





**Data Scientist en Interseguro** 









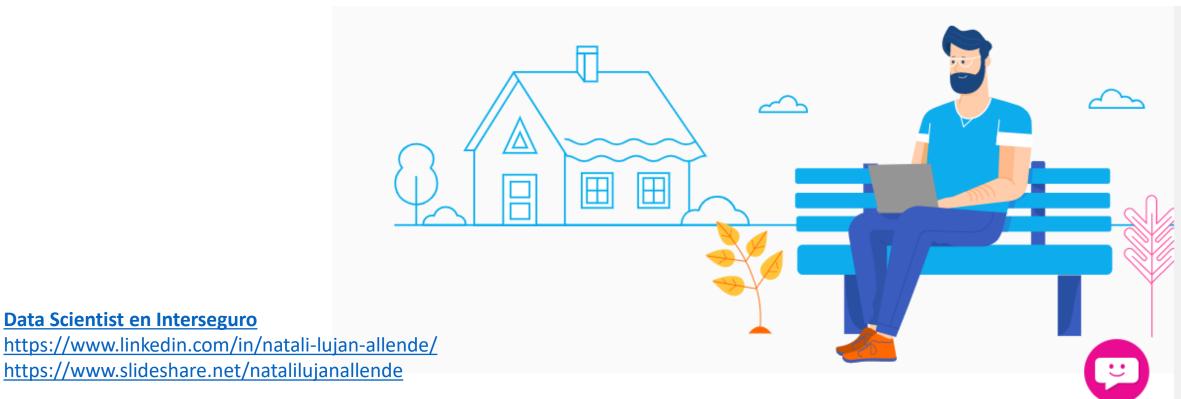




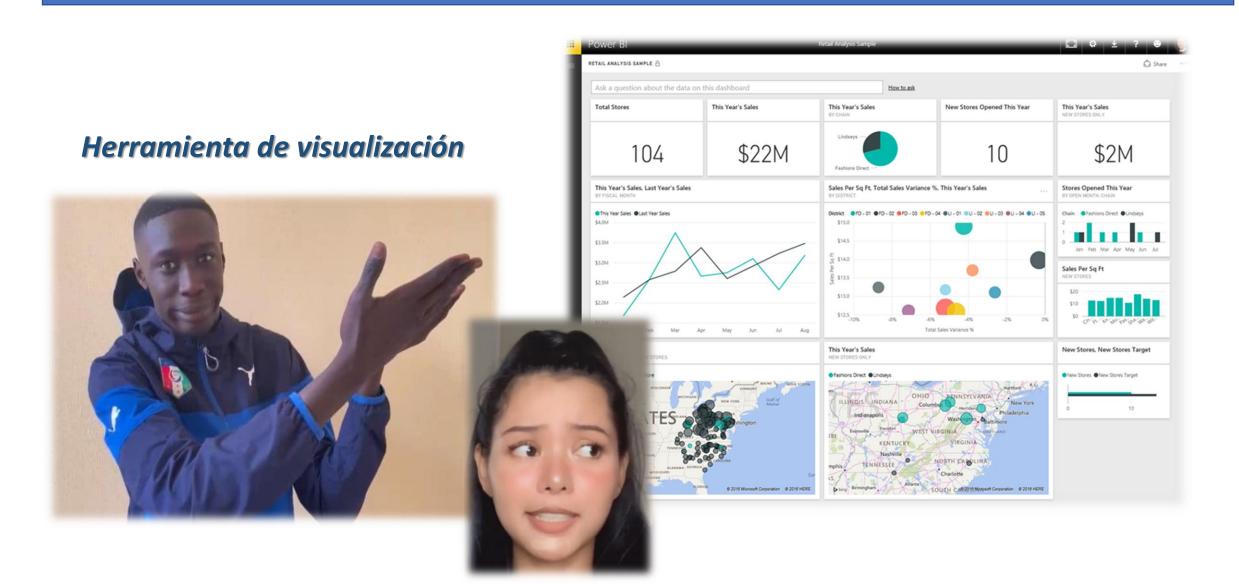


Bachiller en Estadística Informática (UNALM) con especialización en Marketing Digital (UPC)





### ¿Que es Power BI?



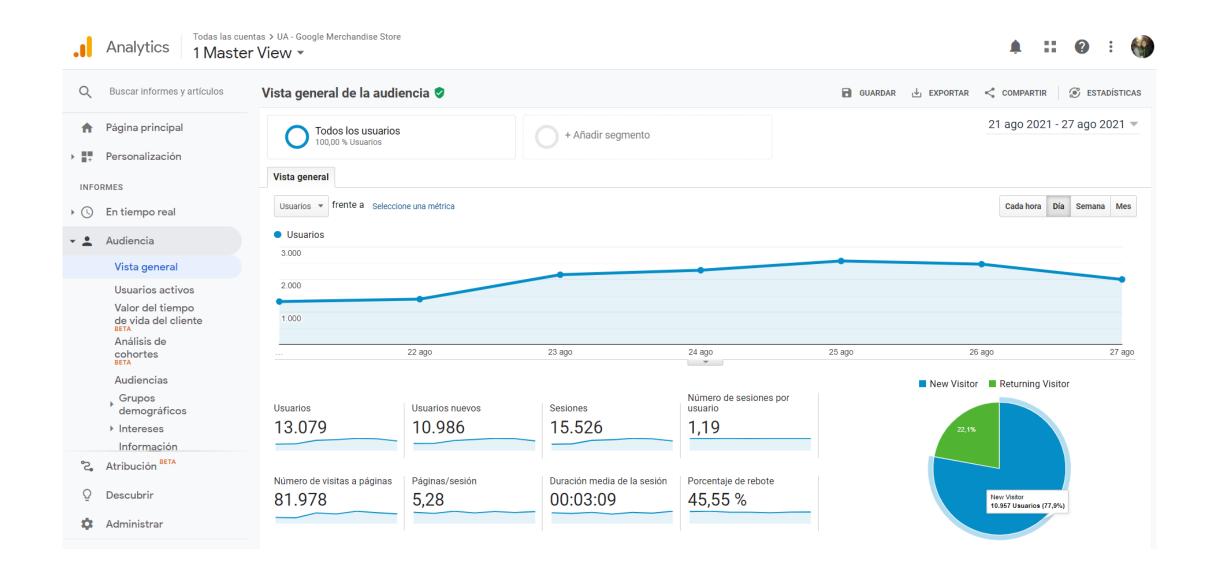
# Caso de uso:

Existe una muestra publica con métricas relacionada a la página Merchandise Store de Google. <a href="https://shop.googlemerchandisestore.com/">https://shop.googlemerchandisestore.com/</a>

Se solicita realizar un dashboard con estadísticas de la página web



# Google Analytics



## Campos a utilizar



Importar datos utilizando R script



# ¿Que es R y R Studio?

R es un lenguaje de programación y entorno computacional dedicado a la estadística.





Es recomendable instalar y usar un entorno integrado de desarrollo como R Studio



Utilizado para minería de textos, procesamiento de imagen, visualizaciones interactivas de datos y procesamiento de Big Data, etc

### Instalar R y sus limitaciones en Power Bi

https://docs.microsoft.com/en-us/power-bi/connect-data/desktop-r-scripts

#### Instalar R

Para ejecutar scripts de R en Power BI Desktop, debe instalar R en su máquina local. Puede descargar e instalar R de forma gratuita desde muchas ubicaciones, incluidos Microsoft R Application Network 2 y CRAN Repository 2. La versión actual admite caracteres Unicode y espacios (caracteres vacíos) en la ruta de instalación.

#### Preparar un script R

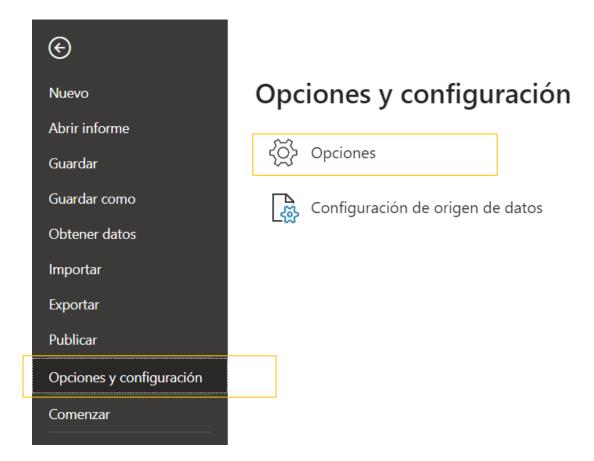
Para ejecutar un script de R en Power BI Desktop, cree el script en su entorno de desarrollo de R local y asegúrese de que se ejecute correctamente.

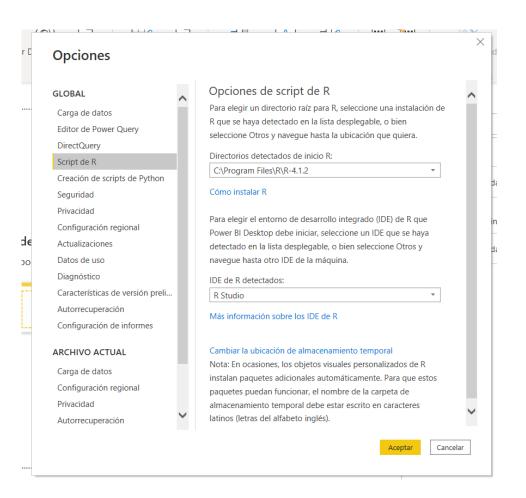
Para ejecutar el script en Power BI Desktop, asegúrese de que se ejecute correctamente en un espacio de trabajo nuevo y sin modificar. Este requisito previo significa que todos los paquetes y dependencias deben cargarse y ejecutarse explícitamente. Puede utilizar source() para ejecutar scripts dependientes.

Cuando prepara y ejecuta un script R en Power BI Desktop, existen algunas limitaciones:

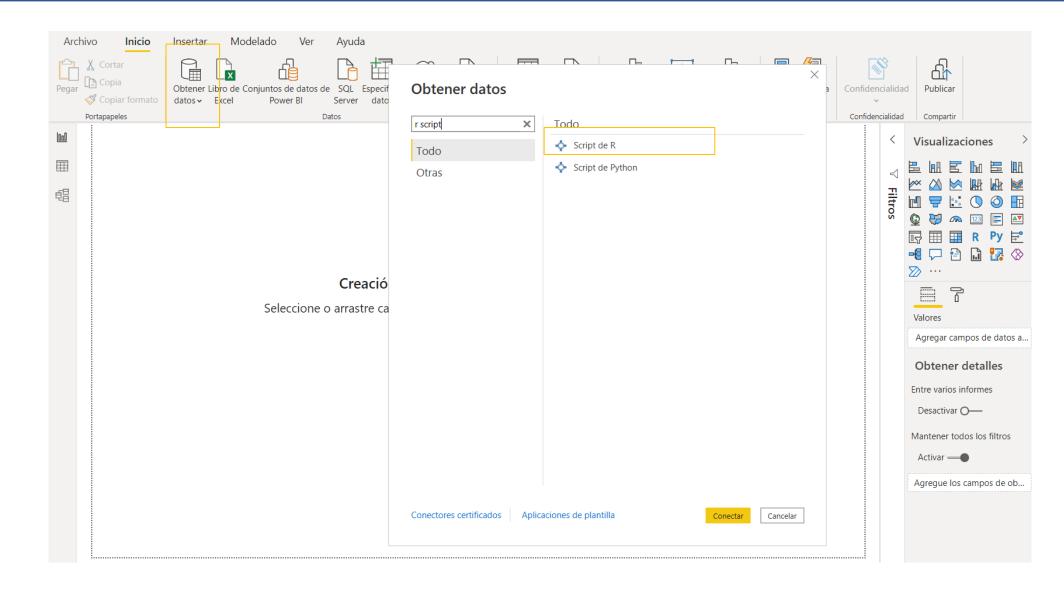
- Dado que solo se importan marcos de datos, recuerde representar los datos que desea importar a Power BI en un marco de datos.
- Las columnas escritas como Complejo y Vector no se importan y se reemplazan con valores de error en la tabla creada.
- Los valores que se N/A traducen a NULL valores en Power BI Desktop.
- Si un script R se ejecuta durante más de 30 minutos, se agota el tiempo de espera.
- Las llamadas interactivas en el script R, como esperar la entrada del usuario, detiene la ejecución del script.
- Al configurar el directorio de trabajo dentro del script R, *debe* definir una ruta completa al directorio de trabajo, en lugar de una ruta relativa.

