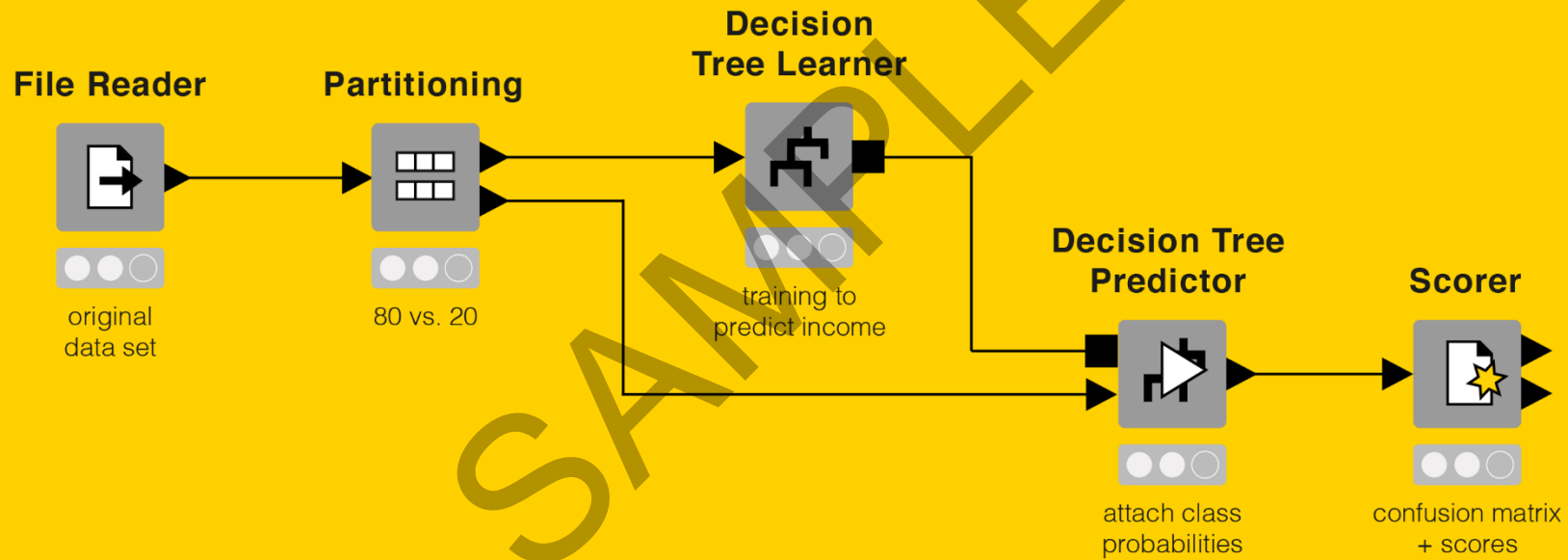


KNIME® BEGINNER'S LUCK



A Guide to KNIME Analytics Platform for Beginners

Authors: Rafael Borneo, Satoru Hayasaka and Rosaria Silipo

Copyright© 2022 by KNIME Press

Reservados todos los derechos. Esta publicación está protegida por derechos de autor, y se debe obtener permiso del editor antes de cualquier reproducción prohibida, almacenamiento en un sistema de recuperación o transmisión en cualquier forma o por cualquier medio, electrónico, mecánico, fotocopiado, grabación o similar.

Este libro ha sido actualizado para KNIME 4.4.

Para obtener información sobre permisos y ventas, escriba a:

KNIME Press
Talacker 50
8001 Zurich
Switzerland

knimepress@knime.com

Table of Contents

Prólogo.....	12
Agradecimientos (versión inglés)	13
Nota sobre la traducción al español.....	14
Capítulo 1. Introducción	16
1.1. Propósito y estructura del libro.....	16
1.2. Comunidad KNIME.....	17
Páginas web útiles.....	17
Cursos, Eventos y Videos.....	18
Libros.....	19
KNIME Hub.....	19
1.3. Descarga e Instalación de KNIME Analytics Platform	23
1.4. Workspace (Ambiente de Trabajo).....	24
La ventana “Workspace Launcher”	25
1.5. KNIME Workflow (flujos de trabajo)	26
¿Qué es un workflow?	26
¿Que es un nodo?	27
1.6. Archivos <i>.knwf</i> and <i>.knar</i>	27
1.7. KNIME workbench (banco de trabajo)	28
El KNIME Workbench.....	30
Menu Principal.....	32
Hotkeys (teclas importantes)	37
Repositorio de Nodos (Node Repository)	37
Cuadro de búsqueda	38

KNIME Explorer (Explorador de KNIME)	38
Mi central KNIME (My-KNIME-Hub)	39
El "EXAMPLES Server"	40
Montaje de servidores en KNIME Explorer	41
El Editor de Workflow	42
Personalización del editor de flujo de trabajo	43
Anotaciones del Workflow	44
Otras personalizaciones del Workbench	44
Vista "Node Monitor"	45
1.8. Descargar las Extensiones de KNIME	45
Instación de KNIME Extensions	46
1.9. Datos y workflows para éste libro	47
1.10. Ejercicios	48
Ejercicio 1	48
Ejercicio 2	49
Ejercicio 3	50
Capítulo 2. Mi primer workflow	53
2.1. Operaciones en el Workflow	53
Crear un Nuevo Workflow Group	54
Crear un nuevo workflow	55
Guardar un workflow	56
Borrar un workflow	56
2.2. Operaciones con Nodos (Node operations)	57
Crear un nuevo nodo	57
Texto del Nodo	58

Configurar un nodo.....	58
Ejecutar un nodo.....	59
Descripcion del Nodo	60
Ver los Datos Procesados por el Nodo.....	61
2.3. Leer datos desde un archivo.....	61
Crear un nodo lector de CSV "CSV Reader"	62
Configure the "CSV Reader node".....	63
Personalizar las propiedades de las columnas.....	64
El protocolo <i>knime://</i>	67
2.4. Estructura de datos y tipos de datos en KNIME.....	68
Estructura de datos en KNIME	70
2.5. Filtrado de columna de Datos.....	71
Crear un "Column Filter" node	71
Configure un nodo "Column Filter"	74
2.6. Filtrado de "Data Rows".....	75
Creando un nodo "Row Filter"	76
Configurando el nodo "Row Filter"	76
Criterios de filtrado (Row filter)	78
2.7. Guardar en un archivo (Write Data to a File).....	80
Creando un nodo "CSV Writer"	81
Configurar el nodo "CSV Writer"	82
2.8. Ejercicios	84
Ejercicio 1.....	84
Ejercicio 2.....	87
Capítulo 3. Mi primera exploración de datos	91

3.1. Introducción	91
3.2. Reemplazar valores en las columnas.....	92
Column Rename (Renombrar una columna)	93
Nodo "Rule Engine "	96
3.3. Separación de cadenas de texto (String Splitting)	98
Cell Splitter por Posición.....	99
Cell Splitter [by Delimiter].....	100
RegEx Split (= Cell Splitter by RegEx).....	101
3.4. Manipulación de cadenas (String Manipulation).....	102
Nodo "String Manipulation"	103
Case Converter	105
String Replacer.....	106
Column Combiner.....	107
Nodo "Column Resorter"	109
3.5. Conversiones de tipos de datos (Type Conversions)	110
Number To String (número a cadena)	111
String To Number (cadena a número)	112
Double To Int.....	113
3.6. Operaciones de Base de Datos (Database Operations).....	114
SQLite Connector.....	116
MySQL Connector.....	117
Workflow Credentials (credenciales)	118
DB Writer	120
Importar un controlador de base de datos JDBC (JDBC Database Driver).....	121
DB Table Selector (.....	124

Lector de DB (DB Reader)	125
3.7. Agregaciones y agrupaciones (Aggregations and Binning)	125
Contenedores numéricos (Numeric Binner).....	127
Agrupando (GroupBy: "Groups" tab)	128
GroupBy: Aggregation tabs.....	129
Pivoting (Pivotando).....	130
3.8. Nodos para Visualización de Datos (Nodes for Data Visualization)	132
3.9. Gráfico de dispersión (Scatter Plot)	133
Gráfico de dispersión: vista interactiva.....	135
3.10. Propiedades gráficas (Graphical Properties).....	136
Gestor de color "Color Manager".....	137
3.11. Gráficos de Líneas (Line Plots) y Coordenadas paralelas (Parallel Coordinates)	139
Line Plot.....	140
Coordenadas paralelas (Parallel Coordinates)	142
3.12. Gráficos de Barras e Histogramas (Bar Charts and Histograms).....	143
Bar Chart	144
Nodo "Table View"	147
3.13. Ejercicios	149
Ejercicio 1	149
Ejercicio 2.....	151
Ejercicio 3.....	151
Chapter 4. Mi Primer Modelo	155
4.1. Introducción	155
4.2. Dividir y combinar conjuntos de datos.....	156
Muestreo por filas (Row Sampling).....	157

Partitioning.....	158
Shuffle	159
Concaternar (Concatenate)	160
4.3. Transformando columnas (Transform Columns)	162
PMML	163
Valores faltantes (Missing Value).....	164
Normalizer (normalizador).....	166
Metodos de normalización (Normalization Methods).....	167
Nodo Normalizer (Apply)	167
4.4. Modelos de Aprendizaje Automático (Machine Learning Models).....	169
Modelo Naïve Bayes.....	170
Scorer (Javascript)	173
Arboles de Decisión (Decision Tree).....	178
Curva ROC	187
Red neuronal artificial (Artificial Neural Network)	189
Write/Read Models to/from file.....	192
Nodo "Statistics"	195
Regresión	197
Clustering	199
Nodo "Cluster Assigner"	201
Test de hipótesis	202
4.5. Ejercicios	203
Ejercicio 1	203
Ejercicio 2.....	204
Ejercicio 3.....	205

Capítulo 5. Preparación de los datos para la elaboración de informes	207
5.1. Introducción	207
5.2. Transformación de filas (Transform Rows)	208
RowID	211
Nodo "Sorter"	214
5.3. Uniendo Columnas.....	214
Joiner.....	216
Nodo Joiner: pestaña "Joiner Settings"	217
Nodo Joiner: pestaña "Column Selection"	218
5.4. Nodos Miscelaneos (Misc Nodes)	220
Java Snippet (simple).....	221
Java Snippet	222
Nodo Math Formula	223
Math Formula (Multi Column)	225
5.5. Limpieza final (Cleaning Up the Final Workflow).....	227
Colapsar nodos preexistentes en un Meta-nodo	227
Crear un Meta-nodo desde cero	228
Expandir y reconfigurar un metanodo.....	230
5.6. Próximo paso: Crear un Reporte.....	231
5.7. Ejercicios	232
Ejercicio 1	232
Ejercicio 2.....	233
Ejercicio 3.....	234
Capítulo 6. Tableros de mando con vistas compuestas.....	237
6.1. El tablero de mando (Dashboard).....	237

