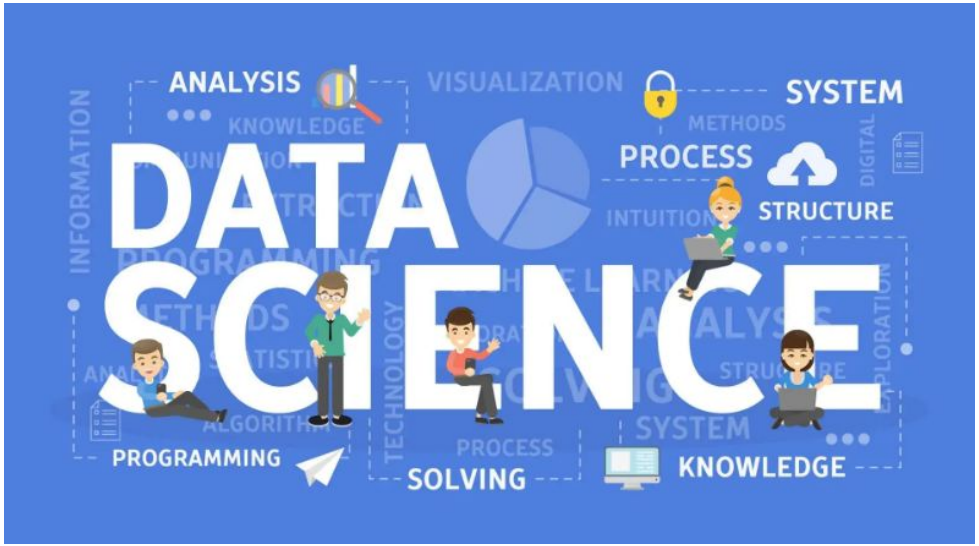
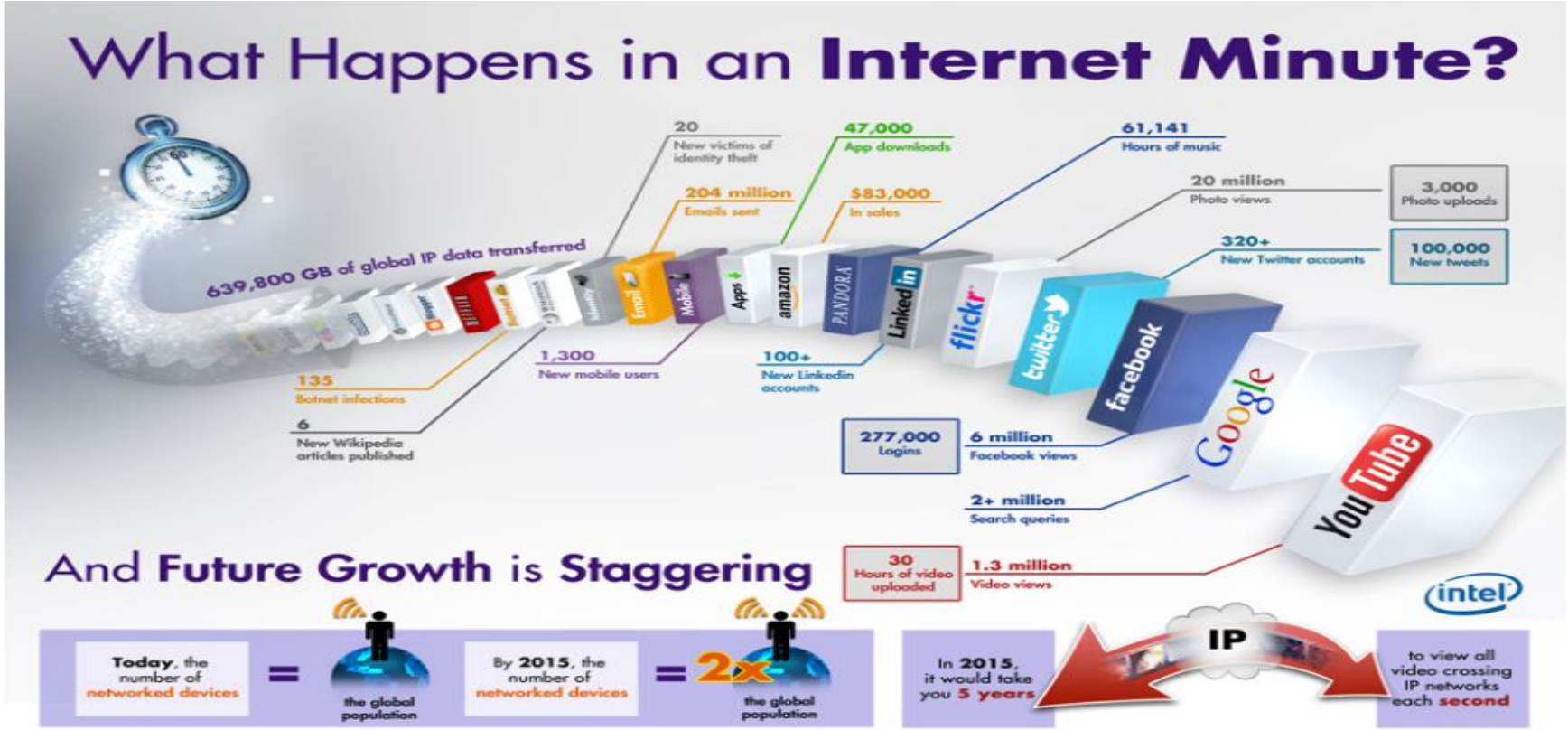
Data Science



