



Executive PG Programme in

# DATA SCIENCE

Get the Whole Picture



*What  
a whale!*

*What  
a bird!!!!*





# Table of Contents

<b>2</b>	About upGrad
<b>3</b>	Why upGrad?
<b>4</b>	Program Highlights
<b>5</b>	Faculty and Industry Experts
<b>7</b>	upGrad Learning Experience
<b>9</b>	Industry Projects
<b>10</b>	Learning Path
<b>11</b>	Executive PG Programme Curriculum
<b>25</b>	Meet the Class
<b>26</b>	Career Support
<b>27</b>	Our Alumni Work at
<b>28</b>	Career Transitions
<b>30</b>	Experience upGrad Offline
<b>31</b>	Hear from our Learners
<b>33</b>	Program Details and Admission Process



# About upGrad

*upGrad has delivered over 20 million hours of learning, delivering programs by collaborating with universities across the world including Duke CE, IIT Madras, IIT Bangalore and Deakin Business School among others.*

Online education is a fundamental disruption that will have a far-reaching impact. **upGrad** was founded taking this into consideration. upGrad is an online education platform to help individuals develop their professional potential in the most engaging learning environment.

Since inception, upGrad has delivered over 20 million hours of learning, delivering programs by collaborating with universities across the world including Duke CE, IIT Madras, IIT Bangalore and Deakin Business School among others.

upGrad is focused on helping working professionals in their bid to learn, grow and move up in their career through a wide-range of programs designed to improve their expertise.

**IITB** is a renowned university offering programs specialising in data science, machine learning and artificial intelligence. The IITB faculty includes an average of 15+ years of experience.

The faculty covers the conceptual depths of topics such as Data Science, Machine Learning and Artificial Intelligence, and Big Data Analytics. These will be complemented by industry relevant case studies from major industry verticals by industry leaders with 8+ years of experience from upGrad's industry network.

Furthermore, our strong placement network, industry mentorship and the credibility of a Executive PG Programme will provide you with just the right push to accelerate your career in Data Science!



# Why upGrad?



**INR 72 Lakh**  
Highest Salary

**400%**  
Highest Hike



**58%**  
Avg Salary Hike



**300+**  
Hiring Partners



**30,000+**  
Learners



**700+**  
Industry Experts





# Program Highlights

## Equivalent to NSQF level 8

Do an Executive PG Programme from IITB that satisfies NSQF level 8 criterion

## Executive PG Programme from IITB and Alumni Status

Get certified by IITB and gain alumni status on successful completion of the program.

## Dedicated Career Assistance

Receive 360 degree career support from mock interviews with hiring managers, resume building, career fairs, industry mentors and much more.

## For the Industry, by the Industry

Learn from 60+ Industry case studies & Projects and industry experts who mentor you throughout the program.

## 6 Specialisations

Choose from 6 specialisations on the basis of your background and career aspirations and get the learning you want.

## Personalised Mentorship

Get unparalleled personalised mentorship and doubt resolution from IITB faculty and our panel of industry experts.

## Blended Learning

Learn with the ease and flexibility of recorded sessions as well as live sessions, designed to ensure a wholesome learning experience.

## NASSCOM future skill certification

India's first Executive PG Programme validated by and recommended by NASSCOM. Avail a participation certificate from NASSCOM on successful program completion.



# Faculty and Industry Experts



**Hindol Basu**  
CEO, Actify Data Labs

An alumnus of IIT and IIM with over 13 years of experience in analytics with industry leaders such as Citigroup and Tata Industries.



**Chandrashekar Ramanathan**  
Dean Academics, IIITB

Prof. Chandrashekar has a PhD from Mississippi State University and experience of over 10 years in several multinational organisations.



**S. Anand**  
CEO, Gramener

A gold medallist from IIM Bangalore, an alumnus of IIT Madras and London Business School, Anand is among the top 10 data scientists in India with 20 years of experience.



**Tricha Anjali**  
Associate Dean, IIITB

Prof. Anjali has a PhD from Georgia Institute of Technology as well as an integrated MTech (EE) from IIT Bombay.



**Sameer Dhanrajani**  
Co-Founder and CEO, AIQRATE

Sameer Dhanrajani is an AI and Analytics evangelist for Fortune 500 companies who has won several industry awards in the field of analytics.



**Prof. S. Sadagopan**  
Director, IIITB

Prof. Sadagopan is currently Director (President) of IIITB. He has a MS and a PhD from Purdue University as well as a BE (Hons.) Degree from Madras University.



**Prof. G. Srinivasaraghavan**  
Professor, IITB

Prof. Srinivasaraghavan has a PhD in Computer Science from IIT-K and 18 years of experience with Infosys Technologies and several other companies.



**Ujjaini Mitra**  
Head of Analytics, Zee5

An alumna of McKinsey and Co., Flipkart, and Bharti Airtel with over 11 years of experience.



**Dinesh Babu Jayagopi**  
Associate Professor, IITB

Prof. Dinesh has a PhD from EPFL Switzerland, MSc from IISc Bangalore in System Science and Signal Processing and BTech.



**Kalpana Subbaramappa**  
Ex-AVP, Genpact

Kalpana is the ex-AVP of Decision Sciences at Genpact with over 20 years of experience.



**Ankit Jain**  
Sr. Research Scientist, Uber

An alumnus of IIT Bombay, UCB and Harvard Business School with over 9 years of experience.



**Sajan Kedia**  
Lead Data Scientist (Pricing), Myntra

An alumnus of IIT with over 7 years of experience at Watson at IBM Research, startups and Myntra.



**Mirza Rahim Baig**  
Lead Business Analytics, Flipkart

Advanced Analytics professional with 8+ years of experience as a consultant in the e-commerce and healthcare domains.



**Bijoy Kumar Khandelwal**  
COO, Actify Data Labs

Bijoy comes with a deep understanding of the private and cloud architectures and has helped numerous companies make the transition.



# upGrad Learning Experience

## Coaching

- Dedicated Student Success Manager
- Weekly real-time doubt clearing sessions
- Live Discussion forum for peer-to-peer doubt resolution monitored by technical experts
- Peer-to-peer networking opportunities with an alumni pool of 10,000+
- Lab walk-throughs of 15+ industry-driven case studies
- 6 Employability Tests for industry readiness
- Access to the program for up to 3 years

## Mentorship

60+ live interactive sessions with industry experts, fortnightly personalised group (1:8) mentorship sessions and a dedicated student mentor for proactive mentoring.

## Format

Online format with weekly live sessions from industry experts to help with topic walk-throughs, doubt resolution and personalised project feedback. Offline sessions such as Basecamps and Hackathons.

## Hands-On Projects and Hackathons

60+ Industry case studies & projects to choose from as well as a Capstone Project and a Hackathon every quarter to apply learnings.