6.2. Los Nodos	238
Text Output Widget.....	239
6.3. El Componente.....	240
6.4. Agregando Colores	243
6.5. La Vista compuesta (Composite View).....	246
6.6. En el WebPortal.....	250
6.7. Ejercicios	251
Ejercicio 1	251
Capítulo 7. Elaboración de informes con BIRT	254
7.1. Informes con BIRT	254
7.2. Instalación de la extensión del diseñador de informes (BIRT).....	255
7.3. Marcado de datos en el flujo de trabajo.....	256
Nodo "Data to Report"	257
7.4. De KNIME a BIRT y viceversa.....	257
7.5. El entorno BIRT.....	259
7.6. La Plantilla (The Layout).....	263
7.7. Las Tablas (The Tables).....	266
Toggle Breadcrumb.....	270
7.8. Salto de página	275
7.9. Los Gráficos (Charts)	275
Seleccionar tipo de gráfico	276
Selecciona "Data"	277
Formatear el Gráfico	279
7.10. Hojas de estilo (Style Sheets)	287
Crear un Nuevo estilo	288

Aplicando un Style Sheet	289
7.11. Generar el documento final	290
7.12. Texto Dinámico	290
7.13. Informes con otras herramientas	293
7.14. Ejercicios	295
Ejercicio 1	295
Referencias	297
Índice de temas	298

SAMPLE

Prólogo

Este es el primer libro que escribí en 2010 para [KNIME Press](#) sobre cómo utilizar KNIME Analytics Platform. Dado que nos acercamos al décimo aniversario de este libro, nosotros (el equipo de prensa de KNIME y yo) pensamos que podría necesitar un nuevo texto de Prólogo. En realidad, no necesita una actualización general, ya que desde su nacimiento se ha actualizado dos veces al año cada año, después de cada nueva versión de KNIME Analytics Platform; no inmediatamente después, pero lo suficientemente cerca.

¡Eso es correcto! KNIME Beginner's Luck, como todos los demás libros electrónicos de KNIME Press, es un libro electrónico en vivo que cambia constantemente para adaptarse a la versión más reciente del software. Esta vivacidad del libro electrónico es también la razón por la que rara vez se ha impreso. ¡Actualizar páginas impresas es, sin duda, más difícil que actualizar un archivo pdf!

Como este es el primer libro, inevitablemente se trata de los conceptos básicos: los conceptos básicos de KNIME Analytics Platform, por supuesto, y también los conceptos básicos de un proyecto de ciencia de datos. Este libro lo guía a través de las funciones de acceso más importantes, las operaciones de transformación de datos y, por supuesto, los nodos de aprendizaje automático disponibles en KNIME Analytics Platform. Complementado con muchos ejemplos de flujos de trabajo, ejercicios y capturas de pantalla, lo familiarizará rápidamente con las funciones básicas del software. Si está buscando temas más avanzados, no los encontrará aquí. En cambio....

Si desea obtener más información sobre algoritmos avanzados de aprendizaje automático, variables de flujo o bucles, consulte la secuela de este libro: "[KNIME Advanced Luck](#)". Si desea obtener más información sobre el procesamiento de texto, consulte el libro "[From Words To Wisdom](#)". Si viene de esa escuela de pensamientos donde la lectura de manuales o instrucciones está sobrevalorada, puede comenzar directamente leyendo sobre soluciones a estudios de casos en varios campos de aplicación en nuestra colección "[Practicing Data Science](#)". Si su trabajo consiste más en integrar y combinar diferentes fuentes de datos y tipos de datos, entonces el libro para usted es "[Will they blend?](#)" collection. Si está realizando la transición de SAS, Excel o Alteryx a KNIME, hay más libros disponibles en la página de [KNIME Press](#) .

Todo esto es para decir que el equipo de KNIME Press y yo hemos estado trabajando arduamente para proporcionarles el material de aprendizaje, los libros y los tutoriales, para que sean cada vez más productivos con el software KNIME y los conceptos de ciencia de datos.

Rosaria Silipo (Autora de varios libros de KNIME Press, PhD)

Agradecimientos (versión inglés)

En primer lugar, me gustaría agradecer a todo el equipo KNIME por su paciencia al tratar conmigo y con mis infinitas preguntas.

Entre todos los demás miembros del equipo de KNIME, me gustaría agradecer específicamente a Peter Ohl por haber revisado este libro para encontrar posibles aspectos que no fueran compatibles con las mejores prácticas de KNIME.

También me gustaría agradecer a Casiana Rimbu por ayudarme a proporcionar las capturas de pantalla más bellas, claras y artísticas que jamás pude imaginar y a Meta Brown por animarme en los primeros pasos para desarrollar la idea embrionaria de escribir este libro.

Muchas gracias finalmente a Heather Fyson por revisar el inglés del libro.

SAMPLE

Nota sobre la traducción al español

La plataforma KNIME es una excelente herramienta para el análisis de datos. Se basa en el uso de programación gráfica en donde el usuario no necesita saber programación en ningún lenguaje computacional para poder utilizarla. Se utilizan nodos que secuencialmente realizan tareas con el fin de realizar el análisis final de los resultados.

La plataforma KNIME se encuentra solamente en el idioma inglés. Todos los nodos, ventanas, menús, asistentes, etc..., están en inglés (el idioma universal del data science). Por tanto, si bien en éste libro se ha hecho en muchos casos la traducción al español de lo que aparece en la plataforma KNIME, cuando se use dicha plataforma todo seguirá apareciendo en el idioma inglés original. Es por esto que todas las figuras aparecen en inglés.

Por ejemplo: "workflow" (base del uso de KNIME) se puede traducir como flujo de trabajo. Sin embargo en la plataforma Ud. seguirá viendo la palabra "workflow" en inglés. NO VERRÁ flujo de trabajo. Otro ejemplo "workspace" se traduce como "espacio de trabajo" pero Ud. seguirá viendo "workspace" en la plataforma.

Lo que se ha hecho es: dejar una mezcla de español e inglés. Así pues, Ud. verá algo así como:

Si ha iniciado KNIME por primera vez, su panel "KNIME Explorer" en la esquina superior izquierda del KNIME workbench (banco de trabajo) contiene solo un grupo de flujo de trabajo (carpeta) llamado "**Example Workflows**". Esta carpeta "Example Workflows" contiene varias subcarpetas, cada una con flujos de trabajo básicos para casos de uso muy comunes:

- **Basic Examples**. Los flujos de trabajo de la subcarpeta "Ejemplos básicos" muestran operaciones generales básicas, como importar datos, combinar datos, ETL, entrenar y evaluar un modelo y, finalmente, mostrar los resultados en un informe simple.
- **Customer Intelligence**. Los flujos de trabajo básicos para la predicción de abandono, la calificación crediticia y la segmentación de clientes están disponibles dentro de la subcarpeta "*Customer Intelligence*".
- **Retail**. Un motor de recomendaciones está integrado en la subcarpeta "*Retail*".
- **Social Media**. Un ejemplo de análisis de redes sociales está disponible en in "*Social Media*".

A medida que transcurre la traducción puede que se use más el español que el inglés siempre con la advertencia de que en la plataforma KNIME todo seguirá apareciendo en el idioma inglés. Por supuesto que las explicaciones, instrucciones, comentarios y sugerencias de cómo utilizar la plataforma han sido traducidos al español.

SAMPLE

Capítulo 1. Introducción

1.1. Propósito y estructura del libro

¡Vivimos en la era de los datos! Cada compra que hacemos se registra debidamente; cada transacción de dinero se registra cuidadosamente; cada clic web termina en un archivo de clics web. Hoy en día todo lleva un chip RFID y puede registrar datos. Tenemos datos disponibles como nunca antes. ¿Qué podemos hacer con todos estos datos? ¿Podemos darle algún sentido? ¿Podemos usarlo para aprender algo útil y rentable? Necesitamos una herramienta, un bisturí quirúrgico que nos permita cortar cada vez más nuestros datos, mirarlos desde muchas perspectivas diferentes, representar su estructura subyacente.

Supongamos entonces que tenemos esta enorme cantidad de datos ya disponibles, esperando ser analizados. ¿Cuáles son las opciones para que un profesional ingrese al mundo de Business Intelligence (BI) y Data Science (DS)? Las opciones disponibles son, por supuesto, múltiples y están creciendo rápidamente. Si nuestro profesional no controla un presupuesto excesivo, podría recurrir al mundo del software de código abierto. El software de código abierto, sin embargo, es más que una elección basada en el dinero. En muchos casos, representa una filosofía de software para compartir y controlar los recursos que muchos profesionales apoyan.

Dentro del mundo del software de código abierto, podemos encontrar algunas herramientas de ciencia de datos y BI. [KNIME Analytics Platform](#) representa una elección fácil para el profesional novato. No requiere el aprendizaje de un script específico y ofrece una interfaz gráfica de usuario (GUI) para implementar y documentar los procedimientos de análisis. Además, y esto no es una ventaja secundaria, KNIME Analytics Platform puede funcionar como una plataforma de integración a la que se pueden conectar muchas otras herramientas de BI y ciencia de datos. Entonces, no solo es posible, sino incluso fácil, analizar datos con KNIME Analytics Platform y luego construir cuadros de mando sobre los mismos datos procesados con una herramienta de BI diferente.

Aunque KNIME Analytics Platform es muy simple e intuitiva en su uso, cualquier principiante se beneficiaría de una orientación acelerada a través de todos los nodos, categorías y configuraciones. Este libro representa la suerte del principiante, porque tiene como objetivo ayudar a cualquier principiante a preparar su proceso de aprendizaje. Este libro no pretende ser una guía exhaustiva de todo el software KNIME. No cubre implementaciones bajo KNIME Server, que no es de código abierto, o temas que se consideran avanzados. Las variables de flujo, por ejemplo, y las implementaciones de consultas de bases de datos SQL se analizan en el libro siguiente. "[KNIME Advanced Luck](#)".

Este libro está dividido en siete capítulos. El primer capítulo cubre los conceptos básicos de KNIME Analytics Platform, mientras que el capítulo dos lleva al lector de la mano a la implementación de la primera aplicación KNIME. Desde el tercer capítulo, comenzamos la

exploración de conceptos de ciencia de datos de una manera más profunda. De hecho, el tercer capítulo explica cómo realizar una exploración y visualización de datos básica, en términos de nodos y flujo de procesamiento. El capítulo cuatro está dedicado al modelado de datos. Cubre algunos enfoques demostrativos del aprendizaje automático, Naïve Bayes, árboles de decisión y redes neuronales artificiales. Por último, los capítulos cinco y seis están dedicados a la presentación de informes. Por lo general, los resultados de una investigación basada en la visualización de datos o, en una fase posterior, en el modelado de datos, deben mostrarse en algún momento a los colegas, la gerencia, los directores, los clientes o los trabajadores externos. Por lo tanto, la presentación de informes es una fase muy importante al final del proceso de análisis de datos. El capítulo cinco muestra cómo preparar los datos para exportarlos a un informe, mientras que el capítulo seis muestra cómo crear el informe en sí.

Cada capítulo guía al lector a través de un [ETL](#) o un proceso de aprendizaje automático (ML) paso a paso. Cada paso se explica en detalle y ofrece algunas explicaciones sobre los empleos alternativos de los nodos actuales. Al final de cada capítulo se proponen al lector varios ejercicios para poner a prueba y perfeccionar lo aprendido hasta el momento.

