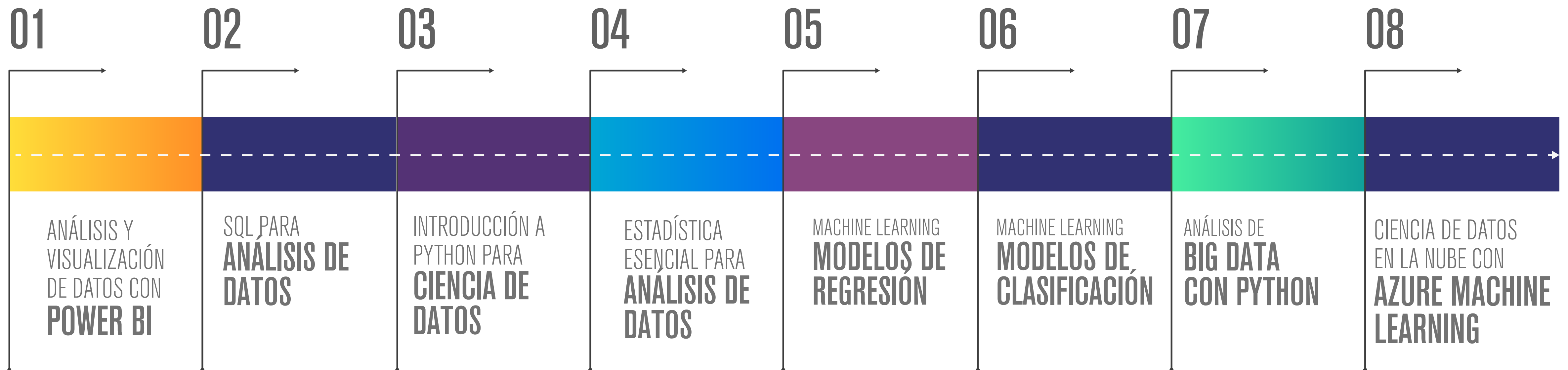


ESPECIALISTA EN  
**CIENCIA DE DATOS Y  
BUSINESS ANALYTICS CON PYTHON**



# CURSOS QUE COMPONEN LA ESPECIALIDAD



¿Necesita crear reportes a partir de miles o millones de datos, incluso de diferentes fuentes pero solo conoce Excel para hacer esto? Con las nuevas herramientas de Microsoft como Power BI usted podrá visualizar y analizar datos con mayor velocidad, eficientemente y de manera sobresaliente. Sea capaz de crear increíbles dashboards que incluyan mapas, velocímetros, tablas, segmentaciones y muchas otras funciones más para mostrar sus datos de manera sorprendente. Incluso podrá actualizar y compartir sus datos en tiempo real y sus dashboards podrán ser consultados en cualquier lugar y prácticamente a través de cualquier dispositivo ya sea en una computadora, Tablet o un teléfono celular al otro lado del mundo.\*

Este curso sobre Power BI impartido por Grow Up le permitirá poner en práctica de inmediato las principales funciones de esta herramienta y crear sus primeros dashboards a partir de ejemplos reales con su propia información. Podrá crear dashboards combinando información que se encuentre en diferentes fuentes como Excel, páginas web, Facebook, Google Analytics o muchas otras más.

\*Las capacidades de Power BI dependerán de la versión de Office con la que cuente cada participante y si dispone de la versión gratuita o cuenta con licencia de Power BI Pro.

### TEMARIO A CUBRIR:

1. Introducción a Power BI
2. Obtención y depuración de datos
3. Modelado de datos
4. Visualizaciones – Dashboards sorprendentes
5. Exploración de datos
6. Power BI y Excel
7. Publicación y uso compartido de la información
8. Introducción a Funciones DAX  
SUM, COUNT, COUNTA, AVERAGE  
DISTINCTCOUNT  
COUNTROWS

CALCULATE  
IF, AND, OR  
FILTER  
ALL  
RANKX  
TOPN  
CALENDAR  
Funciones X (SUMX, AVERAGEX)  
Time Intelligence (DATESMTD, DATESQTD, DATESYTD,  
SAMEPERIODLASTYEAR)

**Precio:** c97,920 / **Duración:** 18 horas / **Requisitos:** Manejo de Excel Avanzado

# ANÁLISIS Y VISUALIZACIÓN DE DATOS CON POWER BI







# SQL PARA ANÁLISIS DE DATOS

La cantidad de datos que generamos diariamente ha crecido de tal manera que ahora para las organizaciones es vital contar con personas con habilidades especiales para tratar estos datos, analizarlos correctamente y transformarlos en posibles decisiones de negocios que se transformen en una ventaja competitiva para la organización. Esto sería tarea para un Analista de Datos o Científico de Datos, que por cierto son de las profesiones más demandadas y atractivas del momento actual, y para lograrlo estas personas deben dominar correctamente el lenguaje SQL para extraer adecuadamente la información de estas inmensas bases de datos.

