**Análisis de Base de Datos**

**Objetivo:**  
Cada grupo deberá seleccionar una base de datos, realizar un análisis exploratorio básico, comprender su estructura y contenido, y presentar un informe con los hallazgos más relevantes.

Cada grupo debe realizar las siguientes tareas con la base de datos seleccionada:

1. **Selección de la Base de Datos**

* Nombre de la base de datos**: Cifras Reclamos No Fraude.**
* Fuente (URL, repositorio, etc.): **Reclamaciones Bancarias. - https://www.banrep.gov.co/es/transparencia/informe-gestion-pqr-primer-trimestre-2025**
* Formato del archivo (.csv, .xlsx.): **Hoja de cálculo de Microsoft Excel (.xlsx)**

1. **Estructura de la Base de Datos**

* Número total de filas (registros): **2511 Registros**
* Número total de columnas (atributos o variables): **25 Columnas**

1. **Descripción de las Columnas**

Por cada columna se debe especificar:

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Nombre de la columna** | **Tipo de dato** | **Descripción** | **Cantidad de datos** | **Valores únicos (si aplica)** |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **VER COMENTARIOS DE LA BBDD** |  |  |  |  |

1. **Calidad de los Datos**

* ¿Hay valores faltantes en alguna columna? ¿Cuántos? Si, en las siguientes columnas:
* **USUARIO RADICADOR: 13 Registros**
* **SEGMENTO\_COMERCIAL: 25 Registros**
* ¿Hay columnas con valores repetidos o inconsistentes?

Si, todas las columnas contienen datos repetidos, excepto la columna NUMERO\_RECLAMO, se evidencia que este número es único para los 2511 registros

* ¿Existen datos fuera de rango o errores evidentes?

No se evidencias errores

* ¿Conoce la última fecha de actualización de los datos?

Existe la columna FECHA\_CIERRE, en esta se evidencia que la última actualización fue el 31/03/2025

1. **Análisis Cuantitativo Básico**

Para columnas numéricas:

* Mínimo:

**Columnas**

**DIAS\_CALENDARIO: Mínimo 0**

**DIAS\_HABILES: Mínimo 0**

* Máximo:

**Columnas**

**DIAS\_CALENDARIO: Máximo 77**

**DIAS\_HABILES: Mínimo 54**

* Media:

**Columnas**

**DIAS\_CALENDARIO: 13,38**

**DIAS\_HABILES: 8,65**

* Mediana:

**DIAS\_CALENDARIO: 13,00**

**DIAS\_HABILES: 8,00**

* Desviación estándar:

**DIAS\_CALENDARIO:7,80**

**DIAS\_HABILES: 5,39**

Para columnas categóricas:

* Frecuencia de cada categoría, COMPARTO EJEMPLOS.

**Ejemplo Frecuencia Absoluta por tipo de solicitud según la BBDD**

|  |  |  |
| --- | --- | --- |
| FRECUENCIA ABSOLUTA | |  |
|  |  |  |
| MESA | (Todas) |  |
|  |  |  |
| Etiquetas de fila | Cuenta de NUMERO\_SS | % Frecuencia Absoluta |
| Reclamación Super | 1773 | 70,6% |
| Empresas y/o PNCN | 209 | 8,3% |
| Solicitud | 194 | 7,7% |
| Novedad Monetaria | 187 | 7,4% |
| Calidad del servicio Super | 140 | 5,6% |
| Solicitud interna | 8 | 0,3% |
| **Total general** | **2511** | 100,0% |

**Ejemplo Frecuencia Absoluta canal origen según la BBDD**

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| MESA | (Todas) |  |  |  |  |  |
|  |  |  |  |  |  |  |
| **Cuenta de NUMERO\_SS** | Etiquetas de columna |  |  |  |  |  |
| Etiquetas de fila | 26/03/2025 | 27/03/2025 | 28/03/2025 | 29/03/2025 | 31/03/2025 | Total general |
| Alcance | 2 |  | 5 |  | 2 | 9 |
| Clientes | 577 | 595 | 542 | 10 | 505 | 2229 |
| Defensoría | 6 | 10 | 2 |  | 5 | 23 |
| Derecho de petición | 54 | 62 | 57 |  | 52 | 225 |
| Superfinanciera | 4 | 11 | 10 |  |  | 25 |
| **Total general** | **643** | **678** | **616** | **10** | **564** | **2511** |

* Número de categorías únicas.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| MESA | (Todas) |  |  |  |  |  |
|  |  |  |  |  |  |  |
| **Cuenta de NUMERO\_SS** | Etiquetas de columna |  |  |  |  |  |
| Etiquetas de fila | 26/03/2025 | 27/03/2025 | 28/03/2025 | 29/03/2025 | 31/03/2025 | Total general |
| Cerrado | 643 | 678 | 616 | 10 | 564 | 2511 |
| **Total general** | **643** | **678** | **616** | **10** | **564** | **2511** |

1. **Conclusiones Iniciales del Grupo**
2. ¿Qué información relevante puede extraerse de esta base de datos?

Existen los motivos por los cuales reclaman los clientes

Se tienen fechas de apertura y de cierre de los reclamos

Se tienen las personas responsables de atender los reclamos

Están los medios en los cuales se enviaron las respuestas a los clientes

Están los días calendario y hábiles que se demoró en atención del reclamo

Existe el segmento de clientes

Existen números únicos para reconocer los reclamos

1. ¿Qué problemas de calidad o estructura podrían afectar un análisis más profundo?

Se evidencia que la columna causa origen tiene muchos motivos, por ende al realizar un análisis de fondo de pueden encontrar diferentes causas raíz del motivo por el cual reclaman los clientes.

1. ¿Qué preguntas podría ayudar a responder esta base de datos?

Segmentación de los clientes

Fechas de apertura y cierre

Medio de respuesta

Motivo del reclamo

Responsable del reclamo

1. ¿Qué mejoras o limpiezas serían necesarias?

Hay columnas repetidas, las cuales se podrían simplificar porque contienen los mismos datos

Columna L y P

1. **Entrega**

Cada grupo debe entregar este documento diligenciado y, si es posible, el archivo de base de datos con sus observaciones o anotaciones.

CRISTIAN CAMILO DURAN GIRALDO