



PROGRAMA DE FORMACIÓN · EDICIÓN 7

Data Science Applied

(DS) · Ciencia de Datos Aplicada al Negocio

- ✓ **Proyectos reales** con casos de empresas líderes del Perú
- ✓ **100 % virtual en vivo** · grabaciones permanentes
- ✓ **Certificación ISO 9001:2015** · mentores activos de la industria

80h

HORAS EN VIVO

4

MÓDULOS

40

SESIONES EN VIVO

26 ene 2027

INICIO

📅 Inicio: **martes 26 de enero de 2027** · martes y jueves · 8:00 – 10:00 p. m. (hora de Perú)

+1,000 profesionales formados | 4.9/5 de valoración | ISO 9001 certificación

80h

HORAS EN VIVO

4

MÓDULOS

40

SESIONES EN VIVO

26 ene 2027

INICIO

01

01

Información general

PROGRAMA	Data Science Applied
CÓDIGO	DS
EDICIÓN	7
INICIO	martes 26 de enero de 2027 (bienvenida: lunes 25 de enero de 2027)
DURACIÓN	80 horas · 40 sesiones en vivo · 5 meses (hasta el 17 de junio de 2027)
HORARIO	martes y jueves · 8:00 – 10:00 p. m. (hora de Perú)
MODALIDAD	100 % virtual · clases en vivo por Google Meet, con grabación
CERTIFICACIÓN	Certificado FUTURA con respaldo ISO 9001:2015
INSTITUCIÓN	FUTURA — Analítica Avanzada

02

02

Presentación del programa

La ciencia de datos es hoy una de las habilidades más demandadas por las empresas. Este programa te forma como un científico de datos completo: con dominio técnico, criterio analítico y capacidad para generar impacto real en una organización.

A lo largo de 4 módulos encadenados y 80 horas en vivo construirás desde cero hasta técnicas avanzadas: Python para analítica, estadística aplicada, análisis multivariante con feature engineering y machine learning con datos reales de negocio.

Con mentores en actividad en **Interbank, Credicorp, IGT y BCP**, aprenderás cómo se trabaja la ciencia de datos en la práctica, no solo en teoría.

Perfil del participante

- > Profesionales en transición hacia roles de ciencia de datos.
- > Analistas que trabajan con Excel o SQL y quieren dar el salto a modelar con Python.
- > Estudiantes y egresados de ingeniería, estadística, economía o carreras afines.
- > Cualquier profesional que quiera aplicar la ciencia de datos a decisiones de negocio.

Metodología

- 1 Clases 100 % virtuales en vivo, con la grabación de cada sesión para repasar cuando quieras.
- 2 Acompañamiento de mentores que trabajan hoy en equipos de datos de la industria.
- 3 Aprendizaje basado en casos reales de banca, seguros, salud y retail.
- 4 Proyectos integradores por módulo que simulan problemas de negocio reales.
- 5 Certificado FUTURA con respaldo ISO 9001:2015 por módulo y por programa.
- 6 Comunidad activa de alumnos y egresados para networking y crecimiento profesional.

Equipo de mentores

Profesionales en actividad en equipos de datos de la industria.

MENTOR	CARGO	EMPRESA	MÓDULOS
Carlos Rondan	Senior Data Scientist	Interbank	M1 · M3
Luis Garayar	Chapter Leader Data Scientist	Credicorp	M1
Manuel Merino	Market Research Analyst	IGT	M2
Miguel Echeverre	Sub Gerente Data Scientist	BCP	M3
Luis Torres	Data Scientist Advanced	BCP	M4
Jonattan Ramos	Data Scientist Advanced	BCP	M4

Estructura curricular

Una ruta encadenada: cada módulo usa lo que construyó el anterior.

MÓD.	NOMBRE	SESIONES	HORAS	FECHAS
1	Python para Data Science	8	16 h	26/01/2027 – 16/02/2027
2	Estadística aplicada	9	18 h	23/02/2027 – 23/03/2027
3	Multivariante & Feature Engineering	11	22 h	30/03/2027 – 04/05/2027
4	Machine Learning	12	24 h	11/05/2027 – 17/06/2027
Programa completo		40	80 h	26/01/2027 – 17/06/2027

Fechas disponibles

Cada módulo se puede llevar por separado y en la fecha que más te convenga.

