

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

Estudiante: Wendy Galvis Larios

Github: https://github.com/WendyGalvis/Documentacion_Dashboard_2.git

En el presente proyecto, nos hemos propuesto realizar un análisis y visualización de los datos recopilados a través de la encuesta "How Much Money Do You Make?" publicada por Ask a Manager. Este conjunto de datos proporciona información relevante sobre los salarios de profesionales en diversas disciplinas, principalmente en los Estados Unidos, y está disponible en un Google Sheets público, que se actualiza constantemente. El objetivo es aplicar técnicas de limpieza, modelado y visualización de datos para transformar la información cruda en insumos útiles para la toma de decisiones e investigaciones internas. A continuación, se describen las características de la fuente de datos, el modelado que se realizó y las acciones necesarias para visualizar adecuadamente la información.

Variables en base de datos original

La exploración inicial nos permite ver que el formulario está en inglés y que hay 18 variables. A continuación, se describen las variables en la base de datos original y el tipo de datos:

Nombre Original	Tipo de Dato	Descripción
Timestamp	Fecha y hora	Almacena la fecha y hora de cada registro en la base de datos.
How old are you?	Texto	Registra el rango de edad en la que se encuentra la persona que diligenció el formulario.
What industry do you work in?	Texto	Describe la industria para la que trabaja el individuo.
Job title	Texto	Describe el cargo de la persona.
If your job title needs additional context, please clarify here	Texto	Campo adicional para aclarar detalles del cargo, solo si es necesario.
What is your annual salary?	Número	Registra el salario anual del individuo en la moneda indicada, considerando un salario anualizado a tiempo completo.
How much additional monetary compensation do you get, if any?	Número	Ingreso adicional promedio anual recibido, como bonos o horas extras.
Please indicate the currency	Texto	Indica la moneda en la que se recibe el salario o compensaciones.

If "Other," please indicate the currency here	Texto	Se diligencia si la moneda no está en la lista de opciones, indicando la moneda manualmente.
If your income needs additional context, please provide it here	Texto	Información adicional sobre el ingreso si la persona considera necesario dar más detalles.
What country do you work in?	Texto	Especifica el país en el que trabaja la persona.
If you're in the U.S., what state do you work in?	Texto	Información opcional, solo si la persona trabaja en los EE. UU., registrando el estado.
What city do you work in?	Texto	Ciudad en la que trabaja la persona (obligatorio para todos los encuestados).
How many years of professional work experience do you have overall?	Texto	Años de experiencia profesional total de la persona.
How many years of professional work experience do you have in your field?	Texto	Años de experiencia profesional en el campo específico del individuo.
What is your highest level of education completed?	Texto	Nivel de educación más alto alcanzado por la persona.
What is your gender?	Texto	Género de la persona.
What is your race?	Texto	Describe la raza o razas del individuo, cuando aplique.

Se puede observar que algunos nombres son muy largos. Por lo tanto, se tomó la decisión de dejar todos los nombres en español y en una versión corta. En la siguiente sección se detallan los nombres iniciales de cada variable y se aclara si hay alguna observación.

Variables luego de modeladas

El modelado se realizó en Looker Studio, el código de todas las variables se encuentra disponible en https://github.com/WendyGalvis/Documentacion_Dashboard_2.git

Nombre Original	Nombre Modificado	Observaciones
How many years of professional work experience do you have in your field?	Años experiencia específica	A partir de esta variable se creó la variable "Grado experiencia específica" que ordena la experiencia de mayor a menor. <u>Archivo código: Años experiencia.txt</u>

How many years of professional work experience do you have overall?	Años Experiencia general	
How much additional monetary compensation do you get, if any (for example, bonuses or overtime in an average year)?	Otras compensaciones	A partir de esta variable se creó la variable "Otras compensaciones (peso colombiano)" que multiplica el valor de esta variable con la Tasa Representativa del Mercado (TRM) de la moneda indicada en el campo "Moneda" a pesos colombianos* <u>Archivo código: Código otras compensaciones pesos colombianos.txt</u>
How old are you?	Rango edad	
If "Other," please indicate the currency here	Moneda – Otra	
If you're in the U.S., what state do you work in?	Estado	
If your income needs additional context, please provide it here	Contexto ingresos	
If your job title needs additional context, please clarify here	Contexto trabajo	
Job title	Cargo	
Please indicate the currency	Moneda	
Timestamp	Fecha y hora	
What city do you work in?	Ciudad	Se cambia el tipo de dato de texto a geográfico (Ciudad). Además, se cambia el nombre a "Ciudad-sucio" y se crea una nueva variable denominada "Ciudad" en la que se ajustan los nombres para que solo se refleje una versión por cada ciudad.** <u>Archivo código: Código ciudad.txt</u>

What country do you work in?	País	Se cambia el tipo de dato de texto a geográfico (Ciudad). Además, se cambia el nombre a "Pais-sucio" y se crea una nueva variable denominada "País" en la que se ajustan los nombres para que solo se refleje una versión por cada ciudad.** <u>Archivo código: Código país.txt</u>
What industry do you work in?	Industria	
What is your annual salary?	Salario anual	A partir de esta variable se creó la variable "Salario (pesos colombianos)" que multiplica el valor de esta variable con la Tasa Representativa del Mercado (TRM) de la moneda indicada en el campo "Moneda" a pesos colombianos* <u>Archivo código: Código salario pesos colombianos.txt</u>
What is your gender?	Género	
What is your highest level of education completed?	Nivel educativo	
What is your race? (Choose all that apply.)	Raza	

*Algunos individuos registraron una moneda y en la variable "Moneda – Otra" indicaron otra moneda. En estos casos se revisó cuando aplicaba el ajuste y se multiplicó el valor del salario y otras compensaciones por la TRM de la moneda correspondiente.