En este curso asumimos que son tus primeros pasos en SQL y que no tienes un conocimiento previo de este lenguaje, así que comenzaremos desde lo más básico con un enfoque en aprender a consultar y modificar datos para analizarlos con objetivos de ciencia de datos, haciendo las preguntas correctas para obtener las respuestas adecuadas de nuestros datos.

El enfoque del curso es totalmente práctico con énfasis en escenarios de la vida real que te permitirán poner en uso esos conocimientos prácticamente de inmediato. Incluso aprenderás cómo llevar esos datos ya modelados a Excel o Power BI para presentar una visualización de la información mucho más enriquecida.

¡Si tu objetivo es trabajar como Analista o Científico de Datos, este curso definitivamente será un excelente comienzo para ti!

## TEMARIO A CUBRIR:

### 1. INTRODUCCIÓN

#### Introducción a Bases de Datos Relacionales

- Modelo Entidad-Relación
- Tablas (Columnas de Tablas)
- Relaciones (Llaves primarias y foráneas)
- Tipos de Datos
- Normalización (Integridad, Consistencia)

#### Lógica de Predicados y Operadores

- Operadores Booleanos
- Operadores de comparación
- Operadores de Proximidad IN, BETWEEN, LIKE
- Conectores Lógicos
- Valores nulos o inexistentes

#### Microsoft SQL Server

- Servidor
- Base de Datos
- Servicios
- Instancias
- Esquema
- Objeto

#### Instalación de MS SQL Server

- Configuración
- Conociendo la interfaz de Desarrollo
- Creación de Bases de Datos
- Respaldo y Recuperación de Bases de Datos

### 2. TRABAJO CON DATOS

#### Exploración de Datos con la sentencia SELECT

- SELECT
- FROM
- WHERE
- ORDER BY

#### Funciones de Conversión de Datos

- Cast
- Convert

#### Funciones de Caracter

- Substring
- Left
- Right
- Len
- CharIndex
- Replace
- Upper
- Lower

#### Funciones de Fecha

- GetDate
- Year
- Month
- Day
- DatePart
- DateDiff
- IsDate

#### Funciones para Valores Nulos

- IsNull
- NullIf
- Coalesce

#### Funciones de Agregación

- Sum()
- Count()
- Avg()
- Min()
- Max()

#### Columnas Calculadas

- Alias de Tablas
- Operadores Aritméticos
- Funciones

#### Agrupar, limitar y filtrar conjunto de resultados

- Top
- Group By
- Having
- Distinct
- Exists

#### Consultar varias tablas

- Inner
- Outer

#### Window Functions

- Ranking
- Aggregate

#### Combinando consultas

- Union
- Intersect

#### Funciones de Fecha

- LTRIM
- RTRIM
- TRIM

#### Otras Funciones Útiles

- Case When
- Choose

#### SubConsultas

- Internas
- Relacionadas

#### Expresiones de Tabla

- Tablas derivadas
- Expresiones de Tabla Común
- Vistas
- Funciones que retornan una tabla

#### Extracción y Carga de Datos en SQL Server

- Archivos CSV delimitados por caracteres especiales
- Archivos de Excel
- Interacción con otras bases de Datos SQL Server en el mismo servidor.

### 3. INTERACCIÓN DE SQL SERVER CON EXCEL Y POWER BI

#### Conectividad entre Excel-PowerBI y SQL Server

- Creación de tablas dinámicas con Excel a partir de consultas realizadas a SQL Server
- Creación de Gráficas con Power BI a partir de consultas realizadas a SQL Server
- Creación de gráficas (histogramas, diagramas de dispersión)

#### Análisis Exploratorio de Datos a partir de Datos procesados en SQL Server

- Conceptos de población, muestra, parámetro, estadística, desviación estándar
- Comprender los conjuntos de Datos (cualitativos, cuantitativos)
- Determinar importancia de las variables (columnas de datos)



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

# INTRODUCCIÓN A PYTHON PARA CIENCIA DE DATOS

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





# 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 optimas 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 (repass)
- 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 displaying data visualizations. The background is dark with blue and white light effects, suggesting a data science or machine learning environment. The text 'MACHINE LEARNING MODELOS DE REGRESIÓN' is overlaid on the image in a white, bold, sans-serif font.

## 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





Todo lo que has aprendido hasta el momento sobre Python es de mucha utilidad, pero ¿qué pasa si tienes que trabajar con Big Data? Cada vez es más común que las empresas cuenten con millones de registros provenientes de diferentes orígenes de datos y en diferentes formatos.

En este curso aprenderás los beneficios de trabajar Python con Apache Spark y utilizarás el paquete PySpark para ejecutar operaciones muchísimo más rápido que con técnicas tradicionales de análisis (hasta 100 veces más rápido).

Aprenderás desde la limpieza y manipulación de datos en entornos Big Data hasta la construcción de modelos de Machine Learning e Ingeniería de Características (Feature Engineering).

Este curso hará la diferencia entre un analista de datos normal y un analista de datos capaz de afrontar situaciones reales que se encontrará en las grandes empresas que generan Big Data, las cuales son cada vez más.

¡Suscríbete en el curso y prepárate para solucionar problemas del mundo real!

## TEMARIO A CUBRIR:

### ▪ Fundamentos de PySpark

Introducción a PySpark  
Análisis de Big Data con Apache Spark  
Programación en PySpark RDDs  
SQL con PySpark

### ▪ Ingeniería de Características (Feature Engineering) con PySpark

Manipulación de Dataframes con PySpark

Mejorar el desempeño de los modelos  
Uso de Pipelines con PySpark  
Análisis Exploratorio de Datos (EDA)  
Manejo de datos faltantes o innecesarios  
Ingeniería de Características (Feature Engineering)  
Construcción de Modelos

### ▪ Machine Learning con PySpark

Modelos de Clasificación  
Modelos de Regresión  
Modelos de Clusterización  
Modelos Conjuntos (Ensemble Models)  
Motores de Recomendación con PySpark

# ANÁLISIS DE BIG DATA CON PYTHON





# CIENCIA DE DATOS EN LA NUBE CON AZURE MACHINE LEARNING

Con tus conocimientos previos en Python, SQL y modelos de Machine Learning, es un buen momento para llevar tus capacidades a otro nivel utilizando las posibilidades que existen para hacer Ciencia de Datos desde la nube con una poderosa herramienta de uno de los fabricantes líderes en la industria. Aprenderás el proceso de inicio a fin para crear, entrenar y evaluar modelos de Machine Learning desde la nube utilizando Azure Machine Learning Studio y Azure Machine Learning Services. Además, conocerás cómo consumir desde otras aplicaciones los resultados de tus experimentos.

Ha sido un largo camino para obtener los conocimientos para llegar hasta acá, pero sin duda valdrá la pena para poder implementar tus modelos de Machine Learning de manera profesional.

¡Anímate a ir más allá como Científico de Datos!

## TEMARIO A CUBRIR:

### 1. Ciencia de Datos en Azure Machine Learning Studio

- Introducción a Azure ML Studio
- Importación y Administración de Datos
- Cantidad de datos necesaria (overfitting y underfitting)
- Transformaciones con el editor de SQL
- Transformaciones con Scripts
- Limpieza de datos y manejo de datos faltantes
- Normalización de datos
- Modelos de Machine Learning de Regresión
- Modelos de Machine Learning de Clasificación
- Evaluación de varios modelos de Machine Learning
- PCA en Azure ML Studio
- Hyper tuning en Azure ML Studio
- Validación cruzada en Azure ML Studio

### 2. Web Service en Azure Machine Learning Studio

- Creación de un Web Service en Azure ML Studio
- Consumir Web Service desde Power BI Desktop

### 3. Fundamentos de Azure

### 4. Introducción a Servicios de Ciencia de Datos en Azure

### 5. Azure Machine Learning Services

- Entrenar un modelo con Azure ML Services
- Automatizar la selección del modelo con Azure ML Services
- Registrar y desplegar modelos de ML con Azure ML Services

### 6. Azure Data Science Virtual Machine

- Tipos de Azure Data Science Virtual Machines
- Provisión y uso de una Azure Data Science Virtual Machine

### 7. Extraer conocimiento con Azure Databricks

- Lectura y escritura de datos con Azure Databricks

**Precio:** c 97.920 / **Duración:** 20 horas / **Requisitos:** Machine Learning – Modelos Predictivos, Machine Learning Modelos de Clasificación, Introducción a SQL para Ciencia de Datos



## INFORMACIÓN GENERAL

### ORDEN DE LOS CURSOS, PRECIOS Y DURACIÓN:

- Análisis y Visualización de Datos con Power BI  
Precio: c97,920 (\$180 USD) / Duración: 18 horas
- SQL para Análisis de Datos  
Precio: c97,920 (\$180 USD) / Duración: 18 horas
- 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 Big Data con Python  
Precio: c97,920 (\$180 USD) / Duración: 20 horas
- Ciencia de Datos en la Nube  
Precio: c97,920 (\$180 USD) / Duración: 20 horas

PRECIO TOTAL: c783,360 por persona IVA incluido (\$1.440 USD)

DURACIÓN TOTAL: 154 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](mailto:info@growupcr.com)

# MECÁNICA DE LA ESPECIALIDAD



INICIA TU RUTA PARA CONVERTIRTE EN UN  
**CIENTIFICO DE DATOS**  
**CON PYTHON**