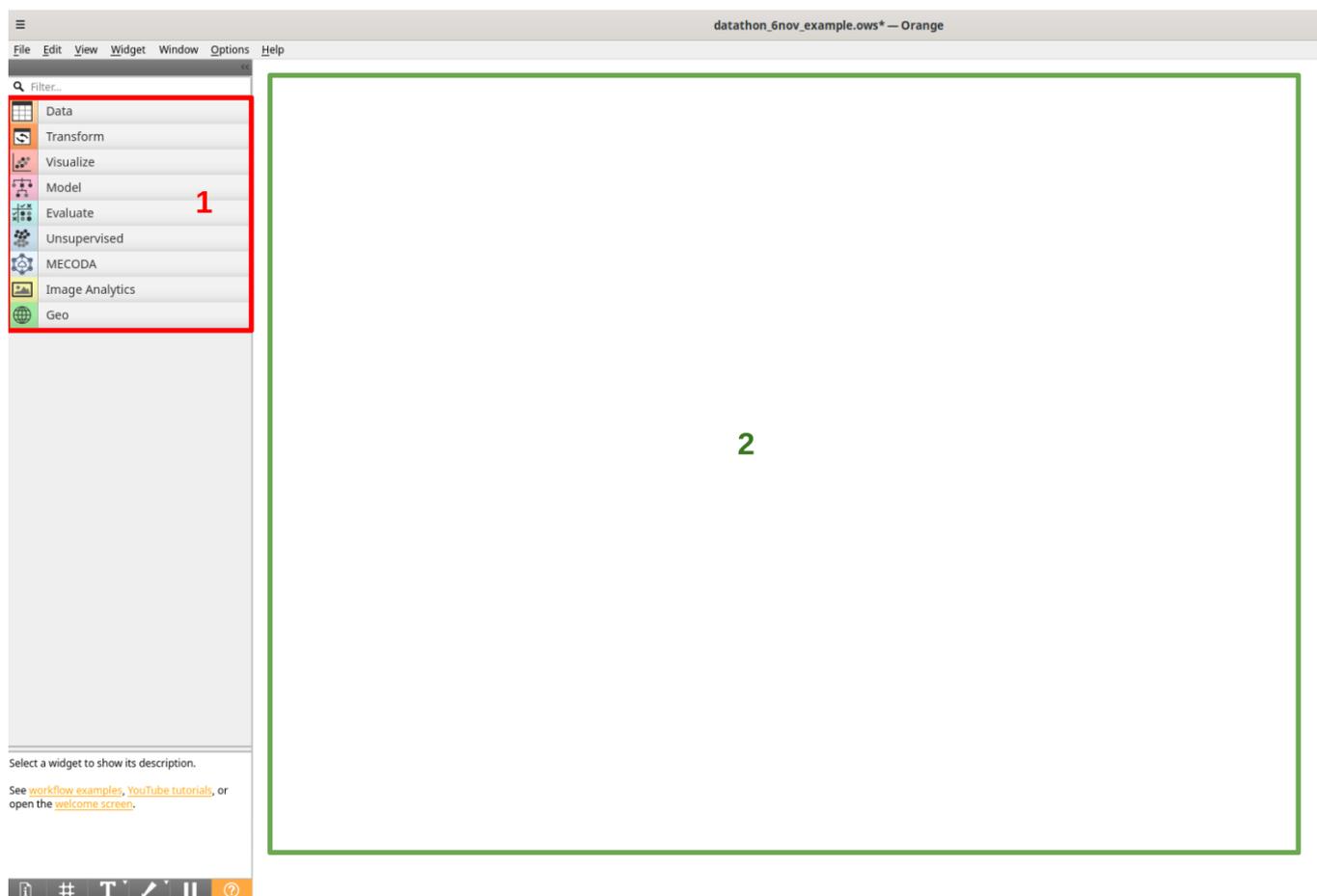


Inicio de la actividad

Las personas participantes en la Datathon deben seguir la [guía de instalación](#) para instalar tanto la aplicación Orange como el paquete mecod-a-orange dentro de esa plataforma y los add-ons "Geo" e "Image Analytics".

La Plataforma de Programación Visual Orange

El aspecto de la aplicación es este:



1. Columna de widgets.
2. Área de creación del flujo de datos.

En la columna de la izquierda, identificada con el número 1, están los widgets agrupados por tipo:

- **Data:** para importación desde diferentes formatos, ver los datos, ver las estadísticas generales, obtener información del dataset o guardar los datos.
- **Transform:** para seleccionar columnas, filas, transposición, combinación de tablas, agregación, valores únicos, randomización, tablas dinámicas, discretizar valores, etc.
- **Visualize:** incluye diferentes formas de representar gráficamente los datos, como árbol, gráfico de violín, de dispersión, de barras, histogramas, etc.
- **Model, Evaluate, Unsupervised:** herramientas para trabajar modelos de datos, propios de machine learning.

A estos añadiremos los add-ons que nos servirán para realizar nuestro análisis de datos de biodiversidad:

- **MECODA:** widgets que conectan con los datos de diferentes observatorios de Ciencia Ciudadana (Minka, OdourCollect), datos de sensores DIY como CanAIRio, Aire Ciudadano o Smart Citizen.
- **Image Analytics:** incluye widgets para el tratamiento de imágenes, que permite verlas dentro de Orange o utilizar modelos de imágenes para encontrar similitudes.
- **Geo:** para la realización de mapas (de puntos o sombreados) y para la geocodificación de valores de latitud y longitud a regiones administrativas.

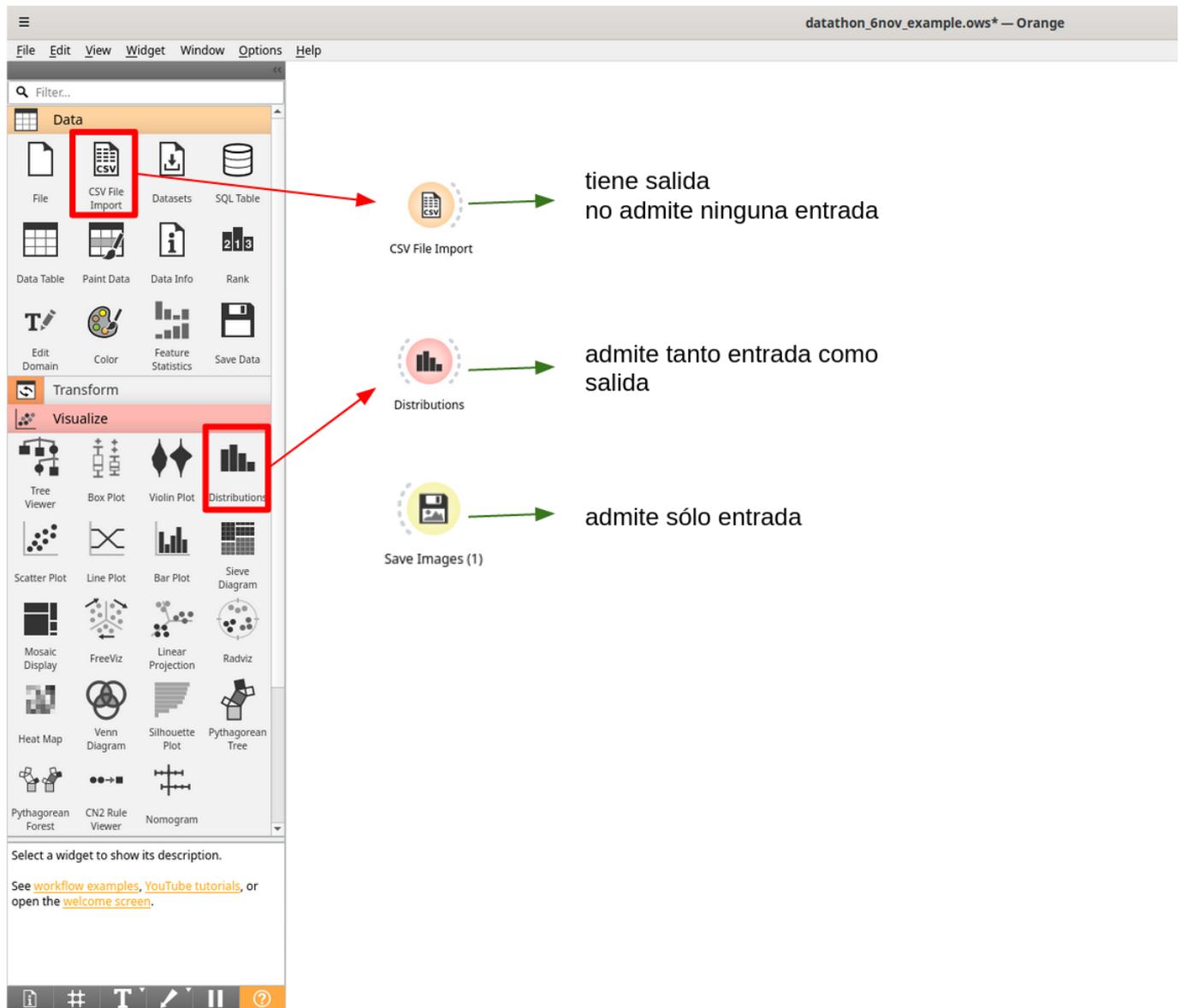
NOTA: Puedes consultar toda la información sobre cómo utilizar cada uno de estos widgets en el [catálogo](#).

Los widgets se seleccionan en esta columna izquierda, y aparecen en el área blanca de la derecha (número 2). Para seleccionar el widget basta con clicar una vez sobre él, en la columna de la izquierda.

Cada widget, cada una de estas herramientas disponibles, puede comunicarse con otros widgets de diferentes formas:

- Un widget puede admitir como entrada la salida de otro widget y dar como resultado una salida, que se pueda comunicar con un siguiente widget.
- Puede generar una salida y no admitir ninguna entrada.
- Puede admitir una entrada, pero no dar ninguna salida.

Aquí vemos un ejemplo:

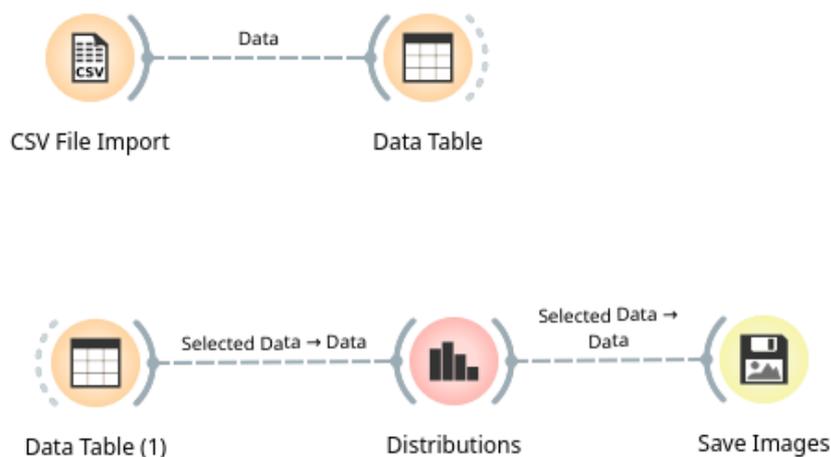


En este caso, vemos que hemos seleccionado del menú de la izquierda tres widgets:

- *CSV File Import*, un widget para importar un csv, tiene una línea curva punteada en el lado derecho. Eso quiere decir que este widget puede dar una salida, pero no admite ninguna entrada (el flujo de datos va de izquierda a derecha).
- *Distributions*, un widget para obtener el histograma o distribución de una columna de nuestros datos, tiene línea curva punteada en ambos lados. Es decir, puede recibir información (una tabla de datos) desde el lado izquierdo y puede conectarse con otro widget por el lado derecho.
- *Save Images*, un widget para guardar imágenes en nuestro ordenador. Este widget sólo tiene línea curva en el lado izquierdo. Eso quiere decir que admite entradas, pero realizará el proceso para el que está diseñado (descargar las fotos) y no tendrá ninguna salida en Orange.

Este ejemplo es para explicar solamente cómo se conectan los diferentes widgets, necesario para construir los flujos de datos o data workflows.

Estos widgets se conectarán de esta forma:



Aquí vemos que se crea una línea entre los widgets, que comunica la información de unos a otros. Si esta línea no es continua, nos indica que no se está transmitiendo ninguna información de uno a otro, como en este caso. Sobre la línea se indica qué se está transmitiendo: Data (todos los datos) o Selected Data (sólo los datos que hayamos seleccionado en el widget inicial).

