

**Dashboard para visualizar la calificación del proceso de peticiones de PQRD de las
entidades de salud en Colombia**

Autores

Juan Carlos Castrillo Gómez
Jairo Humberto Torres

Director

Elio Higinio Cables Pérez, Ph D

Universidad Antonio Nariño
Facultad de Ingeniería de Sistemas
Especialización en Gobierno de Datos
2021

Tabla de Contenidos

1 Introducción..... 3

2 Descripción y formulación del problema 4

2.1 Planteamiento del Problema 4

2.2 Objetivo General..... 4

2.2.1 Objetivos Específicos 5

3 Marco referencial..... 6

3.1 Marco Teórico 6

3.1.1 Supersalud 6

3.1.2 Entidades Prestadoras de Salud 6

3.1.3 Definición y Uso de PRQD en la Supersalud..... 7

3.1.4 Método Ideal de Referencia..... 10

3.2 Descripción Herramientas 14

3.3 Estado del Arte 16

3.4 Impacto 19

3.5 Componente de Innovación 19

4 Metodología..... 20

4.1 Ingestión y Capturas de Datos 21

4.2 Lineamientos de Gobierno de Datos 22

4.3 Preparación y Almacenamiento de los Datos 22

4.4 Transformación de Datos..... 23

4.5 Visualización Datos 23

5 Desarrollo de la propuesta 24

5.1 Ingestión y Captura de Datos..... 24

5.1.1 Orígenes de los Datos 24

5.1.2 Fase Exploratoria de Datos..... 25

5.1.3 Metadatos 27

5.1.4 Implementación de Ambientes 29

5.1.5 Almacenamiento de Datos Stage..... 30

5.2 Definición de los Lineamientos del Gobierno de Datos..... 33

5.2.1 Arquitectura Tecnológica Física..... 34

5.2.2 Diagrama Arquitectura Lógica 35

5.2.3 Flujo de Datos de Contexto 36

5.2.4 Flujo de Datos Lógico 37

5.2.5 Lineamientos y Estandarización de Datos..... 37

5.3 Preparación y Almacenamiento de Datos..... 40

5.3.1 Descripción Problemas de Calidad de Datos..... 40

5.3.2 Datos Críticos 47

5.3.3 Medición de la Calidad de Datos, Indicadores y Métricas 48

5.3.4 Almacenamiento Datos..... 49

5.3.5 Carga Datos Ambiente Productivo 51

5.4 Transformación de los Datos..... 55

5.4.1	Método RIM	56
5.4.2	Implementación de API.....	62
5.5	Visualización de Datos	67
5.5.1	Herramienta Visualización Datos.....	67
5.5.2	Definición Datos a Visualizar	68
5.5.3	Diseño Visualización datos	68
5.5.4	Despliegue y Pruebas Tablero de Control.....	69
6	Conclusiones.....	71
7	Referencias	73
8	Anexos.....	75
8.1	Anexo-1 Diccionario de datos origen PQRD	75
8.2	Anexo-2 Metadatos PQRD.....	76
8.3	Anexo-3 Lineamientos y Estándares.....	78
8.4	Anexo-4 Modelo Flujo Datos Lógico.....	78
8.5	Anexo-5 Flujo de Datos Físico.....	80
8.6	Anexo-6 Modelo Conceptual	81
8.7	Anexo-7 Modelo Lógico Base de Datos.....	82
8.8	Anexo-8 Modelo Físico Base de Datos	83

Lista de tablas

Tabla 1. Datos Críticos PQRD.	28
Tabla 2. Ambientes creados para el proyecto.....	30
Tabla 3. Glosario de Términos.	39
Tabla 4. Problemas Detectados.	40
Tabla 5. Datos Críticos.	48
Tabla 6. Datos Críticos PQRD.	48
Tabla 7. Conteo Total PQRD.	57
Tabla 8. Descripción de Criterios.	57
Tabla 9. Cálculo Vector de Peso	58
Tabla 10. Valores Solución Ideal.	58
Tabla 11. Condiciones del Contexto de Trabajo.	59
Tabla 12. Matriz X Estados PQRD y Entidades de Salud.....	59
Tabla 13. Matriz Normalizada.....	60
Tabla 14. Matriz Normalizada con Peso.	60
Tabla 15. Variación a la Solución Ideal	60
Tabla 16. Variación a la Solución Ideal Negativa.....	61
Tabla 17. Obtención del IR Para Cada Entidad de Salud.....	61
Tabla 18. Calificaciones Proceso PQRD.....	62
Tabla 19. Entidades de Salud.	62
Tabla 20. Diccionario de Datos PQRD.	75
Tabla 21. Metadatos Descriptivos PQRD.	76
Tabla 22. Lineamientos y estándares para el proyecto.....	78

Lista de figuras

Figura 1.	Diagrama Estados PQRD	9
Figura 2.	Diagrama Flujo General Proceso PQRD	10
Figura 3.	Fórmula Obtención Matriz X	11
Figura 4.	Fórmula Calcular Matriz Normalizada.....	12
Figura 5.	Cálculo Referencia Normalizada.....	12
Figura 6.	Formula Calcular Índice Relativo	13
Figura 7.	Algoritmo Obtención Función Lineal	14
Figura 8.	Etapas a Utilizar en la Metodología	21
Figura 9.	Carga Dataset Google Colab	26
Figura 10.	Exploración Datos Google Colab.....	26
Figura 11.	Exploración Tipo de Datos Dataset.....	27
Figura 12.	Descripción Metadatos	28
Figura 13.	Visualización Ambiente Stage.....	31
Figura 14.	Importación Data PQRD	32
Figura 15.	Muestra Datos PQRD	32
Figura 16.	Conteo Registros Ingestados	33
Figura 17.	Data Strategy Proyecto PQRD	34
Figura 18.	Diagrama Arquitectura Tecnológica	35
Figura 19.	Diagrama Arquitectura Lógica.....	36
Figura 20.	Diagrama Flujo Contexto	36
Figura 21.	Ambientes AWS y Microsoft Azure	38
Figura 22.	Creación Usuarios Base de Datos.....	38
Figura 23.	Monitoreo Procesamiento Datos	39
Figura 24.	Valores Null Campo SNS_REGNOMBRE.....	41
Figura 25.	Valores Errados Campo PET_DPTO	41
Figura 26.	Valores Errados Campo DOC_MOTGEN	42
Figura 27.	Tipo de Dato Campo FECHA_CREACION	42
Figura 28.	Valores Máximos y Mínimos	43
Figura 29.	Compleitud Datos.....	43
Figura 30.	Visualización Valores Nulos	44
Figura 31.	Ajuste Tipo Formato Datos Fecha.....	45
Figura 32.	Ajuste de Valores Máximos y Mínimos.....	45
Figura 33.	Ajuste Tipo Dato Código Departamento.....	46
Figura 34.	Identificación Datos Duplicados	46
Figura 35.	Ajuste Datos Duplicados	47
Figura 36.	Eliminación Columnas	47
Figura 37.	Medición de las Dimensiones de Calidad	49
Figura 38.	Creación Tabla DATA_CLEANPQRD.....	49
Figura 39.	Creación Índices	50
Figura 40.	Conexión Base Datos AWS.....	50
Figura 41.	Carga Datos Base datos AWS	51

Figura 42.	Verificación Carga Datos	51
Figura 43.	Base Datos Producción Azure	52
Figura 44.	Verificación Creación Base de Datos	53
Figura 45.	Generación Job Carga Datos	54
Figura 46.	Creación Prejob y Postjob	54
Figura 47.	Flujo General Tablas Primer Job	55
Figura 48.	Carga de Datos por Lotes de Tablas Grandes.....	55
Figura 49.	Codificación PHP Script MCDA PQRD	63
Figura 50.	Plataforma Heroku.....	64
Figura 51.	Codificación Api PHP	64
Figura 52.	Aplicación Método Decisión Multicriterio	65
Figura 53.	Visualización Calificación Entidades.....	66
Figura 54.	Flujo Transformación Datos.....	67
Figura 55.	Diseño Gráfico Estado PQRD	69
Figura 56.	Tablero Comando	69
Figura 57.	Tablero Visualización Calificación Entidades	70
Figura 58.	Modelo Flujo Datos Lógico.....	79
Figura 59.	Flujo de Datos Físico.....	80
Figura 60.	Modelo Conceptual	81
Figura 61.	Modelo Lógico Base de Datos.....	82
Figura 62.	Modelo Físico Base de Datos	83

Abstract

An analysis of the status of the attention of petitions, complaints, claims and complaints (PQRD) filed before the Superhealth by users affiliated with the health system in Colombia, which allowed knowing the variables to deduce the need to develop a dashboard to display the rating of health entities in accordance with PQRD status.

Therefore, it is proposed to implement the RIM method for the transformation of the PQRD data obtained from the open data platforms of Colombia and Supersalud for the years 2018 and 2020, then the data prepared is loaded into a database. data for availability before the Tableau tool, in which a dashboard is built for the visualization of the qualification of health entities for end users.

Resumen

Se realizó el análisis del estado de la atención de las peticiones, quejas, reclamos y denuncias (PQRD) interpuestas ante la Supersalud por los usuarios afiliados al sistema de salud en Colombia, el cual permitió conocer las variables para deducir la necesidad de desarrollar un tablero de comandos para la visualización de la calificación de las entidades de salud de acuerdo con estado de las PQRD.

Por lo tanto, se plantea implementar el método RIM para la transformación de los datos de las PQRD obtenidos de las plataformas de datos abiertos de Colombia y la Supersalud para los años 2018 y 2020, posteriormente se realiza la carga de los datos preparados a una base de datos para la disponibilidad ante la herramienta Tableau, en la cual se construye un tablero de comandos para la visualización de la calificación de las entidades de salud para los usuarios finales.

1 Introducción

Las peticiones, quejas, reclamos y denuncias (PQRD) interpuestas ante la Supersalud, son mecanismos que tienen los usuarios afiliados al sistema de salud para presentar sus inconformidades sobre la atención en los servicios de salud de las diferentes entidades. Las PQRD se interponen a través de los canales de atención que tiene la Supersalud como lo son: página web, atención telefónica, chat personalizado, atención personalizada entre otros.

Durante el desarrollo del proyecto se describe el proceso de calificación de las entidades prestadoras de servicios de salud de Colombia con base al estado de las PQRD interpuestas ante la Supersalud, con la finalidad de presentar indicadores sobre el estado de calificación de las entidades en tablero de comandos para su visualización ante los usuarios.

Este análisis se realiza bajo la información obtenida de las PQRD interpuestas para el año 2018 y 2020 de las plataformas de datos abiertos de Colombia y la Supersalud, a la cual se le aplicó un proceso de tratamiento de calidad de datos en la herramienta Google Colab y la ingestión de datos se realiza por medio de la herramienta Talend a un sistema gestor de base de datos MySQL, donde posteriormente se aplica el modelo de decisión multicriterio MCDA con el método RIM para establecer la calificación de cada entidad acorde al estado de las PQRD interpuestas, para poder presentar dicha calificación ante el público por medio de un tablero de comandos de forma gráfica.

2 Descripción y formulación del problema

En la actualidad existe información publicada en las plataformas de datos abiertos Colombia y la Supersalud, sobre los estados de las PQRD interpuestas ante la Superintendencia de Salud por los diferentes usuarios afiliados al sistema general de salud de Colombia, sobre la atención de los servicios de salud prestados por las entidades, esta información no se encuentra centralizada de manera concisa para presentarla legiblemente ante los potenciales usuarios que desean afiliarse o cambiar de entidad prestadora del servicio de salud.

Los usuarios no tienen el conocimiento de donde se encuentra alojada la información sobre los estados de atención de las PQRD ni como estructurarla para visualizar la calificación de las entidades prestadoras y así poder tomar una decisión para afiliarse a una entidad de salud.

Sobre la base descrita anteriormente se realiza el siguiente planteamiento al problema.

2.1 Planteamiento del Problema

¿Cómo calificar el proceso de gestión de PQRD de las entidades prestadoras de servicios de salud a partir de los datos obtenidos de la plataforma de datos abiertos de Colombia?

2.2 Objetivo General

Implementar un dashboard que permita la visualización de la calificación del

proceso de gestión de las peticiones PQRD de los servicios que prestan las entidades de salud en Colombia sobre la base de datos obtenida de la plataforma de datos abiertos de Colombia y la Supersalud.

2.2.1 Objetivos Específicos

- Identificar las diferentes fuentes de datos abiertos para el análisis del proceso de peticiones PQRD de las entidades prestadoras de servicios de salud en Colombia.
- Determinar las metodologías y buenas prácticas de gobierno de datos necesarias para calificar el proceso de peticiones de PQRD interpuestas ante la Supersalud.
- Definir la infraestructura tecnológica para la visualización de la calificación de las entidades prestadoras del servicio de salud en Colombia.
- Aplicar el modelo el método RIM para la asignación de la calificación del proceso de las peticiones PQRD de las entidades prestadoras del servicio de salud en Colombia.
- Elaborar un tablero de comandos que presente la calificación de las entidades de salud sobre la base de los datos obtenidos de las peticiones PQRD interpuestas ante la Supersalud.

3 Marco referencial

3.1 Marco Teórico

3.1.1 Supersalud

La Supersalud es una entidad del gobierno que se encuentra encargada de la protección de los derechos de los usuarios afiliados al sistema de salud en Colombia de acuerdo con las políticas de control según el decreto 2462, en el cual se estructuró una superintendencia llamada "*protección al usuario*" (Supersalud, Supersalud, 2020).

3.1.2 Entidades Prestadoras de Salud

Son organizaciones encargadas de gestionar la afiliación de usuarios al sistema general de salud y los recaudos económicos establecidos por la ley, los cuales deben realizar los empleadores y trabajadores para poder acceder al servicio de salud (Minsalud, 2021).

Cuando un ciudadano presenta inconformidades relacionadas con la prestación de los servicios de salud debe conocer los trámites y servicios que ofrece la Supersalud para poder interponer las PQRDS por los siguientes canales de atención (Supersalud, Supersalud, 2021).

- Canal web: Este canal permite al usuario interponer PQRDS accediendo a la página web de la Supersalud donde se dispone de un módulo para el registro y consulta del estado de las peticiones.
- Canal telefónico: Este canal permite al usuario interponer PQRDS a través de comunicación telefónica al centro de atención de la Supersalud.

- Personalizado: Este canal permite al usuario interponer PQRDS presencialmente en el centro de atención de la Supersalud.
- Correo: Este canal Permite al usuario interponer PQRDS vía e-mail ante el centro de atención de la Supersalud.

Las PQRDS, son un medio por el cual la población puede ejercer los derechos a cualquier tipo de reclamación, ya sea por inconformidad en la atención o servicios prestados por alguna entidad prestadora de servicios de salud.

3.1.3 Definición y Uso de PRQD en la Supersalud

Las peticiones son el derecho que todo usuario afiliado al servicio de salud en Colombia tiene para presentar las solicitudes ante las respectivas autoridades de control y vigilancia para la resolución de su solicitud, para el caso de la superintendencia de salud se consideran los siguientes tipos de petición.

- Petición de solicitudes de información.
- Petición de consulta.
- Petición preferencial.
- Petición entre autoridades.
- Petición de informes.

Una queja es la expresión de inconformidad que presenta usuario afiliado al servicio de salud ante la Supersalud, ya sea por incumplimiento en la prestación de servicios de salud o vulneración de los derechos sobre los pacientes.

El reclamo es la manifestación que presentan los usuarios por la insatisfacción sobre la atención de los servicios de salud por parte de las entidades, estos reclamos se clasifican de la siguiente manera:

- Reclamo prioritario.
- Reclamo regular.

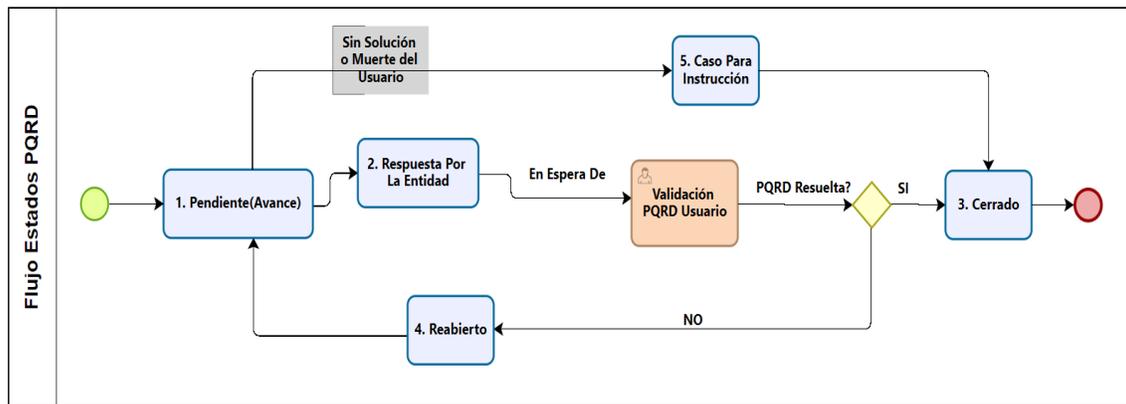
La denuncia es un mecanismo por medio de la cual una persona da a conocer hechos sobre delitos ante las autoridades y por lo tanto deben ser investigados por parte de entidades gubernamentales (Supersalud, Supersalud, 2014).

A continuación, se describe cada uno de los estados de las PQRDS interpuestas ante la Supersalud por los usuarios que se encuentran afiliados al sistema de salud en Colombia, con base a la solicitud número 202131301388951 (Saa, 2021) realizada ante la Supersalud.

- Pendiente: Casos abiertos para gestión del ente vigilado.
- Casos para Instrucción: Se refiere a casos reiterados por los usuarios, los cuales no han tenido respuestas en los términos estipulados por la ley.
- Cerrado: Se refiere a casos cerrados con respuesta por parte del ente vigilado y verificación con el usuario.
- Reabierto: Casos reabiertos por cierre indebido por parte del ente vigilado.
- Respuesta por la entidad: Casos cerrado con respuestas por parte del ente vigilado.

A continuación, se describe el flujo de los estados de las PQRD interpuestas ante la Supersalud, 1. Pendiente, 2. Respuesta Por La entidad, 3. Cerrado, 4. Reabierto y 5. Caso Para Instrucción, (Ver Figura 1).

Figura 1. Diagrama Estados PQRD

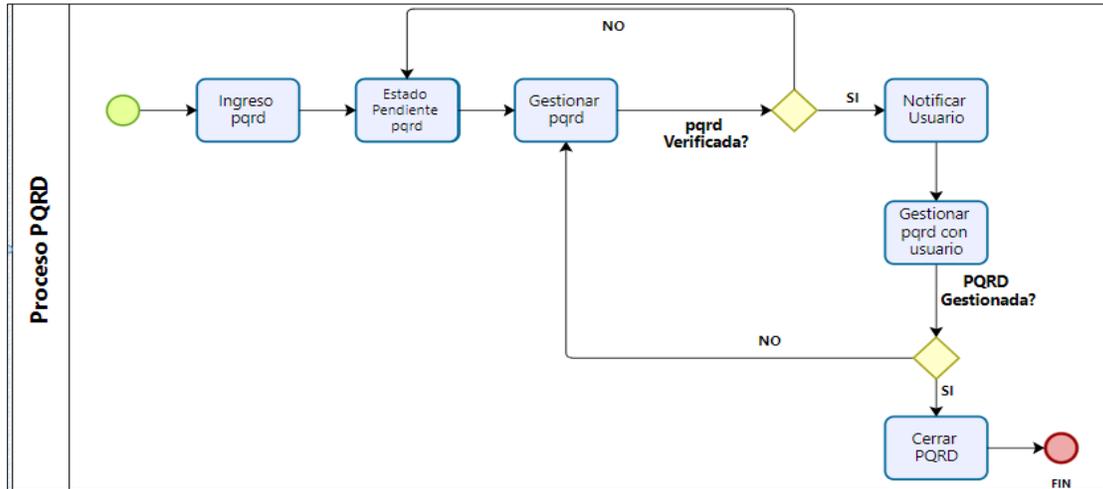


Fuente: Elaboración Propia

Según la Supersalud es importante aclarar que la información de la variable PQRD_ESTADO son cambiantes de acuerdo a la gestión que surta un caso, por ejemplo: un caso se crea en estado Pendiente (Avance) y tan pronto se cierra por parte del vigilador cambia a “Respuesta por la entidad”.

De acuerdo con la descripción anterior de cada estado de las PQRDS, se genera el diagrama de flujo en el cual se indica cada actividad realizada en el proceso de atención de las PQRDS desde su ingreso hasta su cierre (ver figura 2).

Figura 2. Diagrama Flujo General Proceso PQRD



Fuente: Elaboración Propia

3.1.4 Método Ideal de Referencia

El método del ideal de referencia (RIM) fue desarrollado por Cables et al. (Cables, ScienceDirect, 2016), el cual forma parte de los métodos de análisis multicriterio para apoyar el proceso de toma de decisiones. Particularmente, este método (RIM) parte de la identificación del ideal de referencia para cada uno de los criterios que serán implementados para evaluar cada alternativa y se fundamenta en los conceptos siguientes.

- Rango: Es cualquier intervalo o conjunto simple de valores pertenecientes a cualquier universo de discurso.
- Ideal de referencia: Es un intervalo o valores simples que representan la máxima importancia o relevancia en un Rango dado.

Para aplicar el método RIM, se debe seguir los pasos siguientes de acuerdo con el autor (Cables, Information Sciences, 2016).

3.1.4.1 Paso 1 Definición Contexto del Trabajo

En este paso se establecen las condiciones de trabajo para cada criterio y se definen los siguientes aspectos (Cables, RIM-reference ideal method in multicriteria decision making, 2016).

- La gama t_j .
- El ideal de referencia s_j . El ideal de referencia puede ser cualquier conjunto entre el valor mínimo y el valor máximo.
- El peso w_j asociado al criterio

3.1.4.2 Paso 2 Obtener la Matriz X

Se obtiene la matriz de valoración X en correspondencia con los criterios definidos, (ver figura 3).

Figura 3. Fórmula Obtención Matriz X

$$X = \begin{pmatrix} x_{11} & x_{12} & \dots & x_{1n} \\ x_{21} & x_{22} & \dots & x_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ x_{m1} & x_{m2} & \dots & x_{mn} \end{pmatrix}$$

Fuente: Tomado Revista Information Sciences

3.1.4.3 Paso 3 Normalizar la Matriz X

Normalizar la matriz de valoración X con el ideal de referencia donde el ideal de referencia es un conjunto entre el valor mínimo y el valor máximo.

Donde:

$$f(x, t, s) = \begin{cases} 1 & \text{si } x \in [C, D] \\ 1 - \frac{d_{\min}(x, [C, D])}{|A - C|} & \text{si } x \in [A, C] \wedge A \neq C \\ 1 - \frac{d_{\min}(x, [C, D])}{|D - B|} & \text{si } x \in [D, B] \wedge D \neq B \end{cases}$$

$t = [A, B]$, es el rango perteneciente a un universo de discurso

$s = [C, D]$, representa el Ideal de Referencia.

$$x \in [A, B] \quad [C, D] \subset [A, B] \quad d_{\min}(x, [C, D]) = \min(|x - C|, |x - D|)$$

3.1.4.4 Paso 4 Calcular la Matriz Normalizada Ponderada Y

Para calcular la matriz normalizada ponderada y a través de: (ver figura 4).

