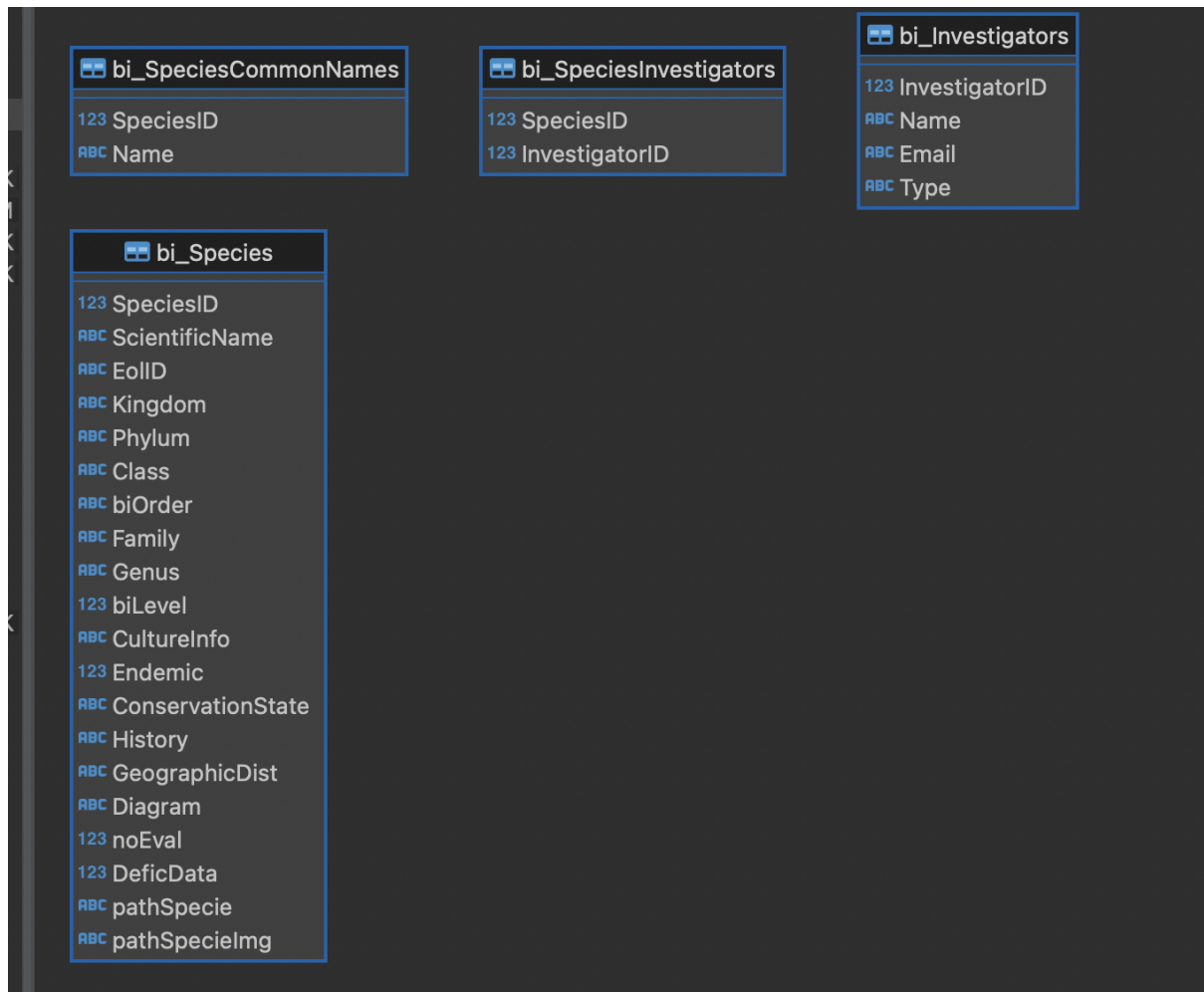


Examen Gestion del Dato

PUNTO 1

Para la ejecución del punto 1, realizamos una conexión previa a la BBDD del ejercicio para entender el modelo ER, para ello usamos el aplicativo dbeaver para ver el diagrama



De él podemos observar los pasos para la construcción de los pasos. Observamos que el modelo tiene una relación 1 a varios entre las fuentes y la tabla **bi_speciesinvestigators** es una tabla puente entre ellos que cumple un formato varios a varios en términos de relación.

