

MÓDULO: *Análisis de Datos con Python*



Data Analysis
with Python



PRÁCTICA DE ANÁLISIS DE DATOS CON EXCEL

Contexto: En esta práctica trabajaremos con un conjunto de datos que recoge información sobre hábitos digitales, bienestar emocional y características personales de distintos individuos. El objetivo es aplicar técnicas básicas de análisis estadístico con Excel para interpretar relaciones entre variables, calcular medidas de centralización y dispersión, y evaluar la idoneidad de ajustes lineales entre pares de variables.

1. Traducción de variables

A continuación, se presenta la traducción al español de las variables del conjunto de datos:

<i>Variable original</i>	Traducción al español
<i>id</i>	id (identificador)
<i>age</i>	edad
<i>gender</i>	género
<i>region</i>	región
<i>income_level</i>	nivel de ingresos
<i>education_level</i>	nivel educativo
<i>daily_role</i>	rol diario
<i>device_hours_per_day</i>	horas de uso del dispositivo por día
<i>phone_unlocks</i>	desbloques del teléfono
<i>notifications_per_day</i>	notificaciones por día
<i>social_media_mins</i>	minutos en redes sociales
<i>study_mins</i>	minutos de estudio
<i>physical_activity_days</i>	días de actividad física
<i>sleep_hours</i>	horas de sueño
<i>sleep_quality</i>	calidad del sueño
<i>anxiety_score</i>	puntuación de ansiedad
<i>depression_score</i>	puntuación de depresión
<i>stress_level</i>	nivel de estrés
<i>happiness_score</i>	puntuación de felicidad
<i>focus_score</i>	puntuación de concentración
<i>high_risk_flag</i>	indicador de alto riesgo
<i>device_type</i>	tipo de dispositivo
<i>productivity_score</i>	puntuación de productividad
<i>digital_dependence_score</i>	puntuación de dependencia digital

2. Medidas de centralización y dispersión

Para cada variable numérica, calcula en Excel:

- **Media**
- **Mediana**
- **Moda**
- **Rango**
- **Desviación estándar**
- **Varianza**
- **Coeficiente de variación**

Aplica estos cálculos a variables como: edad, horas de dispositivo, minutos en redes sociales, puntuaciones emocionales, productividad, dependencia digital, etc.

3. Ajustes lineales entre pares de variables

Selecciona al menos cinco pares de variables que puedan tener relación entre sí. Representa los datos en gráficos de dispersión y aplica una recta de regresión lineal. Evalúa si el ajuste es adecuado observando la forma del gráfico y el coeficiente de correlación.

Ejemplos de pares sugeridos:

Variable X	Variable Y	Posible relación
horas de dispositivo	dependencia digital	Positiva
minutos en redes sociales	desbloques del teléfono	Positiva
horas de sueño	calidad del sueño	Positiva
ansiedad	depresión	Positiva
concentración	productividad	Positiva

4. Análisis por filtros

Realiza comparaciones estadísticas segmentando los datos por:

- **Género:** mujer, hombre, otro
- **Edad:** agrupa por rangos (15–24, 25–34, etc.)
- **Región o país**

Para cada grupo, compara las medias y desviaciones estándar de variables clave como:

- Horas de dispositivo
- Minutos en redes sociales
- Puntuaciones emocionales
- Productividad
- Dependencia digital

Representa los resultados con gráficos de barras, histogramas o diagramas de caja.