# New Additions



## Introduction of a new specialization - Data Science Generalist

1. Specially designed for learners with 0 to 5 years of experience to become job ready.
2. Additional Live sessions, Practice Questions, Quizzes, Career Essentials to be conducted
3. Curriculum Focus Areas: Database, Visualization, Classical Machine Learning, Data Structures & Algorithms

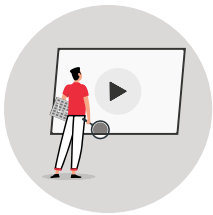
## Career Essential Soft-skills Program

1. Excel your personal & professional life with upGrad's Soft Skills Program
2. Study Three fundamental Skills - Interview & Job Search, Corporate & Business Communication and Problem Solving
3. Get access to 40+ learner hours of soft skills content delivered by the best faculty & Industry experts

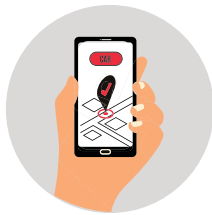
## 30-Hour Programming Bootcamp for Non-Tech Learners

1. Non-tech background? No need to fear of Programming anymore
2. A 30-hour Python Programming bootcamp focusing on developing Basic + Intermediate Python Programming Concepts to assist non-tech learners.
3. A blended learning experience delivered via Interactive live sessions and assessments

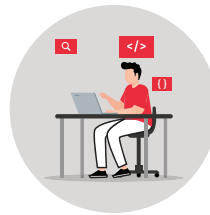
# Industry Projects



IMDb Movie Analysis



Uber Supply-Demand Gap



Lead Scoring



Fraud Detection



Creditworthiness of Customers



Speech Recognition



Image Captioning



Gesture Recognition



Social Media Listening



Telecom Churn



Interactive Market Campaign Analysis



Retail Giant Sales Forecasting



And many more!



# Learning Path



## Preparatory Course

2 weeks

Tools: Python, Excel



## Data Toolkit

14 weeks

Tools: Python, Excel, MySQL



## Machine Learning

9 weeks

Tools: Python, Excel



## Choose any of the 6 Specialisations

25 weeks (with 6 weeks of Capstone)



## Data Science Generalist

Tools: Python, Tableau, SQL



## Natural Language Processing

Tools: Python, Excel



## Deep Learning

Tools: Python, Excel, TensorFlow



## Business Analytics

Tools: Python, MySQL, Excel



## Business Intelligence/ Data Analytics

Tools: Python, Power BI, Excel, MySQL, MongoDB, Shiny, Tableau



## Data Engineering

Tools: Hadoop, HBase, Sqoop, Hive, Flume, PySpark, Spark, Airflow



**Executive PG Programme in Data Science (Data Science Generalist)**



**Executive PG Programme in Data Science (Natural Language Processing)**



**Executive PG Programme in Data Science (Deep Learning)**



**Executive PG Programme in Data Science (Business Analytics)**



**Executive PG Programme in Data Science (Business Intelligence/ Data Analytics)**



**Executive PG Programme in Data Science (Data Engineering)**

# Executive PG Programme in Data Science

## COMMON CURRICULUM

### PRE-PROGRAM PREPARATORY CONTENT

**1**

#### DATA ANALYSIS IN EXCEL

Taught by one of the most renowned data scientists in the country (S.Anand, CEO, Gramener), this module takes you from a beginner level Excel user to an almost professional user.

**2**

#### ANALYTICS PROBLEM SOLVING

This module covers concepts of the CRISP-DM framework for business problem-solving.

**TOOLS COVERED: EXCEL**

**1**

#### DATA TOOLKIT

**2 ASSIGNMENTS**

#### INTRODUCTION TO PYTHON - I

Build a foundation for the most in-demand programming language of the 21st century.

**2**

#### INTRODUCTION TO PYTHON - II

Learn to apply some of the commonly used paradigms of functional programming in Python.

**3**

#### PROGRAMMING IN PYTHON

Learn how to approach and solve logical problems using programming.

**4**

#### DATA ANALYSIS USING SQL

Data in companies is definitely not stored in excel sheets! Learn the fundamentals of database and extract information from RDBMS using the structured query language.

**5**

#### PYTHON FOR DATA SCIENCE

Learn how to manipulate datasets in Python using Pandas which is the most powerful library for data preparation and analysis.

**6**

#### VISUALIZATION IN PYTHON

Humans are visual learners and hence no task related to data is complete without visualisation. Learn to plot and interpret various graphs in Python and observe how they make data analysis and drawing insights easier.

**7**

#### EXPLORATORY DATA ANALYSIS

Learn how to find and analyse the patterns in the data to draw actionable insights.

**8**

#### IMDB MOVIE ASSIGNMENT

Reinforce the concepts learnt in data science through this rigorous assignment involving the past



## 9 MATHS FOR DATA SCIENCE

Build the mathematical foundation required for understanding the Machine Learning Algorithms.

## 10 INFERENCE STATISTICS

Build a strong statistical foundation and learn how to 'infer' insights from a huge population using a small sample.

## 11 HYPOTHESIS TESTING

Understand how to formulate and validate hypotheses for a population to solve real-life business problems.

## 12 ADVANCED SQL

Apply advanced SQL concepts like windowing and procedures to derive insights from data and answer pertinent business questions.

## 13 CREDIT EDA CASE STUDY

Solve a real industry problem through the concepts learnt in exploratory data analysis.

**TOOLS COVERED: PYTHON, MYSQL, EXCEL**

### MACHINE LEARNING

3 ASSIGNMENTS

### 1 INTRODUCTION TO MACHINE LEARNING AND LINEAR REGRESSION

Venture into the machine learning community by learning how one variable can be predicted using several other variables through a housing dataset where you will predict the prices of houses based on various factors.

### 2 LINEAR REGRESSION ASSIGNMENT - BIKE SHARING SYSTEMS

Build a model to understand the factors on which the demand for bike sharing systems vary on and help a company optimise its revenue.

### 3 LOGISTIC REGRESSION

Learn your first binary classification technique by determining which customers of a telecom operator are likely to churn versus who are not to help the business retain customers.

### 4 UNSUPERVISED LEARNING: CLUSTERING

Learn how to group elements into different clusters when you don't have any pre-defined labels to segregate them through K-means clustering, hierarchical clustering, and more.

### 5 BUSINESS PROBLEM SOLVING

Learn how to approach open ended real world problems using data as a lever to draw actionable insights.

### 6 CLUSTERING ASSIGNMENT (OPTIONAL)

Apply the machine learning concepts learnt to help an international NGO cluster countries to determine their overall development and plan for lagging countries.

### 7 CASE STUDY: LEAD SCORING

Help the Sales team of your company identify which leads are worth pursuing through this classification case study.

