

Primera Entrega Proyecto Ingeniería de Datos

González Laura, Mora Diryon, Rincón Laura

28 de Abril de 2022

1. Base de datos

Colombianos registrados en el exterior

Como se puede apreciar, la base de datos seleccionada incluye la información de los colombianos que viven en el extranjero, los datos son totalmente anónimos e incluye información desde el 2016 hasta inicios del 2022. Sin embargo, no cuenta con la fecha en la que decidieron emigrar. Por otro lado, con el fin de mantener la información, se cuenta con el nombre del país donde viven y su respectivo código ISO, junto a la ciudad donde están registrados con su respectiva coordenada geográfica.

Se tiene la edad del colombiano y en función a esta el grupo al que pertenece (infante, adolescente, adulto, etc.). También se incluye el área en la que trabaja y una especificación de la misma. De manera independiente a lo anterior se incluye el nivel académico del individuo. Junto a otros datos personales como lo son: el estado civil, el género, la etnia, la estatura. Para finalizar los datos enumeran la cantidad de personas que cumplen con exactamente las mismas condiciones mencionadas anteriormente.

2. Problema de datos

Se pueden aplicar filtros que permitan conocer la inmigración de poblaciones específicas, dígame, por ejemplo, las personas de género masculino o quienes poseen algún tipo de nivel educativo característico.

3. Reglas de negocio

1. La descripción demográfica esta organizadas por países.
2. Para cada país es necesario saber en cuál continente concreto está ubicado, su nombre y su código iso.
3. Un país puede tener muchas descripciones demográficas.
4. Cada continente tiene un nombre y un código.
5. Un continente posee muchos países.
6. Un área de conocimiento está identificada con un único id y un nombre.
7. Un área de conocimiento puede tener varias especializaciones.
8. Cada especialización se identifica con un único id y con un nombre.
9. Una especialización contiene varias descripciones demográficas.
10. La descripción demográfica hace referencia a aspectos de una persona.
11. La descripción demográfica se identifica por un único id, además es necesario saber la edad, el género, la estatura, la oficina_registro (la ciudad donde se encuentra), el nivel académico y la cantidad de personas que cumplen estas mismas condiciones.
12. Una persona puede tener solo una especialización.

4. Tabla Entidad - Relacion

| Entidad | Atributos | Relaciones |
|----------------------------|--|---|
| CONTINENTE | codigo nombre | PAIS [M] PERTENECER |
| PAIS | codigo nombre | CONTINENTE [1] PERTENECER DESCRIPCION_DEMOGRAFICA [M] TENER |
| DESCRIPCION DEMOGRAFICA | id edad genero estatura cantidad_personas oficina_registro nivel_academico | PAIS [1] TENER ESPECIALIZACION [1] CONOCER |
| ESPECIALIZACION | id nombre | DESCRIPCION_DEMOGRAFICA [M] CONOCER AREA_CONOCIMIENTO [1] INTEGRAR |
| AREA CONOCIMIENTO | id nombre | ESPECIALIZACION [M] INTEGRAR |

Así pues, la información se encuentra específicamente de la siguiente forma:

- CONTINENTE

- codigo

- Tipo de dato: varchar.

- Recibe siglas que identifican el continente, cada uno consta de 2 letras en mayúsculas.

- nombre Tipo de dato: varchar.

- Nombre asociado al codigo de cada continente. Sólo recibe nombres en mayúscula.

- PAIS

- codigo

Tipo de dato: varchar.

Recibe siglas que identifican el país, cada uno consta de 3 letras en mayúsculas.

- nombre Tipo de dato: varchar.

Nombre asociado al codigo_iso de cada país. Sólo recibe nombres en mayúscula.

■ DESCRIPCION_DEMOGRAFICA

- id

Tipo de dato: integer.

Identificador único.

- edad

Tipo de dato: int.

Edad del ciudadano.

- genero

Tipo de dato: varchar.

Permite identificar a una persona como femenina o masculina. Puede ser nulo.

- cantidad_personas

Tipo de dato: smallint.

Cantidad de personas residentes en el exterior que cumplen con las condiciones demográficas.

- oficina_registro

Tipo de dato: varchar.

Incluye el nombre de la ciudad en la que el colombiano se encuentra.

- nivel_academico

Tipo de dato: varchar.

