

- **Introduction to Data Analytics**
- **Modern Data Ecosystem and the Role of Data Analytics**
- **Course Introduction**
- 2 min

- Welcome to this introductory course on Data Analytics, the first in a series of courses designed to prepare you for a career as a junior data analyst.

-

- **Modern Data ecosystem**
- 4

- **Key players in the Data Ecosystem**
- 5

- **Defining Data Analysis**
- 5

Data analysis is the process of gathering, cleaning, analyzing and mining data, interpreting results, and reporting the findings. With data analysis we find patterns within data and correlations between different data points. And it is through these patterns and correlations that insights are generated, and conclusions are drawn. Data analysis helps businesses understand their past performance and informs their decision-making for future actions

- **Viewpoints: What is Data Analytics ?**
- 3

In this video, we will listen to several data professionals talk about how they define data analytics and what this term means to them

- **Data Analytics Vs Data Analysis**

- Reading

The terms Data Analysis and Data Analytics are often used interchangeably, including in this course.

However it is important to note that there is a subtle difference

between the terms and meaning of the words Analysis and Analytics. In fact some people go far as saying that these terms mean different things and should not be used interchangeably. Yes, there is a technical difference...

The dictionary meanings are:

Analysis - detailed examination of the elements or structure of something

Analytics - the systematic computational analysis of data or statistics

Analysis can be done without numbers or data, such as business analysis psycho analysis, etc. Whereas Analytics, even when used without the prefix "Data", almost invariably implies use of data for performing numerical manipulation and inference.

Some experts even say that Data Analysis is based on inferences based on historical data whereas Data Analytics is for predicting future performance. The design team of this course does not subscribe to this view, and you will see why later in the course as you become familiar with the

terms like predictive analytics, prescriptive analytics, etc.

So in this course we take a more liberal view, and use the terms Data Analysis and Data Analytics to mean the same thing. For example, an earlier video is titled Defining Data Analysis, whereas the preceeding video with the viewpoints of several data professionals is titled What is Data Analytics. The difference in these titles is not intentional.

- **Summary and Highlights**

- Reading

In this lesson, you have learned the following information:

The role of a Data Analyst spans across:

- Acquiring data that best serves the use case.
- Preparing and analyzing data to understand what it represents.
- Interpreting and effectively communicating the message to stakeholders who need to act on the findings.
- Ensuring that the process is documented for future reference and repeatability.

In order to play this role successfully, Data Analysts need a mix of technical, functional, and soft skills.

- **Technical**
Skills include varying levels of proficiency in using spreadsheets, statistical tools, visualization tools, programming and querying languages, and the ability to work with different types of data repositories and big data platforms.
- An understanding of Statistics, Analytical techniques, problem-solving, the ability to probe a situation from multiple perspectives, data visualization, and project management skills – all of which come under Functional Skills a Data Analyst needs in order to play an effective role.
- **Soft Skills** include the ability to work collaboratively, communicate effectively, tell a compelling story with data, and garner support and buy-in from stakeholders. Curiosity to explore different pathways and intuition that helps to

give a sense of the future based on past experiences are also essential skills for being a good Data Analyst.

- **Practice Quiz**

Which emerging technology has made it possible for every enterprise to have access to limitless storage and high-performance computing?

Cloud computing

Cloud technologies have made it possible for every enterprise, regardless of its size, to have access to limitless storage and high-performance computing at nominal costs.

Which of the data roles is responsible for extracting, integrating, and organizing data into data repositories?

Data Engineers are responsible for extracting, integrating, and organizing data into data repositories.

When you analyze historical data to predict future outcomes what type of Data Analytics are you performing?

Predictive Analytics is the analysis of historical data to predict future outcomes.

- **Graded Quiz**

A modern data ecosystem includes a network of continually evolving entities. It includes:

Data sources, enterprise data repository, business stakeholders, and tools, applications, and infrastructure to manage data

Data Analysts work within the data ecosystem to:

Gather, clean, mine, and analyze data for deriving insights

The role of a Data Analyst in a data ecosystem is to gather, clean, mine, and analyze data to derive insights.

When we analyze data in order to understand why an event took place, which of the four types of data analytics are we performing?

