

Postgraduate Certificate Data Mining Processing and Transformation



Postgraduate Certificate Data Mining Processing and Transformation

- » Modality: online
- » Duration: 6 weeks
- » Certificate: TECH Global University
- » Credits: 6 ECTS
- » Schedule: at your own pace
- » Exams: online

Website: www.techtute.com/us/information-technology/postgraduate-certificate/data-mining-processing-transformation

Index

01

Introduction

p. 4

02

Objectives

p. 8

03

Course Management

p. 12

04

Structure and Content

p. 16

05

Methodology

p. 20

06

Certificate

p. 28

01

Introduction

Data mining can be simply defined as an analysis prior to the use of machine learning techniques in Data Science. The latter helps computer engineers to extract the maximum value from the data. With the program, students will learn the different statistical inference procedures to analyze and determine the study of variables, their distribution and values. All this is a career advancement for students who wish to specialize in this area.



“

*Develop your skills as a
computer engineer in Data
Science and Data Mining”*

This Postgraduate Certificate will analyze the theoretical bases that help computer engineers develop advanced knowledge about the different existing data preparation techniques for data cleaning, normalization and transformation. It will also present the necessary tools to evaluate different methodologies in search of errors that may cause problems in the work environment.

The entire program is composed of a series of case studies that will favor the learning of students who seek to further advance their professional careers and challenge themselves to achieve excellence.

All this will be feasible thanks to a 100% online program, which adapts to the daily needs of its students. You will only need a device with an Internet connection to start developing a complete professional profile with international projection.

This **Postgraduate Certificate in Data Mining Processing and Transformation** contains the most complete and up-to-date academic program on the market. The most important features of the program include:

- ◆ Practical cases studies are presented by experts in Engineering in data analysis
- ◆ The graphic, schematic, and eminently practical contents with which they are created, provide scientific and practical information on the disciplines that are essential for professional practice
- ◆ Practical exercises where self-assessment can be used 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 an Internet connection



Evaluate the various methodologies presented to identify advantages and drawbacks”

“*Develop the necessary skills for data identification, preparation and transformation”*

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

The multimedia content, developed with the latest educational technology, will provide the professional with situated and contextual learning, i.e., a simulated environment that will provide immersive training programmed to train in real situations.

The design of this program focuses on Problem-Based Learning, which means the student must try to solve the different real-life situations of that arise throughout the academic program. This will be done with the help of an innovative, interactive video system developed by renowned experts with extensive experience in Data Mining Processing and Transformation.

Specify effective and efficient procedures for data processing according to the type of problem presented.

Turn your career around and start developing in-company improvement strategies.



02 Objectives

The knowledge provided on this program will help computer engineers approach the analysis prior to applying machine learning techniques with the objective of extracting the maximum value from the data. That way, they will be able to gain specialized knowledge based on statistical analysis. To that end, TECH has established the following general and specific objectives.



“

*Demonstrate your ability
to interpret data visualization
for descriptive analysis”*

