CTIC UNI

UNIVERSIDAD NACIONAL DE INGENIERÍA

OFICINA DE TECNOLOGÍAS DE LA INFORMACIÓN

Programa de Iniciación Tecnológica PIT - 2025

CERTIFICADO

Otorgado a

ANA ROSA ALVIS HUILLCA

Por haber aprobado el curso de

CIENCIA DE DATOS 1: Exploratory Data Analysis

Realizado del 02 de mayo al 19 de mayo del presente, con una duración de 16 horas.

Lima, junio de 2025

Mag. Ing. Rubén Arturo Borja Rosales
JEFE OTI - UNI

N° Certificado: 017 - 0051421

FIRMA DIGITAL

Firmado digitalmente por: BORJA ROSALES Ruben Arturo FAU 20169004359 hard
Motivo: En señal de conformidad
Fecha: 18/06/2025 21:01:27-0500

UNIVERSIDAD NACIONAL DE INGENIERÍA

OFICINA DE TECNOLOGÍAS DE LA INFORMACIÓN

CTIC UNI

Temario del Curso:

CIENCIA DE DATOS - 1: Exploratory Data Analysis

- ¿Qué es el Análisis Exploratorio de Datos y por qué es importante?.
- Rol del EDA en el ciclo de vida de los datos.
- Tipos de variables: numéricas, categóricas, ordinales.
- Herramientas comunes en EDA (Python, pandas, matplotlib, seaborn, etc.).
- Estadísticas descriptivas y visualización.
- Medidas de tendencia central y dispersión.
- Visualización de variables numéricas: histogramas, boxplots, density plots.
- Visualización de variables categóricas: barplots, countplots.
- Identificación de datos desbalanceados en variables categóricas.
- Detección visual de asimetrías y valores atípicos.
- Correlación: Pearson, Spearman y Kendall.
- Matrices de correlación y mapas de calor.
- Relaciones entre variables numéricas y categóricas.
- Tablas de contingencia y relación entre variables categóricas.
- Identificación y tratamiento de valores faltantes.
- Métodos básicos de imputación.
- Identificación y tratamiento de outliers.
- Estandarización, normalización y codificación de variables.
- Transformaciones estadísticas: log, raíz cuadrada, Box-Cox.
- Creación de nuevas variables a partir de las existentes.
- Transformaciones útiles: ratios, binning, agrupación por categorías.
- Selección de variables basada en EDA: correlación, varianza, redundancia.
- Introducción a técnicas simples para estimar importancia de variables.
- Aplicación práctica de EDA.
- Desarrollo completo de un caso práctico de EDA con dataset real.
- Documentación del proceso exploratorio.
- Identificación de limitaciones de los datos antes del modelado.
- Presentación de resultados: hallazgos, hipótesis, visualizaciones clave.

Nota final: **20 (Veinte)**

N° Certificado: 017 - 0051421