### Ubicar el directorio raíz

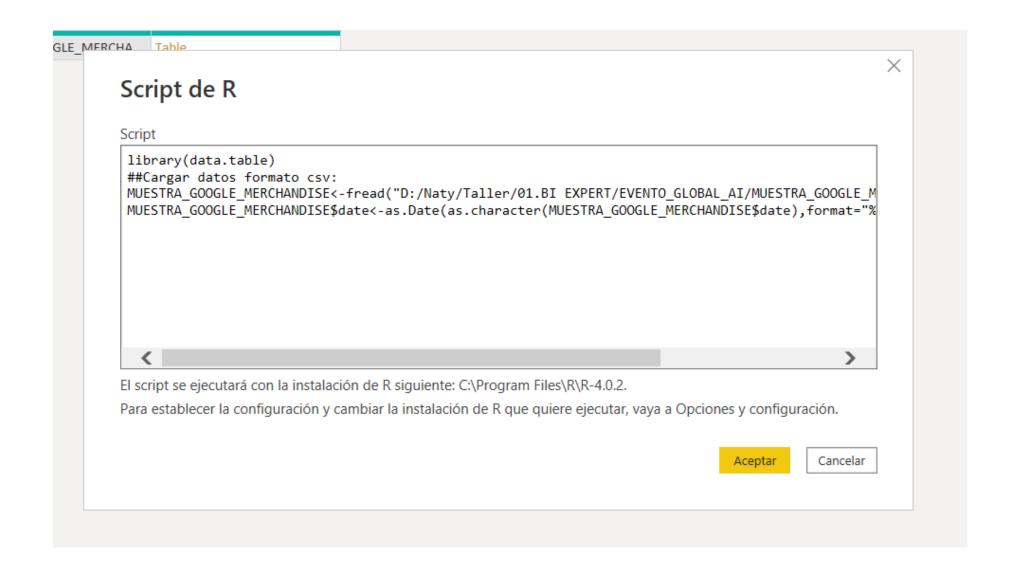




# Cargar datos desde R Script



### Importar datos csv





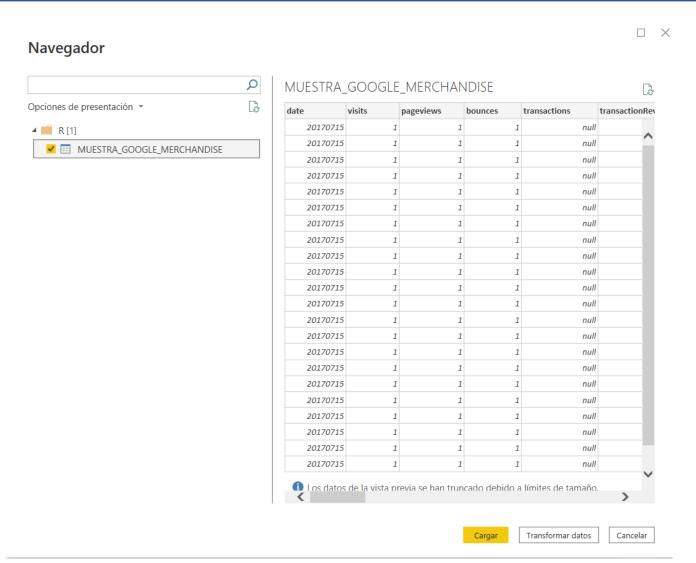
### Script R:

#Importar datos library(data.table) ##Cargar datos formato csv:

GOOGLE\_MERCHANDISE\_R<-fread("D:/Naty/Taller/01.BI EXPERT/EVENTO AZURE DATA ANALYTICS/MUESTRA\_GOOGLE\_MERCHANDISE.csv",sep=",",header=T) GOOGLE\_MERCHANDISE\_R\$date<-as.Date(as.character(GOOGLE\_MERCHANDISE\_R\$date),format="%Y%m%d")

Los paquetes en R son colecciones de funciones y conjunto de datos desarrollados por la comunidad.