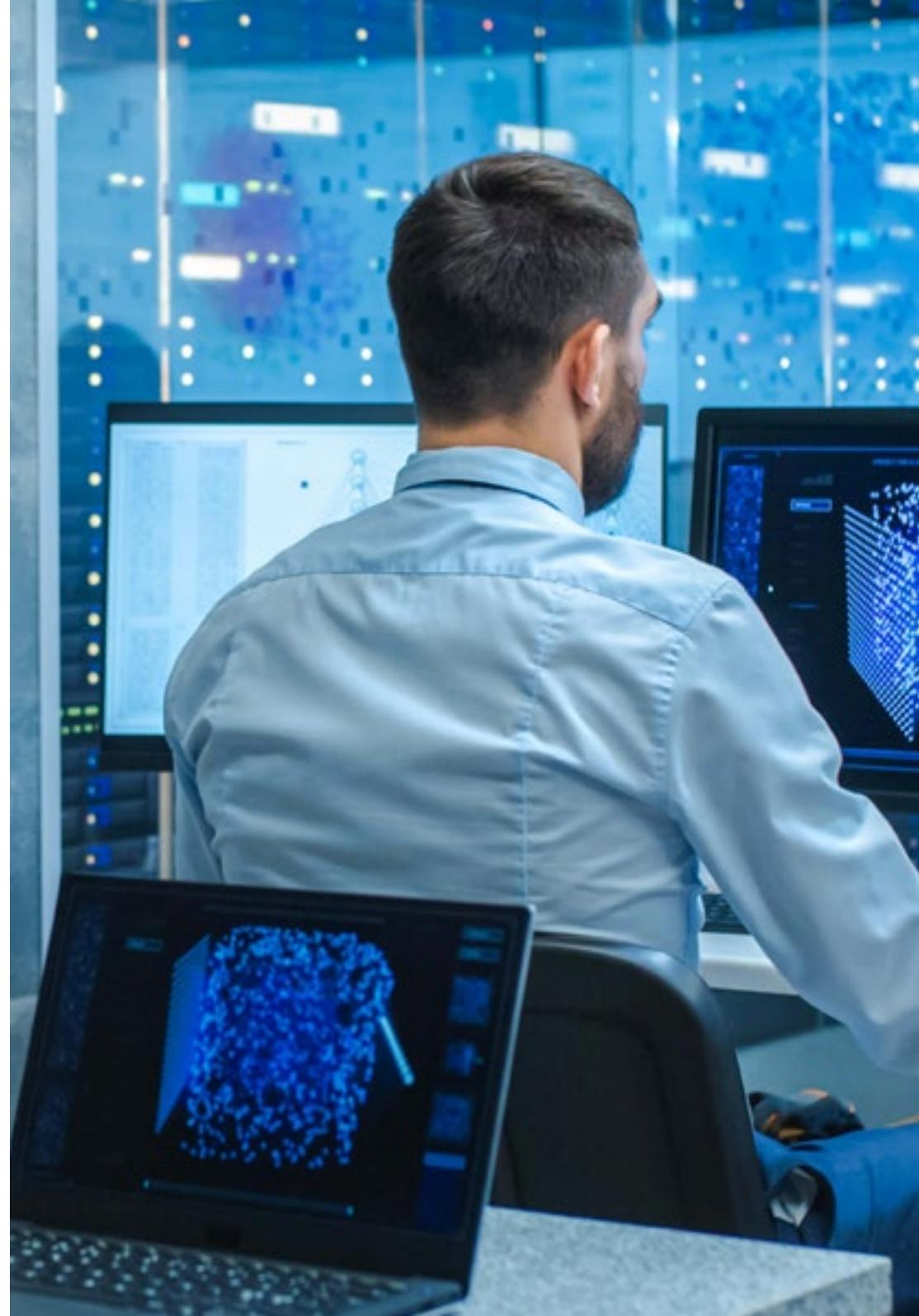


General Objectives

- ◆ Analyze the benefits of applying data analytic techniques in every company department
- ◆ Develop the basis for understanding the needs and applications of each department
- ◆ Generate specialized knowledge to select the right tool
- ◆ Propose techniques and objectives in order to be as productive as possible according to the department



Examine the problems in your work environment and find viable solutions using data as the basis for your proposals”





Specific Objectives

- ◆ Generate specialized knowledge about the statistical prerequisites for any data analysis and evaluation
- ◆ Develop the necessary skills for data identification, preparation and transformation
- ◆ Evaluate the various methodologies presented and identify advantages and drawbacks
- ◆ Examine the problems in high dimensional data environments
- ◆ Implement algorithms used for data preprocessing
- ◆ Demonstrate the ability to interpret data visualization for descriptive analysis
- ◆ Develop advanced knowledge of the different existing data preparation techniques for data cleaning, normalization and transformation



03

Course Management

The Postgraduate Certificate in Data Mining Processing and Transformation brings together a select group of professionals with multiple years of experience in data analysis in the field of business. Students are thus guaranteed that the knowledge imparted comes from professionals capable of answering any questions they may have while providing them with real cases to better exemplify the contents on the program.





