**Data analysts works**

# **What Does a Data Analyst Do?**

A data analyst collects, cleans, and interprets data sets in order to answer a question or solve a problem. They work in many industries, including business, finance, criminal justice, science, medicine, and government.

What kind of customers should a business target in its next ad campaign? What age group is most vulnerable to a particular disease? What patterns in behavior are connected to financial fraud?

These are the types of questions you might be pressed to answer as a data analyst. Read on to find out more about what a data analyst is, what skills you'll need, and how you can start on a path to become one.

## **What is data analysis?**

Data analysis is the process of gleaning insights from data to inform better business decisions. The process of analyzing data typically moves through five iterative phases:

**Identify** the data you want to analyze

**Collect** the data

**Clean** the data in preparation for analysis

**Analyze** the data

**Interpret** the results of the analysis

Data analysis can take different forms, depending on the question you’re trying to answer. You can read more about the [types of data analysis](https://www.coursera.org/articles/what-is-data-analysis-with-examples" \t "https://www.coursera.org/articles/_self) here. Briefly, descriptive analysis tells us what happened, diagnostic analysis tells us why it happened, predictive analytics forms projections about the future, and prescriptive analysis creates actionable advice on what actions to take.

## **Types of data analysis (with examples)**

Data can be used to answer questions and support decisions in many different ways. To identify the best way to analyze your date, it can help to familiarize yourself with the four types of data analysis commonly used in the field.

In this section, we’ll take a look at each of these data analysis methods, along with an example of how each might be applied in the real world.

### **Descriptive analysis**

Descriptive analysis tells us what happened. This type of analysis helps describe or summarize quantitative data by presenting statistics. For example, descriptive statistical analysis could show the distribution of sales across a group of employees and the average sales figure per employee.

**Descriptive analysis** answers the question, “what happened?”

### **Diagnostic analysis**

If the descriptive analysis determines the “what,” diagnostic analysis determines the “why.” Let’s say a descriptive analysis shows an unusual influx of patients in a hospital. Drilling into the data further might reveal that many of these patients shared symptoms of a particular virus. This diagnostic analysis can help you determine that an infectious agent—the “why”—led to the influx of patients.

**Diagnostic analysis** answers the question, “why did it happen?”

### **Predictive analysis**

So far, we’ve looked at types of analysis that examine and draw conclusions about the past. Predictive analytics uses data to form projections about the future. Using predictive analysis, you might notice that a given product has had its best sales during the months of September and October each year, leading you to predict a similar high point during the upcoming year.

**Predictive analysis** answers the question, “what might happen in the future?”

### **Prescriptive analysis**

Prescriptive analysis takes all the insights gathered from the first three types of analysis and uses them to form recommendations for how a company should act. Using our previous example, this type of analysis might suggest a market plan to build on the success of the high sales months and harness new growth opportunities in the slower months.

**Prescriptive analysis** answers the question, “what should we do about it?”

This last type is where the concept of data-driven decision-making comes into play.

**Read more**: [Advanced Analytics: Definition, Benefits, and Use Cases](https://www.coursera.org/articles/advanced-analytics" \t "https://www.coursera.org/articles/_self)

### **What is data-driven decision-making (DDDM)?**

Data-driven decision-making, sometimes abbreviated to DDDM), can be defined as the process of making strategic business decisions based on facts, data, and metrics instead of intuition, emotion, or observation.

This might sound obvious, but in practice, not all organizations are as data-driven as they could be. According to global management consulting firm McKinsey Global Institute, data-driven companies are better at acquiring new customers, maintaining customer loyalty, and achieving above-average profitability [[2](https://www.mckinsey.com/business-functions/marketing-and-sales/our-insights/five-facts-how-customer-analytics-boosts-corporate-performance" \t "https://www.coursera.org/articles/_blank)].

## **Data analyst tasks and responsibilities**

A data analyst is a person whose job is to gather and interpret data in order to solve a specific problem. The role includes plenty of time spent with data but entails communicating findings too.

