

Taller 3 Modelamiento

Samuel Monsalve, Samuel Rodriguez, Miguel Espinosa.

1. Modelaje

- **Modelo dimensional**

FACT_COVID_DIARIA (tabla de hechos)

- **Claves** : date_id, country_sk
- **Medidas aditivas**: new_cases, new_deaths, new_tests, new_vaccinations, weekly_icu_admissions, weekly_hosp_admissions
- **Medidas comparativas** : new_cases_per_million, new_deaths_per_million, new_tests_per_thousand, reproduction_rate, stringency_index
- **Medidas semi-aditivas**: total_cases, total_deaths

DIM_TIEMPO

- **PK**: date_id (YYYYMMDD)
- **Atributos / jerarquías**: date, year, quarter, month, day, week_iso, month_name
(Jerarquía: Año → Trimestre → Mes → Día)

DIM_PAIS

- **PK**: country_sk (surrogate key)
- **Identificadores**: country_id (código OWID/ISO), country, continent
- **Atributos de contexto (último valor conocido, SCD-1)**: population, population_density, median_age, gdp_per_capita, life_expectancy, human_development_index, hospital_beds_per_thousand, diabetes_prevalence

Relaciones (estrella):

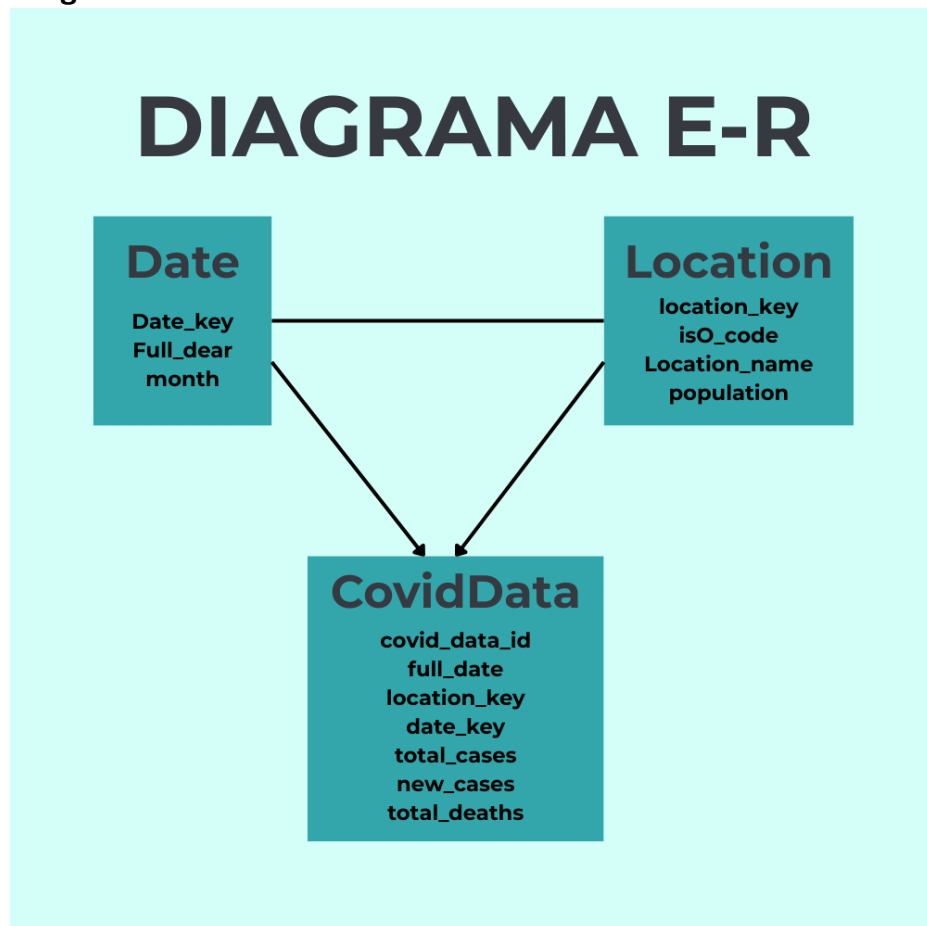
- FACT_COVID_DIARIA[date_id] → DIM_TIEMPO[date_id] (1:*)
- FACT_COVID_DIARIA[country_sk] → DIM_PAIS[country_sk] (1:*)

- **Explicacion Modelo dimensional.**

El modelo se creó de esta forma ya que facilita la organización de los datos de manera clara y estructurada. La tabla principal almacena los datos más relevantes (casos, fallecimientos, evidencias, etc.), en tanto que las tablas de tiempo y país añaden el contexto requerido para comprender dichos números. Se optó por el nivel de detalle de día y país, ya que es el método de registro de los datos que permite examinar la evolución de la situación en cada sitio a través del tiempo.

La disposición en forma de estrella simplifica las comparaciones y consultas, dado que distingue lo que se evalúa (los datos) de las propiedades que caracterizan dichos datos (las dimensiones). Esto permite un análisis más simple, más ágil y sencillo de expandir en el futuro si se requieren nuevas variables.

- **Diagrama de E-R**



2. Visualización

Ver Dashboard

- Analizar la distribución geográfica de los casos de COVID
- Comparar la distribución total de las muertes por COVID a través de los distintos continentes para ver las zonas con mayor impacto
- Identificar los patrones de contagio y la velocidad del impacto

Los tres modismos elegidos forman un buen diseño, ya que cada uno responde a una pregunta diferente y utiliza apropiadamente los canales visuales: El mapa coroplejico posibilita la identificación intuitiva de la distribución espacial de los casos de COVID por país, destacando magnitudes por medio del lugar e intensidad del color; en cambio, el gráfico circular permite ver con claridad y facilidad la proporción de muertes según el continente, lo que ayuda a comparar regiones entre sí sin sobrecargar al usuario con muchas categorías, y el gráfico de áreas resalta la progresión a lo largo del tiempo de los nuevos casos normalizados por millón de habitantes, exhibiendo las olas de contagio con más claridad que un gráfico lineal y garantizando que se puedan comparar entre naciones. En total, estos tres modismos tienen interactividad con los slicers de fecha, continente y país, así como claridad visual que posibilita el análisis de los datos sin crear confusión, cumpliendo de este modo con las buenas prácticas de diseño vistas en clase.