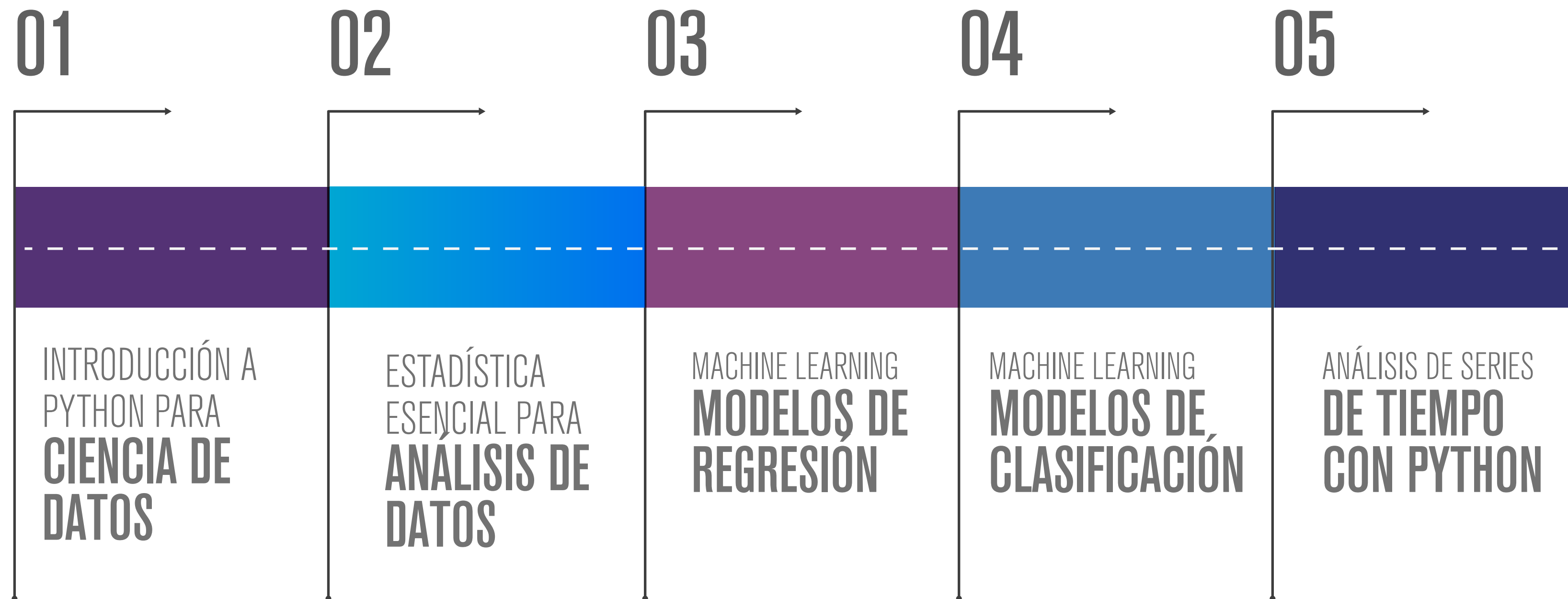


ESPECIALISTA EN
**ANALÍTICA AVANZADA DE
DATOS CON PYTHON**

CURSOS QUE COMPONEN LA ESPECIALIDAD



Python es uno de los lenguajes de programación más utilizados por los Analistas y Científicos de Datos. Nos permite realizar análisis exploratorios de datos, visualizaciones, modelos de machine learning, deep learning y mucho más.

No importa si no tienes una gran experiencia en programación ya que el curso inicia desde el nivel básico y te llevará por los fundamentos de Python, luego aprenderás manipulación y limpieza de datos y por último la construcción de visualizaciones de datos.

¡Aprende este lenguaje y entra al mundo de la analítica avanzada!

TEMARIO A CUBRIR:

1. Introducción a Ciencia de Datos

- a. ¿Qué es un Científico de Datos?
- b. Rango de acción del Científico de Datos

2. Fundamentos de Python

- a. Importando Datos
 - i. Desde archivos Excel (pandas)
 - ii. Desde archivos Planos (pandas)
 - iii. Desde SQL Server (pyodbc)
 - iv. Desde SAS (SAS7BDAT)
 - v. Desde Stata (pandas)
 - vi. Desde Carpetas (glob)
 - vii. Web Scrappin (BeautifulSoup)
 - viii. HTML (pandas + requests)
- b. Tuplas, Listas y Diccionarios
- c. Definiendo Variables
- d. Estableciendo el Directorio (os)
- e. Matrices (numpy)
 - i. Matriz apartir de listas y datos simulados
 - ii. Matrices Unidimensional o Bidimensionales
 - iii. Unión de matrices

