**Realizado por:**

**Juan Camilo Restrepo Velez**

**William Leonardo Andrade Collazos**

**Wilder Valencia Ocampo**

**PRÁCTICA DE CALIDAD DE DATOS 10%**

**DIMENSIONES DE LA CALIDAD DE DATOS**

1. **Completitud:** ¿Está toda la información disponible? ¿Hay datos faltantes o ausentes?
2. **Exactitud:** ¿La información es correcta y libre de error?
3. **Conformidad:** ¿Los valores de los datos están conformes con los formatos esperados? Ejemplo: Una fecha en formato AAAA/MM/DD cuando debería ser DD/MM/AAAA.
4. **Oportunidad:** ¿La información llega cuando se necesita?
5. **Duplicidad:** ¿Existen múltiples instancias, innecesarias de los mismos objetos de datos en el conjunto de datos?
6. **Integridad:** ¿Faltan datos relacionados importantes? ¿Es clara la conectividad y las relaciones con otros datos?
7. Genere el reporte de PYTHON sobre los datos **originales**, adjunte el resultado en HTML.
8. Evalúe cada una de las dimensiones de la calidad de datos, teniendo en cuenta los resultados de PYTHON.
9. Comparar los problemas de calidad encontrados en el perfilado con la preparación de datos realizada en el proyecto de Minería. Indicar los nuevos errores encontrados en los datos.

Entregables:

* Reporte de perfilado en HTML
* Evaluación de las dimensiones de la calidad de datos
* Reporte de los nuevos errores encontrados.