Here’s what many data analysts do on a day-to-day basis:

**Gather data:**Analysts often collect data themselves. This could include conducting surveys, tracking visitor characteristics on a company website, or buying datasets from data collection specialists.

**Clean data:** Raw data might contain duplicates, errors, or outliers. Cleaning the data means maintaining the quality of data in a spreadsheet or through a programming language so that your interpretations won’t be wrong or skewed.

**Model data:**This entails creating and designing the structures of a database. You might choose what types of data to store and collect, establish how data categories are related to each other, and work through how the data actually appears.

**Interpret data:** Interpreting data will involve finding patterns or trends in data that will help you answer the question at hand.

**Present:**Communicating the results of your findings will be a key part of your job. You do this by putting together visualizations like charts and graphs, writing reports, and presenting information to interested parties.

### **What tools do data analysts use?**

During the process of data analysis, analysts often use a wide variety of tools to make their work more accurate and efficient. Some of the most common tools in the data analytics industry include:

Microsoft Excel

Google Sheets

SQL

Tableau

R or Python

SAS

Microsoft Power BI

Jupyter Notebooks

### **Types of data analysts**

As advancing technology has rapidly expanded the types and amount of information we can collect, knowing how to gather, sort, and analyze data has become a crucial part of almost any industry. You’ll find data analysts in the criminal justice, fashion, food, technology, business, environment, and public sectors—among many others.

People who perform data analysis might have other titles such as:

Medical and health care analyst

[Market research analyst](https://www.coursera.org/articles/market-research-analyst" \t "https://www.coursera.org/articles/_self)

[Business analyst](https://www.coursera.org/articles/what-does-a-business-analyst-do-and-how-to-become-one" \t "https://www.coursera.org/articles/_self)

[Business intelligence analyst](https://www.coursera.org/articles/business-intelligence-analysts-what-they-are-and-how-to-become-one" \t "https://www.coursera.org/articles/_self)

Operations research analyst

Intelligence analyst

### **Data analyst vs. data scientist: What’s the difference?**

Data analysts and data scientists both work with data, but what they do with it differs. Data analysts typically work with existing data to solve defined business problems. Data scientists build new algorithms and models to make predictions about the future.

## **How to become a data analyst**

There’s more than one path toward a career as a data analyst. Whether you’re just graduating from school or looking to switch careers, the first step is often assessing what transferable skills you have and building the new skills you’ll need in this new role.

### **Data analyst technical skills**

**Database tools:** Microsoft Excel and [SQL](https://www.coursera.org/articles/sql-certifications-for-your-data-career" \t "https://www.coursera.org/articles/_self) should be mainstays in any data analyst’s toolbox. While Excel is ubiquitous across industries, SQL can handle larger sets of data and is widely regarded as a necessity for data analysis.

### **What is big data?**

The term “[big data](https://www.coursera.org/articles/what-is-big-data-a-laypersons-guide" \t "https://www.coursera.org/articles/_self)” refers to the vast amounts of structured and unstructured data that many businesses have access to on a daily basis. These data sets are typically too large to process using traditional data analysis methods. Big data is characterized by the three Vs: high **volume**, **variety** of data types, and the **velocity** at which the data is received.

**Programming languages:** Learning a statistical programming language like [Python or R](https://www.coursera.org/articles/python-or-r-for-data-analysis" \t "https://www.coursera.org/articles/_self) will let you handle large sets of data and perform complex equations. Though Python and R are among the most common, it’s a good idea to look at several job descriptions of a position you’re interested in to determine which language will be most useful to your industry.

**Data visualization:**Presenting your findings in a clear and compelling way is crucial to being a successful data analyst. Knowing how best to present information through charts and graphs will make sure colleagues, employers, and stakeholders will understand your work. Tableau, Jupyter Notebook, and Excel are among the many tools used to create visuals.

**Statistics and math:** Knowing the concepts behind what data tools are actually doing will help you tremendously in your work. Having a solid grasp of statistics and math will help you determine which tools are best to use to solve a particular problem, help you catch errors in your data, and have a better understanding of the results.