Análisis de un conjunto de datos mayor de 10.000 observaciones de biodiversidad en Minka

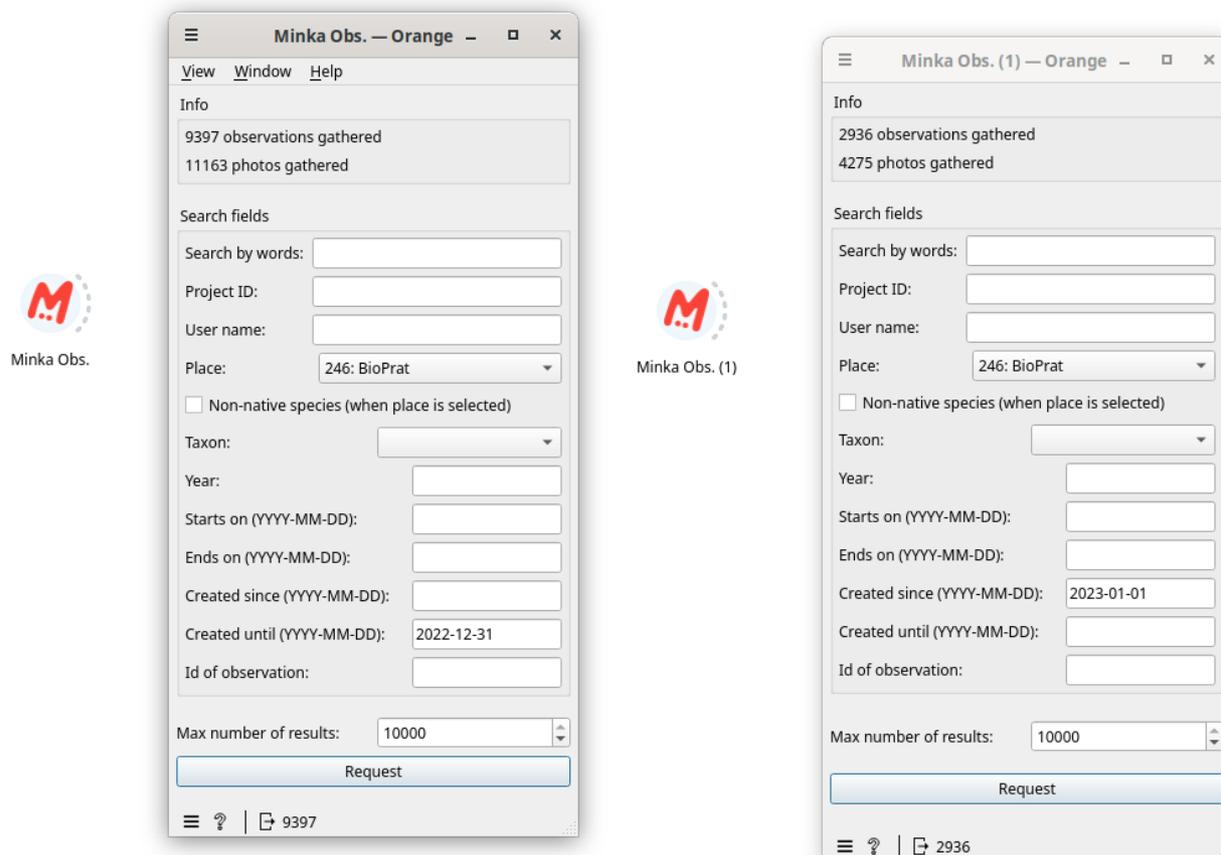
Ejemplo con el área del Ayuntamiento del Prat

Vamos a trabajar con el widget de Minka para obtener las observaciones de biodiversidad registradas en el área de El Prat.

Seleccionamos los datos a analizar

Seleccionamos el widget de Minka y hacemos doble click sobre el widget para desplegar el cuadro de diálogo. Seleccionamos el área geográfica que comprende la localización del ayuntamiento del Prat, que podemos ver aquí en Minka: https://minka-sdg.org/observations?place_id=246

Dado que ese lugar tiene más de 10.000 observaciones y nuestro límite por descarga es de 10.000 podremos descargar todas las observaciones haciendo dos peticiones, una para las observaciones creadas en 2022 y otra para las de 2023. Cada una de las descargas tardará unos minutos en procesarse, porque se trata de peticiones de un volumen grande de observaciones. Ten paciencia. Si te aparece un mensaje de que el proceso está tardando demasiado, simplemente pincha en "Esperar" y deja que termine. Finalmente obtendrás algo así:



Para descargar los datos haz doble clic el widget de Minka, en el primero elige el place BioPrat y el filtro "Created until:" con la fecha "2022-12-31". Con esto obtendremos las observaciones subidas a la plataforma en 2022. Cuando esta petición termine, abre el segundo widget de Minka y selecciona el place BioPrat y ahora el filtro "Created since:" con la fecha "2023-01-01".

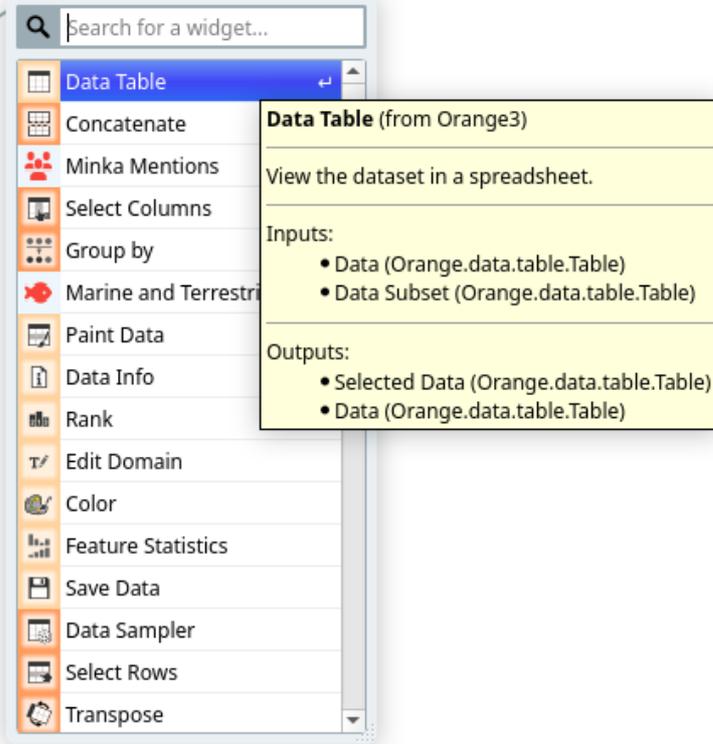
Si la descarga que queremos hacer tiene menos de 10.000 observaciones bastaría con hacer una única descarga con el conjunto del *place* especificado.

Ahora vamos a ver las tablas de datos que hemos obtenido con cada uno. Para ello arrastramos una línea desde el widget de Minka, en la zona donde tiene la línea discontinua. Así podemos "engancharlo" con el widget que nos permitirá ver los datos, llamado "Data Table":

Observaciones 2022

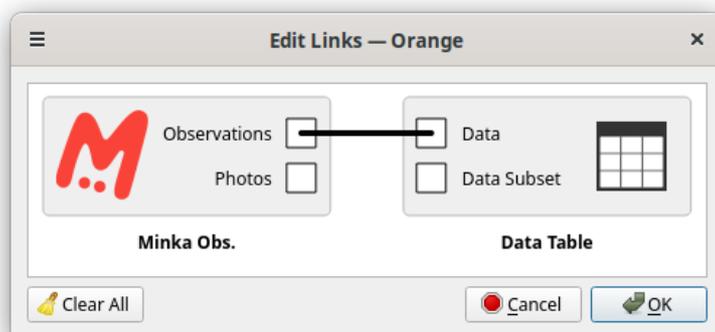


Observaciones 2023



Una vez hayamos pinchado en "Data Table" nos dará la opción de cargar en ese widget dos tipo de datos: "Observations" y "Photos". Esto es porque nuestro widget de Minka tiene dos salidas, una tabla con todas las observaciones y una tabla con todas las fotos. Cada observación puede tener más de una foto. La primera tabla nos permitirá hacer análisis estadístico, la segunda análisis de imágenes:

Observaciones 2022

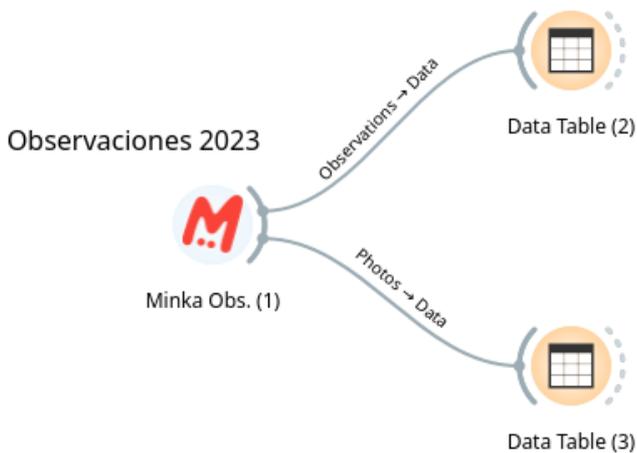


Seleccionamos coger las observaciones (Observations -> Data) y eso cargará las observaciones en la Tabla. Le damos a OK y veremos esto al hacer doble click en "Data Table":

The screenshot shows the Orange Data Mining interface. On the left, a workflow consists of a 'Minka Obs.' widget (with a red 'M' logo) connected to an 'Observations → Data' widget, which is then connected to a 'Data Table' widget. Below this, another 'Minka Obs. (1)' widget is shown. The main window displays the 'Data Table' widget with the following data:

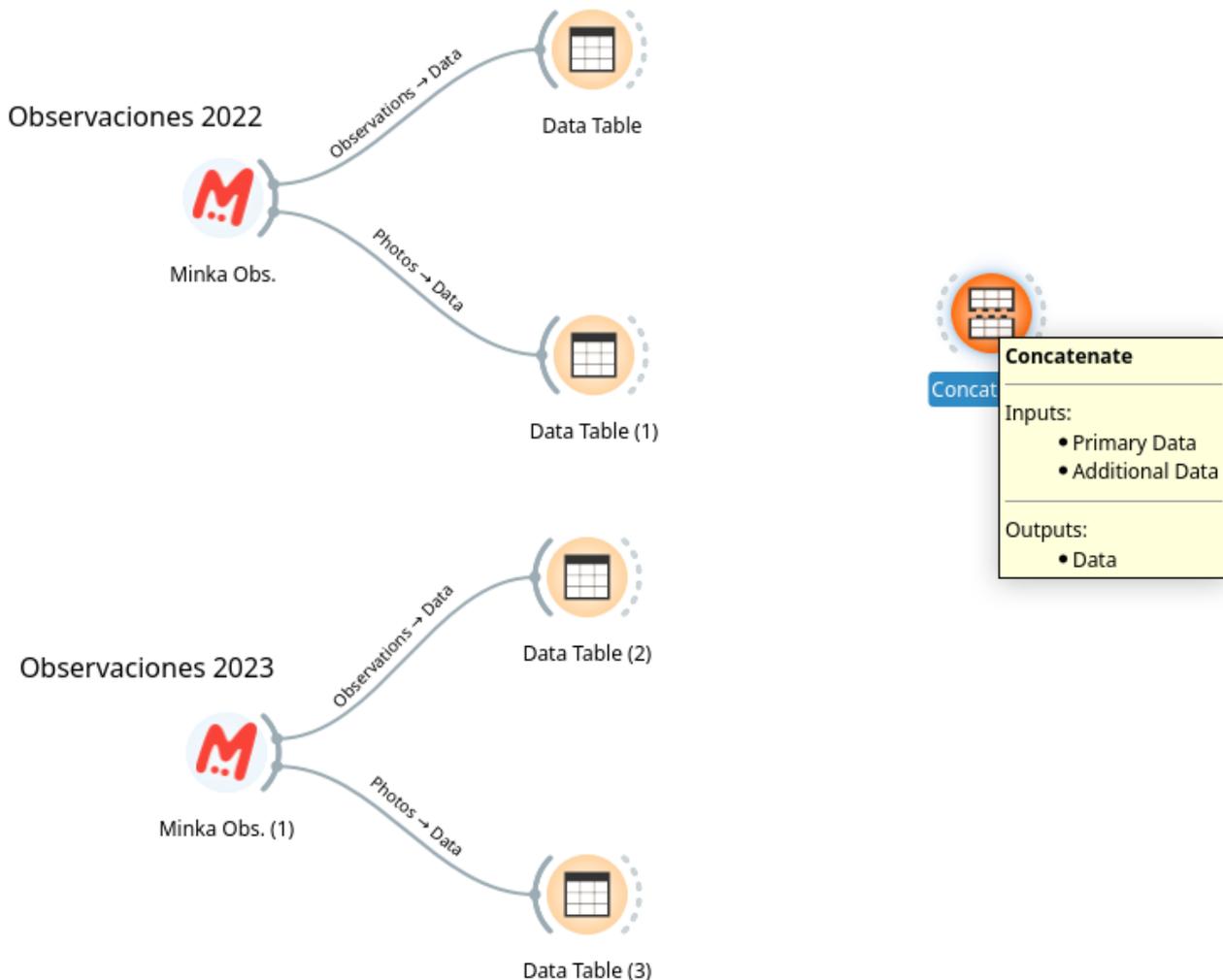
	place_name	id	captive	created_at	updated_at	observed_on	description	iconic_taxon
1	Finca de la Ma...	103909	0	2022-12-23	2022-12-24	2022-12-22	?	insecta
2	Carretera de la...	103676	0	2022-12-21	2022-12-21	2022-12-21	?	aves
3	Camí de cal Tet	103675	0	2022-12-21	2022-12-21	2022-12-21	?	aves
4	Camí alternati...	103674	0	2022-12-21	2023-01-27	2022-12-21	?	fungi
5	Camí alternati...	103673	0	2022-12-21	2022-12-21	2022-12-21	?	?
6	Camí alternati...	103672	0	2022-12-21	2023-01-04	2022-12-21	?	fungi
7	Camí alternati...	103671	0	2022-12-21	2023-01-04	2022-12-21	?	fungi
8	Camí alternati...	103670	0	2022-12-21	2022-12-21	2022-12-21	?	?
9	Camí alternati...	103669	0	2022-12-21	2022-12-21	2022-12-21	?	?
10	Camí alternati...	103668	0	2022-12-21	2022-12-21	2022-12-21	?	?
11	Camí alternati...	103667	0	2022-12-21	2022-12-21	2022-12-21	?	?
12	Camí alternati...	103666	0	2022-12-21	2022-12-21	2022-12-21	?	?
13	Camí alternati...	103665	0	2022-12-21	2022-12-21	2022-12-21	?	?
14	Camí alternati...	103664	0	2022-12-21	2022-12-21	2022-12-21	?	?
15	Camí alternati...	103663	0	2022-12-21	2022-12-21	2022-12-21	?	?
16	Camí alternati...	103662	0	2022-12-21	2023-02-23	2022-12-21	?	insecta
17	Camí alternati...	103661	0	2022-12-21	2022-12-21	2022-12-21	?	?
18	Camí alternati...	103660	0	2022-12-21	2022-12-21	2022-12-21	?	?
19	Camí alternati...	103659	0	2022-12-21	2022-12-21	2022-12-21	?	?
20	Parc Fluvial	103658	0	2022-12-21	2022-07-13	2022-12-21	?	aves
21	Camí alternati...	103657	0	2022-12-21	2022-12-21	2022-12-21	?	?

