

Extracción y transformación de los datos/ Creación del tablero Power BI

Se detallan los pasos que me llevaron al código que está en GitHub y algunas reflexiones

Primero se importan las librerías necesarias y consumimos la API con response, después visualizamos el archivo JSON

```
[1] 1 #Importar las librerías necesarias
    2 import requests
    3 import pandas as pd
    4 import json
    5 from datetime import datetime
```

```
[2] 1 #Consumimos la API
    2 url='https://rickandmortyapi.com/api/character'
    3 response=requests.get(url)
```

```
[3] 1 #Exploramos el archivo json
    2 data= response.json()
    3 data
```

```
{'info': {'count': 826,
          'pages': 42,
          'next': 'https://rickandmortyapi.com/api/character?page=2',
          'prev': None},
 'results': [{'id': 1,
              'name': 'Rick Sanchez',
              'status': 'Alive',
              'species': 'Human',
              'type': '',
              'gender': 'Male',
              ...}]
```

```
[4] 1 #Podemos verlo como un dataframe de pandas
    2 df = pd.json_normalize(data)
    3 df.head(10)
```

	results	info.count	info.pages	info.next	info.prev
0	[{'id': 1, 'name': 'Rick Sanchez', 'status': '...	826	42	https://rickandmortyapi.com/api/character?page=2	None

Maneje el archivo JSON como un dataframe para ver que tenía 5 columnas y el registro de una columna era otro JSON, esa fue uno de las reflexiones que tuve, ya que durante la entrevista se nos presentó un caso parecido y viéndolo ya en la práctica me hizo mucho sentido extraer esa columna

como otro JSON; el resto de columnas fueron desechadas porque no tenían información relevante para el análisis pues solo nos presentaban un conteo de los personajes y las páginas previas, siguiente y total de páginas.

```
[9] 1 df1=pd.json_normalize(data["results"])
```

```
[10] 1 #Obtenemos información de nuestro nuevo dataframe
      2 df1.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 20 entries, 0 to 19
Data columns (total 14 columns):
#   Column              Non-Null Count  Dtype
---  ---
0   id                   20 non-null    int64
1   name                 20 non-null    object
2   status               20 non-null    object
3   species              20 non-null    object
4   type                 20 non-null    object
5   gender               20 non-null    object
6   image                20 non-null    object
7   episode              20 non-null    object
8   url                  20 non-null    object
9   created              20 non-null    object
10  origin.name          20 non-null    object
11  origin.url           20 non-null    object
12  location.name        20 non-null    object
13  location.url         20 non-null    object
```

En una primera inspección la información vimos que tenía todos los campos llenos, sin embargo, haciendo una inspección y con ayuda de la función “unique()”; pude darme cuenta que algunos datos tenían el registro “” y los rellene con “sin (campo)”, tuve que realizar ese proceso para las columnas de “type” y “origin.url” con ayuda de la función replace

```
14] 1 print(df1["type"].unique())
      2 #Sin embargo algunos campos tienen "" en type, mostramos dichos registros
      3 df1[df1["type"]==""]
```

```
['' 'Genetic experiment' 'Superhuman (Ghost trains summoner)' 'Parasite'
 'Human with antennae' 'Human with ants in his eyes']
```

	id	name	status	species	type	gender	image	episode
0	1	Rick Sanchez	Alive	Human		Male	https://rickandmortyapi.com/api/character/avat...	[https://rickandmortyapi.com/api/episode/1, ht...
1	2	Morty Smith	Alive	Human		Male	https://rickandmortyapi.com/api/character/avat...	[https://rickandmortyapi.com/api/episode/1, ht...
2	3	Summer Smith	Alive	Human		Female	https://rickandmortyapi.com/api/character/avat...	[https://rickandmortyapi.com/api/episode/6, ht...
3	4	Beth Smith	Alive	Human		Female	https://rickandmortyapi.com/api/character/avat...	[https://rickandmortyapi.com/api/episode/6, ht...

```
[15] 1 df1["type"]=df1["type"].replace("", "Without type")
```

```
[16] 1 df1
```