Nivel educativo alcanzado por la persona. Sus posibles valores son: PRIMARIA, NO INDICA, BACHILLERATO, PREGRADO - PROFESIONAL, PREGRADO - TECNOLÓGICO, POSTGRADO - ESPECIALIZACIÓN, POSTGRADO

- DOCTORADO, PREGRADO - TÉCNICO PROFESIONAL y POSTGRADO
- MAESTRIA.

■ ESPECIALIZACION

- id

Tipo de dato: serial.

Creado artificialmente para dar un identificador.

- nombre

Tipo de dato: varchar.

Agrupación que se hace de los programas académicos que tienen un área en común.

Sus posibles valores son: "INGENIERÍA MECÁNICA Y AFINES", "PERIODISMO, COMUNICACIÓN SOCIAL Y AFINES", "ARTES PLÁSTICAS, VISUALES Y AFINES", "DISEÑO, NINGUNA, CONTADURÍA PÚBLICA, OTRAS INGENIERÍAS, MEDICINA, "INGENIERÍA ELÉCTRICA Y AFINES", "INGENIERÍA EN SISTEMAS, TELEMÁTICA Y AFINES", ADMINISTRACIÓN, ZOOTECNIA, "OTRO PROGRAMA DE BELLAS ARTES", ARQUITECTURA, AVIACIÓN, "CIENCIA POLÍTICA Y/O RELACIONES INTERNACIONALES", ENFERMERÍA, TERAPIAS, "OTRO PROGRAMA DE SALUD", "INGENIERÍA ELECTRÓNICA, TELECOMUNICACIONES Y AFINES", "PUBLICIDAD Y AFINES", "INGENIERÍA DE PETRÓLEOS", "INGENIERÍA AGRÍCOLA, FORESTAL Y AFINES", EDUCACIÓN, "INGENIERÍA BIOMÉDICA Y AFINES", "INGENIERÍA INDUSTRIAL Y AFINES", "INGENIERÍA CIVIL Y AFINES", "PSICOLOGÍA Y AFINES", "INGENIERÍA DE MINAS, METALURGIA Y AFINES", "INGENIERÍA ADMINISTRATIVA Y AFINES", "COCINA Y CULINARIA", "DERECHO Y AFINES", "OTROS ESTUDIOS EN CIENCIAS SOCIALES Y HUMANAS", ECONOMÍA, "INGENIERÍA AMBIENTAL, SANITARIA Y AFINES", OPTOMETRÍA, ODONTOLOGÍA, "INSTRUMENTACIÓN QUIRÚRGICA", AGRONOMÍA, MÚSICA, "GEOGRAFÍA O HISTORIA", "ANTROPOLOGÍA O ARTES LIBERALES", "OTROS PROGRAMAS DE CIENCIAS NATURALES", BACTERIOLOGÍA, "QUÍMICA Y AFINES", SALUD PÚBLICA, "INGENIERÍA AGROINDUSTRIAL, ALIMENTOS Y AFINES", "FILOSOFÍA, TEOLOGÍA Y AFINES", SALUD OCUPACIONAL (HSEQ), MEDICINA VETERINARIA, GEOLOGÍA, "LENGUAS MODERNAS, FILOLOGÍA, LINGÜÍSTICA Y AFINES", "INGENIERÍA QUÍMICA Y AFINES", "MATEMÁTICAS, ESTADÍSTICA Y AFINES", "SOCIOLOGÍA, TRABAJO SOCIAL Y AFINES", FÍSICA, "BIOLOGÍA, MICROBIOLOGÍA Y AFINES", FORMACIÓN MILITAR, BIBLIOTECOLOGÍA, NUTRICIÓN Y DIETÉTICA, ARTES DRAMÁTICAS Y REPRESENTATIVAS, "DEPORTES, EDUCACIÓN

FÍSICA Y RECREACIÓN“ y “INGENIERÍA AGRONÓMICA, PECUARIA Y AFINES."

■ AREA_CONOCIMIENTO

- id

Tipo de dato: serial.

Creado artificialmente para dar un identificador.

