

Gráficos y Librerías Recomendadas

1. Seaborn: Gráficos Estadísticos

Seaborn es la librería más recomendada por su capacidad para manejar comparaciones y distribuciones estadísticas complejas con código limpio.

Gráfico	Uso para Brecha Digital
Gráfico de Barras (Bar Plot)	Muestra el porcentaje de acceso a internet por grupos categóricos (ej. Urbana vs. Rural, o por nivel educativo). Es la forma más directa de mostrar la diferencia (la brecha).
Gráfico de Violín (Violin Plot)	Excelente para mostrar la distribución del gasto en servicios digitales en función de los ingresos, resaltando cómo la distribución es diferente entre grupos.
Mapa de Calor (Heatmap)	Visualiza la correlación entre diferentes factores de la brecha digital (ej. relación entre la edad y el uso de servicios gubernamentales en línea).

Exportar a Hojas de cálculo

2. Matplotlib: Personalización y Línea Base

Matplotlib ofrece la base para construir cualquier gráfico, siendo ideal para una visualización más detallada o con anotaciones personalizadas.

Gráfico	Uso para Brecha Digital
Gráfico de Líneas (Line Plot)	Muestra la evolución de la brecha digital a lo largo del tiempo (ej. cómo ha cambiado la diferencia de acceso entre hombres y mujeres desde 2010 hasta 2024).
Gráfico de Dispersión (Scatter Plot)	Compara el índice de acceso a internet (eje Y) con el ingreso promedio (eje X) a nivel de provincias o estados, permitiendo identificar <i>clusters</i> geográficos con mayor o menor brecha.

Exportar a Hojas de cálculo

Ejemplo con un Gráfico de Barras (Seaborn/Matplotlib)

Un **Gráfico de Barras Agrupadas** es la forma más efectiva de mostrar la magnitud de la brecha.

Supongamos que tienes datos de porcentaje de acceso a internet por zona (Urbana/Rural) y por grupo de edad.

```

Python
import matplotlib.pyplot as plt
import seaborn as sns
import pandas as pd

# Datos de ejemplo
data = {
    'Zona': ['Urbana', 'Rural', 'Urbana', 'Rural'],
    'Grupo_Edad': ['Jóvenes', 'Jóvenes', 'Adultos', 'Adultos'],
    'Acceso_Internet (%)': [90, 65, 80, 40]
}
df = pd.DataFrame(data)

# Crear el gráfico de barras agrupadas
plt.figure(figsize=(10, 6))
sns.barplot(
    x='Grupo_Edad',
    y='Acceso_Internet (%)',
    hue='Zona',
    data=df,
    palette='viridis'
)

plt.title('Brecha Digital: Acceso a Internet por Edad y Zona', fontsize=16)
plt.ylabel('Porcentaje de Acceso (%)', fontsize=12)
plt.xlabel('Grupo de Edad', fontsize=12)
plt.legend(title='Zona')
plt.grid(axis='y', linestyle='--')
plt.show()
#

```

Este gráfico de barras agrupa la información por edad y usa el color para diferenciar la zona, haciendo que la **diferencia de altura** (la brecha) sea visualmente obvia entre lo Urbano y lo Rural en cada grupo.

La clave es que las tablas de hechos registran los **eventos medibles**, mientras que las de dimensión describen el **contexto** de esos eventos.

Cómo Enfocar el Trabajo y Armar las Tablas

Para construir las tablas de hechos y dimensión, el primer paso es definir el **proceso de negocio** y las **métricas** que quieres analizar.

Paso 1: Definir el Proceso y las Métricas

Pregúntate: ¿Qué voy a medir y analizar?

- **Proceso de Negocio:** El evento que estás siguiendo (e.g., "Acceso a Internet en el Hogar", "Uso de Servicios Gubernamentales Online", "Compra de Dispositivos Electrónicos").
- **Métrica (Hecho):** Lo que vas a sumar, contar o promediar (e.g., **porcentaje_acceso**, **gasto_promedio**, **cantidad_usuarios**).

Paso 2: Construir la Tabla de Hechos (Fact Table)

La tabla de hechos contiene principalmente **claves foráneas** a las tablas de dimensión y los **valores numéricos** (las métricas).

- **Contenido:**
 - **Claves de Dimensión:** **ID_Tiempo**, **ID_Geografia**, **ID_Persona**, **ID_Dispositivo**, etc. (Estas son las FKs que enlazan con las Dimensiones).
 - **Hechos (Métricas):** **Porcentaje_Acceso**, **Gasto_Mensual_Conectividad**, **Minutos_Uso_Promedio**.
 - **Granularidad:** Define el nivel de detalle de cada fila (e.g., ¿es por persona, por hogar, por mes?). **La granularidad debe ser la misma para toda la tabla de hechos.**

Ejemplo de Fila de Hecho: Una fila podría ser: "En el Mes (**ID_Tiempo**) de Enero 2024, el Hogar (**ID_Hogar**) X, en la Ciudad (**ID_Geografia**) de Santiago, tuvo un **Porcentaje_Acceso** de 85%."

3: Construir las Tablas de Dimensión (Dimension Tables)

Las tablas de dimensión proporcionan el **contexto** detallado sobre **quién, qué, dónde, cuándo y cómo** ocurrió el evento registrado en la tabla de hechos.

- **Regla:** Deben ser **descriptivas** y contener los atributos que se usarán para **filtrar, agrupar y etiquetar** los resultados.

Dimensión	Atributos (Columnas Descriptivas)	Uso Analítico
Dim_Tiempo	Año, Trimestre, Mes, Dia_Semana, Es_Fin_Semana.	Analizar tendencias temporales.
Dim_Geografia	Pais, Region, Provincia, Zona (Urbana/Rural), Clima.	Identificar la brecha regional.
Dim_Poblacion	Rango_Edad, Genero, Nivel_Educativo, Nivel_Ingreso, Tamaño_Hogar.	Entender los grupos afectados por la brecha.
Dim_Dispositivo	Tipo_Dispositivo (Smartphone, PC, Tablet), Conexion (Fibra, 3G/4G/5G), Proveedor.	Analizar los medios de acceso.



Lo que el Gráfico Debe Mostrar

Una vez que tengas estas tablas, el enfoque para la visualización es usar los atributos de las dimensiones para **segmentar** y las métricas de la tabla de hechos para **medir la diferencia**.

La Brecha Digital se muestra al:

1. **Agrupar** (usando una Dimensión): Por ejemplo, agrupas por **Zona** (Dim_Geografia).
2. **Calcular** (usando la Tabla de Hechos): Calculas el promedio de **Porcentaje_Acceso** para cada grupo.
3. **Visualizar la Diferencia**: Muestras la gran diferencia entre el promedio de la zona Urbana y el promedio de la zona Rural.

Ejemplo de Consulta para el Gráfico:

```
SQL
SELECT
    d.Zona, -- (Dimensión: Agrupación)
    AVG(f.Porcentaje_Acceso) AS Acceso_Promedio -- (Hecho: Métrica)
FROM
    Fact_Acceso f
JOIN
    Dim_Geografia d ON f.ID_Geografia = d.ID_Geografia
GROUP BY
    d.Zona;
```

Con los resultados de esta consulta, usas **Seaborn** o **Matplotlib** para crear un Gráfico de Barras que muestre la diferencia de acceso entre Urbano y Rural, haciendo visible la brecha.