

# Postgraduate Diploma Big Data Processing



## Postgraduate Diploma Big Data Processing

- » Modality: online
- » Duration: 6 months
- » Certificate: TECH Technological University
- » Dedication: 16h/week
- » Schedule: at your own pace
- » Exams: online

Website: [www.techtitute.com/pk/information-technology/postgraduate-diploma/postgraduate-diploma-big-data-processing](http://www.techtitute.com/pk/information-technology/postgraduate-diploma/postgraduate-diploma-big-data-processing)

# Index

01

Introduction

---

*p. 4*

02

Objectives

---

*p. 8*

03

Course Management

---

*p. 12*

04

Structure and Content

---

*p. 16*

05

Methodology

---

*p. 22*

06

Certificate

---

*p. 30*

# 01

# Introduction

Big Data is in full expansion. The volume of data circulating on the network is immense and its value, in many cases, even more. The processing of this information requires a plan of action in the event of any failure, especially in terms of security. Likewise, compliance with legal and regulatory requirements is essential for companies. This program specializes IT professionals in Azure Storage Cloud and data governance, as well as in the analysis, processing and processing of streaming data. All this in an online mode teaching, with a syllabus available from the first day and downloadable from any electronic device with internet connection.



# Data Driven

“

*Excel in the field of Big Data Processing thanks to this Postgraduate Diploma. Enroll in this program and deepen your knowledge in Data Governance and Cloud Streaming"*

This Postgraduate Diploma is aimed at IT professionals who wish to advance in a growing area of technology. While it is true that the term Big Data is widely used by the general population, few are really aware of the importance of a good job done by a specialist in data processing on the network.

In this teaching, professionals will acquire the necessary skills through theoretical and practical content to implement backups, establish an approach to data governance, applying policies to ensure that organizations and companies comply with legal regulations, and to analyze the process of collecting, structuring, processing and interpreting streaming data.

This program will address cutting-edge Big Data technology. A specialized teaching team that will share their knowledge in this innovative field will accompany the students during the six months of this program.

This is an opportunity to progress professionally while combining work and personal life, thanks to the 100% online mode offered by TECH. In addition, the Relearning system, based on the reiteration of content, and a wide variety of multimedia resources will facilitate learning and the acquisition of solid knowledge.

This **Postgraduate Diploma in Big Data Processing** contains the most complete and up-to-date program on the market. The most important features include:

- ◆ Practical cases presented by experts in digital Transformation
- ◆ The graphic, schematic, and practical contents with which they are created, provide practical information on the disciplines that are essential for professional practice
- ◆ Practical exercises where the self-assessment process can be carried out to improve learning
- ◆ Its special emphasis on innovative methodologies
- ◆ Theoretical lessons, questions to the expert, debate forums on controversial topics, and individual reflection assignments
- ◆ Content that is accessible from any fixed or portable device with Internet connection



*Improve your knowledge about  
Cloud programming in real time  
with this Postgraduate Diploma"*



“*Analyze the different Cloud options available and confidentially face any risk suffered by a company with this Postgraduate Diploma*”

The program's teaching staff includes professionals from the sector who contribute their work experience to this educational program, as well as renowned specialists from leading societies and prestigious universities.

Its multimedia content, developed with the latest educational technology, will allow professionals to learn in a contextual and situated learning environment, i.e., a simulated environment that will provide immersive education programmed to prepare in real situations.

The design of this program focuses on Problem-Based Learning, by means of which professionals must try to solve the different professional practice situations that are presented to them throughout the academic year. For this purpose, the student will be assisted by an innovative interactive video system created by renowned and experienced experts.

*Make the most of your knowledge and learn with TECH how to develop a virtual machine in Azure.*

*Master Apache Spark Streaming, Kafka Stream or Flink Stream and offer a good service to your clients and companies.*



# 02 Objectives

The Postgraduate Diploma in Big Data Processing allows, upon its completion, IT professionals to know how to establish which is the most appropriate Cloud for the company to which the service is offered, detect the main risks that may be faced and the solutions to be applied in each situation. In addition, students will be able to identify the main programming languages in Big Data. In this way, professionals acquire an education with updated content and according to the demands of the sector.







