|  |
| --- |
|  |
| CARAMELO ESCASO |
|  |

|  |
| --- |
|  |

**Tabla** **de** **contenido**

[1.1. Determinar objetivos del proyecto 3](#_Toc62496795)

[2. Requisitos para la ETL 3](#_Toc62496796)

[2.1. Requisitos técnicos 3](#_Toc62496797)

[2.2. Requisitos funcionales 4](#_Toc62496798)

[3. Ejecución 5](#_Toc62496799)

[3.1. Estructura de carpetas 5](#_Toc62496800)

[3.2. Instalación de Python 6](#_Toc62496801)

[3.3. Instalar librerías 7](#_Toc62496802)

[3.4. Ejecución automática 8](#_Toc62496803)

[3.5. Ejecución en GoogleColabs 9](#_Toc62496804)

[3.6. Tableau 11](#_Toc62496805)

**Título del proyecto:** Automatización y visualización de resultados sobre encuesta de la madurez organizacional de una empresa.

**Encargados del proyecto:**Juliana Penagos González  
Yeison Mauricio Muñoz Cardona   
\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

1. **Motivación de proyecto**

Mediante el uso de herramientas de visualización proveer a la empresa Caramelo Escaso información relevante y organizada para que generen conclusiones que ayuden a identificar con mayor precisión el nivel de madurez organizacional de sus clientes (organizaciones).

## Determinar objetivos del proyecto

* Realizar módulo de automatización en python que nos permita ejecutar el ETL (Extraction, Transformation and Loading), de los insumos que se requieren para la vista en Tableau.
* Implementar tablero de control que permita visualizar de forma grafica e interactiva los resultados de la encuesta para identificar insights.

# Requisitos para la ETL

Actualmente la encuesta se presenta con las preguntas en texto plano, lo que implica un gasto operativo significativo ya que se ejecuta manualmente la asignación de los puntajes y la visualización de los resultados. Con el fin de disminuir esta actividad se necesita garantizar los requisitos técnicos y funcionales que se detallan a continuación.

## Requisitos técnicos

**Los requisitos técnicos solo se ejecutan una única vez en la máquina que se debe ejecutar.**

Instalar en la máquina de ejecución los siguiente paquetes y herramientas.

* Python >= 3.6
* pip = 20.3.3
* pandas = 0.25.3
* numpy = 1.16.6
* scipy 1.3.3
* Tableau

## Requisitos funcionales

Para los requisitos funcionales de la operación se estableció lo siguiente:

Generar 2 catálogos que son estáticos hasta que se añadan nuevas preguntas, las cuales se cambiaran en ambos archivos y en función de las necesidades que Caramelo vaya teniendo. Los catálogos contienen:

* Las preguntas tengan asociado un código (identificador de pregunta), un identificador si es pregunta negativa (inversa, 1 si es inversa 0 en otro caso), la actuación a la que está relacionada, ver catalogo\_1.



* Se requiere un catálogo que contenga todas las preguntas y sus puntuaciones. Ver catalogo\_2.



* Estos catálogos se desarrollaron en función de 26 actuaciones y 130 preguntas que se tienen establecidas actualmente.

Los insumos se obtienen de 2 fuentes principales, la primera es el resultado de la encuesta realizada vía *typeforms* la cual está compuesta por las preguntas en las columnas y las respuestas en las filas y el identificador del cliente, ver archivo a continuación. Nota, definir la estructura de los identificadores.



La segunda fuente es la información sociodemográfica de los encuestados ver archivo a continuación. 

# Ejecución

## Estructura de carpetas

A continuación, se muestra la estructura de las carpetas y su contenido del proyecto:

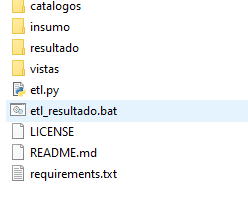
Catálogos: contiene la información sobre los códigos de las preguntas y si son inversas. Para mayor detalle ver requisitos funcionales.