#### **Data analyst workplace skills**

**Problem solving:**A data analyst needs to have a good understanding of the question being asked and the problem that needs to be solved. They also should be able to find patterns or trends that might reveal a story. Having the critical thinking skills will allow you to focus on the right types of data, recognize the most revealing methods of analysis, and catch gaps in your work.

**Communication:**Being able to get your ideas across to other people will be crucial to your work as a data analyst. Strong written and speaking skills to communicate with colleagues and other stakeholders are good assets in data analysts.

**Industry knowledge:**Knowing about the industry you work in—health care, business, finance, or otherwise—will give you an advantage in your work and in job applications. If you’re trying to break into a specific industry, take some time to pay attention to the news in your industry, or read a book on the subject. This can familiarize you with the industry’s main issues and trends.

### **Paths to becoming a data analyst**

Acquiring these skills is the first step to becoming a data analyst. Here are a few routes you can take to get them that are flexible enough to fit in around your life.

**Professional certificate:**Entry-level professional certificate programs usually require no previous experience in the field. They can teach you basic skills like SQL or statistics while giving you the chance to create projects for your portfolio and provide real-time feedback on your work. Several [professional certificate programs on Coursera](https://www.coursera.org/professional-certificates" \t "https://www.coursera.org/articles/_self) do just that.

**Bachelor's degree:**The Bureau of Labor Statistics recommends a bachelor’s degree for jobs that involve data analysis. If you’re considering getting a degree to become a data analyst, focusing your coursework in statistics, math, or computer science can give you a head start with potential employers. Many online bachelor’s degrees have flexible scheduling so you can fit a degree in around your priorities.

**Self-study:** If you want a path that doesn’t include formal training, it’s possible to learn the skills necessary for data analysis on your own. Get started with this [data analytics reading list for beginners](https://www.coursera.org/articles/data-analytics-books-for-beginners" \t "https://www.coursera.org/articles/_self). Once you’re ready to start [building a portfolio](https://www.coursera.org/articles/how-to-build-a-data-analyst-portfolio" \t "https://www.coursera.org/articles/_self), here are some [ideas for data analytics projects](https://www.coursera.org/articles/data-analytics-projects-for-beginners" \t "https://www.coursera.org/articles/_self).

**Information links :**

<https://www.coursera.org/articles/what-is-data-analysis-with-examples>

write an excel formula to add cells B2 to B10

ChatGPT

To add cells B2 to B10 in Excel, you can use the SUM function. The formula would be:

=SUM(B2:B10)

This will return the sum of all the values in cells B2 through B10.

User

can you explain how this function work?

ChatGPT

Sure! The SUM function in Excel is used to add up a range of cells. In this case, we want to add up the values in cells B2 to B10.

The syntax for the SUM function is as follows:

=SUM(number1, [number2], [number3], ...)

Where "number1" is the first number or cell reference you want to add, and the subsequent arguments are optional additional numbers or cell references you want to add.

So, in our formula, we specify the range of cells we want to add using cell references B2:B10. This tells Excel to add up all the values in the range of cells from B2 through B10.

The result of the formula is the sum of all the values in that range. So, if cells B2 to B10 contain the numbers 1, 2, 3, 4, 5, 6, 7, 8, and 9, respectively, then the formula will return the sum of those values, which is 45.

User

write an excel formula to calculate profit .the revenue is in cell A2 and cost is in cell B2

ChatGPT

To calculate profit in Excel based on revenue in cell A2 and cost in cell B2, you can use the following formula:

= A2 - B2

This formula subtracts the cost from the revenue to determine the profit.

For example, if the revenue in cell A2 is $100 and the cost in cell B2 is $50, the formula will calculate the profit as $50 ($100 - $50).

User

write an excel formula to find sugar cookie in a table and return the price

ChatGPT

To find the price of sugar cookies in an Excel table, you can use the VLOOKUP function. The VLOOKUP function allows you to search for a specific value in the first column of a table and return a corresponding value from another column in the same row.