**TOOLS COVERED: PYTHON**

# SPECIALISATION 1:

## DATA SCIENCE GENERALIST

### ADVANCED MACHINE LEARNING AND STORYTELLING

3 ASSIGNMENTS

1

#### TREE MODELS + BOOSTING (OPTIONAL)

Learn how the human decision making process can be replicated using a decision tree and other powerful ensemble algorithms.

2

#### MODEL SELECTION & GENERAL ML TECHNIQUES

Learn the pros and cons of simple and complex models and the different methods for quantifying model complexity, alongwith general machine learning techniques like feature engineering, model evaluation, and many more.

3

#### ML LAB I: CLASSIFICATION

4

#### PRINCIPAL COMPONENT ANALYSIS

Understand important concepts related to dimensionality reduction, the basic idea and the learning algorithm of PCA, and its practical applications on supervised and unsupervised problems.

5

#### ADVANCED REGRESSION - I

In this module, take a more advanced look at regression models and learn the concepts related to regularization.

6

#### ADVANCED REGRESSION - II & ML LAB II: REGRESSION

7

#### TEXT ANALYTICS & PROCESSING + TEXT-BASED PREDICTIVE MODELLING

An introduction to the world of NLP and basic text processing skills. Learn how to build a classification engine that works on (unstructured) textual data.

8

#### BASIC VISUALISATION USING TABLEAU

Learn advanced visualisation techniques using the most in-demand visualization tool in the industry.

9

#### DATA STORYTELLING

Learn how to effectively strategise, communicate, and fine grain your data analysis projects and understand how to optimally present your findings to technical and non-technical stakeholders and upgrade your storytelling skills.

10

#### BUSINESS CASE STUDY

**TOOLS COVERED: PYTHON, TABLEAU**

### ADVANCED PROGRAMMING AND DATABASES

2 ASSIGNMENTS

1

#### DATA MODELLING

In this module, you will learn and use data modelling on a dataset to solve a business problem.



## 2 SQL WEEKLONG LAB

## 3 ADVANCED SQL - WEEK II

Apply advanced SQL concepts like windowing and procedures to derive insights from data and answer pertinent business questions

## 4 ALGORITHM ANALYSIS + RECURSION

Learn how to assess the efficiency your code using algorithm analysis techniques and learn to write recursive algorithms

## 5 SEARCHING AND SORTING (DIVIDE AND CONQUER INCLUDED)

Learn most fundamental searching and sorting algorithms and design techniques

## 6 DATA STRUCTURES - SETS, DICTIONARIES, STACKS, QUEUES

Learn user defined data structures -Stack, Queue, Trees in Python that help in advanced data manipulation

## 7 PYTHON - OOPS

Learn OOP concepts such as Class, Object, Method, Inheritance, Polymorphism, Data Abstraction and Encapsulation.

## 8 PYTHON WEEKLONG LAB

**TOOLS COVERED: PYTHON, SQL**

## CAPSTONE PROJECT WITH VIDEO PRESENTATION

## 1 CAPSTONE PROJECT

Solve an end-to-end real-life industry problem from a wide variety of domains. Make a video presentation of your working demo to showcase in your portfolio.



# SPECIALISATION 2:

## NATURAL LANGUAGE PROCESSING

### ADVANCED MACHINE LEARNING

2 ASSIGNMENTS

1

#### TREE MODELS

Learn how the human decision making process can be replicated using a decision tree and other powerful ensemble algorithms.

2

#### MODEL SELECTION & GENERAL ML TECHNIQUES

Learn the pros and cons of simple and complex models and the different methods for quantifying model complexity, alongwith general machine learning techniques like feature engineering, model evaluation, and many more.

3

#### PRINCIPAL COMPONENT ANALYSIS

Understand important concepts related to dimensionality reduction, the basic idea and the learning algorithm of PCA, and its practical applications on supervised and unsupervised problems.

4

#### ADVANCED REGRESSION

In this module, take a more advanced look at regression models and learn the concepts related to regularization.

5

#### ADVANCED REGRESSION ASSIGNMENT

Build a regularized regression model to understand the most important variables to predict the house prices in Australia.

6

#### BAGGING AND BOOSTING

Learn about ensemble modelling through bagging and boosting and understand how weak algorithms can be transformed into stronger ones.

7

#### TIME SERIES ANALYSIS

In this module, you will learn how to analyse and forecast a series that varies with time.

8

#### TELECOM CHURN CASE STUDY

Solve the most crucial business problem for a leading telecom operator in India and southeast Asia - predicting customer churn.

**TOOLS COVERED: PYTHON**

### NATURAL LANGUAGE PROCESSING

2 ASSIGNMENTS

1

#### TEXT PROCESSING

Do you get annoyed by the constant spams in yor mail box? Wouldn't it be nice if we had a program to check your spellings? In this module learn how to build a spell checker & spam detector using techniques like phonetic hashing, bag-of-words, TF-IDF, etc.

2

#### FEATURE EXTRACTION & MODELLING

This module will help you in understanding how to extract meaningful features from the processed text data. Using these features you will be able to extract entities, classify POS tags, generating similarity score between two question strings.





### 3 **ASSIGNMENT - NATURAL LANGUAGE PROCESSING**

To perform a sentiment analysis on product reviews from Amazon using NLP & Machine learning. This assignment will be focused to give you a business understanding on how to do product optimisation using NLP.

### 4 **INTRO TO DL**

Learn the most sophisticated and cutting-edge technique in machine learning - Artificial Neural Networks & how to apply Deep learning for NLP.

### 5 **NLP INDUSTRIAL APPLICATIONS**

Learn how to use NLP with Neural networks for different industrial applications like text classification, Question pair similarity, Text generation, Topic modelling.

### 6 **CHATBOT CASE STUDY**

Imagine if you could make restaurant booking without opening Zomato. Build your own restaurant-search chatbot with the help of RASA - an open source framework and deploy it on Slack.

### **CAPSTONE**

#### 1 **DEPLOYMENT**

Learn how to productionize your model and deploy it on the server.

#### 2 **CAPSTONE**

Choose from a range of real-world industry woven projects on advanced topics like Recommendation Systems, Fraud Detection, Emotion Detection from faces, Social Media Listening, Speech Recognition among many others.

## **SPECIALISATION 3: DEEP LEARNING**

### **ADVANCED MACHINE LEARNING**

**2 ASSIGNMENTS**

#### 1 **Tree Models**

Learn how the human decision making process can be replicated using a decision tree and other powerful ensemble algorithms.

#### 2 **MODEL SELECTION & GENERAL ML TECHNIQUES**

Learn the pros and cons of simple and complex models and the different methods for quantifying model complexity, alongwith general machine learning techniques like feature engineering, model evaluation, and many more.

#### 3 **PRINCIPAL COMPONENT ANALYSIS**

Understand important concepts related to dimensionality reduction, the basic idea and the learning algorithm of PCA, and its practical applications on supervised and unsupervised problems.