Son nuestros datos de observaciones en una tabla, con todos nuestros campos. Ahora hagamos lo mismo con las fotos, saquémoslas a otra tabla. Y repitamos el proceso con el segundo widget de MINKA donde tenemos las observaciones de 2023:

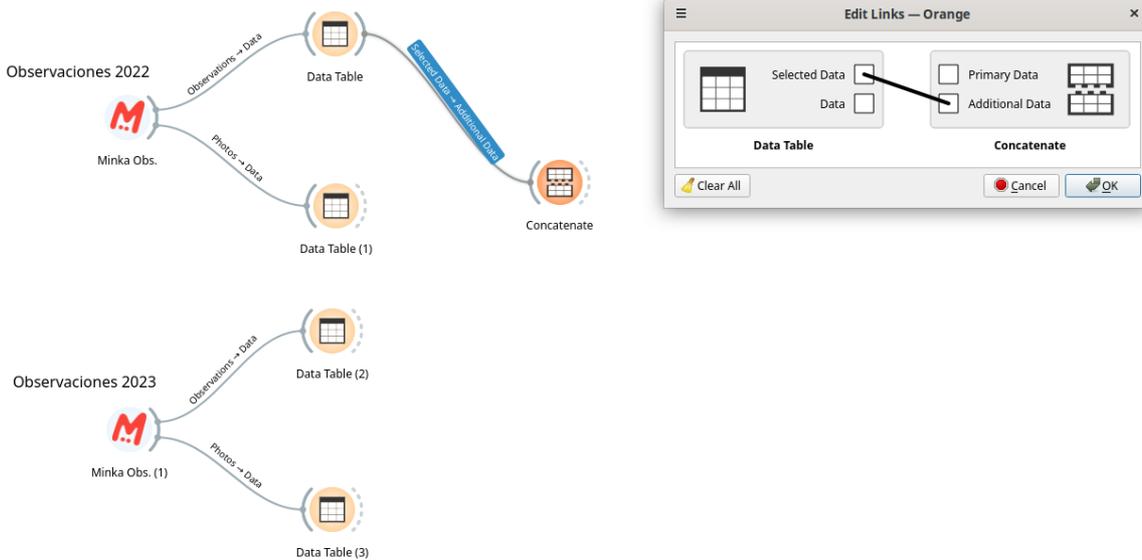


Ya tenemos todos los datos extraídos. Vemos que la conexión de cada una de las dos líneas de Minka a Data Table es distinta y nos indica que en una está llevando "Observations" hacia los datos de la tabla y en el otro caso está llevando "Photos" hacia los datos de la segunda tabla.

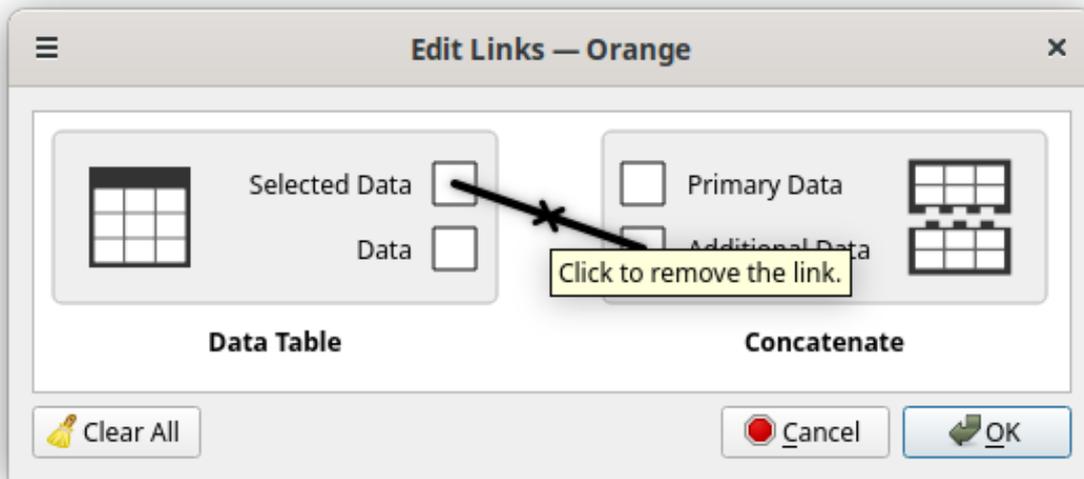
Ahora vamos a unir las tablas de observaciones en una sola tabla para analizarla. Para eso seleccionaremos el widget "Concatenate" que está en la sección "Transform" de la columna de la izquierda. Hacemos click sobre él y se colocará en nuestra área de trabajo.



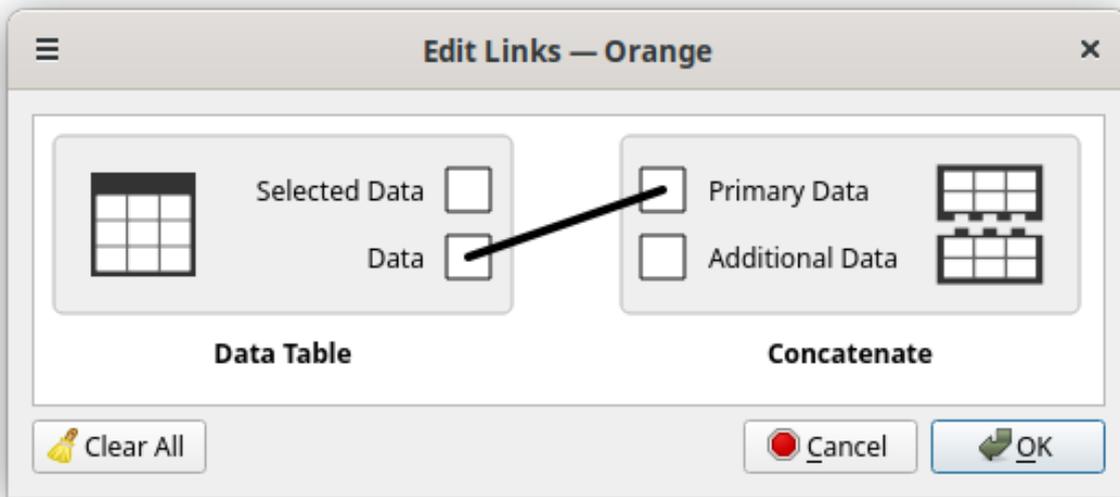
Vamos a pasar por ese widget nuestra dos tablas de observaciones. Sacamos una línea desde la primera hacia "Concatenate". Nos creará una conexión por defecto, pero la vamos a editar haciendo click sobre la línea que comunica los dos widgets:



Quitaremos la conexión de "Selected Data" a "Additional Data" pinchando sobre la línea:



Comunicaremos "Data" con "Primary Data", porque queremos que considere todo el conjunto de nuestros datos como los datos primarios a la hora de concatenarlos. Nos quedará algo así:



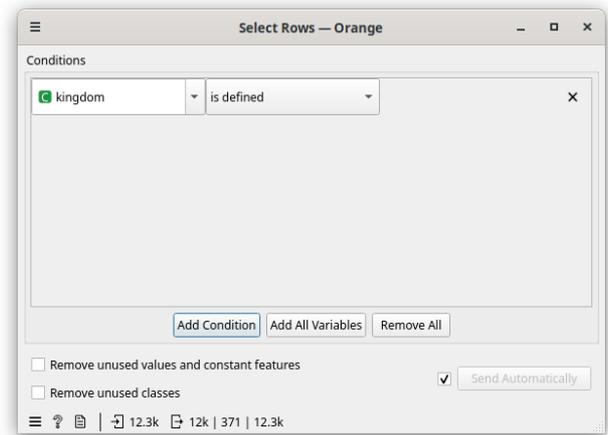
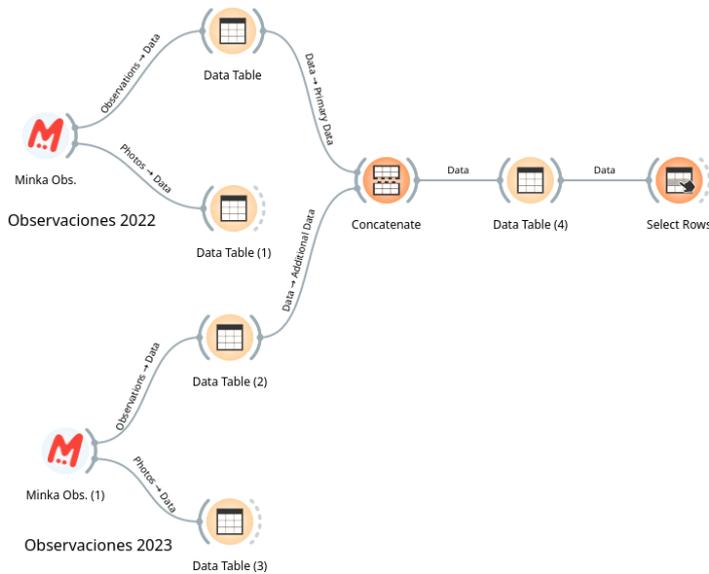
Pinchamos en OK.