MÓDULO	EDICIÓN	INICIO	FIN
M1 · Python para Data Science	7 (este sílabo)	mar 26/01/2027	16/02/2027
M2 · Estadística aplicada	6	jue 12/11/2026	15/12/2026
	7 (este sílabo)	mar 23/02/2027	23/03/2027
M3 · Multivariante & Feature Engineering	6	mar 05/01/2027	09/02/2027
	7 (este sílabo)	mar 30/03/2027	04/05/2027
M4 · Machine Learning	6	jue 18/02/2027	01/04/2027
	7 (este sílabo)	mar 11/05/2027	17/06/2027

Detalle por módulo

M1 Python para Data Science

16h · 8 sesiones

LO QUE APRENDES

- **Fundamentos de Python y programación.** Variables, estructuras, funciones y el hábito de escribir código que otro pueda leer.
- **Manipulación de datos con NumPy y Pandas.** Cargar, limpiar, agrupar y cruzar tablas reales: lo que ocupa la mayor parte del día de un analista.
- **Visualización y entorno profesional.** Gráficos que explican, notebooks ordenados y control de versiones con Git y GitHub.

HERRAMIENTAS

Python

Google Colab

Jupyter

Anaconda

NumPy

Pandas

Matplotlib

Seaborn

Plotly

Git

GitHub

CRONOGRAMA DE SESIONES · 26/01/2027 – 16/02/2027

#	FECHA	SESIÓN	MENTOR
1	Lun 25/01/2027	Bienvenida al Programa Data Science Presentación del programa y de su ruta de 4 módulos encadenados (Python para Data Science, Estadística aplicada, Multivariante & Feature Engineering y Machine Learning), y de las herramientas que se usarán: Python en Colab/Jupyter y Git/GitHub.	Carlos R.
2	Mar 26/01/2027	Introducción a Python, entornos y tipos de datos Primeros pasos en Python para analítica: los entornos de trabajo (Colab y Jupyter), los tipos de datos básicos y las estructuras de datos con las que se trabaja en el resto del programa.	Carlos R.
3	Jue 28/01/2027	Programación Estructurada I Primera parte de programación estructurada: controlar el flujo de un programa con condicionales y bucles.	Carlos R.
4	Mar 02/02/2027	Programación Estructurada II Segunda parte de programación estructurada: se continúa con condicionales y bucles.	Carlos R.
5	Jue 04/02/2027	Funciones Cómo escribir funciones propias en Python para organizar y reutilizar código.	Luis G.
6	Mar 09/02/2027	Numpy Cómputo numérico vectorizado con NumPy: trabajar con arreglos y operaciones numéricas de forma eficiente.	Luis G.
7	Jue 11/02/2027	Pandas Manipulación y análisis de datos con Pandas: cargar, explorar, transformar y resumir tablas de datos.	Luis G.
8	Mar 16/02/2027	Visualización de datos Visualización de datos con Matplotlib, Seaborn y Plotly para explorar los datos y comunicar hallazgos.	Luis G.

✓ **Al terminar:** Un repositorio propio en GitHub con tus notebooks: la primera pieza de tu portafolio.

LO QUE APRENDES

- **Estadística descriptiva e inferencial.** Resumir una población desde una muestra y saber cuánto puedes confiar en ese resumen.
- **Probabilidad, distribuciones y muestreo.** Las distribuciones que aparecen en el negocio y cómo diseñar una muestra que no engañe.
- **Pruebas de hipótesis y experimentos.** Tests A/B, intervalos de confianza y cuasi-experimentos para medir si una acción funcionó.

HERRAMIENTAS

Python

SciPy

Statsmodels

Pingouin

Lifelines

Matplotlib

Seaborn

CRONOGRAMA DE SESIONES · 23/02/2027 – 23/03/2027

#	FECHA	SESIÓN	MENTOR
1	Mar 23/02/2027	Estadística Descriptiva: Centro, Dispersión, Forma y Correlación Cómo resumir y describir datos: medidas de centro (tendencia central), dispersión y forma de una distribución, y correlación entre variables.	Manuel M.
2	Jue 25/02/2027	Fundamentos de Probabilidad y Teorema de Bayes Fundamentos de probabilidad, probabilidad condicional y el Teorema de Bayes.	Manuel M.
3	Mar 02/03/2027	Variable Aleatoria y Funciones de Distribución Qué es una variable aleatoria y cómo se describe con sus funciones de distribución.	Manuel M.
4	Jue 04/03/2027	Distribuciones de Probabilidad y Teorema del Límite Central Las principales distribuciones de probabilidad y el Teorema del Límite Central, base de la inferencia estadística.	Manuel M.
5	Mar 09/03/2027	Inferencia Estadística I: Estimación, Contraste y Prueba t/Z Inferencia estadística, parte 1: estimación, contraste de hipótesis y pruebas t y Z.	Manuel M.
6	Jue 11/03/2027	Inferencia Estadística II: ANOVA, Chi-Cuadrado Inferencia estadística, parte 2: ANOVA para comparar grupos y prueba Chi-cuadrado.	Manuel M.
7	Mar 16/03/2027	Muestreo Correcto: Diseño, Casos Aplicados y Poder Estadístico Cómo diseñar una muestra correctamente, casos aplicados y el poder estadístico de un estudio.	Manuel M.
8	Jue 18/03/2027	Pruebas No Paramétricas: Equivalentes sin Supuesto de Normalidad Pruebas no paramétricas: alternativas para cuando los datos no cumplen el supuesto de normalidad.	Manuel M.
9	Mar 23/03/2027	Diseños Cuasi-Experimentales: Diferencias en Diferencias (DiD) + Bootstrap Diseños cuasi-experimentales para medir el impacto de una intervención con Diferencias en Diferencias (DiD), y Bootstrap.	Manuel M.