## 4 ADVANCED REGRESSION

In this module, take a more advanced look at regression models and learn the concepts related to regularization.

## 5 ADVANCED REGRESSION ASSIGNMENT

Build a regularized regression model to understand the most important variables to predict the house prices in Australia.

## 6 BAGGING AND BOOSTING

Learn about ensemble modelling through bagging and boosting and understand how weak algorithms can be transformed into stronger ones.

## 7 TIME SERIES ANALYSIS

In this module, you will learn how to analyse and forecast a series that varies with time.

## 8 TELECOM CHURN CASE STUDY

Solve the most crucial business problem for a leading telecom operator in India and southeast Asia - predicting customer churn.

**TOOLS COVERED: PYTHON**

## DEEP LEARNING AND NEURAL NETWORKS

2 ASSIGNMENTS

### 1 INTRODUCTION TO NEURAL NETWORKS

Learn the most sophisticated and cutting-edge technique in machine learning - Artificial Neural Networks or ANNs.

### 2 NEURAL NETWORKS ASSIGNMENT

Build a neural network from scratch in Numpy to identify handwritten digits.

### 3 CONVOLUTIONAL NEURAL NETWORKS - INTRODUCTION AND INDUSTRY APPLICATIONS

Learn the basics of CNN and OpenCV and apply it to Computer Vision tasks like detecting anomalies in chest X-Ray scans, vehicle detection to count & categorise them to help the government ascertain the width and strength of the road.

### 4 RECURRENT NEURAL NETWORKS

Ever wondered what goes behind machine translation, sentiment analysis, speech recognition? Learn how RNN helps in these areas having sequential data like text, speech, videos, and a lot more.

### 5 GESTURE RECOGNITION

Make a Smart TV system which can control the TV with user's hand gestures as the remote control.

**TOOLS COVERED: PYTHON, TENSORFLOW, KERAS, OPENCV**

## CAPSTONE

### 1 DEPLOYMENT

Learn how to productionize your model and deploy it on the server.

### 2 CAPSTONE

Choose from a range of real-world industry woven projects on advanced topics like Recommendation Systems, Fraud Detection, Emotion Detection from faces, Social Media Listening, Speech Recognition among many others.



# SPECIALISATION 4:

## Business Analytics

### ADVANCED MACHINE LEARNING

2 ASSIGNMENTS

1

#### TREE MODELS

Learn how the human decision making process can be replicated using a decision tree and other powerful ensemble algorithms.

2

#### TIME SERIES FORECASTING

In this module, you will learn how to analyse and forecast a series that varies with time.

3

#### RETAIL-GIANT SALES FORECASTING ASSIGNMENT

Apply the concepts learnt in time series to solve a forecasting problem for a retail giant.

4

#### MODEL SELECTION & GENERAL ML TECHNIQUES

Learn the pros and cons of simple and complex models and the different methods for quantifying model complexity, alongwith general machine learning techniques like feature engineering, model evaluation, and many more.

5

#### SQL BEST PRACTICES

Learn how to write optimized SQL query that require less memory and execute in lesser amount of time.

6

#### ADVANCED EXCEL

Learn the advanced concepts in Excel and start to perform data analysis like a pro!

7

#### TELECOM CHURN CASE STUDY

Solve the most crucial business problem for a leading telecom operator in India and southeast Asia - predicting customer churn.

**TOOLS COVERED: PYTHON, MYSQL, EXCEL**

### BUSINESS REQUIREMENTS

2 ASSIGNMENTS

1

#### STRUCTURED PROBLEM SOLVING USING FRAMEWORKS

Learn how to attack a business problem using various structured frameworks like 5W, 5WHYs, and SPIN.

2

#### STRUCTURED PROBLEM SOLVING ASSIGNMENT

Apply your learnings from the course to solve a real-life business problem.

3

#### OPERATIONS RESEARCH

Learn about the world of operations research through linear and integer optimisations.

4

#### DATA STORYTELLING

Learn how to effectively strategise, communicate, and fine grain your data analysis projects and understand how to optimally present your findings to technical and non-technical stakeholders and upgrade your storytelling skills.

## 5

**OPERATIONS RESEARCH CASE STUDY**

Understand how a project in the industry is taken up and solved through a comprehensive business case study.

**TOOLS COVERED: PYTHON, EXCEL**

**CAPSTONE PROJECT****5 ASSIGNMENTS**

## 1

**CAPSTONE PROJECT**

Solve an end-to-end real-life industry problem from a wide variety of domains. Available capstone project choices -

- (i) Stock Analysis & Portfolio Management
- (ii) E-Commerce & Marketing
- (iii) Healthcare
- (iv) Supply Chain Optimisation
- (v) Credit Card Fraud Detection

**TOOLS COVERED: PYTHON, EXCEL, TABLEAU**

## SPECIALISATION 5: BUSINESS INTELLIGENCE / DATA ANALYTICS

**SQL AND NOSQL DATABASES****2 ASSIGNMENTS**

## 1

**DATA MODELLING**

In this module, you will learn and use data modelling on a dataset to solve a business problem.

## 2

**SQL BEST PRACTICES**

Learn how to write optimized SQL query that require less memory and execute in lesser amount of time.

## 3

**SQL ASSIGNMENT: IMDB MOVIES**

In this assignment, you will work on a movies dataset using SQL to extract exciting insights.

## 4

**ADVANCED EXCEL**

Learn the advanced concepts in Excel and start to perform data analysis like a pro!"

## 5

**NOSQL DATABASES AND BEST PRACTICES**

Take your knowledge of query languages a step further by learning about MongoDB - a NoSQL database which is becoming more and more popular in the industry.

## 6

**INTRODUCTION TO BIG DATA AND CLOUD**

Understand the basics of big data and cloud and learn to work with an EMR cluster on a cloud-based service.

## 7

**HIVE AND QUERYING**

In this module, you learn about the architecture and features of the Hive Query Language.

## 8

**HIVE CASE STUDY**

Understand how a project in the industry is taken up and solved through a comprehensive business case study.



**TOOLS COVERED: MYSQL, MONGODB, AWS, HQL**

## **STORYTELLING WITH ADVANCED VISUALISATIONS**

**2 ASSIGNMENTS**

### **1 VISUALISATION USING TABLEAU**

Learn advanced visualisation techniques using the most in-demand visualization tool in the industry.

### **2 SPORTS ANALYTICS - IPL VISUALISATION ASSIGNMENT**

Apply the new found Excel and Tableau skills to solve an exciting business assignment.

### **3 VISUALISATION USING POWERBI**

Take your visualization game a step forward by understanding how to operate PowerBI.

### **4 VISUALISATION USING PLOTLY**

Get a brief introduction to another popular open-sourced visualisation library in Python and learn to code and create powerful, pretty, and interactive visualisations.