Diagnostic Analysis

Diagnostic Analysis helps us understand why an event took place—the cause of an outcome.

The first step in the data analysis process is to gain an in-depth understanding of the problem and the desired outcome. What are you seeking answers to at this stage of the data analysis process?

Where you are and where you need to be

As a first step in the Data Analysis process, you are seeking answers to “where you are”, that is, what is the problem that needs to be solved, and “where you need to be”, that is, what is the desired outcome that needs to be achieved.

From the provided list, select the three emerging technologies that are shaping today’s data ecosystem.

Cloud Computing, Machine Learning, and Big Data

- **Discussion Prompt :
Introduce Yourself**

The Data Analyst Role

Responsibilities of Data Analyst

In this video, we learned about the responsibilities and skillsets of a Data Analyst. In the next video, we will walk you through a day in the life of a Data Analyst.

Viewpoints : Qualities & Skills to be a Data Analyst

In this video, we will listen to practicing data professionals talk about the qualities and skills required to become a data analyst

A day in a life of Data Analyst

A day in the life of a Data Analyst can include a number of possibilities — from acquiring data from varied data sources to creating queries for pulling data from data repositories, foraging through rows of data to look for insights, creating reports and dashboards, and interacting with stakeholders for gathering information and presenting the findings, it's a spectrum

View points : Application of Data Analytics

In this video, practicing data professionals talk about some of the applications of data analytics in today's world.

Summary and Highlights

In this lesson, you have learned the following information:

A modern data ecosystem includes a network of interconnected and continually evolving entities that include:

- Data that is available in a host of different formats, structure, and sources.
- Enterprise Data Environment in which raw data is staged so it can be organized, cleaned, and optimized for use by end-users.
- End-users such as business stakeholders, analysts, and programmers who consume data for various purposes.

Emerging technologies such as Cloud Computing, Machine Learning, and Big Data, are continually reshaping the data ecosystem and the possibilities it offers. Data Engineers, Data Analysts, Data Scientists, Business Analysts, and Business Intelligence Analysts, all play a vital role in the ecosystem for deriving insights and business results from data.

Based on the goals and outcomes that need to be

achieved, there are four primary types of Data Analysis:

- Descriptive Analytics, that helps decode “What happened.”
- Diagnostic Analytics, that helps us understand “Why it happened.”
- Predictive Analytics, that analyzes historical data and trends to suggest “What will happen next.”
- Prescriptive Analytics, that prescribes “What should be done next.”

The Data Analysis process involves:

- Developing an understanding of the problem and the desired outcome.
- Setting a clear metric for evaluating outcomes.
- Gathering, cleaning, analyzing, and mining data to interpret results.
- Communicating the findings in ways that impact decision-making.

Practice Quiz

Which of these skills is essential to the role of a Data Analyst?

Statistics

Proficiency in Statistics is essential to the role of a Data Analyst.

What, according to Sivaram Jaladi, goes a long way in lending credibility to your data analysis findings?

Sharing your process of arriving at the findings with your stakeholders

Graded Quiz

Why is proficiency in Statistics an important skill for a Data Analyst?

For identifying patterns and correlations in data

Which of these is one of the soft skills required to be a successful Data Analyst?

Work collaboratively with cross-functional teams

Which of the data analyst functional skills helps research and interpret data, theorize, and make forecasts?

Analytical skills

In “A day in the life of a Data Analyst”, what according to Sivaram Jaladi forms a large part of a Data Analyst’s job?

Cleaning and preparing data

In “A day in the life of a Data Analyst”, what are some of the data points that were useful in analyzing the use case. (Select all that apply)

Serial number of the meters

Serial number of meters was an important data point that helped establish the connection between the meters and overbilling.

Average billing amount of complainants

Correct

The average billing amount of complainants was one of the data points used to check for patterns in subscribers reporting overbilling.

Week 2

THE DATA ECOSYSTEM

The Data Ecosystem and Languages for Data Professionals

Overview of the data analyst ecosystem

A data analyst's ecosystem includes the infrastructure, software, tools, frameworks, and processes used to gather, clean, analyze, mine, and visualize data. In this video, we will go over a quick overview of the ecosystem before going into the details of each of these topics in subsequent videos. We will cover some of the data analytics tools in greater detail in subsequent sections of the course.