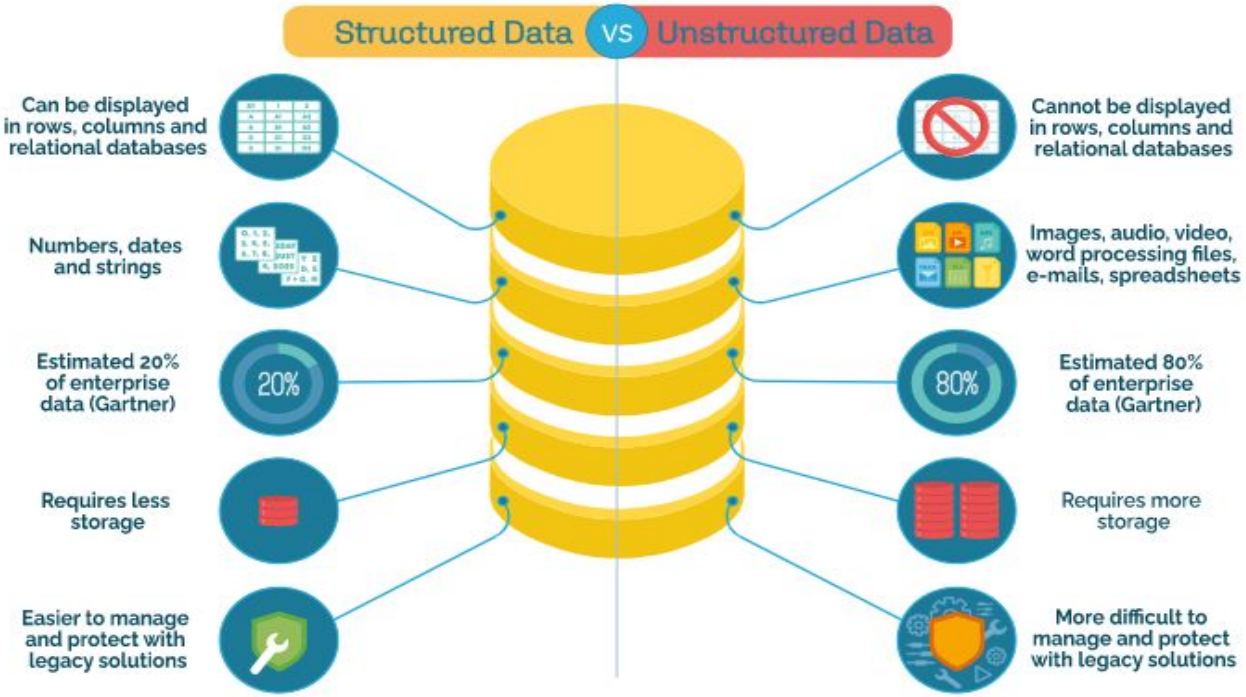
Copyright © 2017, edureka and/or its affiliates. All rights reserved.

Data Explosion



Copyright © 2017, edureka and/or its affiliates. All rights reserved.

Varieties of Data Around Us



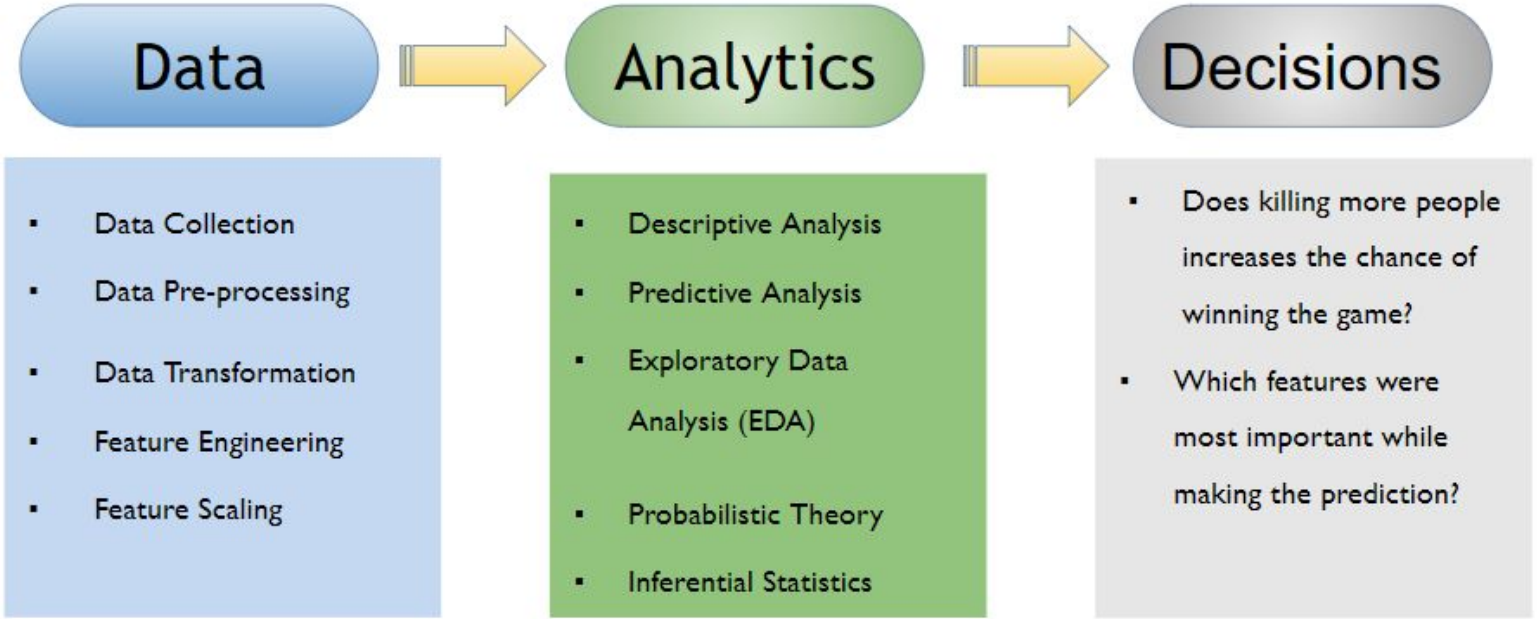
edureka!

