**Máster Universitario en Big Data y Ciencia de Datos**

**ASIGNATURA: 02MBID *Sistemas de almacenamiento y gestión Big Data***

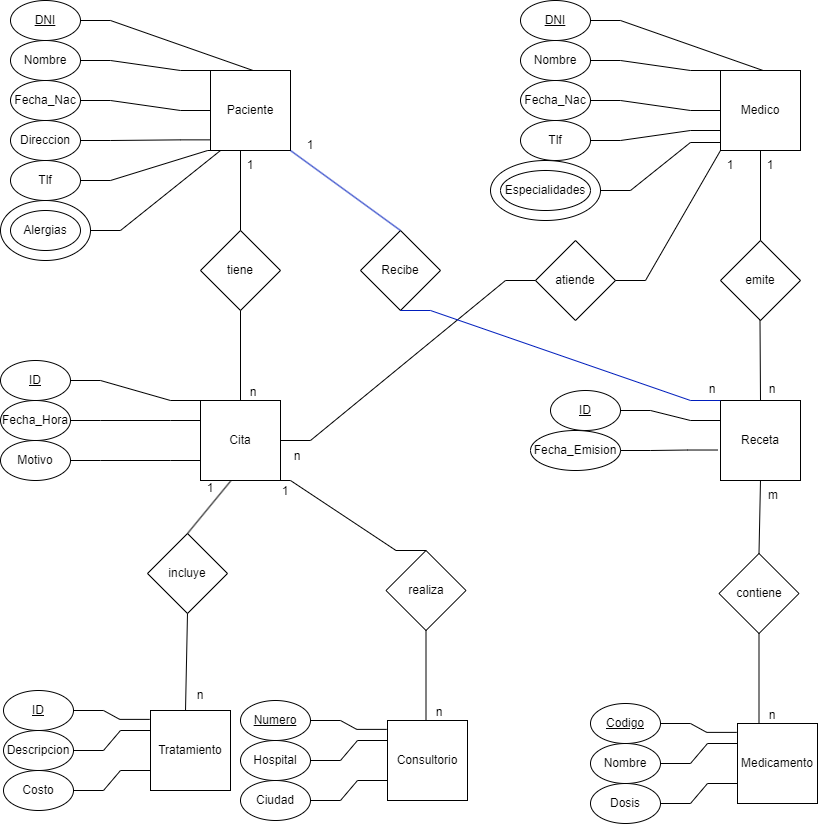
*Actividad 1 - Creación del esquema de una base de datos orientada a columnas*

Alumno: **Bru Montes, Israel**

Edición **Octubre 2024** a 2/11/2024

# Parte 1: Creación del esquema de una base de datos orientada a columnas.

**Modelo conceptual**



**Consultas a satisfacer**

**1. Obtener toda la información de un paciente en base a su nombre.**

|  |  |
| --- | --- |
| **Tabla1** | |
| Paciente\_DNI | Clustering Key |
| Paciente\_Nombre | Partition Key |
| Paciente\_Fecha\_Nac |  |
| Paciente\_Direccion |  |
| Paciente\_Tlf |  |
| Paciente\_Alergias |  |

*Justificación:*

En la tabla 1: contendría la búsqueda por nombre de paciente de todas sus citas, indicamos como “Clustering key” tanto Id de la Cita y DNI del Paciente (sería la clave) y como el campo de búsqueda es el nombre del patiente, lo añadiriamos en la tabla como Partition Key.

**2. Obtener según el DNI de un médico todas las citas que atiende.**

|  |  |
| --- | --- |
| **Tabla2** | |
| Medico\_DNI | Clustering Key |
| Medico\_Nombre |  |
| Medico\_Fecha\_Nac |  |
| Medico\_Tfl |  |
| Medico\_Especialidades |  |
| Cita\_ID | Clustering Key |
| Cita\_Fecha\_Hora | *\*\* Partition Key* |
| Cita\_Motivo |  |

*Justificación:*

En esta tabla de consulta pondría 2 atributos que forman parte del Clustering Key y otros 2 con información de la cita.

Aunque el enunciado no lo especifica, sería interesante añadir los 2 atributos “Fecha y Hora” y Motivo especificando como Partition Key el atributo “Fecha y Hora” para facilitar las búsquedas por fecha y hora en el caso de querer recuperar las citas por fecha y hora.

**3. Obtener según el DNI de un paciente, todos los tratamientos que tiene incluidos en las citas que tuvo.**

|  |  |
| --- | --- |
| **Tabla 3** | |
| Paciente\_DNI | Partition Key |
| Cita\_ID | Clustering Key |
| Tratamiento\_ID | Clustering Key |
| Tratamiento\_Descripcion |  |
| Tratamiento\_Costo |  |

*Justificación:*

En la tabla 3, identificamos como Partition key el atributo DNI de la entidad Paciente. Como Clustering Key los atributos Cita\_ID y Tratamiento\_ID que identifican para cada paciente la relacion de sus citas y los tratamientos por cada cita.

A parte mostramos todos los datos de cada tratamiento como información adicional.

**4. Obtener cuantas citas tiene un paciente.**

|  |  |
| --- | --- |
| **Tabla 4** | |
| Paciente\_DNI | Partition Key |
| num\_citas\_paciente | + |

*Justificación:*

En la tabla 4 asignamos como partition key el atributo “Paciente\_DNI” y el atributo “num\_citas\_paciente” como columna agregada que contiene el número de citas por paciente.

**5. Obtener todas las asociaciones entre recetas y medicamentos a través de la fecha de la receta.**

|  |  |
| --- | --- |
| **Tabla 5** | |
| Receta\_Fecha\_Emision | Partition Key |
| Receta\_ID | Clustering Key |
| Medico\_DNI |  |
| Medico\_Nombre |  |
| Paciente\_DNI |  |
| Paciente\_Nombre |  |
| Medicamento\_Codigo | Clustering Key |
| Medicamento\_Nombre |  |
| Medicamento\_Dosis |  |

*Justificación:*

En la tabla 5 asignamos como partition key el atributo “Paciente\_DNI” y el atributo “num\_citas\_paciente” como columna agregada que contiene el número de citas por paciente (entidad).

**6. Obtener los pacientes que tienen una alergia en concreto**

|  |  |
| --- | --- |
| **Tabla 6** | |
| Paciente\_Alergia | Partition Key |
| Paciente\_DNI | Clustering Key |
| Paciente\_Nombre |  |

**Justificación:**

En la tabla 6 asignamos como Partition key el atributo “Paciente\_Alergia” ya que contiene la información individual de las alergias que puede contenter un Paciente (Clustering Key → Paciente\_DNI), es la representación de un atributo de conjunto (alergias).