- f. Operadores Aritméticos
- g. Introducción a Bucles FOR
- h. Creación y Exploración del DataFrame
- i. Creando DataFrames
- ii. Primeras y Últimas Filas
- iii. Metadata, Estadísticas
- iv. Valores, Columnas, índices y tipos de Datos

3. Manipulación y limpieza de datos

- a. Selección de Variables y Observaciones (columnas y filas)
- b. Validación de Tipo de Datos
- c. Trabajando con el índice
- d. LOC & ILOC
- e. Filtros básicos y avanzados
- f. Función AND – OR – NOT -ISIN - UPPER
- g. Función IF-ELIF-ELSE & y con bucle FOR

- h. Ordenando el Dataset
- i. Columnas calculadas
- j. Eliminando Duplicados
- k. Trabajando con valores vacios (reemplazar, eliminar, rellenar, convertir)
- l. Agrupaciones simples y múltiples con diferentes agregaciones
- m. Pivot Table
- n. Append, concat , merge, join
- o. Merge con % de aceptación (fuzzywuzzy)
- p. Cambiar caracteres especiales, validaciones de tipos de datos, limpiar, trabajando con datos categóricos.
- q. Diferencia entre Tablas de consulta
- r. Valores Perdidos
- s. Trabajando con Fechas (datetime + pandas)
- i. Cambiar el formato de fecha

- ii. Resta entre fecha
- iii. Agrupación de Fechas
- iv. Obtención de Datos a partir de fechas
- t. Validación Cruzada de Filas

4. Visualización de datos (matplotlib)

- a. Gráficos Generales
 - i. Gráficos de Líneas
 - ii. Gráficos de Barras Horizontales y Verticales
 - iii. Gráficos Apilados
 - iv. Gráficos de Dispersión
- b. Visualizaciones Estadísticas (Seaborn)
 - i. Boxplot
 - ii. Histogramas
 - iii. Gráfico Densidad
 - iv. Stripplot
 - v. Swarmplot
 - vi. Lvplot
 - vii. violinplot

growup
DATA ANALYTICS

INTRODUCCIÓN A PYTHON PARA CIENCIA DE DATOS

Precio: c97.920 / Duración: 20 horas / Requisitos: Manejo de Excel Avanzado

 python™



ESTADÍSTICA ESENCIAL PARA ANÁLISIS DE DATOS

En este curso aprenderás como esta disciplina científica es indispensable en nuestra preparación como científicos de datos, utilizando los mejores métodos y procedimientos estadísticos para obtener las conclusiones óptimas con la información suministrada.

El curso cubre el Análisis Exploratorio de los Datos (EDA), el cual nos permite conocer “¿De qué se trata?” la información que estaremos trabajando, y nos ayuda a comprender y vislumbrar distribuciones y patrones en los datos. También estaremos navegando sobre la Inferencia de la estadística que nos facilita obtener las mejores deducciones a partir de las hipótesis planteadas y por último terminaremos construyendo nuestra probabilidad realizando supuestos para que un determinado evento suceda.

Con el curso serás capaz de obtener las mejores conclusiones para la toma de decisiones en tu empresa.

TEMARIO A CUBRIR:

1) Exploración de Datos y Estadística Descriptiva

- a) Variables Cualitativas y Cuantitativas
- b) Medidas e interpretación de Estadística Descriptiva
- c) Exploración y distribución de los Datos

2) Estadística Inferencial

- a) Prueba de Hipótesis
- b) Región Crítica
- c) P Value
- d) Permutaciones
- e) Región Crítica
- f) Intervalos de Confianza

3) Probabilidad

- a) Distribución Binomial
- b) Distribución Hipergeométrica
- c) Distribución Poisson
- d) Distribución Exponencial
- e) Distribución Uniforme
- f) Distribución Normal
- g) Probabilidad Condicional
- h) Teorema de Bayes
- i) Eventos dependientes e independientes

Precio: €97.920 / **Duración:** 20 horas / **Requisitos:** Introducción a Python para Ciencia de Datos



En este curso aprenderás a estandarizar, modelar, predecir y elegir los mejores resultados para tus datos.