Types of data

In the next video, we will learn about the different types of file structures.

Understanding different types of file formats

In this video, we looked at some popular file and data formats. In the next video, we will learn about the different sources of data.

Sources of data

As we touched upon in one of our previous videos, data sources have never been as dynamic and diverse as they are today. In this video, we will look at some common sources such as: Relational Databases, Flat files and XML Datasets, APIs and Web Services, Web Scraping, Data Streams, and Feeds.

Languages of data for professionals

In this video, we will learn about some of the languages relevant to the work of data professionals. These can be categorized as – query languages, programming languages, and shell scripting.

Summary and highlights

In this lesson, you have learned the following information:

A data analyst ecosystem includes the infrastructure, software, tools, frameworks, and processes used to gather, clean, analyze, mine, and visualize data.

Based on how well-defined the structure of the data is, data can be categorized as:

- Structured Data, that is data which is well organized in formats that can be stored in databases.
- Semi-Structured Data, that is data which is partially organized and partially free form.
- Unstructured Data, that is data which can not be organized conventionally into rows and columns.

Data comes in a wide-ranging variety of file formats, such as delimited text files, spreadsheets, XML, PDF, and JSON, each with its own list of benefits and limitations of use.

Data is extracted from multiple data sources, ranging from relational and non-relational databases to APIs, web services, data streams, social platforms, and sensor devices.

Once the data is identified and gathered from different sources, it needs to be staged in a data repository so that it can be prepared for analysis. The type, format, and sources of data influence the type of data repository that can be used.

Data professionals need a host of languages that can help them extract, prepare, and analyze data. These can be classified as:

- Querying languages, such as SQL, used for accessing and manipulating data from databases.
- Programming languages such as Python, R, and Java, for developing applications and controlling application behavior.
- Shell and Scripting languages, such as Unix/Linux Shell, and

PowerShell, for automating repetitive operational tasks.

Practice quiz

Structured Query Language, or SQL, is the standard querying language for what type of data repository?

SQL is the standard querying language for RDBMSs.

In use cases for RDBMS, what is one of the reasons that relational databases are so well suited for OLTP applications?

Support the ability to insert, update, or delete small amounts of data

Which NoSQL database type stores each record and its associated data within a single document and also works well with Analytics platforms?

Document-based NoSQL databases store each record and its associated data within a single document and work well with Analytics platforms.

What type of data repository is used to isolate a subset of data for a particular business function, purpose, or community of users?

Data Mart

A data mart is a sub-section of the data warehouse used to isolate a subset of data for a particular business function, purpose, or community of users.

What does the attribute “Velocity” imply in the context of Big Data?

The speed at which data accumulates

Which of the Big Data processing tools provides distributed storage and processing of Big Data?

Hadoop, a java-based open-source framework, allows distributed storage and processing of large datasets across clusters of computers.

Graded quiz

In the data analyst’s ecosystem, languages are classified by type. What are shell and scripting languages most commonly used for?

Shell and scripting languages are commonly used for automating repetitive operational tasks.

Which of the following is an example of unstructured data?

Video and audio files are examples of unstructured data.

Which one of these file formats is independent of software, hardware, and operating systems, and can be viewed the same way on any device?

PDF

Which data source can return data in plain text, XML, HTML, or JSON among others?

API

According to the video “Languages for Data Professionals,” which of the programming languages supports multiple programming paradigms, such as object-oriented, imperative, functional, and procedural, making it suitable for a wide variety of use cases?

Python

Understanding Data Repositories and Big Data Platforms

Overview of Data Repositories

RDBMS

NoSQL

Data Marts , Data Lakes , ETL and Data Pipelines

Earlier in the course, we examined databases, data warehouses, and big data stores. Now we'll go a little deeper in our exploration of data warehouses, data marts, and data lakes; and also learn about the ETL process and data pipelines

Foundations of Big Data

Big Data Processing Tools

Summary and Highlights

In this lesson, you have learned the following information:

A Data Repository is a general term that refers to data that has been collected, organized, and isolated so that it can be used

for reporting, analytics, and also for archival purposes.