Ahora hacemos lo mismo con la segunda tabla de observaciones, la conectaremos con Concatenate. En este caso no editaremos la conexión, porque estos sí serán el "Additional Data" que queremos unir al primero. Sacamos una tabla ahora con el conjunto de los datos y vemos que el resultado ya tiene más de 12.300 observaciones (es la unión de ambas tablas). Ya podemos trabajar con él:

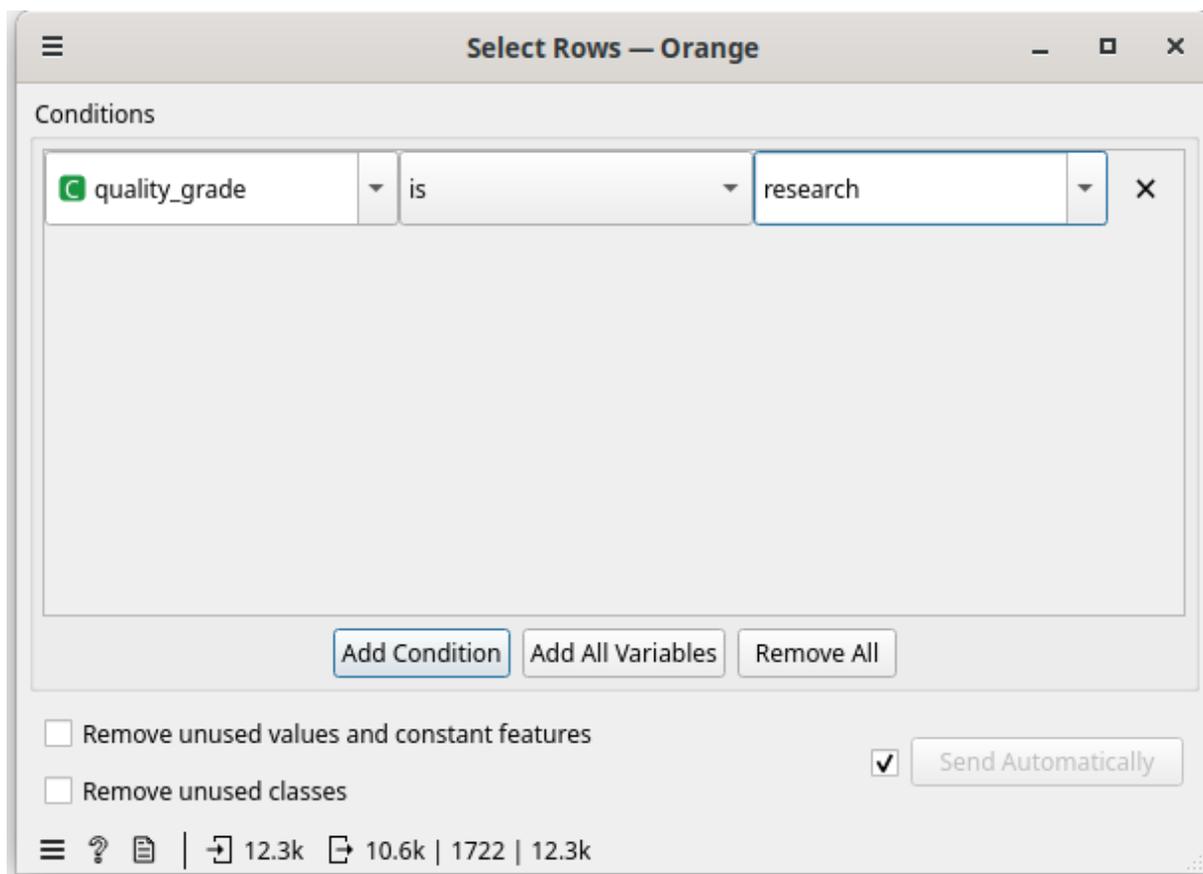
id	captive	created_at	updated_at
1	0	2022-12-23	2022-12-24
2	0	2022-12-21	2022-12-21
3	0	2022-12-21	2022-12-21
4	0	2022-12-21	2023-01-27
5	0	2022-12-21	2022-12-21
6	0	2022-12-21	2023-01-04
7	0	2022-12-21	2023-01-04
8	0	2022-12-21	2022-12-21
9	0	2022-12-21	2022-12-21
10	0	2022-12-21	2022-12-21
11	0	2022-12-21	2022-12-21
12	0	2022-12-21	2022-12-21
13	0	2022-12-21	2022-12-21
14	0	2022-12-21	2022-12-21
15	0	2022-12-21	2022-12-21
16	0	2022-12-21	2023-02-23
17	0	2022-12-21	2022-12-21
18	0	2022-12-21	2022-12-21
19	0	2022-12-21	2022-12-21
20	0	2022-12-21	2023-07-13
21	0	2022-12-21	2022-12-21
22	0	2022-12-21	2022-12-21
23	0	2022-12-21	2022-12-21
24	0	2022-12-21	2022-12-21

Selección de las observaciones con grado de investigación

Lo primero que haremos será quedarnos con aquellas observaciones que hayan alcanzado el grado de investigación (tengan al menos 2 identificaciones y estén de acuerdo en la especie y no le falta ningún dato de los considerados obligatorios). Para eso vamos a conectar nuestra tabla con todas las observaciones con el widget "Select Rows", del bloque "Transform":



Con este widget podremos seleccionar las líneas que queremos simplemente indicando cuál es la condición o condiciones que deben cumplir. En este caso queremos las observaciones que tengan, en la columna "quality_grade" el valor "research":

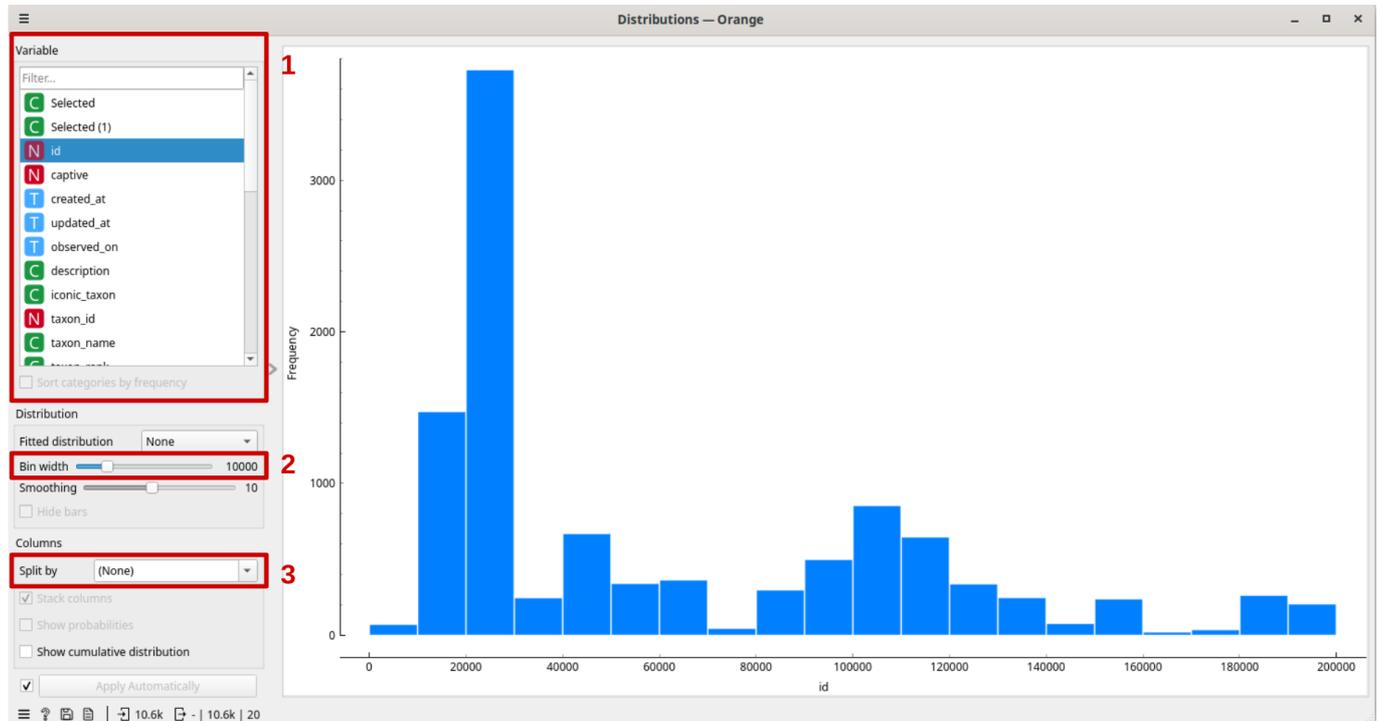


Podemos añadir más de una condición simplemente pinchando en "Add Condition". Una vez establecida nuestra condición podemos ver en la parte de abajo del widget cuál es el resultado de aplicarlo. De las 12.3k observaciones totales está seleccionando 10.6k y descartando 1722. Es decir, en nuestro conjunto de datos teníamos 1722 observaciones que, o bien necesitaban identificación, o bien carecían de algún dato de los considerados obligatorios para alcanzar el nivel de investigación (localización, foto, etc).

Si sacamos una tabla con las observaciones resultantes de "Select Row" veremos las que sí cumplen nuestra condición.

Distribución de nuestros datos

Digamos que queremos saber es cuántas observaciones se han registrado cada mes. Para eso conectaremos nuestra tabla resultante con el widget "Distributions", de la sección "Visualize". Haciendo doble click sobre ese widget, después de haberlo conectado, veremos algo así:



Aquí podremos jugar con varios parámetros. El primero es la columna que queremos analizar. En este caso seleccionaremos "created_at" (para analizar por fecha de subida de la observación) o "observed_on" (para analizar por fecha en la que se ha tomado la observación).

Lo segundo que haremos será seleccionar el periodo temporal que queremos que se use para agrupar los datos (1 mes, 1 año,...). Ahí nos dará más o menos columnas. Si el rango de los datos es muy amplio, nos permitirá crear agrupaciones amplias (año). Si el rango es pequeño nos permitirá bajar a nivel de día o de horas.

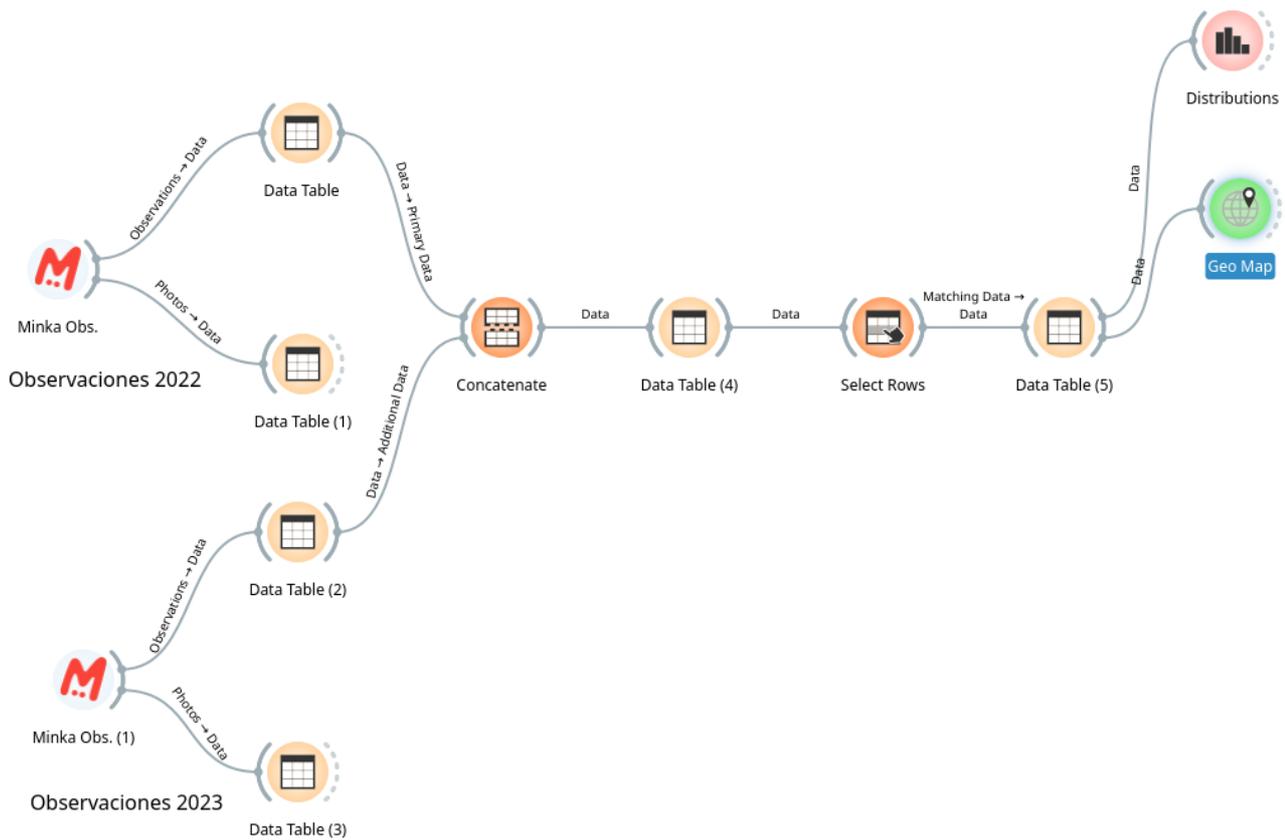
Lo tercero que podemos hacer es dividir nuestras columnas según algún criterio, por ejemplo, el reino al que pertenecen, el usuario que las ha subido, etc. Si elegimos un campo con muchas opciones la división por colores de las columnas las hará difícil de leer.

Podemos guardar el gráfico que nos interese simplemente pinchando en el icono de "disquete" de la parte inferior izquierda del widget "Distribution".