Copyright © 2017, edureka and/or its affiliates. All rights reserved.

How to make sense of Data ?

Copyright © 2017, edureka and/or its affiliates. All rights reserved.

How to Make Sense of Data? (Contd.)

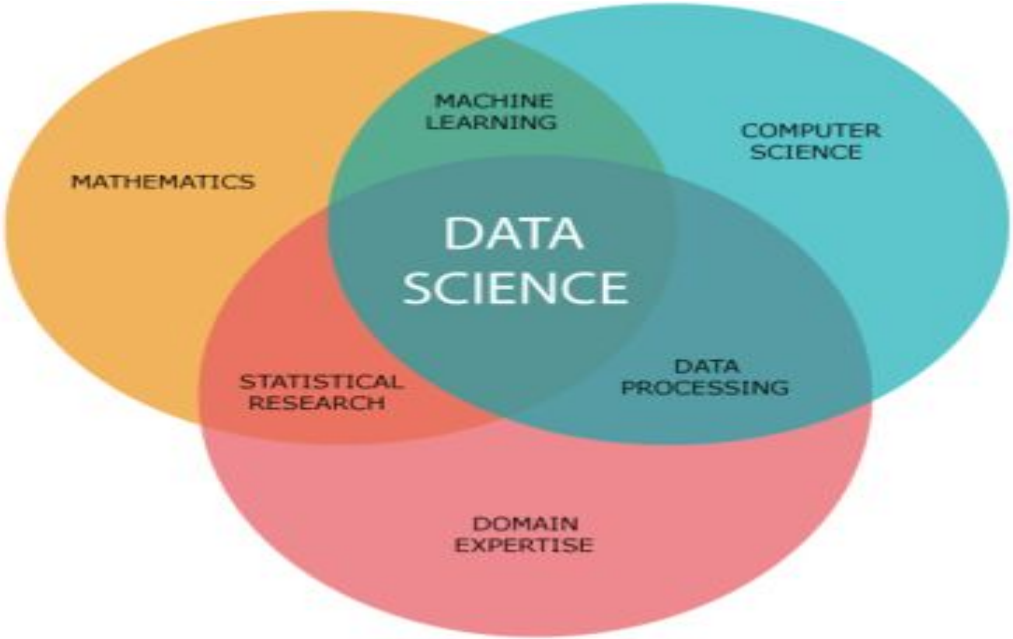


Copyright © 2017, edureka and/or its affiliates. All rights reserved.

What is Data Science?

Data science is the domain of study that deals with vast volumes of data using modern tools and techniques to find unseen patterns, derive meaningful information, and make business decisions

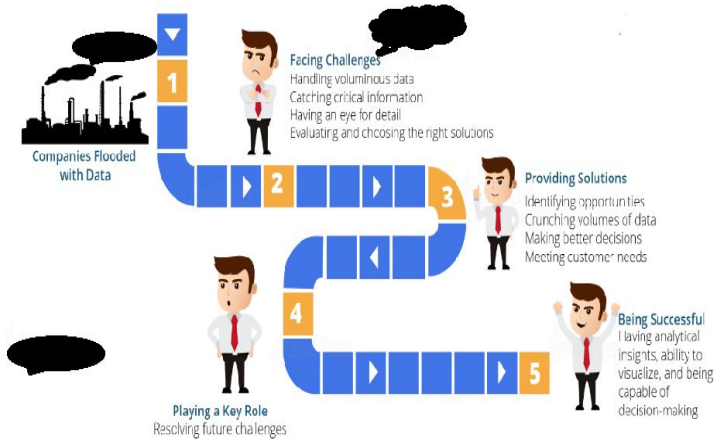
The data used for analysis can come from many different sources and presented in various formats.