### **5 DATA STORYTELLING**

Learn how to effectively strategise, communicate, and fine grain your data analysis projects and understand how to optimally present your findings to technical and non-technical stakeholders and upgrade your storytelling skills.

### **6 PLOTLY CASE STUDY**

Understand how a project in the industry is taken up and solved through a comprehensive business case study.

**TOOLS COVERED: TABLEAU, POWERBI**

## **CAPSTONE PROJECT**

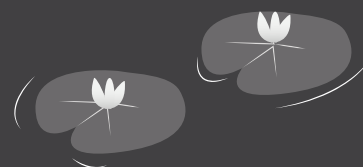
**5 ASSIGNMENTS**

### **1 CAPSTONE PROJECT**

Solve an end-to-end real-life industry problem from a wide variety of domains. Available capstone project choices -

- (i) Web & Social Media Analytics
- (ii) Finance and Risk Analytics
- (iii) Marketing and Retail Analytics
- (iv) Supply Chain Analytics
- (v) Fraud Analytics

**TOOLS COVERED: TABLEAU, PYTHON, EXCEL**



# **SPECIALISATION 6: DATA ENGINEERING**

## **DATA ENGINEERING I**

**4 ASSIGNMENTS (1 Mandatory, 3 optional)**

### **1 INTRODUCTION TO BIG DATA(OPTIONAL)**

This module you will learn what big data is, its various characteristics, and its determining factors. You will also get an idea of the various sources of big data and the wide range of big data applications in different industries such as retail, healthcare, and finance.



## 2 INTRODUCTION TO CLOUD AND AWS SETUP

Understand what is cloud and setup your AWS account which will be required during the program.

## 3 INTRODUCTION TO HADOOP AND MAPREDUCE PROGRAMMING

Understand the world of distributed data processing and storage with Hadoop. Learn to write MapReduce jobs in Python.

## 4 MAPREDUCE PROGRAMMING ASSIGNMENT (OPTIONAL)

Practise MapReduce Programming on a Big Dataset.

## 5 DATA MANAGEMENT AND RELATIONAL DATABASE MODELLING

Understand the concepts of Data Management and learn to model data from a Relational Database.

## 6 NOSQL DATABASES AND APACHE HBASE NOSQL DATABASES AND MONGODB(OPTIONAL)

Learn the concepts of NoSQL databases. Understand the working of Apache HBase.

## 7 DATA WAREHOUSING (OPTIONAL)

Understand the intricacies behind designing a data warehouse and a data lake for use case/s.

## 8 DATA INGESTION WITH APACHE SQOOP AND APACHE FLUME

Get familiar with the challenges involved in data ingestion. Use Sqoop and Flume to ingest structured and unstructured data into Hadoop.

## 9 HIVE & QUERYING

Manage and query a data warehouse with Apache Hive. Learn to write optimized HQL for large scale data analysis.

## 10 HIVE ASSIGNMENT (OPTIONAL)

Use HQL to analyse a Big Dataset

## 11 AMAZON REDSHIFT

Learn to deploy a Redshift cluster and use it for querying data.

## 12 INTRODUCTION TO APACHE SPARK

Get introduced to Apache Spark, a lightning fast big data processing engine.

## 13 NYC PARKING ASSIGNMENT (OPTIONAL)

Practise Apache Spark and its core libraries on the NYC Parking Ticket dataset.

## 14 PROJECT: ETL DATA PIPELINE

Make use of Sqoop, Redshift & Spark to design an ETL data pipeline.

**TOOLS COVERED:** AWS EC2, CLOUDERA DISTRIBUTION OF HADOOP, APACHE HADOOP, APACHE MAPREDUCE, APACHE HBASE, MONGODB, APACHE SQOOP, APACHE FLUME, APACHE HIVE, AMAZON REDSHIFT, APACHE SPARK, MYSQL, AMAZON REDSHIFT



## DATA ENGINEERING - II

3 ASSIGNMENTS (1 Mandatory, 2 optional)

- 1 **OPTIMISING SPARK FOR LARGE SCALE DATA PROCESSING**  
Use PySpark to create large scale data processing applications.
- 2 **APACHE FLINK(OPTIONAL)**  
Get Introduced to Apache Flink and learn query batch data. Use DataStream API to create a stream processing application.
- 3 **REAL-TIME DATA STREAMING WITH APACHE KAFKA**  
Understand the producer-consumer architecture of Apache Kafka. Learn to set up a Kafka cluster for managing real-time data.
- 4 **REAL-TIME DATA PROCESSING USING SPARK STREAMING**  
Learn about the real-time data processing architecture of Apache Spark. Build Spark Streaming applications to process data in real-time.
- 5 **STOCK DATA ANALYSIS ASSIGNMENT (OPTIONAL)**  
This assignment revolves around building Spark structured streaming application to processing stock data in real-time.
- 6 **BUILDING AUTOMATED DATA PIPELINES WITH AIRFLOW**  
Automate Data Pipelines with Airflow.
- 7 **ANALYTICS USING PYSARK**  
Use PySpark to do EDA and Predictive Analysis using Spark's ML library.
- 8 **CLASSIFICATION ASSIGNMENT (OPTIONAL)**  
An assignment related to a classification based problem statement.
- 9 **PROJECT: REAL TIME DATA PROCESSING**  
Build an end-to-end real-time data processing application using Spark Streaming and Kafka.

**TOOLS COVERED: APACHE SPARK, APACHE FLINK, APACHE KAFKA, APACHE SPARK STREAMING, APACHE AIRFLOW, APACHE SPARK**

## CAPSTONE PROJECT

4 ASSIGNMENTS

- 1 **CAPSTONE PROJECT**  
The capstone project will stitch all the components of data engineering together.