	id	name	status	species	type	gender	image	episode
0	1	Rick Sanchez	Alive	Human	Without type	Male	https://rickandmortyapi.com/api/character/avat...	https://rickandmortyapi.com/api/episode/1, ht...
1	2	Morty Smith	Alive	Human	Without type	Male	https://rickandmortyapi.com/api/character/avat...	https://rickandmortyapi.com/api/episode/1, ht...
2	3	Summer Smith	Alive	Human	Without type	Female	https://rickandmortyapi.com/api/character/avat...	https://rickandmortyapi.com/api/episode/6, ht...
3	4	Beth Smith	Alive	Human	Without type	Female	https://rickandmortyapi.com/api/character/avat...	https://rickandmortyapi.com/api/episode/6, ht...
4	5	Jerry Smith	Alive	Human	Without type	Male	https://rickandmortyapi.com/api/character/avat...	https://rickandmortyapi.com/api/episode/6, ht...

```
1 df1["origin.url"]=df1["origin.url"].replace("", "Without URL")
2 df1
```

Posteriormente pude observar que algunos registros eran “uknow” y el formato de nuestros datos es con la primera letra mayúscula, así que convertí el “uknow” a “Uknow” con la misma función replace, pero está la use para toda la tabla y no por columnas como en el caso anterior

```
[18] 1 #Pasamos el unknown a Uknow
```

```
[19] 1 df1=df1.replace("unknown", "Uknow")
```

```
[20] 1 df1
```

4	5	Jerry Smith	Alive	Human	Without type	Male	https://rickandmortyapi.com/api/character/avat...
5	6	Abadango Cluster Princess	Alive	Alien	Without type	Female	https://rickandmortyapi.com/api/character/avat...
6	7	Abradolf Lincler	Uknow	Human	Genetic experiment	Male	https://rickandmortyapi.com/api/character/avat...
7	8	Adjudicator Rick	Dead	Human	Without type	Male	https://rickandmortyapi.com/api/character/avat...
8	9	Agency Director	Dead	Human	Without type	Male	https://rickandmortyapi.com/api/character/avat...

Ya iba mejorando la tabla, sin embargo los registros de “episode” eran listas

```
21] 1 type(df1.episode[0])

list
```

Una de las cosas que aprendí fue el uso de la función “explode” para convertir cada elemento de la lista de “episode” en una fila

```
1 #usamos la funcion explode para convertir cada elemento de la lista de episode en una fila
2 df1=df1.explode("episode")

[23] 1 df1.created=pd.to_datetime(df1.created)
```

```
1 df1
```

	id	name	status	species	type	gender	image	episode
0	1	Rick Sanchez	Alive	Human	Without type	Male	https://rickandmortyapi.com/api/character/avat...	https://rickandmortyapi.com/api/episode/1
0	1	Rick Sanchez	Alive	Human	Without type	Male	https://rickandmortyapi.com/api/character/avat...	https://rickandmortyapi.com/api/episode/2
0	1	Rick Sanchez	Alive	Human	Without type	Male	https://rickandmortyapi.com/api/character/avat...	https://rickandmortyapi.com/api/episode/3
0	1	Rick Sanchez	Alive	Human	Without type	Male	https://rickandmortyapi.com/api/character/avat...	https://rickandmortyapi.com/api/episode/4
0	1	Rick Sanchez	Alive	Human	Without type	Male	https://rickandmortyapi.com/api/character/avat...	https://rickandmortyapi.com/api/episode/5

✓ 0 s completado a las 15:15

Luego renombre las columnas con un mismo formato con ayuda de “rename”

```
1 df1=df1.rename(columns={"id":"ID","name":"Name","status":"Status","species":"Specie",
2                             "type":"Type","image":"Image","episode":"Appearance_episode","url":"URL",
3                             "created":"Created_date","origin.name":"Origin","origin.url":"Origin_URL",
4                             "location.name":"Location","location.url":"Location_URL"
5 })
```

Y finalmente hice el ID nuestro índice de columna

```
1 df1=df1.set_index("ID")

[ ] 1 df1
```

	Name	Status	Specie	Type	gender	Image	Appearance_episode	URL
ID								
1	Rick Sanchez	Alive	Human	Without type	Male	https://rickandmortyapi.com/api/character/avat...	https://rickandmortyapi.com/api/episode/1	https://rickandmortyapi.com/api/character/1
1	Rick Sanchez	Alive	Human	Without type	Male	https://rickandmortyapi.com/api/character/avat...	https://rickandmortyapi.com/api/episode/2	https://rickandmortyapi.com/api/character/1
1	Rick Sanchez	Alive	Human	Without type	Male	https://rickandmortyapi.com/api/character/avat...	https://rickandmortyapi.com/api/episode/3	https://rickandmortyapi.com/api/character/1
1	Rick	Alive	Human	Without	Male	https://rickandmortyapi.com/api/character/avat...	https://rickandmortyapi.com/api/episode/4	https://rickandmortyapi.com/api/character/1

