



Data analytics

2° Trabajo Entregable

Javier Yañez

Comisión 71725

2024

Descripción del tema

La contaminación ambiental es la presencia de componentes nocivos, bien sean de naturaleza biológica, química o de otra clase, en el medioambiente, de modo que supongan un perjuicio para los seres vivos que habitan un espacio, incluyendo, por supuesto, a los seres humanos. Generalmente la contaminación ambiental tiene su origen en alguna actividad humana.

Las partículas en suspensión y gases producidos por el tráfico rodado, la industria y las calefacciones son los principales causantes de la contaminación atmosférica. El incremento de la emisión de CO₂ provoca el calentamiento global que deriva en el cambio climático y por consiguiente provoca los efectos climáticos adversos como olas de calor, sequía, inundaciones, etc. La principal fuente de contaminación atmosférica son los gases ozono troposférico (O₃), óxidos de azufre (SO₂ y SO₃), óxidos de nitrógeno (NO y NO₂), benzopireno (BaP) y las partículas en suspensión (PM). Estos gases se derivan principalmente de las emisiones provocadas por la quema de combustibles fósiles (incluidas las emisiones generadas por el transporte), los procesos industriales, la quema de bosques, el empleo de aerosoles y la radiación.

Objetivo

En el presente proyecto se pretende analizar la evolución de la calidad del aire y la emisión de gases contaminantes, así como el impacto de esta contaminación en la temperatura en los países de Latinoamérica en los últimos años. Se hará un análisis semestral como anual de la temática.

Hipótesis

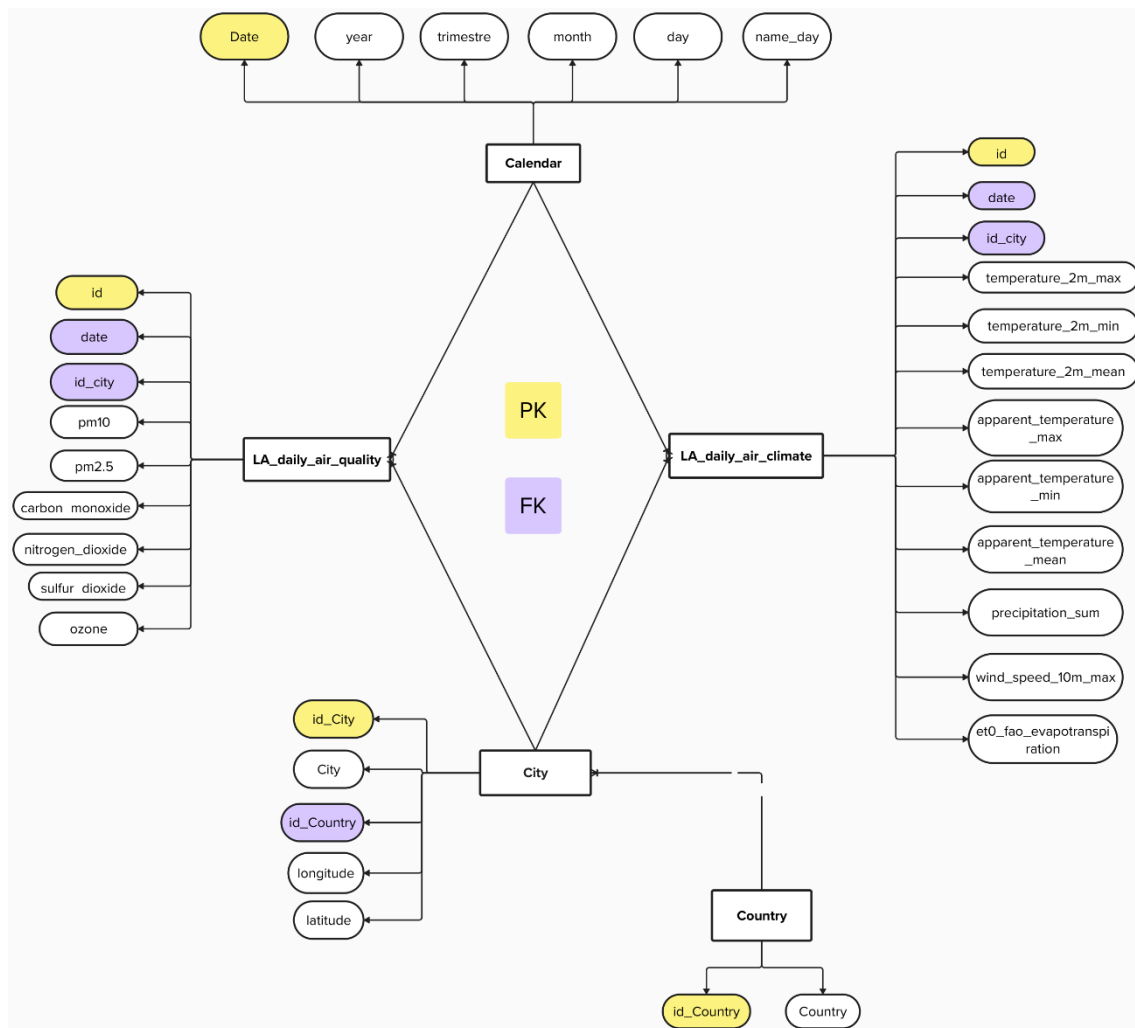
La emisión de gases sigue en aumento, contaminando la atmosfera, lo que contribuye al calentamiento global y, por ende, al aumento de la temperatura.