Assuming the table with the cookie prices is in columns A and B, with the cookie names in column A and the prices in column B, and assuming "Sugar Cookie" is the specific cookie you want to find the price for, the formula would be:

=VLOOKUP("Sugar Cookie", A:B, 2, FALSE)

Here, "Sugar Cookie" is the value you want to search for in column A, A:B is the table range that includes both columns A and B, 2 is the column number of the price column (which is the second column in the table), and FALSE tells Excel to look for an exact match to "Sugar Cookie".

The formula will return the price of the sugar cookie in the same row as the value "Sugar Cookie" in column A. For example, if the price of a sugar cookie is $1.50 and it is listed in row 5 of the table, the formula will return $1.50.

User

Can i use any other function?

ChatGPT

2 / 2

Yes, there are other functions in Excel that you can use to find the price of sugar cookies in a table. Here are some other options:

INDEX-MATCH: You can use the INDEX-MATCH combination to achieve the same result as the VLOOKUP function. The formula would be:

=INDEX(B:B, MATCH("Sugar Cookie", A:A, 0))

Here, the MATCH function searches for the value "Sugar Cookie" in column A and returns the row number where it's found. Then, the INDEX function uses that row number to return the value from column B, which is the price of the sugar cookie.

SUMIFS: If you have a table with multiple columns of information, you can use the SUMIFS function to find the price of sugar cookies based on multiple criteria. For example, if your table has columns for cookie names, flavors, and prices, and you want to find the price of a sugar cookie with a vanilla flavor, the formula would be:

=SUMIFS(C:C, A:A, "Sugar Cookie", B:B, "Vanilla")

Here, the SUMIFS function adds up the values in column C (which contains the prices) that meet the criteria of "Sugar Cookie" in column A and "Vanilla" in column B. The result is the price of a sugar cookie with a vanilla flavor.

These are just a few examples of the many functions in Excel that you can use to search for and retrieve information from a table.

User

write an excel formula to extract all the text before @ character in cell A2

ChatGPT

To extract all the text before the "@" character in cell A2 in Excel, you can use the LEFT function in combination with the FIND or SEARCH function. Here are two possible formulas you can use:

Using the FIND function:

=LEFT(A2, FIND("@", A2) - 1)

This formula uses the FIND function to locate the position of the "@" character in cell A2, and then subtracts 1 from that position to exclude the "@" character from the result. The LEFT function is then used to extract all the characters to the left of the "@" character, up to the specified position.

Using the SEARCH function:

=LEFT(A2, SEARCH("@", A2) - 1)

This formula is similar to the previous one, but it uses the SEARCH function instead of FIND. The difference between the two functions is that SEARCH is not case-sensitive, while FIND is case-sensitive. If the text in cell A2 may contain both uppercase and lowercase letters, you may want to use SEARCH instead of FIND.

Both formulas will return all the text before the "@" character in cell A2. For example, if cell A2 contains the email address "jane.doe@example.com", both formulas will return "jane.doe".

User

write an excel formula to count the number of unique quantity in a list

ChatGPT

To count the number of unique quantities in a list in Excel, you can use the following formula:

=SUM(1/COUNTIF(range,range))

Here, "range" refers to the range of cells that contains the quantities you want to count.

This formula uses the COUNTIF function to count the number of times each quantity appears in the range, and then divides 1 by each count. Since the count of each unique quantity will be 1, the result of the division will be 1 for each unique quantity and 0 for each non-unique quantity.

The SUM function then adds up the results of the division, which is the number of unique quantities in the list.

For example, if your quantities are in cells A2:A10, the formula would be:

=SUM(1/COUNTIF(A2:A10,A2:A10))

This formula will count the number of unique quantities in the range A2:A10 and return the result. Note that this formula must be entered as an array formula by pressing Ctrl + Shift + Enter on your keyboard, since it involves an array of values.