Copyright © 2017, edureka and/or its affiliates. All rights reserved.

Data Science Need ?

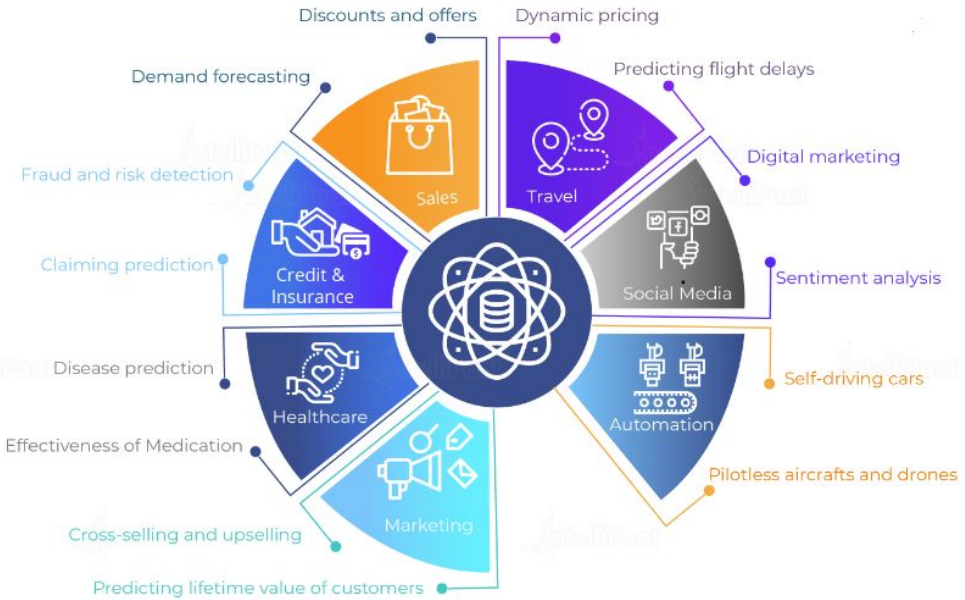
The picture represents the concept of Data Science. It brings together a lot of skills such as statistics, mathematics, and business domain knowledge, and helps organizations find ways to:



- Reduce costs
- Get into new markets
- Tap into different demographics
- Gauge the effectiveness of marketing campaigns
- Launch new products or services

Copyright © 2017, edureka and/or its affiliates. All rights reserved.

Applications of Data Science



edureka!

Copyright © 2017, edureka and/or its affiliates. All rights reserved.

Use Cases of Data Science

- Amazon: Amazon uses a personalized recommendation system to improve customer satisfaction. Amazon analyzes the user's purchase history to recommend more products.
- Spotify: Spotify utilizes Data Science to offer personalized music recommendations to the users.
- Uber: Uber utilizes big data to gain better insights and provide better service to the users.

Copyright © 2017, edureka and/or its affiliates. All rights reserved.

Data Science LifeCycle

1. **Capture:** Data Acquisition, Data Entry, Signal Reception, Data Extraction. This stage involves gathering raw structured and unstructured data

2. **Maintain:** Data Warehousing, Data Cleansing, Data Staging, Data Processing, Data Architecture. This stage covers taking the raw data and putting it in a form that can be used.

3. **Process:** Data Mining, Clustering/Classification, Data Modeling, Data Summarization. Data scientists take the prepared data and examine its patterns, ranges, and biases to determine how useful it will be in predictive analysis.

4. **Analyze:** Exploratory/Confirmatory, Predictive Analysis, Regression, Text Mining, Qualitative Analysis. Here is the real meat of the life cycle. This stage involves performing the various analyses on the data.

5. **Communicate**: Data Reporting, Data Visualization, Business Intelligence, Decision Making. In this final step, analysts prepare the analyses in easily readable forms such as charts, graphs, and reports.

Copyright © 2017, edureka and/or its affiliates. All rights reserved.

Types of Data Science Jobs

Most prominent Data Scientist job titles are:

- Data Scientist
- Data Engineer
- Data Analyst
- Statistician
- Data Architect
- Data Admin
- Business Analyst
- Data/Analytics Manager

Copyright © 2017, edureka and/or its affiliates. All rights reserved.

Who is a Data Scientist ?

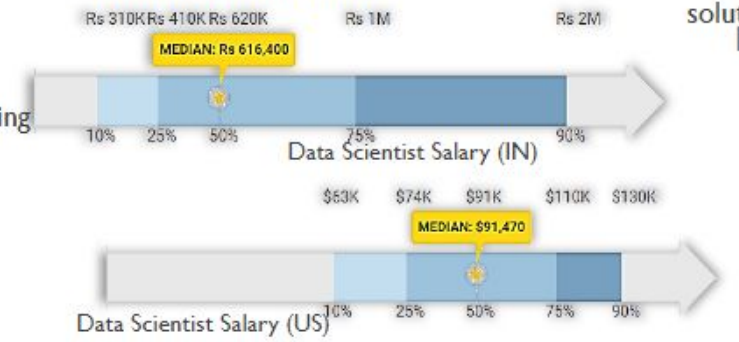
Who Am I?