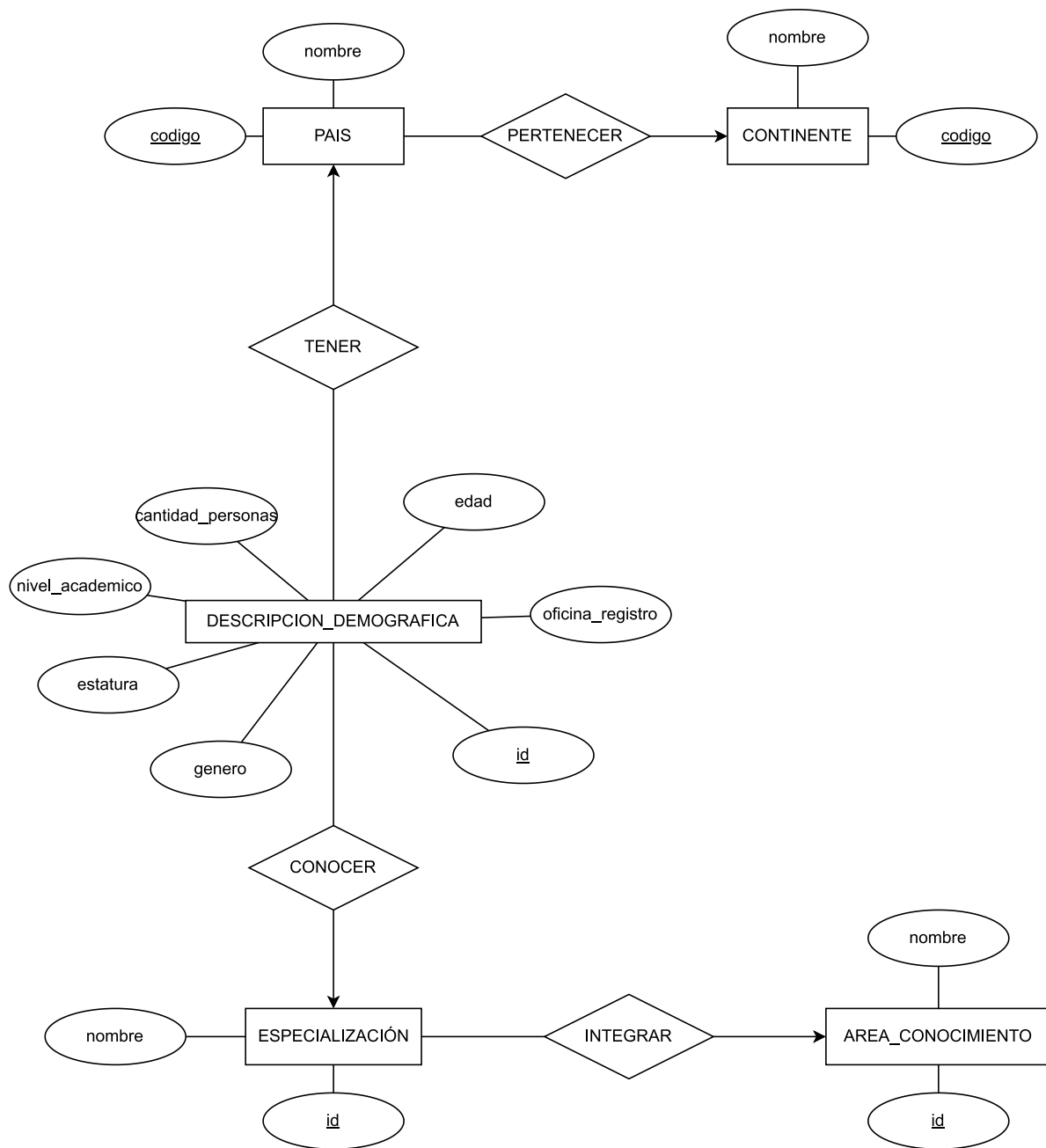
- nombre

Tipo de dato: varchar.

Agrupación que se hace de los programas académicos, teniendo en cuenta cierta afinidad en los contenidos.

Sus posibles valores son: INGENIERÍA, ARQUITECTURA Y AFINES, CIENCIAS SOCIALES Y HUMANAS, BELLAS ARTES, NINGUNA, .^ECONOMÍA, ADMINISTRACIÓN CONTADURIA Y AFINES", CIENCIAS DE LA SALUD, .^AGRONOMÍA, VETERINARIA Y AFINES", CIENCIAS DE LA EDUCACIÓN y "MATEMÁTICAS Y CIENCIAS NATURALES".

5. Diagrama Entidad - Relación



6. Diagrama Relacional