Para la elaboración de las consultas, dejó anexo de las imágenes de la consulta del primer archivo y del segundo archivo.

```

-- archivo 1
SELECT bs.SpeciesID,
       bs.ScientificName,
       bscn.Name as species_common_name,
       bs.Kingdom,
       bi2.Name as name_investigators
FROM Bio.bi_Species bs
LEFT JOIN Bio.bi_SpeciesCommonNames bscn ON (bs.SpeciesID = bscn.SpeciesID)
LEFT JOIN Bio.bi_SpeciesInvestigators bsi ON (bsi.SpeciesID = bs.SpeciesID)
LEFT JOIN Bio.bi_Investigators bi2 ON (bi2.InvestigatorID = bsi.InvestigatorID)
WHERE bs.SpeciesID IN (205667, 279652, 282209, 272325, 285318, 261205, 261205, 275717, 281039, 279652)

--- Archivo 2
SELECT bi2.Name as name_investigators,
       bs.Kingdom,
       COUNT(bs.SpeciesID) as conteo_especies
FROM Bio.bi_Species bs
--LEFT JOIN Bio.bi_SpeciesCommonNames bscn ON (bs.SpeciesID = bscn.SpeciesID)
LEFT JOIN Bio.bi_SpeciesInvestigators bsi ON (bsi.SpeciesID = bs.SpeciesID)
LEFT JOIN Bio.bi_Investigators bi2 ON (bi2.InvestigatorID = bsi.InvestigatorID)
WHERE bs.SpeciesID IN (205667, 279652, 282209, 272325, 285318, 261205, 261205, 275717, 281039, 279652)
GROUP BY bi2.Name, bs.Kingdom

```

Nota: Nótese que para cumplir los objetivos de la consulta tenemos las siguientes condiciones:

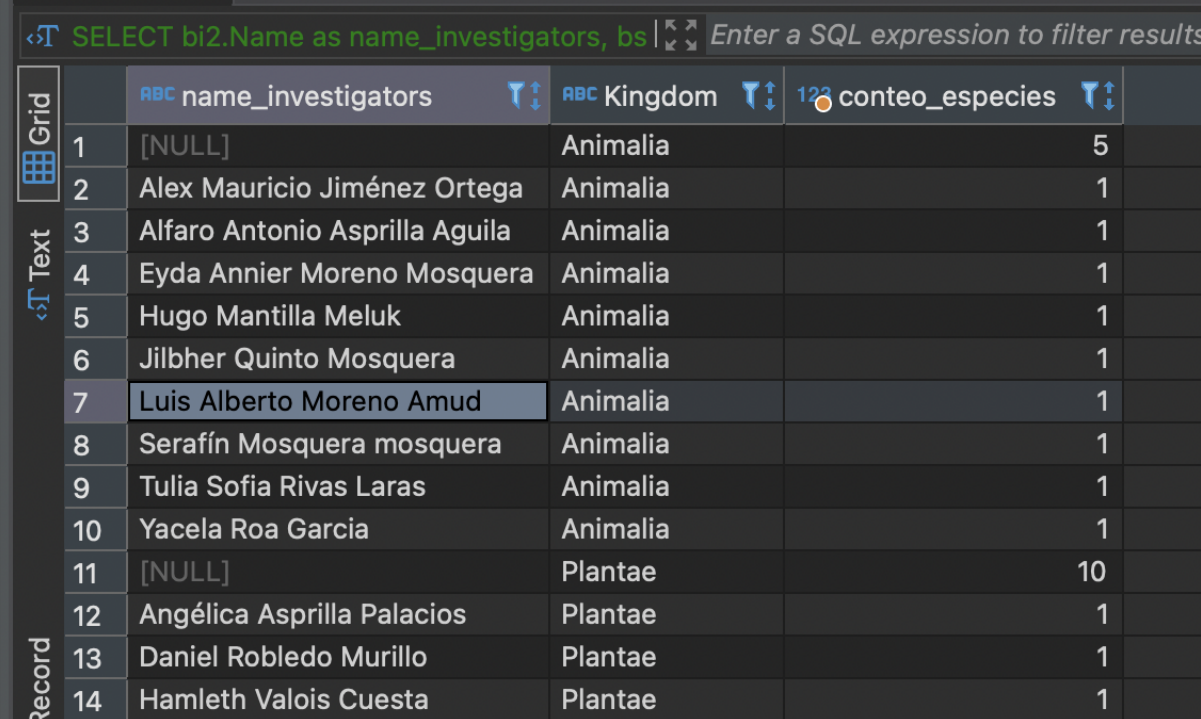
1. Realizamos el filtro de las especies consideradas en el enunciado para trabajar.
2. para no considerar más de una especie, agregamos una relación join left, la cual se podrá identificar elementos que no cumplen estas características.