Finalmente quite las visualizaciones que hace colab y subí el código a GitHub

Power BI

Una vez teniendo la tabla limpia, otra de las cosas que aprendí es que se puede usar un script de Python en power BI, ya que yo lo hacía importando directamente tablas para ello nos dirigimos a la parte de obtener datos y nos vamos a la sección de Python y colocamos nuestro Script

excepción de “Image” y seleccionamos la opción de recuento de filas (que nos dara un recuento de los episodios por personaje)

Agrupar por

Especifique las columnas por las que quiera realizar la agrupación y una o más salidas.

☐ Básico ☒ Uso avanzado

Name
Status
Specie
Type
Gender
Image

Agregar agrupación

Nuevo nombre de columna Operación Columna

Recuento Recuento de filas

Agregar agregación

Una vez hecho esto, nos vamos a la fila de “Image” y en “Herramientas de columnas” en la parte de “Categoría de datos” ponemos “Dirección URL de la imagen” para que podamos visualizar las imagenes

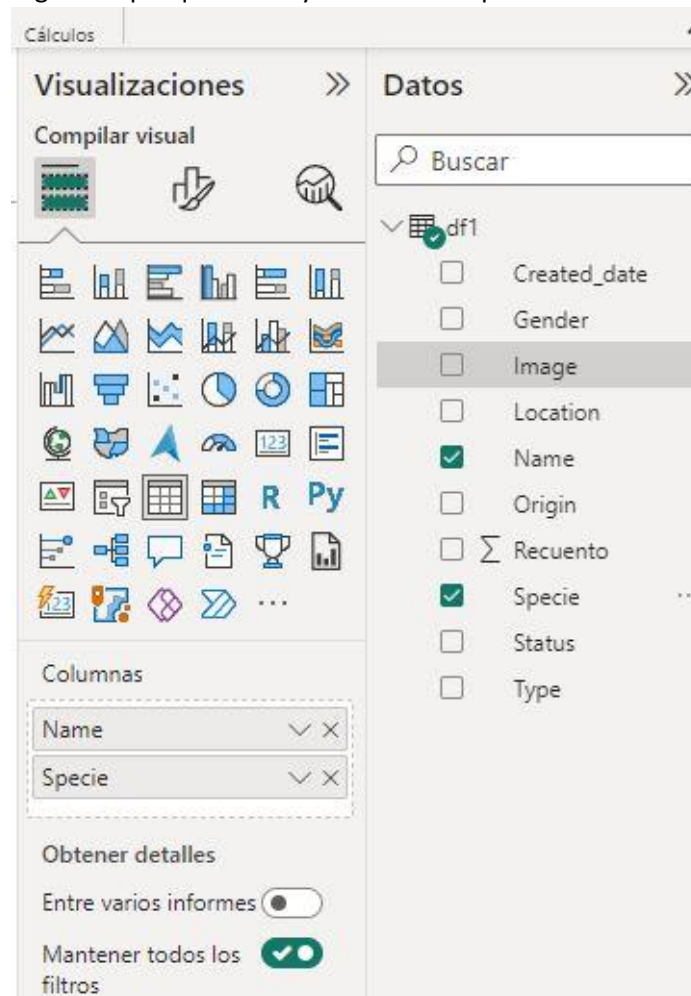
Gender	Image
male	https://rickandmortyapi.com/api/cha
male	https://rickandmortyapi.com/api/cha
male	https://rickandmortyapi.com/api/cha
male	https://rickandmortyapi.com/api/cha
male	https://rickandmortyapi.com/api/cha
male	https://rickandmortyapi.com/api/cha
male	https://rickandmortyapi.com/api/cha
male	https://rickandmortyapi.com/api/cha
male	https://rickandmortyapi.com/api/cha
male	https://rickandmortyapi.com/api/cha
male	https://rickandmortyapi.com/api/cha
male	https://rickandmortyapi.com/api/cha
male	https://rickandmortyapi.com/api/cha
male	https://rickandmortyapi.com/api/cha
male	https://rickandmortyapi.com/api/cha