Insumo: contiene el resultado de las encuestas. Para mayor detalle ver requisitos funcionales.   
Se compone de:



Es el resultado de la encuesta que tiene las preguntas y las respuestas sin los puntajes



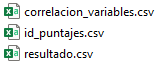


Resultados: contiene el resultado de la ETL que incluye lo siguiente:



Es la información demográfica de los encuestados

Dashboard del reporte



Requerimientos para la ejecución en python

Archivo notebook que se ejecuta en googlecolabs

Matriz de correlación

Matriz de encuesta con puntajes

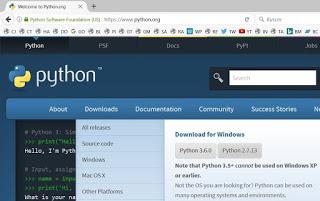
Agregaciones sobre la encuesta con los puntajes

Script funcional ETL

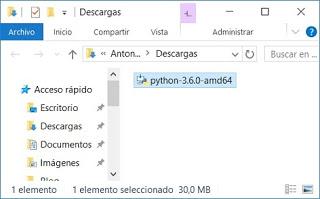
Script de ejecución automatica

## Instalación de Python

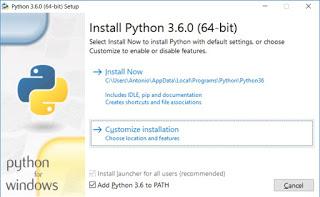
* Ir al [área de descargas para Windows](https://www.python.org/downloads/windows/) del sitio oficial y descargar el instalador ejecutable Python 3.6 de 32 o 64 bit.



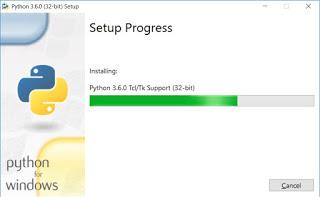
* Iniciar la instalación ejecutando el archivo descargado **Python-3.6.0.exe** o **Python-3.6.0-amd64.exe** con doble clic. Si es necesario, confirmar la ejecución en la ventana de advertencia de seguridad de **Abrir archivo**.

[](https://3.bp.blogspot.com/-np3vzH5bmdU/WJYCE6R9IHI/AAAAAAAADAs/Wv0iwnymH7YFEMuSmAjwo0eHHh1ZvuPzACLcB/s1600/w10-03.jpg)

* Una vez iniciado el instalador, en la ventana**Install Python 3.6.0 (64 bit)** activar las casillas de las opciones: **Install launcher for all users (recommended)** y **Add Python 3.6** to **PATH**. Después, continuar seleccionando la opción **Customize installation. Choose Location and features**

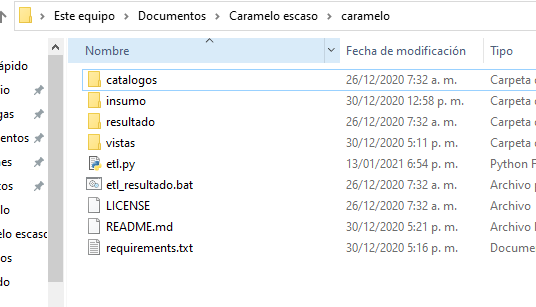


* Finalizar la instalación.

