VISUALIZACIÓN & STORYTELLING

MIAD | 2023-11

**DASHBOARD 2**

JERSSON MORALES

2023-22808

AL KHOBAR, SAUDI ARABIA

12 DE FEBRERO DE 2021

# CARACTERÍSTICAS BASE DE DATOS ORIGINAL

La base de datos usada para la elaboración del tablero (DASHBOARD 2) como entregable de la Semana 3 del curso Visualización & Storytelling resume los resultados de una encuesta realizada a través de la página WEB:

<https://www.askamanager.org/2021/04/how-much-money-do-you-make-4.html>

La encuesta *“Ask A Manager Salary Survey 2021”* pretende tener información anónima relacionada con el salario y compensación anual para personas de diferentes industrias y cargo laborales en el mundo. La encuesta también incluye preguntas de información personal como edad, genero, raza y nivel educativo. Como la encuesta es realizada para todas las personas a nivel mundial, esta permite de manera libre, escribir algunas respuestas, inclusive no responder, y en otros casos, seleccionar una de las respuestas definidas para una pregunta especifica.

Las preguntas de la encuesta definen las columnas de la base de datos, 17 campos en total y sus filas, las encuestas respondidas que, al momento de la preparación de este documento, contaba 28.012. La base datos está disponible en un archivo sin restricciones de Google Sheets®, en el siguiente hipervínculo:

<https://docs.google.com/spreadsheets/d/1IPS5dBSGtwYVbjsfbaMCYIWnOuRmJcbequohNxCyGVw/edit?resourcekey#gid=1625408792>

Todas las columnas de la base datos (las mismas preguntas de la encuesta) se describen a continuación:

| ÍTEM | NOMBRE DEL COLUMNA | TIPO | DESCRIPCIÓN DE LA COLUMNA |
| --- | --- | --- | --- |
| 1 | How old are you? | Categórica | Contiene el rango de edad en 7 categorías: *under 18, 18-24, 25-34, 35-44, 45-54, 55-64, & 65 or over.* |
| 2 | What industry do you work in? | Texto | Contiene la industria en la cual trabaja. Aunque define 26 categorías, permite la libre escritura al escoger la opción: *“Other”.* |
| 3 | Job title | Texto | Contiene información sobre el cargo que desempeña. Es de libre escritura. |
| 4 | If your job title needs additional context, please clarify here: | Texto | Permite escribir libremente el contexto del cargo que se desempeña, si se considera necesario. |
| 5 | 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.) | Numérica | Contiene el valor del salario anual devengado, y clarifica que debe ser prorrateado a una cifra equivalente a 40 horas por semana y 52 semanas al año. |
| 6 | 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. | Numérica | Contiene el valor de las compensaciones anuales adicionales recibidas y pueden incluir bonos o pago de horas extras. |
| 7 | Please indicate the currency | Texto | Permite indicar la moneda usada en el valor del salario y la compensación anuales. Aunque define 10 categorías, permite libre escritura al escoger la opción: *“Other”.* |
| 8 | If "Other," please indicate the currency here: | Texto | Permite indicar la moneda usada si se escoge la opción *“Other”.* |
| 9 | If your income needs additional context, please provide it here: | Texto | Permite escribir libremente el contexto del salario y compensación anuales, si se considera necesario. |
| 10 | What country do you work in? | Texto | Contiene información del país en el que se trabaja. Permite la libre escritura. |
| 11 | If you're in the U.S., what state do you work in? | Categórica | Contiene el estado en el que trabaja, solo si se trabaja en Estados Unidos. Se puede seleccionar uno de los 50 estados o el distrito federal. |
| 12 | What city do you work in? | Texto | Contiene la ciudad en la que se trabaja. Permite libre escritura. |
| 13 | How many years of professional work experience do you have overall? | Categórica | Contiene el número de años de experiencia laboral general en 10 categorías. |
| 14 | How many years of professional work experience do you have in your field? | Categórica | Contiene el número de años de experiencia laboral específica en 10 categorías. |
| 15 | What is your highest level of education completed? | Categórica | Contiene el nivel educativo más alto completado en 6 categorías. |
| 16 | What is your gender? | Categórica | Contiene el género en 4 categorías: *Man, Woman, Non-binary & Other or prefer not to answer.* Sin embargo, el contenido de la columna muestra otras categorías, ej. *Non-binary.* |
| 17 | What is your race? (Choose all that apply.) | Categórica | Contiene la raza del encuestado. Aunque define 7 categorías, permite escoger libremente varias opciones. |