The different types of Data Repositories include:

- Databases, which can be relational or non-relational, each following a set of organizational principles, the types of data they can store, and the tools that can be used to query, organize, and retrieve data.
- Data Warehouses, that consolidate incoming data into one comprehensive storehouse
- Data Marts, that are essentially sub-sections of a data warehouse, built to isolate data for a particular business function or use case.
- Data Lakes, that serve as storage repositories for large amounts of structured, semi-structured, and unstructured data in their native format.
- Big Data Stores, that provide distributed computational and storage infrastructure to store,

scale, and process very large data sets.

ETL, or Extract Transform and Load, Process is an automated process that converts raw data into analysis-ready data by:

- Extracting data from source locations.
- Transforming raw data by cleaning, enriching, standardizing, and validating it.
- Loading the processed data into a destination system or data repository.

Data Pipeline, sometimes used interchangeably with ETL, encompasses the entire journey of moving data from the source to a destination data lake or application, using the ETL process.

Big Data refers to the vast amounts of data that is being produced each moment of every day, by people, tools, and machines. The sheer velocity, volume, and variety of data challenge the tools and systems used for conventional data. These challenges led to the emergence of processing tools and platforms designed specifically for Big Data, such

as Apache Hadoop, Apache Hive, and Apache Spark.

Practice Quiz

What data type is typically found in databases and spreadsheets?

Structured data

Which of these data sources is an example of semi-structured data?

Email

Which one of the provided file formats is commonly used by APIs and Web Services to return data?

JSON

What is one example of the relational databases discussed in the video?

SQL server

Which of the following languages is one of the most popular querying languages in use today?

SQL

Graded Quiz

Data Marts and Data Warehouses have typically been relational, but the emergence of what technology has helped to let these be used for non-relational data?

NoSQL

most significant advantages of an RDBMS?

ACID compliant

Which one of the NoSQL database types uses a graphical model to represent and store data, and is particularly useful for visualizing, analyzing, and finding connections between different pieces of data?

Graph based

Which of the data repositories serves as a pool of raw data and stores large amounts of structured, semi-structured, and unstructured data in their native formats?

Data lakes

What does the attribute “Veracity” imply in the context of Big Data?

Accuracy and conformity of data to facts

Apache Spark is a general-purpose data processing engine designed to extract and process Big Data for a wide range of applications. What is one of its key use cases?

Perform complex analytics in real-time

Week 3 Gathering and Wrangling Data

Gathering Data

Identifying Data for Analysis

At this stage, you have an understanding of the problem and the desired outcome—you know “Where you are” and “Where you want to be.” You also have a well-defined metric—you know “What will be measured,” and “How it will be measured.”

Data Sources

How to Gather and Import Data

In this video, we will learn about the different methods and tools available for gathering data from the data sources

discussed earlier in the course —such as databases, the web, sensor data, data exchanges, and several other sources leveraged for specific data needs.

Summary and Highlights

- The process of identifying data begins by determining the information that needs to be collected, which in turn is determined by the goal you seek to achieve.
- Having identified the data, your next step is to identify the sources from which you will extract the required data and define a plan for data collection. Decisions regarding the timeframe over which you need your data set, and how much data would suffice for arriving at a credible analysis also weigh in at this stage.
- Data Sources can be internal or external to the organization, and they can be primary, secondary, or third-party, depending on whether you are obtaining the data directly from the original source, retrieving it from externally available data sources, or purchasing it from data aggregators.
- Some of the data sources from which you could be gathering data include databases, the web, social media, interactive platforms, sensor devices, data exchanges, surveys and observation studies.
- Data that has been identified and gathered from the various data sources is combined using a variety of tools and methods to provide a single interface using which data can be queried and manipulated.
- The data you identify, the source of that data, and the practices you employ for gathering the data have implications for quality, security, and privacy, which need to be considered at this stage.
-

Practice Quiz

What are the requirements in order for data to be reliable? (Select all that apply)

Data should be free of all errors

Data should be relevant

What type of data is produced by wearable devices, smart buildings, and medical devices?

Sensor data

What type of data is semi-structured and has some organizational properties but not a rigid schema?

Email

Graded Quiz

What are some of the steps in the process of “Identifying Data”? (Select all that apply)

Determine the information you want to collect

Define a plan for collecting data

What type of data refers to information obtained directly from the source?

Primary data

Question 3

Web scraping is used to extract what type of data?

