



Hilo

Iniciar sesión



analitica_ok 18 horas

...

Semana 2 - Día 2: Cómo limpiar datos como un Analista de Datos 🇵🇪🇮🇹

Bienvenidos al Día 2 de la Semana 2 del roadmap para convertirte en Data Analyst en 2025.

🔴 Ayer hablamos sobre la recolección de datos, pero los datos crudos suelen estar llenos de errores.

- ✗ Datos faltantes
- ✗ Duplicados
- ✗ Errores de formato
- ✗ Outliers

Si no los limpias, tu análisis será poco confiable. Hoy te enseño cómo limpiar datos correctamente. 📊💡

👍 13 💬 7 ↻ ⚙️ 2



analitica_ok 18 horas

...

💡 ¿Por qué es clave la limpieza de datos?

Imaginá que tenés una base de datos con clientes y encontrás esto:

- 🔴 Nombre: juan perez / JUAN PEREZ
- 🔴 Fecha de nacimiento: 01/02/2024 / 2024-02-01
- 🔴 País: Argentina / ARG / argentina

👉 Si no normalizás los datos, el sistema los interpretará como valores distintos.

Ahora veamos cómo limpiar datos de forma efectiva.

👍 1 💬 1 ↻ ⚙️



analitica_ok 18 horas

...

1 Eliminación de Valores Faltantes

Cuando hay datos vacíos:

- ✓ Eliminar la fila (si es un porcentaje pequeño del dataset).
- ✓ Rellenar con un valor lógico (promedio, mediana, "Desconocido", etc.).

💡 En una base de datos SQL:

Si hay valores nulos en la columna "edad", podemos rellenarlos con la media:

```
UPDATE clientes
SET edad = (SELECT AVG(edad) FROM clientes)
WHERE edad IS NULL;
```

💡 Ejemplo en Excel:

Rellenar valores vacíos con el promedio:

👉 Fórmula: =SI(A2="", PROMEDIO(A:A), A2)

👍 1 💬 1 ↻ ⚙️



analitica_ok 18 horas

...

2 Eliminación de Duplicados

🔴 Los datos duplicados inflan los resultados y distorsionan el análisis.

✂️ Cómo eliminarlos en distintos entornos:

✅ SQL:

```
DELETE FROM clientes
WHERE id NOT IN (
```



Iniciar sesión

```
SELECT MIN(id) FROM clientes GROUP BY nombre, email
);
```

✅ Excel: Datos -> Quitar Duplicados.

✅ Python:

```
df.drop_duplicates(inplace=True)
```



analitica_ok 18 horas

3 Normalización y Transformación de Datos

🔥 Asegurate de que todos los datos tengan el mismo formato.

✓ Texto: Convertir todo a minúsculas o mayúsculas.

✓ Fechas: Unificar formatos.

✓ Categorías: Si un mismo valor está escrito de distintas formas, hay que normalizarlo.

💡 Ejemplo en SQL: Convertir todos los nombres de clientes a mayúsculas.

```
UPDATE clientes
SET nombre = UPPER(nombre);
```

💡 Ejemplo en Python:

```
df['nombre'] = df['nombre'].str.upper()
```



analitica_ok 18 horas

4 Detección y Manejo de Outliers

🔥 Valores extremos pueden alterar los resultados.

🔍 Cómo detectarlos:

✅ SQL:

```
SELECT * FROM ventas
WHERE monto > (SELECT AVG(monto) + 3 * STDDEV(monto) FROM ventas);
```

✅ Excel:

Usá los percentiles:

👉 =CUARTIL(A:A, 1) (mínimo)

👉 =CUARTIL(A:A, 3) (máximo)



analitica_ok 18 horas

🔥 Una vez detectados, podés:

✓ Eliminarlos (si son errores).

✓ Ajustarlos (si tienen lógica en el negocio).



analitica_ok 18 horas

💡 Limpiar datos es una habilidad clave para cualquier analista.

✅ Eliminá valores faltantes o completalos inteligentemente.

✅ Quitá duplicados.

✅ Normalizá formatos de texto, fechas y números.

✅ Detectá y tratá outliers.

🔥 Mañana seguimos con Normalización de Datos. 🔥

💡 Encuentra más contenido sobre análisis de datos en mi cuenta.

#DataCleaning #SQL #Excel #Python #DataAnalytics

Hilos relacionados



danielsan_avila 1 día

Excelente guía para comprender las bases de datos NoSQL



407 3 33 45



excel_vale 1 hora

"¿Te interesa?", preguntó la de reclutamiento, con esperanza en la voz y un poco de desesperación. "Necesitamos a alguien ya... hay un desastre con los archivos".

Error: Crimen no encontrado microcuento

1 1



alfredovela 22 horas

Os dejo con una infografía sobre los "Elementos para un prompt en Deepseek". ¿Qué te parece Deepseek?

#infografia #inteligenciaartificial #ai #ia #deepseek



20 1 5



imemprende 1 día

7 Herramientas de IA para crear video

1. Adobe AI Video Generator
2. Image to video Ray 2
3. Goku AI
4. Lovable
5. Krea AI
6. Magic 1 for 1
7. Gemini

35 2

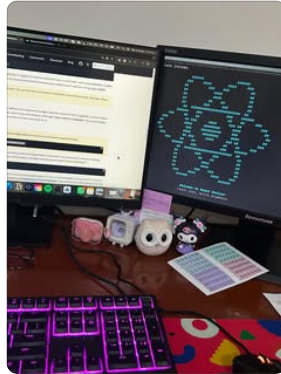
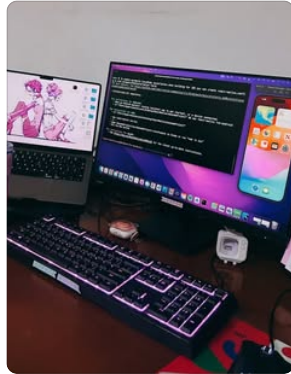


natcrossws_ · 5 horas

...

Les comparto unas fotos de cuando estaba estudiando React Native quisiera retomarlo de nuevo 🤔

Algún consejo o material? 🛠️



Iniciar sesión

© 2025 Condiciones de Threads Política de privacidad Política de cookies Reportar un problema

