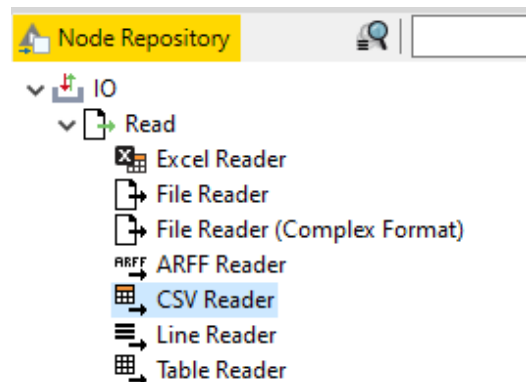
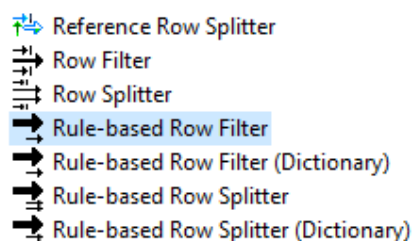


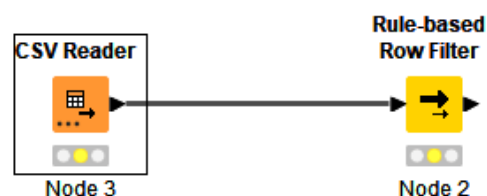
Vamos a crear un proyecto llamado **AnálisisTitanic**. Allí procesaremos los datos del fichero `titanic.csv` sobre pasajeros del Titanic. Para ello, desde el proyecto, vamos a la sección inferior izquierda (repositorio de nodos) y elegimos en la sección *IO > Read* el elemento *CSV Reader* y lo arrastramos al área de trabajo del proyecto.



Haciendo doble clic en el nodo podemos elegir el fichero CSV que queramos cargar. Ahora vamos a conectarlo con ciertos nodos de filtrado. Por ejemplo, vamos a quedarnos sólo con los supervivientes (columna `Survived` del dataset). En el repositorio de nodos, vamos a la sección de *Manipulation > Row* y elegimos *Filter*. Dentro, elegimos el nodo *Rule-based Row Filter* (filtro de fila basado en reglas) y lo arrastramos a la zona de trabajo.

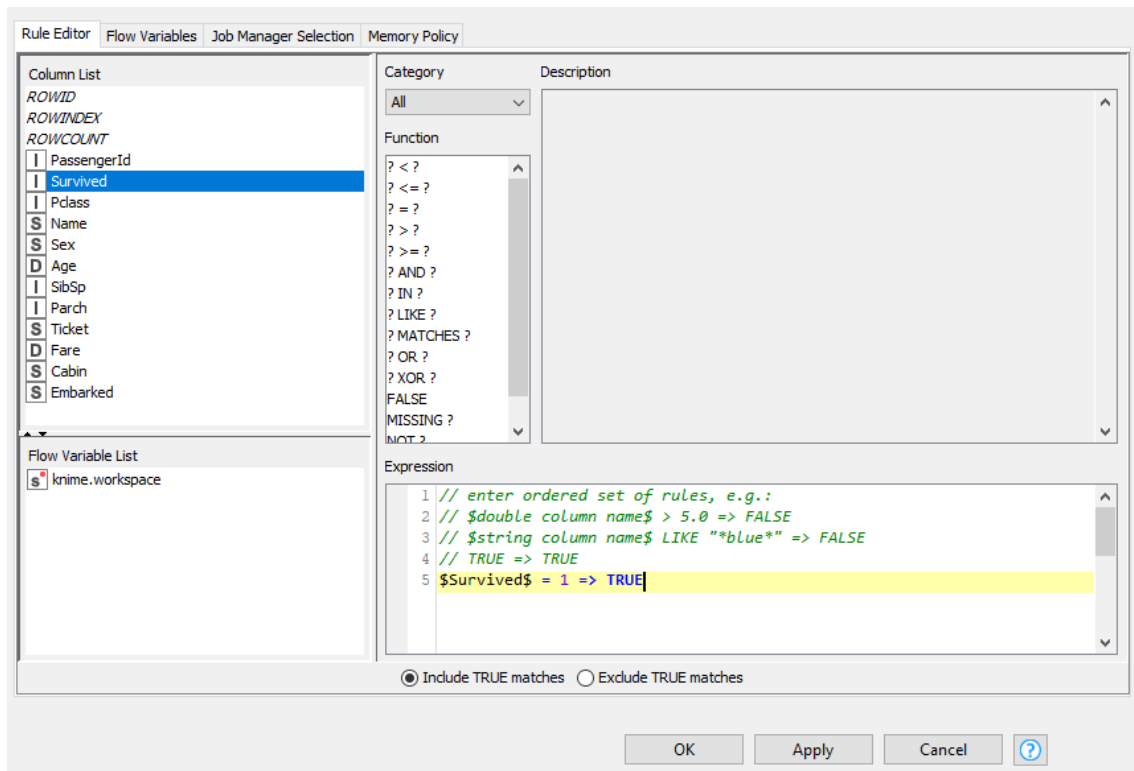


Ahora vamos a unir el nodo anterior de lectura con este nuevo nodo. Arrastramos desde la flecha de salida del nodo de lectura hasta el nodo de filtro.