### Conectarse a la base seleccionada



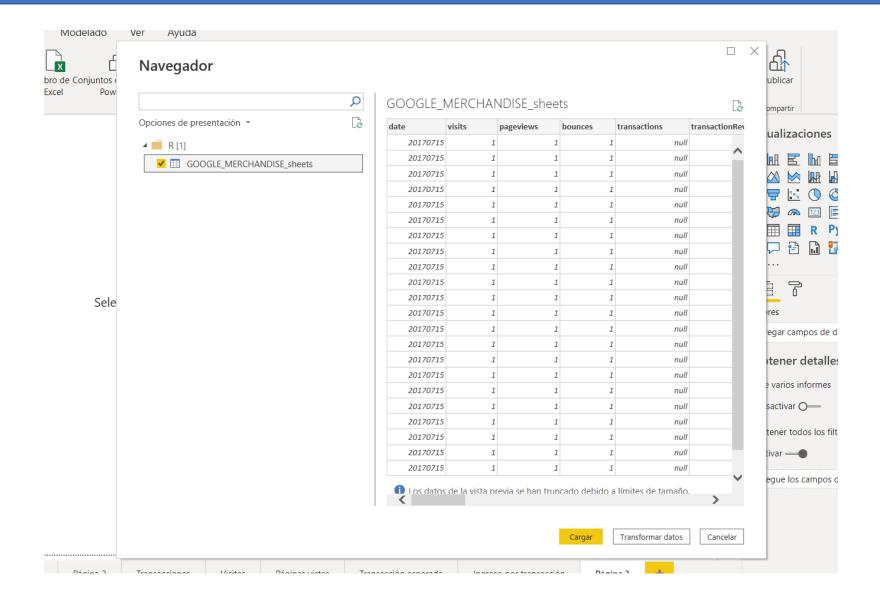
## Importar datos desde Google sheets

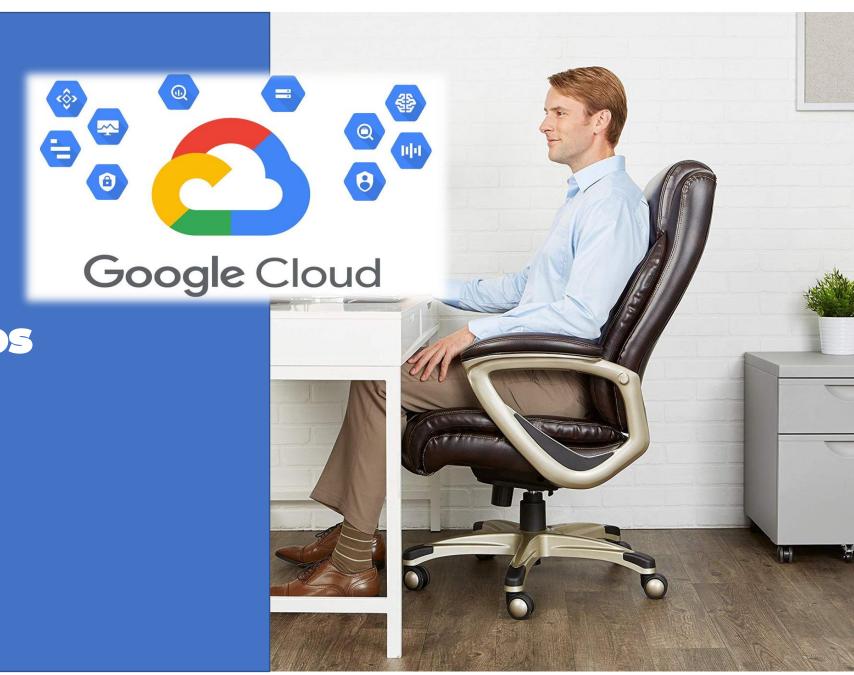




### Script R:

## Importar datos desde Google sheets





Importar datos desde Google Cloud Platform

### Importar datos en la nube-GCP

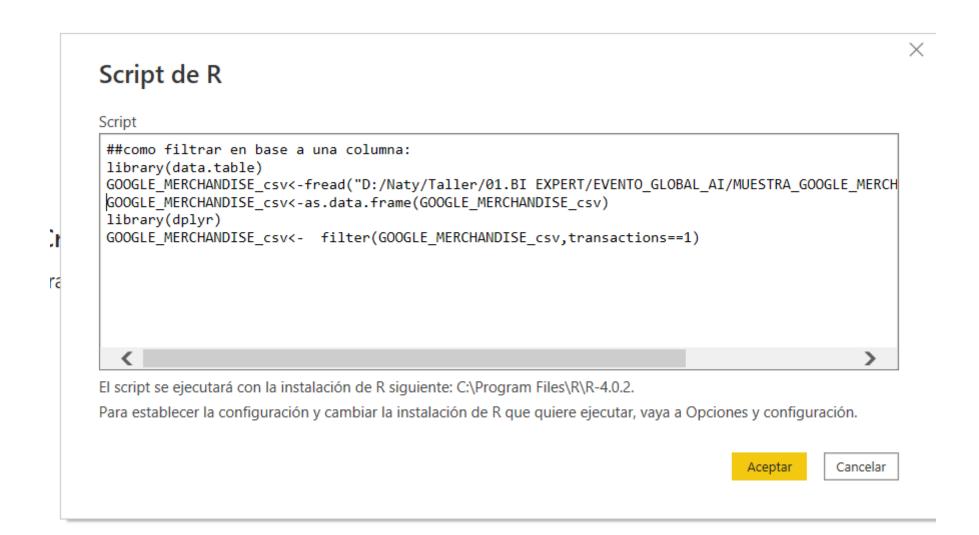




### Script R:

```
#Cargar datos formato google buc
library("bigrquery")
projectid = "data-negocio"
# Set your query
sql <- "select * from data-negocio.data_negocio.vista_google_merchandise"
bq_auth("recomposite = "data-negocio.data_negocio.vista_google_merchandise"
bq_auth("recomposite = "data-negocio.data-negocio.data_negocio.vista_google_merchandise"
bq_auth("recomposite = "data-negocio.data-negocio.data-negocio.data-negocio.data-negocio.data-negocio.data-negocio.data-negocio.data-negocio.data-negocio.data-negocio.data-negocio.data-negocio.data-negocio.data-negocio.data-negocio.data-negocio.data-negocio.data-negocio.data-negocio.data-negocio.data-negocio.data-negocio.data-negocio.data-negocio.data-negocio.data-negocio.data-negocio.data-negocio.data-negocio.data-negocio.data-negocio.data-negocio.data-negocio.data-negocio.data-negocio.data-negocio.data-negocio.data-negocio.data-negocio.data-negocio.data-negocio.data-negocio.data-negocio.data-negocio.data-negocio.data-negocio.data-negocio.data-negocio.data-negocio.data-negocio.data-negocio.data-negocio.data-negocio.data-negocio.data-negocio.data-negocio.data-negocio.data-negocio.data-negocio.data-negocio.data-negocio.data-negocio.data-negocio.data-negocio.data-negocio.data-negocio.data-negocio.data-negocio.data-negocio.data-negoc
```