I am part analyst and a part artist. I use my technical and analytical skills to extract insights out of the data

What did I learn?

- Python/ R
- Statistical Analytics
- Predictive Modelling
- Machine Learning
- Natural Language Processing
- Deep Learning
- Data Warehousing
- Big Data and Cloud – AWS/Azure

How much do I earn?



What do I do?

- Collect data and analyse it from various angles
- Clean existing raw data and build predictive models out of it
- Identify correct business problems and give solution with visualizations

How do I help my organization?

- Cost Optimization
- Develop Strategies
- Improve Operational Efficiency
- Risk Optimization
- Build Recommender System
- Increase data accuracy

edureka!

Copyright © 2017, edureka and/or its affiliates. All rights reserved.

JOB ROLES AND RESPONSIBILITIES

Data Scientist

- Manage, mine, and clean unstructured data to prepare it for practical use.
- Develop models that can operate on Big Data
- Understand and interpret Big Data analysis
- Take charge of the data team and help them towards their respective goals
- Deliver results that have an impact on business outcomes

Data Analyst

- Collecting information from a database with the help of query
- Enable data processing and summarize results
- Use basic algorithms in their work like logistic regression, linear regression and so on
- Possess and display deep expertise in data munging, data visualization, exploratory data analysis and statistics

Data Engineer

- Data Mining for getting insights from data
- Conversion of erroneous data into a useable form for data analysis
- Writing queries on data
- Maintenance of the data design and architecture
- Develop large data warehouses with the help of extra transform load (ETL)

Copyright © 2017, edureka and/or its affiliates. All rights reserved.

JOB ROLES AND RESPONSIBILITIES

ML Engineer

- Designing and developing machine learning and deep learning systems
- Running machine learning tests and experiments
- Implementing appropriate ML algorithms

Computer Vision Engineer

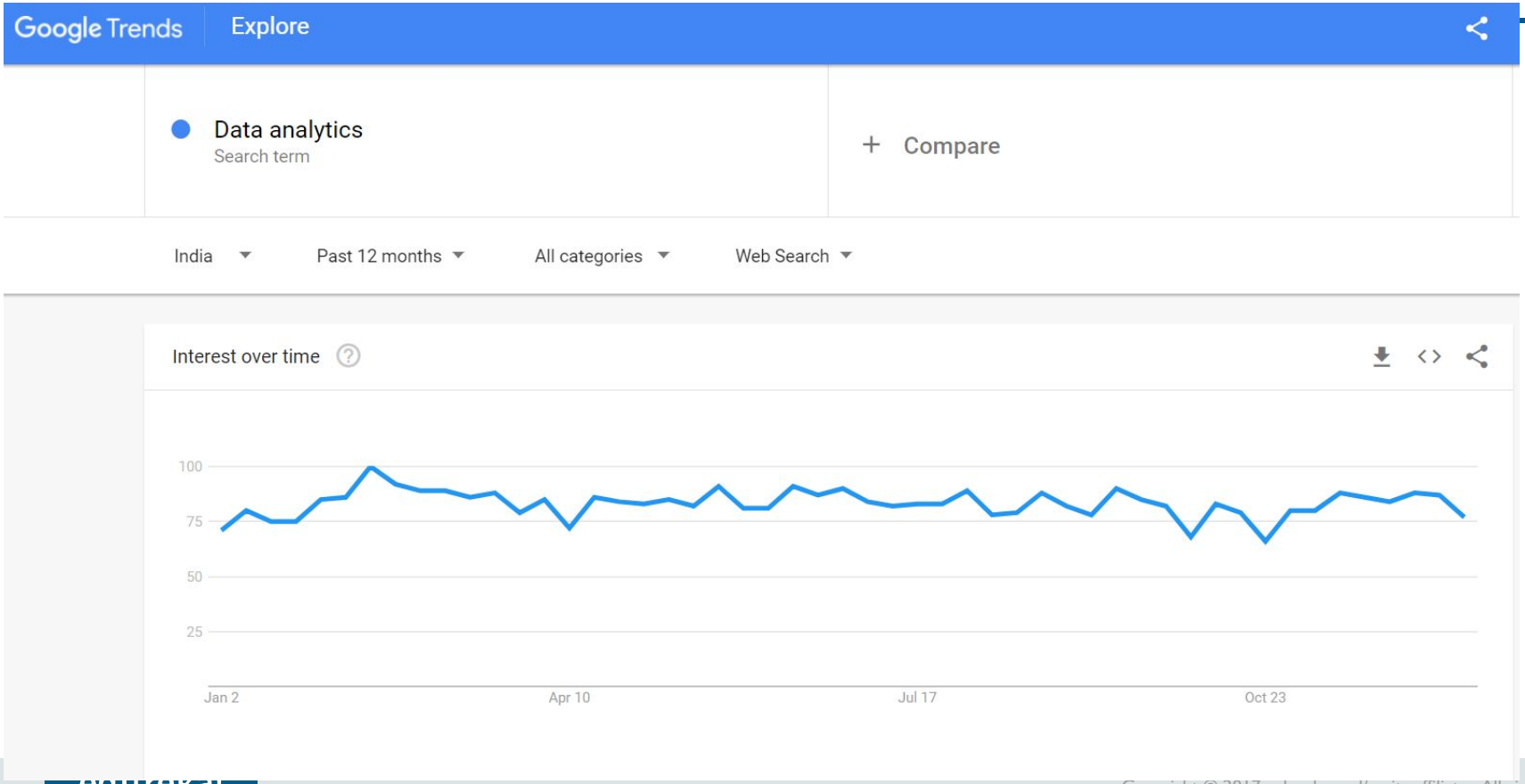
- Develop, test, debug, deploy, and maintain computer vision algorithms and hardware for different environments.
- Develop automated vision algorithms, especially for work with robots and autonomous hardware systems.
- Gather and optimize analytics from computer vision algorithms to improve their performance.

Python Developer

- Developers specializing in writing server-side web application logic.
- Job is to use the Python programming language to develop, debug, and implement application projects.
- They also connect applications with third-party web services and support front-end developers with application integration.

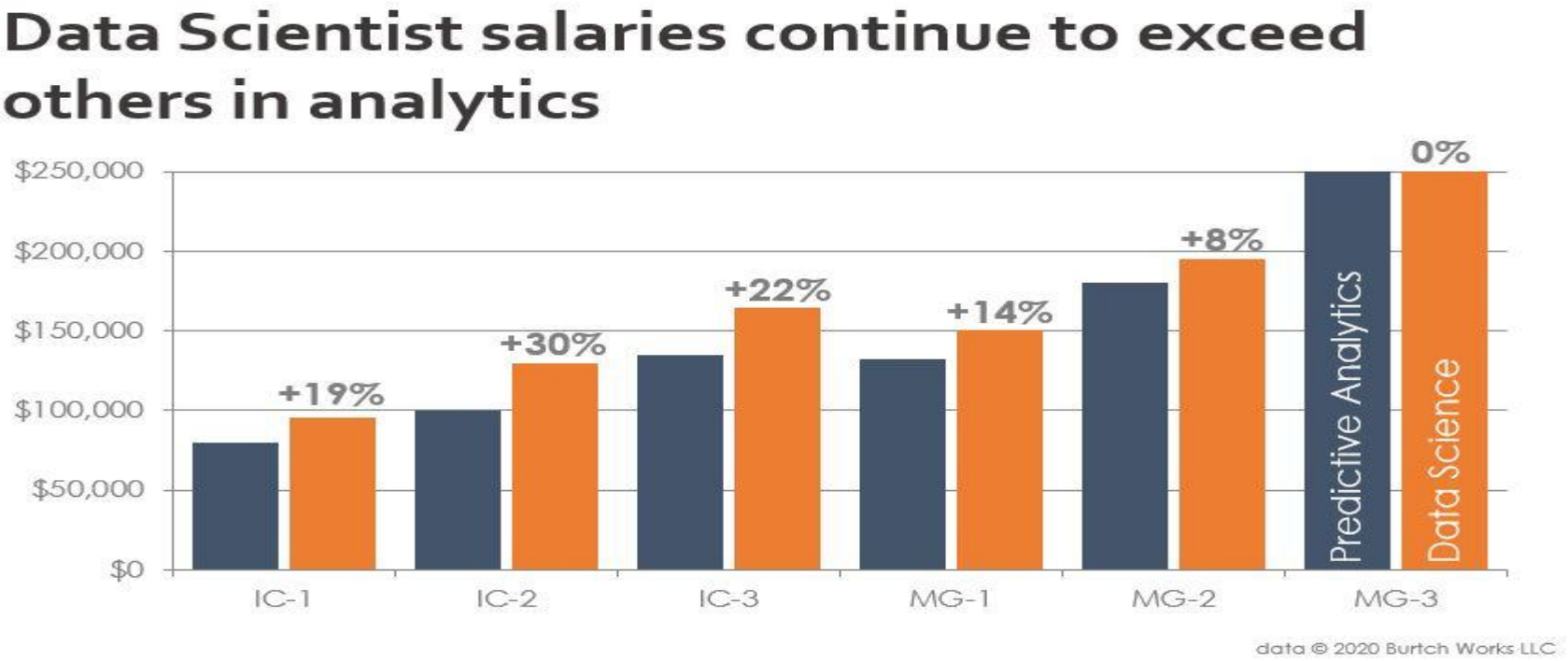
Copyright © 2017, edureka and/or its affiliates. All rights reserved.

Google Trend Statistics



edureka!
Copyright © 2017, edureka and/or its affiliates. All rights reserved.

Important Statistics



edureka!

Copyright © 2017, edureka and/or its affiliates. All rights reserved.

