**Customer engineer:**  *A customer engineer at google is the bridge between the technical and business components or aspects of the corporate America.*

**Note: Data driven career.**

Data professionals are so valuable to their companies. They determine which data streams are most important to specific business projects, challenges and initiatives. They identify key goals for the future and they give their organizations the ability to take meaningful action by reimagining processes and improving operations.

**Data careers** can be loosely categorized into two complementary types of work:

**technical and strategic**

**Technical data professionals.**

Some examples of these professionals are machine learning engineers and statisticians.

These people provide high effort solutions to specific problems. Through their expertise in mathematics, statistics and computing, they build models and make predictions. Using tools such as R and python, they help their teams extract value from business data sets. The result is a solution that has a direct positive impact. Another highly technical role is the expert data analyst whose work involves exploring vast and complex datasets to identify directions worth pursuing in the first place.

They ensure that an organization's data science efforts are directed as efficiently as possible, bridging the gap between other technical data professionals and the strategic work.

**Strategic Data professionals.**

These people include business intelligence professionals and technical project managers to name a couple. Strategic data professionals use their skills to interpret information that affects an organization's operations, finance research and development and so much more. Their work aligns closely to the overall business strategy and involves seeking solutions to problems through data analytics. In short, strategic data professionals maximize information to guide how a business works.

.

**Note**: technical and strategic roles for data analytics professionals. Data analysts in technical roles transform raw data into something useful for decision-making. Strategic data analytics professionals focus on maximizing information to guide the businesses they are working for.

**Key attributes of Data Scientist and Data Analyst**

* **What they do:** Uncover trends, patterns, and insights from data
* **How they do it:** Employ advanced modeling and statistical analytics techniques
* Entry-level data scientist or analyst positions may not require considerable data analysis experience and be less restrictive about requiring those hired to hold a traditional degree.

### ****Sample job titles****

* Data scientist
* Marketing analyst
* Data analyst
* AI analyst
* Business analyst

## **Data management and infrastructure**

Data professionals that work in data management and infrastructural roles are primarily responsible for the systems that distribute data and maintain its integrity.

### ****Key attributes****

* **What they do:** Manage data sources and the overall data infrastructure
* **How they do it:** Work with the tools and databases used to manage data within a business

### ****Sample job titles****

* Data engineer
* Technology engineer
* Data manager
* Data steward
* IT architect

## **Business intelligence**

Data analytics and business intelligence share a lot of commonalities. Both fields have professionals that use data to create insights that inform decision-making. A major difference is that business intelligence is more focused on creating processes and information channels that transform relevant data. Business intelligence professionals create tables, reports, and dashboards that empower stakeholders, giving them access to the data they need to inform the entire decision-making process on a continual basis. These roles often serve as a complement to core data analytics/data science professionals.

### ****Key attributes****

* **What they do:** Perform predictive analysis that enables organizations to determine likely future trends
* **How they do it:** Create tables, reports and dashboards that empower their organization

### ****Sample job titles****

* BI architect
* BI analyst
* BI solution developer
* BI software engineer
* Data viz & BI analyst

## **Product development teams**

The professionals in these roles manage a portfolio of customer and stakeholder analytic projects and initiatives. They often manage the analytical strategy for the organization. In these roles, experience is most likely required, and responsibilities are larger and more global.

### ****Key attributes****

* **What they do:** Manage analytical strategy within a project team
* **How they do it:** They are less hands-on with data analysis, serving as the person a data scientist or analysts would report to

### ****Sample job titles****

* Product manager
* Product developer
* Product lead
* Digital product manager
* Customer product manager

## **C-suite**

This classification of roles covers high-ranking executives within an organization. The ‘C’ in c-suite stands for chief. In general, there's a trend for the c-suite to build data-driven decision making into their processes. Individuals filling these roles within organizational leadership teams are expected to be familiar with data and analytics.