### Utilizar filtro de base en R





### Script R:

##como filtrar en base a una columna:

library(data.table)

GOOGLE\_MERCHANDISE\_csv<-fread("D:/Naty/Taller/01.BI

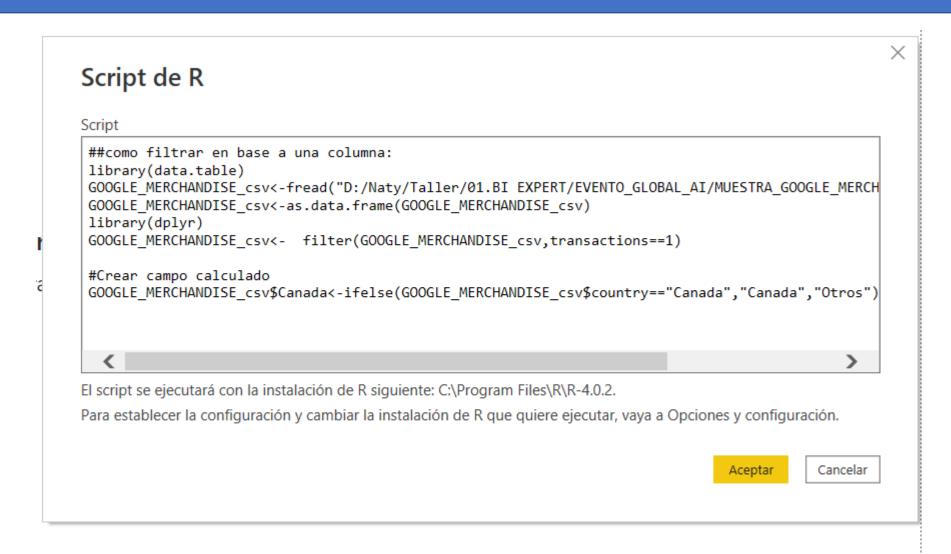
EXPERT/EVENTO\_GLOBAL\_AI/MUESTRA\_GOOGLE\_MERCHANDISE.csv",sep=",",header=T)

GOOGLE\_MERCHANDISE\_csv<-as.data.frame(GOOGLE\_MERCHANDISE\_csv)

library(dplyr)

GOOGLE\_MERCHANDISE\_csv<- filter(GOOGLE\_MERCHANDISE\_csv,transactions==1)

## Crear campo calculado:





### Script R:

```
##como filtrar en base a una columna:
library(data.table)

GOOGLE_MERCHANDISE_csv<-fread("D:/Naty/Taller/01.BI

EXPERT/EVENTO_GLOBAL_AI/MUESTRA_GOOGLE_MERCHANDISE.csv",sep=",",header=T)

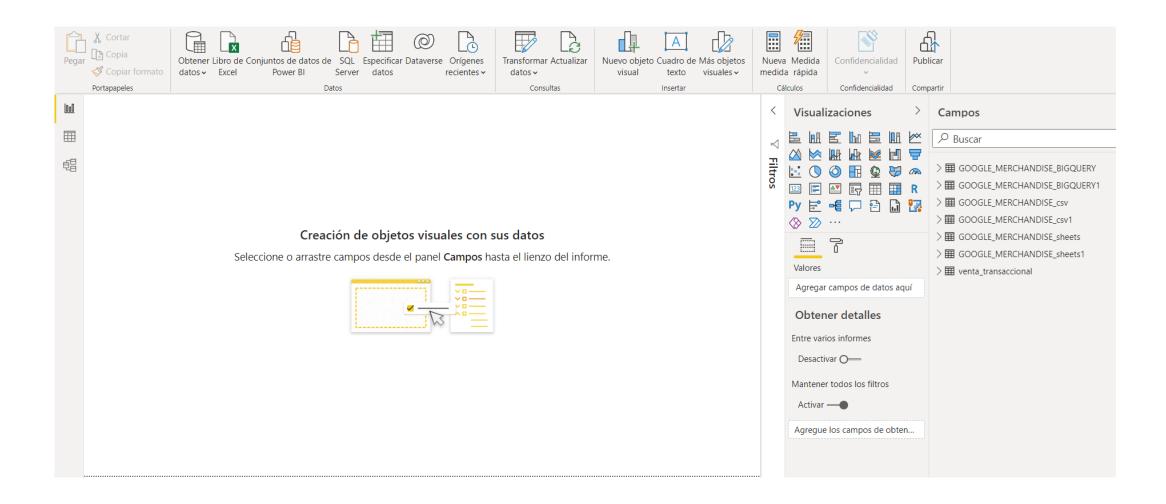
GOOGLE_MERCHANDISE_csv<-as.data.frame(GOOGLE_MERCHANDISE_csv)
library(dplyr)

GOOGLE_MERCHANDISE_csv<- filter(GOOGLE_MERCHANDISE_csv,transactions==1)

#Crear campo calculado

GOOGLE_MERCHANDISE_csv$Canada<-ifelse(GOOGLE_MERCHANDISE_csv$country=="Canada","Canada","Otro ")
```