** Este código puede cambiar a medida que la base de datos tenga más registros y sea necesario añadir nuevas variaciones de países y ciudades.

Además de lo anterior, a partir de la suma de las variables **Código salario pesos colombianos.txt** y **Otras compensaciones (peso colombiano)** se creó la variable **Ingreso total**.

Paso a paso

A continuación, se detalla el procedimiento para actualizar los datos, aplicar el modelado correspondiente y, posteriormente, visualizar la información de manera adecuada, todo utilizando la herramienta Looker Studio:

1. En Looker Studio se cargan los datos disponibles en:
<https://docs.google.com/spreadsheets/d/1IP55dBSGtwYVbjsfbaMCYIWnOuRmJcbequohNxGVw/edit?resourcekey#gid=1625408792>

Hay 2 opciones:

"Cargar archivos": para subir directamente el archivo .csv.

"Google Sheets": para subir el archivo desde Google Drive.

2. Al cargar los datos se ajustan los tipos de variables (Desde la opción Recurso>Gestionar fuentes de datos añadidas>Editar) y se cambian los nombres de acuerdo con lo indicado en [Variables luego de modeladas](#).
3. Se procede a limpiar las variables "Ciudad" y "País", ya que estas al ser diligenciadas manualmente tienen muchas variaciones sobre el mismo dato. En este caso, se hizo directamente el Looker Studio, modificando el nombre de País a Pais-Sucio y creando un nuevo campo calculado (Datos>Añadir Campo>Campo calculado) denominado "País" usando el código disponible en: Código país.txt. Asimismo, se modifica el nombre Ciudad a Ciudad-Sucio y se crea un nuevo campo calculado denominado "Ciudad" usando el código disponible en: Código ciudad.txt. Luego, se edita el tipo de dato de texto a geográfico (Desde la opción Recurso>Gestionar fuentes de datos añadidas>Editar).

Este código puede cambiar a medida que la base de datos tenga más registros y sea necesario añadir nuevas variaciones de países o ciudades.

Teniendo en cuenta que hay registros inválidos para el campo de País que se clasificaron como "NA" se optó por crear un filtro que excluya estos datos. Para esto, se selecciona cualquier visualización y en la opción Propiedades>Añadir un filtro>Crear un filtro se configura:

Nombre
País sin NA

Copia de Ask A Manager Salary Survey ...

Mostrar valores sugeridos al escribir

Excluir

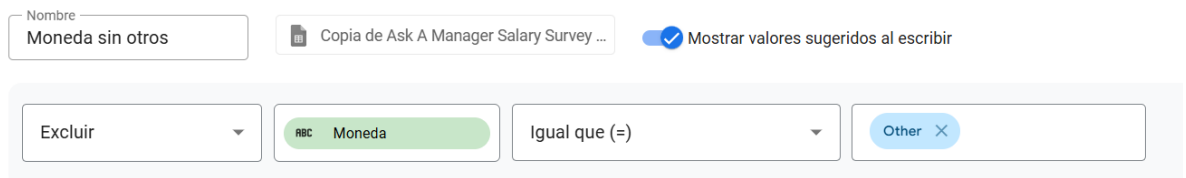
País

Igual que (=)

NA

4. Luego, se procede a crear la variable "Salario-pesos colombianos" para garantizar la comparabilidad de los salarios. Para ello, utilizamos el código disponible en [Código salario pesos colombianos.txt] que multiplica el valor del salario por la TRM de la moneda indicada en el campo "Moneda" en pesos colombianos. Se debe aclarar, que en algunos casos aunque se diligenció el campo "Moneda", se usó la información del campo "Moneda-Otro" ya que había inconsistencias. Teniendo en cuenta que las personas que indicaron en el campo "Moneda" Other son muy pocas y representan un gran trabajo de limpieza, se optó por crear un

filtro que excluya estos datos. Para esto, se selecciona cualquier visualización y en la opción Propiedades>Añadir un filtro>Crear un filtro se configura:



5. El paso anterior se repite para "Otras compensaciones". En este caso, el código se encuentra disponible en Código otras compensaciones pesos colombianos.txt
6. Se crea una variable nueva llamada Ingreso total que suma "Salario-pesos colombianos" con "Otras compensaciones (peso colombiano)".
7. Para facilitar la visualización de los años de experiencia con el salario, se crea la variable "Grado experiencia específica" cuyo código está disponible en Años experiencia.txt que ordena la experiencia de menor a mayor.

Creación de visualizaciones

Para añadir cualquier visualización se selecciona la opción Insertar y se escoge el gráfico deseado.

1. Se añade una tarjeta que usa como métrica el conteo de los registros "Record count".
2. Se añade una tarjeta que usa como métrica el porcentaje de encuestados provenientes de Estados Unidos (campo que se configura así: $\text{SUM}(\text{CASE WHEN País} = \text{"USA"} \text{ THEN } 1 \text{ ELSE } 0 \text{ END}) / \text{COUNT}(\text{País}) * 100$).
3. Se añade un mapa de burbujas cuya dimensión es País, para el tamaño usa como métrica la cantidad de encuestados y para el color se usa el ingreso total promedio.
4. Se añade un gráfico de barras cuya dimensión es Industria y la métrica es el promedio del ingreso total.
5. Se añade un gráfico de barras apiladas cuya dimensión es el grado de experiencia (variable descrita en el ítem 7 del punto anterior), la métrica es el ingreso promedio y se desglosó por género. En este caso, se limitó el análisis solo a hombres y mujeres para que el gráfico fuera mucho más claro para esto se configuró un filtro de la siguiente manera:

