

## Data roles and services

### Introduction

Over the last decade, the amount of data that systems and devices generate has increased significantly. Because of this increase, new technologies, roles, and approaches to working with data are affecting data professionals. Data professionals typically fulfill different roles when **managing, using, and controlling data**. In this module, you'll learn about the various roles that organizations often apply to data professionals, the tasks and responsibilities associated with these roles, and the Microsoft Azure services used to perform them.

### Learning objectives

In this module you will learn how to:

- Identify common data professional roles
- Identify common cloud services used by data professionals

### Data professional roles

There's a wide variety of roles involved in managing, controlling, and using data. Some roles are business-oriented, some involve more engineering, some focus on research, and some are hybrid roles that combine different aspects of data management. Your organization may define roles differently, or give them different names, but the roles described in this unit encapsulate the most common division of tasks and responsibilities.

The three key job roles that deal with data in most organizations are:

- **Database administrators** **manage databases**, assigning permissions to users, storing **backup copies of data and restore data in the event of a failure**.
- **Data engineers** **manage infrastructure** and processes for data integration across the organization, applying data cleaning routines, **identifying data governance rules**, and implementing pipelines to transfer and transform data between systems.
- **Data analysts** explore and **analyze data to create visualizations and charts** that enable organizations to make informed decisions.

**NOTE:** The job *roles* define differentiated tasks and responsibilities. **In some organizations, the same person might perform multiple roles**; so in their role as database administrator they might provision a transactional database, and then in their role as a **data engineer they might create a pipeline to transfer data from the database to a data warehouse for analysis**.

## Database Administrator

A database administrator is responsible for the **design, implementation, maintenance**, and operational aspects of on-premises and cloud-based database systems. They're responsible for the overall availability and **consistent performance and optimizations of databases**. They work with stakeholders to implement policies, tools, and processes for backup and recovery plans to recover following a natural disaster or human-made error.

The database administrator is **also responsible for managing the security of the data** in the database, granting privileges over the data, granting or denying access to users as appropriate.

## Data Engineer

A data engineer collaborates with stakeholders to design and implement data-related workloads, including data ingestion **pipelines, cleansing and transformation activities, and data stores for analytical workloads**. They use a wide range of data platform technologies, including relational and non-relational databases, file stores, and data streams.

They're also responsible for ensuring that the privacy of data is maintained within the cloud and spanning from on-premises to the cloud data stores. They own the management and monitoring of data pipelines to ensure that data loads perform as expected.

## Data Analyst

A data analyst enables businesses to maximize the value of their data assets. They're responsible for exploring data to identify trends and relationships, designing and building analytical models, and enabling advanced analytics capabilities through reports and visualizations.

A data analyst processes raw data into relevant insights based on identified business requirements to deliver relevant insights.

**NOTE:** The roles described here represent the key data-related roles found in most medium to large organizations. There are additional data-related roles not mentioned here, such as *data scientist* and *data architect*; and there are other technical professionals that work with data, including *application developers* and *software engineers*.