Los ejemplos y ejercicios de este libro se han implementado utilizando KNIME 4.3. También deberían funcionar con versiones posteriores de KNIME, aunque puede haber ligeras diferencias en su apariencia.

1.2. Comunidad KNIME

Al ser un software de código abierto, KNIME Analytics Platform se beneficia de una serie de foros y grupos de usuarios de KNIME en todo el mundo. Esta es una buena red de seguridad para consejos, sugerencias y material de aprendizaje. Informamos a continuación los sitios y grupos más populares.

Páginas web útiles

<http://www.knime.com>

La página raíz del sitio web de KNIME.

<https://www.knime.com/software-overview>

El primer lugar para buscar una descripción general de todos los productos KNIME. La plataforma de análisis KNIME de código abierto se puede descargar aquí.

<https://www.knime.com/knime-self-paced-courses>

Esta es la página raíz de los cursos de autoaprendizaje de KNIME (e-learning) donde puede aprender más sobre las funcionalidades específicas de KNIME. Cubre todo el ciclo de creación de la ciencia de datos, desde el acceso y la exploración de datos hasta el aprendizaje automático y las estructuras de control.

<https://hub.knime.com/>

KNIME Hub es un repositorio público de material KNIME: flujos de trabajo, ejemplos, extensiones, plantillas de componentes y nodos. Si desea comenzar con ejemplos prácticos, aquí puede encontrar muchos según el término de búsqueda que ingrese.

<https://forum.knime.com/>

Lo que encuentro particularmente útil es el Foro KNIME. Aquí puede hacer preguntas sobre cómo utilizar KNIME Analytics Platform o sobre cómo ampliarla con nuevos nodos. Alguien de la comunidad KNIME responde siempre y rápidamente.

Cursos, Eventos y Videos

Cursos de KNIME Analytics Platform

KNIME ofrece periódicamente cursos presenciales y en línea para el software KNIME. Esto incluye elementos básicos y avanzados. Para verificar la próxima fecha disponible y registrarse, simplemente vaya a la página web de KNIME Events <https://www.knime.com/learning/events> y seleccione la carpeta Online Courses.

KNIME Webinars

Con frecuencia se encuentran disponibles varios seminarios web sobre temas específicos, como nodos de química, minería de texto, integración con otras herramientas de análisis, aprendizaje automático automatizado, aprendizaje profundo, análisis de series de tiempo, mejores prácticas, etc. Para conocer los próximos seminarios web programados, consulte la página web de KNIME Events en <https://www.knime.com/learning/events> y seleccione el tab de Webinars.

KNIME Meetups y KNIME Summits

KNIME Meetups y KNIME Data Talks se llevan a cabo periódicamente en todo el mundo. Estas son buenas oportunidades para aprender más sobre el software KNIME, inspirarse en nuevos proyectos de ciencia de datos y conocer a otras personas de la comunidad KNIME. (<https://www.knime.com/learning/events>)

Canal KNIME TV en YouTube

KNIME tiene su propio canal de videos en YouTube, llamado KNIMETV. Allí, una serie de videos están disponibles para aprender más sobre muchos temas diferentes y especialmente para actualizarse sobre las nuevas funciones en los nuevos lanzamientos de KNIME. (<http://www.youtube.com/user/KNIMETV>)

Libros

Funciones avanzadas en la plataforma de análisis KNIME

Para uso avanzado:
Rosaria Silipo, Jeanette Prinz, "KNIME Advanced Luck"
(<https://www.knime.com/knimepress/advanced-luck>)

Reporting Suite

KNIME Reporting Suite se basa en BIRT, otra herramienta de código abierto para informes. Aquí hay una guía básica sobre cómo usar BIRT:
D. Peh, N. Hague, J. Tatchell, "BIRT. A field Guide to Reporting.", Addison-Wesley, 2008

Data Science y KNIME

Para obtener una descripción general de la ciencia de datos, la minería de datos y el análisis de datos, consulte:
.Berthold M.R., Borgelt C., Höppner F., Klawonn F., Silipo R., "Guide to intelligent data science", Springer 2020.

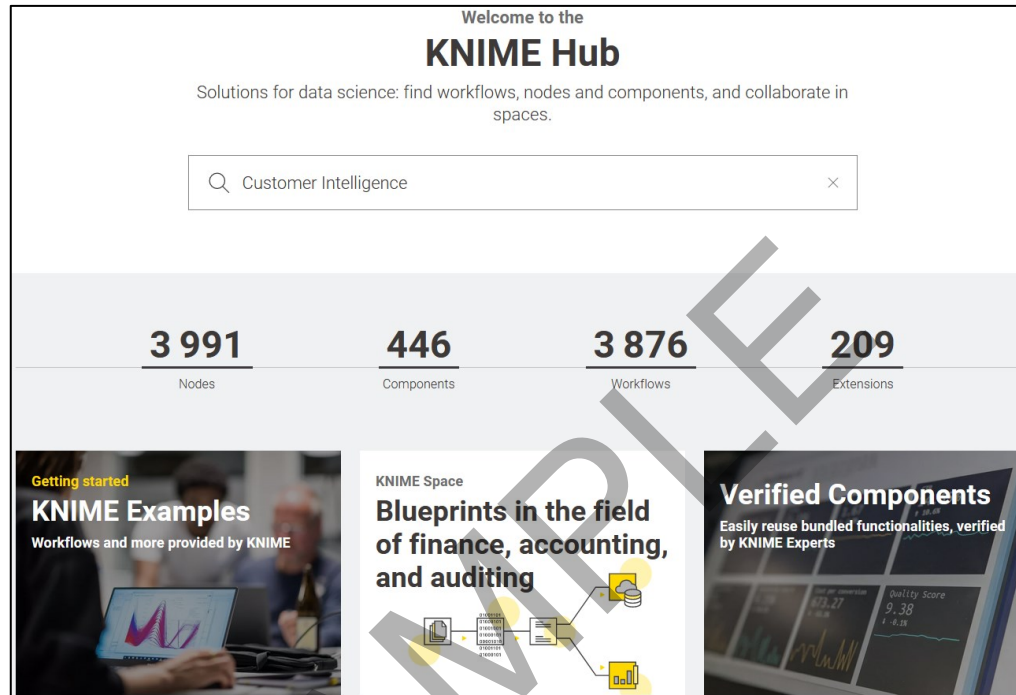
KNIME Hub

Existe un lugar privilegiado donde encontrar información sobre los nodos KNIME y flujos de trabajo, así como ejemplos para sus próximos proyectos: el KNIME Hub (<https://hub.knime.com/>).

KNIME Hub es un repositorio de aplicaciones, componentes y nodos para reciclar, reutilizar y ensamblar en KNIME Analytics Platform. O como dice en la página de inicio: KNIME Hub es "el lugar para buscar y colaborar en flujos de trabajo y nodos de KNIME. Aquí puede encontrar soluciones para sus preguntas sobre ciencia de datos".

Cuando accede a KNIME Hub por primera vez, se encontrará con figura 1.1. Esta página ofrece algunos enlaces a la documentación de la guía de inicio, el Foro KNIME y el blog KNIME. Lo más importante es que en la parte superior ofrece un cuadro de búsqueda para encontrar aplicaciones, nodos y componentes cargados por los usuarios de KNIME en este lugar.

1.1. KNIME Hub home page <https://hub.knime.com/>



Si escribimos "Customer Intelligence" en el cuadro de búsqueda, encontraremos una lista de nodos y flujos de trabajo relacionados con Customer Intelligence. Seleccionemos solo "Flujos de trabajo" en el menú superior. Luego, a continuación, en la figura 1.2, puede ver la lista de aplicaciones (flujos de trabajo) que implementan algunos aspectos de la inteligencia del cliente, y que están debidamente etiquetadas, tal como las cargan los usuarios de la comunidad KNIME. De hecho, puede cargar sus propias aplicaciones desarrolladas en KNIME Hub. Todo lo que necesita es una cuenta con el [KNIME Forum](#).

1.2. Lista de aplicaciones (workflows) relacionados (y etiquetados) con Customer Intelligence disponibles en el KNIME Hub

The screenshot shows the KNIME Hub interface with a search bar containing 'Customer Intelligence'. The results page displays 405 results. A navigation bar at the top of the results section includes 'All', 'Nodes', 'Components', 'Workflows', and 'Extensions'. Four workflow cards are visible:

- B2B Customer Intelligence Use Case**: Showcasing tools and methods available for the Citizen Data Scientist to improve and predict B2B customer behaviour. Path: knime > Examples > 50_Applications > 42_Customer_Intelligence > 01_B2B_Customer_Analytics. Includes a green 'A' badge.
- Training a Churn Predictor**: This workflow is an example of how to build a basic PMML model for a churn prediction using a Decision Tree algorithm. Path: knime > Examples > 50_Applications > 18_Churn_Prediction > 01_Training_a_Churn_Predictor. Includes tags for 'Customer Intelligence' and 'churn'.
- Basic Customer Segmentation**: This workflow implements a basic customer segmentation through a clustering procedure. No input is required from the business analyst. Path: knime > Examples > 50_Applications > 24_Customer_Segmentation_Use_Case > 01_Basic_Customer_Segmentation_Use_Case. Includes tags for 'clustering', 'k-Means', and 'customer segmentation'.
- Customer Segmentation**: This workflow performs customer segmentation by means of clustering k-Means node. The second part of the workflow implements an interactive wizard on the WebPortal to visualize and label (or write no... Path: knime > Examples > 50_Applications > 24_Customer_Segmentation_Use_Case > 02_Customer_Segmentation_Use_Case. Includes tags for 'clustering', 'k-Means', and 'customer segmentation'.

Al hacer clic en una de las aplicaciones de la lista, se abre la página web correspondiente (Fig. 1.3), con una imagen explicativa del flujo de trabajo implementado.

En la esquina superior derecha, puede ver el botón para iniciar sesión con la cuenta de KNIME Forum. Estar conectado le permite cargar, descargar, comentar, dar me gusta y actualizar los espacios para los que tiene permisos. Debajo de eso, puede encontrar la imagen del autor y debajo una serie de botones de utilidad: para descargar el flujo de trabajo, darle Me gusta, arrastrarlo y soltarlo en su instalación de KNIME Analytics Platform, y el enlace permanente corto a este flujo de trabajo para compartir.

Si pasa el cursor sobre la imagen del autor y tiene permisos de edición para este espacio Hub, aparecerá un bolígrafo. Al hacer clic en él, podrá permitir que otros usuarios de KNIME carguen y cambien este espacio.

El KNIME Hub es un repositorio de workflows, pero también de nodos, componentes y extensiones de KNIME.

1.3. La página dedicada a la aplicación denominada "Segmentación de clientes" en el KNIME Hub, con enlace corto <https://kni.me/w/37cHxqr6dbllUeP>

The screenshot shows the KNIME Hub interface for a workflow titled "Customer Segmentation". The page includes a search bar, navigation breadcrumbs, and a workflow diagram. The workflow diagram is divided into three main sections:

- Data Reading:** Includes nodes for Excel Reader (XLS), File Reader, and Contract Data.
- Parameter Selection:** Includes nodes for Joiner, Number To String, Define Cluster Parameters, and k-Means clustering.
- Clustering:** Includes the k-Means clustering node.