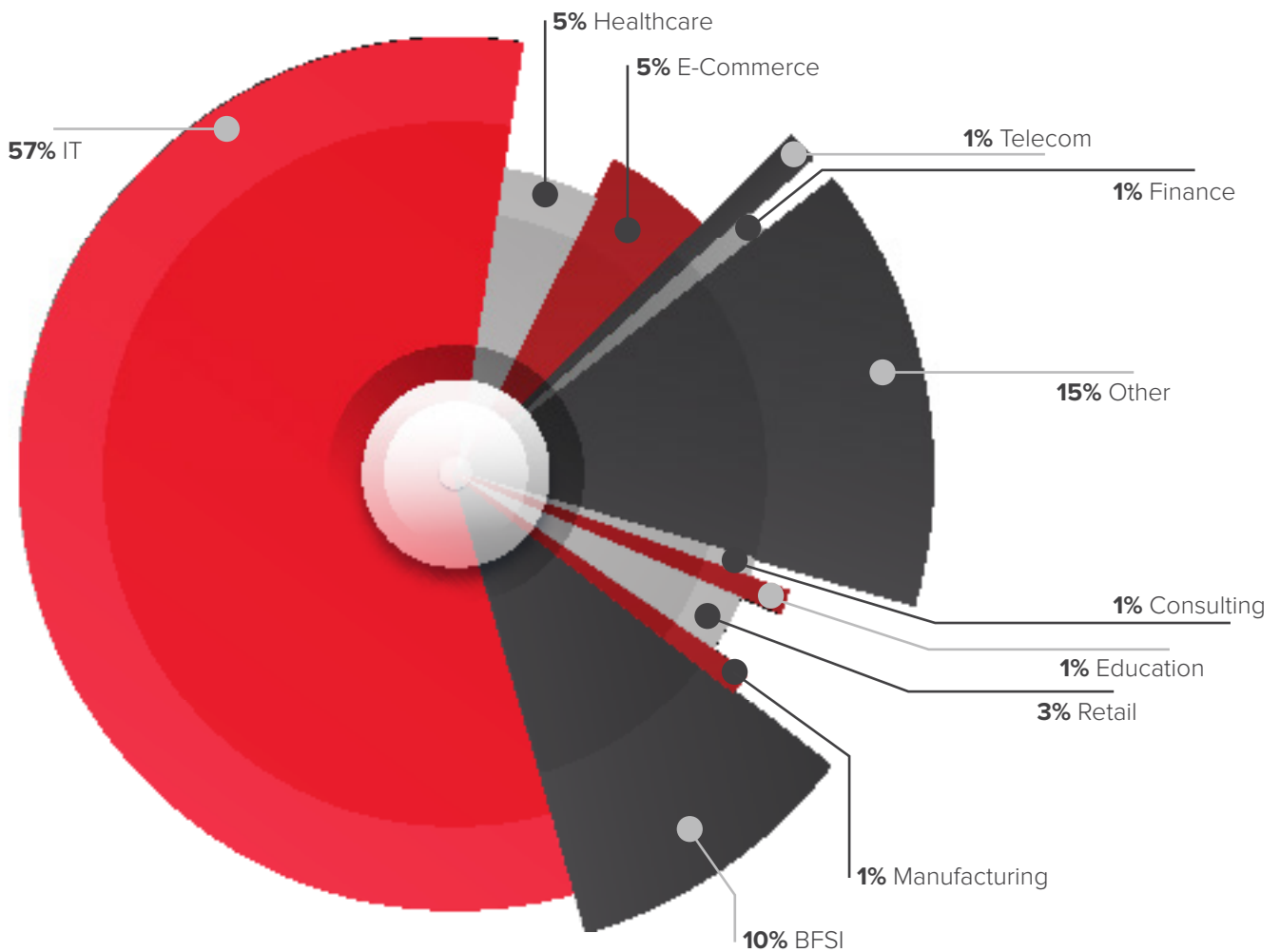
**TOOLS COVERED: APACHE SPARK, MYSQL, HIVE, APACHE SQOOP, SPARK STREAMING, APACHE KAFKA, PYTHON**



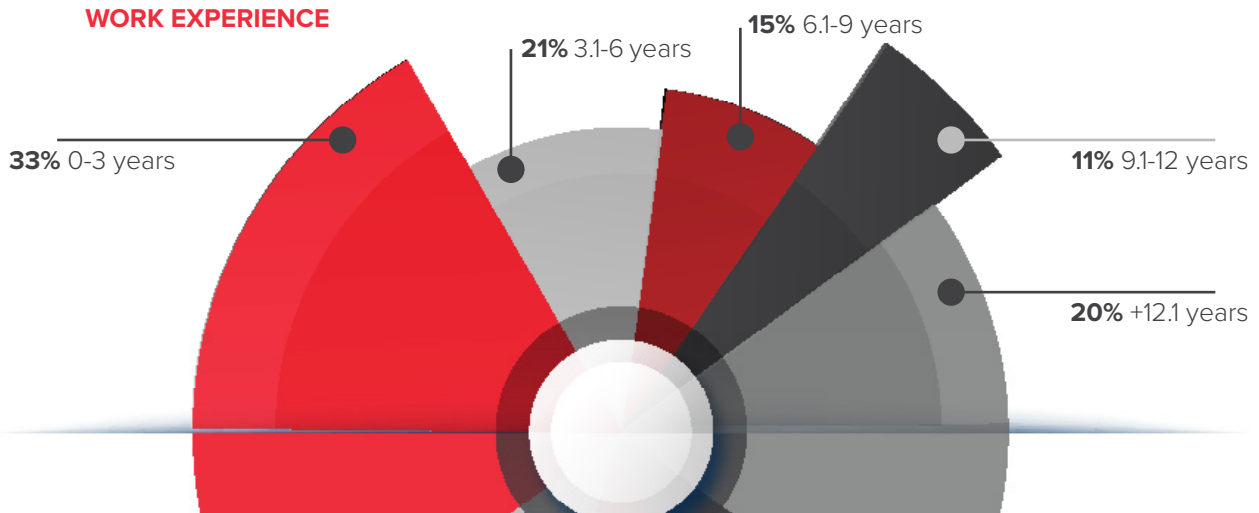


# Meet the Class

## INDUSTRIES OUR STUDENTS COME FROM



## WORK EXPERIENCE







# Career Support

## Just-In-Time Interviews

Get company and role-specific preparation with mock interviews right before your actual interviews.

## Job Opportunities Portal

Gain exclusive access to upGrad's 200 job openings each quarter, giving you the opportunity to be interviewed by upGrad's hiring partners.

## Personalised Mentorship

Get mentored by an experienced data science industry expert and receive personalised feedback with 4 calls spread over a period of 3 months.

## Resume Review

Obtain specific, personalised inputs on your resume structure and content.

## upGrad Career Fairs

Regular hiring drives in major cities across India, giving you the opportunity to interview with upGrad's 300+ hiring partners, ensuring you get every opportunity you deserve.

