

TÉCNICO EN ANALÍTICA DE DATOS

MODALIDAD MIXTA (EN VIVO - VIRTUAL)

En un mundo que genera datos a una velocidad cada vez mayor desde infinidad de fuentes, solo las empresas capaces de analizar esta información correctamente y obtener de ella las conclusiones adecuadas, serán capaces de sobresalir e incluso sobrevivir.

Organizaciones lucrativas o no lucrativas tienen la necesidad cada vez más urgente de contar en sus equipos con Técnicos en Analítica de Datos que sean capaces no solo de decir, con base en los datos, qué sucedió y por qué sucedió, sino que además puedan aportar en la construcción de modelos que permitan predecir qué va a suceder y cómo la empresa puede actuar para aprovechar estas oportunidades detectadas.

¿Cómo puedes aportar a las empresas siendo Técnico en Analítica de Datos?

Serás capaz de construir dashboards y reportes automatizados que permitan entender rápidamente lo que ha sucedido y en dónde están las oportunidades para la empresa.

Podrás colaborar en la creación de modelos de Machine Learning que ayuden a la empresa a predecir lo que va a suceder y que se tomen las mejores decisiones basados en datos.

Podrás aportar con el mantenimiento de la calidad de los datos almacenados en las bases de datos empresariales y podrás consultar y facilitar los datos necesarios para que otros también puedan construir sus reportes y análisis

Esta es sin duda una de las ocupaciones más demandada de la actualidad sin importar el campo o industria en la que te desempeñes. Prepárate para destacar por tus capacidades para hacer hablar a los datos.

Este programa es para ti si:

- Eres profesional de cualquier campo con deseos de especializarse en la Analítica de Datos para extraer conclusiones más poderosas de los datos.
- Eres profesional que desea prepararse para obtener un nuevo rol de tiempo completo en posiciones de alta demanda en el campo del Análisis de Datos.
- Si eres egresado de secundaria con bachillerato completo y deseas incursionar en roles laborales de Analítica de Datos a nivel técnico.

Requisitos:

Bachillerato de secundaria completo.

Manejo del idioma inglés a nivel de lectura.

Disponer de una computadora portátil.

Plan de Estudios:

* La duración estimada del programa completo es de 1 año.

* Los precios mostrados incluyen el 2% de IVA.

* Los cursos los puedes tomar desde cualquier lugar del mundo mediante conexión remota 100% en vivo.

CURSOS QUE COMPONEN EL PROGRAMA TÉCNICO



ANÁLISIS DE DATOS PARA LA TOMA DE DECISIONES

En este curso aprenderás sobre las prácticas que siguen las mejores empresas para basar sus decisiones en el uso de datos adecuadamente tratados. Aprenderás sobre los diferentes tipos de Analítica y cómo se podrían utilizar en la empresa para obtener mayor provecho de cada uno de ellos. Conocerás el proceso para pasar de una pregunta de negocio a una pregunta analítica. Además, aprenderás sobre los términos y tecnologías disponibles como Analista de Datos y por último conocerás cómo se pueden aplicar metodologías como DMAIC y SCRUM a la mejora de procesos y administración de proyectos de Analítica de Datos.

1. Introducción al estilo Analítico

- a. Analítica y Transformación del Negocio
- b. Clasificación de Tipos de Analítica
- c. Aplicaciones comunes de Analítica en Negocios

2. El Proceso de Analítica de Negocios

- a. El Proceso de Analítica
- b. Herramientas en el Proceso de Analítica
- c. Roles en el Equipo de Analítica
- d. De una Pregunta de Negocio a una Pregunta Analítica

3. Las Oportunidades y Retos de los Datos

- a. Orígenes de Datos en una Organización
- b. Pasar de Activos y Actividades a Datos
- c. Procedencia y Calidad de los Datos
- d. Logística de Datos

4. Generalidades de Mecánica de los Datos

- a. Esquemas de Datos y Data sets
- b. Transformaciones de Datos
- c. Limpieza de Datos, Normalización y Mejora

5. Analítica Descriptiva

- a. Explicar qué pasó
- b. Caso de Negocio de Analítica Descriptiva

6. Analítica Predictiva

- a. Flujo de Trabajo de la Analítica Predictiva
- b. Generalidades de Machine Learning
- c. Caso Práctico de Machine Learning: Clasificación
- d. Caso Práctico de Machine Learning: Regresión y Recomendación
- e. Casos de Estudio de Machine Learning

7. Evitar Errores Comunes en Proyectos de Data Analytics

- a. Sesgo de Selección
- b. Fuga de Datos

- c. Correlación vs Causalidad
- d. Modelos Multi-Nivel
- e. Diferencias en los conjuntos de Entrenamiento/Producción

