



UNIVERSIDAD CES

Un compromiso con la excelencia

VIGILADA MINEUCACIÓN

## INSTRUCTIVO PARA GARANTIZAR LA CALIDAD DEL DATO

Código: IN-DE-004

Fecha: 21/05/2024

Versión: 01

PROCESO

Direccionamiento estratégico

### 1. OBJETIVO:

Describir las validaciones destinadas a auditar la calidad del dato, garantizando que la información institucional sea precisa, confiable y útil para la toma de decisiones.

### 2. DEFINICIONES:

- **Atributos de los datos:** Establece aspectos y comportamientos de los datos, determina el nombre, el tipo de dato y sus especificaciones.
- **Compleitud:** deben existir valores asignados en el dato. Todas las filas del conjunto de datos deben estar presentes. Asignación de atributos de restricción u obligatoriedad.
- **Consistencia:** Esta dimensión se asegura de que la misma información que ha sido almacenada y utilizada en múltiples ocasiones coincida.
- **Exactitud:** La precisión de datos es el nivel el cual la información coincide con las métricas establecidas en el mundo real y confirma si vienen de una fuente válida. Los datos de la base pueden ser comparados con otros de una base de datos similares y deben ser los mismos.
- **Fuente Primaria:** Contiene información nueva u original.
- **Integridad:** se refiere a que todas las referencias o relaciones de la tabla a la que pertenece el dato (campo) son válidas en otros conjuntos de datos con los que se pueda relacionar.
- **Singularidad:** Ninguna instancia de entidad se registrará más de una vez con las mismas características. Hace referencia a que ninguna entidad de datos exista más de una vez dentro de un conjunto de datos.
- **Validez:** El dato es válido si está conforme al tipo de dato esperado, rango o estándar según se haya definido. Se refiere a si las instancias de datos se almacenan, intercambian, o si se presentan en un formato que es válido para el dominio de los valores.



**UNIVERSIDAD CES**

Un compromiso con la excelencia

VIGILADA MINEOLUCACION

## INSTRUCTIVO PARA GARANTIZAR LA CALIDAD DEL DATO

Código: IN-DE-004

Fecha: 21/05/2024

Versión: 01

**PROCESO**

**Direccionamiento estratégico**

### 3. CONTENIDO:

#### Observaciones:

- Las facultades y otras dependencias académicas y administrativas asumen la responsabilidad de garantizar la calidad de los datos ingresados en las diversas fuentes de información.
- Para construir una base de datos. Ver **GU-DE-004 Guía de reglas de calidad del dato.**

#### INSTRUCCIONES:

**Validar duplicados:** Validar que ninguna instancia de entidad principal, se registre más de una vez con las mismas características.

Ejemplo: Revisar duplicados de número de cédula.

Se recomienda usar el formato condicional en Excel para buscar y resaltar datos duplicados. De esa manera se puede revisar duplicados y decidir, si desea eliminarlos.

- Paso 1: Seleccione las celdas en las que desea comprobar duplicados.



Código: IN-DE-004

Fecha: 21/05/2024

Versión: 01

PROCESO

Direccionamiento estratégico

PA	ID_TIPO	NUM_DOCUM	CODIGO_ES	NOMB
	TI	1002523463	1002523463	Stefani
	CC	1233493221	1233493221	Andres
	TI	1031640214	1031640214	Valeria
	CC	1006856241	1006856241	Alvaro
	CC	1000783748	1000783748	Stefani
	CC	1000142829	1000142829	Paula
	CC	1000577200	1000577200	Maria
	TI	1000129274	1000129274	Angie
	TI	1000688453	1000688453	Karen

- Seleccione Inicio > Formato condicional > Resaltar reglas de celdas > valores duplicados.

PROGRAMA	AÑO	SEMESTRE	ID TIPO	NUM I
Maestría en Tecnol	2023	2	CC	7126
Maestría en Psicol	2023	2	CC	7131