## Post Graduation Career Support

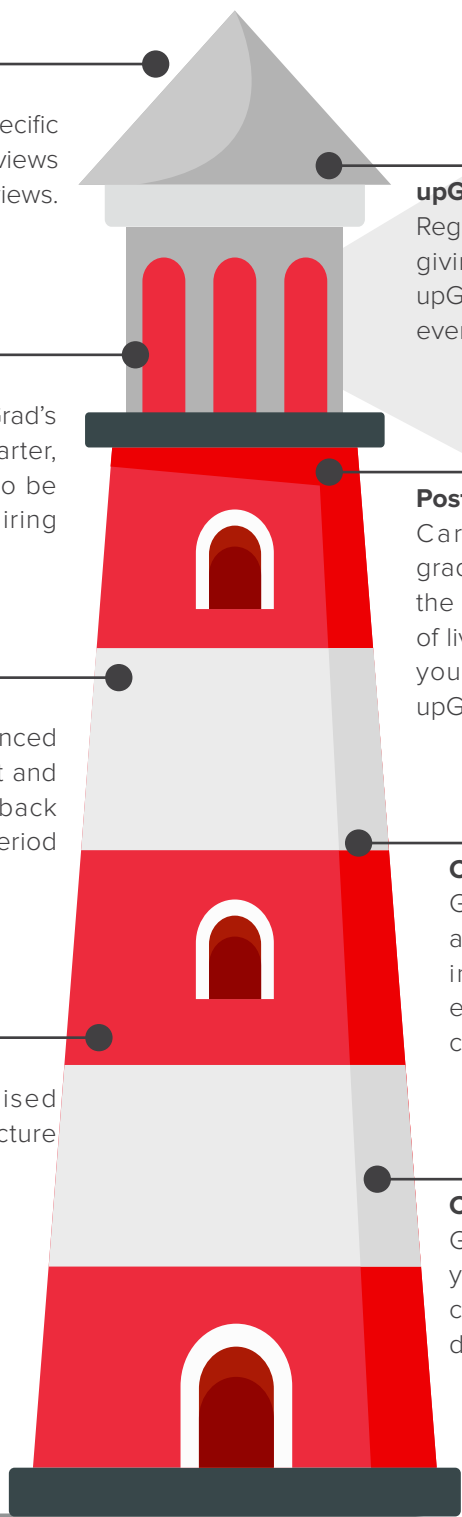
Career sessions are continued post graduation because we want you to have all the help you need. These include: 80 hours of live sessions with industry mentors to guide you, dedicated mentorship, and access to upGrad's career assistance resources.

## Company-Specific Preparation

Get company-specific guidance with access to a carefully curated pool of interview resources per company to ensure you are interview-ready for the company of your choice.

## Career Mentor

Get a dedicated career mentor to help track your weekly company application targets, coach you on your profile, and support you during your career transition journey.



# Our Alumni Work at

## PROGRAM SUMMARY







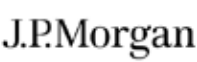







**20,000+**  
Enrolled Learners

**INR72** LPA  
Highest Salary Hike

**400%**  
Highest Salary Hike

**58%**  
Average Salary Hike

upGrad has a network of over 100 companies who look to recruit graduates from our programs. Some of these well-known companies include:



# Career Transitions



**Abhishek Singh**

MIS Executive (F&B)  
Batch: March 2018



Senior Data Analyst, Myntra  
(Oct 2019)



**Ayush Modi**

Associate Consultant (IT)  
Batch: March 2018



Marketing Analyst,  
Globalization Partners  
(July 2019)



**Ashish Y**

Asst. Manager  
(Manufacturing/Production)  
Batch: Sept 2018



Business Analyst, Arvind  
(Sept 2019)



**Jai Krishna**

Fresher  
Batch: Sept 2018



Business Analyst, Quantzig  
(Nov 2019)



**Shadab Hussain**

Data Scientist (IT)  
Batch: Sept 2018



Analyst, TheMathCompany  
(July 2019)



**Damodar Bandi**

Global Supply Chain Analyst  
(Software)  
Batch: Dec 2018



Sr. Business Analyst, Data  
Semantics  
(Aug 2019)



**Shubhadip B**

Technical Project Manager  
(Tech Solutions)  
Batch: June 2018



Sr. Technical Project Manager,  
Aurionpro  
(Oct 2019)



**Deepshikha**

Senior Associate (Tech)  
Batch: March 2019



Data Analyst, Amazon  
(Nov 2019)



**Rohit W**

Quality Assurance Tester  
(BFSI)  
Batch: Jan 2019



Python Automation Engineer,  
Credit Suisse  
(July 2019)



**Hariharan S**

AVP (BFSI)  
Batch: March 2018



Project Manager, HSBC  
(Aug 2019)



**Mohit Mamgain**

Internship (IT)  
Batch: March 2018



Data Analyst, Shine.com (Aug  
2019)



**Aakash Dusane**

Software Engineer (Software)  
Batch: Dec 2018



Data Scientist, Quantzig  
(Aug 2019)

**Sri Harsha Ravi**

Principal Data  
Structure Engineer (IT)  
Batch: Sept 2018



Senior Data Scientist, IHS  
Markit  
(Nov 2019)

**Sujit Nalawade**

Software Engineer (IT)  
Batch: Sept 2018



Data Analyst, Xoriant

**Sudha Choudhary**

Fresher  
Batch: June 2018



Internship - AI Engineer,  
Athancare  
(July 2019)

**Nishant Chalasany**

Project Manager (Agro)  
Batch: March 2019



Analytics Operations Lead,  
Syngenta  
(July 2019)

**Ganesh Varanasi**

Analyst (BFSI)  
Batch: March 2018



Data Scientist, Innominds  
(July 2019)

**Anshul Srivastava**

Business Analyst (BFSI)  
Batch: Sept 2018



Associate Consultant, Fractal  
(Aug 2019)

**Anshul Kumar**

Analyst (BFSI)  
Batch: March 2019



Data Science Intern,  
Merkle Sokrati  
(Aug 2019)

**Sylvester Pinto**

Senior Software Engineer (IT)  
Batch: Dec 2018



Data Associate, J.P. Morgan  
(June 2019)

**Ansuman Das**

Risk Analyst (IT)  
Batch: March 2019



Specialist Data Analyst,  
Novartis  
(Aug 2019)



# Experience upGrad Offline



## UPGRAD BASECAMPS

Held across all major cities in India, upGrad basecamps bring together learners, faculty and industry experts for a power-packed day of activities, career building sessions and live group projects. Get to know your peers and faculty, and hone your networking skills in an exciting environment.

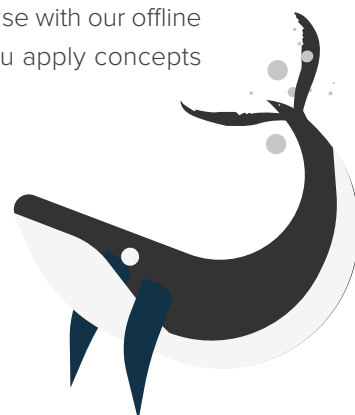
## CAREER FAIRS

Attend regular hiring drives in major cities across India, giving you the opportunity to interview with upGrad's 300+ hiring partners ensuring you get every opportunity you deserve.



## HACKATHONS

Team up and put your learning to use with our offline Hackathons: designed to help you apply concepts and meet, network, and grow!