“

With several years of experience in the sector, this group of experts will at all times help you understand the concepts on the program in an easy and didactic way”

International Guest Director

Dr. Tom Flowerdew is an internationally prominent personality in the field of data science. He has served as Vice President of Data Science at MasterCard in London. In this role, he has been responsible for the preparation, operation and strategy of a consolidated team in this area, with the mission to support a portfolio of innovative products in payments, fight against money laundering (AML) and analyze cryptocurrency use cases.

He has also been Director of Data Science in Cyber Intelligence Solutions, also at MasterCard, where he has led the integration of data to support revolutionary cryptocurrency-based products. In fact, his ability to handle complex data and develop advanced solutions has been instrumental to the success of multiple projects in the cybersecurity and finance space.

Similarly, for the company Featurespace, he has held several crucial roles, including Head of Standardized Product Delivery, in Cambridge, leading a team and a transformation project that has reduced delivery time and effort by more than 75%. In addition, as Head of Delivery, U.S. headquarters, he has managed all of the company's North American delivery functions, significantly improving operational efficiency and strengthening customer relationships.

Additionally, Doctor Tom Flowerdew has demonstrated his ability to build and lead high-performing teams throughout his career, most notably in his role as Data Scientist, both in Atlanta, where he has recruited and managed a group of experts in the field, and in Cambridge. In this way, his focus on innovation and problem solving has left an indelible mark on the organizations where he has worked, establishing himself as an influential leader in the field of data science.



Dr. Flowerdew, Tom

- Vice President of Data Science at MasterCard, London, United Kingdom
- Director of Data Science, Cyber Intelligence Solutions, MasterCard, London
- Head of Standardized Product Delivery at Featurespace, Cambridge
- Delivery Manager for United States, at Featurespace, Cambridge
- Data Scientist at Featurespace, Atlanta, Georgia, United States
- Data Scientist at Featurespace, Cambridge
- Research Fellow in Statistics and Operations Research at the University of Lancaster
- Ph.D. in Operations Research from Lancaster University
- B.S. in Systems Engineering from BAE Systems
- B.Sc. in Mathematics from the University of York



Thanks to TECH, you will be able to learn with the best professionals in the world"

Management



Dr. Peralta Martín-Palomino, Arturo

- CEO and CTO at Prometheus Global Solutions
- CTO at Korporate Technologies
- CTO in AI Shephers GmbH
- Doctorate in Psychology from the University of CastillaLa
- PhD in Economics, Business and Finance from the Camilo José Cela University. Outstanding Award in her PhD
- PhD in Psychology, University of CastillaLa Mancha
- Master's Degree in Advanced Information Technologies from the University of Castilla la Mancha
- Master MBA+E (Master's Degree in Business Administration and Organisational Engineering) from the University of Castilla la Mancha.
- Associate lecturer, teaching undergraduate and master's degrees in Computer Engineering at the University of Castilla la Mancha
- Professor of the Master in Big Data and Data Science at the International University of Valencia
- Lecturer of the Master's Degree in Industry 4.0 and the Master's Degree in Industrial Design and Product Development
- Member of the SMILe Research Group of the University of Castilla la Mancha