“

*Achieve your professional goals in Cloud Programming and specialize in a field with a wide range of job opportunities"*



## General Objectives

---

- ◆ Analyze the different approaches to cloud adoption and their contexts
- ◆ Acquire specialized knowledge to determine the appropriate Cloud
- ◆ Develop a virtual machine in Azure
- ◆ Establish the sources of threats in application development and best practices to apply
- ◆ Evaluate the differences in the specific implementations of different public Cloud vendors
- ◆ Determine the different technologies applied to containers
- ◆ Identify the key aspects of a Cloud - Native adoption strategy
- ◆ Fundamentals and evaluation of the programming languages most commonly used in Big Data, necessary for data analysis and processing





## Specific Objectives

---

### Module 1. Cloud Azure Storage

- ◆ Examine a virtual machine in Azure
- ◆ Establish the different types of storage
- ◆ Evaluate the functions of backup
- ◆ Manage Azure resources
- ◆ Analyze the different types of services
- ◆ Examine the different types of security
- ◆ Generate virtual networks
- ◆ Concretize the different network connections

### Module 2. Cloud Programming: Data Governance

- ◆ Generate specialized knowledge on data management, strategies and processing techniques
- ◆ Develop data governance strategies targeting people, processes, and tools
- ◆ Carry out data governance from ingest to preparation and usage
- ◆ Determine techniques to govern data transmission
- ◆ Establish data protection for authentication, security, backup and monitors

### Module 3. Real-Time Cloud Programming. Streaming

- ◆ Analyze the process of collecting, structuring, processing, analyzing and interpreting streaming data
- ◆ Develop the principles of streaming processing, the current context and current use cases in the national context
- ◆ Develop key fundamentals of statistics, machine learning, data mining and predictive modeling for understanding data analysis and processing
- ◆ Analyze the main Big Data programming languages
- ◆ Examine the fundamentals of Apache Spark Streaming, Kafka Stream and Flink Stream



*Master the main Big Data programming languages. Become the professional that every company wishes in its teams”*



03

# Course Management

TECH has carefully chosen the teaching team for this program, so that students get a quality education according to the demands of an emerging sector. For this purpose, the IT professional will have at his disposal a teaching team that is currently immersed in Cloud projects, and therefore know in detail the latest advances in this technology. In addition, the teaching staff has excellent academic qualifications, thus guaranteeing the students a quality education with enriched and innovative content.



“

*You will have at your disposal a relevant teaching team in the field of Big Data. They will lead you to success”*



## Management



### Mr. Bressel Gutiérrez-Ambrossi, Guillermo

- ♦ Specialist in Systems Administration and Computer Networks
- ♦ Storage and SAN Network Administrator at Experis IT (BBVA)
- ♦ Network Administrator at IE Business School
- ♦ Graduate in Computer Systems and Network Administration at ASIR (ASIR)
- ♦ Ethical Hacking course at OpenWebinar
- ♦ Powershell course at OpenWebinar

## Professors

### D. Bernal de la Varga, Yeray

- ♦ Big Data Solutions Architect at Orange Bank
- ♦ Big Data Architect at Bankia
- ♦ Big Data Engineer at Hewlett-Packard
- ♦ Adjunct Professor in the Master of Big Data at the University of Deusto
- ♦ Degree in IT from the Polytechnic University of Madrid
- ♦ Expert in Big Data by U-TAD

### Ms. Rodríguez Camacho, Cristina

- ♦ Apis consultant and developer of microservices at Inetum
- ♦ Graduate in Health Engineering, with mention in Biomedical Engineering from the University of Malaga
- ♦ Master's Degree in Blockchain and Big Data from the Complutense University of Madrid
- ♦ Expert in DevOps & Cloud at UNIR



# 04

## Structure and Content

The syllabus of this Postgraduate Diploma is structured in three blocks that deepen from the beginning in the Big Data treatment. Initially it addresses the Storage in Cloud Azure, one of the most reliable and economical data storage solutions in the cloud market. Also, in this program, great importance will be given to the correct management of data and its processing in Streaming. All this with a content rich in video summaries, complementary readings and practical examples that will facilitate learning.