Resumen No resumir Ordenar p...
Categoría de datos Sin clasificar Columna







Sin clasificar

- Dirección
- Lugar
- Ciudad
- Condado
- Estado o provincia
- Código postal
- País o región
- Continente
- Latitud
- Longitud
- URL web
- Dirección URL de la imagen
- Código de barras

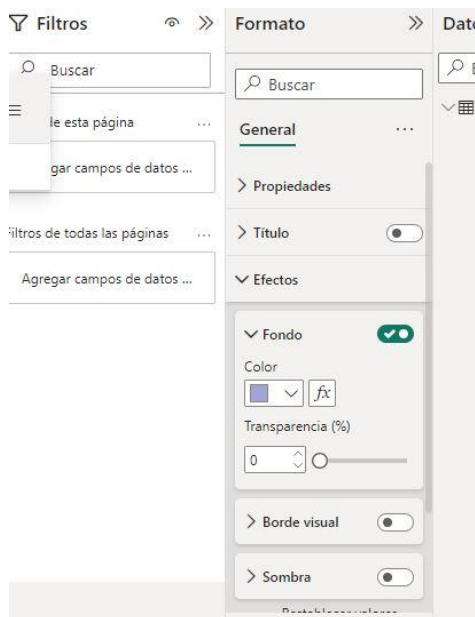
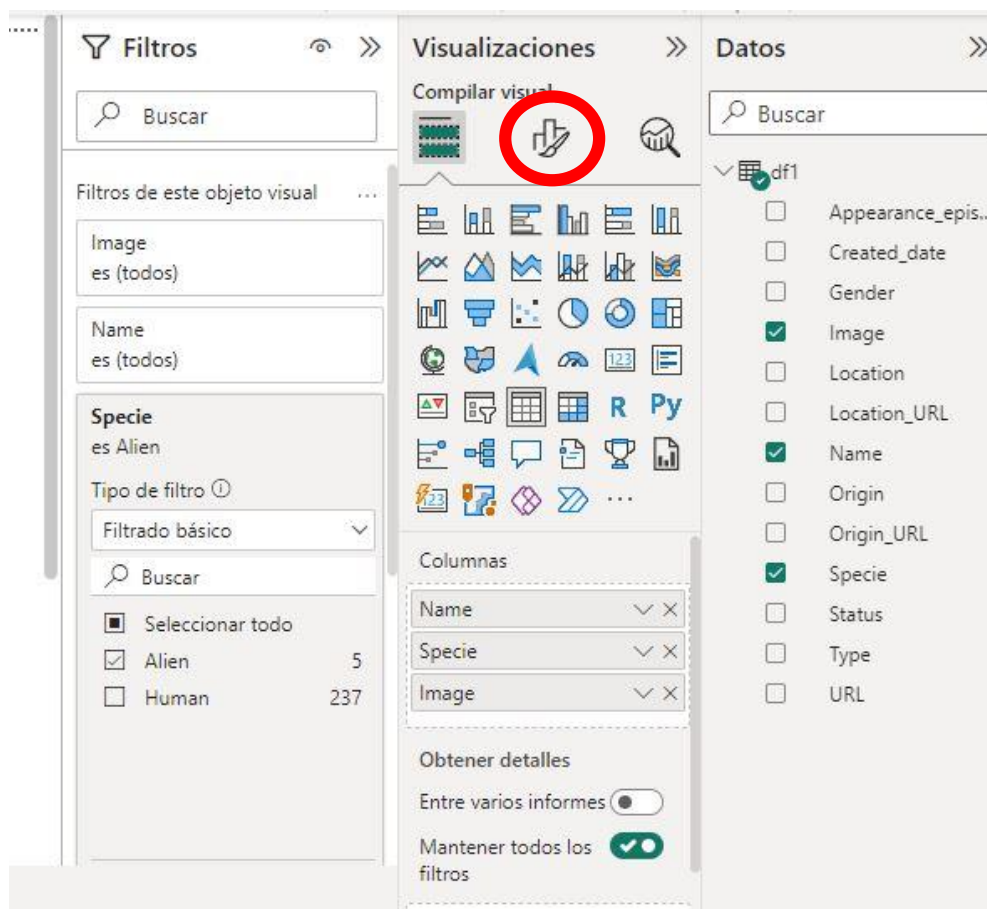
Le damos en aplicar y cerrar y nos vamos a la parte de visualizaciones, ahí elegimos el tipo de gráfica que queremos y las variables que vamos a usar