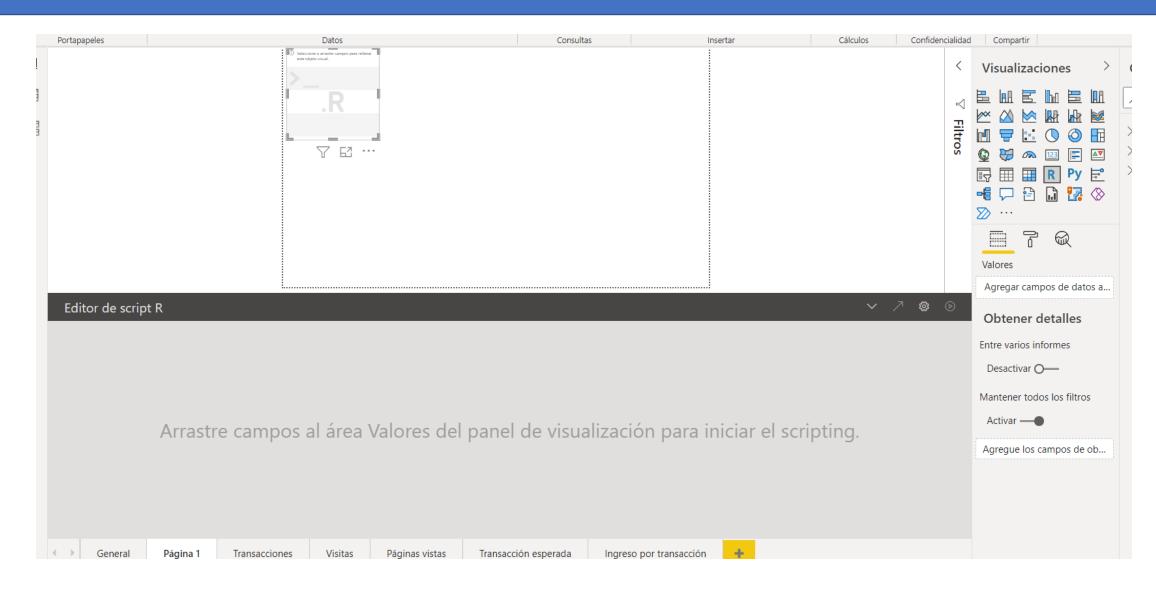
### Cargar varias tablas en una sola consulta



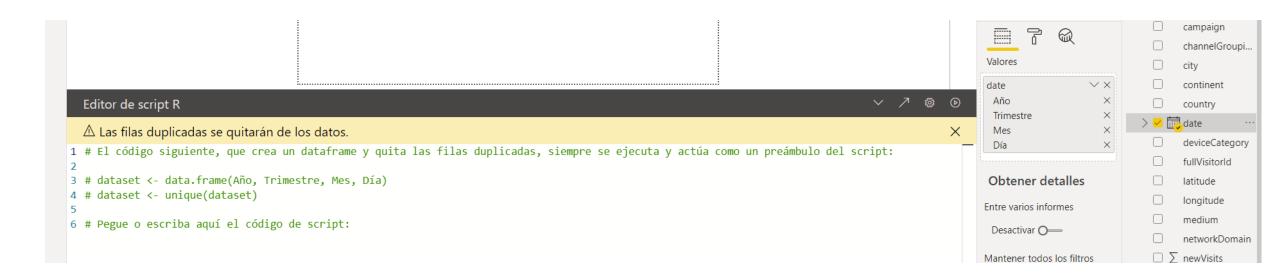


```
#cargar varias tablas en una sola consulta
#Importar datos
library(data.table)
GOOGLE MERCHANDISE csv1<-fread("D:/Naty/Taller/01.BI EXPERT/EVENTO GLOBAL AI/MUESTRA GOOGLE MERCHANDISE.csv",sep=",",header=T)
GOOGLE_MERCHANDISE_csv1$date<-as.Date(as.character(GOOGLE_MERCHANDISE_csv1$date),format="%Y%m%d")
library(googlesheets4)
sheets auth(email = "
                                    com")
GOOGLE MERCHANDISE sheets1 <- read sheet("https://docs.google.com/spreadsheets/d/1G7ccUnndVuONOXMRv3oBC3slmGpTkNPBPx9IHEEWc3M/edit#gid=0",
                  sheet="google")
library("bigrquery")
projectid = "data-negocio"
sql <- "
select * from data-negocio.data negocio.vista google merchandise
bq_auth("natalilujan@gmail.com")
tb <- bq_project_query(projectid, sql)
GOOGLE MERCHANDISE BIGQUERY1 <- bq table download(tb)
```

## Agregar el icono de visualización de R



## Agregar los campos que se van a necesitar para el grafico



## Seleccionar campos pageviews, transactions y visits



## Seleccionar campos pageviews, transactions y visits





### Script R:

```
library(tidyverse)
gg<-ggplot(dataset,aes(x=pageviews,y=transactions)) +
geom_smooth(method = "gam",se=T)+
labs(title="Grafico",
x="pageviews")
plot(gg)</pre>
```

## Seleccionar campos pageviews, transactions y visits





### Script R:

hist(dataset\$pageviews,col="darkolivegreen1")

Importar datos utilizando
Python script



### Instalar R y sus limitaciones en Power Bi

https://docs.microsoft.com/en-us/power-bi/connect-data/desktop-python-scripts

### **Instalar Python**

Para ejecutar scripts de Python en Power BI Desktop, debe instalar Python en su máquina local. Puede descargar Python desde el sitio web de Python . La versión actual de secuencias de comandos de Python admite caracteres y espacios Unicode en la ruta de instalación.

#### Instale los paquetes de Python necesarios

La integración de Power BI con Python requiere la instalación de dos paquetes de Python:

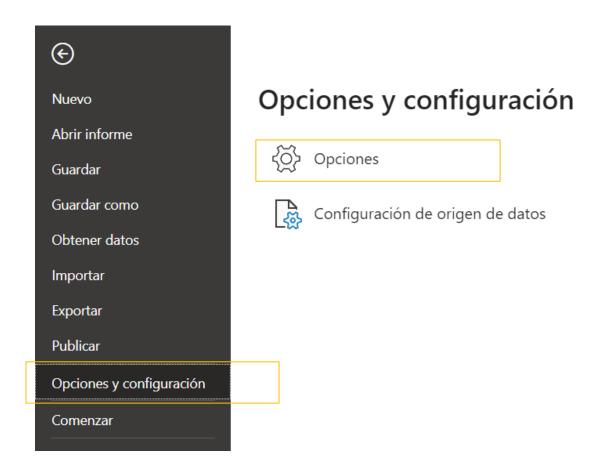
- pandas 🗗 \_ Una biblioteca de software para la manipulación y el análisis de datos. Ofrece estructuras de datos y operaciones para manipular tablas numéricas y series de tiempo. Sus datos importados deben estar en un marco de datos pandas 🗗 . Un marco de datos es una estructura de datos bidimensional. Por ejemplo, los datos se alinean de forma tabular en filas y columnas.