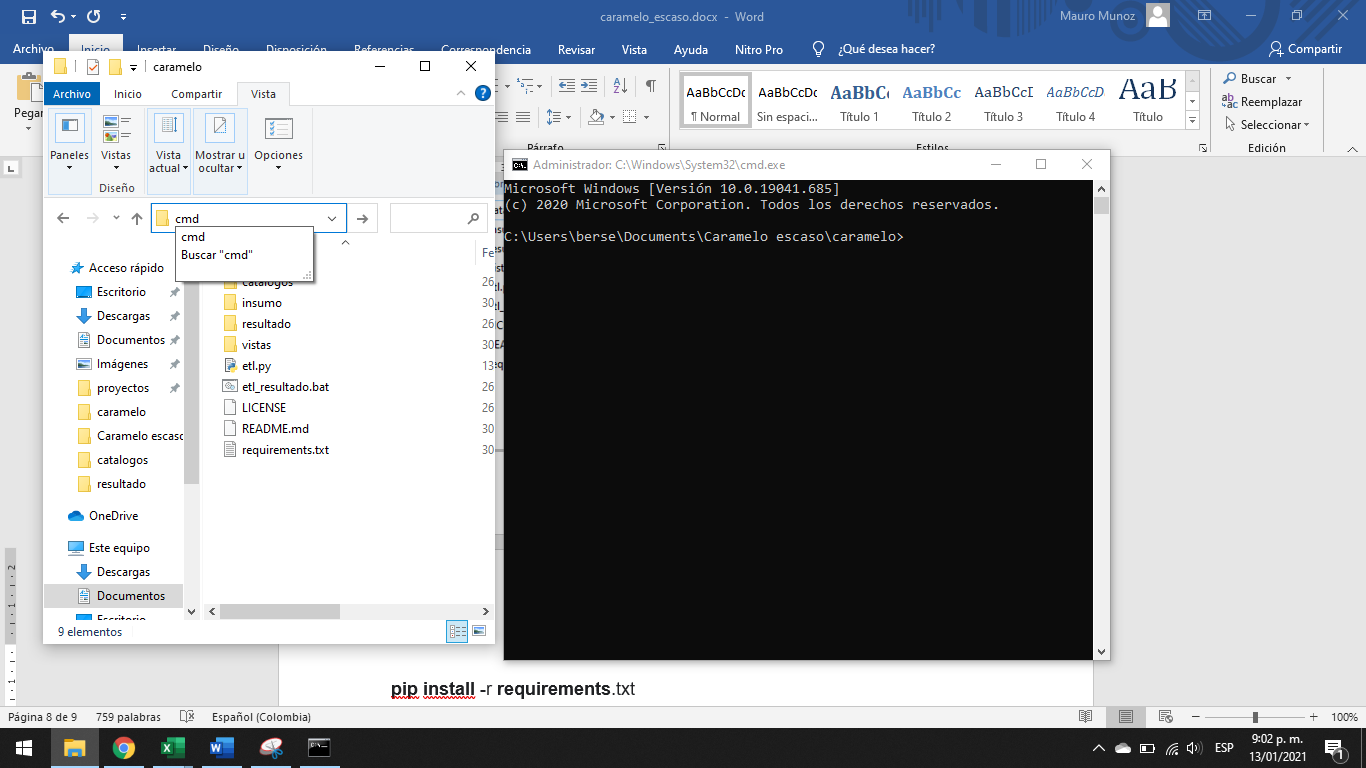


## Instalar librerías

1. Desde el proyecto digitar en la flecha roja cmd



1. Se abrirá una línea de comandos así:



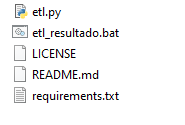
1. Digitar en la venta de comandos (ventana negra), el cual instalará los requerimientos necesarios para poder ejecutar el ETL de manera automática lo siguiente:

pip install -r requirements.txt

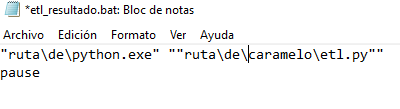
## Ejecución automática

Para la ejecución automática se hace doble -click al siguiente archivo, el cual devolverá los archivos que necesita las vistas de tableau. Nota: el archivo se debe configurar con inicialmente con la ruta del Python instalado y finalmente la ruta del archivo “etl.py”. ver configuración script.