Figura 4. *Fórmula Calcular Matriz Normalizada*

$$Y' = Y \otimes W \begin{pmatrix} y_{11} \cdot w_1 & y_{12} \cdot w_2 & \dots & y_{1n} \cdot w_n \\ y_{21} \cdot w_1 & y_{22} \cdot w_2 & \dots & y_{2n} \cdot w_n \\ \vdots & \vdots & \ddots & \vdots \\ y_{m1} \cdot w_1 & y_{m2} \cdot w_2 & \dots & y_{mn} \cdot w_n \end{pmatrix}$$

Fuente: Tomado Revista Information Sciences

3.1.4.5 Paso 5 Calcular la Variación a la Referencia Normalizada

Calcular la variación a la referencia normalizada ideal para cada alternativa bajo

la siguiente operación, (ver figura 5).

Figura 5. *Cálculo Referencia Normalizada*

$$I_i^+ = \sqrt{\sum_{j=1}^n (y'_{ij} - w_j)^2}, \quad I_i^- = \sqrt{\sum_{j=1}^n (y'_{ij})^2} \quad \text{and } i = 1, 2, \dots, m, \quad j = 1, 2, \dots, n$$

Fuente: Tomado Revista Information Sciences

3.1.4.6 Paso 6 Calcular el Índice Relativo

Calcula el índice relativo de cada alternativo de la expresión (ver figura 6).

Figura 6. Formula Calcular Índice Relativo

$$R_i = \frac{I_i^-}{I_i^+ + I_i^-} \quad \text{where } 0 < R_i < 1, \quad i = 1, 2, \dots, m$$

Fuente: Tomado Revista Information Sciences

3.1.4.7 Paso 7 Clasificación de Alternativas

El método RIM consiste en realizar la agregación de la información a partir de una referencia.

Como se puede observar, el método RIM utiliza el vector de pesos para enseñar la importancia relativa entre los criterios. Para la determinación del vector de pesos existen diversos métodos, sin embargo, en este proyecto se utilizará el método propuesto por Cables y Lamata (Lamata & y Cables, 2009) por ser de fácil implementación y utilizar una función de comportamiento lineal donde la distancia entre pesos adyacentes es constante.

La siguiente función se implementa para realizar el cálculo del vector de pesos según los autores (Lamata & y Cables, 2009) es.

$$f(x) = a \cdot x + \frac{1}{n} - a \cdot \left(\frac{1+n}{2} \right)$$

Donde:

- n, es la cantidad de elementos del vector de pesos.

- a , es la pendiente de la función lineal.
- $x = 1, 2, \dots, n$

3.1.4.8 Algoritmo

Para obtener una función lineal, se debe seleccionar la pendiente de acuerdo con la siguiente fórmula, (ver figura 7).

Figura 7. Algoritmo Obtención Función Lineal

$$a \in \left] \frac{-2}{n(n-1)}, \frac{2}{n(n-1)} \right[.$$

Fuente: Tomado Revista Information Sciences

3.2 Descripción Herramientas

La gobernabilidad de datos consiste en la capacidad de gestionar la calidad de los datos en las organizaciones ya sean públicas o privadas donde se incluye un consejo de gobierno, procedimientos y planes de acción para realizar la ejecución de los procedimientos establecidos. (PowerData, 2014).

La expresión ETL se refiere al proceso de extracción de datos de diferentes fuentes para realizar transformación aplicando la calidad de datos con la finalidad de realizar el proceso de carga a diferentes sistemas de información para su disponibilidad a sistemas destinos. (Maquilme, 2021).

Un dataset es una colección de datos almacenados en diferentes sistemas de información y archivos digitales como; .csv, .txt, .xlsx entre otros, los cuales contienen valores organizados en columnas y filas de acuerdo a su previa construcción (KeepCoding, 2020).

Los tableros de comandos son herramientas que permiten capturar, organizar y visualizar información de forma gráfica en pantallas tanto local como web, con la finalidad de presentar indicadores a diferentes áreas de las organizaciones para el seguimiento de procesos y tomas de decisiones. (Ciberclick, 2020).

Tableau es una herramienta que permite capturar datos desde diferentes fuentes para la generación de indicadores en tableros de comandos, los cuales se pueden diseñar a través de gráficos interactivos para compartir la información con otros usuarios, esta herramienta se implementa en el proyecto para la visualización de la calificación de las entidades de salud en Colombia (USC, 2021).

Python es un lenguaje de programación multipropósito ya que se implementa para diferentes áreas como machine learning, aplicaciones web, internet de las cosas entre otros, este lenguaje se implementa en el proyecto para aplicar calidad a los datos adquiridos de la plataforma datos abiertos de Colombia y Supersalud (Aula21, 2018).

Una base de datos es una recopilación de información estructurada, la cual está dispuesta en un sistema de almacenamiento para lo operación requerida por los procesos de las áreas en las organizaciones (Oracle, Oracle, 2020).

La expresión php (Hypertx Processor) es un lenguaje interpretado el cual se implementa para el desarrollo web y Apis (Aplicacion Interface Programin), este lenguaje es open source y procesados en servidores web (whatis, 2019).

3.3 Estado del Arte

Se realiza la exploración sobre investigaciones relacionadas con datos abiertos de la Superintendencia de Salud de Colombia frente a la información de las PQRD interpuestas por los diferentes usuarios afiliados a las entidades prestadoras del servicio de salud, para lo cual se identificaron las siguientes investigaciones.

Se identificó cronológicamente que las investigaciones encontradas permiten realizar un seguimiento a la información objeto de este estudio, en la cual se encuentra como relevante una publicación de fuente propia de la Superintendencia de Salud, titulada como *“Evaluación y Seguimiento a Procesos y Procedimientos Misionales y de Apoyo con Desarrollo de Verificación al Procedimiento de Peticiones, Quejas y Reclamos de la Superintendencia Nacional de Salud”* (Superintendencia Nacional de Salud, 2013), esta investigación fue realizada para los meses de marzo y abril del año 2013 en el cual se dimensionó el crecimiento de las PQRDS interpuestas ante la Supersalud, donde en el mes de marzo los usuarios interpusieron 54.254 peticiones y 72.895 para abril, donde se obtuvo un crecimiento de 34.36%, evidenciando un desbordamiento en la atención de las respuesta ente los usuarios para lo cual debieron realizar ajuste al proceso de atención de las peticiones.

Una segunda investigación denominada *“Análisis de las Peticiones, Quejas, Reclamos y Denuncias (PQRD) Interpuestas por patología de cáncer ante la Supersalud 2018”* (Naranjo & Rodríguez, 2019) presenta que las PQRD son los mecanismos por medio de los cuales las personas presentan las inconformidades con el sistema de salud,

donde referenciaron la patología de cáncer con mayor número de PQRDS interpuestas en el año 2018 ante la Supersalud (Supersalud).

En el análisis realizado identificaron que los casos de las PQRDS de tipo cáncer y sexo estaban relacionadas con los números de morbilidad, evidenciando que el mayor nivel de inconformidades se presentó por motivos de restricciones, deficiencia en la atención en el servicio de salud e insatisfacción de los usuarios con la gestión de las autorizaciones.

De lo anterior se concluye que la Supersalud tiene determinado un proceso específico de atención al usuario y como parte de ese proceso se encuentra definido un procedimiento para el trámite de las PQRD basado en el cumplimiento de la normatividad colombiana y de la misma entidad, lo que es indispensable para garantizar el derecho de los usuarios del SGSSS; así mismo permitió identificar que el cáncer es la principal causa de muerte en el país y es considerada una enfermedad de alto costo para el SGSSS y además un evento de interés en salud pública.

Los resultados de la monografía tienen bastante relación con las cifras de incidencia, prevalencia, mortalidad, características de la población con cáncer en el país, de acuerdo a los estudios e informes del Ministerio de Salud, INS, ONS, Instituto Nacional de Cancerología y CAC.

Una tercera investigación denominada “*Caracterización de peticiones, quejas y reclamos del SGSSS en Colombia en 2014-2019: retos y perspectivas*” la cual fue realizada por (Foseca & Martínez, 2020) presenta como objetivo caracterizar las PQRD

presentadas por los usuarios afiliados al servicio de salud durante los años 2014-2019 del cual se obtuvo que durante este periodo se interpusieron 2.847.967 PQRD ante la Supersalud, concluyendo que el 58% correspondieron a mujeres, el 22% a personas mayores de 60 años, el 8.26% a personas en condición de especial o constitucional de vulnerabilidad, un 27% fueron relacionadas a usuarios con diagnóstico de cáncer.

En el presente estudio se identifica la información incluida en el sistema de información de PQRDS de la Supersalud, para el análisis de variables de estudio que categorizan los datos en el marco del SGSSS de Colombia. Se reconoce en el análisis de los datos el comportamiento de las PQRD en el país de acuerdo con características sociodemográficas y geográficas de la atención en salud, siendo de especial análisis la falta de oportunidades en las asignaciones de las consultas médicas por especialista y su comportamiento con base a las variables de estudio, como principal motivo de queja durante el periodo de tiempo analizado.

En el último corte de estudio, enero a diciembre de 2019, el departamento con mayor tasa de PQRD por cada 10.000 afiliados fue Quindío y el vigilado con mayor tasa tanto en régimen contributivo como subsidiado es la EPS Medimás.

Una cuarta investigación denominada “*Propuesta para Mejorar la Oportunidad de Atención de las PQRDS en una Empresa Prestadora de Servicios de Salud*”, realizada por (Jurado & Lievano, 2019), donde presentaron una propuesta para la mejora de los tiempos en la atención de las PQRD en una entidad prestadora de servicios de salud, en la cual recopilaban información del porcentaje de participación de los usuarios insatisfechos

por la prestación de los servicios con la finalidad de establecer estándares para la atención de PQRD.

3.4 Impacto

La implementación del proyecto permitirá obtener la calificación de las entidades prestadoras de los servicios de salud en Colombia, sobre el proceso de atención de las PQRDS interpuestas por los usuarios ante la Supersalud, para lo cual se dispondrá la calificación de manera interactiva en gráficos sobre tableros de comandos, con la finalidad de facilitar a los usuarios la visualización del estado de la atención en los servicios de salud de cada entidad.

3.5 Componente de Innovación

El presente proyecto ofrecerá la centralización de los datos de las PQRD de las entidades de salud de los últimos tres años, donde se innovará con la creación de un tablero de control permitiendo visualizar y analizar los datos cuyo propietario es la Supersalud y adicionalmente se obtiene una calificación del proceso de peticiones PQRD de cada entidad de salud en Colombia por medio de la aplicación del modelo de análisis multicriterio RIM.

4 Metodología

Para la gestión de los datos algunos autores establecen determinadas normas y ciclos de vida del dato que se deben tener en cuenta y en las que nos apoyan para organizarlas en diferentes etapas. Por citar un caso, según plantea el autor AWS Public Sector Blog Team (Blog, 2019), la etapa inicial del ciclo de vida del dato comienza con la ingestión y captura de los datos desde las distintas fuentes disponibles, por lo que ésta será la etapa inicial del proyecto. Por otra parte, se establecen los lineamientos de gobierno de datos en cuanto a metadatos, modelamiento y almacenamiento de los datos en base al documento Data Management (DAMA) (DAMA, 2008) y elaboración del estudio, bosquejo y desarrollo de bases de datos hasta su implementación apoyándose en los autores (Nieto Bernal, 2017).

Para las siguientes etapas de ciclo de vida del dato que menciona el autor AWS Public Sector Blog Team (Team, 2019), se encuentran las etapas de almacenamiento y preparación de los datos recopilados desde todas las diferentes fuentes de origen, ya que se debe realizar un proceso de almacenamiento y tratamiento de los datos capturados. A continuación, se iniciará el proceso de transformación de datos para obtener una mayor calidad y entendimiento de los mismos, en el cual se ajustan mediante la aplicación de diferentes procesos como los ETL y se ingesta en almacenes de datos para la implementación de un modelado multidimensional según lo menciona el autor (Carlos Trujillo, 2013).

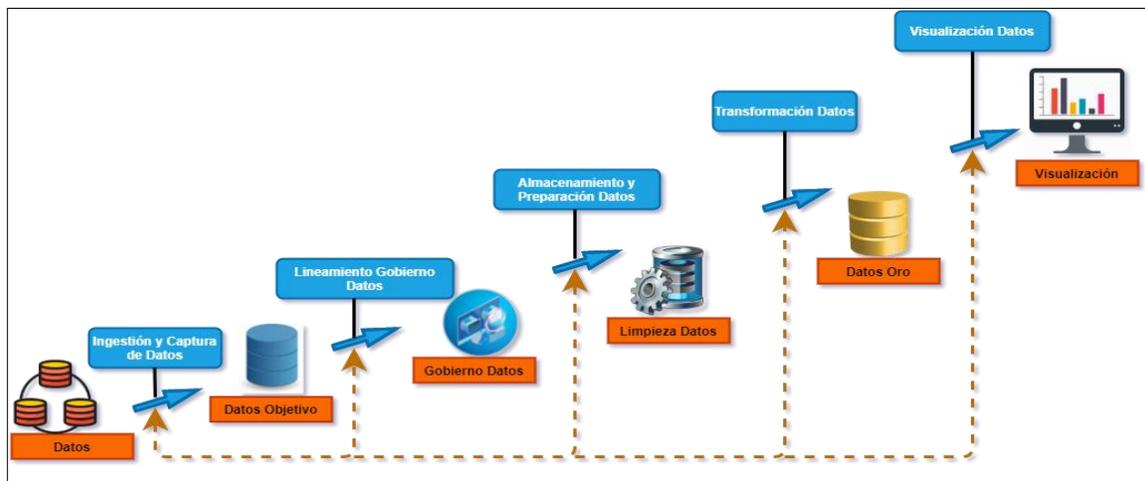
Finalmente se implementa una visualización de todos los datos transformados la cual contendrá filtros y diagramas estadísticos según los autores (Wilke., 2019) y

(Bernard Marr, 2017). Además de las ventajas de obtener información a partir de los datos, se pretende en este proyecto obtener una presentación ordenada, sistemática y de fácil comprensión para los usuarios.

En base a lo expuesto anteriormente, se describen las etapas para el desarrollo del proyecto (ver figura 8).

- Ingestión y captura de datos.
- Lineamiento gobierno datos.
- Almacenamiento y preparación de datos.
- Transformación de datos.
- Visualización de datos.

Figura 8. Etapas a Utilizar en la Metodología



Fuente: Elaboración Propia

A continuación, se describe de forma general cada etapa para el desarrollo de la metodología.

4.1 Ingestión y Capturas de Datos

Consiste en la recolección de todos los datos de las PQRDS interpuestas ante la organización para su ingestión en los sistemas de información para poder realizar los respectivos análisis sobre calidad de datos.

4.2 Lineamientos de Gobierno de Datos

se enfocará en tres aspectos definidos en el framework DAMA.

- Modelado y diseño de datos: Proceso por el cual se realiza el análisis para la representación de los datos de forma precisa en un modelo de datos.
- Almacenamiento de datos y operaciones: Se plantea el diseño e implementación del almacenamiento de los datos y sus operaciones.
- Metadatos: Este proceso incluye acciones como el control, la planificación, la implementación y la definición de flujo de datos.

4.3 Preparación y Almacenamiento de los Datos

En primer lugar, el almacenamiento según (Casas Roma, Nin Guerrero, & Julbe López, 2019) son sistemas de varios tipos donde se guarda y recupera desde pequeños hasta grandes volúmenes de datos de forma eficiente, se incluyen todo tipo de bases de datos, tanto no relacionales como relacionales. En la fase de preparación es fundamental que los datos tengan integridad, debido a que se utilizan para la toma de decisiones estratégicas, las cuales presentan grandes inversiones y son claves para el posicionamiento de las organizaciones en su nicho de mercado, ya que los datos son procedentes de varias fuentes datos, existiendo una alta posibilidad de enfrentarse a errores en caso de no realizar el proceso de limpieza de datos con antelación a la carga en el repositorio (Carlos Trujillo, 2013).

4.4 Transformación de Datos

Esta etapa consiste en aplicar procesos de reducción de dimensionalidad, eliminación de atributos que son insignificantes, redundantes y que no son relevantes para el negocio con base a un previo análisis realizado sobre los datos, los cuales deberán ser cargados posteriormente a un almacén de datos para posteriormente realizar los respectivos análisis para la toma de decisiones estratégicas en las empresas. Estos sistemas son únicamente de consulta y están enfocados a extraer conocimiento de los datos históricos (Carlos Trujillo, 2013). Las herramientas más extendidas para realizar el estudio de estos datos históricos, son OLAP (procesamiento analítico en línea). En esta etapa es necesaria la implementación de herramientas que permitan la transformación de datos como los procesos ETL con los cuales se aplica la calidad de datos y se realiza la disponibilidad de la información en los sistemas finales como tableros de control o aplicaciones front-end (Carlos Trujillo, 2013).

4.5 Visualización Datos

Esta etapa consiste según el autor Claus Wilke (Wilke., 2019) en brindar pautas sobre cómo convertir grandes conjuntos de datos en cifras claras y convincentes. Identificar qué tipo de visualización es mejor para la historia que se quiere contar. Se hacen figuras informativas que sean visualmente agradables. Se exploran los conceptos básicos del color como herramienta para resaltar, distinguir o representar un valor y comprender la importancia de la codificación redundante para asegurarse de proporcionar información clave de múltiples formas.

5 Desarrollo de la propuesta

A continuación, se detalla el desarrollo de la propuesta de acuerdo a la metodología implementada anteriormente y así llevar a cabo la ejecución del proceso.

5.1 Ingestión y Captura de Datos

Para el proyecto se consideró utilizar los datos de las PQRD interpuestas ante la Supersalud por los usuarios afiliados al sistema de salud en Colombia, esta información es entregada por medio de un archivo .csv delimitado por comas (,).

5.1.1 Orígenes de los Datos

La fuente principal de información para el estudio del presente proyecto del análisis de las PQRD se basa en los datos suministrados por esta entidad Supersalud los cuales algunos son publicados por su propia plataforma y otros en datos abiertos del gobierno nacional. Se utilizó como información complementaria datos del portal DANE para la validación de ubicación geográfica de las PQRD y adicionar posicionamiento de geolocalización de su origen (DANE, 2021), a continuación, se describen las fuentes de obtención.

- Base de datos pqrD 2020
 - ✓ URL: <https://www.datos.gov.co/Salud-y-Protecci-n-Social/Base-de-Datos-PQRD-2020/n8qp-zcp9>
 - ✓ Propietario de datos: Superintendencia Nacional de Salud
 - ✓ Sector: Salud y Protección Social
 - ✓ Fecha Creación:24/06/2021
 - ✓ Total, Datos: 851K

- ✓ Formato: .CVS, XML, JSON, API XLS.
- Base de datos pprd 2018
 - ✓ URL: <https://www.supersalud.gov.co/es-co/Paginas/Oficina de Planeación/Datos-abiertos.aspx>
 - ✓ Propietario de datos: Superintendencia Nacional de Salud
 - ✓ Sector: Salud y Protección Social
 - ✓ Fecha Creación: 21/06/2021
 - ✓ Total, Datos: 827959
 - ✓ Formato: .CVS, XML, JSON, API XLS.
- Base de Datos Divipola
 - ✓ URL: <https://geoportal.dane.gov.co/geovisores/territorio/consulta-divipola-division-politico-administrativa-de-colombia/>
 - ✓ Propietario de datos: Dirección Administrativa Nacional de Estadísticas DANE
 - ✓ Sector: Estadísticas y de geografía
 - ✓ Fecha Actualización: 30/09/2021
 - ✓ Total, Datos: 7844
 - ✓ Formato: .CVS, XLS.

5.1.2 Fase Exploratoria de Datos

Una vez identificadas las fuentes de datos se procede a capturar la información desde las fuentes descritas anteriormente, por medio de la herramienta Google Drive para su almacenamiento transitorio de los sets de datos en formato CSV, para ejecutar una fase

inicial exploratoria de los datos, por medio de Google Colab se visualiza la Carga Dataset (ver figura 9 y 10).

Figura 9. Carga Dataset Google Colab



Fuente: Elaboración Propia

Figura 10. Exploración Datos Google Colab



Fuente: Elaboración Propia

Se realiza la visualización estructural del origen del dataset con la finalidad de entender la tipología de los datos como: tamaño total, número de columnas, total filas,

posibles problemas con los datos, campos nulos y planteamiento de las posibles tablas para la conversión a un modelo relacional ver figura (11).

Figura 11. Exploración Tipo de Datos Dataset

```
df_PQRD2020.info()
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 850931 entries, 0 to 850930
Data columns (total 47 columns):
# Column Non-Null Count Dtype
---
0 AÑO 850931 non-null object
1 MES 850931 non-null int64
2 TRIM 850931 non-null object
3 FECHA_CREACION 850931 non-null object
4 PQR_CANAL 850931 non-null object
5 PQR_TIPOPETICION 850931 non-null object
6 PQR_CLASE_SNS 850931 non-null object
7 PET_TIPOPER 850931 non-null object
8 PET_COD_DEPTO 850931 non-null int64
9 PET_DPTO 850931 non-null object
10 PET_COD_MPIO 850931 non-null object
11 PET_MPIO 850931 non-null object
12 AFEC_TIPOPER 850931 non-null object
13 AFEC_PARENTESCO 850931 non-null object
14 AFEC_GENERO 850931 non-null object
15 AFEC_EDADR 850931 non-null object
16 AFEC_EDUC 850931 non-null object
17 AFEC_REGAFILIACION 850931 non-null object
18 AFEC_GETNICO 850931 non-null object
19 AFEC_POBESPECIAL 850931 non-null object
```

Fuente: Elaboración Propia

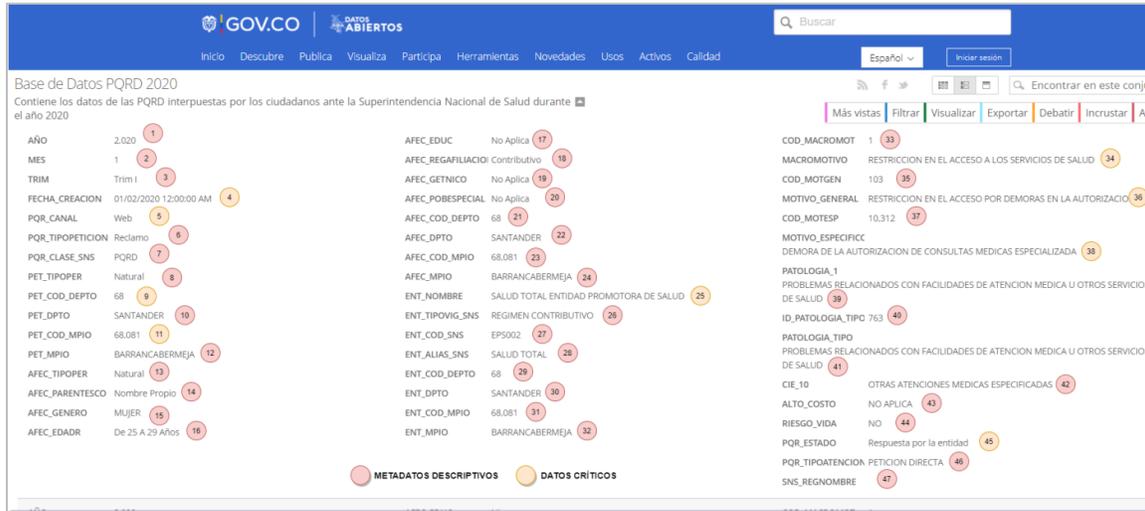
5.1.2.1 Diccionario de Datos Origen

El diccionario de datos origen de la base de datos pqr2018 y pqr2020 es proporcionado por la Supersalud, el cual se toma como referencia para la construcción de las tablas donde se almacenan los registros de las PQRD interpuestas ante la Supersalud, este diccionario de datos se encuentra adjunto en el [Anexo-1](#).