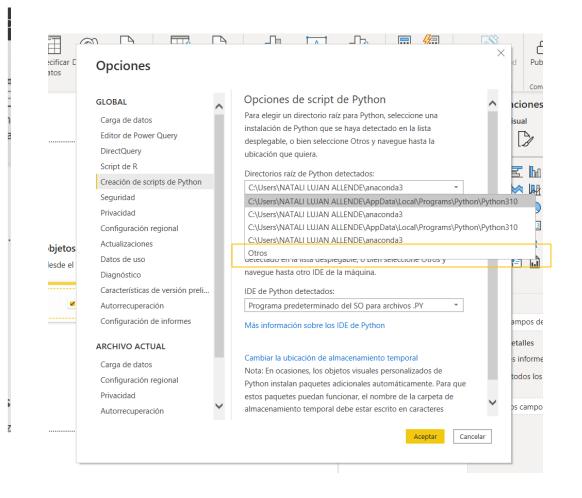
En una consola o shell, use la herramienta de línea de comandos pip para instalar los dos paquetes. 

La herramienta pip está empaquetada con las versiones de Python más recientes.

```
CMD Pip install pandas
pip install matplotlib
```

### Ubicar el directorio raíz





### Ubicar el directorio raíz

```
C:\Users\NATALI LUJAN ALLENDE>python

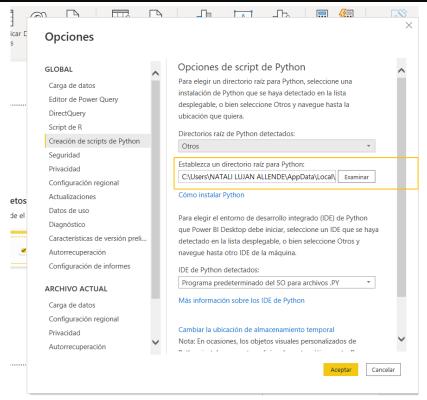
Python 3.9.13 (tags/v3.9.13:6de2ca5, May 17 2022, 16:36:42) [MSC v.1929 64 bit (AMD64)] on win32

Type "help", "copyright", "credits" or "license" for more information.

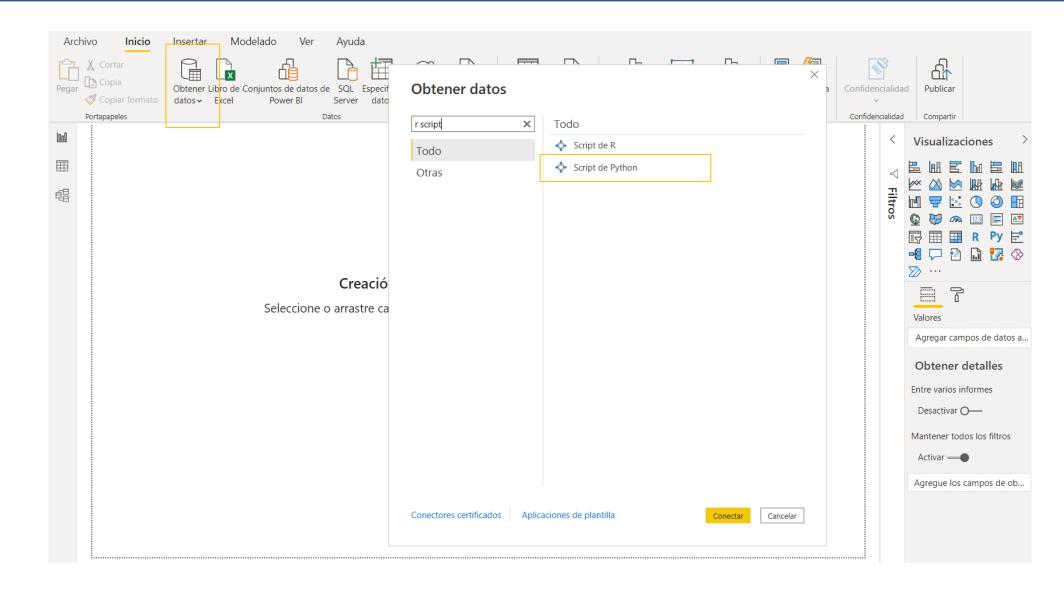
>>> import sys

>>> print(sys.executable)

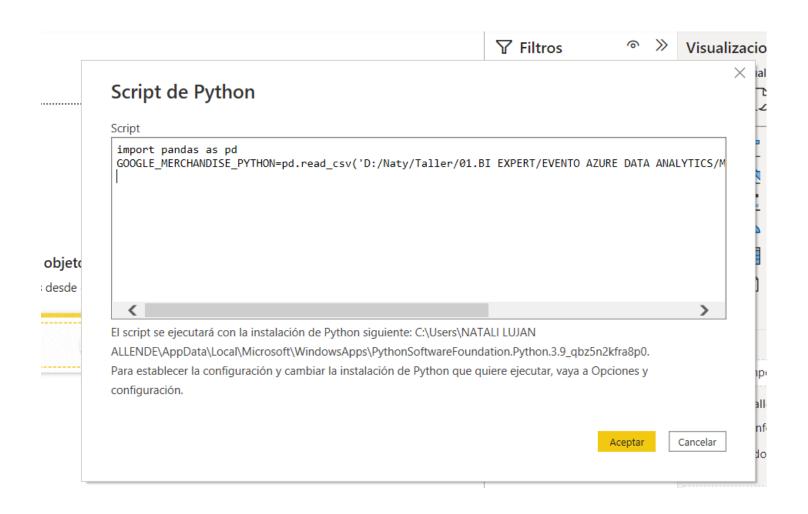
C:\Users\NATALI LUJAN ALLENDE\AppData\Local\Microsoft\WindowsApps\PythonSoftwareFoundation.Python.3.9_qbz5n2kfra8p0\python.exe
```