The workflow description states: "Customer Segmentation. This workflow performs: 1. clustering (k-Means), 2. visualization and labelling of clusters, 3. summary of cluster stats." The workflow is last edited on 20 Feb 2019 and has 1 heart and 4365 views. A short link is provided: <https://kni.me/w/37cHxqr6dbllUeP>.

This workflow performs customer segmentation by means of clustering k-Means node. The second part of the workflow implements an interactive wizard on the WebPortal to visualize and label (or write notes) about the single clusters.

1.3. Descarga e Instalación de KNIME Analytics Platform

Hay dos productos KNIME disponibles:

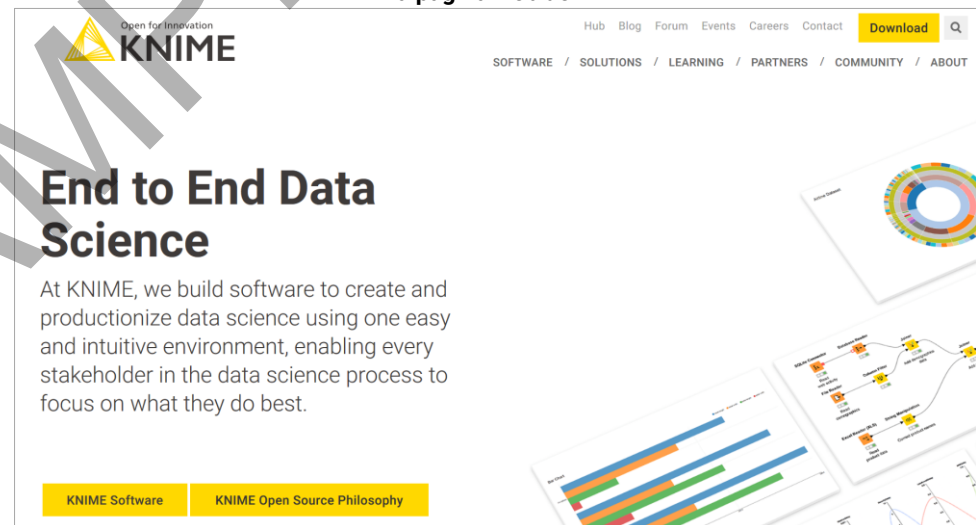
- la [KNIME Analytics Platform \(código abierto\)](https://www.knime.com/software-overview), que puede ser descargada de manera gratuita en: <https://www.knime.com/software-overview> bajo licencia GLP version 3.
- el [KNIME server](https://www.knime.com/knime-server), que se describe en: <https://www.knime.com/knime-server>

Desde el punto de vista analítico, las funcionalidades de los dos productos son las mismas. El servidor KNIME además incluye una serie de funciones TI útiles para la colaboración en equipo, la implementación y la gestión del flujo de trabajo empresarial, el almacenamiento de datos, la integración y la escalabilidad para el laboratorio de ciencia de datos. En este libro, sin embargo, trabajaremos con KNIME Analytics Platform (código abierto). Para comenzar a jugar con KNIME Analytics Platform, primero, debe descargarlo en su máquina.

Como descargar KNIME Analytics Platform

- Vaya a www.knime.com
- En la esquina superior derecha de la página principal, haga clic en "Descargar".
- Proporcione un poco de información sobre usted (que se agradece), luego continúe con el paso 2 "Descargar KNIME"
- Elija la versión que se adapte a su entorno (Windows / Mac / Linux, 32 bits / 64 bits, con o sin instalador para Windows), incluyendo opcionalmente todas las extensiones gratuitas
- Acepta los términos y condiciones.
- Empiece a descargar. Terminará con un archivo comprimido (* .zip), un archivo de almacenamiento autoextraíble (* .exe) o una aplicación de instalación

1.4. La página web de KNIME



- Para archivos .zip y .exe, simplemente descomprímalos en la carpeta de destino. Si seleccionó la versión del instalador, simplemente ejecútela y siga las instrucciones del instalador.

1.4. Workspace (Ambiente de Trabajo)

Para iniciar KNIME Analytics Platform, abra la carpeta donde se instaló KNIME y ejecute `knime.exe` (o `knime` en una máquina Linux / Mac). Si ha instalado KNIME utilizando el instalador, puede simplemente hacer clic en el icono en su escritorio o en el menú principal de Windows.

Si está iniciando KNIME Analytics Platform por primera vez, se le preguntará si desea compartir sus estadísticas de uso con KNIME. Estos números se utilizarán para impulsar el motor de recomendaciones de mejores prácticas proporcionado dentro del KNIME Analytics Platform workbench: Workflow Coach. Ninguna información personal llegará a KNIME y sus estadísticas anónimas nunca se compartirán.

Después de la pantalla de presentación, la ventana "Workspace Launcher" requiere que ingrese la ruta del workspace

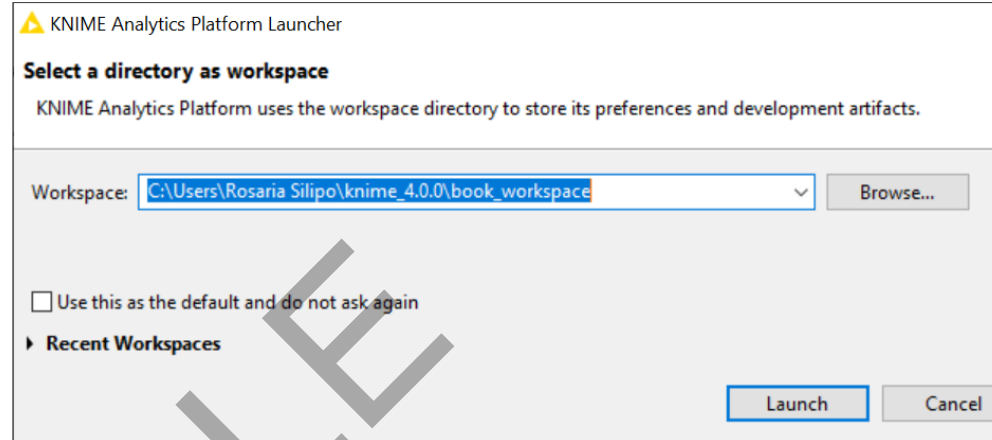
La ventana “Workspace Launcher”

El **workspace** es una carpeta donde todas las preferencias y aplicaciones (flujos de trabajo), tanto desarrolladas como actualmente en desarrollo, se guardan para la próxima sesión de KNIME.

La carpeta del espacio de trabajo se puede ubicar en cualquier lugar del disco duro.

De forma predeterminada, la carpeta del espacio de trabajo es “.. \ knime-workspace”. Sin embargo, puede cambiarlo fácilmente, cambiando la ruta propuesta en la ventana “KNIME Analytics Platform Launcher”, antes de comenzar la sesión de trabajo de KNIME.

1.5. La ventana “Workspace Launcher”



Una vez que la KNIME Analytics Platform se haya abierto, desde dentro del workbench KNIME puede cambiar a otra carpeta del espacio de trabajo, seleccionando “File” en el menú superior y luego “Switch Workspace”. Después de seleccionar el nuevo espacio de trabajo, KNIME Analytics Platform se reinicia y muestra la lista de flujo de trabajo del espacio de trabajo recién seleccionado. Tenga en cuenta que si la carpeta del espacio de trabajo no existe, se creará automáticamente.

Si se tiene una gran cantidad de clientes, por ejemplo, puedo usar un workspace diferente para cada uno de ellos. Esto mantiene mi espacio de trabajo limpio y ordenado y me protege de mezclar información por error. Para este proyecto utilicé el espacio de trabajo “KNIME_4.3.1 \ book_workspace”.

1.5. KNIME Workflow (flujo de trabajo)

KNIME Analytics Platform no funciona con scripts, funciona con flujos de trabajo gráficos.

Pequeños íconos, llamados nodos, están dedicados cada uno a implementar y ejecutar una tarea determinada. Una secuencia de nodos crea un flujo de trabajo para procesar los datos para alcanzar el resultado deseado.

¿Qué es un workflow?

Un workflow es un **flujo de análisis**, es decir, una **secuencia de pasos de análisis** necesarios para alcanzar un resultado dado. Es la tubería del proceso de análisis, algo como:

- Paso 1. Leer datos
- Paso 2. Limpiar datos
- Paso 3. Filtrar datos
- Paso 4. Entrena un modelo

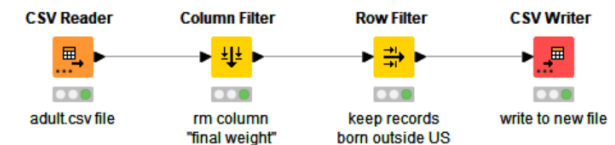
KNIME Analytics Platform implementa sus workflows de forma gráfica. Cada paso del análisis de datos se implementa y ejecuta a través de una pequeña caja, llamada nodo. Una secuencia de nodos crea un flujo de trabajo

En el KNIME whitepaper [1] un workflow is definido como: *“Los flujos de trabajo en KNIME son gráficos que conectan nodos, o más formalmente, gráficos acíclicos directos (DAG).”* (http://www.kdd2006.com/docs/KDD06_Demo_13_Knime.pdf)

Below is an example of a KNIME workflow, with:

- a node to read data from a file
- a node to exclude some data columns
- a node to filter out some data rows
- a node to write the processed data into a file

1.6. Ejemplo de un KNIME “workflow”



Nota. Un flujo de trabajo es una secuencia de análisis de datos, que en un lenguaje de programación tradicional se implementaría mediante una serie de instrucciones y llamadas a funciones. KNIME Analytics Platform lo implementa gráficamente. Esta representación gráfica es más intuitiva de usar, le permite mantener una descripción general del proceso de análisis y también sirve para la documentación.

¿Que es un nodo?

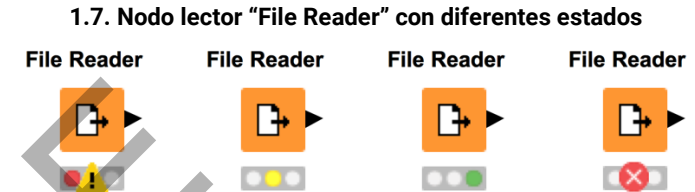
Un nodo (node) es una unidad simple de procesamiento de un workflow

Un nodo toma un conjunto de datos como entrada, lo procesa y lo pone a disposición en su puerto de salida. La acción de "procesamiento" de un nodo varía desde el modelado, como un nodo de aprendizaje de red neuronal artificial, hasta la manipulación de datos, como la transposición de la matriz de datos de entrada, desde herramientas gráficas, como un diagrama de dispersión, hasta operaciones de lectura / escritura. Cada nodo en KNIME puede tener 4 estados:

- Inactivo y no configurado → **red** light
- Configurado pero no ejecutado → **yellow** light
- Ejecutado exitosamente → **green** light
- Ejecutado con errores → **red with cross** light

Nodos que contienen otros nodos se llaman **metanodos** or **components**.

