# Dashboard 2: visualización y documentación para narrar a un equipo de trabajo

Desarrollado por: Juan Pablo Mogollón Avaunza

## Variables en base de datos original

1. timestamp: Marca de tiempo en la que se realizó la encuesta (fecha y hora).
2. How old are you?: Rango de edad del encuestado (texto categórico).
3. What industry do you work in?: Industria en la que trabaja el encuestado (texto categórico).
4. Job title: Título del cargo del encuestado (texto categórico).
5. If your job title needs additional context, please clarify here:: Descripción breve de las responsabilidades del cargo (texto).
6. What is your annual salary? (You'll indicate the currency in a later question. If you are part-time or hourly, please enter an annualized equivalent -- what you would earn if you worked the job 40 hours a week, 52 weeks a year.): Salario anual del encuestado en la moneda declarada (número).
7. What is your annual salary? (You'll indicate the currency in a later question. If you are part-time or hourly, please enter an annualized equivalent -- what you would earn if you worked the job 40 hours a week, 52 weeks a year.): Compensaciones extra anuales (bonos acciones etc.) (número).
8. Please indicate the currency: Moneda en la que se reporta el salario (texto categórico).
9. If "Other," please indicate the currency here:: Moneda adicional en caso de tener ingresos en otra divisa (texto categórico).
10. If your income needs additional context, please provide it here: Información adicional proporcionada por el encuestado (texto).
11. What country do you work in?: País de residencia del encuestado (texto categórico).
12. If you're in the U.S., what state do you work in?: Estado en EE.UU. si aplica (texto categórico).
13. What city do you work in?: Ciudad en la que trabaja el encuestado (texto categórico).
14. How many years of professional work experience do you have overall?: Años de experiencia laboral total (número).
15. How many years of professional work experience do you have in your field?: Años de experiencia en la posición actual o similar (número).
16. What is your highest level of education completed?: Nivel educativo alcanzado (texto categórico).
17. What is your gender?: Identificación de género del encuestado (texto categórico).
18. What is your race?: (Choose all that apply.): Autoidentificación étnica del encuestado (texto categórico).

## Variables luego de transformar y modelar datos

1. momento de captura: Marca de tiempo en la que se realizó la encuesta (fecha y hora).
2. rango de edad: Grupo etario del encuestado (texto categórico).
3. Industria\_es: Industria en la que trabaja el encuestado en español (texto categórico).
4. experiencia total: Rango de años de experiencia laboral total (texto categórico).
5. experiencia especifica: Rango de años de experiencia en el área específica del trabajo (texto categórico).
6. nivel educativo: Nivel de educación alcanzado por el encuestado en español (texto categórico).
7. género: Género con el que se identifica el encuestado (texto categórico).
8. país\_corregido: Nombre del país corregido después de la limpieza de datos (texto categórico).
9. ciudad\_corregido: Nombre de la ciudad corregido después de la limpieza de datos (texto categórico).
10. salario\_anual COP: Salario anual del encuestado convertido a pesos colombianos (número decimal).
11. compensaciones COP: Ingresos adicionales al salario base, convertidos a pesos colombianos (número decimal).
12. ingreso anual total COP: Suma del salario anual y las compensaciones adicionales en pesos colombianos (número decimal).

## Transformación de datos

Los pasos descritos a continuación fueron ejecutados utilizando un Notebook de Jupyter, el nombre del notebook es [transform\_data.ipynb](https://github.com/ZabakJL/ManageSalarySurvey2021/blob/main/Scripts/transform_data.ipynb) y se encuentra alojado en el repositorio de Github del proyecto. Este notebook puede consultarse para mayor detalle del procedimiento.

### Renombrar columnas en español

Se cambio el nombre de las columnas a un nombre más corto y comprensible en español, según el mapeo siguiente:

| **titulo original** | **nuevo titulo** |
| --- | --- |
| Timestamp | momento de captura |
| How old are you? | rango de edad |
| What industry do you work in? | industria |
| Job title | posicion |
| If your job title needs additional context, please clarify here: | descripcion de la posición |
| What is your annual salary? (You'll indicate the currency in a later question. If you are part-time or hourly, please enter an annualized equivalent -- what you would earn if you worked the job 40 hours a week, 52 weeks a year.) | salario anual |
| How much additional monetary compensation do you get, if any (for example, bonuses or overtime in an average year)? Please only include monetary compensation here, not the value of benefits. | compensaciones extra anuales |
| Please indicate the currency | moneda |
| If "Other," please indicate the currency here: | otra moneda |
| If your income needs additional context, please provide it here: | contexto adicional |
| What country do you work in? | pais |
| If you're in the U.S., what state do you work in? | estado us |
| What city do you work in? | ciudad |
| How many years of professional work experience do you have overall? | experiencia total |
| How many years of professional work experience do you have in your field? | experiencia especifica |
| What is your highest level of education completed? | nivel educativo |
| What is your gender? | genero |
| What is your race? (Choose all that apply.) | etnia |

### Verificar formato de fecha

Verificar que la columna “momento de captura” contenga datos en formato dd/mm/aaaa hh:mm:ss y tratar los datos que no correspondan con este formato.