notas consulta 1. podemos observar que existen especies que no tienen un investigador asociado.

| | ID | ScientificName | species_common_name | Kingdom | name_investigators |
|-----|--------|--------------------------------|----------------------------|----------|----------------------------|
| 111 | 2,209 | Bixa orellana L. | Bija blanca | Plantae | Wiston Ivan Murillo Moreno |
| 112 | 2,209 | Bixa orellana L. | Bija de concha blanca | Plantae | Wiston Ivan Murillo Moreno |
| 113 | 2,209 | Bixa orellana L. | Bija, achioté | Plantae | Wiston Ivan Murillo Moreno |
| 114 | 2,209 | Bixa orellana L. | Biza, achioté, onoto, bija | Plantae | Wiston Ivan Murillo Moreno |
| 115 | 2,209 | Bixa orellana L. | Vija | Plantae | Wiston Ivan Murillo Moreno |
| 116 | 31,039 | Cedrela fissilis Vell. | Cedro colorado | Plantae | [NULL] |
| 117 | 31,039 | Cedrela fissilis Vell. | Cedro misionero | Plantae | [NULL] |
| 118 | 31,039 | Cedrela fissilis Vell. | Cedro paraguayo | Plantae | [NULL] |
| 119 | 31,039 | Cedrela fissilis Vell. | Cedro real | Plantae | [NULL] |
| 120 | 31,039 | Cedrela fissilis Vell. | Ygaí | Plantae | [NULL] |
| 121 | 31,039 | Cedrela fissilis Vell. | Ygary | Plantae | [NULL] |
| 122 | 35,318 | Aniba perutilis Hemsl. | Chachajo, comino | Plantae | [NULL] |
| 123 | 9,842 | Ormosia macrophylla Benth. | Amargo | Plantae | [NULL] |
| 124 | 22,319 | Rhaebo haematiticus (Cope, 186 | Sapo | Animalia | [NULL] |

notas consulta 2:

Como lo vimos en la consulta 1, al agrupar, podemos ver que existen especies sin investigador asociada, pero si a un reino.