A continuación se muestran cuatro ejemplos del mismo nodo (un nodo Lector de archivos) en cada uno de los cuatro estados.



1.6. Archivos .knwf and .knar

Los flujos de trabajo de KNIME se pueden empaquetar y exportar en archivos .knwf o .knar. Un archivo .knwf contiene solo un flujo de trabajo, mientras que un archivo .knar contiene un grupo de flujos de trabajo. Dichas extensiones están asociadas con KNIME Analytics Platform. Un doble clic abre el flujo de trabajo dentro de KNIME Analytics Platform.

1.8. Los archivos .knwf and .knar files estan asociados a la plataforma KNIME Analytics

▲ 01_From_Strings_to_Documents.knwf	10/4/2017 9:45 AM	KNIME Workflow ...	18,619 KB
▲ 04_Interaction_Graph.knwf	9/29/2017 8:20 AM	KNIME Workflow ...	9,465 KB
▲ 06_REST_Examples_Google_Geocode.knwf	7/29/2017 7:09 PM	KNIME Workflow ...	62 KB
▲ 06_Semantic_Web_updated.knar	11/3/2016 2:24 PM	KNIME Archive File	178 KB
▲ AzureDemoWorkflowArchive.knar	5/5/2017 11:24 AM	KNIME Archive File	24,104 KB
▲ Building a Simple Classifier_.knwf	2/18/2017 5:46 PM	KNIME Workflow ...	43 KB
▲ Cookbook_Ch5.knar	11/24/2017 10:03 ...	KNIME Archive File	477 KB
▲ Cookbook_Ch6.knar	11/24/2017 10:26 ...	KNIME Archive File	155 KB
▲ Corsair.knwf	7/10/2017 4:20 PM	KNIME Workflow ...	106 KB

1.7. KNIME workbench (banco de trabajo)

Después de aceptar la ruta del espacio de trabajo, el KNIME workbench se abre en una página de "Bienvenido a KNIME". Esta página proporciona algunos enlaces para comenzar, como por ejemplo al KNIME Hub, a cierta documentación básica, a los cursos y eventos actuales, a las actualizaciones disponibles, etc. El "KNIME Workbench" consta de un menú superior, una barra de herramientas y algunos paneles. Los paneles se pueden cerrar, volver a abrir y mover.

1.9. EL KNIME workbench

The screenshot displays the KNIME Analytics Platform interface. The main workspace shows a workflow with four nodes: CSV Reader (input: adult.csv file), Column Filter (action: rm column "final weight"), Row Filter (action: keep records born outside US), and CSV Writer (action: write to new file). The interface includes several panels:

- KNIME Explorer:** A tree view on the left showing project structure, including folders like 'AdvancedLuck', 'BeginnersLuck', and 'Chapter2', and sub-folders like 'Exercises' and '1. My First Workflow(2)'. A yellow box labeled 'Server & Workflows' highlights this area.
- Workflow Coach:** A panel below the Explorer showing recommended nodes. A yellow box labeled 'Workflow Coach' highlights it.
- Node Repository:** A panel at the bottom left listing various nodes like 'DB Reader', 'LDB Reader', 'PDB Reader', etc. A yellow box labeled 'Node Repository' highlights it.
- Workflow Editor:** The central workspace where the workflow is built. A yellow box labeled 'Workflow Editor' highlights the main area.
- Overview:** A small blue box at the bottom left of the workflow editor. A yellow box labeled 'Overview' highlights it.
- Console:** A panel at the bottom right showing system logs and warnings. A yellow box labeled 'Console' highlights it.
- Description:** A panel on the right showing the documentation for the selected 'CSV Reader' node. A yellow box labeled 'Workflow & Node Description' highlights it.
- Search on KNIME Hub:** A search bar at the bottom of the description panel. A yellow box labeled 'Search on KNIME Hub' highlights it.

Watermark: SAMPLE

El KNIME Workbench

Top Menu: File, Edit, View, Node, Help

Tool Bar: New, Save (Save As, Save All), Undo/Redo, Open Report (if reporting was installed), zoom (in %), Align selected nodes vertically/horizontally, Auto layout, Configure, Execute options, Cancel execution options, Reset, Edit node name and description, Open node's first out port table, Open node's first view, Open the "Add Meta node" Wizard, Append IDs to node names, Hide all node names, Loop execution options, Change Workflow Editor Settings, Edit Layout in Components, configure job manager.

KNIME Explorer

Este panel muestra la lista de proyectos de workflows disponibles en el espacio de trabajo seleccionado (LOCAL), en el servidor de EJEMPLOS, en My-KNIME-Hub (su propio espacio en KNIME Hub) o en otros servidores KNIME conectados.

Workflow Coach

Este es un motor de recomendación de nodos. Proporcionará la lista de los principales nodos con mayor probabilidad de seguir al nodo seleccionado actualmente.

Workflow Editor

El área central consta del propio "Workflow Editor".

Se puede seleccionar un nodo del panel "Node Repository" "Repositorio de nodos" y arrastrarlo y soltarlo aquí, en el panel "Workflow Editor"

Los nodos se pueden conectar haciendo clic en el puerto de salida de un nodo y soltando el mouse en el puerto de entrada del siguiente nodo o en el siguiente nodo.

Node Description

Si se selecciona un nodo o un flujo de trabajo, este panel muestra una descripción resumida de las funcionalidades del nodo o la meta información del flujo de trabajo.

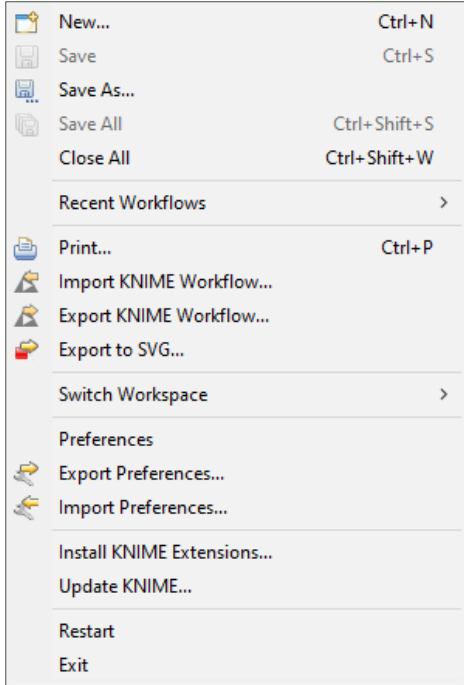
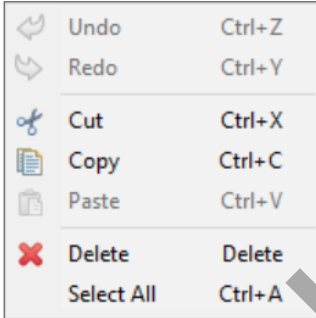
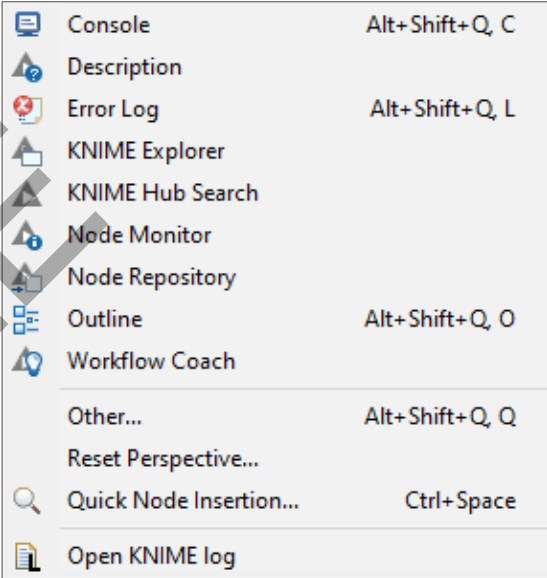
Search box for KNIME Hub

Para buscar material en KNIME Hub

Node Repository	Outline	Console
<p>Este panel contiene todos los nodos que están disponibles en su instalación KNIME. Es algo similar a una paleta de herramientas cuando se trabaja en un informe o con un software de diseño web. Allí usamos herramientas gráficas, mientras que en KNIME usamos herramientas de análisis de datos.</p>	<p>El panel "Outline" contiene una pequeña descripción general del contenido del "Workflow Editor". Es posible que el panel "Outline" no sea de gran interés para flujos de trabajo pequeños. Sin embargo, tan pronto como los flujos de trabajo alcancen un tamaño considerable, es posible que todos los nodos del flujo de trabajo ya no estén visibles en el "Workflow Editor" sin desplazarse. El panel "Outline", por ejemplo, puede ayudarlo a ubicar los nodos recién creados.</p>	<p>El panel "Console" muestra errores y mensajes de advertencia (warnings) al usuario.</p> <p>Este panel también muestra la ubicación del archivo de registro, que puede ser de interés cuando la consola no muestra todos los mensajes.</p> <p>También hay un botón en la barra de herramientas para mostrar el archivo de registro asociado con esta instancia de KNIME.</p>

SAMPLE

Menu Principal

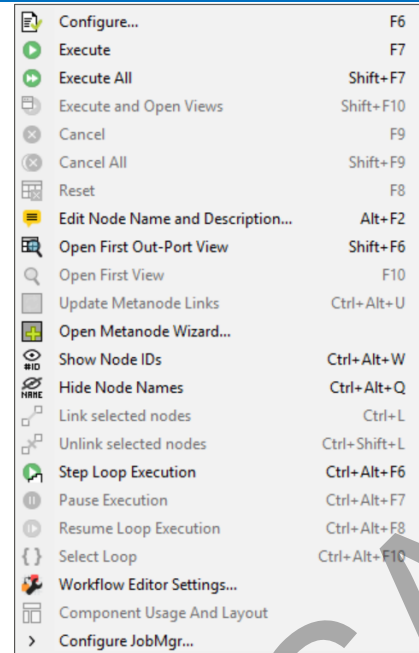
File (archivo)	Edit (editar)	View (Ver)
		
<p>File incluye los comandos de archivo tradicionales, como "Nuevo" y "Guardar", además de algunos comandos específicos de KNIME, como:</p> <ul style="list-style-type: none"> - Importar / Exportar flujo de trabajo KNIME ... - Exportar a SVG - Cambiar espacio de trabajo - Preferencias con preferencias de exportación / importación - Instalar extensiones KNIME 	<p>Edit contiene comandos de edición.</p> <p>Deshacer y Rehacer se refieren a las últimas acciones realizadas.</p> <p>Cortar, Copiar, Pegar y Eliminar hacen referencia a los nodos seleccionados en el flujo de trabajo.</p>	<p>View contiene la lista de todos los paneles que se pueden abrir en el banco de trabajo KNIME.</p> <p>Aquí se puede volver a abrir un panel cerrado.</p> <p>Además, cuando se estropea la disposición del panel, la opción "Restablecer perspectiva" recrea el diseño del panel original cuando se inició el banco de trabajo por primera vez.</p>

- Actualizar KNIME

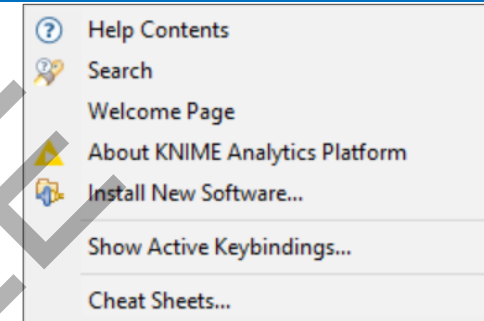
Seleccionar todo selecciona todos los nodos del flujo de trabajo en el editor de flujo de trabajo.