8. DMAIC en Data Analytics

- a. Etapa Define
- b. Etapa Measure
- c. Etapa Analyze
- d. Etapa Improve
- e. Etapa Control

9. SCRUM en Proyectos de Data Analytics

- a. Estructura del Equipo
- b. Valores SCRUM
- c. Product Backlog y Sprint Reviews
- d. Dificultades y Beneficios de SCRUM en Data Analytics

10. Caso Final de Analítica de Negocio



EXCEL INTERMEDIO

FUNCIONES PARA ANÁLISIS DE DATOS



En el curso de Excel: Funciones para Análisis de Datos aprenderás a utilizar muchas funciones de cálculo para una enorme variedad de usos que te permitirán hacer un tratamiento y análisis de datos mucho más rápido y eficiente. Tareas que antes te tomaban días las podrías reducir a tan solo unos minutos (tenemos muchos ejemplos reales de personas que vivieron estos casos y que ahora ahorran muchísimo tiempo, incluso días). Conoceremos de funciones para análisis de condiciones, para búsqueda correcta de datos, para análisis estadístico básico, para manejo de cálculos con bases de datos, o bien, para tratar texto, fechas u horas. Aprenderemos maneras útiles para validar y prevenir que la información ingresada en nuestros reportes sea la correcta, así como formas de proteger nuestra información. Para participar de este curso solo requieres un manejo básico de Excel.

Sin duda los conocimientos adquiridos en este curso te ayudarán enormemente en tu trabajo y/o universidad a la vez que empiezan a abrirte nuevas puertas hacia opciones de crecimiento profesional.

TEMARIO A CUBRIR:

1. Funciones Básicas

Funciones básicas
Función "Si"
Función "Si anidada"
Función "Y"
Función "O"
Función "Contar.Si"
Función "Sumar.Si"
Función "Promedio.Si"
Función "Contar.Si.Conjunto"
Función "Sumar.Si.Conjunto"
Función "Promedio.Si.Conjunto"

2. Funciones de Búsqueda y Referencia

Funciones: HOY, AHORA, MES, AÑO,
DIA SEM, NUM SEMANA, TEXTO
Funciones: HORA, MINUTO, SEGUNDO

Días.lab
Días.lab.intl
Día.lab
Día.lab.intl
Días360

3. Funciones de Texto

Funciones: CONCATENAR,
UNIRCADENAS Y &
Funciones: NOMPROPIO,
MAYUSC,
MINUSC, ESPACIOS
Funciones: IZQUIERDA, DERECHA,
EXTRAE, ENCONTRAR, LARGO
Funciones: SUSTITUIR y
REEMPLAZAR

4. Validaciones de Datos

Creación de Validaciones de Datos
Mensajes de Entrada y de Error
Validación de Número Entero
Listas de Validación y Listas
Dependientes
Validaciones de Fecha y Hora
Validaciones de Largo de Texto
Validaciones con fórmulas
personalizadas

5. Herramientas de Datos y Protección de Información

Mover datos de Texto a Columnas
Remover Duplicados
Proteger y Ocultar Celdas
Proteger Hojas y libros

En el curso de Excel Avanzado aprenderás a manejar enormes cantidades de datos con gran rapidez y a obtener resúmenes ejecutivos de estos datos en solo segundos. Tareas que antes te tomaban horas o incluso días podrían verse reducidas a solo minutos.

Veremos desde la manera de crear nuevas Bases de Datos hasta cómo validar que la información ingresada sea correcta, además conocerás las funciones y cálculos más utilizados con Bases de Datos y la manera de crear y manipular vistosas Tablas Dinámicas y Gráficos Dinámicos.

¡Si quieres destacar en el uso de Excel y ser cada vez más eficiente entonces este curso es para ti!

Nota Importante: Los complementos de Excel BI no funcionan en computadoras con sistema operativo Mac OS.

