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

## Pregunta 1

Vamos a poner en práctica lo estudiado a lo largo de la semana con una base de datos real. Es una fuente que nos habla de los sueldos que tienen profesionales en diferentes disciplinas (principalmente en Estados Unidos) y estos datos son recolectados por “Ask a Manager”, pueden revisar el formulario en <https://www.askamanager.org/2021/04/how-much-money-do-you-make-4.html> y los datos quedan alojados en un Google Sheets público que se actualiza constantemente: <https://docs.google.com/spreadsheets/d/1IPS5dBSGtwYVbjsfbaMCYIWnOuRmJcbequohNxCyGVw/edit?resourcekey#gid=1625408792>

Algunos campos del formulario son campos de texto donde las personas pueden ingresar libremente la información y esto naturalmente produce diferencias en las formas de escribir datos que deberían ser homogéneos para poderlos analizar y visualizar. Por ejemplo: “City” o “Country”. Al mismo tiempo, el campo de “Annual Salary” es propenso a recibir valores que no corresponden con la medida esperada, por ejemplo, alguien puede escribir 80 (refiriéndose a coloquialismos como 80K) para decir que su salario anual es 80,000. El resultado que encuentran en la base de datos es una buena oportunidad para poner en práctica sus habilidades de limpiar, filtrar y en general modelar los datos para ser analizados y visualizados.

Mientras realizan el ejercicio, imaginen que están trabajando en un equipo (empresa, colectivo, proyecto en grupo) y lo que están produciendo son los insumos para que como equipo puedan investigar y diseñar unos objetivos propios a partir de la base de datos. El **dashboard** y **documentación** que van a crear, son materiales de trabajo que comunican **internamente** las potencias de un proyecto. En ese sentido, no necesariamente está pensado para un público general y por ende se busca más la utilidad y claridad para un equipo que esté trabajando con esos datos. Esto nos sirve de simulacro y preparación para las dinámicas del proyecto final y ponemos en práctica las formas de narrar a nosotros mismos y al público especializado.

## Modelar los datos

Para poder visualizar y documentar tienen que realizar los siguientes pasos de modelado primero:

1. Extraer la base de datos “Ask A Manager Salary Survey 2021” de AskAManager.org (<https://docs.google.com/spreadsheets/d/1IPS5dBSGtwYVbjsfbaMCYIWnOuRmJcbequohNxCyGVw/edit?resourcekey#gid=1625408792>)
2. Limpiar los campos de “Country” y “City” para homogenizar los nombres de los lugares.
3. Crear 2 campos nuevos: “salario\_anual” y “compensaciones” convirtiendo sueldos y compensaciones a **Pesos Colombianos** basados en la tasa de cambio del día que hacen el ejercicio.
4. Crear un campo adicional sumando salario anual y compensaciones en pesos colombianos.

## Presentar un dashboard básico:

Después de realizar el modelado de los datos elaboren un dashboard que contenga los siguientes módulos:

* Módulo 1: Contador básico con total de respuestas.
* Módulo 2: Mapa (decidir si lo muestran por ciudades o sólo países).
* Módulo 3: Relación industria y promedio de salario. Ustedes deciden el tipo de visualización que les permita describir esta relación efectivamente, usar los campos nuevos en pesos colombianos que crearon anteriormente.
* Módulo 4: Libre – Ustedes deciden qué quieren visualizar y eligen la estrategia para hacerlo.
* Créditos a la base de datos, incluir la fecha en la que extraen la TRM.