La opción "Otro" abre vistas adicionales útiles para personalizar el banco de trabajo.

Node



Help



Node se refiere a todas las operaciones posibles que se pueden realizar en un nodo. Un nodo puede estar:

- Configured (Configurado)
- Executed (Ejecutado)
- Cancelled (Cancelado -detenido durante la ejecución)
- Reset (restablece los resultados de la última operación "Ejecutar")
- Given a name and description
- Set to show its View (if any)

Help Contents proporciona ayuda general sobre Workbench, BIRT y KNIME.

Search abre un panel a la derecha del panel "Descripción de nodo" para buscar temas de ayuda o nodos específicos.

Welcome Page (re abre) la Welcome Page

Install New Software es la puerta para instalar KNIME Extensions desde KNIME Update sites.

- Reset (restablece los resultados de la última operación "Ejecutar")

- Dado un nombre y una descripción

- Establecer para mostrar su vista (si corresponde)

Las opciones solo están activas si son posibles. Por ejemplo, un nodo ya ejecutado con éxito no se puede volver a ejecutar a menos que se restablezca primero o se haya cambiado su configuración. Entonces la opciones "Cancel" and "Execute" están inactivas.

La opción "Open Meta Node Wizard" inicia el asistente para crear un nuevo metanodo en el editor de flujo de trabajo.

Show Active Keybindings resume todos los comandos del teclado para el editor de flujo de trabajo.

Cheat Sheets ofrece tutoriales sobre temas específicos: la herramienta de informes, csv, Plug-ins.

SAMPLE

Veamos ahora los elementos que se utilizan con más frecuencia en Top Menu.

"File" → "Import KNIME workflow" reads and copies workflows into the current workspace.

Opción "Select root directory", copia el workflow directamente del folder el workspace actual (LOCAL).

Opción "Select archive file", lee un workflow desde una archivo .knwf or .knar en el Los archivos .knwf /.knar pueden ser creados con la opcion "File" → "Export KNIME workflow".

"File" → "Export KNIME workflow" exporta el workflow seleccionado a un archivo .knar

Opción "Reset Workflow(s) before export" exporta los workflows sin la data producida por la ejecucion de los nodos. Esto genera archivos exportables significativamente más pequeños.

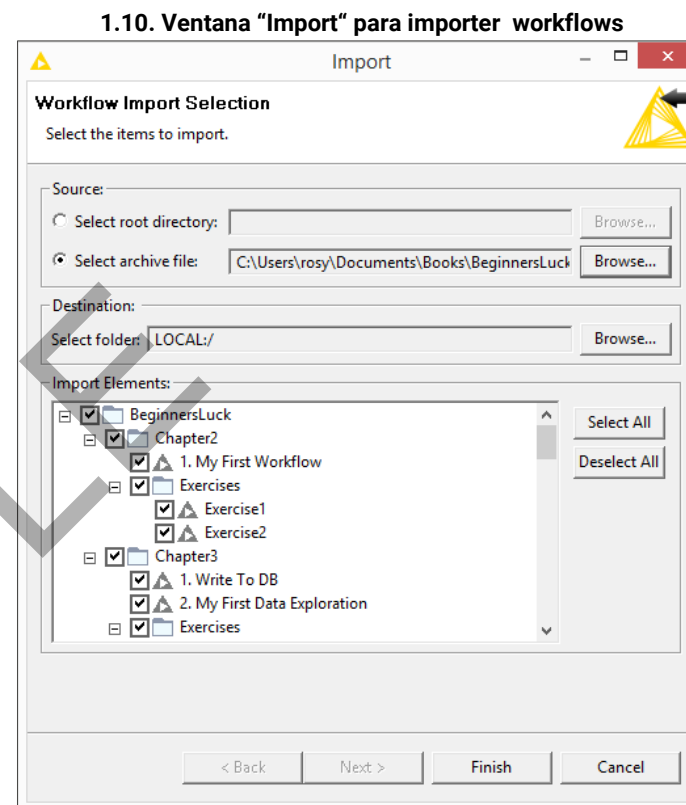
Copiar workflows de una carpeta a otra puede crear ciertos problemas relacionados con las actualizaciones internas de KNIME. Copiar workflows usando la opción "Import KNIME workflow" o mediante double-click es definitivamente más seguro.

"File" → "Install KNIME Extensions" y **"Help" → "Install New Software"** ambos abre un asistente para la instalacion de extensiones de KNIME (ver proximas secciones)

"File" → "Switch Workspace" cambia el workspace actual por uno nuevo.

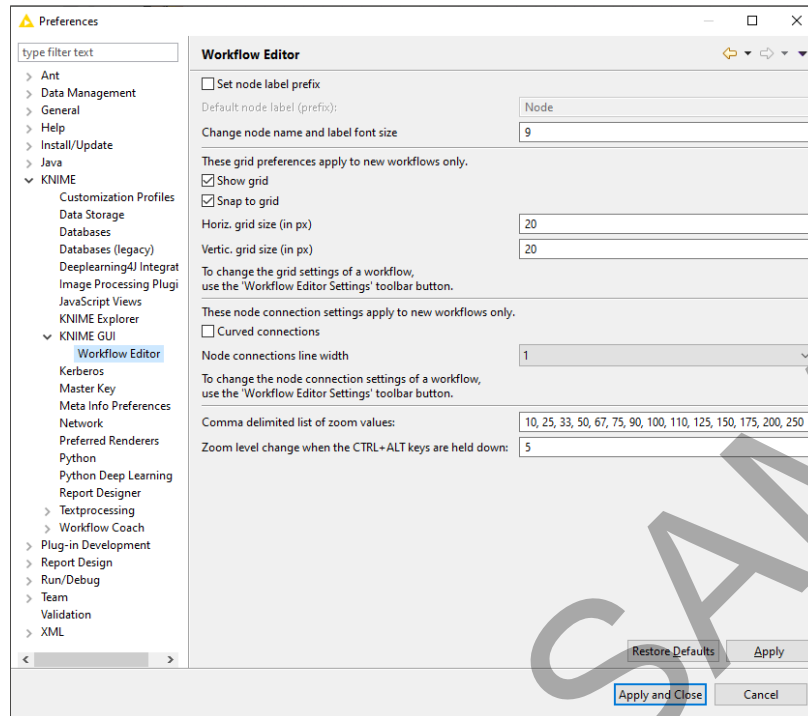
"File" → "Preferences". Abre una ventana donde se pueden personalizar todos los ajustes de KNIME. Se pueden encontrar en el elemento "KNIME". Vamos a comprobarlos.

- **Chemistry** tiene configuraciones relacionadas con los renderizadores KNIME en los módulos de química.



- **Databases** especifica la ubicación de controladores de bases de datos específicos, que aún no están disponibles en KNIME. De hecho, los controladores de bases de datos más comunes y más recientes ya están disponibles en el menú de controladores de los nodos de la base de datos. Sin embargo, si necesita algún archivo de controlador específico, puede establecer su ruta aquí.

1.11. La ventana "Preferences"



- **KNIME Explorer** contiene la lista de los repositorios compartidos a través de KNIME Server.
- **KNIME GUI** permite la personalización de las opciones y el diseño del banco de trabajo KNIME a través de una serie de configuraciones.
- **Master Key** contiene la clave maestra que se utilizará en los nodos con una opción de cifrado, como los nodos de conexión de la base de datos. Dado que las contraseñas de la base de datos KNIME version 2.3 se pasan a través de las variables de flujo de trabajo "Credentials" y la preferencia de clave maestra ha quedado obsoleta. Aún puede encontrarlo en el menú Preferencias para compatibilidad con versiones anteriores.
- In **Meta Info Preferences** puede cargar una plantilla de metainformación para nodos y flujos de trabajo.
- Aquí puede encontrar la configuración de preferencias para los módulos externos, como: H2O, R, Report Designer, Perl, Perl, Open Street Map y otros si los tiene instalados. En particular, para los scripts externos, esta página ofrece la opción de establecer la ruta a la instalación del script de referencia.
- Finalmente, **Workflow Coach** contiene el conjunto de datos que se utilizará para el motor de recomendación de nodos: la comunidad, un espacio de trabajo del servidor o su propio espacio de trabajo local.

Para todos los amantes del teclado, la mayoría de los comandos de KNIME también se pueden ejecutar a través de teclas de acceso rápido. Todas las teclas de acceso rápido se enumeran en los menús de KNIME al lado de los comandos correspondientes o en los mensajes de información sobre herramientas de los iconos de la barra de herramientas en el menú superior. Estas son las teclas de acceso rápido que se utilizan con más frecuencia.

Hotkeys (teclas importantes)

Configuración de nodos

- **F6** abre la ventana de configuración del nodo seleccionado

Ejecución de nodos

- **F7** ejecuta los nodos configurados seleccionados
- **Shift + F7** ejecuta todos los nodos configurados
- **Shift + F10** ejecuta todos los nodos configurados y abre todas las vistas

Para detener la ejecución de nodos

- **F9** cancela los nodos en ejecución seleccionados
- **Shift + F9** cancela todos los nodos en ejecución seleccionados

Para mover nodos

- **Ctrl + Shift + Arrow** mueve el nodo seleccionado en la dirección de la flecha

Reseteo de Nodos

- **F8** restablece los nodos seleccionados

Grabar Workflows

- **Ctrl + S** graba el workflow
- **Ctrl + Shift + S** graba los workflows abiertos
- **Ctrl + Shift + W** cierra todas las ventanas de workflows

Meta-Node

- **Shift + F12** abre el asistente de metanodo

To move Annotations

- **Ctrl + Shift + PgUp/PgDown** mueve la anotación seleccionada al frente o al final de todas las anotaciones superpuestas

Repositorio de Nodos (Node Repository)

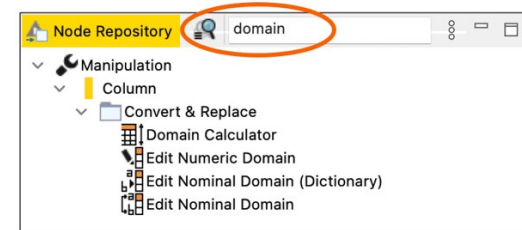
En la esquina inferior izquierda encontramos el Repositorio de nodos, que contiene todos los nodos instalados organizados en categorías y subcategorías. KNIME Analytics Platform ha acumulado hasta ahora más de 1500 nodos. Se ha vuelto difícil recordar la ubicación de cada nodo en el repositorio de nodos. Para resolver este problema, hay dos opciones de búsqueda disponibles: por coincidencia exacta y por coincidencia aproximada, ambas en el cuadro de búsqueda ubicado en la parte superior del panel del repositorio de nodos.