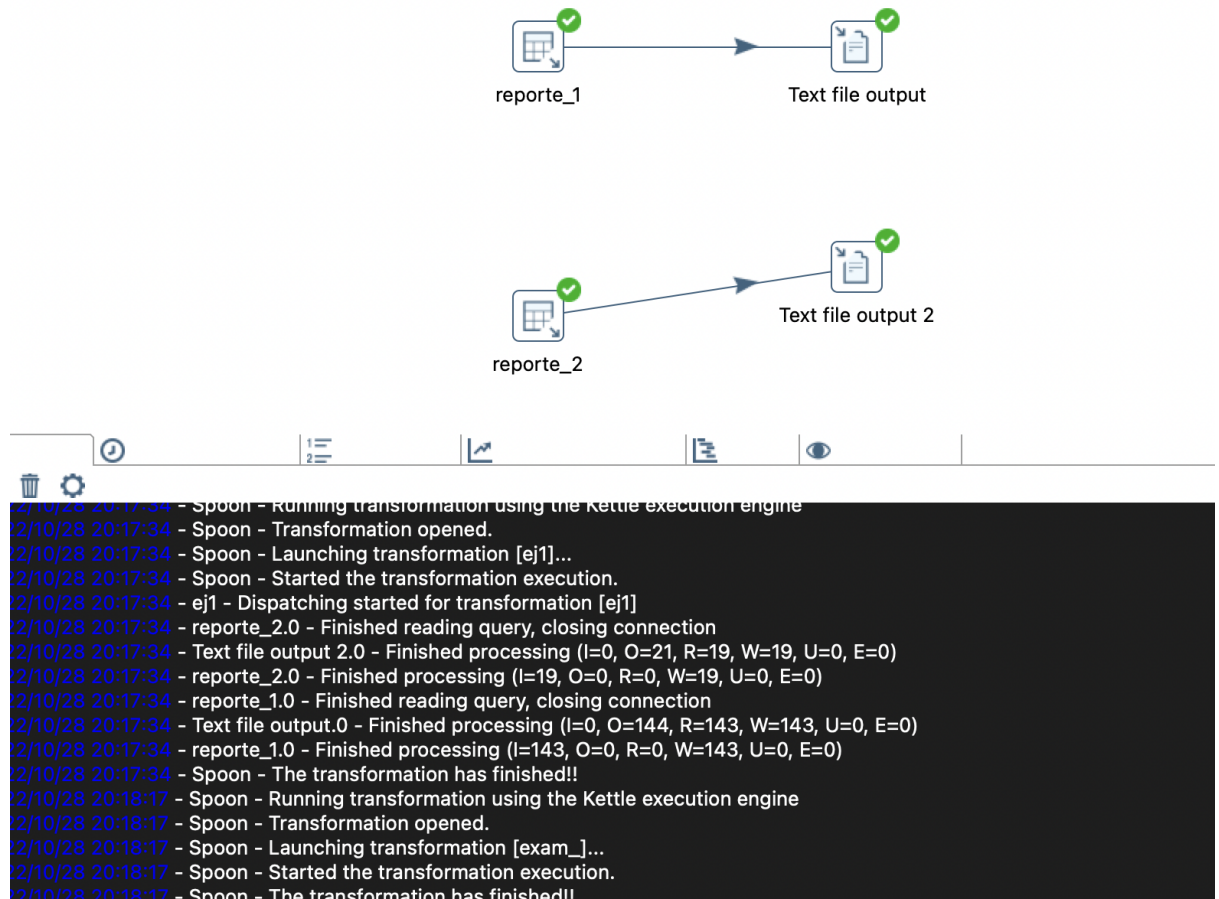
The screenshot shows a SQL query editor with a dark theme. At the top, a SQL query is entered: `SELECT bi2.Name as name_investigators, bs`. Below the query, a table displays the results. The table has four columns: `name_investigators`, `Kingdom`, and `conteo_especies`. The results are grouped by Kingdom, with rows 11 and 12 showing `Plantae` and rows 1-10 showing `Animalia`. Row 1 shows a `[NULL]` value for the investigator. The interface includes a 'Grid' view icon on the left and a 'Text' view icon at the top left of the table. A 'Record' label is visible on the far left.

| | name_investigators | Kingdom | conteo_especies |
|----|--------------------------------|----------|-----------------|
| 1 | [NULL] | Animalia | 5 |
| 2 | Alex Mauricio Jiménez Ortega | Animalia | 1 |
| 3 | Alfaro Antonio Asprilla Aguila | Animalia | 1 |
| 4 | Eyda Annier Moreno Mosquera | Animalia | 1 |
| 5 | Hugo Mantilla Meluk | Animalia | 1 |
| 6 | Jilbher Quinto Mosquera | Animalia | 1 |
| 7 | Luis Alberto Moreno Amud | Animalia | 1 |
| 8 | Serafín Mosquera mosquera | Animalia | 1 |
| 9 | Tulia Sofia Rivas Laras | Animalia | 1 |
| 10 | Yacela Roa Garcia | Animalia | 1 |
| 11 | [NULL] | Plantae | 10 |
| 12 | Angélica Asprilla Palacios | Plantae | 1 |
| 13 | Daniel Robledo Murillo | Plantae | 1 |
| 14 | Hamleth Valois Cuesta | Plantae | 1 |