5.1.3 Metadatos

Se describen los metadatos de tipo descriptivo identificados en el formato de datos PQRD suministrados por la Supersalud desde su plataforma web y desde el sitio web datos abiertos de Colombia, (ver figura 12).

Figura 12. Descripción Metadatos



Fuente: Elaboración Propia

Se realiza la descripción detallada de los metadatos con su respectiva numeración y los datos críticos para la organización ver en [Anexo-2](#)

A continuación, se listan los datos críticos de las PQRD, (ver tabla 1)

Tabla 1. Datos Críticos PQRD.

Campo	Campo
4	FECHA_CREACION
5	PQR_CANAL
9	PET_COD_DEPTO
11	PET_COD_MPIO
25	ENT_NOMBRE
34	MACROMOTIVO
36	MOTIVO_GENERAL
38	MOTIVO_ESPECIFICO
45	PQR_ESTADO

Fuente: Elaboración Propia

5.1.4 Implementación de Ambientes

Para la ejecución de todas las etapas definidas en el presente proyecto se definieron e implementaron dos ambientes, los cuales permiten tener un mejor orden al poder ejecutar pruebas y procesos de limpieza sobre los datos de forma segura sin afectar la integridad o disponibilidad de los datos que estén siendo consumidos por los usuarios finales, por esto, se definieron el ambiente Stage y Producción.

5.1.4.1 Ambiente Stage

Este ambiente se crea para realizar el análisis y ajuste de los datos de acuerdo con el hallazgo de errores encontrados sobre la estructura de la data. Tras realizar el respectivo análisis, en este ambiente se podrán ajustar problemas en la calidad, formato, tipo de datos, etc., sin afectar los datos que utilizan de los usuarios finales.

5.1.4.2 Ambiente Producción

En este ambiente se almacenan los datos previamente preparados desde el ambiente Stage para disponibilizar directamente o aplicar procesos de transformación que generalmente son luego visualizados en tableros de comandos.

A continuación, se describen los ambientes implementados para la ingestión y almacenamiento de los datos (ver tabla 2).

Tabla 2. Ambientes creados para el proyecto.

Ambiente	Descripción
Stage	<p>En este ambiente se ejecutan procesos de limpieza, carga y pruebas de información.</p> <p>Entorno: Cloud.</p> <p>Plataforma utilizada: AWS.</p>
Producción	<p>SGBD: MySQL 8.0.</p> <p>Este ambiente está destinado para la información final ya transformada para disponerla a los consumidores.</p> <p>Entorno: Cloud.</p> <p>Plataforma utilizada: Microsoft Azure</p> <p>SGBD: MySQL 8.0.</p>

Fuente: Elaboración Propia

5.1.5 Almacenamiento de Datos Stage

Una vez se finaliza la exploración de los datos, se procede a la ingesta y almacenamiento de los datos crudos, o sea, sin ningún cambio, implementando una base de datos MySQL en la nube Amazon Web Service AWS. Este proceso se ejecutó de la siguiente forma.

5.1.5.1 Creación Base de Datos

Se crea una cuenta en la nube AWS la cual se implementa una nueva instancia del SGBD MySQL para la ingestión de los datos en la base de datos desde las tres fuentes identificadas.

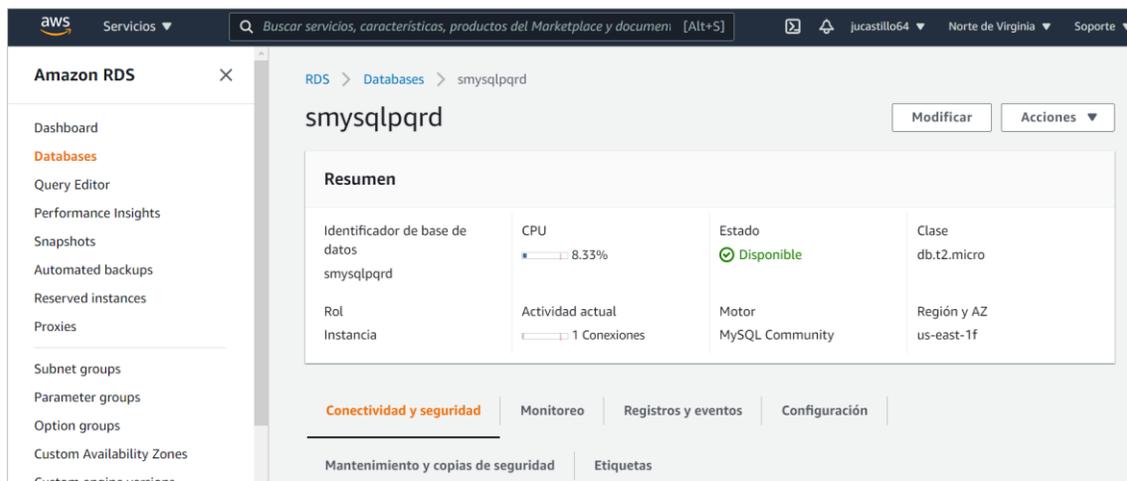
Características de la base de datos:

- ✓ Nombre: smysqlpqrd
- ✓ Disco duro: 20GB con escalado de almacenamiento

- ✓ Versión: 8.0.23 Community
- ✓ vCPU: 1
- ✓ Memoria RAM: 1GB
- ✓ Protección contra eliminación: Si
- ✓ Generación de log: Si
- ✓ Monitoreo y backups: Si

A continuación, se describe la instancia creada para la ingestión y almacenamiento de los datos en el ambiente Stage (ver figura 13).

Figura 13. Visualización Ambiente Stage

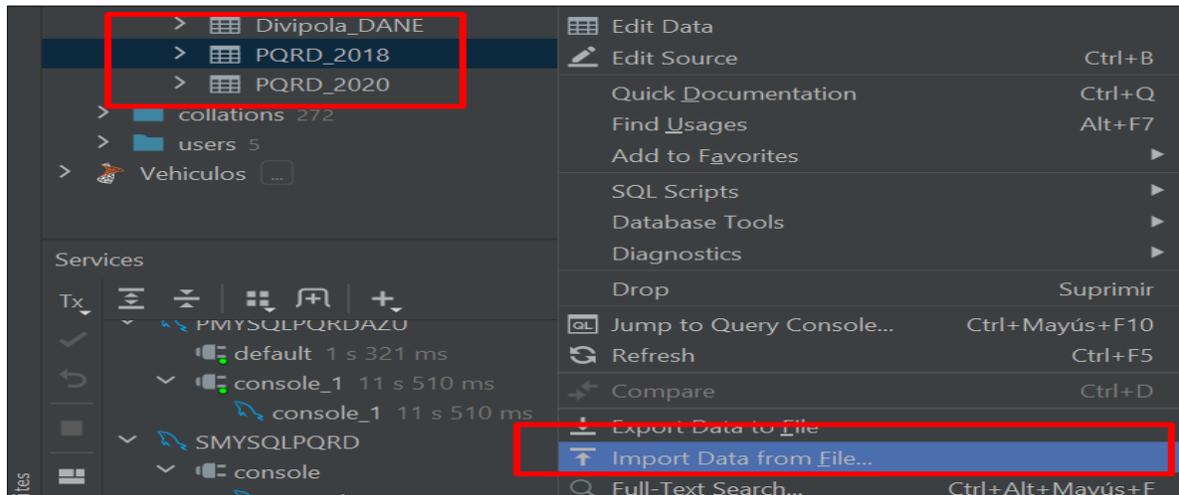


Fuente: Elaboración Propia

5.1.5.2 Importación de Datos

Por medio de la herramienta DataGrip versión 2020 se importan las tres fuentes de datos cada una en archivo formato CSV hacia las tablas: PQRD_2020, PQRD_2018 y Divipola_DANE por medio de la función de importación (ver figura 14).

Figura 14. Importación Data PQRD

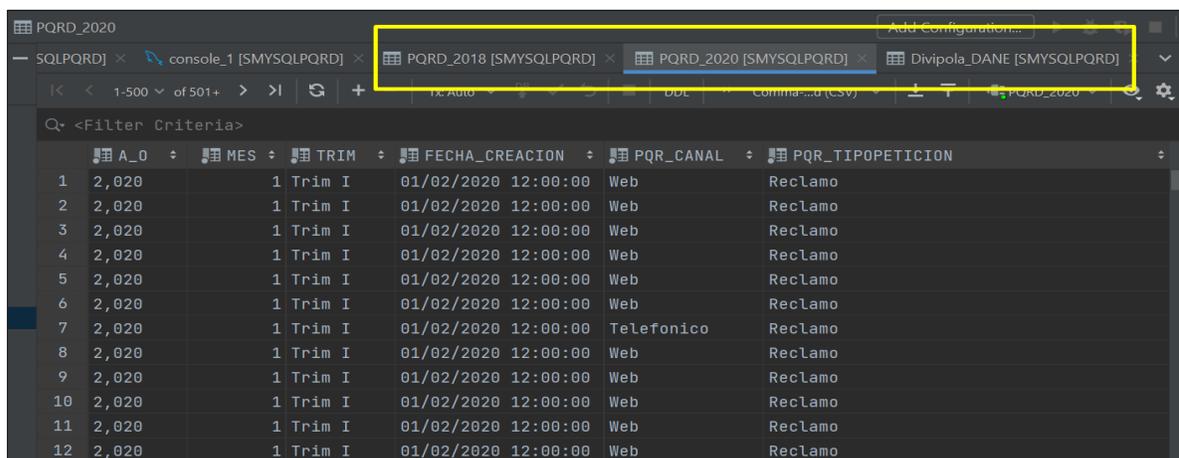


Fuente: Elaboración Propia

5.1.5.3 Muestra Registros Ingestados

Una vez importado los datos a cada una de las tres tablas, se ejecuta una consulta de los datos ingresados para verificar la ingestión (ver figura 15).

Figura 15. Muestra Datos PQRD

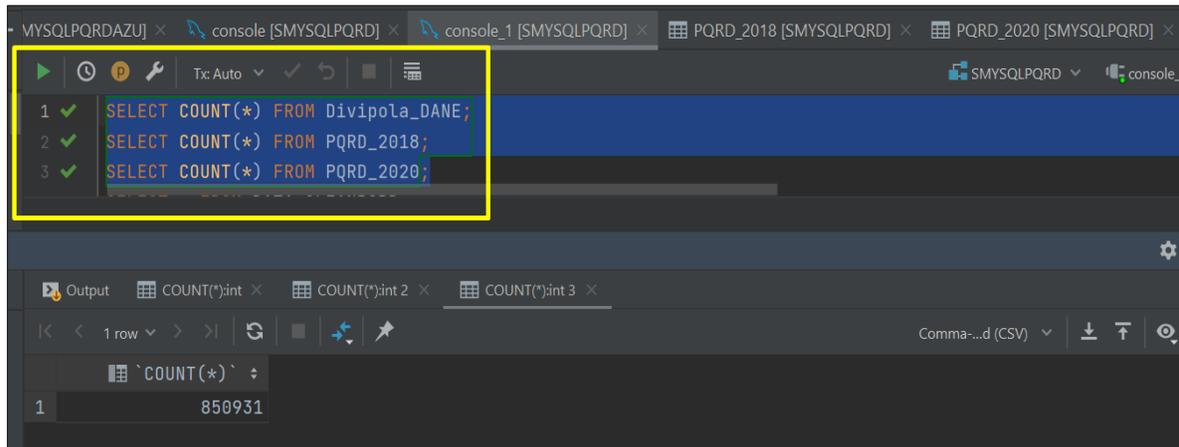


Fuente: Elaboración Propia

A continuación, se presenta el total de datos ingestados en la base de datos smysqlpqr (ver figura 16).

- ✓ Tabla PQRD_2020: 850931
- ✓ Tabla PQRD_2018: 828087
- ✓ Tabla Divipola_DANE: 7809

Figura 16. Conteo Registros Ingestados



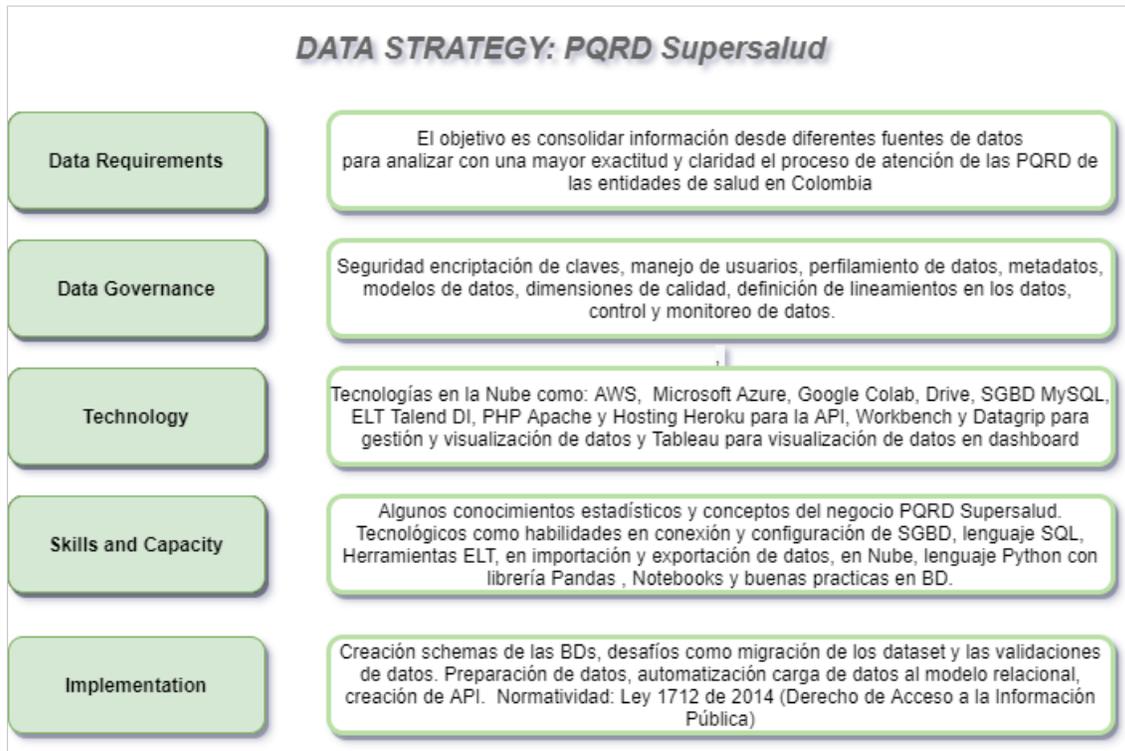
Fuente: Elaboración Propia

5.2 Definición de los Lineamientos del Gobierno de Datos

Para el caso del presente proyecto se definen los lineamientos para la aplicación de gobierno de datos en cuanto a los metadatos, modelamiento y almacenamiento de los datos de las PQRD, a partir de los cuales se toma de referente para ejecutar el análisis, diseño de los modelos y desarrollo de la base de datos tanto para los datos crudos como para los datos transformados siguiendo las prácticas de los autores (Nieto Bernal, 2017) y (Cuadra, 2014).

Se describe la estrategia de datos Data Strategy definida y trabajada a lo largo del presente proyecto sobre el proceso de atención a las PQRD en Colombia donde se nombran las habilidades necesarias, tecnologías, los requerimientos, implementación y gobierno de datos (ver figura 17).

Figura 17. Data Strategy Proyecto PQRD



Fuente: Elaboración Propia

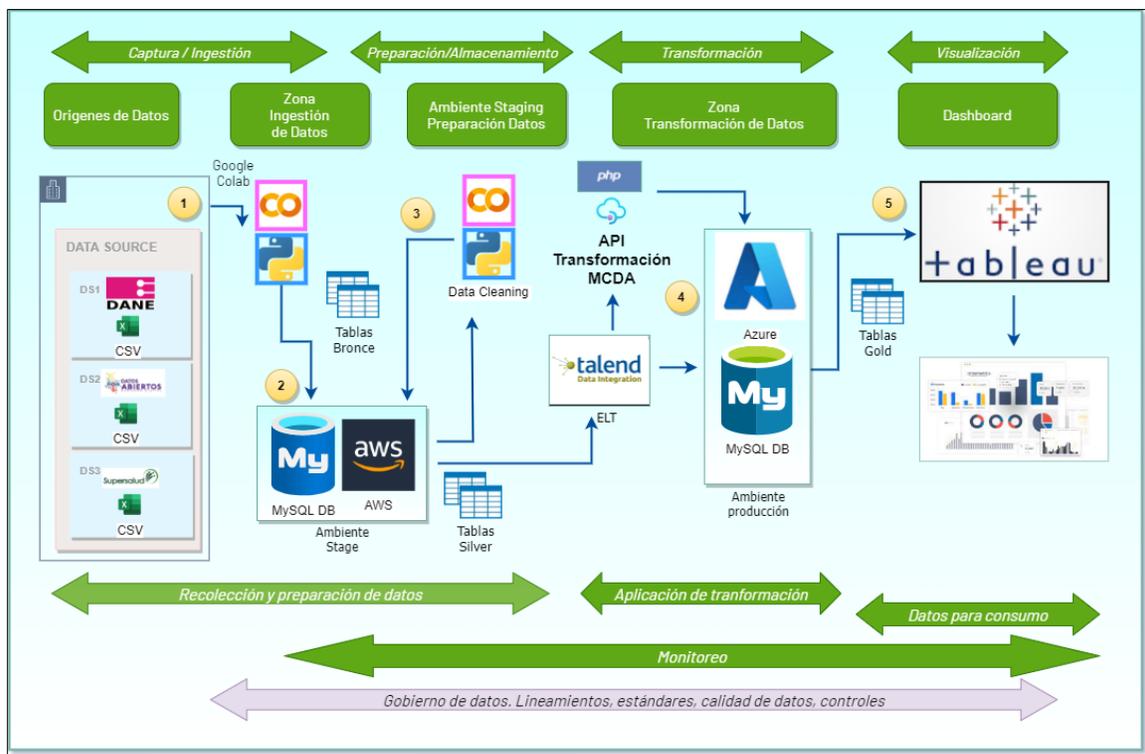
5.2.1 Arquitectura Tecnológica Física

A continuación, se describe la arquitectura tecnológica para la implementación del proyecto (ver figura18).

- Origen de los datos (numeral 1): Se especifica los orígenes de los datos obtenidos como los datos abiertos, Supersalud y DANE.
- Zona ingestión de datos (numeral 2): Se realiza la ingestión de los datos obtenidos en herramientas como Google Colab para el tratamiento de calidad de los datos.
- Zona preparación de datos (numeral 3): Se realiza el proceso de limpieza de los datos ingestados.

- Zona transformación de los datos (numeral 4): Se realiza el proceso de transformación de la estructura de los datos.
- Zona visualización de los datos (numeral 5): Se realiza el proceso de disponibilidad de la información en tableros de comando para la respectiva visualización.

Figura 18. Diagrama Arquitectura Tecnológica

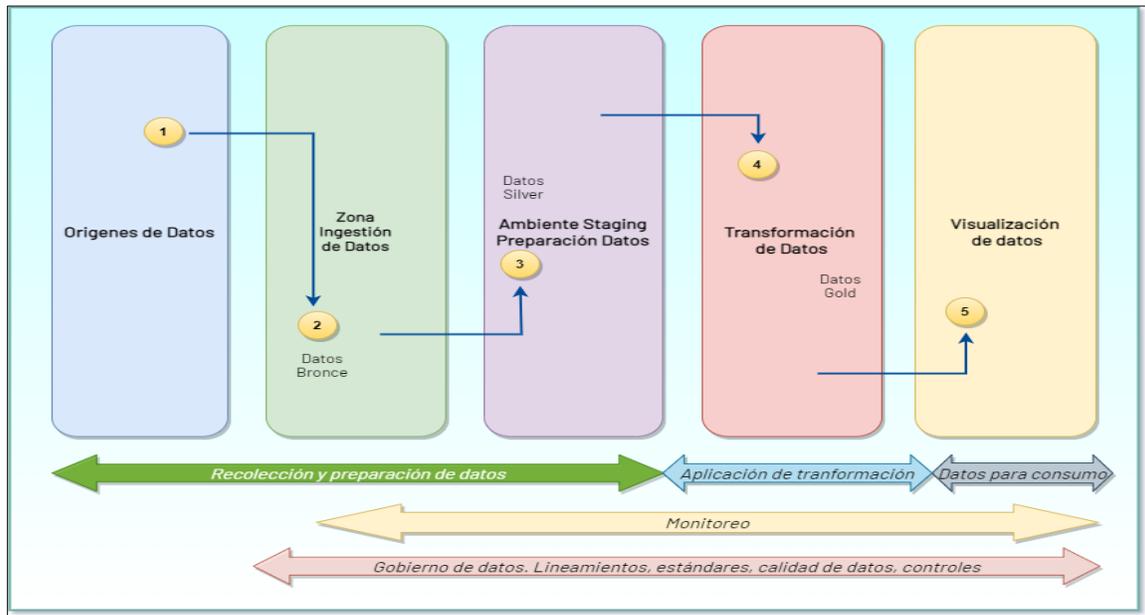


Fuente: Elaboración Propia

5.2.2 Diagrama Arquitectura Lógica

En el siguiente diagrama se representa el flujo lógico del proceso de tratamiento de los datos desde su captura desde los sistemas origen hasta su visualización en los tableros de comandos (ver figura 19).

Figura 19. Diagrama Arquitectura Lógica

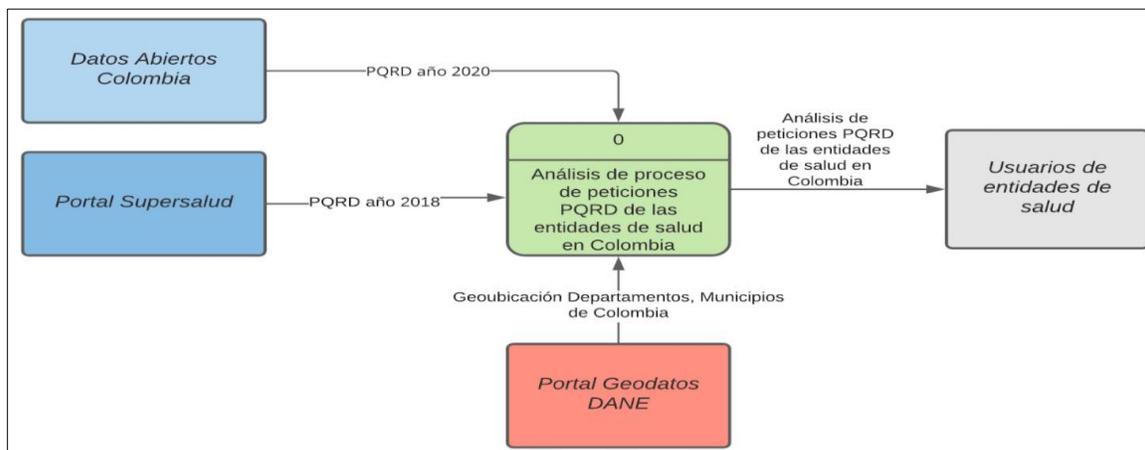


Fuente: Elaboración Propia

5.2.3 Flujo de Datos de Contexto

En el siguiente diagrama se visualiza a modo general el proceso que tienen las PQRD desde su captura del sistema origen y su tratamiento de datos hasta la disponibilidad para la visualización de los datos ante la entidad solicitante (ver figura 20).