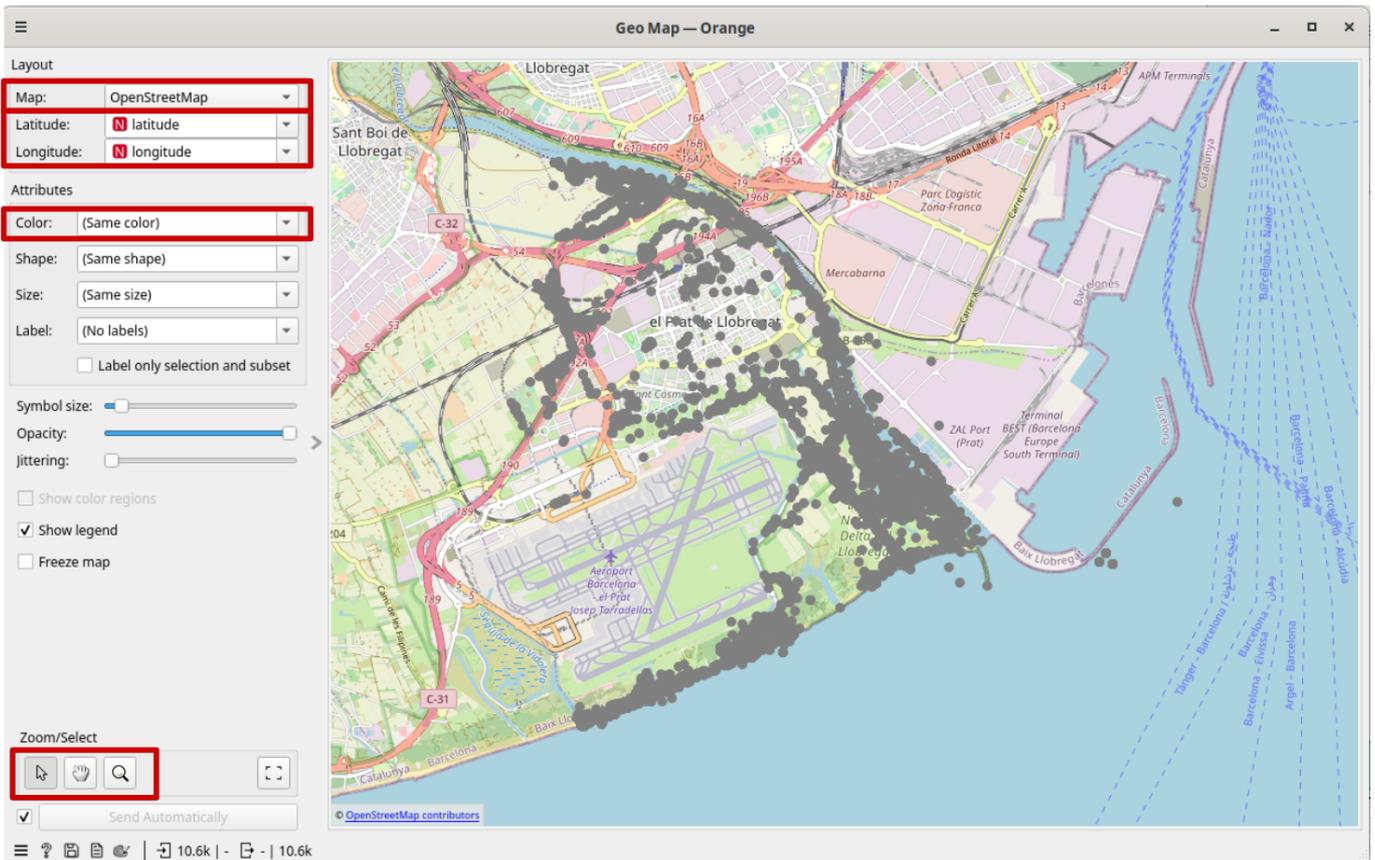
Podemos usar este widget también para ver la distribución en otras columnas, como la de "user login", que nos agruparía las observaciones por nombre de usuario que las ha reportado.

Mapeo de las observaciones

Ahora conectaremos nuestra tabla de datos con el widget "Geo Map", de la sección "Geo" de la columna izquierda:



Haciendo doble click sobre él veremos qué opciones tenemos para trabajar los datos a un nivel geográfico:



Lo primero en lo que debemos fijarnos es que la latitud y longitud han sido leídas correctamente en nuestra tabla de datos. En este caso, como le hemos dado a esas columnas el nombre exacto, el widget ha sabido

dónde encontrar los datos. Pero si usamos otro conjunto de datos podríamos tener que indicarle manualmente cuáles son esas columnas.

Encima de los datos de lat-lon tenemos el tipo de mapa en el que visualizar nuestras observaciones. Por defecto, se abrirá en Open Street Map, pero podemos cambiar la visualización a Satellite.

Vemos que podemos usar atributos como el color de los puntos para ver los datos de nuestra tabla. Podríamos elegir que el color de los puntos estuviera en función del usuario (columna `user_login`), de la clase taxonómica (columna `class`), etc.

Si queremos interactuar con el mapa tenemos tres botones en la sección Zoom/Select:

- Flecha: nos permite seleccionar áreas del mapa
- Mano: nos permite desplazar el mapa
- Lupa: nos permite hacer zoom

Podemos extraer el mapa resultante como una imagen usando el botón de "disquete" de la parte inferior.

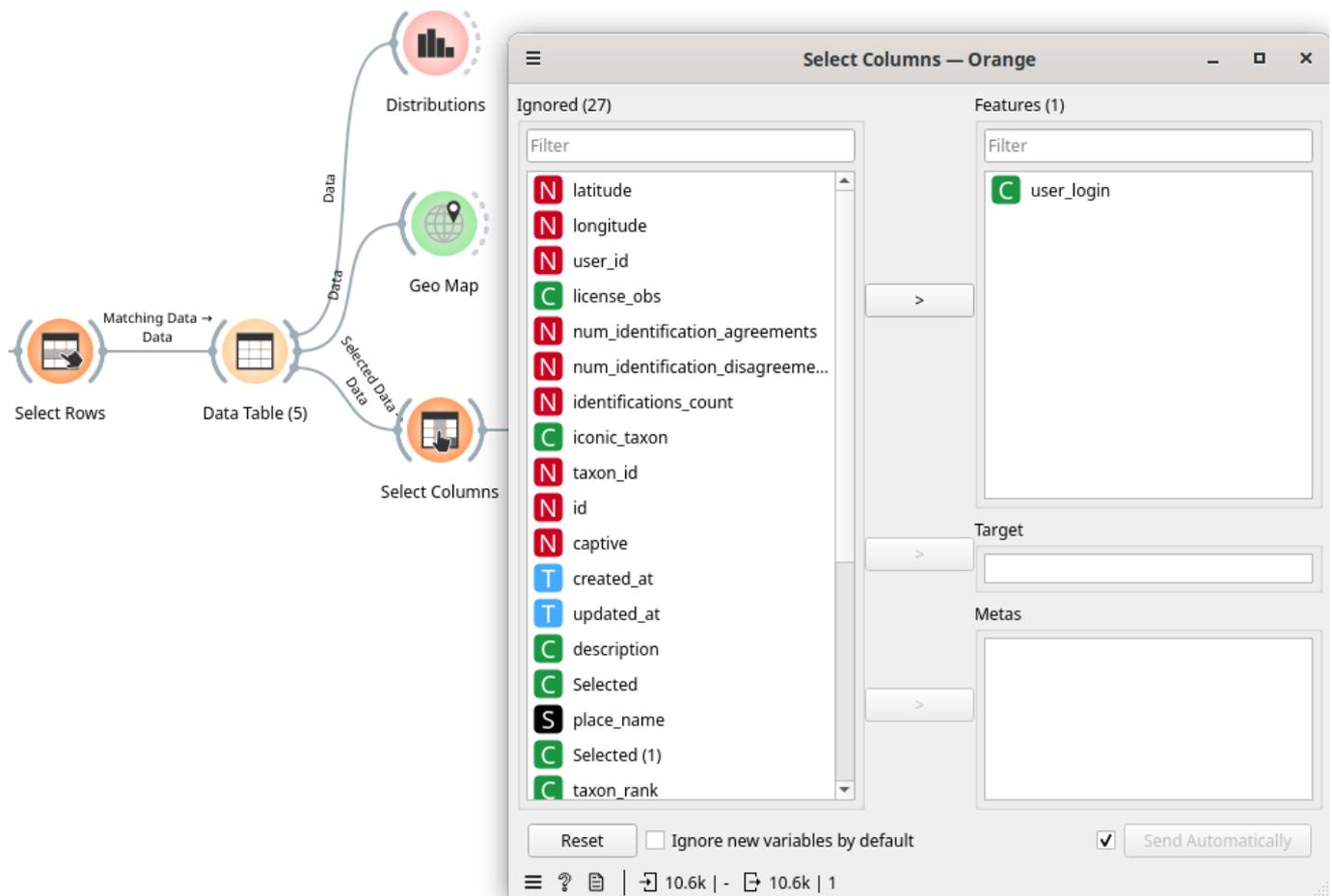
Y podemos usar el mapa como filtro para seleccionar observaciones y extraer una nueva tabla con las seleccionadas.

Estadísticas de nuestros datos

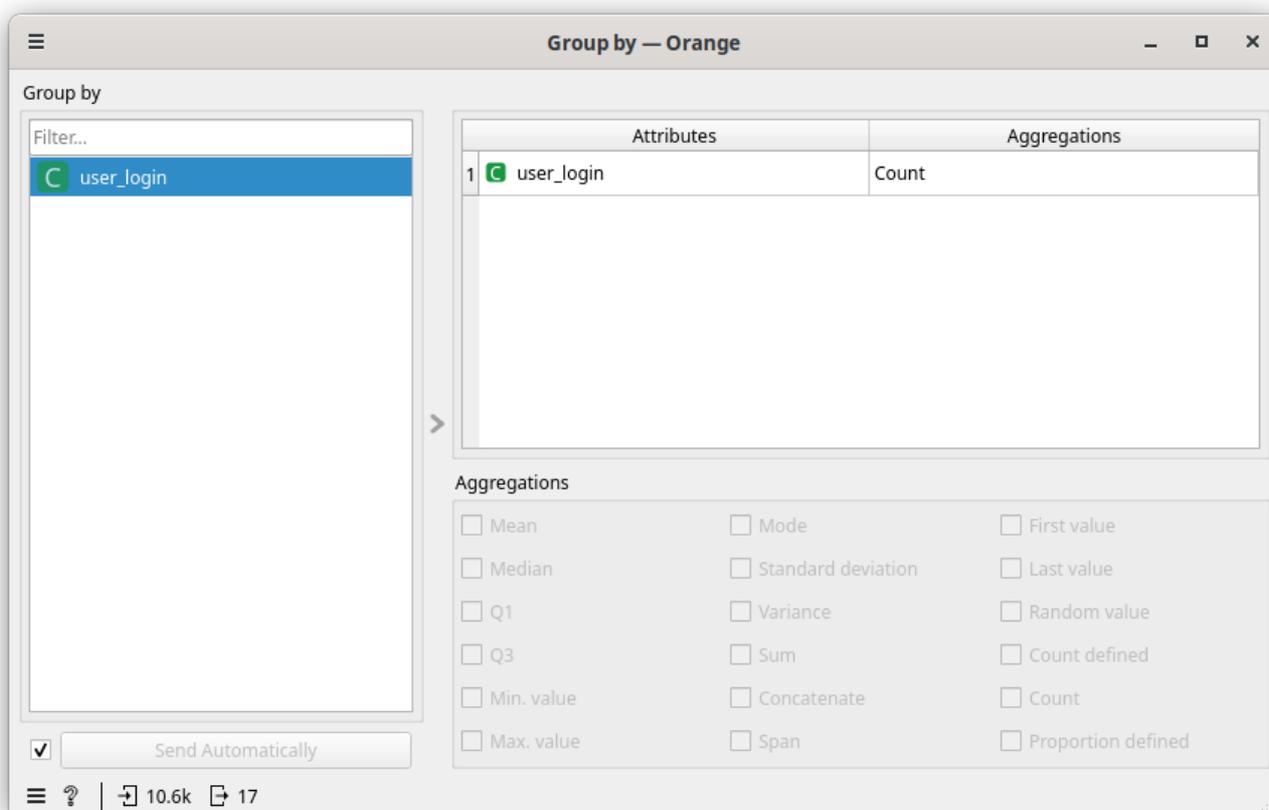
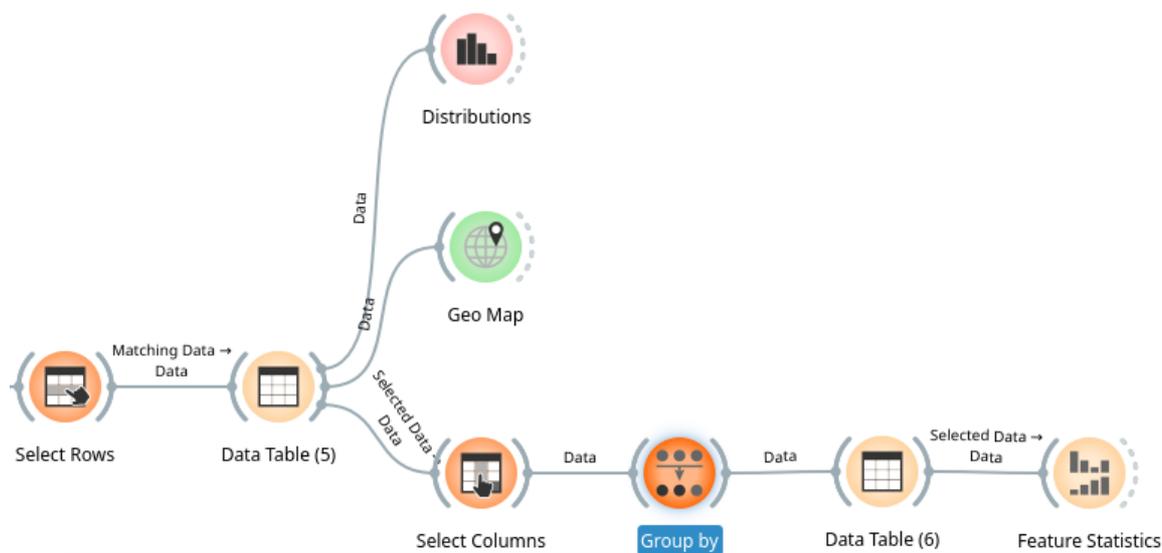
Digamos que ahora queremos analizar cuántos usuarios han aportado datos en el área que estamos analizando, la media de observaciones por usuario, la mediana, etc. En nuestro conjunto de datos tenemos la columna "user_login", que nos da el nombre, pero no tenemos el número de observaciones de cada usuario en ninguna tabla. Vamos a crear esta tabla numérica.