Script de ejecución automatica

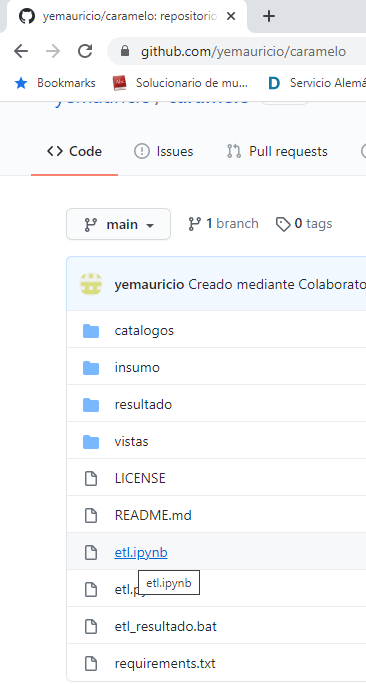


Configuración script

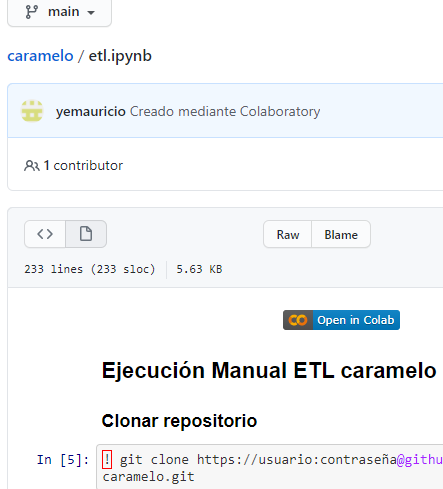


## Ejecución en GoogleColabs

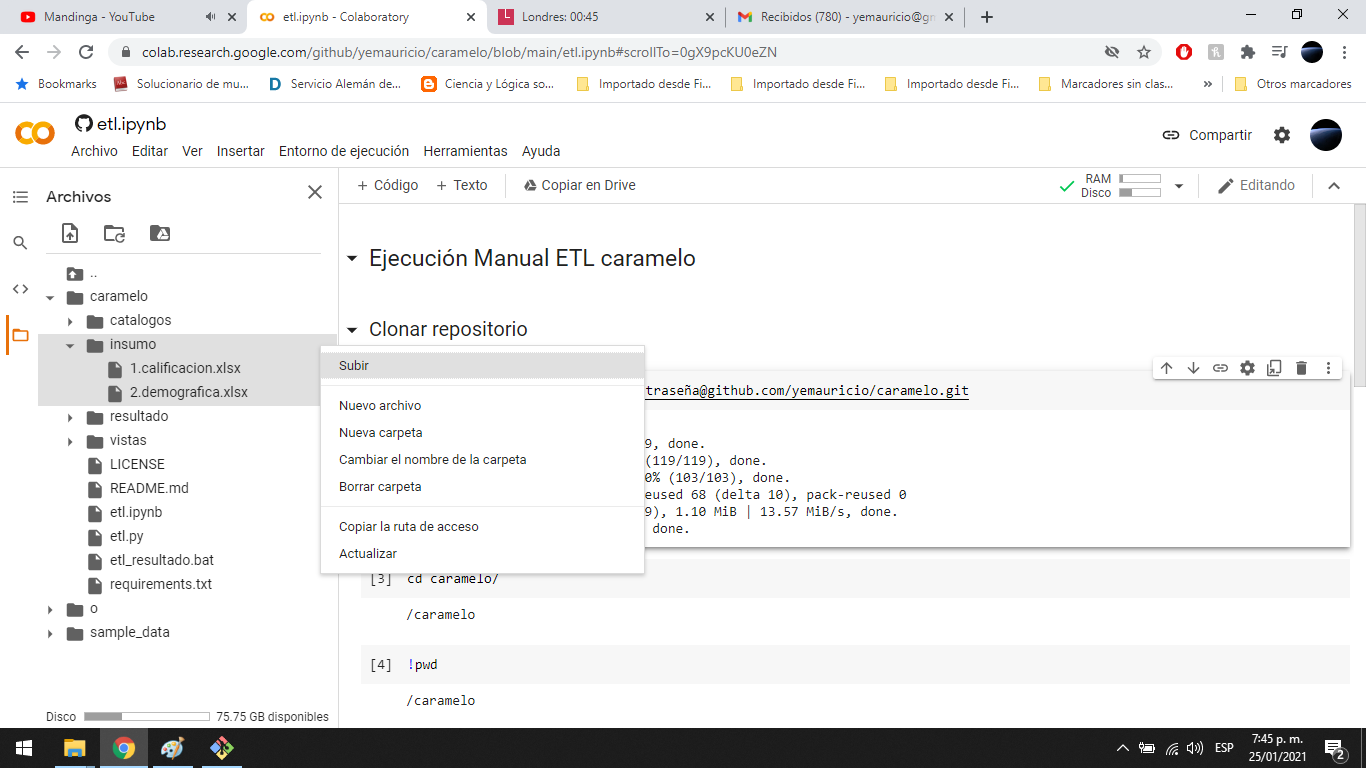
* Abrir github del proyecto y buscar el archivo “*etl.ipynb”* tal como se muestra a continuación.