Cuadro de búsqueda

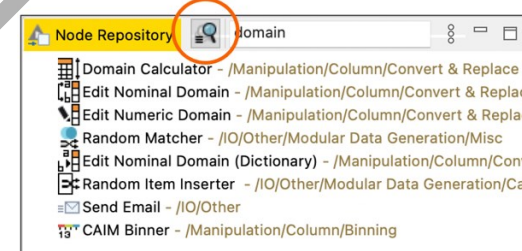
En la parte superior del panel "Repositorio de nodo" hay un cuadro de búsqueda. Si escribes una palabra clave en el cuadro de búsqueda y presiona "Enter", obtendrá la lista de nodos que contienen una coincidencia exacta de esa palabra clave. Presione la tecla "Esc" para ver todos los nodos nuevamente.

Al hacer clic en la lente a la izquierda del cuadro de búsqueda, se ejecuta un algoritmo de búsqueda difusa (fuzzy) que conduce a una lista de resultados coincidentes más amplia que la que se encuentra en la figura anterior.

1.13. Búsqueda de palabras en el panel del repositorio de nodos: modo de coincidencia exacta



1.14. Búsqueda de palabras en el panel del repositorio de nodos: modo de búsqueda difusa (fuzzy)



KNIME Explorer (Explorador de KNIME)

En la esquina superior izquierda del banco de trabajo KNIME, encontramos el panel KNIME Explorer. Este panel contiene:

- En LOCAL los workflows que se han desarrollado en el espacio de trabajo seleccionado
- Los puntos de montaje a varios servidores KNIME
- Los workflows contenidos en el espacio de trabajo de referencia de dichos servidores.
- El acceso a My-KNIME-Hub, es decir, a su espacio en KNIME Hub. Recuerda que necesitas una cuenta en KNIME Forum para acceder a este espacio

Al principio, el panel KNIME Explorer solo contiene LOCAL, My-KNIME-Hub y EXAMPLES. Como ya dijimos, LOCAL muestra el contenido del espacio de trabajo seleccionado. EXAMPLES apunta a un servidor público de solo lectura, accesible mediante inicio de sesión anónimo.

Este servidor aloja una serie de flujos de trabajo de ejemplo que puede utilizar para iniciar un nuevo proyecto. My-KNIME-Hub permite acceder a su espacio en KNIME Hub.

Cuando abra KNIME Analytics Platform por primera vez, encontrará una carpeta denominada "Example Workflows" que contiene las soluciones para algunos casos de uso de ciencia de datos comunes.

Las carpetas en "KNIME Explorer", que contienen workflows, también se denominan "Workflow Groups".

Nota. El panel KNIME Explorer también puede albergar datos. Simplemente cree una carpeta debajo de la carpeta del área de trabajo, llénela con archivos de datos en la máquina y seleccione "Refresh" en el menú contextual (clic derecho) del panel "KNIME Explorer".

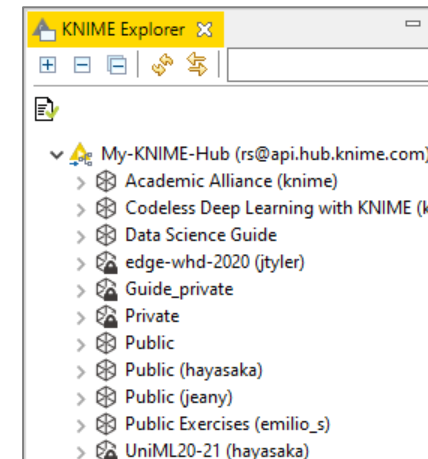
Mi central KNIME (My-KNIME-Hub)

Desde el panel KNIME Explorer, puede acceder a sus espacios en KNIME Hub y cargar y actualizar contenido nuevo o existente allí.

De forma predeterminada, un usuario KNIME autenticado tiene un espacio público, para que el material se comparta públicamente, y un espacio privado para estacionar su propio material para su uso posterior. Sin embargo, se pueden crear nuevos espacios públicos o privados con un clic derecho en My-KNIME-Hub en el panel KNIME Explorer y luego una selección de la opción "Crear nuevo espacio...".

De forma predeterminada, eres el único propietario de tus propios espacios. Sin embargo, al acceder a este espacio desde un navegador web, después de colocar el cursor sobre su imagen en la esquina superior derecha, aparece un bolígrafo. Esto permitirá agregar colegas y compañeros de equipo como contribuyentes al espacio.

1.15. Panel KNIME Explorer. En la parte superior, el contenido del servidor de EXAMPLES; debajo del contenido del espacio de trabajo LOCAL



El "EXAMPLES Server"

Un enlace al servidor público KNIME (EXAMPLES) está disponible en el panel "KNIME Explorer". Este es un servidor proporcionado por KNIME a todos los usuarios para tutoriales y demostraciones. Allí puede encontrar una serie de ejemplos útiles sobre cómo implementar tareas específicas con KNIME. Para conectarse al servidor de EJEMPLOS:

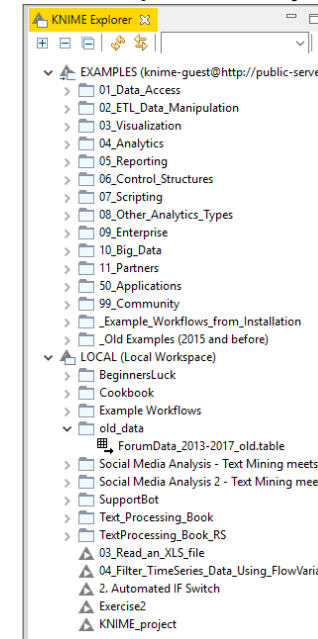
- haga doble clic en "EJEMPLOS" en el panel "KNIME Explorer"
- haga doble clic en "Haga doble clic para conectarse ..."

Debería iniciar sesión automáticamente como invitado.

Para transferir ejemplos de flujos de trabajo del servidor de EXAMPLES a su espacio de trabajo LOCAL, simplemente arrastre y suelte o copie y pegue (Ctrl-C, Ctrl-V en Windows) de "EXAMPLES" a "LOCAL".

También puede abrir los workflows de EXAMPLES en el editor de flujo de trabajo, pero solo temporalmente y en modo de solo lectura. Un cuadro de advertencia amarillo en la parte superior advierte que esta copia del flujo de trabajo no se guardará.

1.16. Panel KNIME Explorer. En la parte superior, el contenido del servidor de EXAMPLES; debajo del contenido del espacio de trabajo LOCAL



El panel KNIME Explorer puede albergar más de un servidor KNIME. Es suficiente agregar puntos de montaje del servidor a la lista de servidores KNIME disponibles.

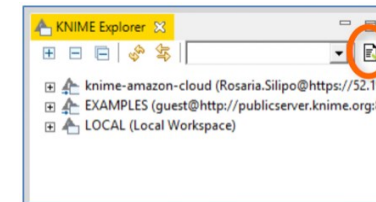
Montaje de servidores en KNIME Explorer

Para agregar el KNIME servers al panel "KNIME Explorer" :

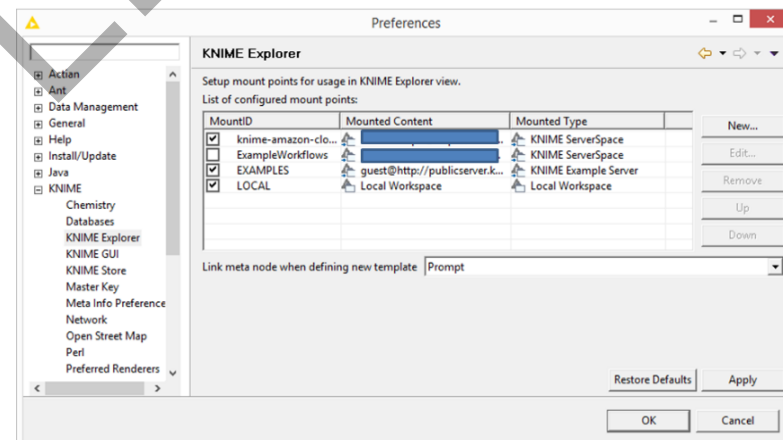
- Seleccione el panel "KNIME Explorer"
- Haga Click en "Configure Explorer View"
- La ventana The "Preferences (Filtered)" se abre en la página "KNIME Explorer" y enumera todos los espacios KNIME ya montados en esta instancia KNIME. Los tres espacios KNIME disponibles de forma predeterminada en cada instancia KNIME son el espacio de trabajo local "LOCAL", el servidor público KNIME "EXAMPLES" y el My-KNIME-Hub ubicado en el servidor KNIME Hub (hub.knime.com).
- Utilice el botón "New" y "Remove" para agregar / eliminar conexiones a servidores remotos.
- Después de hacer clic en el botón "New", complete la información requerida sobre el servidor en la ventana "Select New Content"
- Utilice el botón "Test Connection" para recuperar automáticamente el punto de montaje predeterminado para el servidor seleccionado.

El la pagina KNIME Explorer "Preferences" puede ser accedida mediante "File" → "Preferences" → "KNIME Explorer".

1.17. El boton "Configure KNIME Explorer"



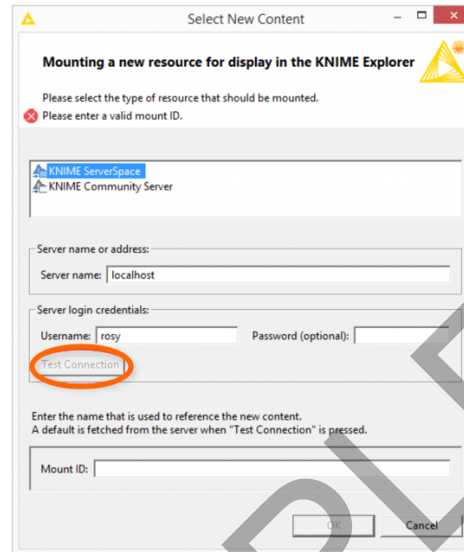
1.18. La ventana "Preferences (Filtered)"



Para iniciar sesión en cualquiera de los servidores disponibles en el panel "KNIME Explorer":

- haga clic con el botón derecho o haga doble clic en el nombre del servidor
- proporcionar las credenciales

1.19. La ventana "Select New Content"



El Editor de Workflow

La pieza central del banco de trabajo KNIME consiste en el propio editor de flujo de trabajo. Este es el lugar donde se construye un flujo de trabajo agregando un nodo tras otro. Los nodos se insertan en el editor de flujo de trabajo arrastrando y soltando o haciendo doble clic desde el repositorio de nodos o el entrenador de flujo de trabajo. El proceso de construcción del flujo de trabajo se describirá ampliamente en las próximas secciones de este libro. Aquí, describiremos cómo personalizar y probablemente mejorar la función del canvas del espacio del editor de flujo de trabajo.

Describiremos dos opciones:

- cambiar la apariencia del canvas con cuadrículas y diferentes visualizaciones para las conexiones;
- introducción de anotaciones para comentar el trabajo.