Figura 20. Diagrama Flujo Contexto



Fuente: Elaboración Propia

5.2.4 Flujo de Datos Lógico

Se describe el paso a paso del proceso de transformación que tienen los datos desde su captura de los sistemas origen hasta su disponibilidad para la visualización en tableros de comandos, ver [Anexo-4](#).

A continuación, se describe el proceso de forma física de la captura, ingestión, transformación y disponibilidad de la información para la visualización de la calificación de las entidades prestadoras de servicios de salud ver [Anexo-5](#).

Se implementan los modelos conceptual, lógico y físico para el desarrollo de la base de datos para el almacenamiento de la información de las PQRD.

5.2.4.1 Modelo Conceptual

En este modelo se identifican las relaciones de alto nivel entre las diferentes entidades de la base de datos en las cuales no se incluye llaves primarias ni atributos ver [Anexo-6](#).

5.2.4.2 Modelo Lógico

Para este modelo se identifican las relaciones entre las entidades, atributos, llaves primarias y llaves foráneas (Información, 2018), ver [Anexo-7](#).

Se identifican las estructuras de las tablas como son: nombres, atributos tipos de datos, restricciones, llaves principales y foráneas y relaciones que contiene la base de datos (Información, 2018), ver [Anexo-8](#).

5.2.5 Lineamientos y Estandarización de Datos

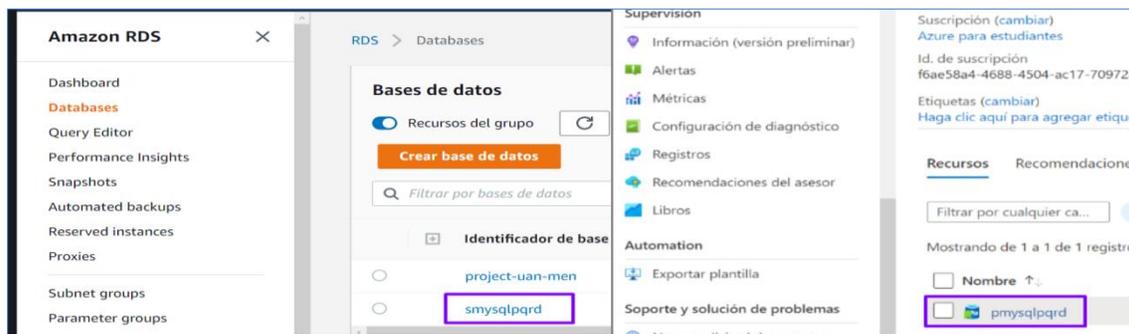
Para la realización y ejecución de las diferentes etapas del proyecto se

establecieron lineamientos para lograr tener y garantizar un lenguaje común en el proyecto y aplicar buenas prácticas en las bases de datos ver [Anexo-3](#).

A continuación, se muestran las evidencias de la aplicación de los principales lineamientos establecidos para el presente proyecto.

Evidencia de creación y ambientes, izquierda AWS (Stage) y derecha Microsoft Azure (Producción) (ver figura 21).

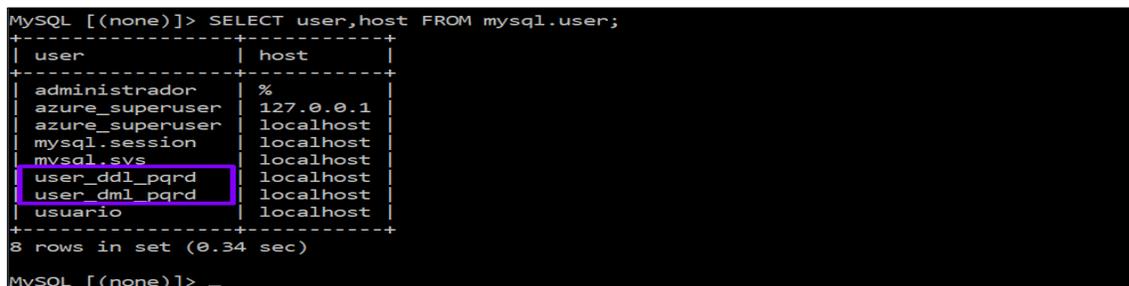
Figura 21. Ambientes AWS y Microsoft Azure



Fuente: Elaboración Propia

A continuación, se evidencia la creación de usuarios para interactuar con la base de datos productiva (ver figura 22).

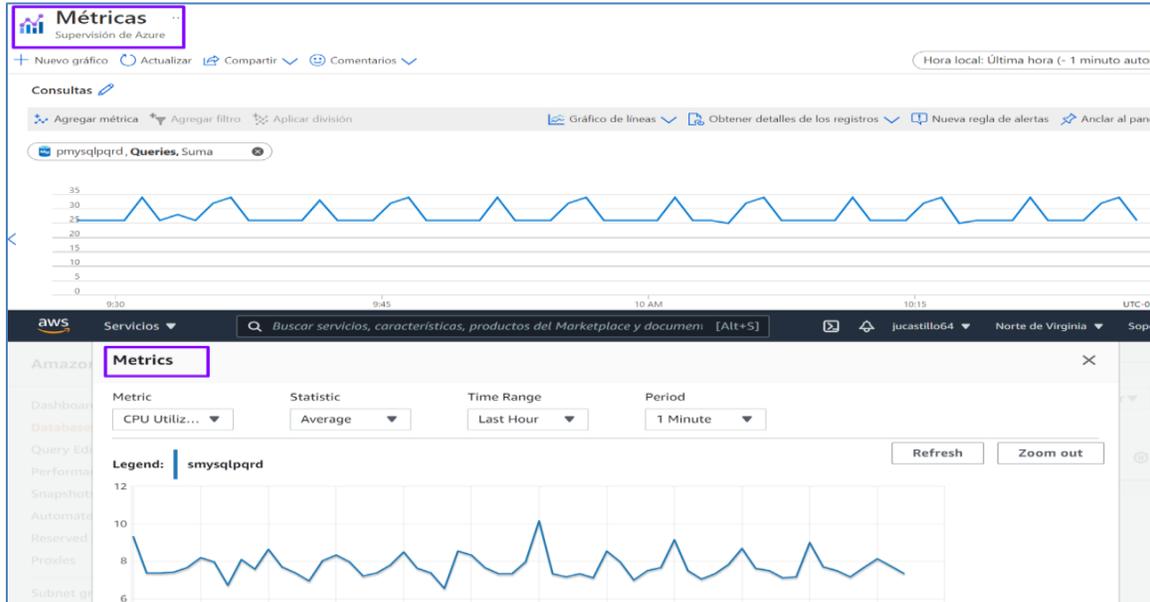
Figura 22. Creación Usuarios Base de Datos



Fuente: Elaboración Propia

Se realiza la parametrización de métricas para el monitoreo del procesamiento de los datos en las plataformas AWS Amazon y Microsoft Azure (ver figura 23).

Figura 23. Monitoreo Procesamiento Datos



Fuente: Elaboración Propia

5.2.5.1 Glosario de Términos

Se identificaron los siguientes términos o siglas utilizadas a lo largo del proyecto para un mejor entendimiento del lector (ver tabla 3).

Tabla 3. Glosario de Términos.

Problema	Columna
BD	Base de Datos
ELT	Extract Load Transformation
DML	Data Manipulation Language
DDL	Data Definition Language
API	Application Programming Interface
FK	Foering Key
PK	Primary Key
SQL	Structured Query Language
SGBD	Sistema Gestor Base de Datos
AWS	Amazon Web Services
MCDA	MultiCriteria Decision Analysis

Fuente: Elaboración Propia

5.3 Preparación y Almacenamiento de Datos

En el presente proyecto, para esta etapa se aplican procesos de almacenamiento y preparación de los datos de las pqr obtenidos en la etapa anterior aplicando tratamiento de datos con la ayuda de herramientas como Google Colab.

5.3.1 Descripción Problemas de Calidad de Datos.

Se ejecutó un análisis de perfilamiento de datos donde se encontraron los siguientes problemas, (ver tabla 4).

Tabla 4. Problemas Detectados.

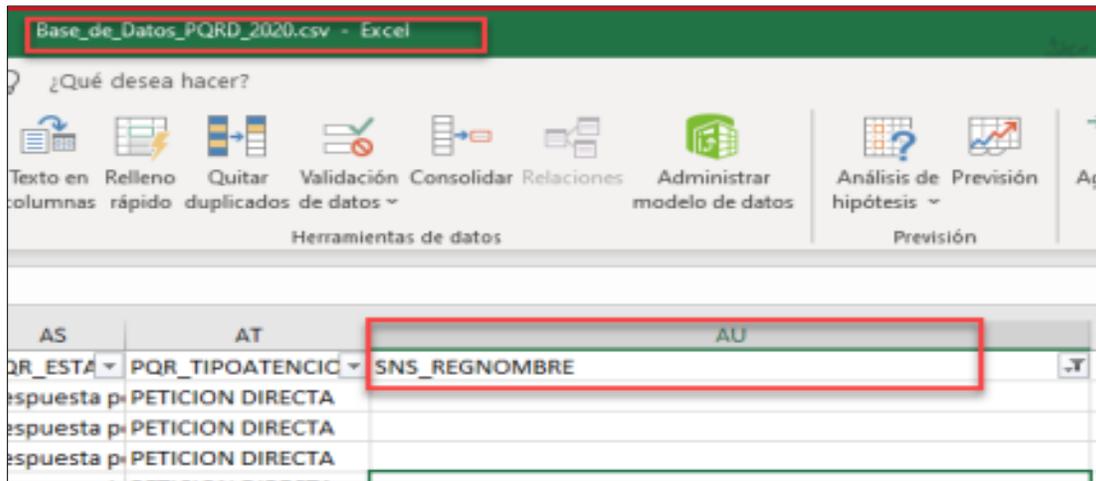
Problema	Columna
Campos nulos	SNS_REGNOMBRE, PET_DPTO
Duplicidad (nombres de entidades)	ENTIDAD_NOMBRE
Tipo de datos inconsistentes	COD_MOTGEN, FECHA_CREACION, AÑO, PET_CODIG_MPIO, AFECTADO_CODIG_MPIO y ENT_COD_MPIO.
Valor máximo y mínimo	PET_COD_DEPTO

Fuente: Elaboración Propia

- Conteo de Datos Nulos

Se realiza el conteo de datos sobre el campo SNS_REGNOMBRE donde se evidencian 736303 datos nulos de un total de 1678890, lo cual equivale a un 43.85 % del total de los datos para el campo SNS_REGNOMBRE, (ver figura 24).

Figura 24. Valores Null Campo SNS_REGNOMBRE



Fuente: Elaboración Propia

Para el campo PET_DPTO, se hallaron 62564 registros sin datos equivalente al 3.72% del total de los datos, valores errados campo PET_DPTO, (ver figura 25).

Figura 25. Valores Errados Campo PET_DPTO

The image shows a screenshot of a data analysis tool displaying the results of a 'value_counts()' operation on the 'PET_DPTO' field. The table lists various departments and their corresponding counts. Two rows are highlighted with red boxes: '0' with a count of 60940 and '(Sin información)' with a count of 1624.

Departamento	Cantidad
BOGOTÁ D.C.	257656
ANTIOQUIA	154718
VALLE DEL CAUCA	117157
0	60940
SANTANDER	47774
ARMUCA	1000
CAQUETÁ	1660
(Sin información)	1624
PUTUMAYO	1298

Fuente: Elaboración Propia

- Tipo Formato

Se realiza revisión del campo COD_MOTGEN donde se hallan datos tipo decimales, los cuales no son acordes al tipo de dato entero del campo, para este caso la cantidad de datos que presentan esta incidencia son 13848, lo cual equivale 0.82% del

total de los datos del campo COD_MOTGEN, valores errados campo COD_MOTGEN, (ver figura 26).

Figura 26. Valores Errados Campo DOC_MOTGEN

MACROM	COD_MOTGEN	MOTIVO_GENERAL
PETICIONES,	1,001	FALTA DE CONTRATACION
NO RECONO	502	NCUMPLIMIENTO DE LAS PRESTACIONES ECON
NO RECONO	502	NCUMPLIMIENTO DE LAS PRESTACIONES ECON
NO RECONO	502	NCUMPLIMIENTO DE LAS PRESTACIONES ECON
NO RECONO	502	NCUMPLIMIENTO DE LAS PRESTACIONES ECON
PETICIONES,	1,004	NCUMPLIMIENTO PARA LA PRESTACION DE LAS
NO RECONO	502	NCUMPLIMIENTO DE LAS PRESTACIONES ECON
NO RECONO	502	NCUMPLIMIENTO DE LAS PRESTACIONES ECON
PETICIONES,	1,001	FALTA DE CONTRATACION
NO RECONO	502	NCUMPLIMIENTO DE LAS PRESTACIONES ECON
PETICIONES,	1,001	FALTA DE CONTRATACION
NO RECONO	502	NCUMPLIMIENTO DE LAS PRESTACIONES ECON
PETICIONES,	1,001	FALTA DE CONTRATACION
NO RECONO	502	NCUMPLIMIENTO DE LAS PRESTACIONES ECON
PETICIONES,	1,001	FALTA DE CONTRATACION
NO RECONO	502	NCUMPLIMIENTO DE LAS PRESTACIONES ECON
PETICIONES,	1,001	FALTA DE CONTRATACION
NO RECONO	502	NCUMPLIMIENTO DE LAS PRESTACIONES ECON
PETICIONES,	1,001	FALTA DE CONTRATACION

Fuente: Elaboración Propia

Igualmente se evidencia que los campos FECHA_CREACION y ANIO también tienen problemas de Tipo formato definidos como tipo de dato string, tanto para el campo FECHA_CREACION como ANIO se presenta este problema en 850931 registros lo cual equivale al 50.68% del total de datos, campo fecha, (ver figura 27).

Figura 27. Tipo de Dato Campo FECHA_CREACION

#	Column	Non-Null Count	Dtype
0	AÑO	850931 non-null	object
1	MES	850931 non-null	int64
2	TRIM	850931 non-null	object
3	FECHA_CREACION	850931 non-null	object

Fuente: Elaboración

- Valores Máximos y Mínimos

Se encuentran valores atípicos del campo PET_COD_DEPTO donde, se evidencia que el código del departamento de Antioquia presenta un número de código que sobrepasa el límite estipulado por el DANE, teniendo en cuenta que los códigos de los

departamentos en Colombia de acuerdo con Divipola, tienen un valor máximo de 99, para este caso, se evidencian un total de 32526 datos, para un equivalente del 1.93% del total de los datos de este campo, (ver figura 28).

Figura 28. Valores Máximos y Mínimos

PTO	E	F	G	H	I	J	K	L	M	N
	ACIC - PQR_CANAL	PQR_TIPC	PQR_CLAS	PET_TIPOI	PET_COD_DEPTO	PET_DPTC	PET_COD	PET_MPIC	AFEC_TIPC	AFEC_PAIS
1/2020	Web	Reclamo	PQRD	Natural	656656465465	ANTIOQUIA	5001	MEDELLÍN	Natural	Nombre Proj
1/2020	Personalizado	Reclamo	PQRD	Natural	656656465465	ANTIOQUIA	5001	MEDELLÍN	Natural	Nombre Proj
1/2020	Telefonico	Reclamo	PQRD	Natural	656656465465	ANTIOQUIA	5088	BELLO	Natural	Nombre Proj
1/2020	Web	Reclamo	PQRD	Natural	656656465465	ANTIOQUIA	5088	BELLO	Natural	Nombre Proj
1/2020	Personalizado	Reclamo	PQRD	Natural	656656465465	ANTIOQUIA	5001	MEDELLÍN	Natural	Madre
1/2020	Telefonico	Reclamo	PQRD	Natural	656656465465	ANTIOQUIA	5001	MEDELLÍN	Natural	Nombre Proj
1/2020	Telefonico	Reclamo	PQRD	Natural	656656465465	ANTIOQUIA	5001	MEDELLÍN	Natural	Nombre Proj

Fuente: Elaboración Propia

- Duplicidad

Se identifica que existen algunas duplicidades en los datos del campo ENTIDAD_NOMBRE, donde se evidencia que se ingresaron entidades con nombres muy similares, pero en alguna parte del texto se diferencia de la otra pero que finalmente tiene que ver con la misma entidad. Para lo cual se presentan seis entidades afectando a 47,540 registros equivalente al 2.83% del total de los datos, (ver figura 29).

Figura 29. Completitud Datos

NUEVA EPS
Nueva EPS S.A
PIJAOS SALUD EPSI
PIJAOS SALUD EPSI

Fuente: Elaboración Propia

5.3.1.1 Ajustes Ejecutados

Por medio de la herramienta Google Colab, se procede a realizar los ajustes de los problemas de calidad presentados en los datos.

- Valores Nulos

Se definió asignar por defecto el texto ‘No definido’ para todos los campos con datos nulos en PET_DPTO y SNS_REGNOMBRE como se muestra a continuación, (ver figura 30).

Figura 30. Visualización Valores Nulos

```

✓ [21] # Nulos
0 s df_PQRD2020['SNS_REGNOMBRE']=df_PQRD2020['SNS_REGNOMBRE'].fillna('No definido')
df_PQRD2020['PET_DPTO']=df_PQRD2020['PET_DPTO'].fillna('No definido')

✓ [22] df_PQRD2020.SNS_REGNOMBRE
0 s
0      No definido
1      No definido
2      No definido
3      No definido
4      No definido
...
850926  No definido
850927  No definido
850928  No definido
850929  No definido
850930  No definido
Name: SNS_REGNOMBRE, Length: 850931, dtype: object
    
```

Fuente: Elaboración Propia

- Tipo Formato

Se ajustaron todos los tipos de datos a las columnas identificadas con un formato inadecuado para el campo fecha y año como se muestra a continuación, (ver figura 31).

Figura 31. Ajuste Tipo Formato Datos Fecha

```
# Formatos de datos
df_PQRD2020['AÑO']=df_PQRD2020['AÑO'].replace('2,020', '2020')
df_PQRD2020['PET_COD_MPIO']=df_PQRD2020['PET_COD_MPIO'].replace(',',' ', regex=True)
df_PQRD2020['AFEC_COD_MPIO']=df_PQRD2020['AFEC_COD_MPIO'].replace(',',' ', regex=True)
df_PQRD2020['ENT_COD_MPIO']=df_PQRD2020['ENT_COD_MPIO'].replace(',',' ', regex=True)
df_PQRD2020['FECHA_CREACION'] = pd.to_datetime(df_PQRD2020['FECHA_CREACION'])
df_PQRD2020 = df_PQRD2020.astype({"AÑO": int, "FECHA_CREACION": 'datetime64[ns]', "PET_COD_MPIO": int})
df_PQRD2020.dtypes
```

AÑO	int64
MES	int64
TRIM	object
FECHA_CREACION	datetime64[ns]
PQR_CANAL	object
PQR_TIPOPETICION	object
PQR_CLASE_SNS	object
PET_TIPOPER	object
PET_COD_DEPTO	int64
PET_DPTO	object
PET_COD_MPIO	int64
PET_MPIO	object

Fuente: Elaboración Propia

- Valores Máximos y Mínimos

Para ejecutar este ajuste se utilizó el set de datos de Divipola del Dane para la validación de datos de ubicación de los departamentos y municipios ya que algunos presentaban inconsistencias para lo cual se generó nuevamente los datos de la columna PET_COD_DEPTO tomando como referencia el nombre del departamento por medio de una función lambda, (ver figura 32 y 33).

Figura 32. Ajuste de Valores Máximos y Mínimos

```
# Funcion ajuste valores maximos y minimos: columna PET_COD_DEPTO
def ajuste_cod(nombre):
    var=df_CODDIV[df_CODDIV.Cod_departamento == nombre ]
    if (var.Cod_departamento.empty == False):
        return var.Cod_departamento.iloc[0]
    else:
        return 0

[ ] df_PQRD2018_["PET_COD_DEPTO "]=df_PQRD2018.apply( lambda x: ajuste_cod(x.PET_DPTO) ,axis=1)
df_PQRD2018_.head(2)
```

Fuente: Elaboración Propia

Figura 33. Ajuste Tipo Dato Código Departamento

PET_TIOPER	PET_COD_DEPTO	PET_DPTO	PET_MPIO	PQ
Natural	11	BOGOTÁ D.C.	BOGOTÁ D.C.	
Natural	76	VALLE DEL CAUCA	CALI	
Natural	76	VALLE DEL CAUCA	CALI	

Fuente: Elaboración Propia

- Duplicidad

Se verificó e identificó en el listado de las entidades de salud los datos duplicados y se ajustaron con la finalidad de tener solamente un registro de cada entidad, (ver figura 34).

Figura 34. Identificación Datos Duplicados

```

# Duplicados, entidades de salud
Entidades=pd.unique(df_PQRD2018.ENT_NOMBRE)
Entidades
.....
'ESE HOSPITAL SAN VICENTE DE PAUL',
'HOSPITAL GENERAL DE MEDELLIN - LUZ CASTRO DE GUTIERREZ',
'CENTRO MEDICO Y DE REHABILITACION VALLE SALUD SAS',
'CENTRO HOSPITAL DE LA FLORIDA E.S.E.', 'ESE HOSPITAL SANTA CLARA',
'DEPARTAMENTO DEL GUAINIA',
'CAJA DE COMPENSACION FAMILIAR COMFENALCO DEL VALLE DEL CAUCA',
'HOSPITAL UNIVERSITARIO HERNANDO MONCALEANO PERDOMO',
'DENTIX COLOMBIA S.A.S.', 'ESE HOSPITAL LA MISERICORDIA',
'FUNDACION', 'ALTA VISION LTDA', 'VINZETA S.A.',
'ESE HOSPITAL PRIMER NIVEL GRANADA', 'FAMISALEM IPS LIMITADA',
  
```

Fuente: Elaboración Propia

Ejecutando los ajustes en los duplicados se logró ajustar todos los campos. A continuación, la entidad de salud Nueva EPS S.A en un único registro, (ver figura 35).

Figura 35. Ajuste Datos Duplicados

```
'MEDIMAS EPS S.A.S.': 1,
'MEDPLUS CENTRO DE RECUPERACION INTEGRAL SAS': 1,
'MEDPLUS MEDICINA PREPAGADA S.A.': 1,
'MUNDIAL DE SEGUROS S.A.': 1,
'NO ABILICA': 1,
'NUEVA EPS S.A.': 1,
'PIJAOS SALUD EPSI': 1,
'PLAN U.H.C.M. MEDICINA PREPAGADA COMFENALCO VALLE': 1,
'POSITIVA COMPAÑIA DE SEGUROS SA': 1,
```

Fuente: Elaboración Propia

Se realiza la eliminación de columnas no necesarias del dataset, de las cuales no serán utilizados los datos para la visualización en el tablero de comandos, para lo cual de 47 columnas se depuran 11 columnas para un total de 36 columnas después de la limpieza, (ver figura 36).