## Cargar datos desde R Script



## Importar datos csv



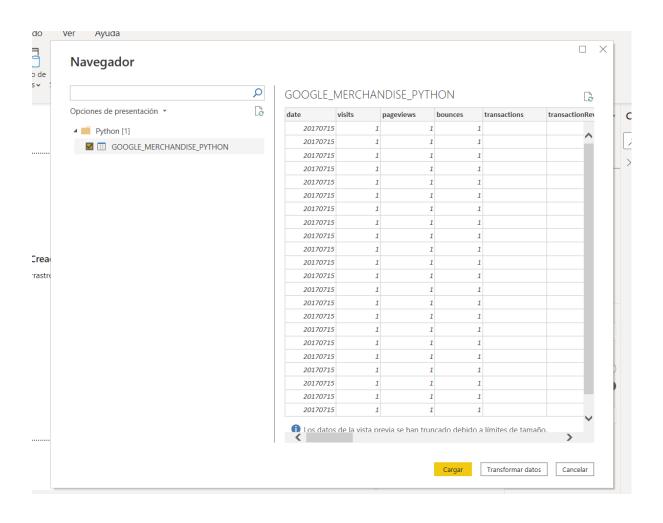


# BLOQUES DE CÓDIGO

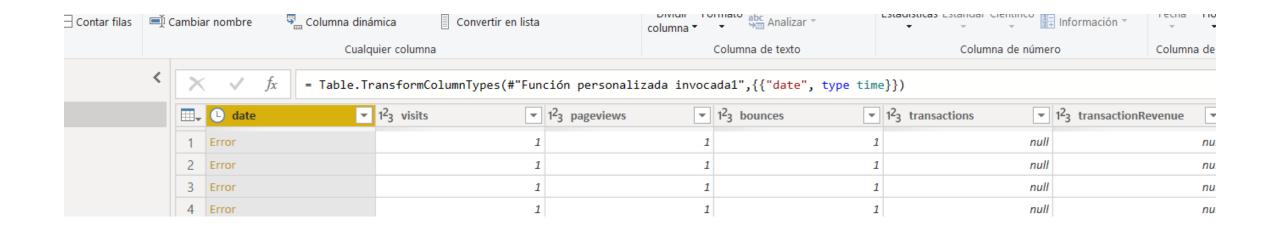
#### Script Python:

import pandas as pd GOOGLE\_MERCHANDISE\_PYTHON=pd.read\_csv('D:/Naty/Taller/01.BI EXPERT/EVENTO AZURE DATA ANALYTICS/MUESTRA\_GOOGLE\_MERCHANDISE.csv',encoding='UTF-8',sep=',')

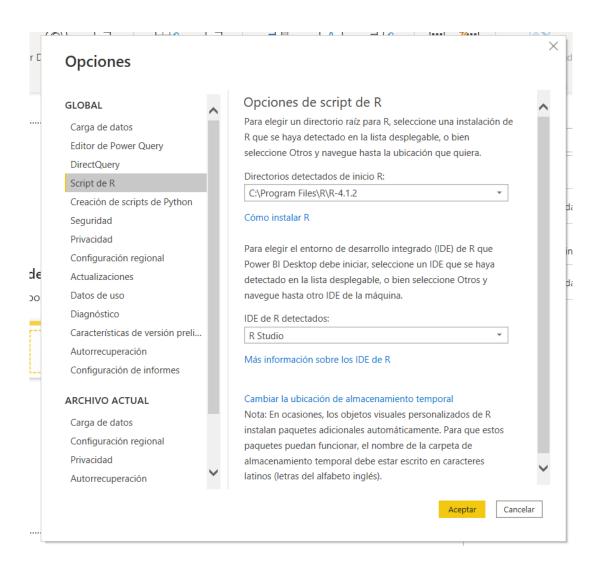
## Cargar los datos



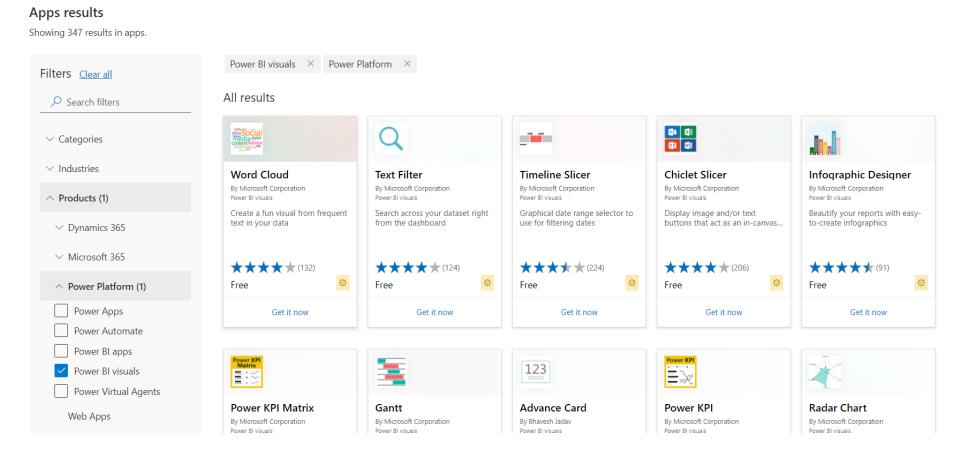




### Revisar los Errores por tipo de dato



Revisar el directorio raíz



Sign in

Microsoft | AppSource More V

Explorar las gráficos que existen para importar a tu informe

https://appsource.microsoft.com/en-us/marketplace/apps?page=1&product=power-bi-visuals



