



Ciclo 1 Fundamentos de programación

Reto 4

Descripción del problema:

El departamento de estadística del ICFES lo ha contratado a usted como científico de datos para la limpieza y transformación de datos alojado en un DataFrame con el fin de utilizar estos datos para el desarrollo de un nuevo modelo de machine learning que permita realizar mejores estimaciones estadísticas para poder comprender mejor los resultados de los estudiantes de bachillerato de 11 grado y aplicar medidas que mejoren la realización de las Pruebas Saber 11 con el fin de evaluar y mejorar la educación del país.

La información recogida está organizada en un DataFrame donde los nombres de las columnas son: Resultados, ID, SEXO, donde la columna Resultados se refiere a los parámetros de entrada para el modelo de machine learning el cual solo acepta valores enteros positivos, ID, representa el identificador del estudiante, y SEXO que se refiere al sexo del estudiante.

Usted debe construir una función que, dados los resultados de los estudiantes, calcule los parámetros válidos que sirvan de entrada para el entrenamiento del modelo de machine learning, tenga en cuenta que los datos necesitan limpieza, es decir, no todos los valores son válidos para el modelo, por tanto, los parámetros entrenamiento se calculan de la siguiente manera:

- Se toman únicamente los valores que sean enteros positivos.
- A los valores obtenidos anteriormente se les calcula su cuadrado correspondiente, por ejemplo, si el número es 7 su cuadrado correspondiente es 49.

La función debe retornar un diccionario. Para las llaves que debe tener: “parametros” cuyo valor es una lista que contiene los parámetros transformados, valores los cuales se les halló el cuadrado correspondiente, “cantidad_errores” que indica la cantidad de valores que no fueron tomados en cuenta para la inteligencia artificial, “errores” que indica los valores, en



una lista, que no fueron tomados en cuenta para el entrenamiento del modelo. Si ningún valor pudo ser transformado entonces el valor para la llave parámetros será un diccionario vacío.

Nota: Queda **prohibido** el uso de cualquier estructura cíclica, por el contrario, use las funciones de orden superior: map, filter, reduce, lambda.

Ejemplo:

	Resultados	ID	SEXO
0	1	5595	F
1	2	2603	F
2	3	4865	F
3	4	1502	M
4	5	7191	M
5	6	6341	M
6	7	8141	M
7	-2	4620	F
8	-3	5333	M
9	4.9	2725	F
10	No hay registro	8497	M
11		8197	M

(Ejemplo 1)

El retorno de realizar las respectivas transformaciones al DataFrame debe ser:

```
{'parametros': [1, 4, 9, 16, 25, 36, 49], 'cantidad_errores': 5, 'errores': [-2, -3, 4.9, 'No hay registro', '']}
```



	Resultados	ID	SEXO
0	51	982	F
1	32	193	M
2	29	123	M
3	-1	3	M
4	-1230	491	M
5	12.3	1912	F
6	59	1432	F
7		9012	F
8		8129	M
9	No hay registro	4	F

(Ejemplo 2)

El retorno de realizar las respectivas transformaciones al DataFrame debe ser:

```
{'parametros': [2601, 1024, 841, 3481], 'cantidad_errores': 6, 'errores': [-1, -1230, 12.3, '', ''],  
'No hay registro']}
```

Entradas:

Nombre	Tipo	Descripción
data	pandas.DataFrame	DataFrame que contiene la información de los resultados parcialmente transformados del departamento de Estadística del ICFES

Salidas:

Tipo del retorno	Descripción
dict	Un diccionario con los datos que se le pudo aplicar transformaciones y aquellos que no. Las llaves que debe tener: “parametros” cuyo valor es una lista que contiene los parámetros



	transformados, valores los cuales se les halló el cuadrado correspondiente, “cantidad_errores” que indica la cantidad de valores que no fueron tomados en cuenta para la inteligencia artificial, “errores” que indica los valores, en una lista, que no fueron tomados en cuenta para el entrenamiento del modelo.
--	---

Esqueleto:

```
def lista_cuadrados(data)->dict:  
    #data: DataFrame que contiene la información de los resultados  
    parcialmente transformados del departamento de Estadística del ICFES
```