Ahora para el paso de pentaho, con las credenciales previamente predefinidas, realizamos un job con dos flujos que cumplen la siguiente condición:

- Extracción de una consulta SQL a la base de datos, previamente enviada y desarrollada.

- Extracción de la información con los pasos mencionados en el documento, de formato, y separador.



Los archivos resultantes son:

archivo 1

| | A | B | C | D | E |
|---|--------------|------------------------------------|---------------------------|----------|--------------------------------|
| | consulta_1_1 | | | | |
| | SpeciesID | ScientificName | species_common_name | Kingdom | name_investigators |
| 2 | 241523 | Artibeus lituratus (Olfers, 1818) | Gran Murci@lago frutero | Animalia | Alex Mauricio Jiménez Ortega |
| 3 | 241523 | Artibeus lituratus (Olfers, 1818) | Murci@lago frutero grande | Animalia | Alex Mauricio Jiménez Ortega |
| 4 | 241523 | Artibeus lituratus (Olfers, 1818) | Gran Murci@lago frutero | Animalia | Jilbher Quinto Mosquera |
| 5 | 241523 | Artibeus lituratus (Olfers, 1818) | Murci@lago frutero grande | Animalia | Jilbher Quinto Mosquera |
| 3 | 241523 | Artibeus lituratus (Olfers, 1818) | Gran Murci@lago frutero | Animalia | Alfaro Antonio Asprilla Aguila |
| 7 | 241523 | Artibeus lituratus (Olfers, 1818) | Murci@lago frutero grande | Animalia | Alfaro Antonio Asprilla Aguila |
| 3 | 241523 | Artibeus lituratus (Olfers, 1818) | Gran Murci@lago frutero | Animalia | Hugo Mantilla Meluk |
| 3 | 241523 | Artibeus lituratus (Olfers, 1818) | Murci@lago frutero grande | Animalia | Hugo Mantilla Meluk |
| 0 | 241523 | Artibeus lituratus (Olfers, 1818) | Gran Murci@lago frutero | Animalia | Eyda Annier Moreno Mosquera |
| 1 | 241523 | Artibeus lituratus (Olfers, 1818) | Murci@lago frutero grande | Animalia | Eyda Annier Moreno Mosquera |
| 2 | 241523 | Artibeus lituratus (Olfers, 1818) | Gran Murci@lago frutero | Animalia | Yacela Roa Garcia |
| 3 | 241523 | Artibeus lituratus (Olfers, 1818) | Murci@lago frutero grande | Animalia | Yacela Roa Garcia |
| 4 | 241523 | Artibeus lituratus (Olfers, 1818) | Gran Murci@lago frutero | Animalia | Luis Alberto Moreno Amud |
| 5 | 241523 | Artibeus lituratus (Olfers, 1818) | Murci@lago frutero grande | Animalia | Luis Alberto Moreno Amud |
| 6 | 241523 | Artibeus lituratus (Olfers, 1818) | Gran Murci@lago frutero | Animalia | Serafin Mosquera mosquera |
| 7 | 241523 | Artibeus lituratus (Olfers, 1818) | Murci@lago frutero grande | Animalia | Serafin Mosquera mosquera |
| 8 | 244210 | Xenodon rabdocephalus (Wied, 1824) | Culebra | Animalia | Tulia Sofia Rivas Laras |
| 9 | 244210 | Xenodon rabdocephalus (Wied, 1824) | Falsa equis | Animalia | Tulia Sofia Rivas Laras |
| 0 | 261205 | Ammandra decasperma O.F.Cook | Antá | Plantae | Hamleth Valois Cuesta |
| 1 | 261205 | Ammandra decasperma O.F.Cook | Cabecilla | Plantae | Hamleth Valois Cuesta |
| 2 | 261205 | Ammandra decasperma O.F.Cook | Cabecita | Plantae | Hamleth Valois Cuesta |
| 3 | 261205 | Ammandra decasperma O.F.Cook | Ñume | Plantae | Hamleth Valois Cuesta |
| 4 | 261205 | Ammandra decasperma O.F.Cook | Ñume blanco | Plantae | Hamleth Valois Cuesta |
| 5 | 261205 | Ammandra decasperma O.F.Cook | Tagua | Plantae | Hamleth Valois Cuesta |
| 6 | 261205 | Ammandra decasperma O.F.Cook | Yumal | Plantae | Hamleth Valois Cuesta |
| 7 | 261205 | Ammandra decasperma O.F.Cook | Yume | Plantae | Hamleth Valois Cuesta |
| 8 | 261205 | Ammandra decasperma O.F.Cook | Tagua blanca | Plantae | Hamleth Valois Cuesta |
| 9 | 261205 | Ammandra decasperma O.F.Cook | Tagua de agua | Plantae | Hamleth Valois Cuesta |

