

DQ Problems	DQ Dimension	DQ Factor	DQ Metric	DQ method	Applied DQ method
	<p>ID: D4_Unicidad Name: Unicidad Description: Evalúa si los registros en un conjunto de datos son únicos y no presentan duplicaciones no deseadas Suggested by = {RN2, RN5}</p>	<p>ID: Name: No duplicación Description: Porcentaje de datos que no están duplicados en forma exacta Represents = {RN2, RN5}</p>	<p>ID: M_duplicate_ratio Name: duplicate_ratio Description: Da el porcentaje de valores duplicados. Influenced by = {RN2, RN5} Granularity: Columna Result domain = [0...1]</p>	<p>ID: metodo_duplicate_ratio Name: met_duplicate_ratio Description: Implementa la metrica duplicate_ratio sobre un atributo dado. Uses = {RN2, RN5} Input data types: String Output data types: Float Algorithm: <pre>"def duplicate_ratio(data, column): total = len(data) duplicated = data.duplicated(subset=[column]).sum() return duplicated / total"</pre> </p>	<p>ID: MA_duplicate_ratio Type: Medición Descriptinon: Dados los datos y el atributo, calcula el porcentaje de valores duplicados en ese atributo. AppliedTo: Atributos «ISBN», «AuthorID», «PublisherID» de la tabla NL_Books. Atributos «ID» de la tabla NL_USERS.</p>

DQ Problems	DQ Dimension	DQ Factor	DQ Metric	DQ method	Applied DQ method
P11	<p>ID: D2_Completitud Name: Completitud Description: Indica si el sistema de información contiene toda la información de interés. Suggested by = {RN1, RN2, RN5, RQ3, RQ7, F2}</p>	<p>ID: F2_Dens Name: Densidad Description: Indica cuanta información se tiene y cuanta falta sobre las entidades del sistema de información. Represents = {RN1, RN2, RN5, RQ3, RQ7, F2}</p>	<p>ID: M_contar_nulls Name: contar_nulls Description: Calcula el porcentaje de entradas vacias en una columna Influenced by = {RN1, RN5, RQ3, RQ7, F2} Granularity: Columna Result domain = [0...1]</p>	<p>ID: metodo_contar_nulls Name: met_contar_nulls Description: Implementa la metrica contar_nulls. Uses = {RN1, RN5, RQ3, RQ7, F2} Input data types: String Output data types: Float Algorithm: <pre>"def porcentaje_nulos(atributo): nulos = atributo.isnull() contador_nulos = nulos.sum() porcentaje = contador_nulos / len(atributo) return porcentaje"</pre> </p>	<p>ID: MA_contar_nulls Type: Medición Descriptinon: Dado un atributo de una tabla, calcula el porcentaje de entradas vacias. AppliedTo: Cualquier atributo de cualquier tabla.</p>
			<p>ID: M_check_RN1 Name: check_RN1 Description: Da el porcentaje de entradas de la tabla que tienen al menos un campo vacío entre los atributos isbn, título, autor y editor. Influenced by = {RN1, RN2, RQ3, RQ7, F2} Granularity: Conjunto de columnas Result domain = [0...1]</p>	<p>ID: metodo_check_RN1 Name: met_check_RN1 Description: Implementa la metrica Check_RN1 sobre una tabla dada. Uses = {RN1, RN2, RQ3, RQ7, F2} Input data types: String Output data types: Float Algorithm: <pre>"def Check_RN1(datos): # Seleccionar solo las columnas relevantes (isbn, título, autor, editor) datos_relevantes = datos[['isbn', 'título', 'autor', 'editor']] # Verificar cuántas filas tienen al menos un campo NULL entre estas columnas incompletos = datos_relevantes.isnull().any(axis=1).sum() # Calcular la proporción de filas incompletas return incompletos / len(datos)"</pre> </p>	<p>ID: MA_check_RN1 Type: Medición Descriptinon: Dada una tabla, calcula el porcentaje de entradas que tienen al menos un campo vacío entre los atributos isbn, título, autor y editor. AppliedTo: Tabla NL_Books.</p>