# MODELADO DE LA BASE DE DATOS

Con el propósito de elaborar un tablero que sirva las veces de reporte interactivo, fue necesario modelar las columnas originales. El modelado consistió principalmente, en unificar las respuestas en nuevos campos que permitían la libre escritura, por ejemplo, el país donde trabaja, la industria en la que trabaja y la moneda usada para las cifras reportadas, de manera que se pudiera facilitar su comunicación a través del tablero. Asimismo, fue necesario crear nuevos campos que permitieran representar los valores de salarios y compensación anual en pesos colombianos (COP) para facilitar su dimensionamiento por parte del público que interactuará con el reporte, que se entiende colombiano, mayormente.

Las columnas modeladas y los campos creados se describen a continuación:

| ÍTEM | NOMBRE DEL CAMPO | TIPO | DESCRIPCIÓN DEL CAMPO |
| --- | --- | --- | --- |
| 1 | COMPENSACIÓN | Numérica | Contiene el valor de la compensación anual en pesos colombianos. Es una variable calculada a partir de la tasa de cambio y la columna original. |
| 2 | FECHA\_HOY | Fecha | Contiene la fecha actual. Se usa como referencia para compararla con la fecha del día de la tasa de cambio usada en las conversiones. |
| 3 | FECHA\_TRM | Fecha | Contiene la fecha del día de las diferentes tasas de cambio usadas en las conversiones. Depende del campo creado: *fecha\_trm\_text* que contiene una fecha en tipo cadena de texto. |
| 4 | Fecha\_trm\_text | Fecha | Contiene la fecha de la tasa de cambio en tipo cadena de texto a partir del parámetro creado: *trm\_date.* |
| 5 | GENERO | Categórica | Contiene el género de los encuestados agrupado en 4 categorías obtenidas a partir de la columna original: *Woman, Non-Binary, Other, Man.* |
| 6 | INDUSTRIA | Categórica | Contiene la industria en la que trabaja en encuestado agrupado en 32 categorías para facilitar su visualización. |
| 7 | INGRESO TOTAL | Numérica | Contiene el ingreso total anual. Se calcula a partir de la suma entre el SALARIO ANUAL y la COMPENSACIÓN. |
| 8 | OTHER\_CURRENCY | Texto | Contiene las diferentes monedas escogidas en las categorías definidas como en las escritas libremente al tomar la opción de *“Other”.* Se obtiene a partir de las columnas originales y permite agrupar de acuerdo con las siglas oficiales de las monedas corrigiendo errores. |
| 9 | PAIS | Texto | Contiene los diferentes países en donde trabajan los encuestados. Se obtiene a partir de la columna original y permite agrupar de acuerdo con los nombres oficiales de los países corrigiendo errores de ortografía y palabras mal escritas. |
| 10 | SALARIO\_ANUAL | Numérica | Contiene el valor del salario anual en pesos colombianos. Es una variable calculada a partir de la tasa de cambio y la columna original. |

# MANIPULACIÓN DE LA BASE DE DATOS PARA FUTURAS ACTUALIZACIONES

Este documento incluye las indicaciones para que un sucesor sea el encargado de las actualizaciones requeridas en un futuro para comunicar con veracidad de los resultados de la encuesta a través del tablero descrito en este documento.

Las indicaciones se dan a continuación:

## Base de Datos (Archivo Google Sheets®)

El tablero carga los datos del archivo directamente desde el vínculo proporcionado para el archivo de Google Sheets®, y los actualiza periódicamente, por lo que la actualización de la base de datos no es necesaria, y los ajustes para modelar y actualizar las cifras son mínimos.

## País donde Trabaja

La escritura libre de este campo permite errores de ortografía y palabras mal escritas para responder el nombre del país en donde el encuestado trabaja, por consiguiente, el sucesor encargado deberá identificar nuevas palabras e incluirlas en el código empleado para modelar la información de la columna. Abajo se muestra parcialmente el código usado.

A screenshot of a computer

Description automatically generated

Quizás en el futuro, el encargado puede valerse de métodos ReGex para optimizar el código, darle flexibilidad, y disminuir el número de palabras a ser incluidas.