## Documentar el modelado

Ahora, elaboren un documento escrito en español con las siguientes secciones:

1. Variables en base de datos original. Tengan en cuenta que, al describir esta base de datos, los nombres de las variables deben estar escritas en inglés, tal cual las encuentran en la base de datos original, mientras que las descripciones sí deben estar escritas en español. Esto permite a los otros miembros del grupo volver a la base original y reconocer las variables que están describiendo:
   1. Tipo de variable (número, texto, etc.).
   2. Descripción de variable.
2. Variables luego de modeladas. Durante el modelado, ustedes pueden cambiar el nombre de las variables, las pueden dejar en inglés o traducirlas al español. En cualquier caso, por buena práctica, decidan qué idioma van a usar y sean consistentes con esa decisión en todas las variables:
   1. Tipo de variable.
   2. Descripción corta, máximo 1 párrafo.
3. Describir paso a paso lo que una persona debería hacer para actualizar los datos y aplicar el modelado que han diseñado. Piensen esta sección como una ayuda a alguien que los va a reemplazar en el trabajo. Esta persona debe entender con precisión cómo replicar lo que diseñaron. Este paso a paso asume que la estructura de la base de datos original no cambia en nuevas versiones. Puede ser tan detallada como ustedes quieran, pero debe ser muy claro cómo crear los nuevos campos y procedimientos para limpiar.

## ¡Puntos extra!

Ustedes pueden alcanzar el 100% de la nota cumpliendo con los requisitos anteriormente mencionados. Pero, esta semana pueden acumular 10% extra que se suma a la entrega con la nota más baja que tengan al final del curso.

Para sumar estos puntos extras, simplemente deben presentar su trabajo de esta semana completamente unificado en un sitio web. Es decir, presentar el dashboard y la documentación en una página web como si fuera el sitio oficial de documentación para un proyecto de Analytics. Pueden hacerlo tan sencillo como una entrada de blog con el tablero embebido o crear un micro-sitio propio (también pueden crear un wiki o repositorio en Github, las formas de presentar documentación en sitios web es diversa, busquen una que les parezca cómoda). La idea es que exploren autónomamente las maneras de hacer publicaciones web y voy a premiar aquellos que decidan invertir tiempo y esfuerzo en hacer una presentación de algo que normalmente es de uso interno a un equipo, en un espacio web. En las empresas con equipos grandes, es una muy buena práctica hacer micro-sitios o blogs de documentación para los recursos internos de la empresa ya que son fáciles de diseminar y mantener actualizados en un solo lugar.

## Rúbrica

| **Dimensión** | **Peso** | **Deficiente** | **Bien** | **Muy bien** | **Excelente** |
| --- | --- | --- | --- | --- | --- |
| Dashboard: módulos 1-3 | 30% | No entrega | Falta algún módulo, o, la estrategia de visualización no es muy clara para cumplir con el objetivo. | Módulos completos, cumple con cada uno de los objetivos. La estrategia de visualización podría mejorar. | Cumple claramente con los objetivos y las estrategias de visualización sintetizan con claridad los objetivos. |
| Dashboard: módulo 4 | 10% | No entrega | No es del todo claro lo que busca mostrar con esta visualización. | Se entiende el objetivo de la visualización, la estrategia de visualización o presentación podría mejorar. | La propuesta es clara y revela información implícita en los datos. La estrategia de visualización es clara y efectiva. |
| Documentación: Descripción de variables tabla original | 10% | No entrega | Hay errores en la tipología de las variables o las descripciones no son claras. | Bien tipificadas las variables, las descripciones podrían ser más claras. | Bien tipificadas las variables, las descripciones son efectivas y claras. |
| Documentación: Modelado | 25% | No entrega | No están todos los campos nuevos o hay errores en la conversión de variables. Hay errores en la tipología de las variables o las descripciones no son claras. | Bien tipificadas las variables, las descripciones podrían ser más claras. | Bien tipificadas las variables, las descripciones son efectivas y claras. |
| Documentación: paso a paso | 25% | No entrega | No es muy claro cómo replicar el modelado sobre datos actualizados. | La descripción es clara, pero faltan pasos o la redacción es confusa. | Es perfectamente claro cómo repetir el modelado sobre datos actualizados. |

## Entregables

Si realiza su entrega a través de enlaces, escríbalos en el campo a continuación. De lo contrario, utilice el espacio que está más adelante para adjuntar los archivos que dan cuenta de la visualización y la documentación. La plataforma recibirá un único archivo, por lo que es necesario que los comprima.