TEMARIO A CUBRIR:

1. Tablas Dinámicas

- Creación de Tablas Dinámicas en Excel
- Creación de Sub-Tablas
- Actualización de datos
- Creación de Tablas Dinámicas con fuente de origen externa
- Configuración de Campos
 - Formas de Resumir Valores por
 - Formas de Mostrar valores como
 - Personalizar nombres
 - Aplicar formatos numéricos
- Campos calculados

- Diseño de la Tabla Dinámica
- Práctica solución de casos con Tablas Dinámicas

3. Gráficos Dinámicos

- 4. Segmentación de Datos
- Creación de Dashboards con Tablas Dinámicas

- 5. Complementos de Excel para Inteligencia de Negocios
- Power Pivot

- Creación y edición de Relaciones
- Creación de Nuevas Columnas y Medidas con fórmulas
- DAX básicas
- Tabla de Fechas
- Power Query
 - Carga de datos al Modelo Relacional
 - Carga de datos desde diferentes orígenes
 - Edición de consultas

EXCEL
AVANZADO

ANÁLISIS DE DATOS CON TABLAS DINÁMICAS Y EXCEL BI



¿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 | COUNTROWS |
| 2. Obtención y depuración de datos | CALCULATE |
| 3. Modelado de datos | IF, AND, OR |
| 4. Visualizaciones –
Dashboards sorprendentes | FILTER |
| 5. Exploración de datos | ALL |
| 6. Power BI y Excel | RANKX |
| 7. Publicación y uso compartido
de la información | TOPN |
| 8. Introducción a Funciones DAX
SUM, COUNT, COUNTA, AVERAGE
DISTINCTCOUNT | CALENDAR |
| | Funciones X (SUMX, AVERAGEX) |
| | Time Intelligence (DATESMTD,
DATESQTD, DATESYTD,
SAMEPERIODLASTYEAR) |

Modalidad: Sincrónica / Precio: c\$97.920 / Duración: 18 horas / Requisito: 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 trasformarlos 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 Boleanos
- 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)



INTELIGENCIA DE NEGOCIOS APLICADA CON POWER BI

Este curso está diseñado para quienes quieren ir más allá con sus habilidades de Power BI al trabajar seriamente como Analistas de Datos. Nos concentraremos en aprender muchas funciones DAX para tareas específicas que nos permitirán hacer análisis más elaborados.

Además aprenderemos combinaciones de técnicas (Tips &Tricks) con las cuales lograremos hacer análisis mucho más vistosos y útiles al tiempo que abriremos las puertas a la verdadera Inteligencia de Negocios.

Como parte de este curso se incluye paralelamente el modelado de datos con herramientas de Excel BI (Power Pivot y Power Query) para aquellas personas que por alguna razón no puedan tener acceso a Power BI.

¡Este curso no es para cualquier persona . . . es para quienes quieren ir más allá y liderar la Inteligencia de Negocios en sus empresas!

TEMARIO A CUBRIR:

Modelado y Funciones DAX para Power BI (casos aplicados)

- Funciones DAX más avanzadas en Power BI (Calendarios con DAX, Variables en DAX, Jerarquía de Funciones, FILTER, ALL, ALLSELECTED, ALLEXCEPT, RANK, TOPN, EARLIER, SUMMARIZE, SWITCH, DATEADD, DATEDIFF, USERELATIONSHIP, ISFILTERED, RAND, RELATED, RELATEDTABLE, HASONEVALUE, HASENEFILTER, SEARCH, FIND, FORMAT, UNICHAR, REPT, VALUES, UNION, SELECTCOLUMNS, IN, CONTAINSROW, DATESINPERIOD, DATESBETWEEN, LOOKUPVALUE, SUBSTITUTE,

CROSSFILTER, NETWORKDAYS, SELECTEDVALUE, SUMIF y COUNTIF, GENERATESERIES, CALCULATETABLE, VALUE, VALUES, EXCEPT, INTERCEPT.

- Casos de Negocio
Refinar el Uso de Mapas en Power BI
Agregar Indicadores KPI
Cambiar Medidas con una Segmentación de Datos
Cambiar Tipos de Moneda con una Segmentación de Datos
Análisis de Escenarios de Utilidad con Parámetros
Análisis de Clientes Destacados con Parámetros
Clasificación de Clientes según su desempeño
Análisis de Clientes Nuevos vs Clientes Perdidos

Modelos de Pronósticos Básicos con Power BI
Análisis de Cohortes (tiempo entre eventos)
Análisis de Venta Cruzada
Estado de Pérdidas y Ganancias P&L
Parámetros de Consultas (Query Parameters)
Conexiones a otros orígenes de datos (Web, SQL, Google Analytics, Facebook, GitHub)

- Manipulación y transformación de datos con Power Query
- Introducción a Análisis Estadístico con Power BI

Este curso es la puerta de entrada a Python, uno de los lenguajes más populares en ciencia de datos. Iniciando desde los fundamentos, exploraremos la historia de la ciencia de datos y su relevancia en la actualidad. A través de clases prácticas, aprenderás a manejar estructuras de datos (listas, tuplas, diccionarios), realizar operaciones en álgebra lineal, y desarrollar habilidades en control de flujo y funciones personalizadas. También nos adentraremos en el manejo de DataFrames y la manipulación avanzada de datos: importación desde diversas fuentes, limpieza de datos, y visualización mediante gráficos dinámicos. Este curso establece una base sólida para todas las siguientes etapas de la especialidad.