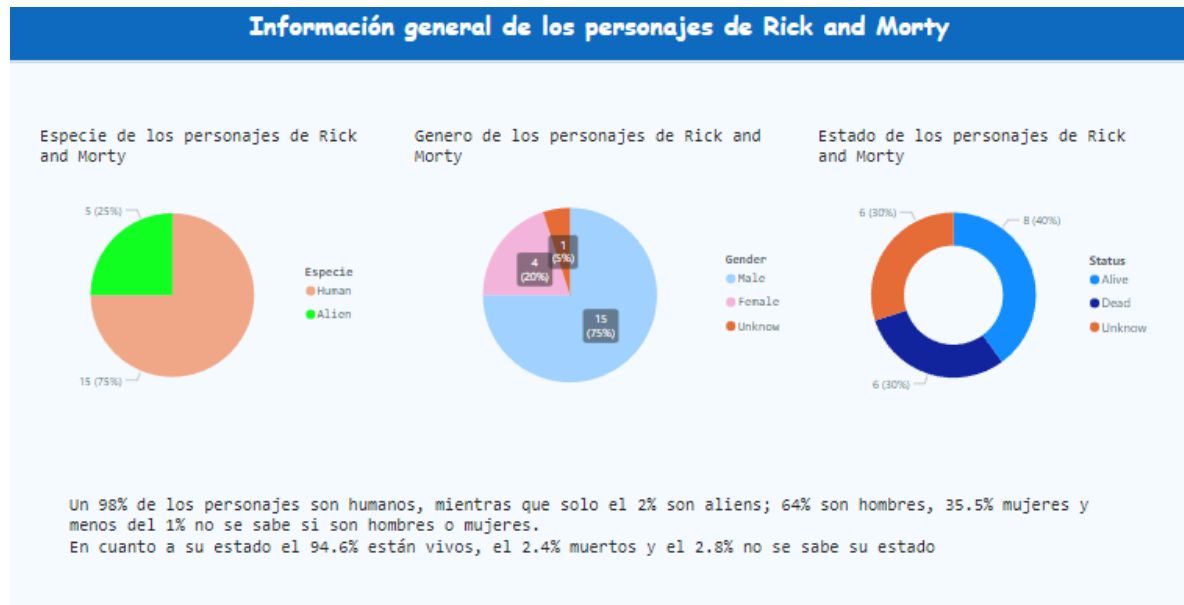
Agregamos un cuadro de texto y lo posicionamos, damos fuente a la letra y colores de fondo

Name	Specie	Image
Abadango Cluster Princess	Alien	
Abradolf Lincler	Human	
Adjudicator Rick	Human	
Agency Director	Human	
Alan Rails	Human	
Albert Einstein	Human	

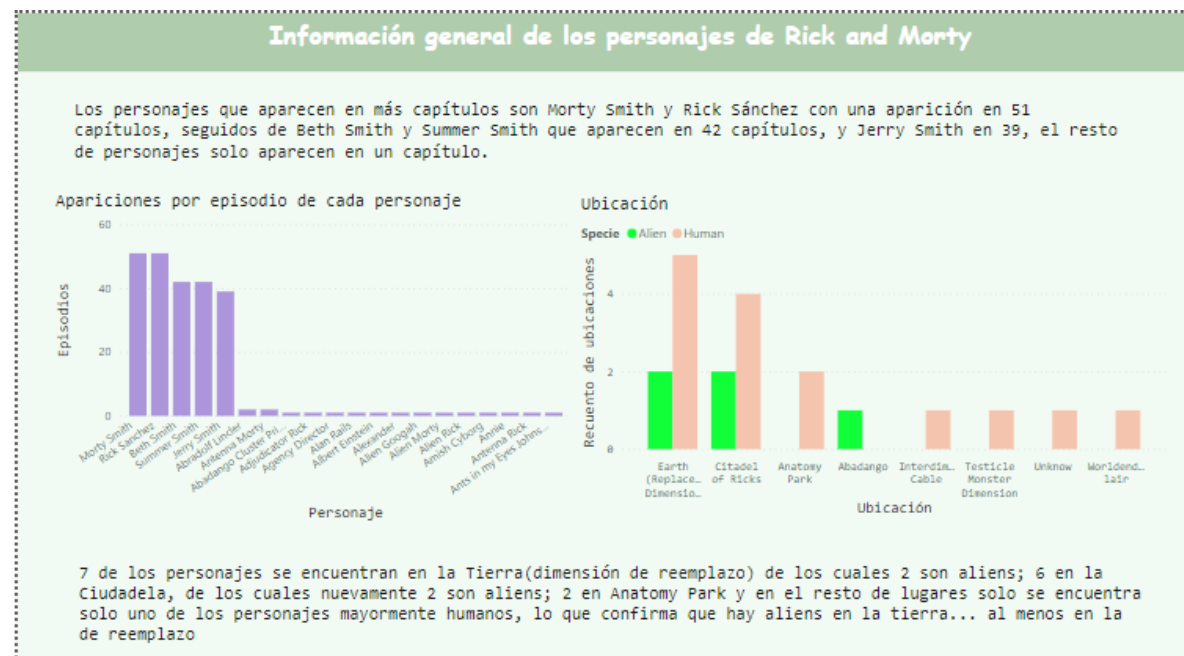
Importante: Para dar formato como fondo, títulos de los gráficos, colores de los gráficos, fondo del lienzo, etc. Seleccionamos el elemento que queremos modificar y le damos click en el siguiente icono