- ✓ **Al terminar:** Saber leer un resultado con su incertidumbre y defender una conclusión con datos frente a un gerente.

LO QUE APRENDES

- **Limpieza y tratamiento de datos.** Nulos, atípicos, codificación de categorías y particiones train/test sin fugas de información.
- **Análisis multivariante y reducción.** PCA y MCA para entender la estructura de cientos de variables y quedarte con las que importan.
- **Feature engineering avanzado.** Transformaciones, WOE, selección de variables y un pipeline reproducible listo para el modelo.

HERRAMIENTAS

Python

Pandas

Scikit-Learn

Prince

Imbalanced-learn

OptBinning

Statsmodels

CRONOGRAMA DE SESIONES · 30/03/2027 – 04/05/2027

#	FECHA	SESIÓN	MENTOR
1	Mar 30/03/2027	Análisis Multivariante: Observando Datos Multivariados Introducción al análisis multivariante: cómo observar y visualizar datos con muchas variables a la vez.	Carlos R.
2	Jue 01/04/2027	Reducción de Dimensionalidad: PCA (Teoría) Teoría de la reducción de dimensionalidad con Análisis de Componentes Principales (PCA).	Carlos R.
3	Mar 06/04/2027	PCA Práctica 1 Primera práctica de PCA: aplicar la reducción de dimensionalidad con Python.	Carlos R.
4	Jue 08/04/2027	PCA Práctica 2: Caso de Negocio Banca Retail Segunda práctica de PCA, aplicada a un caso de negocio de banca retail.	Carlos R.
5	Mar 13/04/2027	Reducción de Dimensionalidad: MCA (Teoría) Teoría del Análisis de Correspondencias Múltiples (MCA) para reducir la dimensionalidad de variables categóricas.	Carlos R.
6	Jue 15/04/2027	MCA Práctica: Caso de Negocio Banca Retail Práctica de MCA aplicada a un caso de negocio de banca retail.	Carlos R.
7	Mar 20/04/2027	Fundamentos de Preprocesamiento y Feature Engineering Fundamentos del preprocesamiento de datos (imputación y escalado) y del feature engineering: preparar las variables antes de modelar.	Miguel E.
8	Jue 22/04/2027	Pipeline: División Train / Test (muestreo estratificado + bietápico) + EDA Inicio del pipeline de preprocesamiento: división de los datos en entrenamiento y prueba con muestreo estratificado y bietápico, y análisis exploratorio (EDA).	Miguel E.
9	Mar 27/04/2027	Pipeline: Filtros duros + outliers multivariante + WOE + Feature Engineering Continuación del pipeline: filtros duros, tratamiento de outliers multivariantes, transformación WOE y creación de variables (feature engineering).	Miguel E.
10	Jue 29/04/2027	Pipeline: Selección de Variables + PCA/MCA + Balanceo Datos Cierre del pipeline: selección de variables, uso de PCA/MCA y balanceo de datos (SMOTE).	Miguel E.
11	Mar 04/05/2027	Pipeline para Regresión: Segmentación, Test de Chow, Árboles de Inferencia y PLS Pipeline para modelos de regresión: segmentación, Test de Chow, árboles de inferencia y PLS.	Miguel E.

✓ **Al terminar:** Un pipeline de preparación de datos de un caso real de banca, reproducible y listo para tu portafolio.

LO QUE APRENDES

- > **Modelos supervisados: regresión y clasificación.** Regresión lineal y logística, árboles y sus límites: cuándo usar cada uno y cómo explicarlo.
- > **Clustering, segmentación RFM y recomendación.** Agrupar clientes sin etiquetas, segmentar por valor y recomendar el siguiente producto.
- > **Ensamblados y evaluación.** Random Forest y XGBoost, validación out-of-time, ROC-AUC y Gini: las métricas que pide el negocio.