archivo 2

consulta_1_2

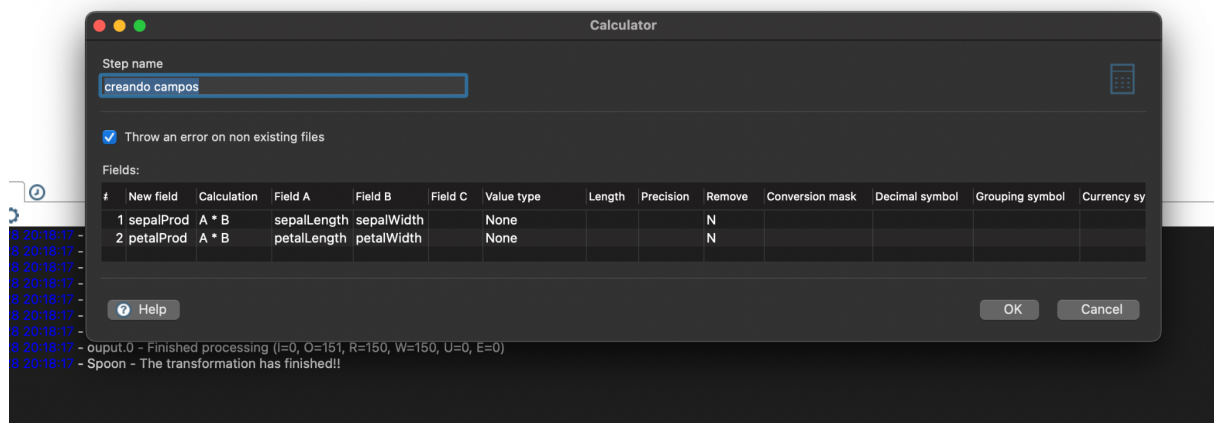
Table data was imported. [Adjust](#)

| name_investigators | Kingdom | conteo_especies |
|--------------------------------|----------|-----------------|
| | Animalia | 5 |
| Alex Mauricio Jiménez Ortega | Animalia | 1 |
| Alfaro Antonio Asprilla Aguila | Animalia | 1 |
| Eyda Annier Moreno Mosquera | Animalia | 1 |
| Hugo Mantilla Meluk | Animalia | 1 |
| Jilbher Quinto Mosquera | Animalia | 1 |
| Luis Alberto Moreno Amud | Animalia | 1 |
| Serafín Mosquera mosquera | Animalia | 1 |
| Tulia Sofia Rivas Laras | Animalia | 1 |
| Yacela Roa Garcia | Animalia | 1 |
| | Plantae | 10 |
| Angélica Asprilla Palacios | Plantae | 1 |
| Daniel Robledo Murillo | Plantae | 1 |
| Hamleth Valois Cuesta | Plantae | 1 |
| Helcias Jose Ayala Mosquera | Plantae | 1 |
| Jovanny Mosquera Pino | Plantae | 1 |
| Luis Javier Mosquera Ramos | Plantae | 1 |
| Nayive Pino Benitez | Plantae | 1 |
| Wiston Ivan Murillo Moreno | Plantae | 1 |
| name_investigators | Kingdom | conteo_especies |

PUNTO 2.