“

*Deepen into the knowledge of Cloud Azure Storage and get the most out of your virtual network. Improve your professional career”*



## Module 1. Cloud Azure Storage

- 1.1. MV Installation in Azure
  - 1.1.1. Creation Commands
  - 1.1.2. Visualization Commands
  - 1.1.3. Modification Commands
- 1.2. Azure Blobs
  - 1.2.1. Types of Blobs
  - 1.2.2. Container
  - 1.2.3. Azcopy
  - 1.2.4. Reversible Blob Suppression
- 1.3. Managed Disk and Storage in Azure
  - 1.3.1. Managed Disk
  - 1.3.2. Security/Safety
  - 1.3.3. Cold Storage
  - 1.3.4. Replication
    - 1.3.4.1. Local Redundancy
    - 1.3.4.2. Redundancy in a Zone
    - 1.3.4.3. "Georedundant"
- 1.4. Azure Tables, Queues, Files
  - 1.4.1. Tables
  - 1.4.2. Queues
  - 1.4.3. Files
- 1.5. Azure Encryption and Security
  - 1.5.1. Storage Service Encryption (SSE)
  - 1.5.2. Access Codes
    - 1.5.2.1. Shared Access Signature
    - 1.5.2.2. Container-Level Access Policies
    - 1.5.2.3. Access Signature at Blob Level
  - 1.5.3. Azure AD Authentication
- 1.6. Azure Virtual Network
  - 1.6.1. Subnetting and Matching
  - 1.6.2. Vnet to Vnet
  - 1.6.3. Private Link
  - 1.6.4. High Availability
- 1.7. Types of Azure Connections
  - 1.7.1. Azure Application Gateway
  - 1.7.2. Site-to-Site VPN
  - 1.7.3. Point-to-Site VPN
  - 1.7.4. ExpressRoute
- 1.8. Azure Resources
  - 1.8.1. Blocking Resources
  - 1.8.2. Resource Movement
  - 1.8.3. Removal of Resources
- 1.9. Azure Backup
  - 1.9.1. Recovery Services
  - 1.9.2. Azure Agent Backup
  - 1.9.3. Azure Backup Server
- 1.10. Solutions Development
  - 1.10.1. Compression, Deduplication, Replication
  - 1.10.2. Recovery Services
  - 1.10.3. Disaster Recovery Plan



## Module 2. Cloud Programming: Data Governance

- 2.1. Data Management
  - 2.1.1. Data Management
  - 2.1.2. Data Handling Ethics
- 2.2. Data Governance
  - 2.2.1. Classification. Access Control
  - 2.2.2. Data Processing Regulation
  - 2.2.3. Data Governance Value
- 2.3. Data Governance. Data Science
  - 2.3.1. Lineage
  - 2.3.2. Metadata
  - 2.3.3. Data Catalog Business Glossary
- 2.4. User and Processes in Data Governance
  - 2.4.1. Users
    - 2.4.1.1. Roles and Responsibilities
  - 2.4.2. Processes
    - 2.4.2.1. Data Enrichment
- 2.5. Data Life Cycle in the Enterprise
  - 2.5.1. Data Creation
  - 2.5.2. Data Processing
  - 2.5.3. Data Storage
  - 2.5.4. Data Use
  - 2.5.5. Data Destruction
- 2.6. Data Quality
  - 2.6.1. Quality in Data Governance
  - 2.6.2. Data Quality in Analytics
  - 2.6.3. Data Quality Techniques
- 2.7. Data Governance in Transit
  - 2.7.1. Data Governance in Transit
    - 2.7.1.1. Lineage
  - 2.7.2. The Fourth Dimension

- 2.8. Data Protection
  - 2.8.1. Access Levels
  - 2.8.2. Classification
  - 2.8.3. Compliance Regulations
- 2.9. Data Governance Monitoring and Measurement
  - 2.9.1. Data Governance Monitoring and Measurement
  - 2.9.2. Lineage Monitoring
  - 2.9.3. Data Quality Monitoring
- 2.10. Data Governance Tools
  - 2.10.1. Talend
  - 2.10.2. Collibra
  - 2.10.3. IT specialist

## Module 3. Real-Time Cloud Programming. Streaming

- 3.1. Processing and Structuring of Streaming Information
  - 3.1.1. Data Collection, Structuring, Processing, Analysis, and Interpretation Process
  - 3.1.2. Streaming Data Processing Techniques
  - 3.1.3. Streaming Processing
  - 3.1.4. Streaming Processing Use Cases
- 3.2. Statistics for Understanding Streaming Data Flows
  - 3.2.1. Descriptive Statistics
  - 3.2.2. Probability Calculation
  - 3.2.3. Inference
- 3.3. Programming with Python
  - 3.3.1. Typology, Conditionals, Functions and Loops
  - 3.3.2. Numpy, Matplotlib, Dataframes, CSV Files and JSON Formats
  - 3.3.3. Sequences: Lists, Loops, Files and Dictionaries
  - 3.3.4. Mutability, Exceptions and Higher-Order Functions