Figura 36. Eliminación Columnas

```
# Eliminacion columnas innecesarias
df_PQRD2020.drop(['PQR_CLASE_SNS'], axis=1, inplace=True)
df_PQRD2018.drop(['AFEC_TIPODOC'], axis=1, inplace=True)
df_PQRD2018.drop(['AFEC_INSTITUCION'], axis=1, inplace=True)
df_PQRD2018.drop(['ENT_CODTIPOVIG'], axis=1, inplace=True)
df_PQRD2018.drop(['ENT_TIPOVIG'], axis=1, inplace=True)
df_PQRD2018.drop(['ENT_AMBITOIVC'], axis=1, inplace=True)
df_PQRD2018.drop(['ENT_MPIO_DANE'], axis=1, inplace=True)
df_PQRD2018.drop(['PQR_CLASE_SNS'], axis=1, inplace=True)
df_PQRD2018.drop(['ENTIDAD_FUNC'], axis=1, inplace=True)
df_PQRD2018.drop(['ALERTA_TIPO'], axis=1, inplace=True)
df_PQRD2018.drop(['FECHA_RADICACION'], axis=1, inplace=True)
df_PQRD2018.drop(['POS'], axis=1, inplace=True)
df_PQRD2018.drop(['REMITIDO'], axis=1, inplace=True)
df_PQRD2018.drop(['PQR_FUNCIONARIOSNS'], axis=1, inplace=True)
```

Fuente: Elaboración Propia

5.3.2 Datos Críticos

A continuación, se describen los datos críticos de las PQRD interpuestas ante la Supersalud, (ver tabla 5).

Tabla 5. Datos Críticos.

Campo	Descripción
PQR_CANAL	Se trata del medio por donde los usuarios interponen las pqrq ante la Supersalud
FECHA_CREACION	Fecha de creación de la PQRD por parte del usuario afectado.
ENT_NOMBRE	Nombre de la entidad de salud
PQR_ESTADO	El estado de la PQRD que se compone de 5: Cerrado, Respuesta por la entidad, Pendiente (Avance), Casos para instrucción y Reabierto.
PET_COD_MPIO	El código del municipio donde se interpone la PQRD.

Fuente: Elaboración Propia

5.3.3 Medición de la Calidad de Datos, Indicadores y Métricas.

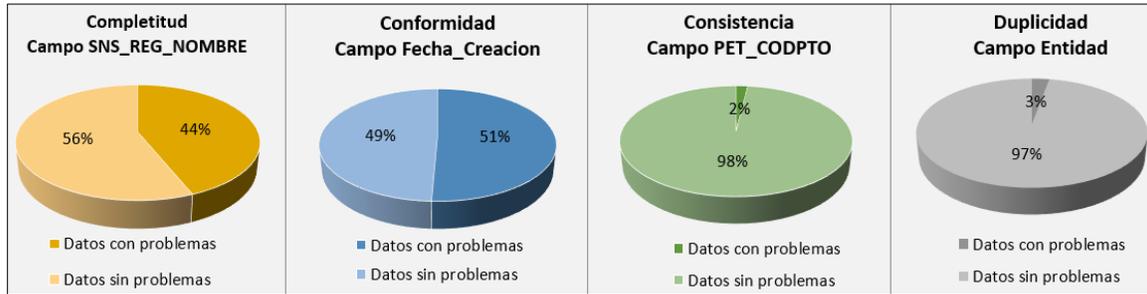
En la siguiente tabla se describen las dimensiones de calidad, las reglas de negocio, los atributos o datos críticos, las medidas, las métricas de cada una de las dimensiones, el código del indicador “IND” y el indicador de estado donde se describe el porcentaje de tolerancia de error establecido, (ver tabla 6 y figura 37).

Tabla 6. Datos Críticos PQRD.

Dimensión y Reglas de Negocio	Atributo	Medida	Métricas	Indicador de Estado
Complejidad (Vacíos). El lugar en el cual se recepciona la PQRD de acuerdo con sedes regionales dispuestas por la Supersalud debe ser obligatorio.	SNS_Reg Nombre	Conteo del número de registros diligenciados, compararlo con el número total de registros. Total identificados: 736303 Total campos: 1678890	Dividir el número de registros por el número total de registros de la tabla y multiplíquelo por 100. Porcentaje (736303 /1678890)*100=43.85%	IND1 % tolerancia: 12 Valor obtenido: 43.85 % Inaceptable: por encima del 2%
Conformidad (Formato). El formato establecido para fechas es: YYYY-MM-DD	Fecha creación	Conteo del número de campos que no contienen el formato de fecha definido. Mostrar el porcentaje de registros con mal formato de fecha. Total identificados: 850931 Total campos: 1678890	Dividir el número de registros por el número total de registros de la tabla y multiplíquelo por 100. Porcentaje (850931 /1678890)*100=50.68%	IND2 % tolerancia: 10 Valor obtenido: 50.68 % Inaceptable: por encima del 5%
Consistencia. al hacer el cruce de información con la fuente de información de la Supersalud con datos abiertos no debe existir información contradictoria	Pet_Cod departamento	Conteo del número de campos inconsistentes con otra fuente de datos. Mostrar el porcentaje de registros inconsistentes. Total identificados: 32526 Total campos: 1678890	Dividir el número de registros por el número total de registros de la tabla y multiplíquelo por 100. Porcentaje (32526/1678890)*100=1.93%	IND3 % tolerancia: 0 Valor obtenido: 1.93 % Inaceptable: por encima del 0%
Duplicación/Unicidad. Debe haber un único registro por entidad de salud en la tabla.	Entidad Nombre	Conteo del número de registros repetidos identificados. Mostrar el porcentaje de registros duplicados. Total identificados: 47540 Total campos: 1678890	Dividir el número de registros por el número total de registros de la tabla y multiplíquelo por 100. Porcentaje (47540 /1678890)*100=2.83%	IND4 % tolerancia: 0 Valor obtenido: 2.83 % Inaceptable: por encima del 0%

Fuente: Elaboración Propia

Figura 37. Medición de las Dimensiones de Calidad



Fuente: Elaboración Propia

5.3.4 Almacenamiento Datos

Una vez ejecutado todo el proceso de preparación se procede a almacenar los datos limpiados a una tabla en el SGBD Mysql alojado en una instancia de AWS por medio de Google Colab. Inicialmente se crea la tabla en el ambiente Stage llamada “DATA_CLEANPQRD” como se muestra a continuación, (ver figura 38).

Figura 38. Creación Tabla DATA_CLEANPQRD

```

130
131 CREATE TABLE IF NOT EXISTS DATA_CLEANPQRD (
132     AÑO TINYINT,
133     MES TINYINT,
134     TRIM varchar(15) ,
135     FECHA_CREACION DATE,
136     PQR_CANAL varchar(150) ,
137     PQR_TIPOPETICION varchar(150) ,
138     PET_TIPOPER varchar(100) ,
139     PET_COD_DEPTO INT,
140     PET_COD_MPIO INT,
141     AFEC_TIPOPER varchar(150) ,
142     AFEC_PARENTESCO varchar(150) ,
143     AFEC_GENERO varchar(150) ,
144     AFEC_EDADR varchar(150) ,

```

Fuente: Elaboración Propia

Se realiza la creación de índices sobre la tabla data_cleanpqr para su optimización en la inserción y consulta de datos, (ver figura 39).

Figura 39. Creación Índices

Table:	Comment:
DATA_CLEANPQRD	
Columns (36)	Keys
Indexes (5)	
Foreign Keys	
DATA_CLEANPQRD_AÑO_index INDEX DATA_CLEANPQRD (AÑO ASC)	
DATA_CLEANPQRD_FECHA_CREACION_index INDEX DATA_CLEANPQRD (FECHA_CREACION ASC)	
DATA_CLEANPQRD_PQR_CANAL_index INDEX DATA_CLEANPQRD (PQR_CANAL ASC)	
DATA_CLEANPQRD_PET_COD_DEPTO_index INDEX DATA_CLEANPQRD (PET_COD_DEPTO ASC)	
DATA_CLEANPQRD_PQR_ESTADO_index INDEX DATA_CLEANPQRD (PQR_ESTADO ASC)	

Fuente: Elaboración Propia

Proceso de inserción de los datos en la tabla data_cleanpqr de la base de datos SMYSQLPQRD en el motor de base datos MYSQL en la plataforma AWS Amazon por medio de la herramienta Google Colab en dos partes, (ver figura 40 y 41).

Figura 40. Conexión Base Datos AWS

```

# Conexión a BD Mysql-AWS (STAGE)
connection = pymysql.connect(host='smysqlpqr.c8cyqfmrolpx.us-east-1.rds.amazonaws.com',
                             user='[REDACTED]',
                             password='[REDACTED]',
                             db='SMYSQLPQRD',
                             charset='utf8mb4',
                             port=3310,
                             cursorclass=pymysql.cursors.DictCursor)

# Almacenamiento primer DataFrame
try:
    with connection.cursor() as cursor:
        # Limpiar tabla antes de insertar datos
        sql1 = "TRUNCATE TABLE `DATA_CLEANPQRD`"
        cursor.execute(sql1)
        result = cursor.fetchone()
        for index, row in df_PQRD2020.iterrows():
            sql = "INSERT INTO `DATA_CLEANPQRD` (AÑO, MES, TRIM, FECHA_CREACION, PQR_CANAL, PQR_TIPOPETICION, PET_T)
            cursor.execute(sql, (row.AÑO, row.MES, row.TRIM, row.FECHA_CREACION, row.PQR_CANAL, row.PQR_TIPOPETICION))
            connection.commit()

```

Fuente: Elaboración Propia

Figura 41. Carga Datos Base datos AWS

```
#Almacenamiento segundo DataFrame

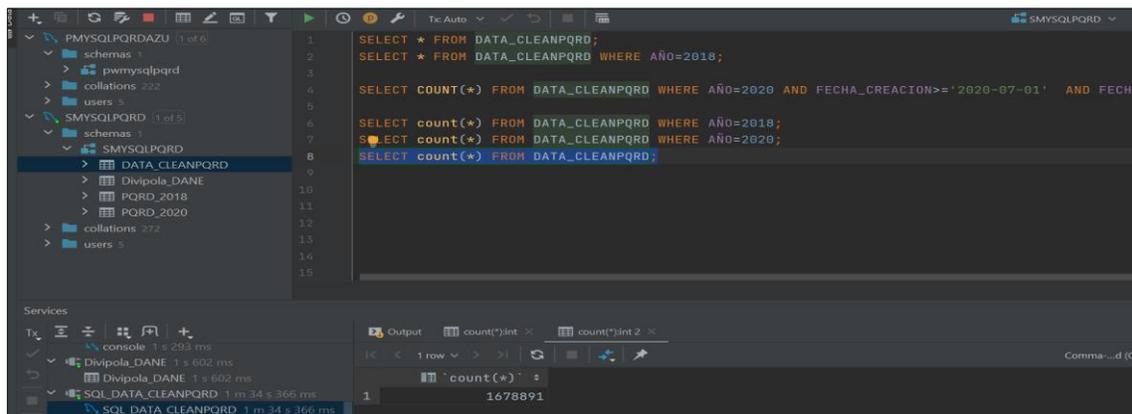
connection = pymysql.connect(host='smysqlpqrdr.c8cyqfmrolpx.us-east-1.rds.amazonaws.com',
                             user=██████████,
                             password=██████████,
                             db='SMYSQLPQRD',
                             charset='utf8mb4',
                             port=3310,
                             cursorclass=pymysql.cursors.DictCursor)

try:
    with connection.cursor() as cursor:
        for index, row in df_PQRD2018.iterrows():
            sql = "INSERT INTO `DATA_CLEANPQRD` (AÑO, MES, TRIM, FECHA_CREACION, PQR_CANAL, PQR_TIPOPETICION, PET_T
            cursor.execute(sql, (2018, row.MES, row.TRIM, row.FECHA_CREACION, row.PQR_CANAL, row.PQR_TIPOPETICION, r
            connection.commit()
        finally:
            cursor.close()
            connection.close()
```

Fuente: Elaboración Propia

Una vez finalizada la carga de los datos a la base de datos alojada en AWS, se realiza la verificación de los datos en la base de datos, (ver figura 42).

Figura 42. Verificación Carga Datos



Fuente: Elaboración Propia

5.3.5 Carga Datos Ambiente Productivo

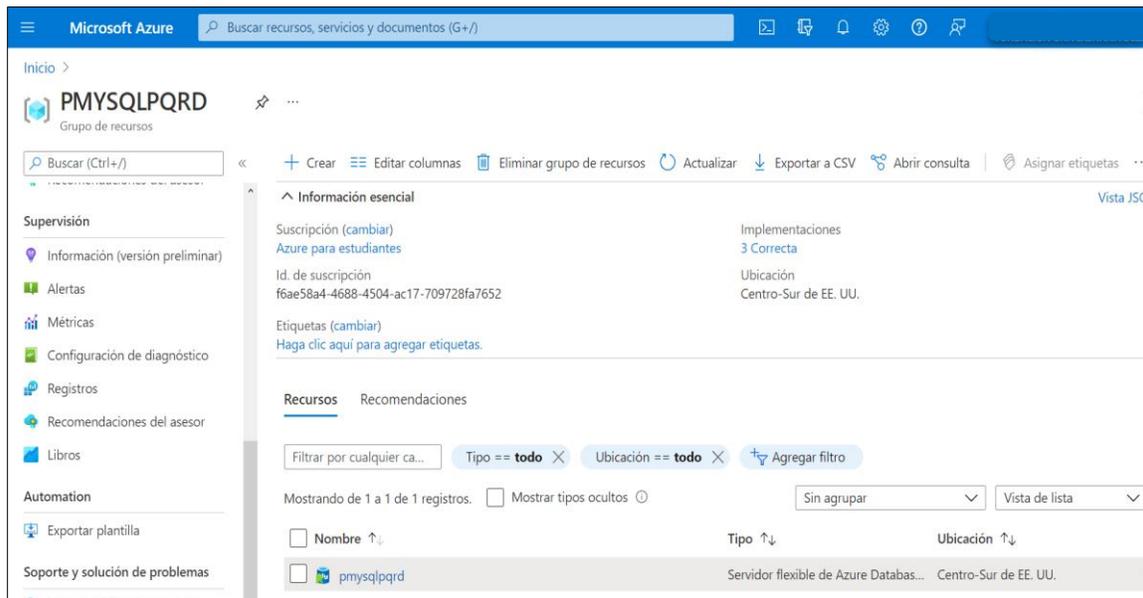
Se creó inicialmente una instancia del SGBD MySQL en la plataforma Microsoft Azure para almacenar los datos limpiados en la etapa 3 y se ejecutó las sentencias SQL para la creación de las tablas que habían sido diseñadas y elaboradas en la etapa 2 del

proyecto. Con los datos obtenidos aquí se aplicó el proceso de transformación para obtener la calificación del proceso de atención a las PQRD de las entidades de salud.

A continuación, se presentan las características del servidor de base de datos en la plataforma Microsoft Azure para el ambiente productivo (ver figura 43).

- ✓ Nombre: **pmysqlpqr**d
- ✓ Disco duro: 20 GB
- ✓ Versión: 5.7 Community
- ✓ vCPU: 2
- ✓ Memoria RAM: 4 GB
- ✓ Generación de log: Si
- ✓ Monitoreo y backups: Si

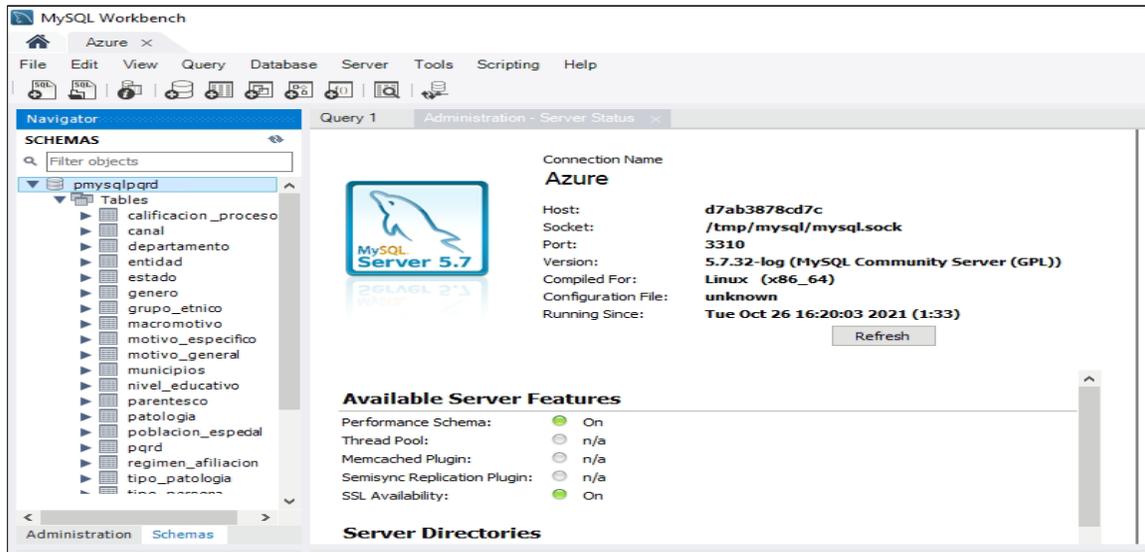
Figura 43. Base Datos Producción Azure



Fuente: Elaboración Propia

Se crea el esquema de base de datos pmysqlpqrđ con las respectivas tablas de acuerdo a la definición durante el modelamiento de datos en la plataforma Azure, (ver figura 44).

Figura 44. Verificación Creación Base de Datos



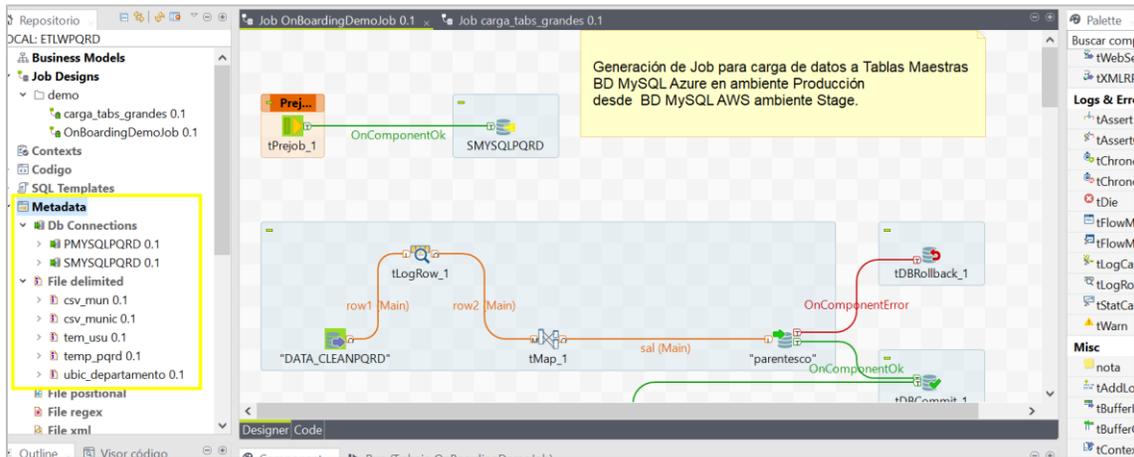
Fuente: Elaboración propia

Una vez se tuvieron las tablas creadas se procedió a cargarlas comenzando con las tablas Maestras y terminando con las tablas Transaccionales. Para este proceso se utilizó la herramienta Talend para extraer y carga los datos desde la BD AWS Stage hacia Azure BD Producción. A continuación, se describe el proceso ejecutado.

5.3.5.1 Creación Job en Talend para Extraer y Cargar Datos

Se realiza la parametrización de la Metadata para que las conexiones y las plantillas CSV del proceso de extracción queden de forma global (ver figura 45).

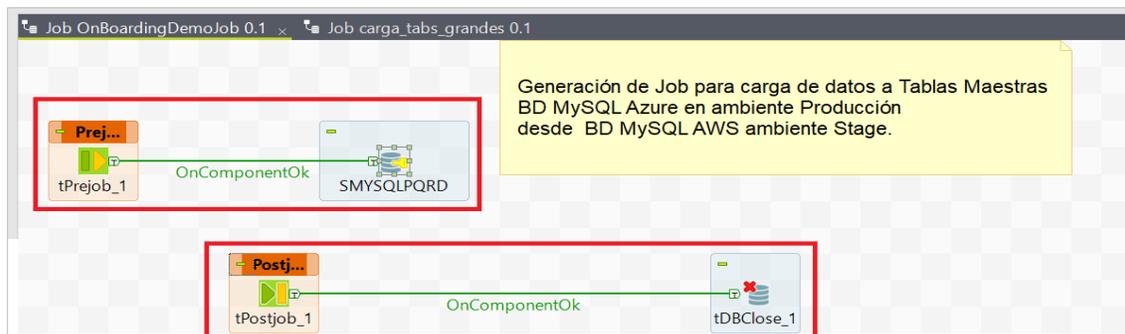
Figura 45. Generación Job Carga Datos



Fuente: Elaboración Propia

Al tener definidas las conexiones y plantillas CSV se crean los componentes PreJob que inicia la conexión a la BD y PostJob que cierra la conexión una vez termina la ejecución de todos los componentes del Job, (ver figura 46).

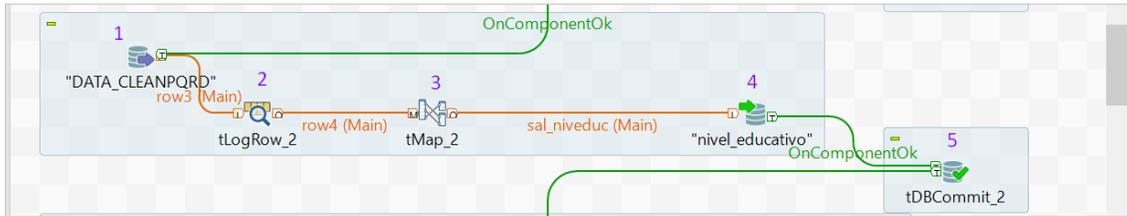
Figura 46. Creación Prejob y Postjob



Fuente: Elaboración Propia

Luego, para el primer Job que carga las tablas maestras, se crea los componentes siguiendo un orden general en casi todos los flujos de componentes, (ver figura 47).

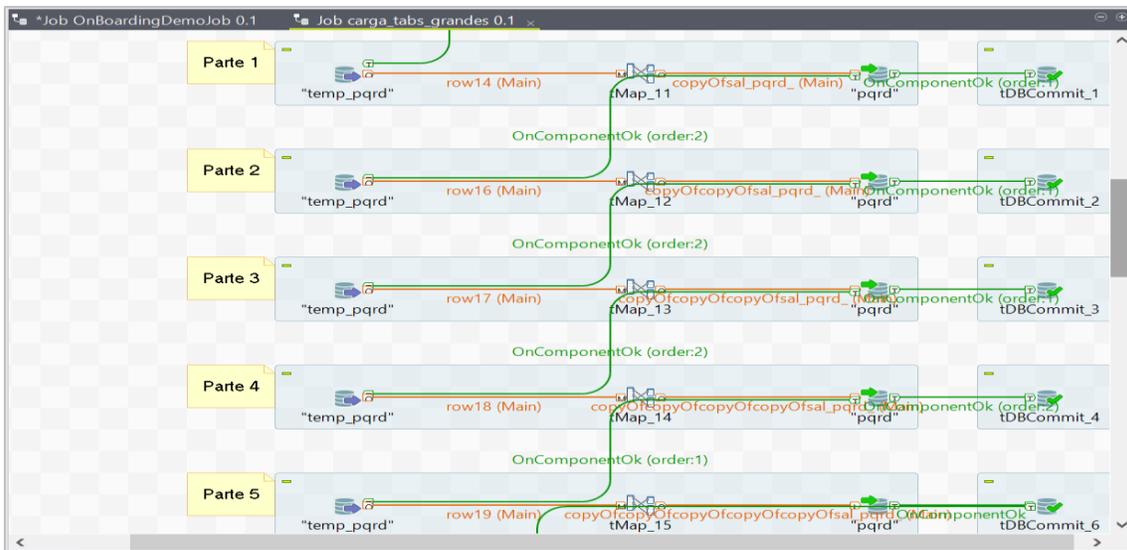
Figura 47. Flujo General Tablas Primer Job



Fuente: Elaboración Propia

Y finalmente para la carga de las tablas transnacionales se dividió la carga de datos en 6 partes de 300.000 registros y la ultima de 178.891aproximadamente para un total de 1.678.891 registros. De esta forma se aplicó la carga a las 2 tablas con mayor volumen en la base de datos “pmysqlpqrđ” de Producción tomando menos tiempo de carga (ver figura 48).