User

write an excel macro to send emails. Use the following subjects "Kevin Cookie Company invoice".Use the following text."You owe the kevin Cookie Company X".Take the value X from column A .send to the email address listed in column B

ChatGPT

Here's an example Excel macro that sends emails using the subject and text you provided, and takes the value of X from column A and the email address from column B:

vbnet

Copy code

Sub SendEmails()

Dim OutApp As Object

Dim OutMail As Object

Dim LastRow As Long

Dim i As Long

Set OutApp = CreateObject("Outlook.Application")

LastRow = ActiveSheet.Cells(Rows.Count, 1).End(xlUp).Row ' Determine the last row with data

For i = 2 To LastRow ' Start from row 2 to exclude header row

Set OutMail = OutApp.CreateItem(0)

With OutMail

.To = Cells(i, 2).Value ' Email address from column B

.Subject = "Kevin Cookie Company invoice" ' Subject

.Body = "You owe the Kevin Cookie Company " & Cells(i, 1).Value ' Text with value of X from column A

.Send ' Send the email

End With

Set OutMail = Nothing

Next i

Set OutApp = Nothing ' Clear the Outlook object

End Sub

Before running the macro, make sure that Outlook is open and you have the appropriate permission to send emails. Also, adjust the range of cells to match the location of your data in columns A and B. Finally, test the macro with a small number of emails first to ensure that it is working as expected.

User

The names of sopranose series character

ChatGPT

Here are the names of some of the main characters in the "Sopranos" TV series:

Tony Soprano

Carmela Soprano

Meadow Soprano

Anthony Soprano Jr. (also known as "AJ")

Christopher Moltisanti

Paulie "Walnuts" Gualtieri

Silvio Dante

Dr. Jennifer Melfi

Bobby "Bacala" Baccalieri

Janice Soprano

Livia Soprano

Junior Soprano

Richie Aprile

Ralph Cifaretto

Adriana La Cerva

Note that there are many other characters in the series, including recurring characters and those that appear for only a few episodes.

User

give me excel formula to find which app was most downloaded in a table

ChatGPT

To find which app was most downloaded in a table in Excel, you can use the following formula:

=INDEX(apps,MATCH(MAX(downloads),downloads,0))

Here, "apps" refers to the range of cells that contains the app names, and "downloads" refers to the range of cells that contains the download counts for each app.

This formula uses the MAX function to find the highest download count in the "downloads" range, and then uses the MATCH function to find the position of that value in the "downloads" range. The INDEX function then returns the corresponding app name from the "apps" range.

For example, if your app names are in cells A2:A10 and download counts are in cells B2:B10, the formula would be:

=INDEX(A2:A10,MATCH(MAX(B2:B10),B2:B10,0))

This formula will return the name of the app that has the highest download count in the table. If there are multiple apps with the same highest download count, the formula will return the name of the first one it encounters in the "apps" range.

User

give me excel formula to find in which month was most downloaded app in a table

ChatGPT

To find in which month the most downloaded app was downloaded in a table in Excel, you can use the following formula:

=INDEX(months,MATCH(MAX(downloads),downloads,0))

Here, "months" refers to the range of cells that contains the month names, and "downloads" refers to the range of cells that contains the download counts for each app in each month.

This formula uses the MAX function to find the highest download count in the "downloads" range, and then uses the MATCH function to find the position of that value in the "downloads" range. The INDEX function then returns the corresponding month name from the "months" range.

For example, if your month names are in cells A1:L1 and download counts for each app are in cells A2:L10, the formula would be:

=INDEX(A1:L1,MATCH(MAX(A2:L10),A2:L10,0))

This formula will return the name of the month in which the most downloaded app was downloaded. If there are multiple months with the same highest download count, the formula will return the name of the first one it encounters in the "months" range.

User

in October 2021 which app has most 1Star ratings?