El curso cubre los algoritmos principales de Aprendizaje Supervisado en modelos de Regresión llevando a cabo diferentes etapas en proyectos de Ciencia de Datos. Aprenderás la lógica y razonamiento matemático en cada algoritmo, la importancia de los HiperParámetros y como evaluar los resultados

TEMARIO A CUBRIR:

1) Introducción a Machine Learning

- a) Que es Machine Learning
- b) Tipos de Aprendizaje en ML
- c) Conceptos Básicos de Regresión

2) Preprocesamiento de Datos

- a) EDA (repaso)
- b) Manejo de Variables categóricas
- c) Train - test - split
- d) Método Scale
- e) Método StandarScaler
- f) Método SimpleImputer
- g) Método RobustScaler

- h) Método MinMaxScaler / MaxAbsScaler

3) Modelos de Regresión

- a) Regresión Lineal
- b) Regresión Lineal Múltiple
- c) Regularización Ridge
- d) Regularización Lasso
- e) Regularización ElasticNet
- f) Extreme Gradient Boosting Regressor
- g) Regresión de Soporte Vectorial (SVR)

4) Hiperparámetros & Pipelines

- a) Underfitting & Overfitting

- b) RandomizedSearchCV
- c) GridSearchCV
- d) Introducción a Pipelines

5) Desempeño Predictivo

- a) Puntuación r^2
- b) Error promedio cuadrático (RMSE)
- c) Puntuación de Varianza Explicada
- d) Error promedio Absoluto (MAE)

A man with glasses is looking at a computer screen. The screen displays various data visualizations, including a line graph and a scatter plot. The background is dark with blue and white light effects, suggesting a high-tech or data-driven environment. The overall tone is professional and focused.

MACHINE LEARNING MODELOS DE REGRESIÓN

Precio: €97.920 / **Duración:** 20 horas / **Requisitos:** Introducción a Python para Ciencia de Datos

MACHINE LEARNING MODELOS DE CLASIFICACIÓN

En este curso aprenderás el uso de algoritmos de Machine Learning de Clasificación, con los cuales podremos utilizar nuestros datos históricos para clasificar diferentes categorías. Con el uso de Machine Learning Modelos de Clasificación podrás responder a preguntas de gran utilidad en los negocios como por ejemplo:

¿Es una transacción fraudulenta o no?, ¿Un vuelo llegará a tiempo o no?, ¿El cliente pagará su crédito o no?, ¿El tipo de tumor es benigno o maligno?, ¿Al cliente le gustará esta película o no?, ¿Una máquina fallará o no?, ¿Qué tipo de deporte realiza un usuario de un dispositivo? (Caminar, nadar, correr, etc), ¿Qué tipo de campaña publicitaria dirigimos a cada cliente?

Esto lo lograrás a través del uso de algunos algoritmos que estudiaremos en este curso, entre ellos:

Naive Bayes, k Nearest Neighbors, Regresión Logarítmica, Árboles de Decisión, Clustering Jerárquico, Clustering kmeans, Análisis de Componentes Principales, Máquinas de Soporte Vectorial, entre otros

Adquirir estas nuevas habilidades te permitirán responder las interrogantes del negocio con analítica avanzada de datos. ¡Empieza ya con este curso y aprende técnicas utilizadas en la Ciencia de los Datos!

TEMARIO A CUBRIR:

1) Introducción a Machine Learning

- a) Qué es Machine Learning
- b) Tipos de Aprendizaje en ML
- c) Conceptos Básicos de Clasificación

2) Preprocesamiento de Datos

- a) EDA (repass)
- b) Manejo de Variables categóricas y continuas
- c) Train - test - split
- d) Normalización de Datos (StandardScaler, MinMaxScaler, entre otros)
- e) Probabilidades y Distribuciones
- f) Correlación

3) Modelos de Clasificación

- a) Árboles de Decisión
- b) K-Nearest Neighbors
- c) Naive Bayes
- d) Random Forest
- e) Regresión Logística
- f) Máquinas de Soporte Vectorial (RBF, Lineal, Polinomial)

4) Aprendizaje No Supervisado

- a) Análisis de Componentes Principales (PCA)

- b) Clúster K-Means
- c) Clúster Jerárquico

5) Métricas o Métodos de evaluación

- a) Reporte de Clasificación
- b) Matriz de confusión
- c) Técnicas de precisión del modelo
- d) Feature Importance
- e) Curva ROC /AUC
- f) Método del Codo
- g) Dendrograma

Precio: € 97.920 / **Duración:** 20 horas / **Requisitos:** Introducción a Python para Ciencia de Datos



A hand holding a tablet with a futuristic data visualization overlay. The overlay features a line graph with blue and yellow points, a bar chart with blue bars, and a grid of blue and yellow lines. The background is dark blue with a subtle pattern of plus signs and dots.

ANÁLISIS DE SERIES DE TIEMPO CON PYTHON

Una serie de tiempo es un modo estructurado de representar datos. Por lo que visualmente, es una curva que evoluciona a lo largo del tiempo. En general, se pueden representar cualquier tipo de datos cronológicamente almacenados.

El objetivo central de trabajar con series de tiempo en cualquier sector es identificar los patrones de los datos, lo que se busca es entender el acontecimiento a medida que avanza el tiempo para predecir y pronosticar el futuro.

Es importante mencionar que los análisis de series de tiempo suelen seguir algunas características especiales, entre ellas que suelen tener una tendencia, estar auto-correlacionadas y son dependientes del tiempo.