Los pasos a seguir son:

- Conectar nuestra tabla de datos con "Select column" para quedarnos únicamente con la columna que nos interesa:



- Conectar la columna seleccionada con "Group by", que nos permitirá agrupar los usuarios y sacar el número de cada uno:



- Conectar el resultado de Group by con una tabla para ver qué tenemos:

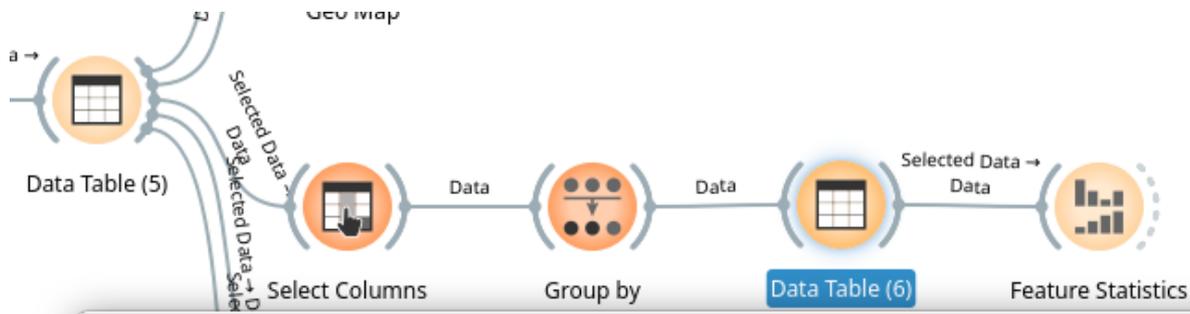
The screenshot shows a workflow in Orange data mining software. It starts with a 'Data Table (5)' widget, followed by 'Select Columns', 'Group by', 'Data Table (6)', and 'Feature Statistics'. The 'Data Table (6)' window is open, displaying the following data:

	user_login	user_login - Count
8	mediambient_ajelprat	8849
13	pau_esteban	553
15	seaprat	522
9	mestralet	153
14	piripip	151
7	jordiruestesalba	100
2	andres_valverde_valera	83
17	xasalva	64
5	helenamarrugat	23
16	serapio	16
4	educacioambientalfundesplai	9
3	ederoa	5
10	nozcc	4
11	oriol_d	3
1	andrea	1
6	jaume-piera	1
12	patopi	1

The 'Data Table (6)' window also shows the following information on the left side:

- Info:** 17 instances (no missing data), 1 feature, No target variable, 1 meta attribute.
- Variables:**
 - Show variable labels (if present)
 - Visualize numeric values
 - Color by instance classes
- Selection:**
 - Select full rows
- Buttons:** Restore Original Order, Send Automatically (checked).
- Status:** 17 | 17 | 17

- Conectar la tabla con Feature Statistics para obtener las estadísticas (media, mediana, moda,...) de nuestro resultado:



Data Table (6) — Orange

Info
 17 instances (no missing data)
 1 feature
 No target variable.
 1 meta attribute

Variables
 Show variable labels (if present)
 Visualize numeric values
 Color by instance classes

Selection
 Select full rows

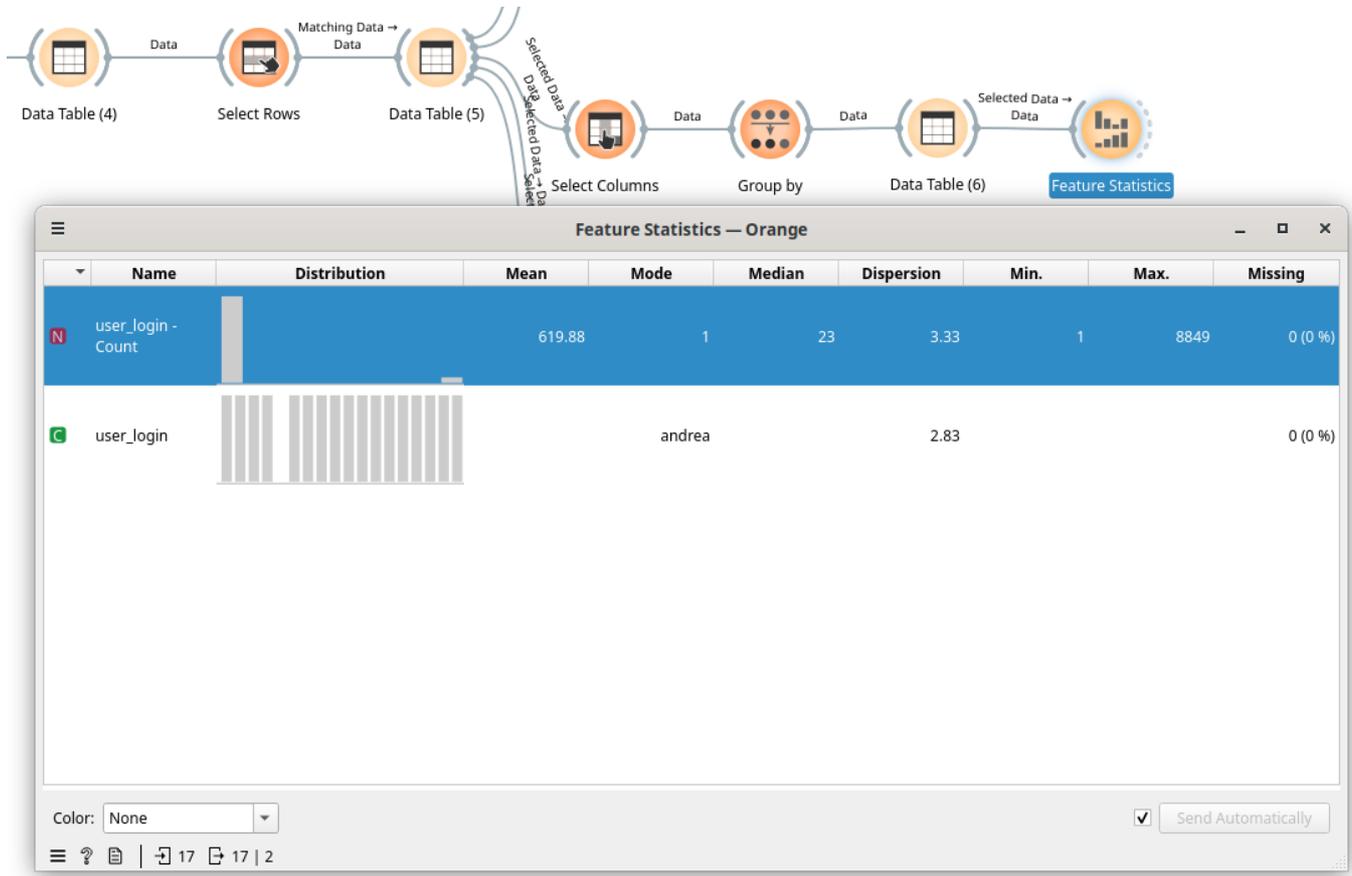
Restore Original Order

Send Automatically

	user_login	user_login - Count
8	mediambient_ajelprat	8849
13	pau_esteban	553
15	seaprat	522
9	mestralet	153
14	piripip	151
7	jordiruestesalba	100
2	andres_valverde_valera	83
17	xasalva	64
5	helenamarrugat	23
16	serapio	16
4	educacioambientalfundesplai	9
3	ederoa	5
10	nozcc	4
11	oriol_d	3
1	andrea	1
6	jaume-piera	1
12	patopi	1

17 | 17 | 17

Podemos hacer el mismo proceso usando la columna de "taxon name" y sacar el número de observaciones de cada especie:



Incluso saber cuándo ha sido la primera vez que se ha reportado una especie, indicando que incluya el dato del valor mínimo de la columna "observed_on" que tiene las fechas:

The screenshot shows the Orange data mining interface. The top window, 'Group by (2) — Orange', is configured with the following settings:

- Group by:** A list of variables including `taxon_name`, `taxon_rank`, `kingdom`, `phylum`, `class`, `order`, `family`, and `observed_on`.
- Attributes:** A list of variables from 2 to 8, including `taxon_rank`, `kingdom`, `phylum`, `class`, `order`, `family`, and `observed_on`.
- Aggregations:** A grid of aggregation options where `Min. value` is selected for the `observed_on` variable.

The bottom window, 'Data Table (8) — Orange', displays the resulting data table:

	taxon_name	taxon_rank	taxon_name - Count	observed_on - Min. value
1	acanthocardia tuberculata	species	7	2022-11-23
2	accipiter gentilis	species	1	2021-12-27
3	acer negundo	species	1	2019-03-19
4	acherontia atropos	species	4	2018-11-14
5	achyra nudalis	species	1	2020-08-03
6	acontia lucida	species	7	2018-06-13
7	acrida ungarica	species	16	2018-07-17
8	acrida ungarica ungarica	subspecies	5	2022-06-28
9	acrocephalus arundinaceus	species	11	2017-05-18
10	acrocephalus melanopogon	species	1	2022-03-09
11	acrocephalus scirpaceus	species	3	2022-09-06
12	acronicta rumicis	species	4	2019-10-08
13	acrotylus insubricus	species	1	2022-06-03
14	actitis hypoleucos	species	30	2016-04-21
15	adalia bipunctata	species	2	2020-05-16
16	aedes albopictus	species	2	2020-06-23
17	aegithalos caudatus	species	15	2017-02-14
18	aequorea forskalea	species	2	2022-06-03
19	aeshna cyanea	species	1	2018-10-25
20	aeshna mixta	species	14	2017-11-28

Relación entre variables

Si queremos ver la relación entre dos columnas, podemos utilizar Pivot Table, que es similar a la tabla dinámica en Excel. Nos permite crear una tabla con el número de observaciones agrupadas por dos de las columnas, por ejemplo, entre los grupos taxonómicos y los usuarios:

Pivot Table — Orange

Rows: **iconic_taxon**

Columns: **user_login**

Values: (None)

Aggregations: Count

Count	andrea	andres_valverde_valera	ederoa	educacioambientalfundesplai	helenamarrugat	jaume-piera	jordiruestesalba	mediambient_ajelprat	mestralet	nozc	oriol_d	patopi	pau_esteban	piripip	seaprat	serapio	xasalva	Total
actinopterygii	0.0	0.0	0.0	0.0	0.0	0.0	1.0	75.0	0.0	0.0	1.0	0.0	3.0	0.0	2.0	0.0	4.0	86.0
amphibia	0.0	0.0	0.0	0.0	0.0	0.0	0.0	54.0	1.0	0.0	0.0	0.0	3.0	0.0	0.0	0.0	0.0	58.0
animalia	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	2.0	0.0	0.0	0.0	12.0	14.0
annelida	0.0	1.0	0.0	0.0	0.0	0.0	0.0	1.0	0.0	0.0	1.0	0.0	1.0	0.0	0.0	0.0	2.0	6.0
arachnida	0.0	2.0	2.0	0.0	1.0	0.0	2.0	76.0	0.0	0.0	0.0	0.0	23.0	2.0	0.0	0.0	0.0	108.0
aves	0.0	9.0	0.0	1.0	0.0	0.0	16.0	4444.0	139.0	1.0	0.0	1.0	171.0	115.0	71.0	14.0	0.0	4982.0
bryozoa	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	6.0	6.0
chromista	0.0	0.0	0.0	0.0	0.0	0.0	0.0	1.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	1.0	2.0
cnidaria	0.0	0.0	0.0	0.0	0.0	0.0	0.0	47.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	19.0	66.0
crustacea	0.0	0.0	0.0	0.0	0.0	0.0	0.0	68.0	0.0	0.0	1.0	0.0	0.0	0.0	0.0	1.0	0.0	70.0
echinodermata	0.0	0.0	0.0	0.0	0.0	0.0	0.0	12.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	10.0	22.0
elasmobranchii	0.0	0.0	0.0	0.0	0.0	0.0	0.0	5.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	5.0
fungi	0.0	13.0	0.0	0.0	1.0	0.0	0.0	121.0	0.0	0.0	0.0	0.0	6.0	0.0	7.0	0.0	0.0	148.0
insecta	0.0	21.0	3.0	0.0	7.0	0.0	67.0	2891.0	7.0	2.0	0.0	0.0	280.0	29.0	56.0	0.0	0.0	3363.0
mammalia	0.0	0.0	0.0	2.0	0.0	0.0	0.0	66.0	0.0	1.0	0.0	0.0	3.0	0.0	34.0	0.0	0.0	106.0
mollusca	0.0	3.0	0.0	0.0	1.0	0.0	0.0	220.0	0.0	0.0	0.0	0.0	8.0	0.0	5.0	0.0	6.0	244.0
plantae	1.0	33.0	0.0	6.0	11.0	0.0	8.0	349.0	3.0	0.0	0.0	0.0	46.0	3.0	328.0	0.0	1.0	789.0
platyhelminthes	0.0	0.0	0.0	0.0	0.0	0.0	0.0	1.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	1.0
porifera	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	3.0	3.0
reptilia	0.0	1.0	0.0	0.0	2.0	1.0	5.0	380.0	3.0	0.0	0.0	0.0	4.0	2.0	15.0	1.0	0.0	414.0
Total	1.0	83.0	5.0	9.0	23.0	1.0	100.0	8811.0	153.0	4.0	3.0	1.0	550.0	151.0	518.0	16.0	64.0	10493.0

Aquí podemos ver si los distintos usuarios se especializan en un grupo taxonómico o sus observaciones son diversas.