Professors

Mr. Montoro Montarroso, Andrés

- ◆ Researcher in the SMILe Group at the University of Castilla-La Mancha.
- ◆ Data Scientist at Prometheus Global Solutions
- ◆ Graduate in Computer Engineering from the University of Castilla-La Mancha.
- ◆ Master's Degree in Data Science and Computer Engineering from the University of Granada
- ◆ Guest lecturer in the subject of Knowledge-Based Systems at the Escuela Superior de Informática de Ciudad Real, Giving the Lecture "Advanced Artificial Intelligence Techniques: Search and Analysis of Potential Social Media Radicals".
- ◆ Guest lecturer in the subject of Data Mining at the Escuela Superior de Informática de Ciudad Real giving the lecture: "Applications of Natural Language Processing: Fuzzy Logic to the analysis of messages in social networks"
- ◆ Speaker at the Seminar on Corruption Prevention in Public Administrations and Artificial Intelligence. Faculty of Law and Social Sciences of Toledo. Conference entitled "Artificial Intelligence Techniques" Speaker at the first International Seminar on Administrative Law and Artificial Intelligence (DAIA). Organised by Centro de Estudios Europeos Luis Ortega Álvarez and Institut de Recerca TransJus Conference entitled "Sentiment Analysis for the prevention of hate speech on social media"

04

Structure and Content

The modules on this program provide a theoretical and practical perspective to examine state-of-the-art data cleaning techniques, transformation, dimensionality reduction, as well as feature and instance selection. Thus will the objectives of the program to train professional, integral and prestigious engineers be fulfilled.