### Mapear géneros a español y agruparlos

Mapear los géneros así:

* "Woman": "Mujeres",
* "Man": "Hombres",
* "Non-binary": "Otros",
* "Other or prefer not to answer": "No informado",
* "Prefer not to answer": "No informado"

### Mapear los niveles educativos a español

Realizar el mapeo de los niveles educativos así:

* "Master's degree":"Maestría",
* "College degree":"Título universitario",
* "PhD":"Doctorado",
* "Some college":"Estudios universitarios incompletos",
* "High School":"Bachillerato",
* "Professional degree (MD, JD, etc.)":"Título profesional"

### Borrar espacios

Eliminar los espacios al inicio y final de las cadenas de texto de las columnas: 'ciudad', 'pais', 'otra moneda', 'industria'.

### Mapear los códigos de las monedas

Crear la columna ‘moneda corregida’ y utilizando la información del archivo [codigosMoneda.csv](https://github.com/ZabakJL/ManageSalarySurvey2021/blob/main/Data/Processed/codigosMoneda.csv) mapear las monedas, primero las monedas en la columna moneda, y luego las incluidas en la columna otra moneda. El valor mapeado se debe cargar en la columna ‘moneda corregida’.

* En primera instancia se copian los valores de la columna ‘moneda’ en la columna ‘moneda corregida’ si y solo si el valor de la fila no es igual a ‘Other’ o no es igual a ‘AUD/NZD’. Si el valor de la fila es igual a ‘AUD/NZD’ el valor de la columna ‘moneda corregida’ debe ser igual a ‘AUD’.
* Para las filas en el que el valor de la columna ‘moneda’ es igual a ‘Other’ se buscan las coincidencias de la columna ‘otra moneda’ en el archivo codigosMoneda.csv y se trae el valor de la columna ‘Moneda’ de este archivo a la columna ‘moneda corregida’. En caso de que no hayan coincidencias se debe agregar el registro al archivo [codigosMoneda.csv](https://github.com/ZabakJL/ManageSalarySurvey2021/blob/main/Data/Processed/codigosMoneda.csv).

### Mapear las tasas de cambio

Crear la columna ‘COP’ y rellenar con el valor correspondiente a la tasa de cambio de cada moneda en pesos colombianos. El archivo [TasasDeCambio.csv](https://github.com/ZabakJL/ManageSalarySurvey2021/blob/main/Data/Processed/TasasDeCambio.csv) contiene la equivalencia de cada moneda en pesos colombianos, el valor del cambio se obtuvo utilizando la función monedas de Excel el día 9 de febrero de 2025. En caso de encontrar una nueva moneda, se debe crear en el archivo [TasasDeCambio.csv](https://github.com/ZabakJL/ManageSalarySurvey2021/blob/main/Data/Processed/TasasDeCambio.csv) con su correspondiente equivalencia en pesos colombianos.

### Crear las columnas 'salario\_anual COP', 'compensaciones COP', 'ingreso anual total COP'

* Crear la columna 'salario\_anual COP', igual al producto de la columna ‘COP’ y la columna ‘salario anual’,
* Crear la columna 'compensaciones COP', igual al producto de la columna ‘COP’ y la columna ‘compensaciones extra anuales’,
* Crear la columna 'ingreso anual total COP', igual a la suma de la columna ‘'salario\_anual COP' y la columna 'compensaciones COP'.
* Todas las columnas deben redondearse a cero decimales y tratarse como enteros.

### Mapear los nombres de las industrias a español

Utilizando el archivo [mapearIndustrias.csv](https://github.com/ZabakJL/ManageSalarySurvey2021/blob/main/Data/Processed/mapearIndustrias.csv) convertir los nombres de las industrias a español y almacenarlos en la columna industrias\_es, para ello:

* Cree la nueva columna ‘industrias\_es’
* Cruze el valor de la columna ‘industria’ con la misma columna del archivo mapearIndustrias.csv y obtenga en este archivo el valor de la columna ‘industria\_es’ para la fila correspondiente.

### Mapear los países y las ciudades

Utilizando el archivo [mapearPaisCiudad.csv](https://github.com/ZabakJL/ManageSalarySurvey2021/blob/main/Data/Processed/mapearPaisCiudad.csv) homologar los nombres de países y ciudades. La clave del archivo mapearPaisCuidad.csv corresponde al resultado de concatenar las columnas ‘pais, ‘estado us’ y ‘ciudad’ separadas por ‘|’.

### Seleccionar las columnas de interés

Seleccionar las columnas 'momento de captura', 'rango de edad', 'industria\_es', 'experiencia total', 'experiencia especifica', 'nivel educativo', 'genero', 'pais\_corregido', 'ciudad\_corregido', 'salario\_anual COP', 'compensaciones COP', 'ingreso anual total COP' y crear un archivo csv para usarlo como fuente del tablero.

## Tablero

El tablero se encuentra disponible en: <https://lookerstudio.google.com/reporting/063a1df1-c242-465c-b9f7-55dfe7fe2508>

### Documentación

La documentación está disponible en <https://github.com/ZabakJL/ManageSalarySurvey2021.git>

En este repositorio además se encuentran disponibles todos los archivos csv mencionados en este documento.