Reconocimiento en el uso de los datos

Si vais a utilizar este conjunto de datos para alguna publicación o investigación, deberéis reconocer la contribución de aquellos usuarios y usuarias que hayas hecho aportaciones a él. Para facilitar el trabajo a la hora de citarlos podéis usar el widget "Minka Mentions", simplemente conectándolo con nuestra tabla con el conjunto de los datos y mostrando su salida en un Data Table:

Data Table (9) — Orange

Info: 18 instances, 1 feature, No target variable, 2 meta attributes (2.8 % missing data)

Variables: Show variable labels (if present), Visualize numeric values, Color by instance classes

Selection: Select full rows

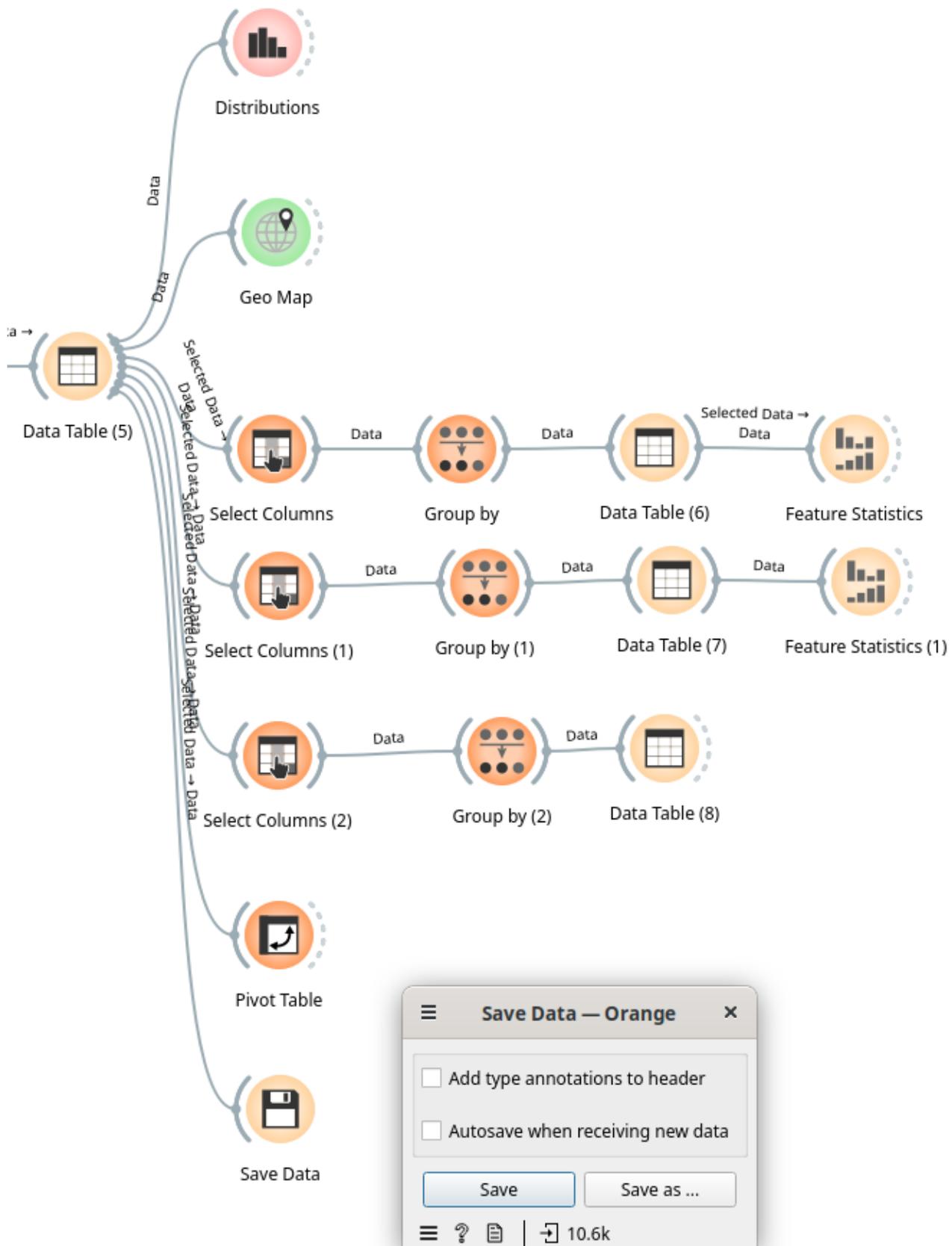
Buttons: Restore Original Order, Send Automatically

	user	user_url	observations
1	mediambient_ajelprat	https://minka-sdg.org/users/mediambient_ajelprat	8849
2	pau_esteban	https://minka-sdg.org/users/pau_esteban	553
3	seaprat	https://minka-sdg.org/users/seaprat	522
4	mestralet	https://minka-sdg.org/users/mestralet	153
5	piripip	https://minka-sdg.org/users/piripip	151
6	jordiruestesalba	https://minka-sdg.org/users/jordiruestesalba	100
7	andres_valverde_valera	https://minka-sdg.org/users/andres_valverde_valera	83
8	xasalva	https://minka-sdg.org/users/xasalva	64
9	helenamarrugat	https://minka-sdg.org/users/helenamarrugat	23
10	serapio	https://minka-sdg.org/users/serapio	16
11	educacioambientalfundesplai	https://minka-sdg.org/users/educacioambientalfundesplai	9
12	ederoa	https://minka-sdg.org/users/ederoa	5
13	nozc	https://minka-sdg.org/users/nozc	4
14	oriol_d	https://minka-sdg.org/users/oriol_d	3
15	andrea	https://minka-sdg.org/users/andrea	1
16	patopi	https://minka-sdg.org/users/patopi	1
17	jaume-piera	https://minka-sdg.org/users/jaume-piera	1

Obtenemos una tabla con el nombre de los usuarios, el enlace a su perfil, donde tenemos el resto de sus datos de usuario que han querido hacer públicos, y el número de observaciones con las que han contribuido.

Guardar nuestro conjunto de datos

Para guardar en nuestro ordenador el conjunto el widget "Save Data" y lo abrimos. Desmarcamos la opción "Add type annotations to header".



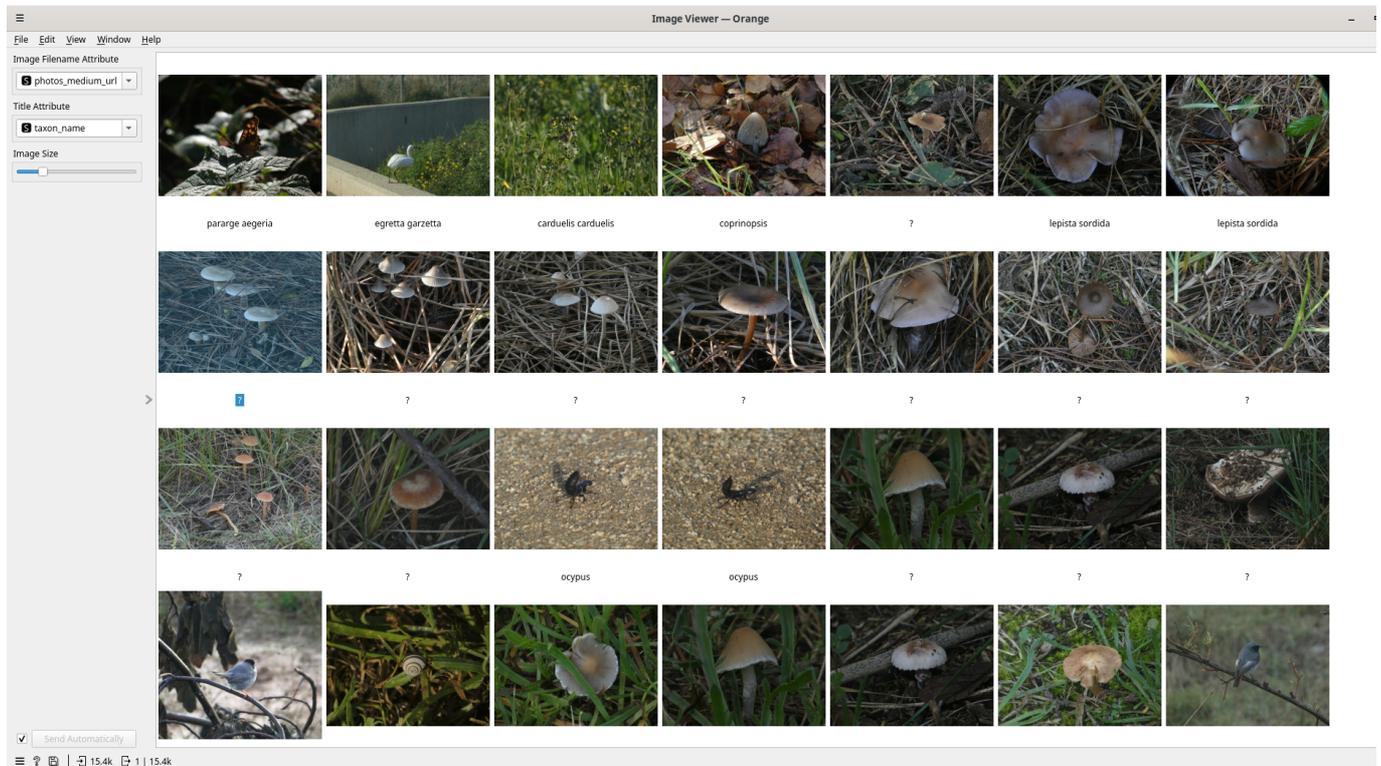
Al pinchar en "Save As" nos dará la opción de guardarlo en cualquier formato de tabla habitual (xlsx, csv, tab). Así podremos guardarnos la descarga con los datos actualizados a fecha de hoy.