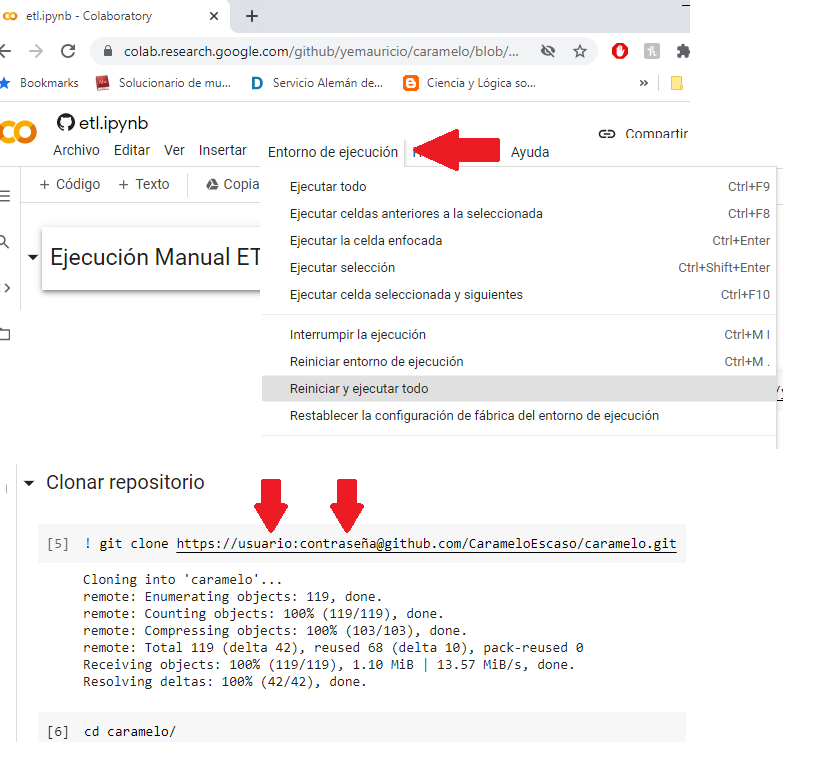
* Al abrirlo los lleva a la siguiente pestaña la cual nos permite usar “googlecolab”, presionando en la el icono que se muestra a continuación.



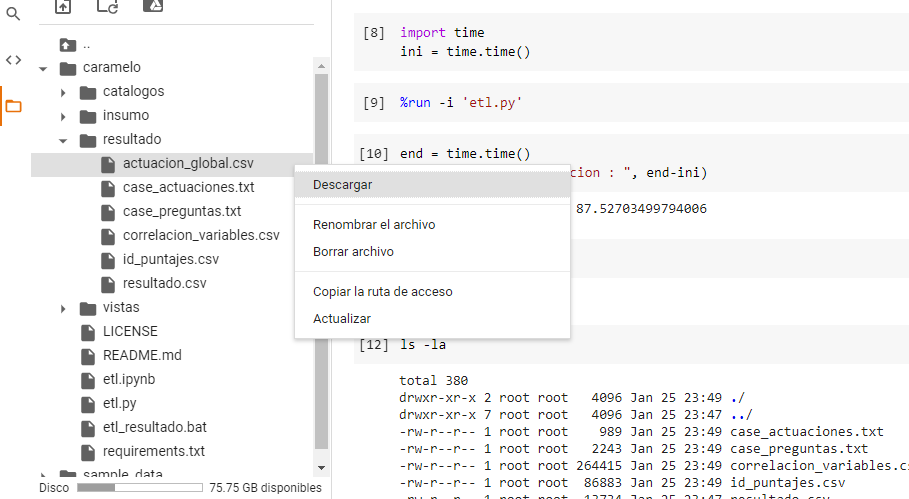
* Lo cual despliega la siguiente pestaña sin embargo es necesario actualizar los archivos de la carpeta insumo (borrar los que aparecen dentro de esa carpeta).



* Después es necesario actualizar la contraseña y el usuario en las flechas en rojo. Nota, las contraseñas no debería tener el signo @.



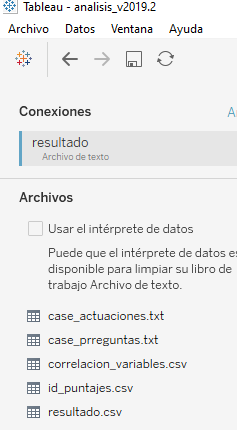
* Finalmente, se observa la actualización del archivo al final del script y se procede a descargar los archivos de la carpeta “resultados” para poder actualizar las vistas. Nota: se observa que la actualización esta en el horario UTC – 0 es decir 5 horas más que Colombia.