Data Science vs BI

Feature	Data Science	Business Intelligence
Definition	Data science uses scientific methods, processes, algorithms and systems to extract knowledge and insights from structured and unstructured data, and apply algorithms and actionable insights from data across a wide range of application domains.	Business intelligence comprises the strategies and technologies used by organizations for the data analysis of business information. BI technologies provide historical, current, and predictive views of business operations.
Data	It deals with both Structured & Unstructured Data.	It deals majorly with Structured Data.
Method	It is a Scientific method.	It is an Analytical method.
Complexity	Highly complex	Comparatively simpler
Flexible	Data science is much more flexible as data sources can be added as per requirement.	It is less flexible as data sources need to be pre-planned in case of business intelligence.

Copyright © 2017, edureka and/or its affiliates. All rights reserved.

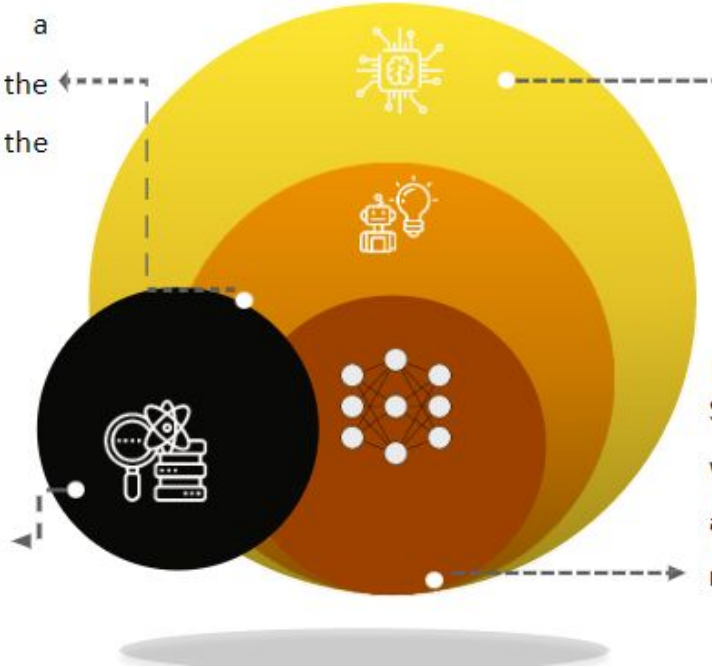
DS vs ML vs AI vs DL

Machine Learning

A subset of AI which gives a machine the ability to use the stat model to learn from the data.

Data Science

Data science is not exactly a subset of ML, but it uses ML and DL to gain insights from both structured and unstructured data.



Artificial Intelligence

Area of computer science that emphasizes on the creation of intelligent machines that work and react like humans.

Deep Learning

Subset of machine learning concerned with algorithms inspired by the structure and function of the brain called artificial neural networks.

Copyright © 2017, edureka and/or its affiliates. All rights reserved.

Prerequisites for Data Science

For Technical:

- Mathematical modeling
- Understanding of Programming
- Data Visualization
- Machine Learning
- Deep Learning
- Database understanding

For Non Technical:

- Business Problem Solving
- Critical Thinking
- Communication Skills
- Mathematics
- Statistics

Copyright © 2017, edureka and/or its affiliates. All rights reserved.

Tools for Data Science

Copyright © 2017, edureka and/or its affiliates. All rights reserved.

What Companies are asking?

Data Analyst 2

PayPal ★★★★☆ 1,648 reviews
Bengaluru, Karnataka
You must create an Indeed account before continuing to the company website to apply

Apply on company site

- Proficiency in SQL and Excel
- Proficiency in at least one statistical analysis tool: SAS / R / Python/ Hadoop
- Strong analytical skills: ability to build quick estimates using back-of-the-envelope analysis, structure (and, if needed, execute) more complex analyses, pull together business cases, navigate multidimensional sets of tradeoffs. Above all, the job calls for comfort with data – ability to manipulate it, question its validity, interpret it, and develop recommendations based on it
- Strong written, oral, and interpersonal skills, including the ability to explain and/or present analysis
- Dedicated, proactive, curious and eager to learn new approaches / methodologies – a must

Data Analyst

Shell ★★★★☆ 9,011 reviews
Bengaluru, Karnataka
You must create an Indeed account before continuing to the company website to apply

Apply on company site

What we need from you

Qualifications:

- Expertise to Azure data platform components: Azure Data Lake Store, Azure Data Factory, Data bricks, SQL PAAS
- Experience in SQL programming required.
- Experience in developing visual reports, dashboards using Power BI or other visualization tool.
- Experience in Azure data platform components: Azure Data Lake Store, Azure Data Factory, Data bricks, SQL PAAS
- Expertise in at least one Data Science language such as Python (Pandas, Scikit Learn, etc.), and R. Knowledge on ML methods like Decision forest, NLP, Time Series etc.
- Statistical skills such as knowledge of statistical tests, distributions, regression etc.

Copyright © 2017, edureka and/or its affiliates. All rights reserved.

Which companies are hiring the most Data Scientists ?

