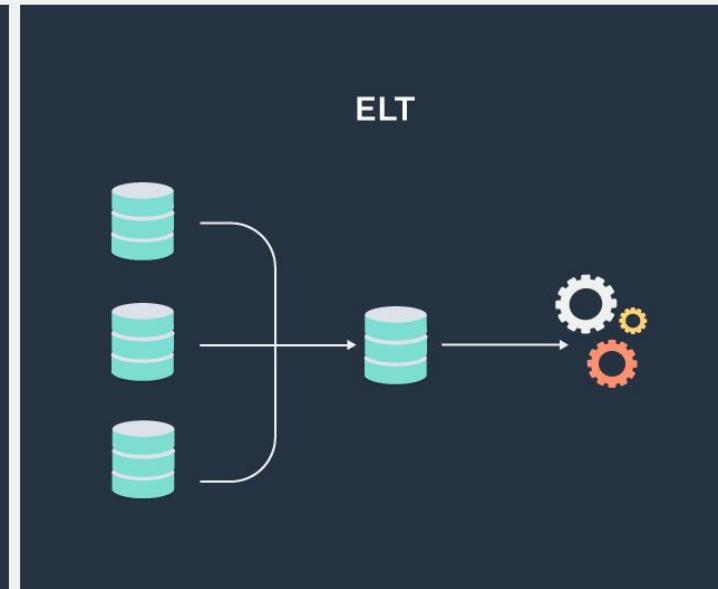
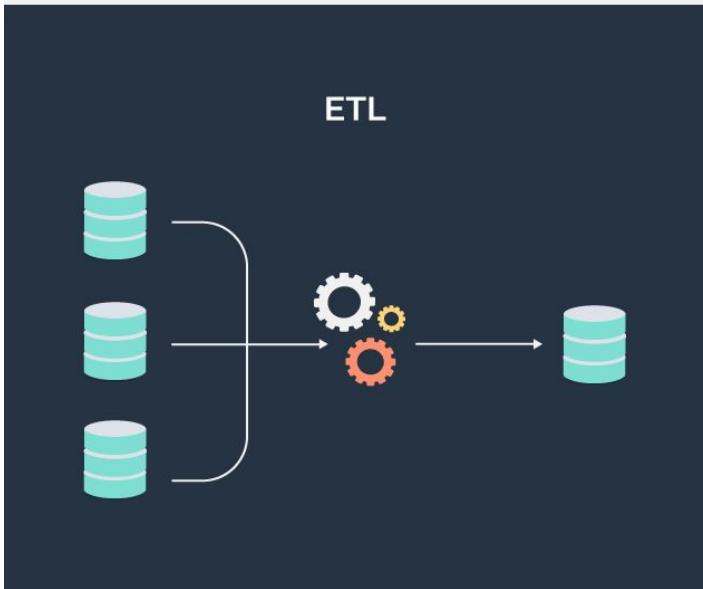
# ELT



# VS



# Evolution of ETL to ELT



# ETLT

It comprises two levels of data transformation, including ETL and ELT both.

1. **Extract** the raw, unprepared data from source applications and databases and load it into a staging area.
2. **Transform** data “lightly” while it’s in the staging area (usually to remove/mask/encrypt PHI, PII, or other sensitive data). The first transformation stage only applies to one data source at a time. These transformations are fast and simple because they transform each source independently of other sources. There is no attempt to integrate two data sources into one until after loading. Transformations for this first stage relate to data formats, data cleansing, and masking/removing sensitive data for compliance purposes.

1. **Load** the prepared data into the data warehouse.
2. **Transform** and integrate data more completely within the data warehouse using the data warehouse to process those transactions. The second transformation stage relates to integrating multiple data sources and other transformations that apply to data from multiple sources at the same time.

ETLT allows you to ingest data from diverse sources faster because the pre-load transformation stage only performs light-duty transformations. These light-duty transformations are quick to set up and fast to process, and they overcome the risks of ELT by satisfying essential data compliance requirements and preserving data quality. Meanwhile, you can leave the complex, multi-source transformations for the data warehouse to process later.

Steps involved in data exploration:

- Data identification
- Univariate Analysis
- Multivariate Analysis
- Filling Null values
- Feature engineering

# **Model Building**

- In this phase data science team needs to develop data sets for training, testing, and production purposes. These data sets enable data scientist to develop analytical method and train it, while holding aside some of data for testing the model.
- Team develops datasets for testing, training, and production purposes. In addition, in this phase, the team builds and executes models based on work done in the model planning phase. The team also considers whether its existing tools will suffice for running the models, or if it will need more robust environment for executing models and workflows (Example – fast hardware and parallel processing).

## **Free or open-source Tools :**

Rand PL/R, Octave, WEKA, Python

## **Commercial Tools –**

Matlab, STASTICA

### **Python :**

It is programming language that provides toolkits for machine learning and analysis, such as scikit-learn, NumPy, scipy, Pandas, and related data visualization using matplotlib.

### **SQL :**

SQL in database implementations, such as MADlib, provides an alternative to memory desktop analytical tools.

## **Lifecycle of Model Building –**

- Select variables
- Balance data
- Build models
- Validate
- Deploy
- Maintain
- Define success
- Explore data
- Condition data

# Result Communication and Publication

- This is the phase where you have to communicate the data analysis with your clients. It requires several intricate processes where you how to present information to clients in a lucid manner. Your clients don't have enough time to determine which data is essential. Therefore, you must do an impeccable job to grab the attention of your clients.
- Check the data accuracy
- Is the data provide information as expected? If not, then you have to run some other processes to resolve this issue. You need to ensure the data you process provides consistent information. This will help you build a convincing argument while summarizing your findings.
- Highlight important findings

- Well, each data holds a significant role in building an efficient project. However, some data inherits more potent information that can truly serve your audience's benefits. While summarizing your findings, try to categorize data into different key points.
- Determine the most appropriate communication format
- How you communicate your findings tells a lot about you as a professional. We recommend you to go for visuals presentation and animations as it helps you to convey information much faster. However, sometimes you also need to go old-school as well. For instance, your clients may have to carry the findings in physical format. They may also have to pick up certain information and share them with others.

# Operationalize

- As soon you prepare a detailed report including your key findings, documents, and briefings, your data analytics life cycle almost comes close to the end. The next step remains the measure the effectiveness of your analysis before submitting the final reports to your stakeholders.
- In this process, you have to move the sandbox data and run it in a live environment. Then you have to closely monitor the results, ensuring they match with your expected goals. If the findings fit perfectly with your objective, then you can finalize the report. Otherwise, you have to take a step back in your data analytics lifecycle and make some changes.