- 3.4. R Programming
  - 3.4.1. R Programming
  - 3.4.2. Vector and Factors
  - 3.4.3. Matrix and Array
  - 3.4.4. Lists and Data Frame
  - 3.4.5. Functions
- 3.5. SQL Database for Streaming Data Processing
  - 3.5.1. SQL Databases
  - 3.5.2. Entity-Relationship Model
  - 3.5.3. Relational Model
  - 3.5.4. SQL
- 3.6. Non-SQL Database for Streaming Data Processing
  - 3.6.1. Non-SQL Databases
  - 3.6.2. MongoDB
  - 3.6.3. MongoDB Architecture
  - 3.6.4. CRUD Operations
  - 3.6.5. Find, Projections, Index Aggregation and Cursors
  - 3.6.6. Data Model
- 3.7. Data Mining and Predictive Modeling
  - 3.7.1. Multivariate Analysis
  - 3.7.2. Dimension Reduction Techniques
  - 3.7.3. Cluster Analysis
  - 3.7.4. Sets
- 3.8. Machine Learning for Streaming Data Processing
  - 3.8.1. Machine Learning and Advanced Predictive Modeling
  - 3.8.2. Neural Networks.
  - 3.8.3. Deep Learning
  - 3.8.4. Bagging and Random Forest
  - 3.8.5. Gradient Bosting
  - 3.8.6. SVM
  - 3.8.7. Assembly Methods





- 3.9. Streaming Data Processing Technologies
  - 3.9.1. Spark Streaming
  - 3.9.2. Kafka Streaming
  - 3.9.3. Flink Streaming
- 3.10. Apache Spark Streaming
  - 3.10.1. Apache Spark Streaming
  - 3.10.2. Spark Components
  - 3.10.3. Spark Architecture
  - 3.10.4. RDD
  - 3.10.5. SPARK SQL
  - 3.10.6. Jobs, Stages and Tasks

“

*Stay ahead of your competition With this Postgraduate Diploma, you will obtain knowledge with up-to-date content and a flexible online methodology”*

# 05 Methodology

This academic program offers students a different way of learning. Our methodology uses a cyclical learning approach: **Relearning**.

This teaching system is used, for example, in the most prestigious medical schools in the world, and major publications such as the **New England Journal of Medicine** have considered it to be one of the most effective.





“

*Discover Relearning, a system that abandons conventional linear learning, to take you through cyclical teaching systems: a way of learning that has proven to be extremely effective, especially in subjects that require memorization"*



## Case Study to contextualize all content

Our program offers a revolutionary approach to developing skills and knowledge. Our goal is to strengthen skills in a changing, competitive, and highly demanding environment.

“

*At TECH, you will experience a learning methodology that is shaking the foundations of traditional universities around the world”*



*You will have access to a learning system based on repetition, with natural and progressive teaching throughout the entire syllabus.*



*The student will learn to solve complex situations in real business environments through collaborative activities and real cases.*

### A learning method that is different and innovative

This TECH program is an intensive educational program, created from scratch, which presents the most demanding challenges and decisions in this field, both nationally and internationally. This methodology promotes personal and professional growth, representing a significant step towards success. The case method, a technique that lays the foundation for this content, ensures that the most current economic, social and professional reality is taken into account.

“*Our program prepares you to face new challenges in uncertain environments and achieve success in your career”*

The case method has been the most widely used learning system among the world's leading Information Technology schools for as long as they have existed. The case method was developed in 1912 so that law students would not only learn the law based on theoretical content. It consisted of presenting students with real-life, complex situations for them to make informed decisions and value judgments on how to resolve them. In 1924, Harvard adopted it as a standard teaching method.

What should a professional do in a given situation? This is the question that you are presented with in the case method, an action-oriented learning method. Throughout the course, students will be presented with multiple real cases. They will have to combine all their knowledge and research, and argue and defend their ideas and decisions.