# Hear from Our Learners

**Kunwar Alok, Experience: 15+ Years**

*"You may not believe it but I had never done coding in my life. I did it during this course and was thrilled to see the outcomes coming out of those codes. Just the way I used to get happy after solving a good (tough) math problem during my school age. Thanks to upGrad for providing a great service to people like us who at the age of 43 can dream to study with budding talents around."*

**Sachin Aggarwal, Experience: 18+ Years**

*"Learning with IITB and upGrad has been an experience like no other. Being an online program, you have your worries about how the program and teaching methods will be. My favourite part about the learning experience has been programming through well designed and thoughtful content shared by IITB professors and industry experts on upGrad platforms. Kudos to upGrad."*

**Savita Upadhaya, Experience: 4 Years**

*"It has been an amazing journey with upGrad till now. Starting with their course material to live sessions to mentor support, each helps you to always be on track and progress efficiently with the Data Science course. My sincere thanks to the entire team of upGrad and Professors of IITB for showing me the path and direction for my dream to become a Data Analyst."*

**Sidharth Mahapatra, Experience: 3 Years**

*"The concepts of R programming and Machine Learning will be taught by Prof. Chandrasekhar Ramanathan and Prof. G Srinivasaraghavan respectively. Both of them have been listed in the list of the top twenty most prominent Data Science academics published by Analytics India Magazine. So you need not worry about quality of teaching in this program."*



**Tuhin Pal, Experience: 5 Years**

*"I appreciate the platform upGrad has provided and how they have arranged modules and assignments. Modules are locked until you complete the previous one, so it feels like clearing a semester and going to the next one."*

**Harkirat Dhillon, Experience: 8 Years**

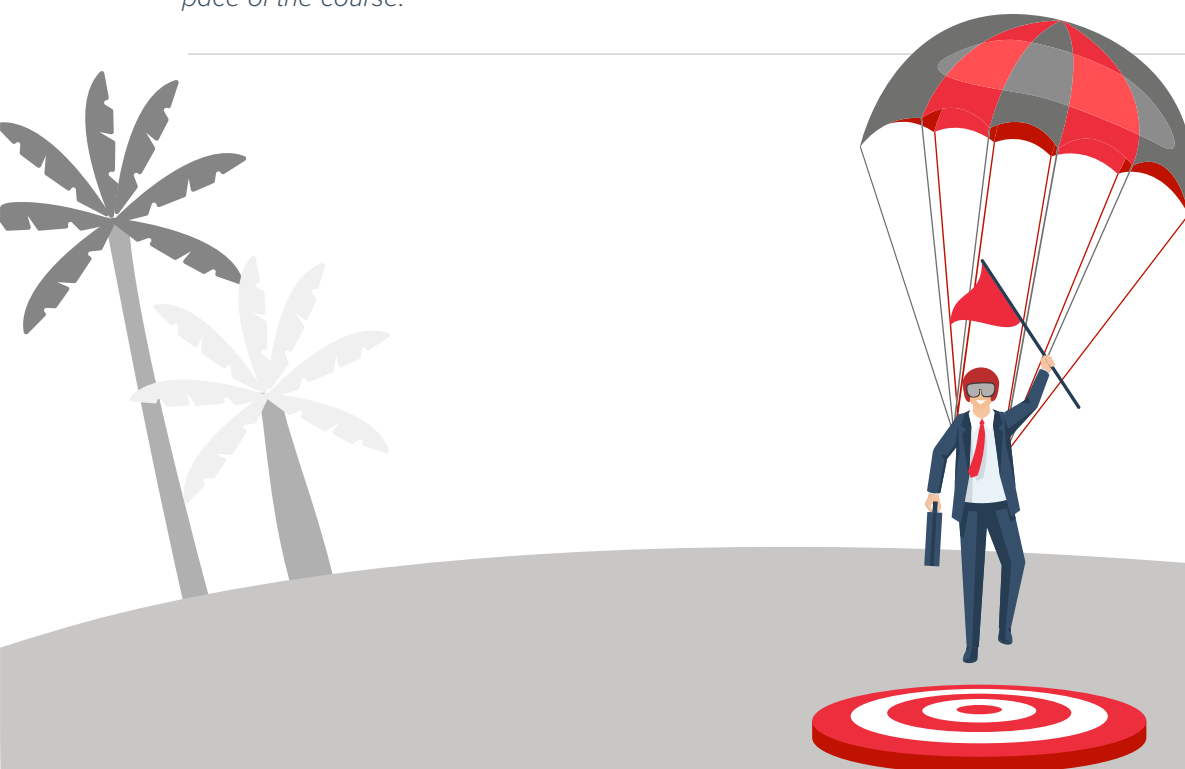
*"A dedicated studying regime is the key to be successful and pass the program. This program will help build a strong foundation for a successful transition into Data Science. Additionally, participating in Hackathons and Kaggle competitions to solve real-world problems will definitely give you an edge and land a job if one is willing to work hard."*

**Shravani Shahapure, Experience: 16 Years**

*"For someone who really wants to pursue a career in the field of Data Science, it is worth opting for the complete course by IIITB and upGrad. IIITB and upGrad's online program on Data Science gives many opportunities and develops students for their future as they provide the best professors, thought-provoking assignments and case studies."*

**Sagar Tekwani, Experience: 2 Years**

*"A very well-structured and well-balanced program content which you won't get in other programs/nano-degrees. Being a beginner in DS, I found the structure of the Executive PG Programme from IIITB and upGrad most helpful. They even teach you most of the prerequisites with prep sessions before you even start the course. Being a working professional, it was neither too difficult nor too easy to keep up with the pace of the course."*



# Program Details and Admission Process

## PROGRAM DURATION AND FORMAT

12 Months | Online

## PROGRAM FEE

INR 2,85,000 (Incl. of all taxes.)

## PROGRAM START DATES

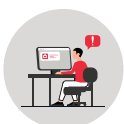
Please refer to the website for program start dates.

[www.upgrad.com/data-science-pgd-iiitb/](http://www.upgrad.com/data-science-pgd-iiitb/)

## ELIGIBILITY

Bachelor's Degree with 50% or equivalent passing marks. No coding experience required.

## WEEKLY COMMITMENT (12-15 hours/week)



### 6-7 HOURS

Asynchronous learning time.



### 6-7 HOURS

Assignments and projects.



### 1 LIVE SESSION

Every two weeks.

## SELECTION PROCESS



### STEP 1: Selection Test

Fill out an application and take a short 17-minute online test with 11 questions.



### STEP 2: Review and Shortlisting of Suitable Candidates

Our faculty will review all applications, considering the educational and professional background of an applicant and review the test scores where applicable. Following this, Offer Letters will be rolled out so you are assured a great peer group to learn and network with.



### STEP 3: Enrollment for Access to Prep Content

Make a quick block payment with assistance from our loan partners where required, receive immediate access to the prep content and begin your upGrad journey.

## FOR FURTHER INFORMATION, CONTACT

### PRIYANKA PRAJAPATI

Program Marketing Manager, Data Programs

[admissions@upgrad.com](mailto:admissions@upgrad.com)

1800 210 2020

## COMPANY INFORMATION

upGrad Education Private Limited  
Nishuvi, 75, Annie Besant Road,  
Worli, Mumbai - 400018.