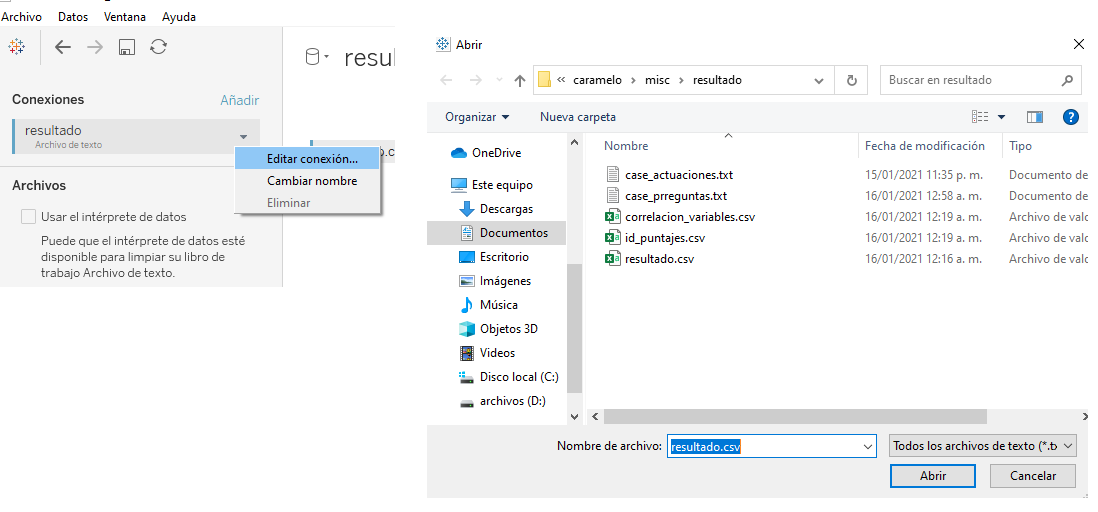
Archivo actualizado en don de se ve la hora 23:49 la cual corresponde a UTC-0, para Colombia es 6:49 pm

## Tableau

Para actualizar las vistas se debe seleccionar el botón que señala la flecha roja, a continuación:

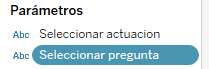


En caso de que la actualización no funcione se puede realizar editando la fuente de conexión lo que despliega la siguiente pantalla para seleccionar la fuente, la cual siempre va a ser resultado.csv, id\_puntajes.csv y correlacion\_variables.csv.



Para la vista se crearon

* Parámetros: se crearon los siguientes 2 parámetros



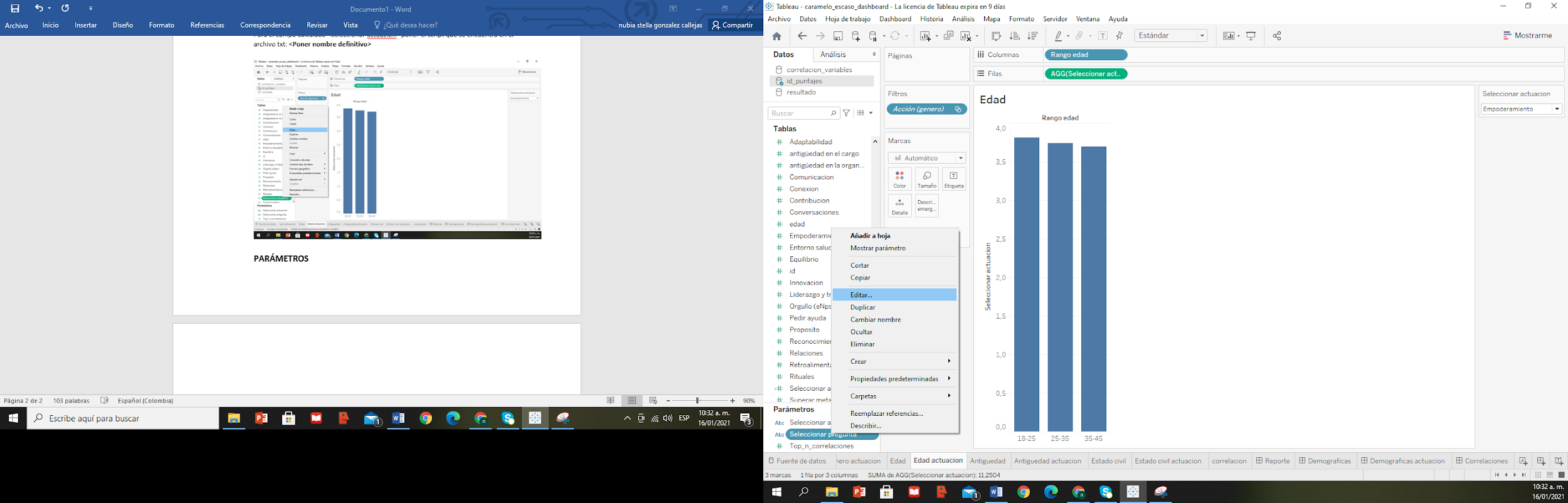
* Campos calculados: se crearon los siguientes 2 campos calculados:





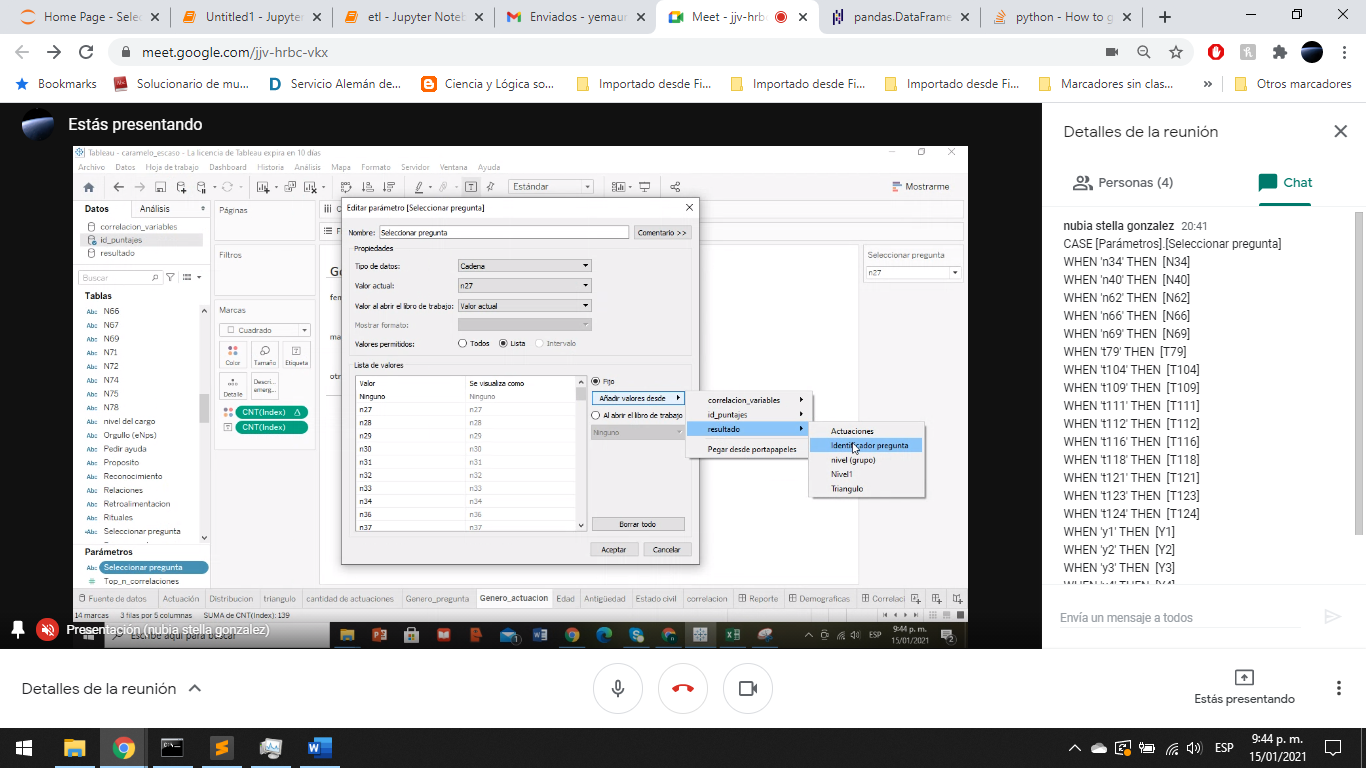
**Parámetros**

Se selecciona el parámetro, click derecho y editar como en la siguiente pantalla:

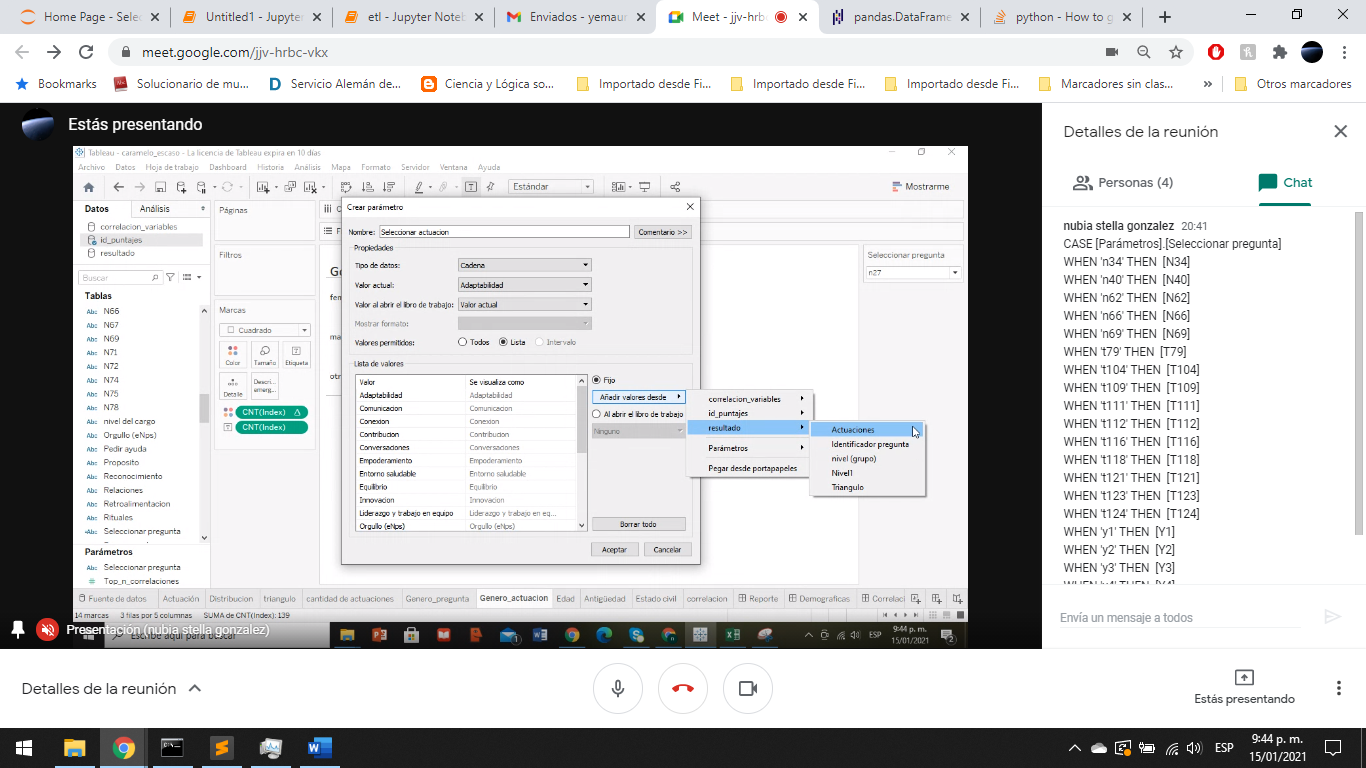


Sale una pantalla y se da clic en el botón borrar todo.

Para el parámetro <Seleccionar pregunta> dar clic en Añadir valores desde luego resultado y luego identificador de pregunta y finalmente aceptar.



Para el parámetro <Seleccionar actuacion> dar clic en Añadir valores desde, luego resultado y luego actuaciones y finalmente aceptar.



**CAMPOS CALCULADOS**

Para actualizar los campos calculados dar clic en el campo y dar clic derecho, y clic en editar como lo muestra la siguiente pantalla. Sale un recuadro donde hay que modificar el script que sale de la ejecución de la ETL.

Para el campo calculado <Seleccionar pregunta> poner el script que se encuentra en el archivo txt: **case\_preguntas.txt**

Para el campo calculado <Seleccionar actuacion> poner el script que se encuentra en el archivo txt: **case\_actuacion.txt**