ChatGPT

To determine which app had the most 1-star ratings in October 2021, you would need to have a table with the names of the apps, their corresponding ratings for each month, and the months. Assuming you have such a table, you can use the following formula:

=INDEX(apps,MATCH(MAXIFS(OCT\_2021, OCT\_Ratings, 1),OCT\_2021,0))

Here, "apps" refers to the range of cells that contains the app names, "OCT\_2021" refers to the range of cells that contains the number of 1-star ratings for each app in October 2021, and "OCT\_Ratings" refers to the range of cells that contains the ratings for each app in October 2021.

This formula uses the MAXIFS function to find the highest number of 1-star ratings in October 2021, based on the criteria that the rating in the "OCT\_Ratings" range must be equal to 1. The MATCH function then finds the position of that value in the "OCT\_2021" range, and the INDEX function returns the corresponding app name from the "apps" range.

For example, if your app names are in cells A2:A10, the number of 1-star ratings for each app in October 2021 are in cells D2:D10, and the ratings for each app in October 2021 are in cells E2:E10, the formula would be:

=INDEX(A2:A10,MATCH(MAXIFS(D2:D10, E2:E10, 1),D2:D10,0))

This formula will return the name of the app that had the most 1-star ratings in October 2021. If there are multiple apps with the same highest number of 1-star ratings, the formula will return the name of the first one it encounters in the "apps" range.

User

continue

ChatGPT

It's important to note that for the formula to work correctly, you need to ensure that your data is properly formatted and that the ranges in the formula are correctly defined.

For example, the "OCT\_2021" range should only include the number of 1-star ratings for each app in October 2021, and not include any other ratings or months. Similarly, the "OCT\_Ratings" range should only include the ratings for each app in October 2021, and not include any other ratings or months.

Additionally, if there are ties for the highest number of 1-star ratings, the formula will return the name of the first app it encounters in the "apps" range that has that number of ratings. In such cases, you may want to consider using additional criteria to break ties and ensure that you are getting the desired result.

Overall, using the MAXIFS, MATCH, and INDEX functions in combination allows you to quickly and easily determine which app had the most 1-star ratings in October 2021, based on your data.

User

continue

ChatGPT

To further analyze your data, you may want to create additional formulas that calculate the number of 1-star ratings, 2-star ratings, and so on, for each app in each month. You can then use these formulas to create charts and pivot tables to visualize and explore your data.

For example, you could use the following formulas to calculate the number of 1-star ratings, 2-star ratings, and 3-star ratings for each app in October 2021:

Number of 1-star ratings: =COUNTIFS(A2:A10,"October 2021",B2:B10,"App Name",C2:C10,1)

Number of 2-star ratings: =COUNTIFS(A2:A10,"October 2021",B2:B10,"App Name",C2:C10,2)

Number of 3-star ratings: =COUNTIFS(A2:A10,"October 2021",B2:B10,"App Name",C2:C10,3)

Here, "A2:A10" refers to the range of cells that contains the months, "B2:B10" refers to the range of cells that contains the app names, and "C2:C10" refers to the range of cells that contains the ratings for each app in each month.

You can then use these formulas to create a pivot table that summarizes the number of 1-star, 2-star, and 3-star ratings for each app in October 2021. This will allow you to quickly compare the ratings of different apps and identify any trends or patterns in your data.

Overall, using a combination of formulas, charts, and pivot tables can help you analyze your data and gain insights into which apps are most popular and which ones may need improvement.

User

what are the best downloads figures for each app?

ChatGPT

To determine the best download figures for each app, you can use the MAX function in Excel. Assuming that your download figures are in a range of cells named "Download\_Figures" and your app names are in a range of cells named "Apps", you can use the following formula:

=MAX(Download\_Figures\*(Apps="App Name"))

Replace "App Name" with the name of the app you want to find the best download figures for.

This formula multiplies the range of download figures by a logical test that checks if the corresponding cell in the "Apps" range matches the name of the app you're interested in. This returns an array of values that contains only the download figures for the specified app, and zeroes for all other apps.

The MAX function then returns the largest value in this array, which corresponds to the best download figure for the specified app.

You can use this formula for each app in your data by replacing "App Name" with the name of each app in turn. This will allow you to quickly determine the best download figures for all of your apps and make comparisons between them.