Exploración de las imágenes

Hemos estado analizando los datos que teníamos en nuestra tabla de observaciones. Pero el widget de Minka nos dio dos tablas con resultados: la de observaciones y otra con las fotos. La tabla de fotos tendrá este aspecto:

Data Table (10) — Orange												
type	Selected	photos_id	taxon_name	photos_medium_url	path	obs_url	id	iconic_taxon	user_login	latitude	longitude	license_photo
1	No	126358	pararge aegeria	https://minka-sdg.org/attachments/local_photos/files/126358/medium.jpeg	103909_12635...	https://minka...	103909	Insecta	mediambient...	41.313	2.11558	cc-by-nc (c)
2	No	126005	egretta garzetta	https://minka-sdg.org/attachments/local_photos/files/126005/medium.jpeg	103676_12600...	https://minka...	103676	aves	mediambient...	41.317	2.11092	cc-by-nc (c)
3	No	126007	carduelis card...	https://minka-sdg.org/attachments/local_photos/files/126007/medium.jpeg	103675_12600...	https://minka...	103675	aves	mediambient...	41.3079	2.11441	cc-by-nc (c)
4	No	126002	copropnopsis	https://minka-sdg.org/attachments/local_photos/files/126002/medium.jpeg	103674_12600...	https://minka...	103674	fungi	mediambient...	41.2938	2.10629	cc-by-nc (c)
5	No	126001	?	https://minka-sdg.org/attachments/local_photos/files/126001/medium.jpeg	103673_12600...	https://minka...	103673	?	mediambient...	41.2938	2.10629	cc-by-nc (c)
6	No	126000	lepista sordida	https://minka-sdg.org/attachments/local_photos/files/126000/medium.jpeg	103672_12600...	https://minka...	103672	fungi	mediambient...	41.2938	2.10629	cc-by-nc (c)
7	No	125999	lepista sordida	https://minka-sdg.org/attachments/local_photos/files/125999/medium.jpeg	103671_12599...	https://minka...	103671	fungi	mediambient...	41.2933	2.10699	cc-by-nc (c)
8	No	125998	?	https://minka-sdg.org/attachments/local_photos/files/125998/medium.jpeg	103670_12599...	https://minka...	103670	?	mediambient...	41.2933	2.10699	cc-by-nc (c)
9	No	125997	?	https://minka-sdg.org/attachments/local_photos/files/125997/medium.jpeg	103669_12599...	https://minka...	103669	?	mediambient...	41.2933	2.10699	cc-by-nc (c)
10	No	126004	?	https://minka-sdg.org/attachments/local_photos/files/126004/medium.jpeg	103668_12600...	https://minka...	103668	?	mediambient...	41.2933	2.10699	cc-by-nc (c)
11	No	125995	?	https://minka-sdg.org/attachments/local_photos/files/125995/medium.jpeg	103667_12599...	https://minka...	103667	?	mediambient...	41.2933	2.10699	cc-by-nc (c)
12	No	125994	?	https://minka-sdg.org/attachments/local_photos/files/125994/medium.jpeg	103666_12599...	https://minka...	103666	?	mediambient...	41.2938	2.10629	cc-by-nc (c)
13	No	125993	?	https://minka-sdg.org/attachments/local_photos/files/125993/medium.jpeg	103665_12599...	https://minka...	103665	?	mediambient...	41.2933	2.10699	cc-by-nc (c)
14	No	125992	?	https://minka-sdg.org/attachments/local_photos/files/125992/medium.jpeg	103664_12599...	https://minka...	103664	?	mediambient...	41.2933	2.10699	cc-by-nc (c)
15	No	125991	?	https://minka-sdg.org/attachments/local_photos/files/125991/medium.jpeg	103663_12599...	https://minka...	103663	?	mediambient...	41.2933	2.10699	cc-by-nc (c)
16	No	125990	?	https://minka-sdg.org/attachments/local_photos/files/125990/medium.jpeg	103662_12599...	https://minka...	103662	?	mediambient...	41.2933	2.10699	cc-by-nc (c)
17	No	125987	ocypus	https://minka-sdg.org/attachments/local_photos/files/125987/medium.jpeg	103662_12598...	https://minka...	103662	Insecta	mediambient...	41.2933	2.10699	cc-by-nc (c)
18	No	125988	ocypus	https://minka-sdg.org/attachments/local_photos/files/125988/medium.jpeg	103661_12598...	https://minka...	103661	?	mediambient...	41.2933	2.10699	cc-by-nc (c)
19	No	125986	?	https://minka-sdg.org/attachments/local_photos/files/125986/medium.jpeg	103660_12598...	https://minka...	103660	?	mediambient...	41.2933	2.10699	cc-by-nc (c)
20	No	125985	?	https://minka-sdg.org/attachments/local_photos/files/125985/medium.jpeg	103659_12598...	https://minka...	103659	?	mediambient...	41.2932	2.10672	cc-by-nc (c)
21	No	125983	?	https://minka-sdg.org/attachments/local_photos/files/125983/medium.jpeg	103658_12598...	https://minka...	103658	aves	mediambient...	41.3364	2.10025	cc-by-nc (c)
22	No	125984	curruca melan...	https://minka-sdg.org/attachments/local_photos/files/125984/medium.jpeg	103657_12598...	https://minka...	103657	?	mediambient...	41.2932	2.10672	cc-by-nc (c)
23	No	125982	?	https://minka-sdg.org/attachments/local_photos/files/125982/medium.jpeg	103656_12598...	https://minka...	103656	?	mediambient...	41.2932	2.10672	cc-by-nc (c)
24	No	125981	?	https://minka-sdg.org/attachments/local_photos/files/125981/medium.jpeg	103655_12598...	https://minka...	103655	?	mediambient...	41.2932	2.10672	cc-by-nc (c)
25	No	125980	?	https://minka-sdg.org/attachments/local_photos/files/125980/medium.jpeg	103654_12598...	https://minka...	103654	?	mediambient...	41.2932	2.10672	cc-by-nc (c)
26	No	125979	?	https://minka-sdg.org/attachments/local_photos/files/125979/medium.jpeg	103653_12597...	https://minka...	103653	?	mediambient...	41.2932	2.10672	cc-by-nc (c)
27	No	125978	?	https://minka-sdg.org/attachments/local_photos/files/125978/medium.jpeg	103652_12597...	https://minka...	103652	aves	mediambient...	41.333	2.08518	cc-by-nc (c)
28	No	125977	phoenicurus o...	https://minka-sdg.org/attachments/local_photos/files/125977/medium.jpeg	103649_12597...	https://minka...	103649	aves	mediambient...	41.2884	2.10298	cc-by-nc (c)
29	No	125976	passer monta...	https://minka-sdg.org/attachments/local_photos/files/125976/medium.jpeg	103648_12596...	https://minka...	103648	?	mediambient...	41.2884	2.10298	cc-by-nc (c)
30	No	125967	russula	https://minka-sdg.org/attachments/local_photos/files/125967/medium.jpeg	103647_12596...	https://minka...	103647	?	mediambient...	41.2883	2.10164	cc-by-nc (c)
31	No	125965	?	https://minka-sdg.org/attachments/local_photos/files/125965/medium.jpeg	103646_12596...	https://minka...	103646	?	mediambient...	41.2884	2.10298	cc-by-nc (c)
32	No	125972	?	https://minka-sdg.org/attachments/local_photos/files/125972/medium.jpeg	103645_12597...	https://minka...	103645	?	mediambient...	41.2884	2.10298	cc-by-nc (c)
33	No	125964	?	https://minka-sdg.org/attachments/local_photos/files/125964/medium.jpeg	103644_12596...	https://minka...	103644	?	mediambient...	41.2884	2.10298	cc-by-nc (c)
34	No	125963	?	https://minka-sdg.org/attachments/local_photos/files/125963/medium.jpeg	103643_12596...	https://minka...	103643	?	mediambient...	41.2884	2.10298	cc-by-nc (c)
35	No	125961	?	https://minka-sdg.org/attachments/local_photos/files/125961/medium.jpeg	103642_12596...	https://minka...	103642	?	mediambient...	41.289	2.10273	cc-by-nc (c)
36	No	125971	?	https://minka-sdg.org/attachments/local_photos/files/125971/medium.jpeg	103641_12597...	https://minka...	103641	?	mediambient...	41.315	2.09741	cc-by-nc (c)
37	No	125973	?	https://minka-sdg.org/attachments/local_photos/files/125973/medium.jpeg	103640_12597...	https://minka...	103640	?	mediambient...	41.3152	2.09783	cc-by-nc (c)
38	No	125957	omphalina pyx...	https://minka-sdg.org/attachments/local_photos/files/125957/medium.jpeg	103640_12595...	https://minka...	103640	?	mediambient...	41.3152	2.09783	cc-by-nc (c)
39	No	125956	?	https://minka-sdg.org/attachments/local_photos/files/125956/medium.jpeg	103640_12595...	https://minka...	103640	?	mediambient...	41.3152	2.09783	cc-by-nc (c)
40	No	125955	?	https://minka-sdg.org/attachments/local_photos/files/125955/medium.jpeg	103640_12595...	https://minka...	103640	?	mediambient...	41.3152	2.09783	cc-by-nc (c)
41	No	125969	?	https://minka-sdg.org/attachments/local_photos/files/125969/medium.jpeg	103640_12595...	https://minka...	103640	?	mediambient...	41.3152	2.09783	cc-by-nc (c)
42	No	125970	?	https://minka-sdg.org/attachments/local_photos/files/125970/medium.jpeg	103640_12595...	https://minka...	103640	?	mediambient...	41.3152	2.09783	cc-by-nc (c)
43	No	125954	?	https://minka-sdg.org/attachments/local_photos/files/125954/medium.jpeg	103640_12595...	https://minka...	103640	?	mediambient...	41.3152	2.09783	cc-by-nc (c)
44	No	125953	?	https://minka-sdg.org/attachments/local_photos/files/125953/medium.jpeg	103640_12595...	https://minka...	103640	?	mediambient...	41.3152	2.09783	cc-by-nc (c)

Uno de los campos es la url de la imagen, y eso nos permite verlas dentro de Orange, conectado esa tabla al widget "Image viewer" de la sección de "Image Analytics" y ver algo así:

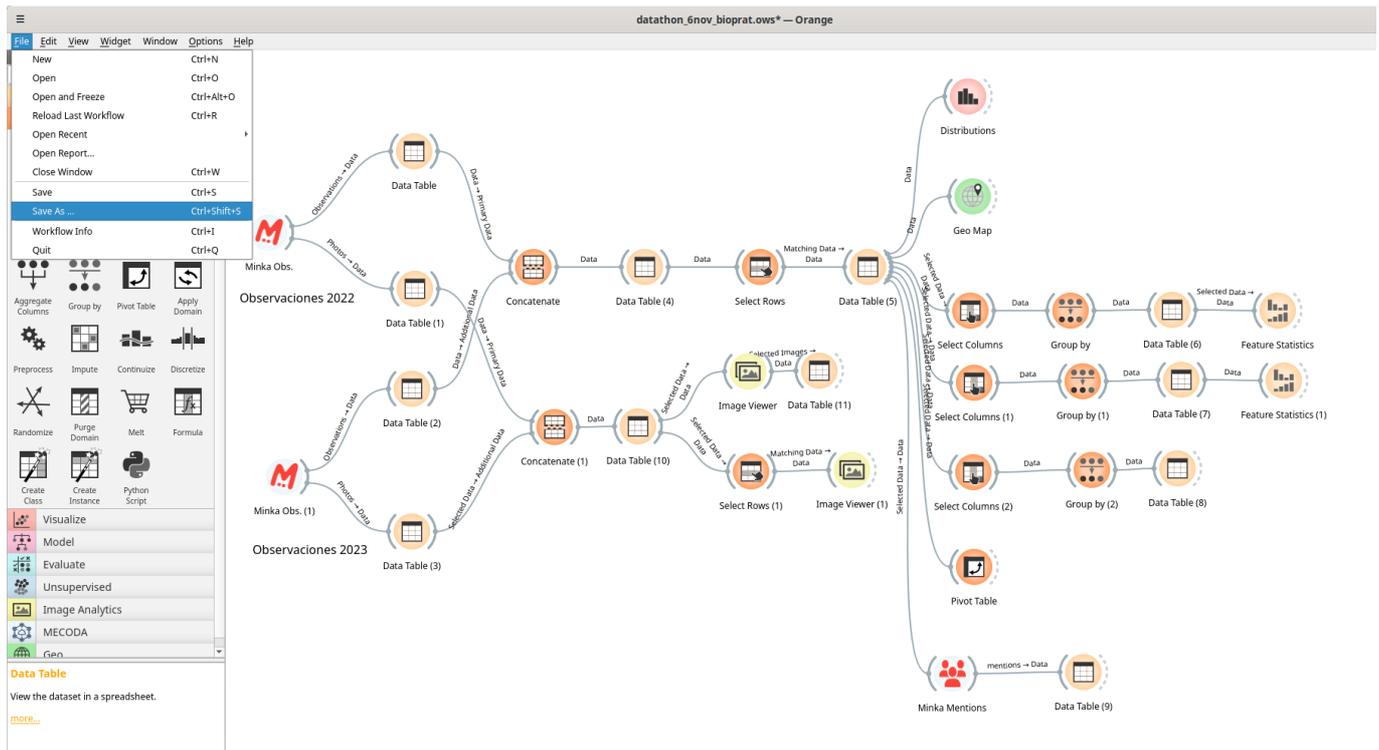


Aquí podemos modificar el título que nos aparece debajo de cada imagen, por cualquiera de los campos que tenemos en nuestra tabla, explorar todo el conjunto de imágenes. O también hacer una selección conectando la tabla a "Select Rows" y seleccionar solo las de una taxonomía (columna taxon_name) y ver las imágenes de esa taxonomía.

O podemos seleccionar todas aquellas que no están identificadas, para intentar hacerlo de un vistazo rápido, usando "Select Rows" y la condición "taxon_name" is not defined. Así podemos ver si podemos identificar alguna de las imágenes. Simplemente la seleccionamos, sacamos una tabla, y tenemos la url de la observación en Minka para acceder a ella e identificarla.

Guardar nuestro workflow para volver a ejecutarlo

El árbol de procesos que hemos creado en Orange se puede guardar en el formato nativo de Orange (.ows). De esta forma podemos volver a abrirlo y reproducir el proceso. Para guardar un workflow simplemente nos vamos al menú superior de Orange, dentro de File y pinchamos en Save As:



El workflow resultante de este ejercicio lo podéis encontrar [aquí](#) y descargarlo para abrirlo dentro de vuestro Orange.

Otros ejemplos de workflows:

- [Parc Natural Cap de Creus](#)