En este curso aprenderás a manipular, desplazar, descomponer y visualizar las series de tiempo para obtener resultados utilizando técnicas tales como Medias Móviles, ARIMA, entre otros .

Matricúlate y construye un modelo paso a paso, que te permita predecir la variable en estudio. Además, aprenderás a realizar visualizaciones dinámicas para presentar en reportes formales.

TEMARIO A CUBRIR:

1) Introducción a Series de Tiempo

- a) Google Colab
- b) Qué son Series de Tiempo
- c) Exploración de Datos
- d) Correlación / Correlogramas
- e) Procesos Estacionarios
- f) Descomposición de Series de Tiempo

2) Modelos Según Metodología Box Jenkins

- a) Modelo Autorregresivo
- b) Modelo Media Móvil
- c) Modelo ARMA & ARIMA
- d) Modelo Suavización Exponencial

- e) Modelo de Aislamiento Exponencial
- f) Modelo Estacional ARIMA (SARIMA)
- g) Modelo GARCH & ARCH
- h) Volatilidad

3) Análisis de Suavizamiento

- a) Múltiples Modelos
- b) Enfoque Suavizamiento
- c) Modelo de Predicción ETS

4) Visualización

- a) Técnicas de visualización de series de tiempo

Modalidad: Asincrónica / **Precio Modalidad Asincrónica:** c52,865 / **Duración:** 9 horas / **Requisitos:** Introducción a Python para Ciencia de Datos



INFORMACIÓN GENERAL

ORDEN DE LOS CURSOS Y PRECIOS:

- Introducción a Python para Ciencia de Datos
Precio: c97,920 (\$180 USD) / Duración: 20 horas
- Estadística Esencial para Análisis de Datos
Precio: c97,920 (\$180 USD) / Duración: 20 horas
- Machine Learning – Modelos de Regresión
Precio: c97,920 (\$180 USD) / Duración: 20 horas
- Machine Learning – Modelos de Clasificación
Precio: c97,920 (\$180 USD) / Duración: 20 horas
- Análisis de Series de Tiempo con Python / Duración: 9 horas
Precio Modalidad Asincrónica: c52,865 (\$97 USD)

PRECIO TOTAL: c444,545 (\$817 USD) por persona IVA incluido

DURACIÓN TOTAL: 90 horas

REQUISITOS DE APROBACIÓN:

Se obtiene un certificado de aprovechamiento al aprobar cada curso con una nota mayor a 70. Si se culminaron todos los cursos con nota mayor a 70 se entrega el certificado final como Especialista en Ciencia de Datos y Business Analytics con Python.

EXAMEN DE UBICACIÓN VS EXAMEN DE CONVALIDACIÓN:

La persona puede realizar una prueba de ubicación para determinar si tiene el nivel adecuado para entrar directamente a uno de los cursos en específico, sin embargo, para completar el programa de especialidad debe hacer examen de convalidación para validar los conocimientos en este tema, el cual tiene un costo de c15000 (\$25 USD) y una duración de 2 horas. O bien, puede presentar certificados de cursos similares de otras instituciones o universidades.

CERTIFICACIÓN INTERNACIONAL:

Una vez el estudiante haya completado la totalidad del programa, podrá optar por aplicar el examen de certificación internacional de Microsoft (esto es opcional y tendrá un costo por aparte a cancelar directamente con el ente evaluador):

DP-100: Designing and Implementing a Data Science Solution on Azure

Al aprobar el examen de certificación de Microsoft, el estudiante obtendrá la certificación:

Microsoft Certified: Azure Data Scientist Associate

SEDE:

SAN PEDRO

Del Mall San Pedro, 300 m Norte, 50 m Oeste, Edificio Omala, 2do piso, local #2.

MODALIDADES:

- **Sincrónica** con instructor en vivo.
- **In-House:** El curso se brindaría en las instalaciones que asigne la empresa, esta debe contar con sala de capacitación con proyector y equipo de cómputo con los programas requeridos previamente instalados. Si la empresa está fuera del GAM se cobrará un adicional por concepto de viáticos.
- **Escuela Virtual:** Ofrecemos opciones para brindar el curso de manera virtual utilizando la mejor tecnología para capacitaciones asincrónicas

Programa respaldado por la política de calidad de cursos Grow Up:

<https://www.growupcr.com/politicadecalidad>

¿QUIERES LLEVAR ESTA ESPECIALIDAD EN MODALIDAD ONLINE A TU PROPIO RITMO?

Puedes estudiar esta Especialidad desde nuestra Escuela Virtual tomando los cursos en el horario que desees y desde donde desees.

Más detalles: info@growupcr.com

growup
DATA ANALYTICS

MECÁNICA DE LA ESPECIALIDAD