Figura 48. Carga de Datos por Lotes de Tablas Grandes



Fuente: Elaboración Propia

5.4 Transformación de los Datos

Para este caso particular, en esta etapa se aplican los procesos de transformación de los datos de las PQRD por medio de procesos ETL, los cuales serán posteriormente

cargados en una base de datos para finalmente disponerlos para las consultas previamente definidas en base al autor Carlos Trujillo (Carlos Trujillo, 2013). Adicionalmente, se aplica el modelo de análisis multicriterio a través del modelo RIM (E. Cablesa, 2016) sobre los datos en este caso particular, sobre las PQRD de las entidades prestadoras de servicio de salud, aplicando el método RIM con el objetivo de obtener una calificación del estado del proceso de peticiones de cada una.

Los criterios utilizados en el método RIM no necesariamente tienen los mismos dominios asociados, Por tales razones, para cada criterio seleccionado se determina la frecuencia relativa. También es necesario identificar ciertos conceptos como el rango y la referencia Ideal (E. Cablesa, 2016), para lo cual se tienen en cuenta los siguientes puntos.

- Aplicación del modelo multicriterio RIM a los estados de la PQRD de las entidades de salud.
- Aplicación de procesos ETL a los datos de la PQRD.
- Definición de indicadores, KPIs y consultas SQL de los datos para la disposición en el tablero de control.

5.4.1 Método RIM

Se establecen seis pasos para la aplicación de este método en base al autor (Cables, ScienceDirect, 2016) y un paso previo donde se calculan y definen algunos valores necesarios para su implementación, las cuales se describirán a continuación tomando tres entidades de salud para el estudio ejemplo.

Se inicia con la organización de las entidades por el estado de la PQRD, conteo total de cada estado PQRD, conteo total de PQRD por cada entidad y se calcula su frecuencia relativa, (ver tabla 7).

Tabla 7. Conteo Total PQRD.

COD_ENT	ENTIDAD	ESTADO PQRD	TOTAL PC	TOTAL	FRECUENCIA
999998	ALCALDIA CARTAGENA DE INDIAS DISTRITO TURISTICO	Pendiente (Avance)	31	66	0.4696997
999998	ALCALDIA CARTAGENA DE INDIAS DISTRITO TURISTICO	Casos para instruccion	3	66	0.04545455
999998	ALCALDIA CARTAGENA DE INDIAS DISTRITO TURISTICO	Cerrado	32	66	0.48484848
999998	ALCALDIA CARTAGENA DE INDIAS DISTRITO TURISTICO	REABIERTO	0	66	0
999998	ALCALDIA CARTAGENA DE INDIAS DISTRITO TURISTICO	Respuesta por la entidad	0	66	0
EPS001	ALIANZASALUD ENTIDAD PROMOTORA DE SALUD S.A	Casos para instruccion	194	4175	0.04646707
EPS001	ALIANZASALUD ENTIDAD PROMOTORA DE SALUD S.A	Pendiente (Avance)	142	4175	0.03401198
EPS001	ALIANZASALUD ENTIDAD PROMOTORA DE SALUD S.A	Respuesta por la entidad	3434	4175	0.82251497
EPS001	ALIANZASALUD ENTIDAD PROMOTORA DE SALUD S.A	REABIERTO	18	4175	0.00431138
EPS001	ALIANZASALUD ENTIDAD PROMOTORA DE SALUD S.A	Cerrado	387	4175	0.09269461
EPSS40	ALIANZA MEDELLIN ANTIOQUIA EPS S.A.S. - SAVIA SALUD	Pendiente (Avance)	7281	20719	0.35141657
EPSS40	ALIANZA MEDELLIN ANTIOQUIA EPS S.A.S. - SAVIA SALUD	Respuesta por la entidad	5444	20719	0.26275399
EPSS40	ALIANZA MEDELLIN ANTIOQUIA EPS S.A.S. - SAVIA SALUD	Casos para instruccion	1623	20719	0.0783339
EPSS40	ALIANZA MEDELLIN ANTIOQUIA EPS S.A.S. - SAVIA SALUD	REABIERTO	176	20719	0.00849462
EPSS40	ALIANZA MEDELLIN ANTIOQUIA EPS S.A.S. - SAVIA SALUD	Cerrado	6195	20719	0.29900092

Fuente: Elaboración Propia

Una vez organizadas las entidades se procede a asignar códigos a cada uno de los estados PQRD para una mejor identificación y aplicación de las fórmulas, (ver tabla 8).

Tabla 8. Descripción de Criterios.

Descripción de los criterios	
Código	Criterios
C1	Cerrado
C2	Respuesta por la entidad
C3	Pendiente (Avance)
C4	Casos para instrucción
C5	Reabierto

Fuente: Elaboración Propia

Se procede a continuación a calcular el vector de Pesos de cada uno de los estados PQRD a partir de la cantidad de criterios (n=5), el intervalo de la pendiente para este caso va desde -0.1 hasta 0 y el valor para el pendiente identificado es de -0.9. A continuación,

se muestra la figura (Cálculo Vector de Peso) donde está la aplicación de la formula y el valor del Peso obtenido para cada estado PQRD, cuya sumatoria debe dar 1, (ver tabla 9).

Tabla 9. Cálculo Vector de Peso

CALCULO VECTOR DE PESO				
Cantidad de Criterios (n):	5	(ESTADOS PQRD)		
	Ext. Izq.	Ext. Der.		
Intervalo de la pendiente:	-0.1	0		
Valor de la pendiente (a):	-0.09		x	F(x)
			1	0.38
			2	0.29
			3	0.20
			4	0.11
			5	0.02
			Suma:	1.00

Fuente: Elaboración Propia

Realizado estos cálculos, se procede a definir la solución ideal para cada criterio. Se estableció como solución ideal para el criterio C1(estado cerrado) el intervalo seria de 0,9 a 1 para el C2(Respuesta por la entidad) de 0 a 0,05, C3(Pendiente) de 0 a 0,05, para C4(Casos para instrucción) y para los criterios C5(Reabierto) de 0 (ver tabla 10).

Tabla 10. Valores Solución Ideal.

Descripción de los criterios				
Código	Criterios	Pesos	Dominio	Solución Ideal (Intervalo)
C1	Cerrado	0.38	[0,1]	[0.9, 1]
C2	Respuesta por la entidad	0.29	[0,1]	[0, 0.05]
C3	Pendiente (Avance)	0.2	[0,1]	[0, 0.05]
C4	Casos para instrucción	0.11	[0,1]	0
C5	Reabierto	0.02	[0,1]	0
	Suma	1.00		

Fuente: Elaboración Propia

Una vez se tienen calculados los pesos, definida la solución ideal, los códigos asignados a cada estado PQRD y los datos organizados se inicia la aplicación de los seis pasos.

5.4.1.1 Paso 1

Se establece las condiciones del contexto de trabajo como se describe en la siguiente tabla, (ver tabla 11).

Tabla 11. Condiciones del Contexto de Trabajo.

PASOS		Ideal de Referencia					
1	Ext. I.	0.9	0	0	0	0	
	Ext. D.	1	0.05	0.05	0	0	
	Intervalos o Rango (Dominios de trabajo)						
	Ext. I.	0	0	0	0	0	
	Ext.D	1	1	1	1	1	
	Pesos						
		0.38	0.29	0.20	0.11	0.02	1.00

Fuente: Elaboración Propia

5.4.1.2 Paso 2

Se conforma la matriz de decisión, la cual contiene la frecuencia relativa para cada entidad de salud en los diferentes criterios que presentan los estados, (ver tabla 12).

Tabla 12. Matriz X Estados PQRD y Entidades de Salud.

Entidad de Salud (Alternativas)	Matriz de Decisión (Matriz X)				
	C1	C2	C3	C4	C5
ALCALDIA CARTAGENA DE INDIAS DISTRITO TURISTICO Y CULTURAL	0.48484848	0	0.46969697	0.045454545	0
ALIANSA SALUD ENTIDAD PROMOTORA DE SALUD S.A	0.09269461	0.822515	0.034011976	0.046467066	0.00431138
ALIANZA MEDELLIN ANTIOQUIA EPS S.A.S. - SAVIA SALUD EPS	0.29900092	0.262754	0.351416574	0.078333896	0.00849462

Fuente: Elaboración Propia

5.4.1.3 Paso 3

Para este paso se aplica la fórmula del método RIM definido en el marco conceptual en el paso tres, para regularizar la matriz X con el ideal de referencia (IR) o solución ideal. A continuación, se muestra la nueva matriz Y obtenida (ver tabla 13).

Tabla 13. Matriz Normalizada.

Entidad de Salud (Alternativas)	Matriz Normalizada (Matriz Y)				
	C1	C2	C3	C4	C5
ALCALDIA CARTAGENA DE INDIAS DISTRITO TURISTICO Y CULTURAL	0.53872	1.00000	0.55821	0.95455	1.00000
ALIANSA SALUD ENTIDAD PROMOTORA DE SALUD S.A	0.10299	0.18683	1.00000	0.95353	0.99569
ALIANZA MEDELLIN ANTIOQUIA EPS S.A.S. - SAVIA SALUD EPS	0.33222	0.77605	0.68272	0.92167	0.99151

Fuente: Elaboración Propia

5.4.1.4 Paso 4

Se procede en este paso a ponderar la matriz normalizada mediante el vector de peso definido, obteniendo los siguientes valores tal y como se visualiza, (ver tabla 14).

Tabla 14. Matriz Normalizada con Peso.

Entidad de Salud (Alternativas)	Matriz normalizada con influencia de pesos (vij)				
	C1	C2	C3	C4	C5
ALCALDIA CARTAGENA DE INDIAS DISTRITO TURISTICO Y CULTURAL	0.20471	0.29000	0.11164	0.10500	0.02000
ALIANSA SALUD ENTIDAD PROMOTORA DE SALUD S.A	0.03914	0.05418	0.20000	0.10489	0.01991
ALIANZA MEDELLIN ANTIOQUIA EPS S.A.S. - SAVIA SALUD EPS	0.12624	0.22505	0.13654	0.10138	0.01983

Fuente: Elaboración Propia

5.4.1.5 Paso 5

En este paso se calcula la variación al ideal de referencia normalizado para cada alternativa, según la fórmula descrita en el apartado. Luego se aplica la raíz cuadrada a la sumatoria de cada fila para determinar la variación de la solución ideal positiva, (ver tabla 15).

Tabla 15. Variación a la Solución Ideal

Entidad de Salud (Alternativas)	Variación a la Solución Ideal +					I+
	C1	C2	C3	C4	C5	
ALCALDIA CARTAGENA DE INDIAS DISTRITO TURISTICO Y CULTURAL	0.03073	0.00000	0.00781	0.00002	0.00000	0.19636
ALIANSA SALUD ENTIDAD PROMOTORA DE SALUD S.A	0.11619	0.05561	0.00000	0.00003	0.00000	0.41452
ALIANZA MEDELLIN ANTIOQUIA EPS S.A.S. - SAVIA SALUD EPS	0.06439	0.00422	0.00403	0.00007	0.00000	0.26965

Fuente: Elaboración Propia

Ahora, se ejecutan los cálculos esta vez para el indicador (I-) que es la variación de la solución ideal negativa con la raíz cuadrada de la resta de los valores de cada fila, (ver tabla 16).

Tabla 16. Variación a la Solución Ideal Negativa

Entidad de Salud (Alternativas)	Variación a la Solución Ideal -					
	C1	C2	C3	C4	C5	I-
ALCALDIA CARTAGENA DE INDIAS DISTRITO TURISTICO Y CULTURAL	0.04191	0.08410	0.01246	0.01103	0.00040	0.38717
ALIANSA SALUD ENTIDAD PROMOTORA DE SALUD S.A	0.00153	0.00294	0.04000	0.01100	0.00040	0.23636
ALIANZA MEDELLIN ANTIOQUIA EPS S.A.S. - SAVIA SALUD EPS	0.01594	0.05065	0.01864	0.01028	0.00039	0.30968

Elaboración: Propia

5.4.1.6 Paso 6

Para este último paso, se calcula el Índice Relativo IR a partir de los valores de la variación a la solución ideal positiva (I+) y negativa (I-) del paso 5, dividiendo (I-) entre la suma de (I+) con (I-), dando como resultado el índice relativo (IR) para alternativa que para este caso son las entidades de salud. Cada valor de agregación obtenido se encuentra entre el intervalo (0-1), lo que significa que mientras el IR se encuentre más próximo a 1 estará más cercano al Ideal de Referencia definido y mientras más se aleje de 1 (se acerque a cero) no coincidirá con el ideal, (ver tabla 17).

Tabla 17. Obtención del IR Para Cada Entidad de Salud.

Entidad de Salud (Alternativas)	
ALCALDIA CARTAGENA DE INDIAS DISTRITO TURISTICO Y CULTURAL	0.66349349
ALIANSA SALUD ENTIDAD PROMOTORA DE SALUD S.A	0.36313933
ALIANZA MEDELLIN ANTIOQUIA EPS S.A.S. - SAVIA SALUD EPS	0.53455122

Fuente: Elaboración Propia

Ahora se propone para el proceso de las PQRD de las entidades de la salud las siguientes calificaciones lingüísticas y su intervalo asociado, pero se tiene la libertad de elegir y proponer intervalos y calificaciones lingüísticas diferente, (ver tabla 18).

Tabla 18. Calificaciones Proceso PQRD.

Intervalo	Calificación
I.R ≥ 0 y < 0.30	MUY MALO
I.R ≥ 0.30 y < 0.50	MALO
I.R ≥ 0.50 y < 0.60	ACEPTABLE
I.R ≥ 0.60 y < 0.75	BUENO
I.R ≥ 0.75 y < 0.90	MUY BUENO
I.R ≥ 0.9 y < 1	EXCELENTE

Fuente: Elaboración Propia

A partir del paso 6, por medio del índice relativo (IR) calculado para cada alternativa, se aplica el intervalo propuesto de las calificaciones descrito a cada una de las entidades de salud (ver tabla 18), para finalmente asignar una calificación en el proceso de atención de peticiones de PQRD, (ver tabla 19).

Tabla 19. Entidades de Salud.

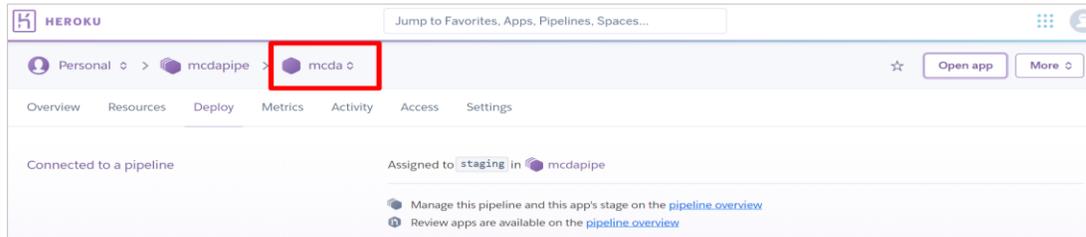
Entidad de Salud (Alternativas)		CALIFICACION
ALCALDIA CARTAGENA DE INDIAS DISTRITO TURISTICO Y CULTURAL	0.66349349	Bueno
ALIANSA ENTIDAD PROMOTORA DE SALUD S.A	0.36313933	Malo
ALIANZA MEDELLIN ANTIOQUIA EPS S.A.S. - SAVIA SALUD EPS	0.53455122	Aceptable

Fuente: Elaboración Propia

5.4.2 Implementación de API

Para el proceso de aplicación de los 6 pasos comprendidos dentro del MCDA con el método RIM se determinó la creación de una Application Programming Interface API en lenguaje Hypertext Preprocessor PHP que se encargó de ejecutar todos los cálculos

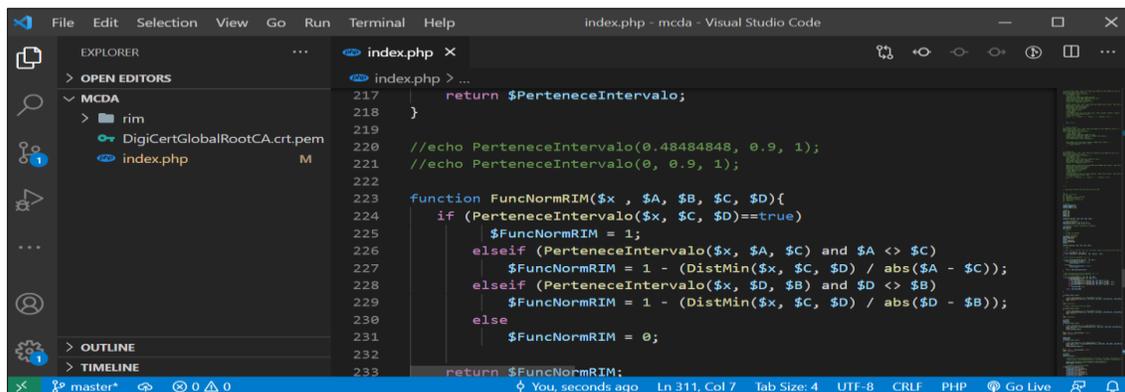
Figura 50. Plataforma Heroku



Fuente: Elaboración Propia

Para el siguiente paso de la transformación se codificó cada una de las ecuaciones matemáticas de los 6 pasos necesarios en los procesos del MCDA por medio del lenguaje PHP 8.0 procesando la información preparada de las PQRD de cada una de las entidades de salud desde la BD MySQL del ambiente producción configurado en Azure, ver figura 51).

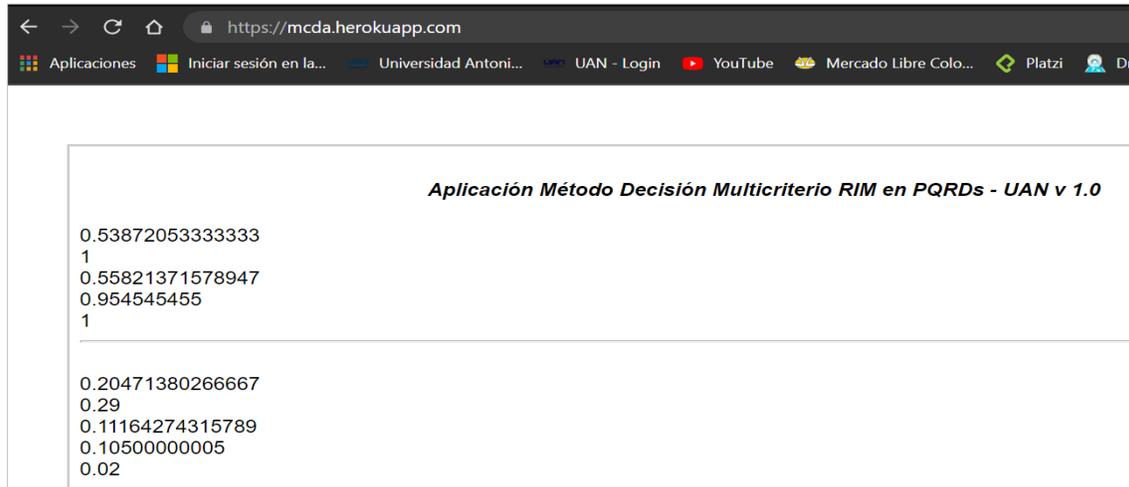
Figura 51. Codificación Api PHP



Fuente: Elaboración Propia

Se verificó la carga y se abrió la API por medio de la URL generada con certificado SS de seguridad por Heroku ingresando el respectivo usuario y contraseña: <https://mcda.herokuapp.com> (ver figura 52).

Figura 52. Aplicación Método Decisión Multicriterio



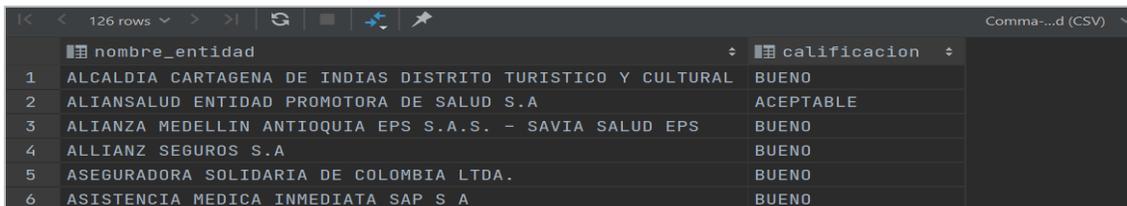
Fuente: Elaboración Propia

Al activar la API con el proceso de MCDA creada y desplegada en Heroku, se ejecutan las siguientes tareas:

- ✓ Completa creando un registro por cada Entidad de salud en la que falte alguno de los cinco estados de PQRD para que cada entidad quede con 5 registros. Las nuevas quedan con valor cero en su total.
- ✓ Calcula el total de estados PQRD por cada entidad
- ✓ Calcula la Frecuencia relativa por cada estado y entidad de las PQRD
- ✓ Asigna un código de 1 a 5 según el peso definido de cada estado PQRD
- ✓ Luego se aplican los 6 pasos definidos de MCDA en base al método RIM sobre los datos con sus respectivas ecuaciones.
- ✓ Finalmente, en base a los intervalos de calificaciones definidas para el proyecto (Aquí referencia tabla Estados PQRD) se aplica a cada una de las entidades según el resultado del modelo MCDA con el método RIM.

Al ser activada la API, la tabla entidades es actualizada con los valores de la calificación en base a la aplicación de MCDA con el método RIM en esta etapa de transformación de datos (ver figura 53).