Y realizamos varias gráficas que nos puedan brindar información de los personajes de acuerdo a los datos que tenemos, y les damos formato como se explicó arriba

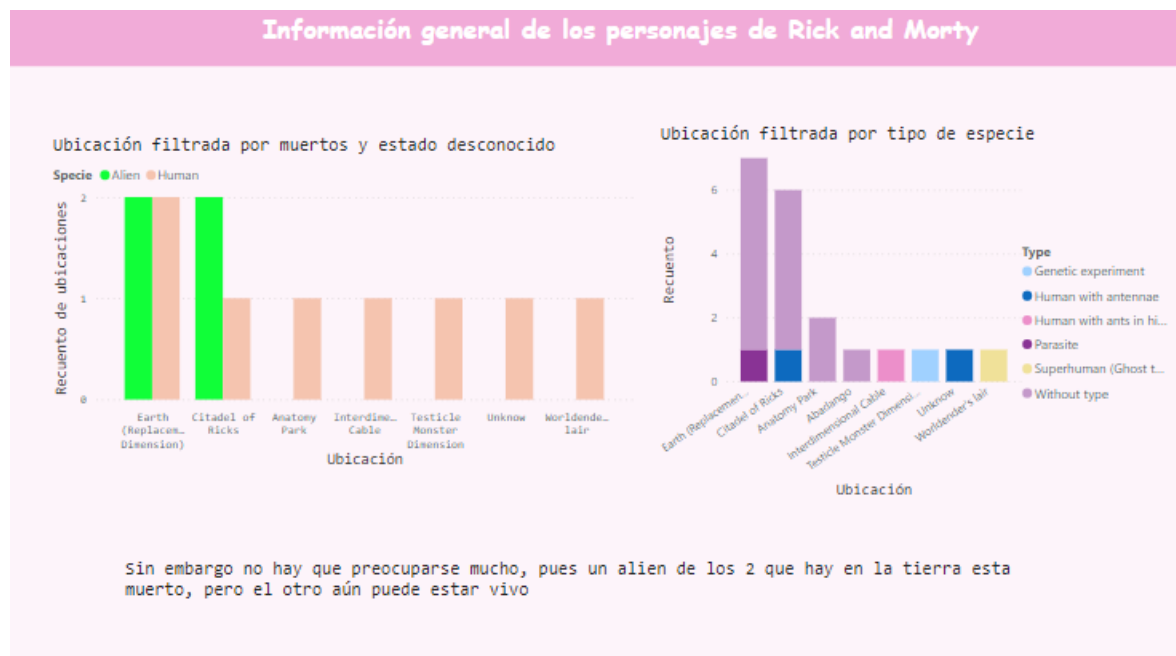


Hice una visualización para ver cuantos personajes eran humanos, aliens; cuantos eran mujeres, hombres o de genero desconocido; también de los que estaban vivos, muertos o no se conoce su estado; obteniendo que la mayoría son humanos, hombres, y están vivos, respectivamente.