HERRAMIENTAS

Python

Scikit-Learn

XGBoost

LightGBM

CatBoost

Imbalanced-learn

SHAP

Optuna

CRONOGRAMA DE SESIONES · 11/05/2027 – 17/06/2027

#	FECHA	SESIÓN	MENTOR
1	Mar 11/05/2027	Regresión Lineal: Simple, Múltiple y Diagnóstico Regresión lineal simple y múltiple, y diagnóstico de los modelos.	Luis T.
2	Jue 13/05/2027	Análisis Discriminante y Regresión Logística (Teoría) Teoría del Análisis Discriminante y de la Regresión Logística para problemas de clasificación.	Luis T.
3	Mar 18/05/2027	Evaluación, Validación Cruzada y Optimización de Hiperparámetros Cómo evaluar modelos, validación cruzada y optimización de hiperparámetros.	Luis T.
4	Jue 20/05/2027	Regresión Regularizada: Teoría y Práctica Regresión regularizada (Ridge, Lasso y ElasticNet): teoría y práctica.	Luis T.
5	Mar 25/05/2027	Árboles, Random Forest y K-NN Árboles de decisión, Random Forest y K-NN (vecinos más cercanos).	Luis T.
6	Jue 27/05/2027	Práctica Integradora de Clasificación Práctica integradora: aplicar los modelos de clasificación vistos en el módulo a un caso completo.	Luis T.
7	Mar 01/06/2027	Clustering: K-means, PAM, Jerárquico y DBSCAN Clustering o agrupamiento: K-means, PAM, clustering jerárquico y DBSCAN.	Jonattan R.
8	Jue 03/06/2027	Clustering Aplicado: Caso de Negocio y Comparación de Técnicas Clustering aplicado a un caso de negocio, comparando las técnicas vistas.	Jonattan R.
9	Mar 08/06/2027	Segmentación RFM y PAM Segmentación de clientes con RFM (recencia, frecuencia y monto) y PAM.	Jonattan R.
10	Jue 10/06/2027	Market Basket Analysis y Sistemas de Recomendación Market Basket Analysis (qué productos se compran juntos) y sistemas de recomendación.	Jonattan R.
11	Mar 15/06/2027	Máquina de Soporte Vectorial (SVM) Máquinas de Soporte Vectorial (SVM).	Jonattan R.
12	Jue 17/06/2027	Modelos Ensamblados: Bagging, Random Forest y Boosting Completo Modelos ensamblados: Bagging, Random Forest y Boosting (XGBoost, LightGBM y CatBoost).	Jonattan R.
✓ Al terminar: Un modelo entrenado, validado y explicado de punta a punta: la pieza que más pesa en una entrevista técnica.			

Herramientas y tecnologías

Las mismas que usan hoy los equipos de datos.

L1 Lenguaje, entorno y versionado

Python NumPy Pandas Jupyter Anaconda Google Colab Git GitHub

L2 Análisis estadístico y visualización

SciPy Statsmodels Pingouin Lifelines Prince Matplotlib Seaborn Plotly

L3 Machine learning

Scikit-Learn XGBoost LightGBM CatBoost Imbalanced-learn OptBinning SHAP Optuna

Casos de estudio

Proyectos aplicados sobre problemas reales de la industria.

BANCA

Credit scoring y probabilidad de default

BANCA

Marketing bancario y propensión de compra

SEGUROS

Detección de fraude y pricing

SALUD

Predicción de readmisión y no-show

RETAIL

Churn, segmentación RFM y demanda

TELECOM

Fuga de clientes

COBRANZAS

Contactabilidad y cobranza preventiva

RECURSOS HUMANOS

People analytics y rotación

Sistema de evaluación

COMPONENTE	PESO
Participación en clases en vivo	10 %
Ejercicios prácticos por módulo	30 %
Proyecto integrador / caso de estudio	40 %
Presentación final	20 %

AL TERMINAR EL PROGRAMA

Lo que obtendrás

> Portafolio real

Proyectos con datos de casos de negocio, listos para mostrar en una entrevista.

> Certificado ISO

Certificación FUTURA con respaldo ISO 9001:2015, verificable en línea.

> Red profesional

Comunidad de alumnos, egresados y mentores de la industria.

> Grabaciones

Acceso permanente a las clases grabadas y a los materiales del programa.

INSCRIPCIONES ABIERTAS

¿Listo para empezar?

Inicio el **26 de enero de 2027** · cupos limitados. Te ayudamos a elegir la ruta y las fechas que te convienen.

 **Escríbenos por WhatsApp**

 +51 908 814 045

 informes@futuradata.pe

 futuradata.pe



ANALÍTICA AVANZADA

Transforma datos en decisiones de alto impacto

✉ informes@futuradata.pe

📞 +51 908 814 045

🌐 futuradata.pe

📍 Guayaquil 184, Urb. Santa Patricia, La Molina — Lima

🏅 Certificación ISO 9001:2015