TEMARIO A CUBRIR:

Introducción a Python y Ciencia de Datos

- a. Historia de la Ciencia de Datos
- b. Historia de Python y su aplicación a la Ciencia de Datos
- 2. Fundamentos de Programación con Python
- Variables y sus tipos
- b. Estructuras de Datos
- Listas
- ii. Tuplas
- iii. Diccionarios
- iv. Conjuntos
- c. Algebra Lineal
- Escalares
- ii. Vectores
- iii. Matrices
- iv. Tensores
- d. Funciones Personalizadas
- e. Estructuras de Control de Flujo

- Condicionales (If, elif, else)
- ii. Bucles (for, while)
- iii. Control de Bucles (break, continue)
- iv. Control de Excepciones (try, except, finally)
- f. DataFrames
- 3. Importación de Datos
- a. Excel
- b. Archivos Planos
- c. Carpetas
- d. SQL Server
- e. Web Scrapping
- 4. Manipulación y transformación de Datos
- Gestión de Índices (LOC – ILOC)
- b. Selección de Variables
- c. Filtros de Datos
- d. Operadores Lógicos – Búsqueda
- e. Condicionales

- f. Ordenamiento de Datos
- g. Agrupaciones
- h. Funciones de Agregación
- Pivot Table
- j. Combinación de Tables (Merge – JOIN)
- k. Fuzzy Wuzzy (Combinación de similitud)
- l. Unión de Tablas
- m. Funciones de Fecha
- n. Expresiones Regulares
- o. Funciones de Texto
- 5. Limpieza de Datos
- a. Validación y Gestión de Duplicados
- b. Coerción de Formato en los Datos
- c. Validación Cruzada (Integridad de los datos)

- d. Identificación de Valores Nulos
- e. Eliminación - Reemplazar Vacíos
- f. Función Anónima Lambda para rellenar valores
- 6. Visualización de Datos
- a. Gráfico de líneas
- b. Gráficos de Barras
- c. Gráficos Apilados
- d. Gráficos Combinados
- e. Gráficos de Dispersión
- f. Grafico Circular y de Anillo
- g. MitoSheet (Tablas y Gráficos Dinámicos)
- h. Gráfico de Pirámide



Probablemente ya conoces sobre Excel y Power BI pero en tus proyectos de Analítica de Datos o Inteligencia de Negocios aún gastas la mayor parte del tiempo en la fase de preparación de datos (usualmente más del 80% del tiempo del proyecto), tratando de obtener una tabla lo suficientemente limpia como para poder utilizarla en tus análisis.

Si es así, definitivamente deberías dominar todo el potencial que te ofrece Power Query como herramienta de ETL para Extraer datos de una gran variedad de orígenes de datos, Transformarlos de manera ágil y sencilla y Cargarlos (Load) a tu modelo de datos. Power Query cada vez viene incorporado en más herramientas de análisis como Excel, Power BI, Analysis Services y otras.

El gran beneficio para ti al aprender Power Query es que ahorrarás horas en preparación de datos y tendrás más tiempo disponible para hacer análisis que realmente agreguen valor.

¡Reserva ya tu espacio y dispara tu productividad al transformar datos!

TEMARIO A CUBRIR:

- Combinación de datos Avanzada
 - Combinación de data desestructurada
 - Combinación de Múltiples Hojas de Excel
 - Transformación de datos en tablas NxM
 - Dinamización de Columnas
 - Combinación de Columnas (Merge)
 - Lógica Difusa (Fuzzy Luck Up)
 - Agrupando y Sumarizando Data
 - Índices por categoría
 - Table.Profile
 - SIERROR Y SI en Power Query
- Manejo de fechas en diferente región
 - Columnas a partir de Ejemplos
 - Introducción a Lenguaje M (Sintaxis y Escritura)
 - Formulas del Lenguaje M
 - Modificaciones en M
 - Tipos de Datos Básicos y Complejos en M
 - Parámetros de Power Query
 - Escritura de Datos con Power Query
 - Funciones Personalizadas
 - Querys Avanzados y Seguros

Modalidad: Sincrónica / **Precio:** c71.400 / **Duración:** 12 horas / **Requisito:** Análisis y Visualización de Datos con Power BI

growup
DATA ANALYTICS

MANIPULACIÓN Y TRANSFORMACIÓN DE DATOS CON POWER QUERY



Este curso llamado Introducción a R para Ciencia de Datos te permitirá empezar desde cero a manejar el lenguaje principal de los Científicos de Datos: “R”. No importa si no tienes una gran experiencia en programación ya que el curso inicia desde el nivel básico y te llevará hasta la creación de tus primeros gráficos avanzados incluyendo exploración y limpieza de datos.