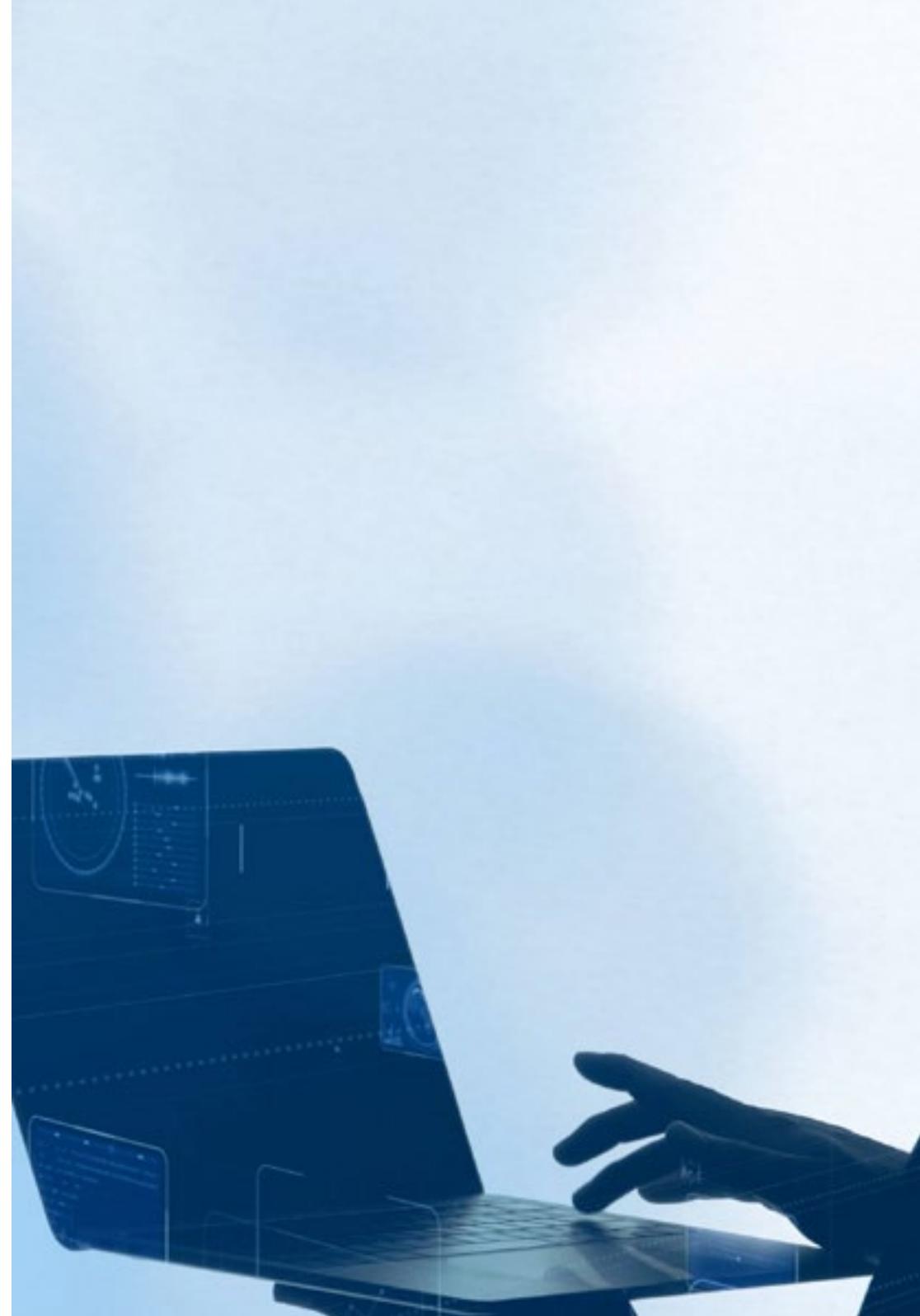




“Develop advanced knowledge of the different existing data preparation techniques for data cleaning, normalization and transformation”

Module 1. Data Mining: Selection, Pre-Processing and Transformation

- 1.1. Statistical Inference
 - 1.1.1. Descriptive Statistics vs. Statistical Inference
 - 1.1.2. Parametric Procedures
 - 1.1.3. Non-Parametric Procedures
- 1.2. Exploratory Analysis
 - 1.2.1. Descriptive Analysis
 - 1.2.2. Visualization
 - 1.2.3. Data Preparation
- 1.3. Data Preparation
 - 1.3.1. Integration and Data Cleaning
 - 1.3.2. Normalization of Data
 - 1.3.3. Transforming Attributes
- 1.4. Missing Values
 - 1.4.1. Treatment of Missing Values
 - 1.4.2. Maximum Likelihood Imputation Methods
 - 1.4.3. Missing Value Imputation Using Machine Learning
- 1.5. Noise in the Data
 - 1.5.1. Noise Classes and Attributes
 - 1.5.2. Noise Filtering
 - 1.5.3. The Effect of Noise
- 1.6. The Curse of Dimensionality
 - 1.6.1. Oversampling
 - 1.6.2. Undersampling
 - 1.6.3. Multidimensional Data Reduction
- 1.7. From Continuous to Discrete Attributes
 - 1.7.1. Continuous vs. Discrete Data
 - 1.7.2. Discretization Process



- 1.8. The Data
 - 1.8.1. Data Selection
 - 1.8.2. Prospects and Selection Criteria
 - 1.8.3. Selection Methods
- 1.9. Instance Selection
 - 1.9.1. Methods for Instance Selection
 - 1.9.2. Prototype Selection
 - 1.9.3. Advanced Methods for Instance Selection
- 1.10. Data Preprocessing in Big Data Environments
 - 1.10.1. Big Data
 - 1.10.2. "Conventional" Vs. Mass Pre-Processing
 - 1.10.3. Smart Data

“

*Completing this program
will allow students to better
understand data selection
methods”*

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.