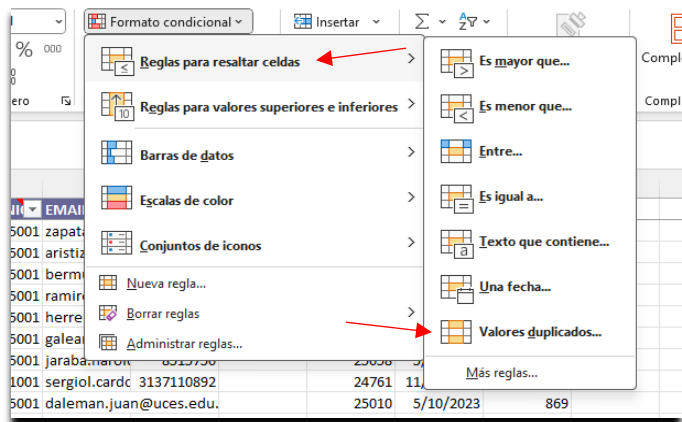
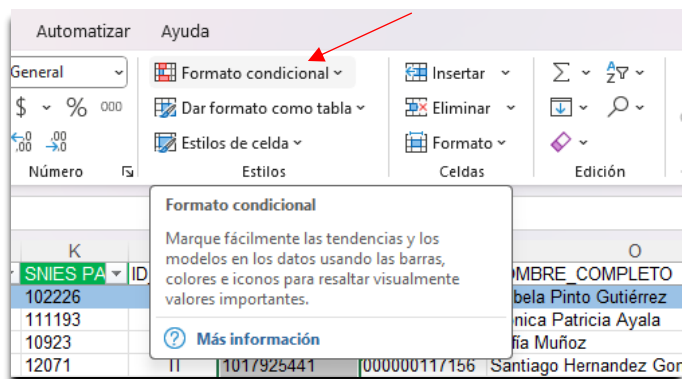
Código: IN-DE-004

Fecha: 21/05/2024

Versión: 01

PROCESO

Direccionamiento estratégico



- Paso 2: En el cuadro que se despliega llamado “valores duplicados” aplicar el formato a las celdas que contiene: elija el formato *(Relleno rojo con texto rojo oscuro)* y, a continuación, seleccione **Aceptar**.



UNIVERSIDAD CES

Un compromiso con la excelencia

VIGILADA MINEUCACION

## INSTRUCTIVO PARA GARANTIZAR LA CALIDAD DEL DATO

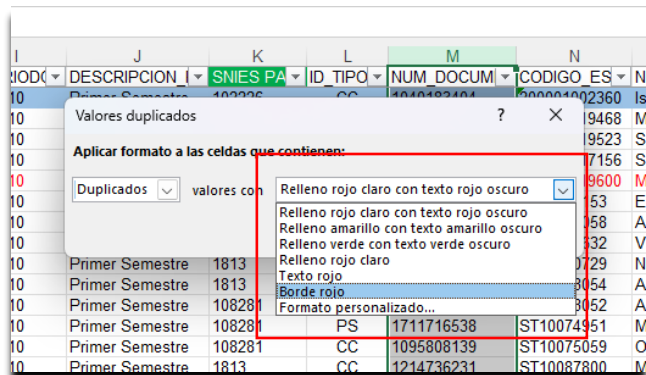
Código: IN-DE-004

Fecha: 21/05/2024

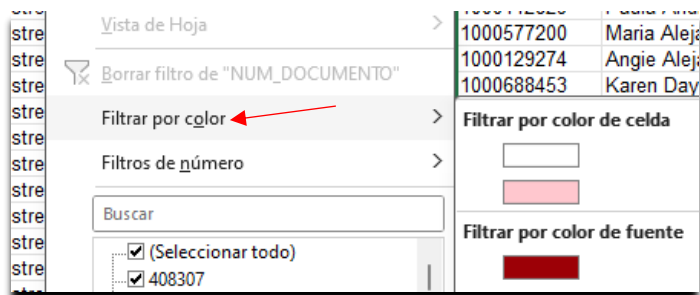
Versión: 01

PROCESO

Direccionamiento estratégico



Paso 3: Luego en la columna que se aplicó la condición, realizar un filtro por color (*Filtrar por color de celda rosada*) para que nos muestre las cédulas duplicadas.





UNIVERSIDAD CES

Un compromiso con la excelencia

VIGILADA MINEUCACION

### INSTRUCTIVO PARA GARANTIZAR LA CALIDAD DEL DATO

Código: IN-DE-004

Fecha: 21/05/2024

Versión: 01

**PROCESO**

**Direccionamiento estratégico**

Ordenar la columna de menor a mayor para que la vista de duplicados sea más rápida de comparar.

ID_TI	NUM_DOCUMENTO	NOMBRE_COMPLETO
CE	509392	Miguel Ángel Galk
CE	509392	Miguel Ángel Galk
CE	7925636	Karen Danelia He
CE	7925636	Karen Danelia He
CE	7925655	Diana Jimena Ror
CE	7925655	Diana Jimena Ror

Paso 4: Comprobar si el dato duplicado cumple con los criterios establecidos de duplicidad en la instancia de entidad principal y eliminar fila duplicada de la base de datos, si es necesario.