Figura 53. Visualización Calificación Entidades

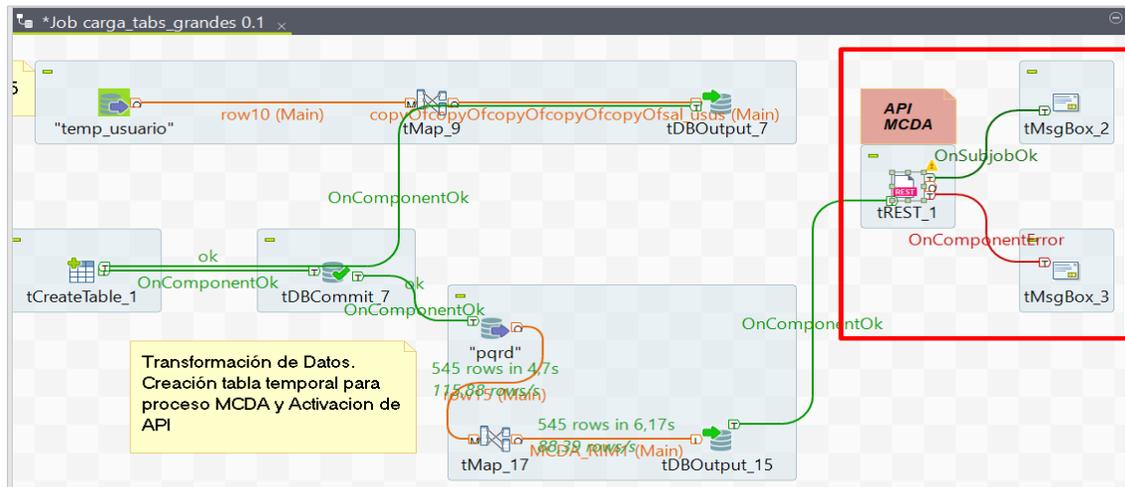


	nombre_entidad	calificacion
1	ALCALDIA CARTAGENA DE INDIAS DISTRITO TURISTICO Y CULTURAL	BUENO
2	ALIANSA SALUD ENTIDAD PROMOTORA DE SALUD S.A	ACEPTABLE
3	ALIANZA MEDELLIN ANTIOQUIA EPS S.A.S. - SAVIA SALUD EPS	BUENO
4	ALLIANZ SEGUROS S.A	BUENO
5	ASEGURADORA SOLIDARIA DE COLOMBIA LTDA.	BUENO
6	ASISTENCIA MEDICA INMEDIATA SAP S A	BUENO

Fuente: Elaboración Propia

Para el proceso de automatización se utilizó la aplicación Talend Open Studio donde se generó el flujo del job para la API con los componentes necesarios donde finalmente se procesó la información y se almacenó en una tabla en la base de datos de producción alojada en Microsoft Azure, la cual contiene los datos de cada una de las entidades de salud y la calificación obtenida por medio del método MCDA aplicado en la API. Una vez se ejecuta la API se valida si está todo correcto y genera una ventana informativa de confirmación, de lo contrario la ventana mostrará de tipo error para reintentar el proceso (ver figura 54).

Figura 54. Flujo Transformación Datos



Fuentes: Elaboración Propia

Con este proceso se termina la presente etapa de transformación de datos y se dispone la información para la etapa final de visualización.

5.5 Visualización de Datos

En esta etapa se definen las siguientes tareas para generar la visualización de los datos de los procesos de las pqrq en tableros de comandos en la herramienta Tableau.

- Definición de la herramienta para la visualización de los datos.
- Diseño de la visualización de los datos.
- Definición de los datos a visualizar.
- Despliegue y pruebas del tablero de control.

5.5.1 Herramienta Visualización Datos

Para la visualización de los datos en tableros de comandos se implementa la

herramienta Tableau, con la cual se realiza la captura de los datos oro desde la base de datos origen y se presentan de manera gráfica, de acuerdo a la definición de cada indicador a visualizar.

5.5.2 Definición Datos a Visualizar

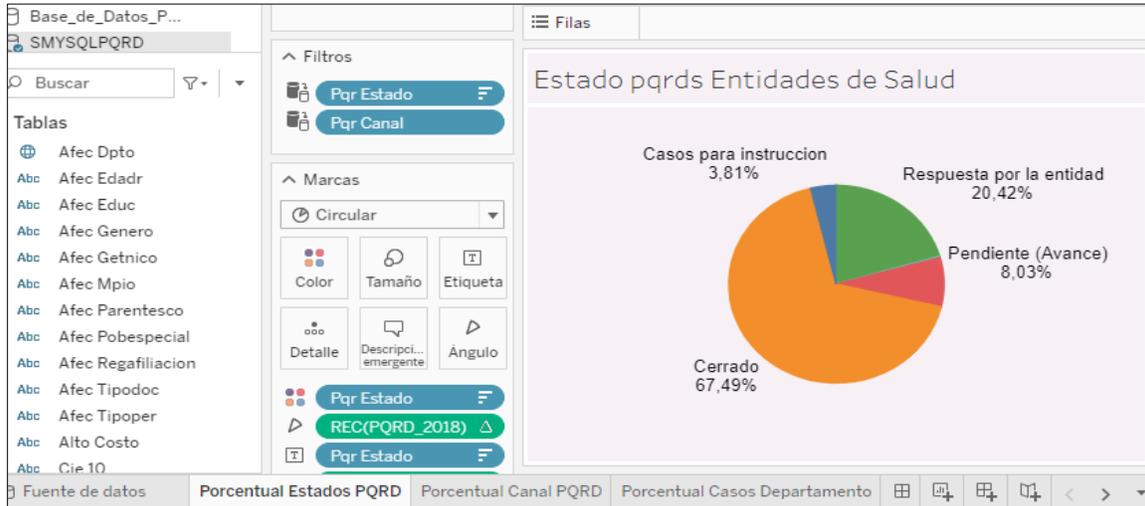
Se establecen los siguientes datos a visualizar en el tablero de comandos.

- Visualizar la calificación de las entidades prestadoras de servicios de salud en Colombia
- Visualizar Total PQRD interpuestas por departamentos de Colombia según su estado.
- Visualizar el total de PQRD por entidad según su estado.
- Visualizar la distribución de PQRD por canal y entidad de salud
- Visualizar la distribución de PQRD interpuestas por tipo de patología
- Visualizar la cantidad de PQRD interpuestas por mes según su estado.

5.5.3 Diseño Visualización datos

Durante esta etapa se define el diseño del tablero de comandos donde se visualizarán los resultados de los datos, a continuación, se visualiza una descripción sobre la construcción del gráfico cantidad porcentual por cada estado de PQRD interpuesta, (ver figura 55).

Figura 55. Diseño Gráfico Estado PQRD

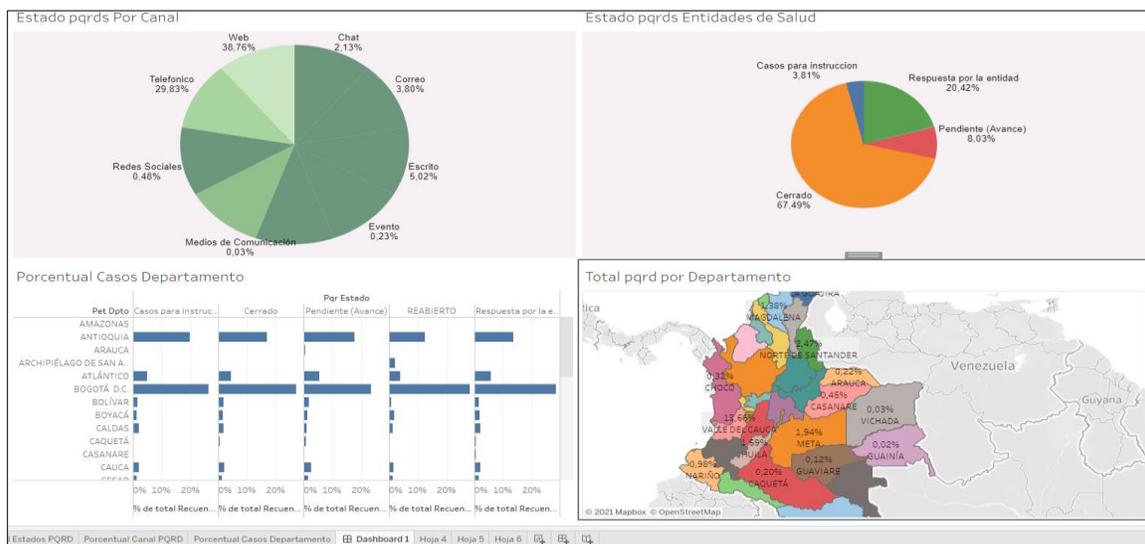


Fuente: Elaboración Propia

5.5.4 Despliegue y Pruebas Tablero de Control.

Una vez desarrollados los tableros de control con los gráficos de acuerdo con la definición de datos a visualizar, se procede a realizar el despliegue web de los tableros para la visualización de la información, (ver figura 56).

Figura 56. Tablero Comando



Fuente: Elaboración Propia

Una vez terminado el procesamiento de los datos y la aplicación del modelo MCDA con el método RIM se logró determinar la calificación del proceso de atención de las PQRD de las entidades de salud en Colombia interpuestas ante la Supersalud. El beneficio para los usuarios es que podrán buscar la entidad de salud a la que están afiliados y conocer por medio de la calificación obtenida si dicha entidad de salud se preocupa por la atención de las quejas de sus usuarios para darles pronta respuesta o por el contrario el proceso de atención de PQRD no es muy bien atendido por la entidad y puedan tomar la decisión si desean continuar, afiliarse o cambiar de entidad prestadora del servicio de salud. (ver figura 57).

Figura 57. Tablero Visualización Calificación Entidades

nombre_entidad	calificacion	Abc
ALCALDIA CARTAGENA DE INDIAS DISTRITO TURISTICO Y CULTURAL	BUENO	Abc
ALIANZASALUD ENTIDAD PROMOTORA DE SALUD S.A	ACEPTABLE	Abc
ALIANZA MEDELLIN ANTIOQUIA EPS S.A.S. - SAVIA SALUD EPS	BUENO	Abc
ALLIANZ SEGUROS S.A	BUENO	Abc
ASEGURADORA SOLIDARIA DE COLOMBIA LTDA.	BUENO	Abc
ASISTENCIA MEDICA INMEDIATA SAP S A	BUENO	Abc
ASISTENCIA MEDICA SAS SERVICIO DE AMBULANCIA PREPAGADO	EXCELENTE	Abc
ASOCIACION DE CABILDOS INDIGENAS DEL CESAR Y LA GUAJIRA DUSAKAWI EPSI	ACEPTABLE	Abc
ASOCIACIÓN INDÍGENA DEL CAUCA A.I.C. EPSI	BUENO	Abc
ASOCIACIÓN MUTUAL BARRIOS UNIDOS DE QUIBDÓ AMBUQ EPS-S ESS	ACEPTABLE	Abc
ASOCIACIÓN MUTUAL EMPRESA SOLIDARIA DE SALUD EMSSANAR E.S.S	BUENO	Abc
ASOCIACIÓN MUTUAL LA ESPERANZA "ASMET SALUD E.S.S EPS S	ACEPTABLE	Abc
ASOCIACIÓN MUTUAL SER EMPRESA SOLIDARIA DE SALUD ESS	ACEPTABLE	Abc
AXA COLPATRIA MEDICINA PREPAGADA S.A.	BUENO	Abc
AXA COLPATRIA SEGUROS DE VIDA SA	ACEPTABLE	Abc
CAJA COMPENSACIÓN FAMILIAR COMPENSAR	BUENO	Abc
CAJA DE COMPENSACION FAMILIAR DEL VALLE DEL CAUCA COMFENALCO VALLE	BUENO	Abc

Fuente: Elaboración Propia

6 Conclusiones

En base a la investigación realizada sobre la información relacionada con las PQRD interpuestas ante la Supersalud, se identificaron dos fuentes de datos provenientes de datos abiertos de Colombia y la Supersalud, donde se identificó que la Supersalud publicó la información del año 2020 y del 2018. Estas dos fuentes fueron de estudio para el desarrollo del proyecto. Se tomó una tercera fuente de datos DANE Divipola para la validación de las ubicaciones de las entidades de salud, usuarios e interposición de PQRD en Colombia.

Para el proceso de calificación de las PQRD se implementó en la plataforma AWS Amazon donde se almacenó la información en su estado original, para aplicarle los procesos de limpieza y preparación por medio de las herramientas Google Colab, una vez obtenida la data limpia se procedió a cargarla a la base de datos relacional creada en la plataforma Microsoft Azure por medio de la ELT Talend DI para disponerla hacia la plataforma Tableau que permitió para la visualización de los datos en un tablero de comandos de las PQRD y la calificación del proceso de atención de peticiones PQRD por parte de las entidades de salud.

Para todo el proceso y ciclo de vida de los datos de las PQRD se aplicaron buenas prácticas como la creación de ambiente de desarrollo, documentación de los datos, creación de cuentas de usuario, modelamiento de datos conceptual, lógico y físico en base al DAMABOOK, se ejecutaron tareas de perfilamiento de datos, dimensiones de calidad, monitoreo y control en las bases de datos, definición de lineamientos y estándares para los datos y componentes utilizados, tareas de automatización y definición

de la arquitectura de datos . Todo lo anterior permitió que la información de las PQRD de las entidades de salud tuviera mayor calidad, libre de inconsistencias, disponible para los usuarios y adecuada para aplicar el método RIM que define las calificaciones en el proceso de atención de las PQRD.

El método RIM fue posible aplicarlo en la programación de un script en lenguaje PHP el cual se dispuso en una API en un hosting gratuito y es activada automáticamente desde un Job con la herramienta ELT Talend. Se logró que esta API aplicará cada uno de los 6 pasos del método utilizando los datos de las entidades de salud y sus estados PQRD, con lo cual, se permitió asignar una calificación al proceso de atención de las PQRD de cada una de las entidades de salud en Colombia.

Con todos datos de las PQRD preparados y transformados almacenados en la nube Azure, se realizó los indicadores y consultas necesarias para la visualización de los datos en un tablero de control en la plataforma Tableau, la cual contiene filtros y los datos más relevantes de la investigación del proyecto como la información por entidades y sus calificaciones en el proceso de atención a las PQRD en Colombia. Estos filtros facilitan y generan mayor entendimiento sobre los datos permitiendo segregarlos por ubicación geográfica, motivos de la PQRD, canal de la PQRD, género de la persona solicitante, fecha de la petición PQRD, intervalo de edades de los usuarios que solicitan PQRD ante la Supersalud de los años 2018 y 2020, entre otras.

7 Referencias

- Aula21. (08 de 2018). *Aula21*. Obtenido de Aula21: <https://www.cursosaula21.com/que-es-python/>
- Bernard Marr, D. S. (2017). *Ecoe Ediciones*. Obtenido de Ecoe Ediciones: <http://ezproxy.uan.edu.co:2071/?il=6317&pg=137>
- Blog, A. P. (24 de 10 de 2019). AWS. Obtenido de AWS: <https://aws.amazon.com/es/blogs/publicsector/building-a-data-analytics-practice-across-the-data-lifecycle/>
- Cables, E. (2016). Information Sciences. *Information Sciences*, 5.
- Cables, E. (2016). RIM-reference ideal method in multicriteria decision making. *Information Sciences*, 5.
- Cables, E. (10 de 04 de 2016). *ScienceDirect*. Obtenido de ScienceDirect: <https://www.sciencedirect.com/science/article/abs/pii/S0020025515009007>
- Cables, E. (19 de 04 de 2016). *ScienceDirect*. Obtenido de ScienceDirect: <https://www.sciencedirect.com/science/article/abs/pii/S0020025515009007?via%3Dihub>
- Carlos Trujillo, J. (2013). *Diseño y explotación de almacenes de datos*. Obtenido de conceptos básicos de modelado multidimensional: <https://ezproxy.uan.edu.co:2830/es/ereader/bibliouan/62323?page=21>
- Casas Roma, J., Nin Guerrero, J., & Julbe López, F. (2019). *Big Data: Análisis de datos en entornos masivos*. Obtenido de Big Data: Análisis de datos en entornos masivos.: <https://ezproxy.uan.edu.co:2830/es/ereader/bibliouan/117744?page=95>
- Ciberclick. (29 de 06 de 2020). *Ciberclick*. Obtenido de Ciberclick: <https://www.cyberclick.es/numerical-blog/que-es-un-dashboard>
- DAMA. (2008). *FDOC*. Obtenido de FDOC: <https://qdoc.tips/dmbok-pdf-free.html>
- DANE. (30 de 09 de 2021). *Geovisor de Consulta de Codificación de la Divipola*. Obtenido de Geovisor de Consulta de Codificación de la Divipola: <https://geoportal.dane.gov.co/geovisores/territorio/consulta-divipola-division-politico-administrativa-de-colombia/>
- E. Cablesa, *. M. (2016). *RIM-Reference ideal method in multicriteria decision making*. Elsevier.
- Información, T. d. (2018). *Tecnologías de la Información*. Obtenido de Tecnologías de la Información: <https://www.tecnologias-informacion.com/modelos-datos.html>
- KeepCoding. (27 de 02 de 2020). *KeepCoding*. Obtenido de KeepCoding: <https://keepcoding.io/blog/que-son-datasets/>
- Lamata, M. T., & y Cables, E. (2009). *OWA weights determination by means of linear functions*.
- Maquilme, M. (25 de 09 de 2021). *Web y Empresas*. Obtenido de Web y Empresas: <https://www.webyempresas.com/en-que-consiste-el-proceso-etl/#:~:text=ETL%20es%20la%20abreviatura%20de%20Extracci%C3%B3n%20Transformaci%C3%B3n%20y,lectura%20de%20datos%20de%20una%20base%20de%20datos.>

- Minsalud. (2021). *Minsalud*. Obtenido de Minsalud:
<https://www.minsalud.gov.co/salud/Pservicios/Paginas/relaciones-entidades-prestadoras-pagadoras-servicios-salud.aspx>
- Nieto Bernal, W. y. (2017). *Diseño de base de datos*. Obtenido de Diseño de base de datos: <https://ezproxy.uan.edu.co:2830/es/ereader/bibliouan/70030?page=69>
- Oracle. (2020). *Oracle*. Obtenido de Oracle: <https://www.oracle.com/co/database/what-is-database/>
- Oracle. (2021). *Oracle*. Obtenido de Oracle: <https://www.oracle.com/co/database/what-is-database/>
- PowerData. (2014). *PowerData*. Obtenido de PowerData:
<https://www.powerdata.es/data-governance>
- Saa, M. S. (29 de 09 de 2021). *Supersalud*. Obtenido de Supersalud:
<https://superargo.supersalud.gov.co/2/bodega/2021/21000/docs/2021210000142488100002.pdf?v=72027>
- Supersalud. (01 de 09 de 2014). *Supersalud*. Obtenido de Supersalud:
<https://docs.supersalud.gov.co/PortalWeb/planeacion/AdministracionSIG/AUPDO1.docx>
- Supersalud. (10 de 07 de 2020). *Supersalud*. Obtenido de Supersalud:
<https://www.supersalud.gov.co/es-co/nuestra-entidad/estructura-organica-y-talento-humano/mision-y-vision>
- Supersalud. (27 de 08 de 2021). *Supersalud*. Obtenido de Supersalud:
<https://www.supersalud.gov.co/es-co/atencion-ciudadano/contactenos>
- Team, A. P. (Octubre de 2019). *Desarrollar una práctica de análisis de datos a lo largo del ciclo de vida de los datos*. Obtenido de Desarrollar una práctica de análisis de datos a lo largo del ciclo de vida de los datos.:
<https://aws.amazon.com/es/blogs/publicsector/building-a-data-analytics-practice-across-the-data-lifecycle/>
- USC. (29 de 04 de 2021). *USC*. Obtenido de USC:
<https://www.uscmarketingdigital.com/como-hacer-un-dashboard-con-power-bi/>
- whatis. (2019). *My PHP.net*. Obtenido de My PHP.net:
<https://www.php.net/manual/es/intro-whatism.php>
- Wilke., C. O. (Abril de 2019). *Fundamentos de Visualización de Datos*. Obtenido de Fundamentos de Visualización de Datos:
<https://clauswilke.com/dataviz/index.html>

8 Anexos

A continuación, se relacionan tablas, diccionario de datos y flujogramas del proyecto.

8.1 Anexo-1 Diccionario de datos origen PQRD

Tabla 20. Diccionario de Datos PQRD.

Campo	Formato	Descripción
Año	Numero	Año en que se interpone la PQRD
Mes	Numero	Mes en que se interpone la PQRD
Trimestre	Texto	Trimestre en que se interpone la PQRD
Fecha Creacion	Fecha	Fecha en que se interpone la PQRD
PQR Canal	Texto	Canal por el cual se interpone la PQRD
PQR Tipo Peticion	Texto	Tipología de la PQRD
Clase PQR	Texto	Clasificación de la PQRD
Tipo Persona	Texto	Tipo de persona que interpone la PQRD
Codigo Departamento	Numero	Código DANE del Departamento en el que se encuentra la persona que interpone la PQRD (Peticionario)
Departamento	Texto	Departamento en el que se encuentra la persona que interpone la PQRD (Peticionario)
Codigo Municipio	Numero	Código DANE del Municipio en el que se encuentra la persona que interpone la PQRD (Peticionario)
Municipio	Texto	Municipio en el que se encuentra la persona que interpone la PQRD (Peticionario)
Tipo Persona	Texto	Tipo de persona del afectado (natural o Jurídica)
Parentesco	Texto	Relación entre el peticionario y el afectado
Genero	Texto	Sexo del afectado
Rango Edad	Texto	Rango de edad en el que se encuentra el afectado
Nivel Academico	Texto	Nivel de estudios del afectado
Regimen Afiliacion	Texto	Régimen de afiliación del afectado
Grupo Etnico	Texto	Grupo Étnico del afiliado (si aplica)
Grupo Poblacion	Texto	Grupos poblacional de protección especial del afectado (si aplica)
Codigo Departamento	Numero	Código DANE del Departamento en el que se encuentra la persona afectada o que se le vulnera sus derechos a la salud (Afectado)
Departamento	Texto	Departamento en el que se encuentra la persona afectada o que se le vulnera sus derechos a la salud (Afectado)
Codigo Municipio	Numero	Código DANE del Municipio en el que se encuentra la persona afectada o que se le vulnera sus derechos a la salud (Afectado)
Municipio	Texto	Municipio en el que se encuentra la persona afectada o que se le vulnera sus derechos a la salud (Afectado)
Nombre Entidad	Texto	Nombre de la ENTIDAD/VIGILADO sobre la cual se interpone la PQRD
Tipo Regimen	Texto	Clasificación por grupos de vigilados de la ENTIDAD/VIGILADO sobre la cual se interpone la PQRD
Codigo Entidad Salud	Texto	Código asignado por la Superintendencia a la ENTIDAD/VIGILADO sobre la cual se interpone la PQRD (si aplica)
Alias Entidad	Texto	Nombre corto asignado por la Superintendencia a la ENTIDAD/VIGILADO sobre la cual se interpone la PQRD
Codigo Departamento	Numero	Código DANE del Departamento en el que se encuentra la ENTIDAD/VIGILADO sobre la cual se interpone la PQRD
Departamento	Texto	Departamento en el que se encuentra la ENTIDAD/VIGILADO sobre la cual se interpone la PQRD
Código Municipio	Numero	Código DANE del Municipio en el que se encuentra la ENTIDAD/VIGILADO sobre la cual se interpone la PQRD
Municipio	Texto	Municipio en el que se encuentra la ENTIDAD/VIGILADO sobre la cual se interpone la PQRD
Codigo Macromotivo	Numero	Código para la categoría del MACROMOTIVO en el cual se encuentra el motivo específico por el cual se interpuso la PQRD

Macromotivo	Texto	Categoría del MACROMOTIVO en el cual se encuentra el motivo específico por el cual se interpuso la PQRD
Codigo Motivo General	Numero	Código para la categoría del MOTIVO GENERAL en el cual se encuentra el motivo específico por el cual se interpuso la PQRD/
Motivo General	Texto	Categoría del MOTIVO GENERAL en el cual se encuentra el motivo específico por el cual se interpuso la PQRD
Codigo Especifico	Numero	Código del MOTIVO ESPECIFICO por el cual se interpuso la PQRD
Motivo Especifico	Texto	MOTIVO ESPECIFICO por el cual se interpuso la PQRD
Patología	Texto	Descripción de la enfermedad relacionada con el motivo por el cual se interpuso la PQRD
Codigo Patología	Numero	Identificador asignado a la descripción de la enfermedad relacionada con el motivo por el cual se interpuso la PQRD
Tipo Patología	Texto	Detalle de la enfermedad relacionada con el motivo por el cual se interpuso la PQRD
CIE_10	Texto	Descripción de la enfermedad relacionada con el motivo por el cual se interpuso la PQRD de acuerdo con la nomenclatura internacional para la clasificación de enfermedades CIE-10
Alto Costo	Texto	Descripción que relaciona si el motivo por el cual se interpuso la PQRD está relacionado con una enfermedad de alto costo (si aplica)
Riesgo Vida	Texto	Indicador que relaciona si el motivo por el cual se interpuso la PQRD pone en riesgo de vida del paciente
PQRD Estado	Texto	Estado relacionado con la gestión del trámite para la PQRD
Tipo Atencion	Texto	Tipo de atención sobre el cual se recepciona la PQRD
Nombre Regional	Texto	Descripción del lugar en el cual se recepciona la PQRD de acuerdo con sedes regionales dispuestas por la Superintendencia Nacional de Salud en su organización en el territorio colombiano

Fuente: Elaboración Propia

8.2 Anexo-2 Metadatos PQRD

Tabla 21. Metadatos Descriptivos PQRD.