Nivel de aplicación del análisis

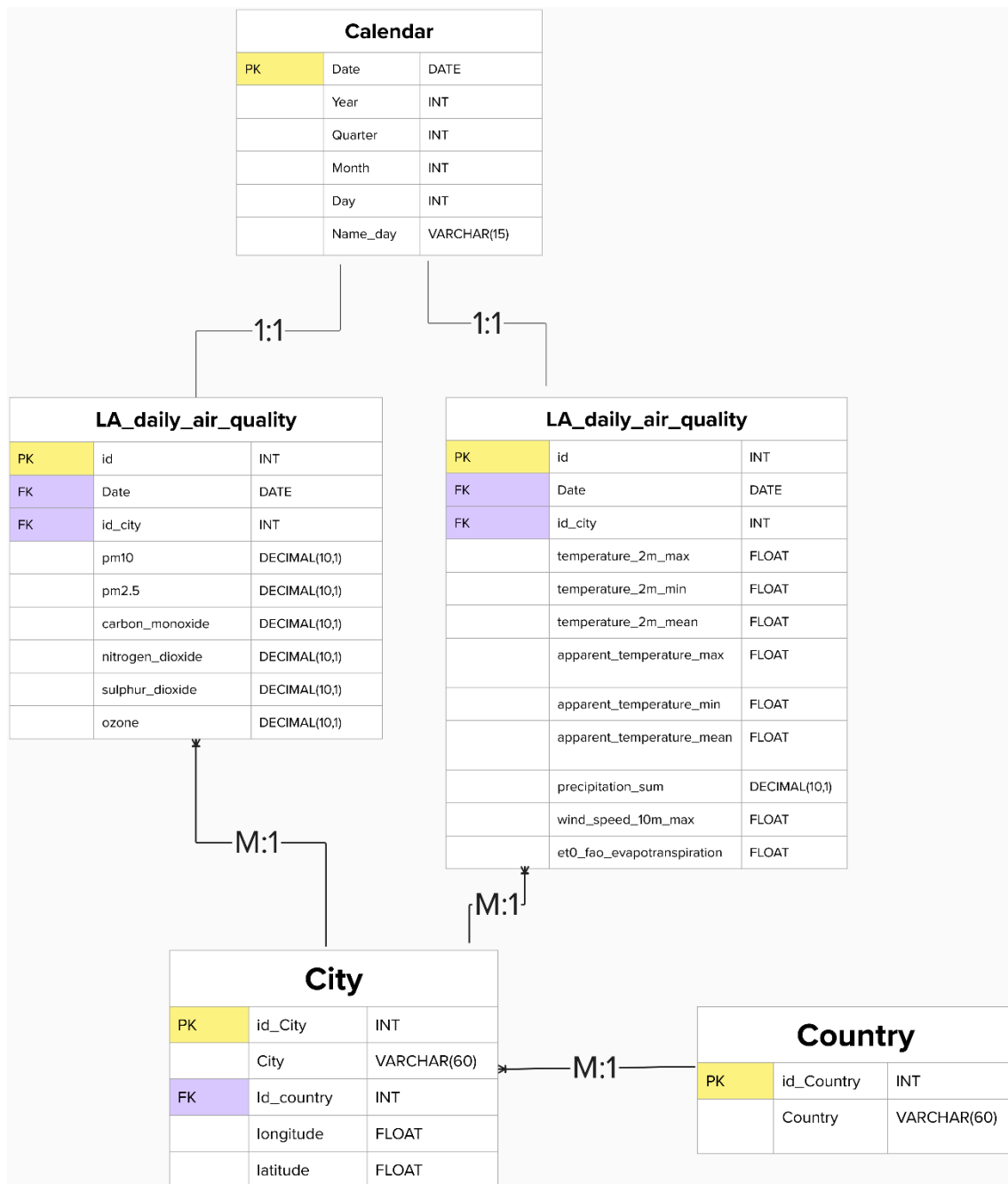
Los tableros serán para el uso a nivel estratégico, siendo los dirigentes de cada país los destinatarios de los mismos para su análisis.

Diagramas

A continuación, se presentará el DER del proyecto a realizar. Se buscará exponer de manera clara las relaciones entre las distintas tablas de hecho y dimensionales de los datos disponibles. Al centro se ubican las dos tablas de hecho, arriba el calendario y abajo la información normalizada, siguiendo la tercera forma normal, sobre la locación del hecho observado dentro de las tablas City y Country



En el siguiente gráfico, se mostrará este diagrama ahondando en el tipo de dato de las tablas y el tipo de relación que hay entre ellas.