Text, videos, and images

Data obtained from an organization’s internal CRM, HR, and workflow applications is classified as:

Primary data

Which of the provided options offers simple commands to specify what is to be retrieved from a relational database?

SQL

Week 3 : Wrangling Data

What is Data Wrangling

Tools for Data Wrangling

In this video, we will look at some of the popularly used data wrangling software and tools, such as: Excel Power Query / Spreadsheets, OpenRefine, Google DataPrep, Watson Studio Refinery, Trifacta Wrangler, Python and R. Let’s begin with the most basic software used for manual wrangling—Spreadsheets.

Data Cleaning

Viewpoints : Data Preparation and Reliability

In this segment, data professionals share what portion of their job involves gathering, cleaning, and preparing data for analysis.

Summary Highlights

Once the data you identified is gathered and imported, your next step is to make it analysis-ready. This is where the process of Data Wrangling, or Data Munging, comes in.

Data Wrangling is an iterative process that involves data exploration, transformation, and validation.

Transformation of raw data includes the tasks you undertake to:

- Structurally manipulate and combine the data using Joins and Unions.
- Normalize data, that is, clean the database of unused and redundant data.
- Denormalize data, that is, combine data from multiple tables into a single table so that it can be queried faster.
- Clean data, which involves profiling data to uncover quality issues, visualizing

data to spot outliers, and fixing issues such as missing values, duplicate data, irrelevant data, inconsistent formats, syntax errors, and outliers.

- Enrich data, which involves considering additional data points that could add value to the existing data set and lead to a more meaningful analysis.

A variety of software and tools are available for the Data Wrangling process. Some of the popularly used ones include Excel Power

Query, Spreadsheets, OpenRefine, Google DataPrep, Watson Studio

Refinery, Trifacta Wrangler, Python, and R, each with their own set of characteristics, strengths, limitations, and applications.

Practice Quiz

What is one of the common structural transformations used for combining data from one or more tables?

Joins are one of the common structural transformations used for combining data from multiple tables.

What tool allows you to discover, cleanse, and transform data with built-in operations?

Watson Studio Refinery has built-in features that allow you to discover, cleanse, and transform data.

What is data called that does not fit within the context of the use case?

Data that does not fit within the context of the use case is called irrelevant data.

Graded Quiz

What does a typical data wrangling workflow include?

Validating the quality of transformed data is an essential step in a data wrangling workflow.

OpenRefine is an open-source tool that allows you to:

Using OpenRefine, you can transform data into a wide variety of formats such as TSV, CSV, XLS, XML, and JSON.

What is one of the steps in a typical data cleaning workflow?

Inspecting data to detect issues and errors is one of the first steps in a typical data cleaning workflow.

When you're combining rows of data from multiple source tables into a single table, what kind of data transformation are you performing?

Unions are a common structural transformation used for combining rows of data from multiple source tables.

When you detect a value in your data set that is vastly different from other observations in the same data set, what would you report that as?

Outliers are values in your data set that may be vastly different from other values in the same data field.

Week 4

Analysing and mining data

overview of statistical analysis

Before we understand Statistical Analysis, its relation to Data Analysis, and specifically data mining, let's first examine what Statistics is. Statistics is a branch of mathematics dealing with the collection, analysis, interpretation, and presentation of numerical or quantitative data

what is data mining

Data mining essentially helps separate the noise from the real information and helps businesses focus their energies on only what is relevant.

tools for Data Mining

In this video, we will learn about some of the commonly used software and tools for data mining, such as: Spreadsheets, R-Language, Python, IBM SPSS Statistics, IBM Watson Studio; and SAS. Spreadsheets, such as Microsoft Excel and Google Sheets, are commonly used for performing basic data mining tasks.

summary and highlights

Statistics is a branch of mathematics dealing with the collection, analysis, interpretation, and presentation of numerical or quantitative data.

Statistical Analysis involves the use of statistical methods in order to develop an understanding of what the data represents.

Statistical Analysis can be:

- Descriptive; that which provides a summary of what the data represents. Common measures include Central Tendency, Dispersion, and Skewness.
- Inferential; that which involves making inferences, or generalizations, about data. Common measures include Hypothesis Testing, Confidence Intervals, and Regression Analysis.

Data Mining, simply put, is the process of extracting knowledge from data. It involves the use of pattern recognition technologies, statistical analysis, and mathematical techniques, in order to identify correlations, patterns, variations, and trends in data.

There are several techniques that can help mine data, such as, classifying attributes of data, clustering data into groups, establishing relationships between events, variables, and input and output.

A variety of software and tools are available for analyzing and

mining data. Some of the popularly used ones include Spreadsheets, R-Language, Python, IBM SPSS Statistics, IBM Watson Studio, and SAS, each with their own set of characteristics, strengths, limitations, and applications.

practice quiz

What is one of the common measures of Central Tendency?

Mean

Mean, or the mathematical average

What technique is used to help identify the nature of the relationship between two variables?

Regression

What Python libraries are commonly used for data mining?

Pandas

NumPy

graded quiz

What is a branch of mathematics dealing with the collection, analysis,

interpretation, and presentation of numerical or quantitative data?

Statistics

Data Mining is the process of extracting knowledge from data.

R has libraries explicitly built for data mining operations such as classification of data.

When you're calculating the middle value of a data field in a data set, what are you really calculating?

Median

What is the general tendency of a set of data to change over time called?

Trend

communicating data analysis findings

overview of communicating and sharing data analysis findings

viewpoints storytelling and data analysis

In this video, we will listen to data professionals talk about

the role storytelling plays and the life of a data analyst.

introduction to data visualisation

introduction to visualisation and Dashboarding software

In this video, we will look at some of the most commonly used data visualization software and tools. These include: Spreadsheets, Jupyter Notebook and Python libraries, R-Studio and R-Shiny, IBM Cognos Analytics, Tableau and Microsoft Power BI.

viewpoints visualisation tools

In this video, data professionals talk about the visualization tools they rely on the most and why.

summary and highlights

In this lesson, you have learned the following information:

Data has value through the stories that it tells. In order to communicate your findings impactfully, you need to:

- Ensure that your audience is able to trust you, understand you, and relate to your findings and insights.

- Establish the credibility of your findings.
- Present the data within a structured narrative.
- Support your communication with strong visualizations so that the message is clear and concise, and drives your audience to take action.

Data visualization is the discipline of communicating information through the use of visual elements such as graphs, charts, and maps. The goal of visualizing data is to make information easy to comprehend, interpret, and retain.

For data visualization to be of value, you need to:

- Think about the key takeaway for your audience.
- Anticipate their information needs and questions, and then plan the visualization that delivers your message clearly and impactfully.

There are several types of graphs and charts available for you to be able to plot any kind of data, such as bar charts,

column charts, pie charts, and line charts.

You can also use data visualization to build dashboards. Dashboards organize and display reports and visualizations coming from multiple data sources into a single graphical interface. They are easy to comprehend and allow you to generate reports on the go.

When deciding which tools to use for data visualization, you need to consider the ease-of-use and purpose of the visualization. Some of the popularly used tools include Spreadsheets, Jupyter Notebook, Python libraries, R-Studio and R-Shiny, IBM Cognos Analytics, Tableau, and Power BI.

practice quiz

Data visualizations such as graphs and charts are a great way to bring data to life.

Question 2

You can use dashboards to present operational data such as daily progress data, as well as analytical data, such as the overall health of a business function.

Google Sheets is preferred over other spreadsheets when multiple users need to collaborate.

graded Quiz

A presentation is not a data dump”. What is the one thing you would do to ensure your presentation is not a data dump?

Include only that information as is needed to address the business problem

What is the discipline of communicating information through the use of visual elements?

Data visualization

Matplotlib is a widely used Python data visualization library.

What is the goal of Data Visualization?

The goal of Data Visualization is to make information easy to comprehend, interpret, and retain.

What can you do to help your audience trust you?