¡Si te gusta el análisis de datos, anímate a inscribirte al curso y a empezar en este apasionante mundo de la Ciencia de Datos!

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. Entorno R básico

- a. Introducción a R
- b. Instalación de R
- c. Utilización de Consola de R
- d. Instalación de R Studio
- e. Manejo de archivos, consola, vistas y plots
- i. Proyectos de R
- l. Uso de la Ayuda de R Studio

3. Trabajando con R usos prácticos

- a. Utilizando R de forma práctica
- b. Instalando Paquetes.
- c. Variables (Tipos de Datos, numéricos, alfanuméricos)
- d. Matrices y arreglos
- e. Funciones
- f. Control de Flujo
- g. Bucles
- h. Matrices
- i. DataFrames
- j. Listas

- k. Subsetting Datasets
- l. Tablas

4. Importación y Exportación de Datos

- a. Definiendo y Obteniendo directorios de trabajo
- b. Importando CSV File
- c. Importando Excel File
- d. Conexión a Bases de Datos
- d. Exportando Datos en diferentes formatos

5. Exploración de Datos

- a. Identificación de tipos de datos
- b. Identificación de Nulos o errores
- c. Identificación de valores atípicos

6. Limpieza de Datos

- a. Transformación de tipos de datos
- b. Reemplazo o eliminación

- de Nulos o errores
- c. Manejo de valores atípicos

7. R Distribución y Gráficas: usos prácticos

- a. Paquetes de distribución stats
- b. Encadenamiento lógico
- c. Trabajando con Paquetes Comunes (ggplot2, Janitor,etc)
- d. Manejo de Gráficas
- e. Comprendiendo Matrices de Confusión
- f. Operaciones Aritméticas Avanzadas
- g. Diagramas de Barra
- h. Histogramas
- i. Boxplots
- j. Plots
- k. Gráficos en 3D

10. Proyecto del curso

- R Distribución y Gráficas: usos prácticos
- a. Paquetes de distribución stats

- b. Encadenamiento lógico
- c. Trabajando con Paquetes Comunes (ggplot2, Janitor,etc)
- d. Manejo de Gráficas
- e. Comprendiendo Matrices de Confusión
- f. Operaciones Aritméticas Avanzadas
- g. Diagramas de Barra
- h. Histogramas
- i. Boxplots
- j. Plots
- k. Gráficos en 3D
- l. Gráficos de Dispersión
- m. Otros gráficos avanzados

INTRODUCCIÓN A R PARA
CIENCIA DE DATOS





ESTADÍSTICA ESENCIAL PARA ANÁLISIS DE DATOS CON PYTHON

La estadística es el corazón de cualquier análisis de datos, y en este curso aprenderás cómo aplicar tanto estadística descriptiva como inferencial en proyectos de ciencia de datos. Desde medidas de tendencia central hasta técnicas avanzadas para detectar valores atípicos (como Isolation Forest y DBSCAN), este módulo te prepara para explorar y comprender patrones en los datos. Además, cubriremos métodos de muestreo, intervalos de confianza, y el uso de pruebas de hipótesis (como ANOVA y T-Student), fundamentales para tomar decisiones basadas en datos. Con esta capacitación, serás capaz de analizar datos de manera crítica y estructurada.

TEMARIO A CUBRIR:

1. Estadística Descriptiva

- a. Medidas de Rango
- b. Medidas de Tendencia Central
- c. Medidas de Dispersión
- d. Covarianza
- e. Datos Cuantitativos y Cualitativos

2. Análisis Exploratorio de Datos

- a. Histograma
- b. Densidad
- c. Boxplot
- d. Matriz de Correlación

3. Valores Atípicos

- a. Técnicas detección Univariante
 - i. Método Rango Intercuartílico
 - ii. Método Desviación Estándar
 - iii. Método Puntuación Z
- b. Técnicas detección Multivariante
 - i. Isolation Forest
 - ii. DBSCAN
 - iii. Factor Atípico Local

4. Análisis de Normalidad

- a. Representaciones Gráficas
 - i. Histograma
 - ii. QQ-Plot
- b. Contraste de Hipótesis
 - i. Shapiro Wilk
 - ii. D'Angostino
 - iii. K-Squared
- c. Métodos Analíticos
 - i. Asimetría
 - ii. Curtosis