Para el desarrollo del punto 1, iniciamos con una lectura general del archivo csv iris y usando el objeto **calculator** podemos construir los campos como indica la imagen de abajo. Finalmente para expulsar los elementos, consideramos el objeto de extraer:

- Para ello en el objeto de salida realizamos un intercambio de campos, para poder expulsar los campos en sus objetos.
- Y ajustamos los tipos de datos.



Los resultados son:

iris.data

| sepalLength | sepalWidth | petalLength | petalWidth | sepalProd | petalProd | class |
|-------------|------------|-------------|------------|--------------------|-----------|-------------|
| 5.1 | 3.5 | 1.4 | 0.2 | 000000000000017.85 | 0.28 | Iris-setosa |
| 4.9 | 3 | 1.4 | 0.2 | 000000000000014.70 | 0.28 | Iris-setosa |
| 4.7 | | 1.3 | 0.2 | 000000000000015.04 | 0.26 | Iris-setosa |
| 4.6 | | 1.5 | 0.2 | 000000000000014.26 | 0.3 | Iris-setosa |
| 5 | 3.6 | 1.4 | 0.2 | 000000000000018.00 | 0.28 | Iris-setosa |
| 5.4 | 3.9 | 1.7 | 0.4 | 000000000000021.06 | 0.68 | Iris-setosa |
| 4.6 | 3.4 | 1.4 | 0.3 | 000000000000015.64 | 0.42 | Iris-setosa |
| 5 | 3.4 | 1.5 | 0.2 | 000000000000017.00 | 0.3 | Iris-setosa |
| 4.4 | 2.9 | 1.4 | 0.2 | 000000000000012.76 | 0.28 | Iris-setosa |
| 4.9 | 3.1 | 1.5 | 0.1 | 000000000000015.19 | 0.15 | Iris-setosa |
| 5.4 | 3.7 | 1.5 | 0.2 | 000000000000019.98 | 0.3 | Iris-setosa |
| 4.8 | 3.4 | 1.6 | 0.2 | 000000000000016.32 | 0.32 | Iris-setosa |
| 4.8 | 3 | 1.4 | 0.1 | 000000000000014.40 | 0.14 | Iris-setosa |
| 4.3 | | 1.1 | 0.1 | 000000000000012.90 | 0.11 | Iris-setosa |
| 5.8 | | 1.2 | 0.2 | 000000000000023.20 | 0.24 | Iris-setosa |
| 5.7 | 4.4 | 1.5 | 0.4 | 000000000000025.08 | 0.6 | Iris-setosa |
| 5.4 | 3.9 | 1.3 | 0.4 | 000000000000021.06 | 0.52 | Iris-setosa |
| 5.1 | 3.5 | 1.4 | 0.3 | 000000000000017.85 | 0.42 | Iris-setosa |
| 5.7 | 3.8 | 1.7 | 0.3 | 000000000000021.66 | 0.51 | Iris-setosa |
| 5.1 | 3.8 | 1.5 | 0.3 | 000000000000019.38 | 0.45 | Iris-setosa |
| 5.4 | 3.4 | 1.7 | 0.2 | 000000000000018.36 | 0.34 | Iris-setosa |
| 5.1 | 3.7 | 1.5 | 0.4 | 000000000000018.87 | 0.6 | Iris-setosa |
| 4.6 | 3.6 | 1 | 0.2 | 000000000000016.56 | 0.2 | Iris-setosa |
| 5.1 | 3.3 | | 0.5 | 000000000000016.83 | 0.85 | Iris-setosa |
| 4.8 | 3.4 | | 0.2 | 000000000000016.32 | 0.38 | Iris-setosa |