### ****Key attributes****

* **What they do:** Responsible for data and data professionals across an entire organization
* **How they do it:** They are decision makers found at the top end of a company’s hierarchy

### ****Examples of job titles****

* Chief marketing officer
* Chief data officer
* Chief analytics officer
* Chief information officer
* Chief data scientist

| **Industry** | **Overview** | **How data is used** |
| --- | --- | --- |
| App-driven business (sharing economy service) | Facilitates users acquiring, providing, or sharing access to goods and services, often through online or app-based communities | * Maintaining functioning mobile applications * Delivering customized content based on user history including discounts * Using machine learning models to send notifications at key times or even locations |
| Automotive | Includes industries associated with the production, wholesaling, retailing, and maintenance of motor vehicles | * Gaining greater control over their supply chains * Improving production line performance, and designing new and more efficient vehicles * Enhancing vehicle safety and new features |
| Cybersecurity | Protects networks, devices, and data from unauthorized access or criminal use and the practice of maintaining confidentiality, integrity, and availability of information | * Locating weak points within networks and systems using predictive analytics * Defending against security attacks * Detecting data breaches through logic, models, and data tools * Improving the ability to identify attacks and respond to them with Artificial Intelligence (AI) |
| Digital marketing | Assists in advertising and promotional efforts of companies using the internet and online technologies | * Translating customer interaction into actionable business data * Predicting user behaviors to personalize content and offers * Identifying patterns and trends that guide innovations * Determining the return on investment (ROI) of marketing efforts |
| Energy | Includes companies that explore, produce, refine, market, store, and transport both renewable and non-renewable energy resources | * Analyzing real-time data from power systems and monitoring devices * Optimizing technologies, monitoring power grids, and predicting failures * Preventing accidents and malfunctions |
| Gaming | Hosts an estimated 2.7 billion gamers worldwide, facilitating the interaction of players across the globe | * Designing world-building and character creation systems * Monitoring character engagement and how the environment reacts to player input * Optimizing game-play by identifying potential new features or upgrades * Regulating in-game purchases and fraud detection systems * Personalizing marketing campaigns |
| Streaming media and entertainment | Provides access to live and recorded content on-demand, delivered via the internet to computers, smart devices, and mobile devices | * Analyzing and monitoring user interactions to better understand customer sentiment * Matching users with advertisers with real-time analytics * Guiding future content decisions * Personalizing marketing campaigns |
| Telecommunications | Primarily involves operating and providing access to facilities for the transmission of voice, data, text, sound, and video | * Assisting the deployment, optimization, and predictive maintenance of telecommunications networks * Optimizing pricing models * Targeting advertisement and incentive campaigns, as well as detecting fraudulent activity * Analyzing customer data to customize subscriber plans |
| Travel and tourism | Encompasses a variety of services from transportation, accommodations, attractions, booking, and much more | * Marketing to individuals based on their previous travel or searched destinations * Directing machine learning systems that can adjust a traveler’s itinerary based on set factors including weather and availability * Generating recommendations based on personal preferences and location-based discounts * Managing reservations and processing transactions |

## **Data trends for the future**

As you can already tell, data analytics is an emerging field with a wide range of exciting opportunities. And, even more exciting is the fact that big data is getting bigger. The need for people to understand, prioritize, manage, and analyze that information is not slowing down in any industry. Businesses will continue to rely on data-driven decision-making, fueled by both simple trend analyses and more complex techniques like predictive modeling and forecasting.

Additionally, more companies are storing all of their raw data within large repositories accessible across the organization. As companies become more reliant on insights generated by this data, there will be many opportunities for data analytics professionals to use their skills and knowledge to organize that information and make it useful.

## **Innovative technologies**

Innovations in accessing this data are leading to new approaches in making data interconnected—meaning that there are still new and evolving ways businesses in different industries are going to use data in the future. Technology is also ever-changing and adapting to these new needs. Because of this, there will always be exciting new tools and data solutions to explore.

Artificial intelligence will continue to have a large impact on business, helping to streamline many areas. For example, an increase in sales is understood to be a direct result of forecasting product demand. Artificial intelligence helps companies ensure warehouse supply, keep items in stock, reduce delivery time, and boost operational efficiency through automating processes.

Additionally, artificial intelligence will combine with machine learning, business intelligence, and automation to deliver more personalized services to customers.

Offering additional services will push forward innovation, bringing computer applications and the sources of stored data closer together physically. This concept is referred to as **edge computing**. By closing the gap between data and computation, speed improves. This results in greater support of real-time analytics and the automation necessary to support the increasing number of devices that are becoming linked through the internet of things.

An increasing number of data analytics tasks will be automated by creating, managing, and analyzing data in edge environments. Artificial intelligence and machine learning systems are only as equitable and inclusive as the people who create and train these systems. You will learn more about how you, as a data analytics professional, will need to take steps to ensure equity in the future.

Note: PII personal identifiable information.

**Sample**: A good sample is a segment of a population that is representative of that entire population.

## **Privacy matters**

Data privacy means preserving a data subject’s information and activity any time a data transaction occurs. This is also called information privacy or data protection. Data privacy is concerned with the access, use, and collection of personal data. For the people whose data is being collected, this means they have the right to:

* Protection from unauthorized access to their private data
* Freedom from inappropriate use of their data
* The right to inspect, update, or correct their data
* Ability to give consent to data collection
* Legal right to access the data

In order to maintain these rights, businesses and organizations have to put privacy measures in place to protect individuals’ data. This is also a matter of trust. The public’s ability to trust companies with personal data is important. It’s what makes people want to use a company’s product, share their information, and more.

## **Protecting privacy with data anonymization**

Organizations use a lot of different measures to protect the privacy of their data subjects, like incorporating access permissions to ensure that only the people who are supposed to access that information can do so. Another key strategy to maintaining privacy is data anonymization.

**Data anonymization** is the process of protecting people's private or sensitive data by eliminating PII. Typically, data anonymization involves blanking, hashing, or masking personal information, often by using fixed-length codes to represent data columns, or hiding data with altered values.

Data professionals can take additional measures to protect users and their data. **Data aggregation**, for example, is the process of collecting and combining details from a significant number of users in terms of totals or summary. Aggregating data ensures that information contained within datasets is shown in groups; when coupled with other anonymization techniques, data professionals can ensure compliance with data privacy and anonymization standards.

Data anonymization is used in just about every industry. As a data analytics professional, you probably won’t personally be performing anonymization, but it’s useful to understand what kinds of data are often anonymized before you start working with it. This data might include:

* Telephone numbers
* Names
* License plates and license numbers
* Social security numbers
* IP addresses
* Medical records
* Email addresses
* Photographs
* Account numbers

Imagine a world where we all had access to each other’s addresses, account numbers, and other identifiable information. That would invade a lot of people’s privacy and make the world less safe. Data anonymization is one of the ways we can help keep data private and secure!

**Data stewardship** is the practice of ensuring that data is accessible, usable, and safe. Making data stewardship a normal part of your work habits will benefit everyone who relies on your analysis, both inside and outside of your organization.

**The GDPR (European Union law)** is described on their website as the toughest privacy and security law in the world. It imposes obligations onto organizations anywhere, so long as they target or collect data related to people in the European Union.

**The LGPD** **(Brazil’s general law for the protection of personal data)** is a data protection law that governs how companies collect, use, disclose, and process personal data belonging to people in Brazil. LGPD applies to companies that process data about individuals in Brazil.

**The CCPA (Privacy rights for California consumers)** gives consumers more control over the personal information that businesses collect about them. These regulations provide guidance on how to implement the law.

* Additionally, states like Virginia, Colorado, New York, Utah, and Connecticut have enacted similar legislation to protect consumer privacy in their states.

**Glossary Terms**

**Aggregate information**:Data from a significant number of users that has eliminated personal information

**Artificial intelligence (AI):** Refers to computer systems able to perform tasks that normally require human intelligence

**Data anonymization**: The process of protecting people's private or sensitive data by eliminating PII

**Data stewardship**: The practices of an organization that ensure that data is accessible, usable, and safe

**Edge computing**: A way of distributing computational tasks over a bunch of nearby processors (i.e., computers) that is good for speed and resiliency and does not depend on a single source of computational power

**Hackathon:** An event where programmers and data professionals come together and work on a project

**Nonprofit:** A group organized for purposes other than generating profit; often aims to further a social cause or provide a benefit to the public

**Open data**: Data that is available to the public and free to use, with guidance on how to navigate the datasets and acknowledge the source

**Personally identifiable information (PII)**: Information that permits the identity of an individual to be inferred by either direct or indirect means

**Sample:** A segment of a population that is representative of the entire population

**Data professional**: Any individual who works with data and/or has data skills

**Data science**: The discipline of making data useful

**Data stewardship:** The practices of an organization that ensure that data is accessible, usable, and safe

**Edge computing**: A way of distributing computational tasks over a bunch of nearby processors (i.e., computers) that is good for speed and resiliency and does not depend on a single source of computational power

**Jupyter Notebook:** An open-source web application used to create and share documents that contain live code, equations, visualizations, and narrative text

**Machine learning:** The use and development of algorithms and statistical models to teach computer systems to analyze patterns in data

**Metrics**: Methods and criteria used to evaluate data

**Python**: A general-purpose programming language

**Tableau**: A business intelligence and analytics platform that helps people visualize, understand, and make decisions with data