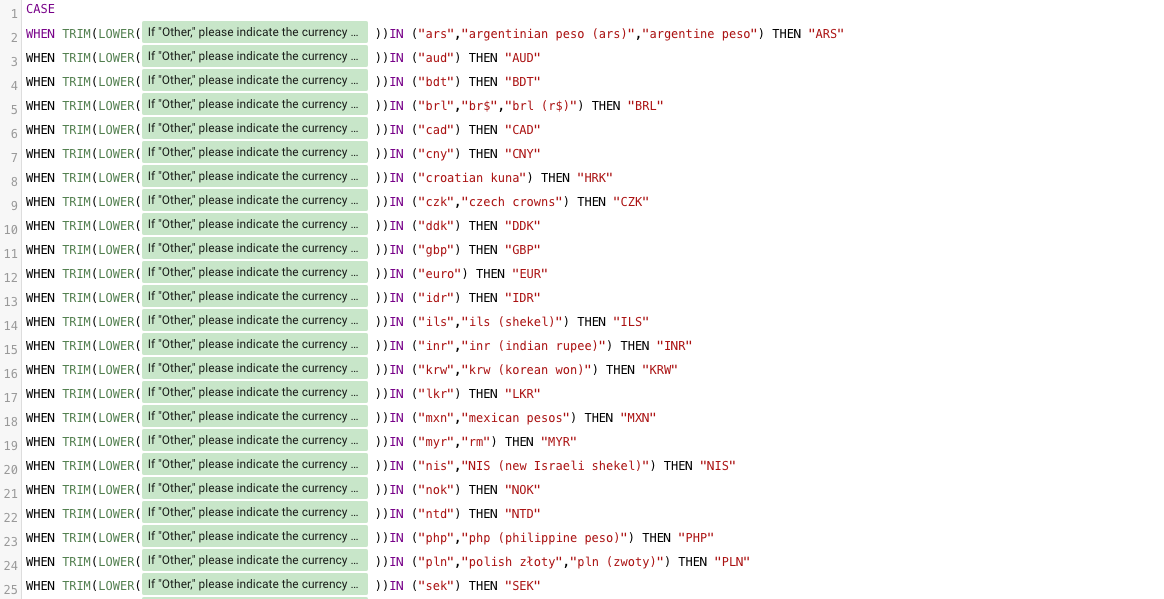
## Moneda y Tasa de Conversión a Pesos Colombianos (COP).

Al igual que el nombre del país, la moneda usada para las cifras permite libre escritura cuando se escoge la opción *“Other”;* entonces, el sucesor encargado deberá identificar siglas de monedas mal escritas, y nuevas monedas para convertir las cifras en pesos colombianos apropiadamente.

De igual manera, el sucesor encargado deberá actualizar periódicamente la tasa de conversión de cada moneda a pesos colombianos (COP) de manera manual, con la información disponible en la página WEB:

<https://www.xe.com/currencyconverter/>

Quizás en el futuro, el sucesor puede valerse de tablas relacionales vinculadas que puedan proveer de tasas de conversión de moneda con cierta frecuencia directamente desde la página para evitar su ajuste manual. Abajo se muestra parcialmente el código usado.



A screenshot of a computer

Description automatically generated

## Industria en la que Trabaja

El campo para la industria donde trabaja el encuestado también permite libre escritura cuando se escoge la opción *“Other”.* Para maximizar los diferentes cargos y agruparlos en un mínimo de sectores de la industria, se usó el código abajo (parcial) que incluye métodos ReGex. Esto reduce considerablemente la actualización de este campo, por lo que el sucesor debe estar atento solo a nuevas industrias o ramas de ellas que deban ser incluidas en el código.

A screenshot of a computer screen

Description automatically generated

## Fecha de la Tasa de Conversión

Mientras no se realice de manera automática, el sucesor encargado deberá actualizar el parámetro: *trm\_date,* creado en el tablero para comunicar la fecha en la cual los valores de las diferentes tasas de cambio fueron tomados para la conversión a Pesos Colombianos (COP) de las cifras de SALARIO\_ANUAL y COMPENSACIÓN.

A screenshot of a computer

Description automatically generated

# TABLERO (DASHBOARD2)

El tablero creado para la comunicación a un público con ciertos conocimientos matemáticos, del ingreso total anual (salario y compensación) de trabajadores en el mundo para ciertas industrias agrupado además por nivel de educación, género y edad, está disponible en el siguiente hipervínculo:

<https://lookerstudio.google.com/s/gzYXjO-40q8>