Share your data sources, hypotheses, and validations

Week 5 : Career Opportunities and Data Analysis in Action

Opportunities and Learning Paths

Carrier opportunity in data analysis

viewpoints career options for data professionals

In this video, we will listen to practicing data professionals talk about the various career options available in this field.

viewpoints get into data profession

In this video, we will listen to data professionals talk about how they got into this profession

viewpoints advice for aspiring data analyst

In this video, we will listen to data professionals giving advice to aspiring data analysts.

viewpoints what do employers look for in a data analyst

In this video, we will listen to data professionals talk about what employers look for in a Data Analyst

view points woman and data professions

In this video, we will listen to women share their experience of being a data professional, and their advice to women aspiring to enter this field.

many path to data analysis

summary and highlights

Data Analyst roles are sought after in every industry, be it

Banking and Finance,
Insurance, Healthcare, Retail,
or Information Technology.

Currently, the demand for skilled data analysts far outweighs the supply, which means companies are willing to pay a premium to hire skilled data analysts.

Data Analyst job roles can be broadly classified as follows:

- Data Analyst Specialist roles - On this path, you start as a Junior Data Analyst and move up to the level of a Principal Analyst by continually advancing your technical, statistical, and analytical skills from a foundational level to an expert level.
- Domain Specialist roles - These roles are for you if you have acquired specialization in a specific domain and want to work your way up to be seen as an authority in your domain.
- Analytics-enabled job roles - These roles include jobs where having analytic skills can up-level your performance and

differentiate you from your peers.

Other Data Professions - There are several other roles in a modern data ecosystem, such as Data Engineer, Big Data Engineer, Data Scientist, Business Analyst, or Business Intelligence Analyst. If you upskill yourself based on the required skills, you can transition into these roles.

There are several paths you can consider in order to gain entry into the Data Analyst field. These include:

- An academic degree in Data Analytics or disciplines such as Statistics and Computer Science.
- Online multi-course specializations offered by learning platforms such as Coursera, edX, and Udacity.
- Mid-career transition into Data Analysis by upskilling yourself. If you have a technical background, for example, you can focus on developing the technical skills specific to Data Analysis. If you do not have a technical

background, you can plan to skill your self in some basic technologies and then work your way up from an entry-level position.

practice quiz

On the Data Analyst Specialist path, you could be starting your career as an Associate or Junior Data Analyst and working your way up to a Principal Analyst role. What are some of the factors that influence your growth on this path?

The experience and exposure you gain in the different areas within Data Analysis

Skills such as problem-solving, communication, and storytelling are critical to the role of a Data Analyst. And like most soft skills, you're either good at them, or you're not; these skills cannot be acquired over time.

False

These skills can be acquired over time.

graded quiz

Which of the following statement describes Data Analyst Specialist Roles?

Analysts who advance technical, statistical, and analytical skills, over time, to expert levels

A Principal Data Analyst is responsible for:

Establishing processes in the team

Job roles such as Project Managers, Marketing Managers, and HR Managers, can achieve greater efficiency and effectiveness in their current roles by acquiring data analysis skills, and are therefore known as analytics-enabled job roles.

Which of these is essential for getting started and growing as a Data Analyst?

Love for numbers, a curious mind, and openness to learn

Question 5

What Data Analysis roles may be best suited for people with little or no technical training?

Functional analyst

Week 5 : Final Assignment

Final Exam :

Introduce yourself.
1 point

I am undergraduate student in USA, Who aspires to undertake a technology track career path in Data Science, A Data Analyst or A Data Scientist.

Why do you want to learn about Data Analytics?
1 point

A data analyst's job is to take data and use it to help companies make better business decisions. I'm good with numbers, collecting data, and market research. I chose this role because it encompasses the skills I'm good at, and I find

data and marketing research interesting

In the Week 1 video titled "What is Data Analytics", Asha Barnes defines Data Analysis as "the use of information around you to make decisions" and shares examples of how we use data analysis to make everyday decisions. What is one of the examples she shares?
Using the weather report to decide what to wear and what activities you can do

In Week 1 video titled "Key Players in the Data Ecosystem", which of these is shared as an example of the kind of problems you can solve using Data Analysis?
1 point

Understanding how the recent marketing campaigns have impacted the sales of your products

5. In Week 4 video titled "Overview of Communicating and Sharing Data Analysis Findings", you learned that "The success of your

communication depends on how well others can understand and trust your insights to take further action.”

What are the key elements of a successful communication?

1 point

Data, Narrative, and Visuals

Using Data Analysis for
Detecting Credit Card Fraud I
Coursera PDF

Peer Graded Assignment Q
PDF