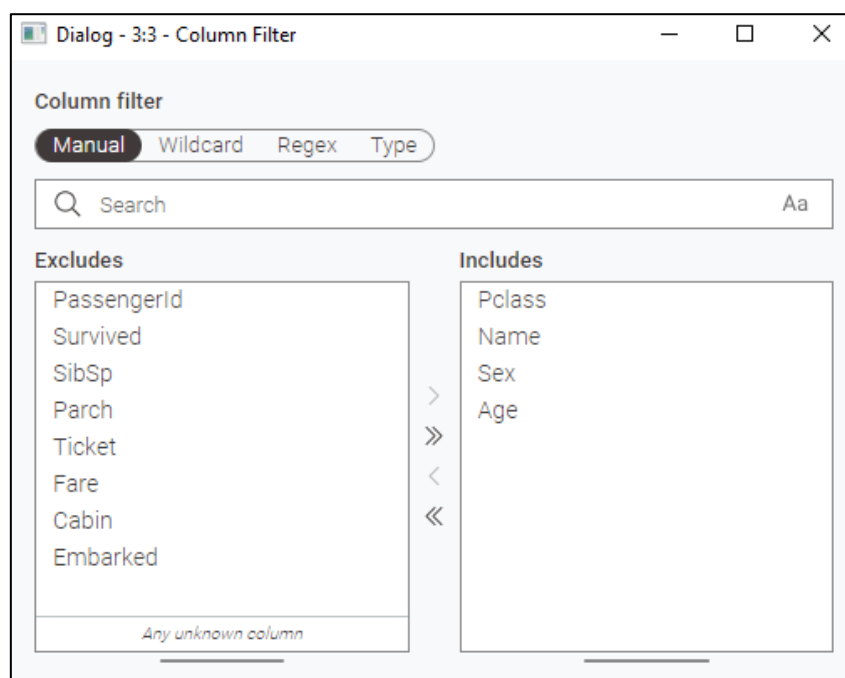


Hacemos a continuación doble clic en el nodo de filtro, y definimos el filtro. Es **IMPORTANTE** hacer esto después de unir el nodo con el de lectura, para que así este nodo ya reconozca los campos que hay en el CSV. Hacemos doble clic

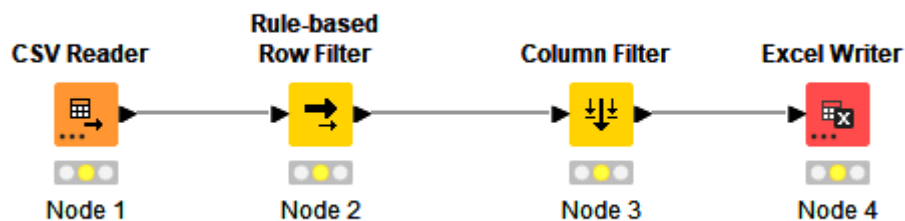
en el campo *Survived* del panel izquierdo y completamos la instrucción indicando que queremos las filas que tengan este campo a 1:



Ahora vamos a quedarnos con el nombre, sexo, edad y clase a la que pertenecen los supervivientes. Añadimos un filtro de columna, desde la sección *Manipulation* > *Column* > *Filter* del repositorio de nodos, eligiendo *Column Filter*. Conectamos este nodo con el anterior, hacemos doble clic en él y dejamos en la sección *Include* las columnas con las que queramos quedarnos.



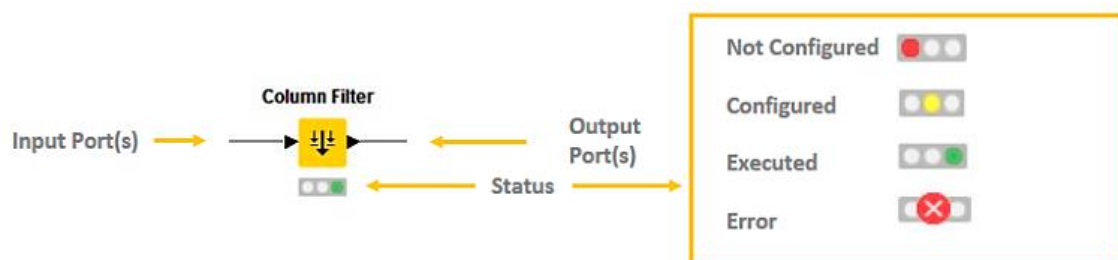
Finalmente, vamos a guardar el resultado en una hoja de cálculo Excel. Añadimos para ello un nuevo nodo de la categoría *IO > Write*. En este caso elegimos *Excel Writer*, pero notar que también podemos exportar a otros formatos, como por ejemplo generar otro CSV con los datos procesados (*CSV Writer*). Conectamos el nodo con el anterior, y hacemos doble clic para elegir el fichero de salida.



Ahora podemos poner en marcha el proceso desde el menú *Node > Execute All* o bien pulsando *Shift + F7*, o con el icono verde de la doble flecha de la barra de herramientas. Podremos comprobar que se genera el fichero de salida esperado con los datos que hemos filtrado.

Nota

Podemos ejecutar todo el *workflow* de golpe o ejecutar nodo a nodo (haciendo clic derecho en cada nodo y eligiendo *Execute*). Las luces bajo cada nodo nos indican su estado:



El proceso se puede ampliar todo lo que queramos. Por ejemplo, podemos añadir otra rama paralela que filtre sólo la columna de sexo, y con ello construya un gráfico circular de porcentajes por sexo (apartado *Views > JavaScript*, nodo *Pie/Donut Chart*). También podemos hacer doble clic en los nombres de los nodos (*Node 1*, *Node 2...* etc) y cambiarlo por otros más representativos.