Metadatos descriptivos		
No.	Metadato	Descripción
1	Año	Año en que se interpone la PQRD
2	Mes	Mes en que se interpone la PQRD
3	Trimestre	Trimestre en que se interpone la PQRD
4	Fecha Creacion	Fecha en que se interpone la PQRD
5	PQR Canal	Canal por el cual se interpone la PQRD
6	PQR Tipo Peticion	Tipología de la PQRD
7	Clase PQR	Clasificación de la PQRD
8	Tipo Persona	Tipo de persona que interpone la PQRD
9	Codigo Departamento PQRD	Código DANE del Departamento en el que se encuentra la persona que interpone la PQRD (Peticionario)
10	Departamento PQRD	Departamento en el que se encuentra la persona que interpone la PQRD (Peticionario)
11	Codigo Municipio PQRD	Código DANE del Municipio en el que se encuentra la persona que interpone la PQRD (Peticionario)
12	Municipio PQRD	Municipio en el que se encuentra la persona que interpone la PQRD (Peticionario)
13	Tipo Persona	Tipo de persona del afectado (natural o Jurídica)
14	Parentesco	Relación entre el peticionario y el afectado
15	Genero	Sexo del afectado
16	Rango Edad	Rango de edad en el que se encuentra el afectado
17	Nivel Academico	Nivel de estudios del afectado
18	Regimen Afiliación	Régimen de afiliación del afectado
19	Grupo Etnico	Grupo Étnico del afiliado (si aplica)
20	Grupo Población	Grupos poblacional de protección especial del afectado (si aplica)
21	Código Departamento del Afectado	Código DANE del Departamento en el que se encuentra la persona afectada o que se le vulnera sus derechos a la salud (Afectado)

22	Departamento del Afectado	Departamento en el que se encuentra la persona afectada o que se le vulnera sus derechos a la salud (Afectado)
23	Codigo Municipio del Afectado	Código DANE del Municipio en el que se encuentra la persona afectada o que se le vulnera sus derechos a la salud (Afectado)
24	Municipio del Afectado	Municipio en el que se encuentra la persona afectada o que se le vulnera sus derechos a la salud (Afectado)
25	Nombre Entidad	Nombre de la ENTIDAD/VIGILADO sobre la cual se interpone la PQRD
26	Tipo Regimen	Clasificación por grupos de vigilados de la ENTIDAD/VIGILADO sobre la cual se interpone la PQRD
27	Codigo Entidad Salud	Código asignado por la Superintendencia a la ENTIDAD/VIGILADO sobre la cual se interpone la PQRD (si aplica)
28	Alias Entidad	Nombre corto asignado por la Superintendencia a la ENTIDAD/VIGILADO sobre la cual se interpone la PQRD
29	Codigo Departamento Entidad	Código DANE del Departamento en el que se encuentra la ENTIDAD/VIGILADO sobre la cual se interpone la PQRD
30	Departamento Entidad	Departamento en el que se encuentra la ENTIDAD/VIGILADO sobre la cual se interpone la PQRD
31	Codigo Municipio Entidad	Código DANE del Municipio en el que se encuentra la ENTIDAD/VIGILADO sobre la cual se interpone la PQRD
32	Municipio Entidad	Municipio en el que se encuentra la ENTIDAD/VIGILADO sobre la cual se interpone la PQRD
33	Codigo Macromotivo	Código para la categoría del MACROMOTIVO en el cual se encuentra el motivo específico por el cual se interpuso la PQRD
34	Macromotivo	Categoría del MACROMOTIVO en el cual se encuentra el motivo específico por el cual se interpuso la PQRD
35	Codigo Motivo General	Código para la categoría del MOTIVO GENERAL en el cual se encuentra el motivo específico por el cual se interpuso la PQRD/
36	Motivo General	Categoría del MOTIVO GENERAL en el cual se encuentra el motivo específico por el cual se interpuso la PQRD
37	Codigo Especifico	Código del MOTIVO ESPECIFICO por el cual se interpuso la PQRD
38	Motivo Especifico	MOTIVO ESPECIFICO por el cual se interpuso la PQRD
39	Patologia	Descripción de la enfermedad relacionada con el motivo por el cual se interpuso la PQRD
40	Codigo Patologia	Identificador asignado a la descripción de la enfermedad relacionada con el motivo por el cual se interpuso la PQRD
41	Tipo Patología	Detalle de la enfermedad relacionada con el motivo por el cual se interpuso la PQRD
42	CIE_10	Descripción de la enfermedad relacionada con el motivo por el cual se interpuso la PQRD de acuerdo con la nomenclatura internacional para la clasificación de enfermedades CIE-10
43	Alto Costo	Descripción que relaciona si el motivo por el cual se interpuso la PQRD está relacionado con una enfermedad de alto costo (si aplica)
44	Riesgo Vida	Indicador que relaciona si el motivo por el cual se interpuso la PQRD pone en riesgo de vida del paciente
45	PQRD Estado	Estado relacionado con la gestión del trámite para la PQRD
46	Tipo Atencion	Tipo de atención sobre el cual se recepciona la PQRD
47	Nombre Regional	Descripción del lugar en el cual se recepciona la PQRD de acuerdo con sedes regionales dispuestas por la Superintendencia Nacional de Salud en su organización en el territorio colombiano

Fuente: Elaboración Propia

8.3 Anexo-3 Lineamientos y Estándares.

Tabla 22. Lineamientos y estándares para el proyecto.

Lineamiento	Descripción
Creación de ambientes	Se deben crear 2 ambientes, Stage para pruebas y procesos de ajustes y Producción para contener los datos con calidad y listos para disponer.
Nombres de instancias y schemas	El nombre debe comenzar con la primera letra del ambiente en el que se encuentra, seguido del nombre del motor de BD y finalmente 4 siglas descriptivas del tema del proyecto: “ smysqlpqr d” y “ pmysqlpqr d”.
Nombre de llaves primarias y foráneas	Las llaves PK y FK deben comenzar con el sufijo “ cod_ ”.
Contraseñas de acceso	Deben tener al menos 8 caracteres, contener al menos una mayúscula, un carácter especial y números y debe encriptarse como un HASH.
Puertos de acceso a las Bases de Datos diferentes al default	Deben cambiarse los puertos por defecto al 3306 del SGBD MySQL.
Creación usuarios y roles en el SGBD	Se deben crear además del usuario administrador, 2 adicionales en base a funciones de DML y DDL: “ user_ddl_pqr d” y “ user_dml_pqr d”.
Monitoreo y control en los datos	Se debe monitorear el funcionamiento del SGBD en cada uno de los ambientes e instancias creadas en las nubes AWS y Microsoft Azure para control de fallos y ejecución de backups semanales.
Bitácora para almacenamiento de diccionario de datos, metadatos y el modelo de datos	Se debe crear una ubicación de almacenamiento donde se centralicen los Metadatos, Diccionario de datos, Modelos de datos y demás componentes relevantes.
Índices en tablas	Las tablas del modelo de Producción deben contener índices.

Fuente: Elaboración Propia

8.4 Anexo-4 Modelo Flujo Datos Lógico

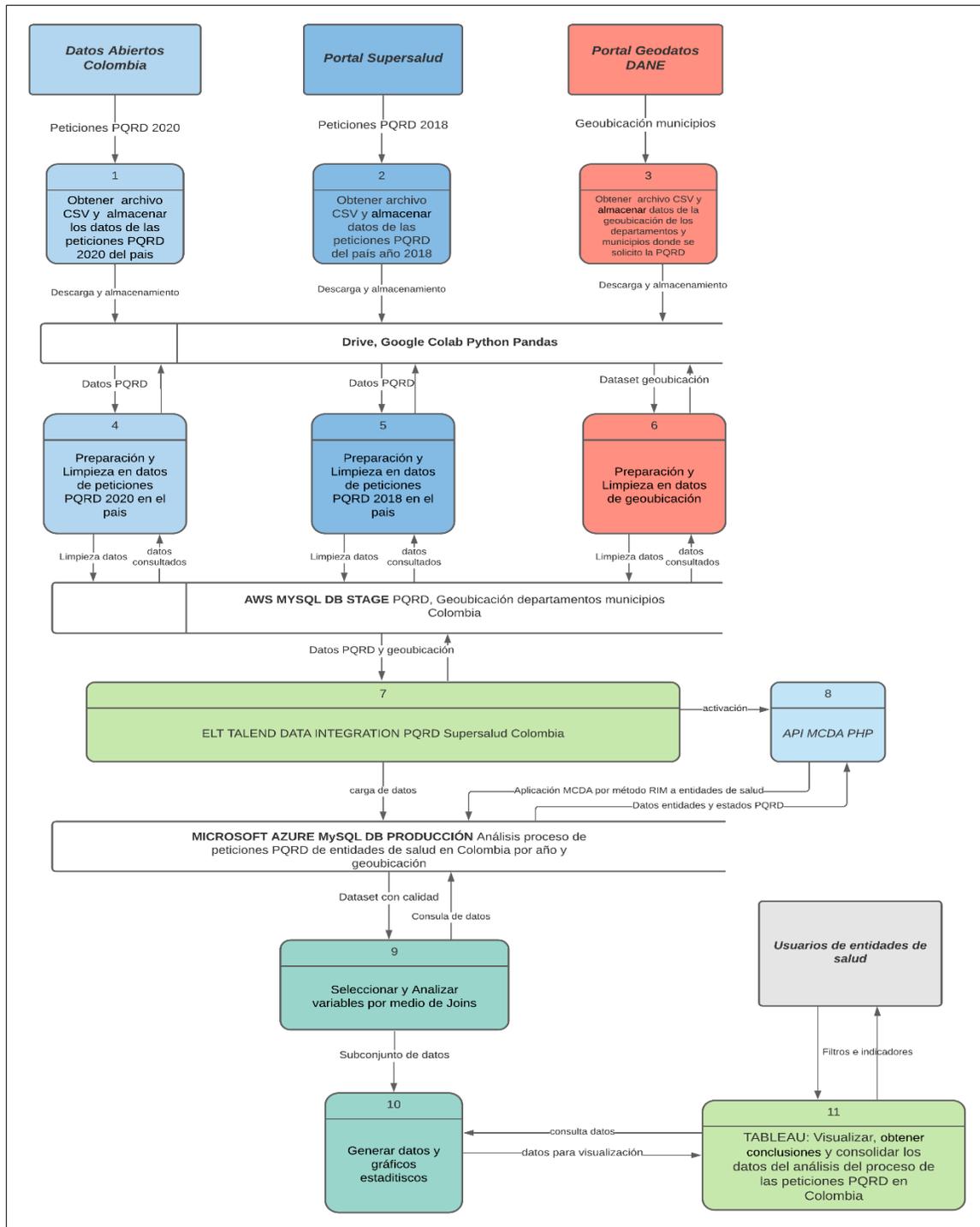
Figura 58. Modelo Flujo Datos Lógico



Fuente: Elaboración Propia

8.5 Anexo-5 Flujo de Datos Físico

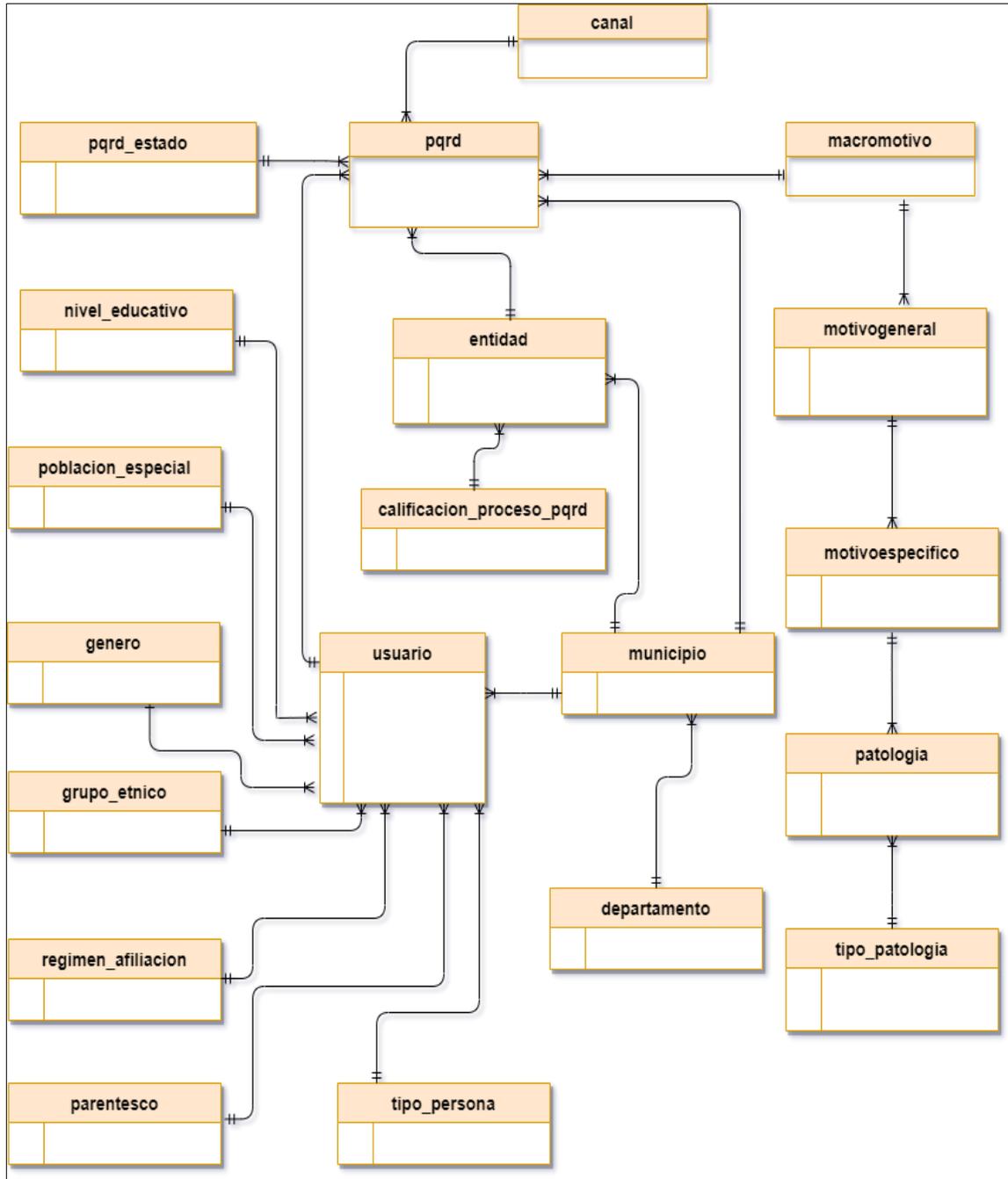
Figura 59. Flujo de Datos Físico



Fuente: Elaboración Propia

8.6 Anexo-6 Modelo Conceptual

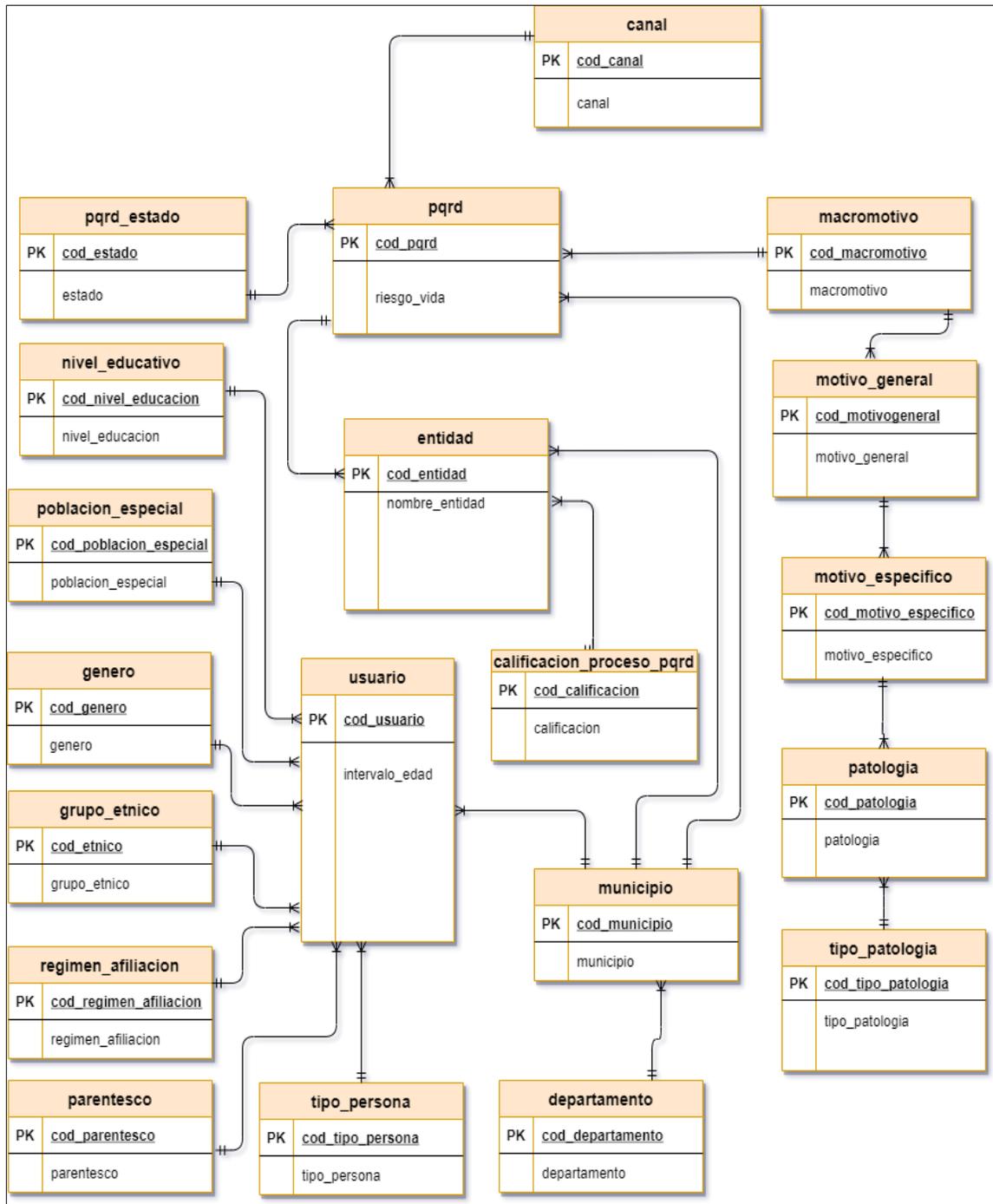
Figura 60. Modelo Conceptual



Fuente: Elaboración Propia

8.7 Anexo-7 Modelo Lógico Base de Datos

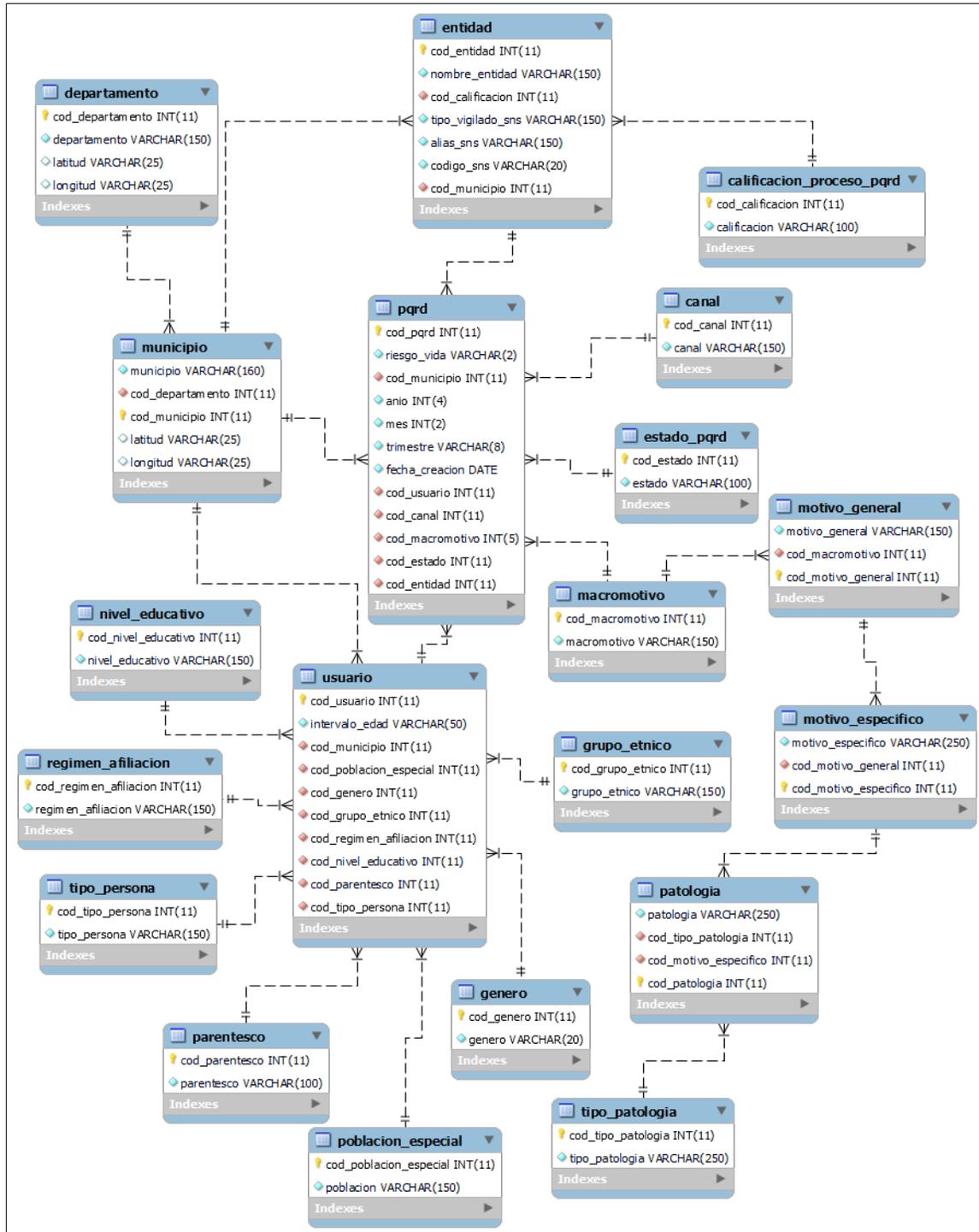
Figura 61. Modelo Lógico Base de Datos



Fuente: Elaboración Propia

8.8 Anexo-8 Modelo Físico Base de Datos

Figura 62. Modelo Físico Base de Datos



Fuente: Elaboración Propia