## Relearning Methodology

TECH effectively combines the Case Study methodology with a 100% online learning system based on repetition, which combines different teaching elements in each lesson.

We enhance the Case Study with the best 100% online teaching method: Relearning.

*In 2019, we obtained the best learning results of all online universities in the world.*

At TECH you will learn using a cutting-edge methodology designed to train the executives of the future. This method, at the forefront of international teaching, is called Relearning.

Our university is the only one in the world authorized to employ this successful method. In 2019, we managed to improve our students' overall satisfaction levels (teaching quality, quality of materials, course structure, objectives...) based on the best online university indicators.





In our program, learning is not a linear process, but rather a spiral (learn, unlearn, forget, and re-learn). Therefore, we combine each of these elements concentrically.

This methodology has trained more than 650,000 university graduates with unprecedented success in fields as diverse as biochemistry, genetics, surgery, international law, management skills, sports science, philosophy, law, engineering, journalism, history, and financial markets and instruments. All this in a highly demanding environment, where the students have a strong socio-economic profile and an average age of 43.5 years.

*Relearning will allow you to learn with less effort and better performance, involving you more in your training, developing a critical mindset, defending arguments, and contrasting opinions: a direct equation for success.*

From the latest scientific evidence in the field of neuroscience, not only do we know how to organize information, ideas, images and memories, but we know that the place and context where we have learned something is fundamental for us to be able to remember it and store it in the hippocampus, to retain it in our long-term memory.

In this way, and in what is called neurocognitive context-dependent e-learning, the different elements in our program are connected to the context where the individual carries out their professional activity.



This program offers the best educational material, prepared with professionals in mind:



#### Study Material

All teaching material is produced by the specialists who teach the course, specifically for the course, so that the teaching content is highly specific and precise.

These contents are then applied to the audiovisual format, to create the TECH online working method. All this, with the latest techniques that offer high quality pieces in each and every one of the materials that are made available to the student.



#### Classes

There is scientific evidence suggesting that observing third-party experts can be useful.

Learning from an Expert strengthens knowledge and memory, and generates confidence in future difficult decisions.



#### Practising Skills and Abilities

They will carry out activities to develop specific skills and abilities in each subject area. Exercises and activities to acquire and develop the skills and abilities that a specialist needs to develop in the context of the globalization that we are experiencing.



#### Additional Reading

Recent articles, consensus documents and international guidelines, among others. In TECH's virtual library, students will have access to everything they need to complete their course.







#### Case Studies

Students will complete a selection of the best case studies chosen specifically for this program. Cases that are presented, analyzed, and supervised by the best specialists in the world.



#### Interactive Summaries

The TECH team presents the contents attractively and dynamically in multimedia lessons that include audio, videos, images, diagrams, and concept maps in order to reinforce knowledge.

This exclusive educational system for presenting multimedia content was awarded by Microsoft as a "European Success Story".



#### Testing & Retesting

We periodically evaluate and re-evaluate students' knowledge throughout the program, through assessment and self-assessment activities and exercises, so that they can see how they are achieving their goals.



# 06 Certificate

This Postgraduate Diploma in Big Data Processing guarantees students, in addition to the most rigorous and up-to-date education, access to a Postgraduate Diploma issued by TECH Technological University.



“

*Successfully complete this program and receive your postgraduate certificate without having to travel or fill out laborious paperwork”*



This **Postgraduate Diploma in Big Data Processing** contains the most complete and up-to-date program on the market.

After the student has passed the assessments, they will receive their corresponding **Postgraduate Diploma**, issued by **TECH Technological University** via tracked delivery\*.

The certificate issued by **TECH Technological University** will reflect the qualification obtained in the Postgraduate Diploma, and meets the requirements commonly demanded by labor exchanges, competitive examinations, and professional career evaluation committees.

Title: **Postgraduate Diploma in Big Data Processing**

Official N° of Hours: **450 h.**



\*Apostille Convention. In the event that the student wishes to have their paper certificate issued with an apostille, TECH EDUCATION will make the necessary arrangements to obtain it, at an additional cost.



## Postgraduate Diploma Big Data Processing

- » Modality: **online**
- » Duration: **6 months**
- » Certificate: **TECH Technological University**
- » Dedication: **16h/week**
- » Schedule: **at your own pace**
- » Exams: **online**

# Postgraduate Diploma Big Data Processing