NUM_DOCUMENTO	CODIGO_ES	NOMBRE_COMPLETO
1020103364	ST10077045	Manuela Foronda López
1102802003	ST10084100	Camila Guerra Florian
487476	ST10081828	Nicole Valentina Prado
487476	ST10081828	Nicole Valentina Prado
1066598349	ST10086072	Stephany Paulin Pérez Ba
107417032	ST10087670	Martina Rafaela Barros Gr
994217903	ST10089637	David Leonel Lenis Aragur

Código: IN-DE-004

Fecha: 21/05/2024

Versión: 01

**PROCESO**

**Direccionamiento estratégico**

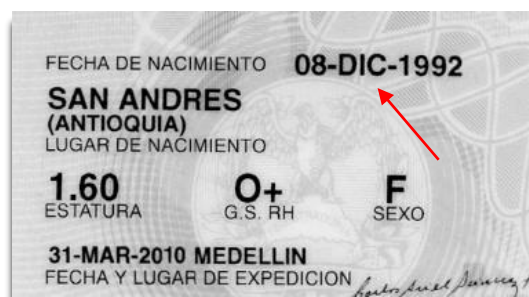
### Validar fuente de información:

Validar que los datos provengan de una fuente válida es fundamental. La precisión de los datos se logra cuando estos coinciden con las métricas establecidas en el mundo real. Para garantizar esta precisión, es importante comparar la información de la base con otras bases de datos similares o con archivos digitales originales, y asegurarse de que sean idénticos.

Ejemplo: Revisar Fecha de nacimiento de un estudiante.

Paso: Revisar la información ingresada en la base de datos y comprobar directamente con la fuente, en este caso el documento de identidad del estudiante.

DO	J	K	ID_M
FECHA_NACIMIENTO	ID_PAIS		
8/12/1992	170		



### Validar formato:

Validar que los datos sean válidos implica asegurarse de que se ajusten al tipo de formato esperado, al rango especificado o a los estándares definidos. Esto se refiere a si las instancias de datos se almacenan, intercambian o se presentan en un formato que sea válido para el dominio de los valores.

Ejemplo: Revisar reglas y formato esperado de estado civil de participante.

Código: IN-DE-004

Fecha: 21/05/2024

Versión: 01

**PROCESO**

**Direccionamiento estratégico**

Regla: Formato establecido de celda= numérico.

Numérico (1).  
Estado civil del participante.  
Los códigos posibles son:

1. Soltero(a)
2. Casado(a)
3. Divorciado(a)
4. Viudo(a)
5. Unión libre
6. Religioso(a)
8. Separado(a)

Paso 1: Crear un diccionario en una pestaña nueva en Excel, con las características relacionadas de “estado civil”, y nombrarla DICCIONARIO.

A	B
Nº Estado Civil	Estado Civil
1	Soltero
2	Casado
3	Divorciado
4	Viudo
5	Unión Libre
6	Religioso
8	Separado





Código: IN-DE-004

Fecha: 21/05/2024

Versión: 01

PROCESO

Direccionamiento estratégico

Paso 2: En una columna continua a la columna AA "estado civil", nombrada "BUSCAR ESTADO", relacione cada celda mediante la formula BUSCARV, con el diccionario creado para la validación . Recuerde que el valor de búsqueda debe estar siempre en la primera columna del rango para que la fórmula= BUSCARV funcione correctamente.

fx =BUSCARV(AA2;[Libro1]Diccionario!\$A:\$B;2;0)

Z	AA	AB	AC
PROGRAMA	ESTADO CIVIL	BUSCAR ESTADO	INCULACION
5001	3	Divorciado	
5001	3		
5001	0		
5001	4		
5001	4		
5001	4		

fx =BUSCARV(AA2;[Libro1]Diccionario!\$A:\$B;2;0)

Z	AA	AB	AC	AD	AE
PROGRAMA	ESTADO CIVIL	BUSCAR ESTADO	INCULACION	ACION_SEXU	UPO_ETNICI
5001	3	2;0)	1	0	0

Argumentos de función

BUSCARV

Valor\_buscado AA2 = 3

Matriz\_tabla [Libro1]Diccionario!\$A:\$B = {...} *Buscar en Diccionario*

Indicador\_columnas 2 = 2

Rango 0 = FALSO

= "Divorciado"