5. Estadística Inferencial

- a. Muestreos Probabilísticos
 - i. No Probabilísticos
 - ii. Estimación de Parámetros
 - iii. Estimación Puntual
 - iv. Estimación por Intervalo
- b. Remuestreo
 - i. Resampling
 - ii. Bootstrap
 - iii. Permutacion
- c. Pruebas de

- Hipótesis
 - i. Contraste de Hipótesis
 - 1. Conformidad
 - 2. Homogeneidad
 - 3. Independencia
 - ii. Prueba T-Student
 - iii. Prueba ANOVA
 - iv. Prueba Chi-Cuadrado
 - v. Prueba Kolmogorov-Smirnov
 - vi. Prueba Mann-Whitney U
 - vii. Prueba Wilcoxon Signed Rank
 - viii. Prueba Kruskal-Wallis H
 - ix. Prueba Correlación Spearman
- e. Intervalos de Confianza
 - i. IC para Media
 - ii. IC para Proporción
 - iii. IC diferencia de

- Medias
 - i. IC para Varianza
 - ii. Probabilidad Estadística
- Distribuciones Discretas
 - i. Distribuciones Continuas
 - ii. Distribución Multivariada
 - iii. Generación de Variables Aleatorias
- Simulación de Datos según Distribución
 - i. Modelo Monte Carlo
 - ii. Teorema Central del Limite
 - iii. Ajuste de Distribución Paramétrica y No Paramétrica

Modalidad: Sincrónica / **Precio:** €126.480 / **Duración:** 24 horas / **Requisito:** Introducción a Python

Enfocado en el aprendizaje supervisado, este curso te guía en la construcción de modelos de regresión para predecir variables continuas. Comenzaremos con un análisis exploratorio para identificar relaciones en los datos, luego aprenderás técnicas de preprocesamiento de datos como la estandarización y la imputación. Exploraremos algoritmos de regresión lineal (Ridge, Lasso, ElasticNet) y métodos no paramétricos como KNN y Random Forest. Además, trabajaremos en la optimización de hiperparámetros y la validación de modelos para lograr resultados precisos y confiables. Este curso es ideal para quienes buscan una comprensión profunda de la predicción en ciencia de datos.

TEMARIO A CUBRIR:

1. EDA para Machine Learning

- a. Distribución Variables Cualitativas
- b. b. Correlaciones de
- c. Variables Numéricas
- d. c. Transformación
- e. Logarítmica o Raíz
- f. Cuadrática
- g. d. Ajuste Distribución
- h. AIC-BIC – DIC
- i. e. Relación de Variables
- j. Independiente contra la
- k. Dependiente

2. Preprocesamiento de Datos

- a. División Train-Test
- b. b. Estandarización y
- c. Escalado de Variables

- a. c. Imputación de Valores

- b. Vacíos

- c. d. Binarización de Variables

- d. Cualitativas

3. Algoritmos de regresión lineal

- a. Ridge

- b. Lasso

- c. ElasticNet

- d. OLS

4. Algoritmos No Paramétricos

- a. KNN

- b. b. Árbol de Decisión

- c. c. Random Forest

- d. d. Gradient Boosting

- e. e. Super Learner

5. Validación

- a. Validación Cruzada

- b. Repetida

- c. Diagnóstico de

- d. Residuos

6. Métricas de Precisión

- a. MAE

- b. MSE

- c. RMSE

- d. Coeficiente

- e. Determinación

7. Búsqueda de HiperParámetros

- a. GridSearch

- b. RandomSearch

- c. Optimización

- d. Bayesiana

MACHINE LEARNING MODELOS DE REGRESIÓN CON PYTHON

Este curso profundiza en los modelos de machine learning que permiten clasificar datos en categorías. Cubriremos desde algoritmos de clasificación binaria y multiclase (como Regresión Logística y Random Forest) hasta técnicas de agrupamiento (k-Means y DBSCAN). El curso incluye el desarrollo de técnicas de ingeniería de características, balanceo de muestras, y selección de características. También aprenderás a evaluar el rendimiento del modelo usando métricas como precisión, F1-Score, y curvas ROC/AUC. Es una capacitación esencial para quienes desean desarrollar habilidades prácticas en clasificación y segmentación de datos.

