

Curso : ANÁLISIS Y EXPLOTACIÓN DE BASES DE DATOS CON POWER BI (NIVEL BÁSICO)

Tema : Modelado de dataset

Caso: Gestión de Residuos Orgánicos en Restaurantes – Miraflores

Se dispone de una base de datos que contiene información sobre la generación de residuos orgánicos en restaurantes del distrito de Miraflores durante el año 2025. Esta información incluye variables relacionadas con fechas, tipo de restaurante, cantidad de residuos generados, tipo de residuo, turnos de operación, método de disposición y otras características relevantes.

Sin embargo, la base se encuentra en formato plano (una sola tabla), por lo que no es adecuada para realizar análisis eficientes ni para la construcción de dashboards.

Usa el archivo: Residuos Organicos Miraflores – 2025.xlsx, las variables son:

Variable	Descripción
ID	Identificador único de cada registro. Permite diferenciar cada observación y evitar duplicados.
Fecha	Fecha en la que se registró la generación de residuos. Base para el análisis temporal.
Año	Año correspondiente al registro. Facilita el análisis por periodos anuales.
Mes	Mes en el que se generaron los residuos. Permite identificar patrones mensuales.
Día de la semana	Indica el día en que se registró el dato. Útil para análisis por comportamiento semanal.
Restaurante	Nombre del establecimiento que genera los residuos. Permite comparaciones entre negocios.
Tipo de restaurante	Clasificación del restaurante (gourmet, comida rápida, criollo, etc.).
Distrito	Ubicación geográfica del restaurante. En este caso, Miraflores.
Zona	Subdivisión dentro del distrito (ejemplo: costa, centro). Permite análisis espacial.
Tipo de residuo	Tipo de residuo orgánico generado (pescado, vegetales,

Variable	Descripción
	pan, mixtos).
Cantidad (kg)	Cantidad de residuos generados en kilogramos. Variable principal de análisis.
Turno	Momento del día en que se generan los residuos (mañana, tarde, noche).
Método disposición	de Forma en que se gestionan los residuos (recolección municipal, compostaje).
Programa municipal	Indica si el restaurante participa en programas ambientales (Sí/No).
Capacidad restaurante	del Número de personas que puede atender el restaurante. Relacionado al volumen de residuos.
Nivel de congestión	Nivel de actividad del restaurante (bajo, medio, alto).
Observaciones	Campo de texto para registrar comentarios adicionales o situaciones especiales.

Ustedes deben de realizar como analista de datos:

- Revisar la estructura de la base de datos.
- Identificar las variables que deben separarse en tablas.
- Diseñar un modelo de datos adecuado (tipo estrella).
- Construir las tablas necesarias para el análisis.
- Establecer relaciones entre las tablas.

Para el modelado de la base de datos, el participante debe crear un conjunto de tablas estructuradas bajo un esquema tipo estrella, separando la información en una tabla principal de hechos y varias tablas de dimensiones. En primer lugar, se debe construir la tabla Hechos_Residuos, la cual contendrá los registros principales del evento, incluyendo el identificador, la fecha, el identificador del restaurante, el tipo de residuo, la cantidad en kilogramos, el turno y el método de disposición; esta tabla concentra las variables cuantitativas y será la base para los análisis. Asimismo, se debe crear la tabla Dim_Tiempo, que incluirá la fecha junto con sus atributos derivados como año, mes y día de la semana, permitiendo realizar análisis temporales detallados. Por otro lado, la tabla Dim_Restaurante debe contener la

Instituto Nacional de Estadística e Informática	Escuela Nacional de Estadística e Informática
campusvirtual@inei.gob.pe	

información descriptiva de los establecimientos, como el nombre del restaurante, tipo de restaurante, capacidad y zona, facilitando la segmentación por características del negocio. También se debe generar la tabla Dim_Residuo, que incluirá los diferentes tipos de residuos orgánicos, lo cual permitirá clasificar y analizar la composición de los desechos. De igual forma, la tabla Dim_Turno contendrá los distintos momentos del día en que se generan los residuos, permitiendo evaluar variaciones horarias. Finalmente, se debe crear la tabla Dim_Disposicion, que agrupe el método de disposición de los residuos y la participación en programas municipales, lo que permitirá analizar prácticas de gestión ambiental. Todas estas tablas deben relacionarse con la tabla de hechos mediante claves adecuadas, garantizando un modelo organizado, eficiente y apto para el análisis.