Busca un valor en la primera columna de la izquierda de una tabla y luego devuelve un valor en la mis



UNIVERSIDAD CES

Un compromiso con la excelencia

VIGILADA MINEUCACIÓN

## INSTRUCTIVO PARA GARANTIZAR LA CALIDAD DEL DATO

Código: IN-DE-004

Fecha: 21/05/2024

Versión: 01

PROCESO

Direccionamiento estratégico

Paso 3: En la nueva columna AB nombrada "BUSCAR ESTADO", realizar un filtro de texto, y validar las palabras que no cruzan efectivamente y se registran con el error "#N/D"

A	AB	AC
0 CIV	BUSCAR ESTADO	/INCLACI
3	Divorciado	
3	Divorciado	
0	#N/D	
4	Viudo	
4	Viudo	
4	Viudo	
3	Divorciado	
4	Viudo	
3	Divorciado	

Filtros de texto

Buscar

- (Seleccionar todo)
- Casado
- Divorciado
- Religioso
- Soltero
- Unión Libre
- Viudo
- #N/D



UNIVERSIDAD CES

Un compromiso con la excelencia

VIGILADA MINEUCACIÓN

## INSTRUCTIVO PARA GARANTIZAR LA CALIDAD DEL DATO

Código: IN-DE-004

Fecha: 21/05/2024

Versión: 01

PROCESO

Direccionamiento estratégico

	AA	AB	AC
AN	ESTADO CIV	BUSCAR ESTADO	INCUA
5001	0		#N/D
5001	0		#N/D
5001	0		#N/D
5001	0		#N/D
5001	0		#N/D
5001	0		#N/D
5001	0		#N/D
5001	0		#N/D
5001	0		#N/D

Se observa que el parámetro 0, no está definido en los criterios de selección de estado civil, por lo tanto se debe revisar con la fuente de información y corregir.

### Validar completitud en los datos:

Validar si las reglas de cada variable tienen obligatoriedad, es decir, que exista valores asignados en el dato. Todas las filas del conjunto de datos deben estar presentes.

Ejemplo: Revisar reglas de completitud y obligatoriedad para la variable código prueba saber 11.

Regla 1: Si es un estudiante de pregrado, debe ingresar código de Prueba Saber 11.



**UNIVERSIDAD CES**

Un compromiso con la excelencia

VIGILADA MINEUCACIÓN

## INSTRUCTIVO PARA GARANTIZAR LA CALIDAD DEL DATO

Código: IN-DE-004

Fecha: 21/05/2024

Versión: 01

**PROCESO**

**Direccionamiento estratégico**

### COD PRUEBA SABER 11

Obligatorio si el programa al cual se esta inscribiendo es de modalidad de pregrado.

Alfanumérico (14 dígitos).

Código asignado por el Icfes en la prueba Saber presentada por el estudiante.

Paso 1: Revisar si los datos la columna “W” nivel académico son igual a “pregrado”, y así mismo la columna X cod\_prueba\_saber\_11, debe de tener un código ingresado. Para esta regla vamos a utilizar la formula =SI, que comprueba si se cumple una condición y devuelve un valor como verdadero o falso.

	W	X
	Nivel de Académico	COD_PRUEBA_SABER_11
	Pregrado	EXTRANJERO2024

Se construye la condición en la columna Y, donde la formula ingresada es: Si el nivel académico es pregrado y el código prueba saber 11 es diferente a campo vacío, “cumple la condición”, de lo contrario “no cumple”. Después, copie y pegue la fórmula en otras celdas y se revisa cuales no cumplen.



UNIVERSIDAD CES

Un compromiso con la excelencia

VIGILADA MINEUCACIÓN

### INSTRUCTIVO PARA GARANTIZAR LA CALIDAD DEL DATO

Código: IN-DE-004

Fecha: 21/05/2024

Versión: 01

PROCESO

Direccionamiento estratégico

SI

Prueba\_lógica  = VERDADERO

Valor\_si\_verdadero  = "Cumple"

Valor\_si\_falso  = "No Cumple"

= "Cumple"

Comprueba si se cumple una condición y devuelve un valor si se evalúa como VERDADERO y otro valor si se evalúa como FALSO.

Prueba\_lógica es cualquier valor o expresión que pueda evaluarse como VERDADERO o FALSO.