Acerca del dataset

Clima y calidad del aire en Latinoamérica

Este dataset contiene 2 archivos CSV generados a partir de la API Open-Meteo y contiene información sobre varios países de Latinoamérica. Los datasets incluyen nombre del país, capital, clima y calidad del aire. Los países incluidos en estos datos son:

Argentina - Buenos Aires

Bolivia - Sucre

Brazil - Brasilia

Chile - Santiago

Colombia - Bogota

Costa Rica - San Jose

Cuba - Havana

Dominican Republic - Santo Domingo

Ecuador - Quito

El Salvador - San Salvador

Guatemala - Guatemala City

Haiti - Port-au-Prince

Honduras - Tegucigalpa

Mexico - Mexico City

Nicaragua - Managua

Panama - Panama City

Paraguay - Asuncion

Peru - Lima

Uruguay - Montevideo

Venezuela - Caracas

Tablas

En este apartado se detallará cada tabla de la base de datos, incluyendo nombre, descripción, tipo de dato y si es key o no cada campo.

Calendario

Contiene información sobre cada fecha de cada año.

Campo	Descripción	Tipo de dato	Key
Date	Fecha	DATE	PK
Year	Año	INT	-
Quarter	Trimestre del año	INT	-
Month	Mes del año	INT	-
Day	Día del año	INT	-
Name_day	Nombre del día de la semana	VARCHAR(15)	-

LA daily air quality

Tabla que contiene las mediciones de los distintos agentes que contaminan el medioambiente en la zona en cada día.

Campo	Descripción	Tipo de dato	Key
Id	Identificador del suceso	INT	PK
Date	Fecha del suceso	DATE	FK
Id_City	Identificador del lugar de medición	INT	FK
Pm10	Cantidad de partículas de materia inferior a 10 µm hasta a 10 mts del suelo	DECIMAL(10,1)	-
Pm2.5	Cantidad de partículas de materia inferior a 2.5 µm hasta a 10 mts del suelo	DECIMAL(10,1)	-
Carbon_monoxide	monóxido de carbono cerca de la superficie (10mts del suelo)	DECIMAL(10,1)	-
Nitrogen_dioxide	dióxido de nitrógeno cerca de la superficie	DECIMAL(10,1)	-
Sulfur_dioxide	dióxido de azufre cerca de la superficie	DECIMAL(10,1)	-
Ozone	Ozono cerca de la superficie	DECIMAL(10,1)	-

LA daily air climate

Tabla que contiene las mediciones de las temperaturas, sensación térmica, precipitaciones, viento y evapotranspiración registradas en la zona en cada día.

Campo	Descripción	Tipo de dato	Key
Id	Identificador del suceso	INT	PK
Date	Fecha del suceso	DATE	FK
Id_City	Identificador del lugar de medición	INT	FK
Temperature_2m_max	temperatura máxima a 2 metros del suelo	FLOAT	-
Temperature_2m_min	temperatura mínima a 2 metros del suelo	FLOAT	-
Temperature_2m_mean	temperatura media a 2 metros del suelo	FLOAT	-
Apparent_temperature_max	Sensación térmica máxima	FLOAT	-
Apparent_temperature_min	Sensación térmica mínima	FLOAT	-
Apparent_temperature_mean	Sensación térmica media	FLOAT	-
Precipitation_sum	Precipitación total del día	DECIMAL(10,1)	-
Wind_speed_10m_max	Velocidad máxima del viento a 10 metros del suelo	FLOAT	-
Et0_fao_evapotranspiration	evapotranspiración potencial calculada	FLOAT	-

City

Contiene el nombre de la ciudad, sus coordenadas y el país al que pertenecen.

Campo	Descripción	Tipo de dato	Key
Id_City	Identificador de la ciudad	INT	PK
City	Nombre de la ciudad	VARCHAR(60)	-
Id_Country	Identificador del país	INT	FK
Longitude	Coordenadas de longitud de la ciudad	FLOAT	-
Latitude	Coordenadas de latitud de la ciudad	FLOAT	-

Country

Contiene el identificador del país y el nombre del país

Campo	Descripción	Tipo de dato	Key
Id Country	Identificador del país	INT	PK
Country	Nombre del país	VARCHAR(60)	-

Unidades

Temperatura -> C°

Precipitación -> mm (1 milímetro de lluvia es igual a 1 litro por metro cuadrado)

Evapotranspiración -> mm (1 mm de evapotranspiración por hora equivale a 1 litro de agua por metro.) El valor presentado es la sumatoria del día

PM -> $\mu\text{g}/\text{m}^3$ (microgramos/metro cúbico)

Gases atmosféricos -> $\mu\text{g}/\text{m}^3$

Fuente y links relacionados

https://www.kaggle.com/datasets/anycaroliny/latin-america-weather-and-air-quality-data?resource=download&select=LA_daily_climate.csv

https://github.com/anycarolinys/kaggle_weather_dataset