TEMARIO A CUBRIR:

1. Feature Engineering

- a. Imputación de Valores Faltantes
- b. Normalización y Estandarización
- c. Pipelines
- d. Generación de Características Temporales
- e. Binning (Agrupamiento de Variables)
- f. Tratamiento de Outliers en Pipelines
- g. Descomposición de Componentes Principales
- h. Feature Selection
- i. Encoding
- j. Balance de Muestras
- k. (Oversampling y Undersampling)

2. Modelos Clasificación Binaria y Multiclase

- a. Regresión Logística
 - b. Máquinas de Soporte Vectorial
 - c. K Vecinos Mas Cercanos
 - d. Árbol de Decisión
 - e. Bosques Aleatorios
 - f. Naive Bayes
 - g. Categórico
 - ii. Bernoulli
 - iii. Gaussiano
- #### 3. Precision del Modelo
- a. Exactitud
 - b. Precision
 - c. Recall
 - d. F1-Score

- e. Matriz de Confusion
- f. Reporte de Clasificación
- g. AUC/ROC
- h. Cohen's Kappa
- Log Loss

4. Búsqueda de HiperParametros

- a. GridSearch
- b. RandomSearch

5. Serialización del modelo

- a. Guardar el Modelo
- b. Predicciones a datos nuevos

6. Modelos de Clusterización

- a. Cluster Particional y Basados

- a. en Densidad K-Means
- ii. DBSCAN
- iii. Affinity Propagation
- iv. Spectral Clustering
- b. Cluster Jerárquicos Jerárquico Aglomerativo
- ii. Jerárquico Divisivo
- iii. HDBSCAN
- c. Técnicas de Validación
- Metodo del Codo
- ii. Coeficiente de Silueta
- iii. Dendrograma

CURSO MACHINE LEARNING MODELOS DE CLASIFICACIÓN Y CLUSTERIZACIÓN

Aprende a procesar y analizar grandes volúmenes de datos con PySpark en Azure Databricks, la plataforma líder para Big Data en la nube. Este curso combina la configuración de Databricks en Azure con el procesamiento distribuido en PySpark, desde la ingesta y transformación de datos hasta la implementación de modelos de Machine Learning con MLlib y la automatización de pipelines bajo la arquitectura Medallón (Bronze, Silver y Gold). Ideal para científicos, analistas, ingenieros de datos y desarrolladores que buscan dominar Big Data en un entorno moderno, eficiente y colaborativo.

TEMARIO A CUBRIR:

Módulo 1. Fundamentos de Big Data y Databricks

- ¿Qué es Big Data y qué resuelve PySpark?
- Arquitectura de Apache Spark y entorno colaborativo de Databricks.
- Creación de cuenta en Azure, costos y configuración inicial.
- Notebooks en Databricks: primeros pasos en PySpark.

Módulo 2. Ingesta y Gestión de Datos en la Nube

- Conexión a Azure Data Lake Storage (ADLS).
- Montaje con claves, SAS y Secret Scope.
- Lectura de múltiples formatos (CSV, JSON, TXT, Parquet).

Módulo 3. Transformaciones y Consultas con PySpark

- DataFrames, SQLContext y Spark SQL.
- Operaciones con columnas: withColumn, select, drop.

- Filtrado, ordenamiento y conversiones de tipos.
- Funciones de ventana y consultas distribuidas.

Módulo 4. Agrupaciones y Estadística Descriptiva

- Uso de groupBy, agg, count y funciones de agregación.
- Cálculo de métricas descriptivas.
- Tablas dinámicas (pivot tables) y operaciones avanzadas.

Módulo 5. Optimización y Delta Lake

- Lectura/escritura Delta y Time Travel.
- Upserts, limpieza de datos y Z-Ordering.
- Estrategias de optimización para grandes volúmenes de datos.

Módulo 6. Preparación para BI

- Creación de tablas Gold para análisis empresarial.
- Casos de negocio: Employees, Customers, Products y Sales.
- Preparación de datasets para dashboards en Power BI/Tableau.

Módulo 7. Ingeniería de Características para Machine Learning

- Preparación de datos categóricos y numéricos.
- Codificación: StringIndexer, OneHotEncoder.
- Escalado y normalización (StandardScaler).
- Creación de pipelines de transformación en Spark.

Módulo 8. Machine Learning en Big Data con MLlib

- Introducción a MLlib y flujo de modelos en Spark.
- Modelos de regresión y clasificación.
- Evaluación de modelos: precisión, recall y AUC.

Módulo 9. Automatización y Orquestación de Pipelines

- Canalizaciones de datos (Bronze > Silver > Gold).
- Jobs y programación de ejecuciones en Databricks.
- Reducción de costos y buenas prácticas de operación.

Modalidad: Asincrónica / Precio: €50.440 / Duración: 24 horas / Requisito: Conocimientos básicos de Python y análisis de datos. SQL básico. Cuenta Azure (se ofrece crédito inicial de \$200)

BIG DATA CON PYTHON
EN AZURE DATABRICKS





Cualquier persona podría crear algunos gráficos a partir de datos, pero no todas las personas son capaces de contar verdaderas historias a través de las visualizaciones de datos. De hecho, es una habilidad escasa y sumamente apreciada.