W	X	Y
Nivel de Académico	COD_PRUEBA_SABER_11	condición
	AC202340892935	No Cumple
Pregrado	NSL05031400001	Cumple
Posgrado		No Cumple
Pregrado		No Cumple
Pregrado	EXTRANJERO2024	Cumple

Código: IN-DE-004

Fecha: 21/05/2024

Versión: 01

**PROCESO**

**Direccionamiento estratégico**

Regla 2: El formato es alfanumérico de 14 dígitos.

Ejemplo: la longitud de los datos ingresados en la columna `cod_prueba_saber_11`, debe ser igual a 14 dígitos. Por lo tanto se puede utilizar la fórmula `=LARGO` (celda para contar los caracteres en más de una celda), introduzca la fórmula en una columna continua y después, copie y pegue la fórmula en otras celdas.

FECHA DE INGRESO	COD_PRUEBA_SABER_11	Contar digitos
	AC202140921163	<code>=LARGO(R2)</code>
	AC202243351624	
	AC202144143335	

COD_PRUEBA_SABER_11	Contar digitos
AC202140921163	14
AC202243351624	
AC202144143335	

### Validar valores extremos o atípicos:

Validar e identificar valores extremos o atípicos que puedan indicar errores en los datos. Estos valores pueden ser inusuales en comparación con el rango esperado o con otros datos similares.

Ejemplo: Se espera que el valor de la matrícula de los estudiantes de primer periodo sea mayor o igual a \$4.000.000

Regla 1: Si es un el valor de la matrícula es mayor o igual a \$4.000.000, si cumple, de lo contrario, no cumple.

- Revisar la columna donde se encuentra el valor de la matrícula.



UNIVERSIDAD CES

Un compromiso con la excelencia

VIGILADA MINEUCACIÓN

## INSTRUCTIVO PARA GARANTIZAR LA CALIDAD DEL DATO

Código: IN-DE-004

Fecha: 21/05/2024

Versión: 01

PROCESO

Direccionamiento estratégico

Mátricula	Valor de Matrícula
10309690	\$ 4.982.400
10063651	\$ 4.972.400
10212772	\$ 4.932.840
10219211	\$ 4.932.400
10192867	\$ 4.932.400
10186514	\$ 5.905.900
10241745	\$ 4.932.400
10142243	\$ 4.932.400
10150481	\$ 2.262.585
10012054	\$ 4.932.400
10210593	\$ 5.905.900
10169464	\$ 4.905.900
10463158	\$ 3.262.585

Revisar la primera regla: crear una nueva variable en la columna continua a el valor de la matrícula, con la función =SI (que permite realizar comparaciones lógicas entre un valor y un resultado que espera)

Prueba_lógica	C2>4000000	↑	= FALSO
Valor_si_verdadero	"SI"	↑	= "SI"
Valor_si_falso	"NO"	↑	= "NO"
			= "NO"



**UNIVERSIDAD CES**

Un compromiso con la excelencia

VIGILADA MINEUCACIÓN

### INSTRUCTIVO PARA GARANTIZAR LA CALIDAD DEL DATO

Código: IN-DE-004

Fecha: 21/05/2024

Versión: 01

**PROCESO**

**Direccionamiento estratégico**

=SI(C2>=4000000;"SI";"NO")		
B	C	D
Mátricula	Valor de Matrícula	Cumple Valor Matrícula
10309690	\$ 4.982.400	SI
10063651	\$ 4.972.400	SI
10212772	\$ 4.932.840	SI
10219211	\$ 4.932.400	SI
10192867	\$ 4.932.400	SI
10186514	\$ 5.905.900	SI
10241745	\$ 4.932.400	SI
10142243	\$ 4.932.400	SI
10150481	\$ 2.262.585	NO
10012054	\$ 4.932.400	SI
10210593	\$ 5.905.900	SI
10169464	\$ 4.905.900	SI
10463158	\$ 3.262.585	NO

	ELABORACIÓN	REVISIÓN TÉCNICA	APROBACIÓN
<b>Nombre</b>	Sandra Milena Arteaga Cardona	Daniela María Restrepo Acevedo	Mario Alexander Orozco Arcila
<b>Cargo</b>	Analista de Información	Analista de Información	Jefe de Planeación
<b>Fecha</b>	10/05/2024	15/05/2024	21/05/2024

CONTROL DE CAMBIOS		
FECHA DE ACTUALIZACIÓN	CAMBIO REALIZADO	RESPONSABLE DEL CAMBIO
21/05/2024	No aplica por ser versión inicial.	No aplica.