Accenture hired 63,206 data employees

TCS hired 40,525 data employees

IBM hired 35,986 data employees

Citigroup hired 30,548 data employees

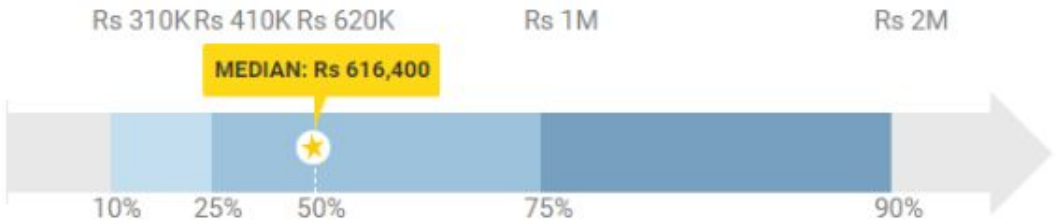
Deloitte hired 29,436 data employees

edureka!

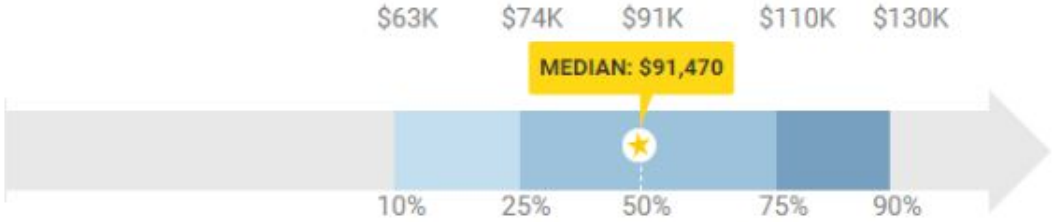
Copyright © 2017, edureka and/or its affiliates. All rights reserved.

Salary Trends

Data Scientist Salary (IND)



Data Scientist Salary (US)



edureka!

Copyright © 2017, edureka and/or its affiliates. All rights reserved.

FAQs

Can I pursue my Data Science course after Graduation?

Yes, our Data Science Training is suitable for fresh graduates and experienced professionals who want to start a career in data science. There are no particular eligibility requirements to take up this Training Program.

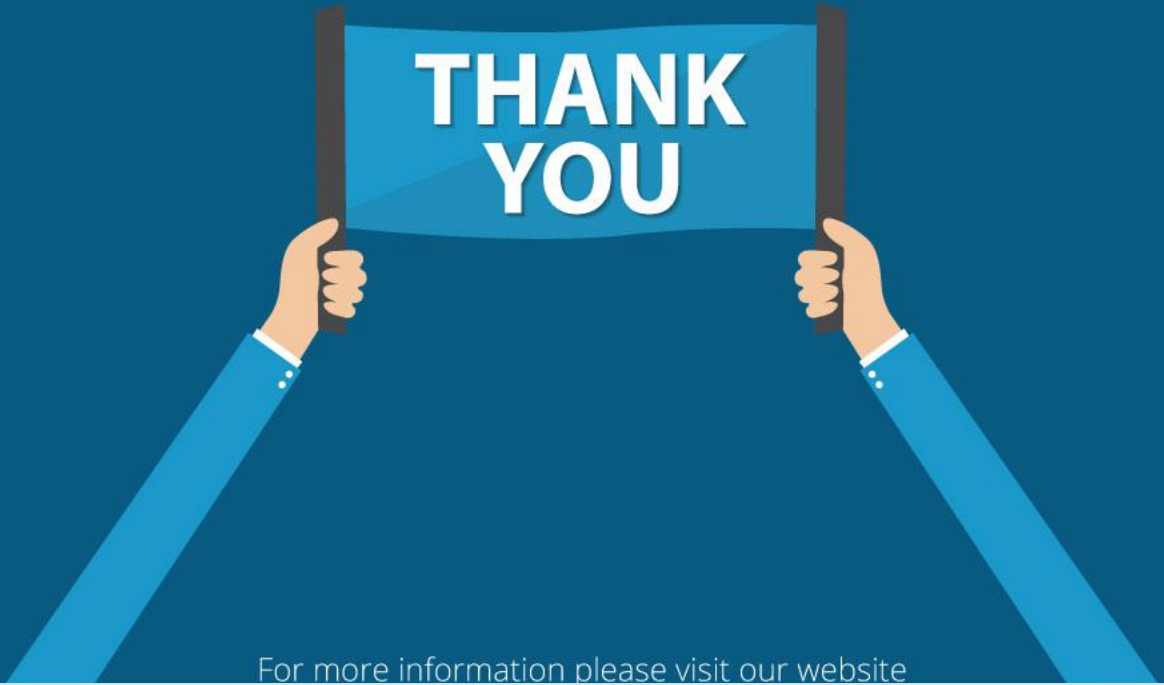
What is the Salary of a Data Scientist as fresher?

According to Ambition Box, the Average salary for a Data Scientist in India is 10.5 Lakhs per year and freshers earn an average of Rs. 5,71,493 annually.

What are the industries which use Data Science?

Data Science is used in all major industrial sectors like banking, healthcare, finance, automotive, marketing, manufacturing, retails, and government agencies.

Copyright © 2017, edureka and/or its affiliates. All rights reserved.



For more information please visit our website

edureka!

Copyright © 2017, edureka and/or its affiliates. All rights reserved.