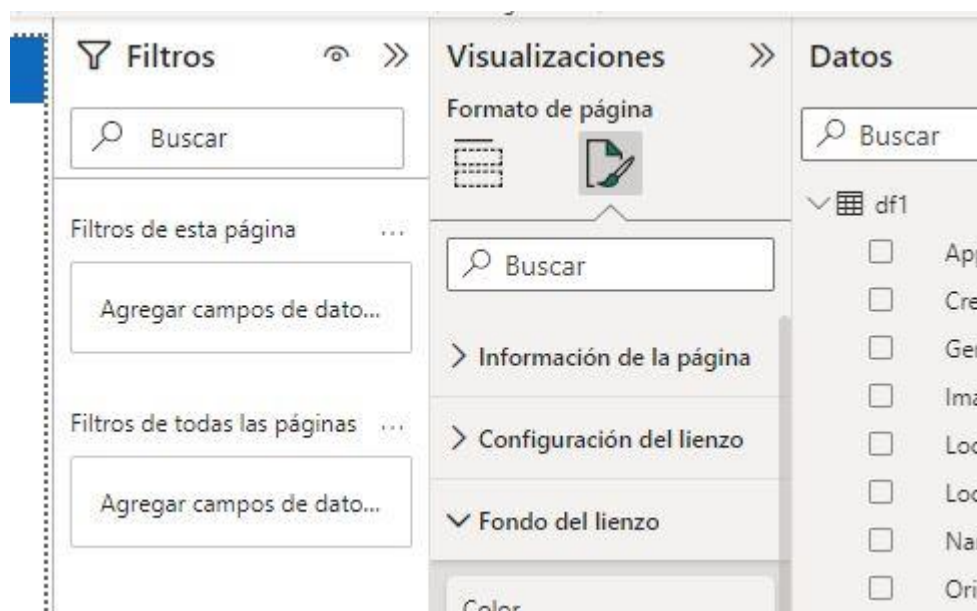


También me pareció interesante ver quiénes son los personajes que más aparecen en la serie y son los protagonistas pero sin ver la serie puedo intuir que su familia es Beth, Summer y Jerry porque son personajes muy recurrentes y además tienen el mismo apellido.

Además podemos ver otro gráfico que nos muestra los registros por ubicación, separados por humanos y aliens, podemos ver que en la tierra hay más humanos que aliens pero hay aliens! Y en algunos lugares no hay aliens y en uno solo no hay humanos.










Sin embargo, si nos quedamos preocupados porque hay aliens en la tierra, podemos ver su estado filtrando las columnas, hacemos un gráfico parecido al azul pero con la condición de que no nos enseñe a los vivos



Así obtenemos la primer tabla del lienzo rosita, que nos dice que los dos aliens están o muertos o no se desconoce su estado, sin embargo había una categoría de tipo de especie, para ver si había alguna otra criatura extraña en la tierra de reemplazo, y solo había parásitos que eran los aliens que mencionamos.

Pero en Worldender's lair hay un superhumano! De acuerdo a la segunda gráfica.

Finalmente podemos ver una tabla con foto de los registros que queramos, seleccionando en tipo de gráfica el icono de la tabla

Personajes de Rick and Morty con nombre, especie y foto		
Name	Specie	Image
Abadango Cluster Princess	Alien	
Abradolf Lincler	Human	
Adjudicator Rick	Human	
Agency Director	Human	
Alan Rails	Human	
Albert Einstein	Human	
Alexander	Human	


De igual forma podemos filtrar información para ver quiénes son aliens...





Personajes de Rick and Morty que humanos o aliens		
Name	Specie	Image
Amish Cyborg	Alien	
Alien Rick	Alien	
Alien Morty	Alien	
Alien Googah	Alien	
Abadango Cluster Princess	Alien	

Name	Specie	Image
Abradolf Lincler	Human	
Adjudicator Rick	Human	
Agency Director	Human	
Alan Rails	Human	
Albert Einstein	Human	

O quienes son mujeres y hombres y conocerlos físicamente. El límite es la creatividad.

Personajes de Rick and Morty por género

Name	Gender	Image
Summer Smith	Female	
Beth Smith	Female	
Annie	Female	
Abadango Cluster Princess	Female	

Name	Gender	Image
Abradolf Lincler	Male	
Adjudicator Rick	Male	
Agency Director	Male	
Alan Rails	Male	

Name	Gender	Image
Alien Googah	Unknow	