No hay mucha diferencia entre contar una historia de una película y contar una historia correctamente utilizando gráficos y tablas. El secreto está en que el presentador logre transmitir al espectador el mensaje de manera clara y contundente en poco tiempo, para que de esta forma el observador logre tomar una decisión completamente informado.

El Data Storytelling une las habilidades de la Visualización de Datos + Narrativa y en este curso nos enfocaremos en ambas temáticas utilizando muchísimos ejemplos y ejercicios para que pongas en práctica de inmediato lo que has aprendido.

Conocerás desde el tipo de gráfica adecuada para cada aplicación y cómo hacer que estas transmitan un mensaje efectivamente, hasta técnicas de narrativa y diseño para hacer que tu historia corta lleve un orden lógico en el que tu espectador mantenga la atención y finalmente tome la decisión informada que buscas.

Eres bienvenido a participar de este curso sin importar la herramienta de visualización de datos que utilices (Excel, Power BI, Tableau, R, Python u otra), ya que los conocimientos que obtendrás te servirán para aplicarlos desde el día uno en cualquier programa que selecciones.

TEMARIO A CUBRIR:

1. Introducción a Data Storytelling

2. La importancia del contexto

- Análisis Exploratorios vs Explicativos
- Quién, Qué y Cómo
- Obtener el contexto adecuado: preguntas a realizar
- Historias de 3 minutos e Idea Principal
- Storyboarding

3. Seleccionar el objeto visual adecuado

- Textos Simples, Tablas, Gráficos, Puntos, Líneas, Barras, Áreas
- Otros tipos de visualizaciones
- Qué visualizaciones evitar

4. Técnicas para evitar el desorden

- Carga Cognitiva
- Desorden
- Principios de Percepción Visual
- Uso No Estratégico del Contraste
- Reordenamiento Paso a Paso
- Enfocar la atención de la audiencia
- Preparación para ver con el cerebro
- Obtener enfoque de la memoria
- Atributos de preatención visual
- Tamaño, color, posición en la página

5. Enfocar la atención de la audiencia

- Preparación para ver con el cerebro
- Obtener enfoque de la memoria
- Atributos de preatención visual
- Tamaño, color, posición en la página

6. Principios de diseño

- Posibilidades
- Accesibilidad
- Estética
- Aceptación

7. Buenas Prácticas en Storytelling

- La magia de la historia
- Construyendo la historia
- La estructura narrativa
- El poder de la repetición
- Tácticas para asegurar que nuestra historia es clara

8. Proyecto Final de Data Storytelling

INFORMACIÓN GENERAL

ORDEN DE LOS CURSOS Y PRECIOS:

- Analítica de Datos para la Toma de Decisiones
Precio: c29.640 / Duración: 20 horas
- Excel Intermedio: Funciones para Análisis de Datos
Precio: c61.200 / Duración: 15 horas
- Excel Avanzado: Tablas dinámicas y Excel BI
Precio: c71.400 / Duración: 18 horas
- Análisis y Visualización de Datos con Power BI
Precio: c97.920 / Duración: 18 horas
- SQL para Análisis de Datos
Precio: c97.920 / Duración: 18 horas
- Inteligencia de Negocios Aplicada con Power BI
Precio: c117.300 / Duración: 24 horas
- Introducción a Python para Ciencia de Datos
Precio: c126.480 / Duración: 24 horas
- Manipulación y Transformación de Datos con Power Query
Precio: c71.400 / Duración: 12 horas
- Introducción a R para Ciencia de Datos
Precio: c50.440 / Duración: 24 horas
- Estadística Esencial para Análisis de Datos con Python
Precio: c126.480 / Duración: 24 horas
- Machine Learning: Modelos de Regresión con Python
Precio: c110.160 / Duración: 20 horas
- Machine Learning: Modelos de Clasificación y Clusterización con Python
Precio: c110.160 / Duración: 20 horas
- Big Data con Python en Azure Databricks
Precio: c50.440 / Duración: 24 horas
- Data Storytelling de Impacto
Precio: c81.600 / Duración: 15 horas

PRECIO TOTAL: c1.202.540 por persona IVA incluido

PRECIO EQUIVALENTE EN DÓLARES:

DURACIÓN TOTAL: 276 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 Técnico en Analítica de Datos.

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 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): Microsoft Data Analyst.

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

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.

<https://escuela.growupcr.com/p/tecnico-analitica-de-datos>

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

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

Más información al:  [8414-4646](https://api.whatsapp.com/send?phone=5299984144646)