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

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

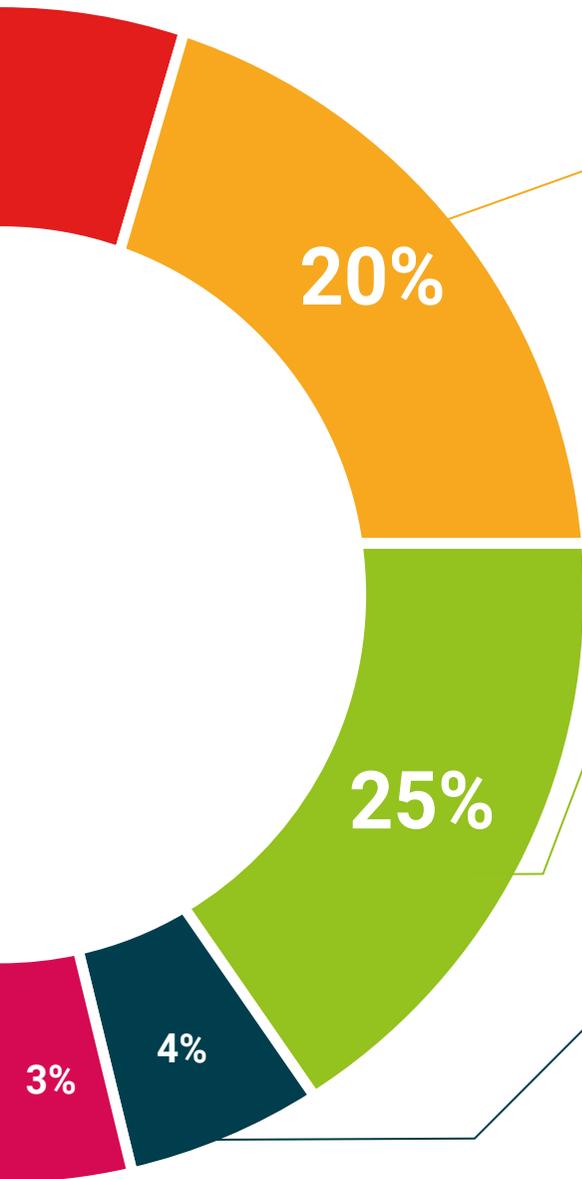
They will carry out activities to develop specific skills and abilities in each subject area. Exercises and activities to acquire and develop the skills and abilities that a specialist needs to develop in the context of the globalization that we are experiencing.



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

The Postgraduate Certificate in Data Mining Processing and Transformation guarantees students, in addition to the most rigorous and up-to-date education, access to a Postgraduate Certificate issued by TECH Global University.



“

Successfully complete this program and receive your university qualification without having to travel or fill out laborious paperwork”

This program will allow you to obtain your **Postgraduate Certificate in Data Mining Processing and Transformation** endorsed by **TECH Global University**, the world's largest online university.

TECH Global University is an official European University publicly recognized by the Government of Andorra (*official bulletin*). Andorra is part of the European Higher Education Area (EHEA) since 2003. The EHEA is an initiative promoted by the European Union that aims to organize the international training framework and harmonize the higher education systems of the member countries of this space. The project promotes common values, the implementation of collaborative tools and strengthening its quality assurance mechanisms to enhance collaboration and mobility among students, researchers and academics.

This **TECH Global University** title is a European program of continuing education and professional updating that guarantees the acquisition of competencies in its area of knowledge, providing a high curricular value to the student who completes the program.

Title: **Postgraduate Certificate in Data Mining Processing and Transformation**

Modality: **online**

Duration: **6 weeks**

Accreditation: **6 ECTS**



health future
confidence people
education information tutors
guarantee accreditation teaching
institutions technology learning
community commitment
personalized service innovation
knowledge present
online training
development language
classroom



Postgraduate Certificate Data Mining Processing and Transformation

- » Modality: online
- » Duration: 6 weeks
- » Certificate: TECH Global University
- » Credits: 6 ECTS
- » Schedule: at your own pace
- » Exams: online

Postgraduate Certificate Data Mining Processing and Transformation

```
ngSwitchWatchAction(valu  
elements.length; i < ii; ++i)  
move());  
length = 0;  
selectedScopes.length; i < ii; ++i) {  
selected = selectedElements[i];  
selectedScopes[i].$destroy();  
previousElements[i] = selected;  
animate.leave(selected, function() {  
previousElements.splice(i, 1);  
});  
}  
selectedElements.length = 0;  
selectedScopes.length = 0;  
if ((selectedTranscludes = ngSwitchController.selectedTranscludes
```