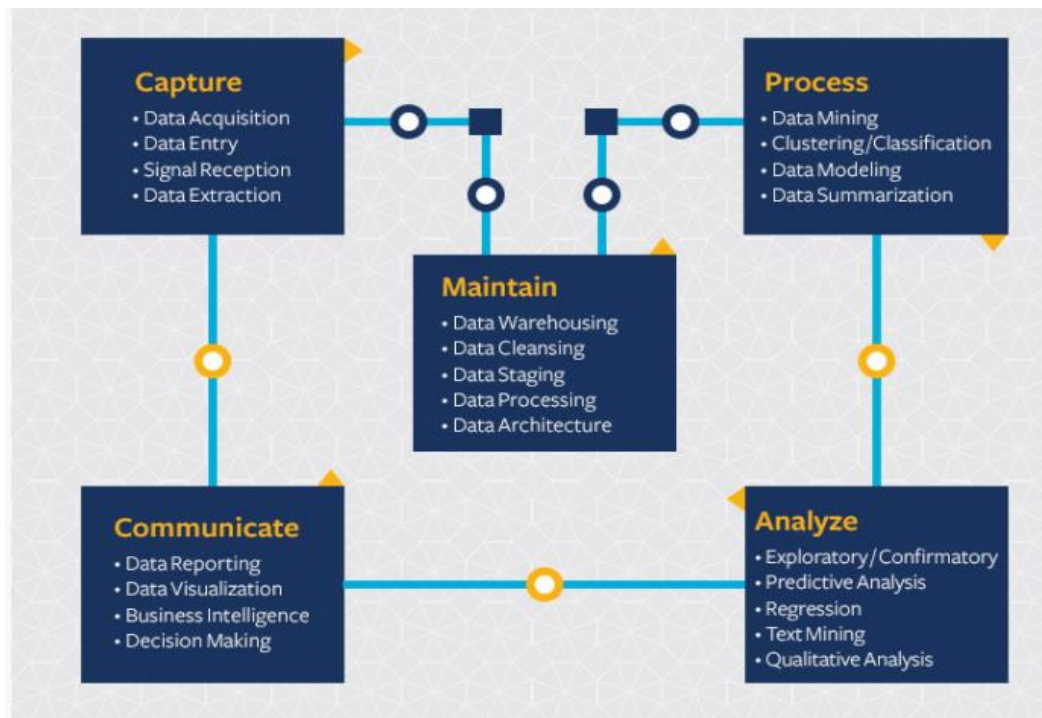# Data science

What is Data Science?

Data science continues to evolve as one of the most promising and in-demand career paths for skilled professionals. Today, successful data professionals understand that they

must advance past the traditional skills of analyzing large amounts of data, data mining, and programming skills. In order to uncover useful intelligence for their organizations, data scientists must master the full spectrum of the data science life cycle and possess a level of flexibility and understanding to maximize returns at each phase of the process.

The Data Science Life Cycle:-

Capture
- Data Acquisition
- Data Entry
- Signal Reception
- Data Extraction

Process
- Data Mining
- Clustering/Classification
- Data Modeling
- Data Summarization

Maintain
- Data Warehousing
- Data Cleansing
- Data Staging
- Data Processing
- Data Architecture

Communicate
- Data Reporting
- Data Visualization
- Business Intelligence
- Decision Making

Analyze
- Exploratory/Confirmatory
- Predictive Analysis
- Regression
- Text Mining
- Qualitative Analysis

The term "data scientist" was coined as recently as 2008 when companies realized the need for data professionals who are skilled in organizing and analyzing massive amounts of data.[1] In a 2009 McKinsey&Company article, Hal Varian, Google's chief economist and UC Berkeley professor of information sciences, business, and economics, predicted the importance of adapting to technology's influence and reconfiguration of different industries.

"The ability to take data — to be able to understand it, to process it, to extract value from it, to visualize it, to communicate it — that's going to be a hugely important skill in the next decades."

Effective data scientists are able to identify relevant questions, collect data from a multitude of different data sources, organize the information, translate results into solutions, and communicate their findings in a way that positively affects business

decisions. These skills are required in almost all industries, causing skilled data scientists to be increasingly valuable to companies.

WHAT DOES A DATA SCIENTIST DO?

In the past decade, data scientists have become necessary assets and are present in almost all organizations. These professionals are well-rounded, data-driven individuals with high-level technical skills who are capable of building complex quantitative algorithms to organize and synthesize large amounts of information used to answer questions and drive strategy in their organization. This is coupled with the experience in communication and leadership needed to deliver tangible results to various stakeholders across an organization or business.

Data scientists need to be curious and result-oriented, with exceptional industry-specific knowledge and communication skills that allow them to explain highly technical results to their non-technical counterparts. They possess a strong quantitative background in statistics and linear algebra as well as programming knowledge with focuses in data warehousing, mining, and modeling to build and analyze algorithms.

They must also be able to utilize key technical tools and skills, including:

R	Apache Spark	Apache Pig
Python	NoSQL databases	Tableau
Apache Hadoop	Cloud computing	iPython notebooks
MapReduce	D3	GitHub

WHY BECOME A DATA SCIENTIST

Glassdoor ranked data scientist among the top three jobs in America since 2016. As increasing amounts of data become more accessible, large tech companies are no longer the only ones in need of data scientists. The growing demand for data science professionals across industries, big and small, is being challenged by a shortage of qualified candidates available to fill the open positions.

The need for data scientists shows no sign of slowing down in the coming years. LinkedIn listed data scientist as one of the most promising jobs in 2021, along with multiple data-science-related skills as the most in-demand by companies.

The statistics listed below represent the significant and growing demand for data scientists.

28%	5,971	$127,500	#2
Demand Increase by 2026[7]	Number of Job Openings	Average Base Salary	Best Job in America 2021

WHERE DO YOU FIT IN DATA SCIENCE

Data is everywhere and expansive. A variety of terms related to mining, cleaning, analyzing, and interpreting data are often used interchangeably, but they can actually involve different skill sets and complexity of data.

Data Scientist

Data scientists examine which questions need answering and where to find the related data. They have business acumen and analytical skills as well as the ability to mine, clean, and present data. Businesses use data scientists to source, manage, and analyze large amounts of unstructured data. Results are then synthesized and communicated to key stakeholders to drive strategic decision-making in the organization.

Skills needed: Programming skills (SAS, R, Python), statistical and mathematical skills, storytelling and data visualization, Hadoop, SQL, machine learning

Data Analyst

Data analysts bridge the gap between data scientists and business analysts. They are provided with the questions that need answering from an organization and then organize and analyze data to find results that align with high-level business strategy. Data analysts are responsible for translating technical analysis to qualitative action items and effectively communicating their findings to diverse stakeholders.

Skills needed: Programming skills (SAS, R, Python), statistical and mathematical skills, data wrangling, data visualization

Data Engineer

Data engineers manage exponential amounts of rapidly changing data. They focus on the development, deployment, management, and optimization of data pipelines and infrastructure to transform and transfer data to data scientists for querying.

Skills needed: Programming languages (Java, Scala), NoSQL databases (MongoDB, Cassandra DB), frameworks (Apache Hadoop

DATA SCIENCE CAREER OUTLOOK AND SALARY OPPORTUNITIES

Data science professionals are rewarded for their highly technical skill set with competitive salaries and great job opportunities at big and small companies in most industries. With almost 6,000 open positions listed on Glassdoor, data science professionals with the appropriate experience and education have the opportunity to make their mark in some of the most forward-thinking companies in the world.

Below are the average base salaries for the following positions:

Data analyst: $69,517	Data scientist: $117,212	Senior data scientist: $142,258	Data engineer: $112,493

Gaining specialized skills within the data science field can distinguish data scientists even further. For example, machine learning experts utilize high-level programming skills to create algorithms that continuously gather data and automatically adjust their function to be more effective..

Thank You...