Agregar una cuadrícula al canvas y conexiones curvas a los workflows

Casi hacia el final, a la derecha de la barra de herramientas, puede ver el botón "Change Workflow Editor Settings" Si hace clic en él, se abre la ventana "Workflow Editor Settings".

1.19. Boton "Change Workflow Editor Settings" en la barra de herramientas

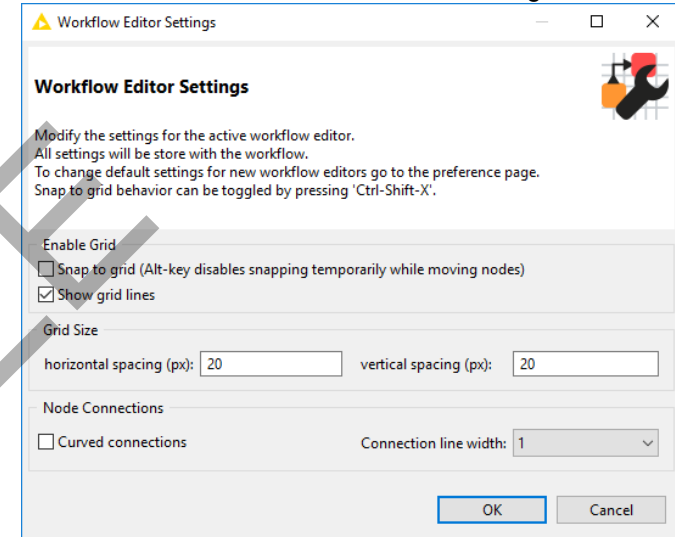


Personalización del editor de flujo de trabajo

La función de cuadrícula contiene algunas opciones:

1. "Show grid lines" Esto muestra líneas de cuadrícula en el editor de flujo de trabajo y permite alinear mejor los nodos y anotaciones manualmente.
2. "Snap to grid". Esta opción adjunta nodos y anotaciones a la esquina disponible más cercana a la cuadrícula. Le brinda menos libertad manual, pero el resultado es más limpio y ordenado en menos tiempo.
3. "Curved Connections". Aquí puede permitir que las conexiones de nodo sigan una curva en lugar de una línea recta. Esto podría dar lugar a gráficos de flujo de trabajo más atractivos.

1.20. La ventana "Workflow Editor Settings"



Agregar anotaciones al canvas

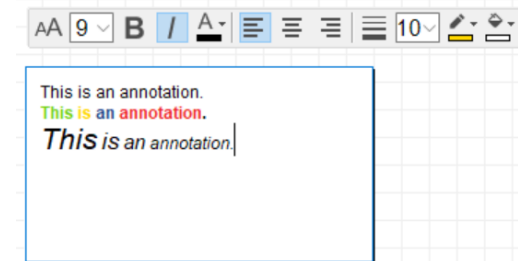
También es posible incluir anotaciones en el editor de flujo de trabajo. Las anotaciones pueden ayudar a explicar la tarea workflow y la función de cada nodo o grupo de nodos. El resultado es una descripción general mejorada, similar a la documentación, de la tarea general del workflow y de las subtarefas individuales.

Anotaciones del Workflow

Para insertar una nueva anotación:

- haga clic con el botón derecho en cualquier lugar del editor de workflow y “New Workflow Annotation”
- aparece un pequeño marco de color amarillo pálido; este es el marco de anotaciones predeterminado
- haga doble clic en el marco para editar su contenido
- Observe la barra de herramientas que aparece en la parte superior para editar el estilo del texto, el color de la fuente, el color de fondo, la alineación del texto y las propiedades del borde (color, grosor).
- Para volver a abrir una anotación, simplemente haga doble clic en la esquina superior izquierda, donde está el icono de lápiz.

1.21. El editor Annotations



Otras personalizaciones del Workbench

Otra posibilidad de personalización consiste en agregar vistas. Las vistas disponibles se encuentran en el elemento "View" del menú superior.

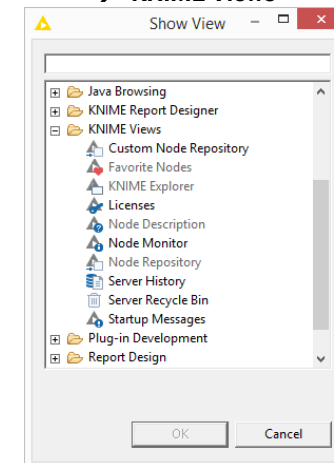
Algunas vistas populares son las vistas "Node Monitor", "Custom Node Repository", y "Licenses" y "Server" si ud esta conectado al servidor. Estas puede invocarse siguiendo el esquema: Top menu under "View" → "Other" → "KNIME Views".

"Node Monitor" ayuda, especialmente durante la fase de desarrollo, a monitorear y depurar la ejecución del flujo de trabajo.

"Custom Node Repository" permite una personalización del "Node Repository" con solo un conjunto de nodos.

"Licenses" permite monitorear la situación de sus licencias .

1.22. Vistas adicionales "View" → "Other" → "KNIME Views"

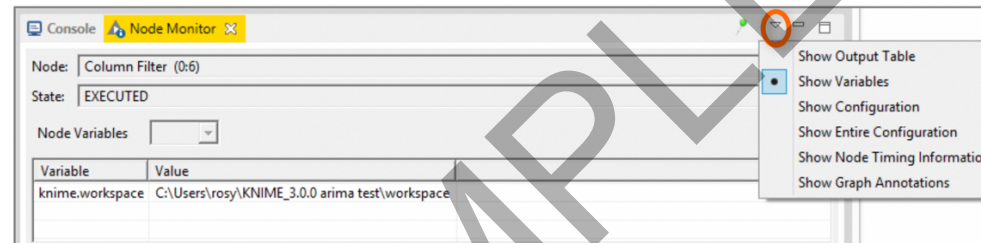


Vista “Node Monitor”

Para insertar la vista “Node Monitor” en el workbench:

- Seleccione “View” → “Other...” en el menu principal
- Expanda el elemento “KNIME Views” y haga doble clic en “Node Monitor” un panel, llamado “Node Monitor”, aparece en el lateral del panel “Console”; el panel muestra los valores de las variables de flujo de salida, los datos de salida o los valores de configuración del nodo seleccionado en el editor de flujo de trabajo.
- Aquí puede decidir qué mostrar (datos, configuración, variables), a través del menú en la esquina superior derecha.

1.23. La vista Node Monitor



1.8. Descargar las Extensiones de KNIME

KNIME Analytics Platform es un producto de código abierto. Como todo producto de código abierto, se beneficia de los comentarios y las funcionalidades que desarrolla la comunidad. Hay varias extensiones disponibles para KNIME Analytics Platform. Si ha descargado e instalado KNIME Analytics Platform, incluidas todas sus extensiones gratuitas, verá las categorías correspondientes en el panel Node Repository, como KNIME Labs, Text Processing, R Integration y muchas otras. Sin embargo, si en el momento de la instalación ha elegido instalar la plataforma KNIME Analytics sin las extensiones gratuitas, es posible que deba instalarlas por separado en algún momento de una instancia en ejecución.

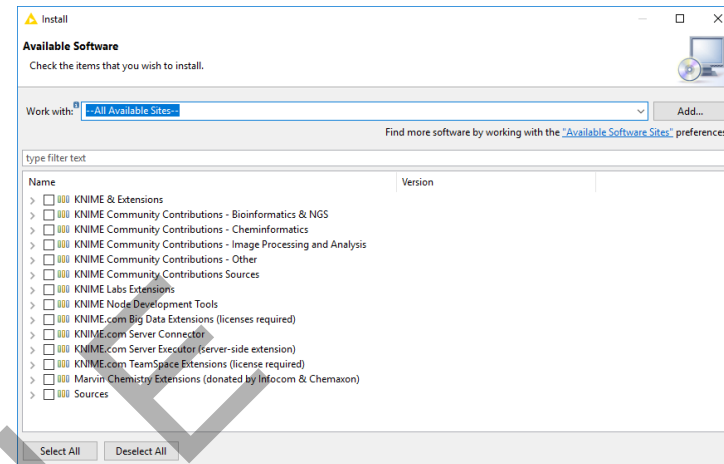
Instación de KNIME Extensions

Para instalar una nueva extensión KNIME desde KNIME Analytics Platform, hay tres opciones.

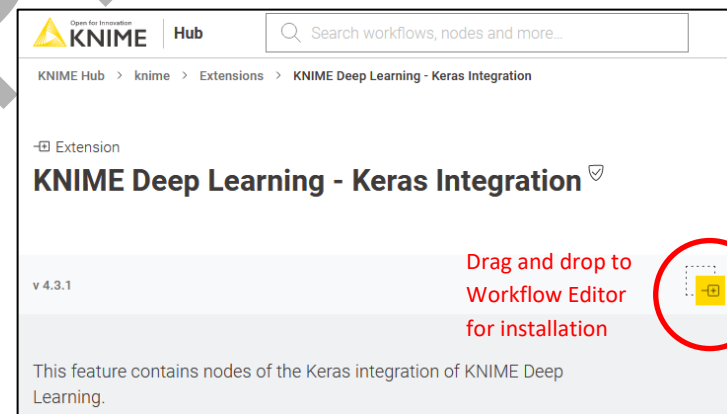
1. Desde el Top Menu, seleccione **"File" → "Install KNIME Extensions"**, seleccione la extensión deseada y haga click en el boton **"Next"**. Siga las instrucciones del asistente
2. Desde el Top Menu, seleccione **"Help" → "Install New Software"**. En la ventana "Available Software", en el cuadro de texto "Work with", seleccione el URL de actualización de KNIME (usualmente llamado "KNIME Analytics Platform 4.x Update Site" - <http://update.knime.com/analytics-platform/4.x>). Entonces seleccione la extensión, haga click the **"Next"** button y siga las instrucciones del asistente
3. Busque en el KNIME Hub, en un navegador web o desde el panel KNIME Hub. Cuando encuentre la extensión deseada, arrastre y suelte el icono de la extensión desde el navegador al Workflow Editor.

Una vez que se hayan instalado las extensiones KNIME seleccionadas y se haya reiniciado KNIME, debería ver la nueva categoría, correspondiente a la extensión instalada, en el "Node Repository".

1.24. La ventana "Available Software"



1.25. Una extensión del KNIME Hub



En la ventana "Software disponible" puede encontrar algunos grupos de extensiones: KNIME & Extensions, KNIME Labs Extensions, KNIME Node Development Tools, Sources y más. "KNIME & Extensions" contiene todas las extensiones proporcionadas para la versión actual; "KNIME Labs Extensions" contiene una serie de extensiones listas para usar, pero que aún no tienen la calidad de lanzamiento x.1; "KNIME Node Development Tools" contiene paquetes con algunas herramientas útiles para que los programadores de Java desarrollen nodos; "Sources" contiene el código fuente de KNIME. Los paquetes específicos donados por terceros o entidades comunitarias también pueden estar disponibles en la lista de extensiones. Por lo general, se agrupan en categorías de "Community". Mi consejo es instalar todas las

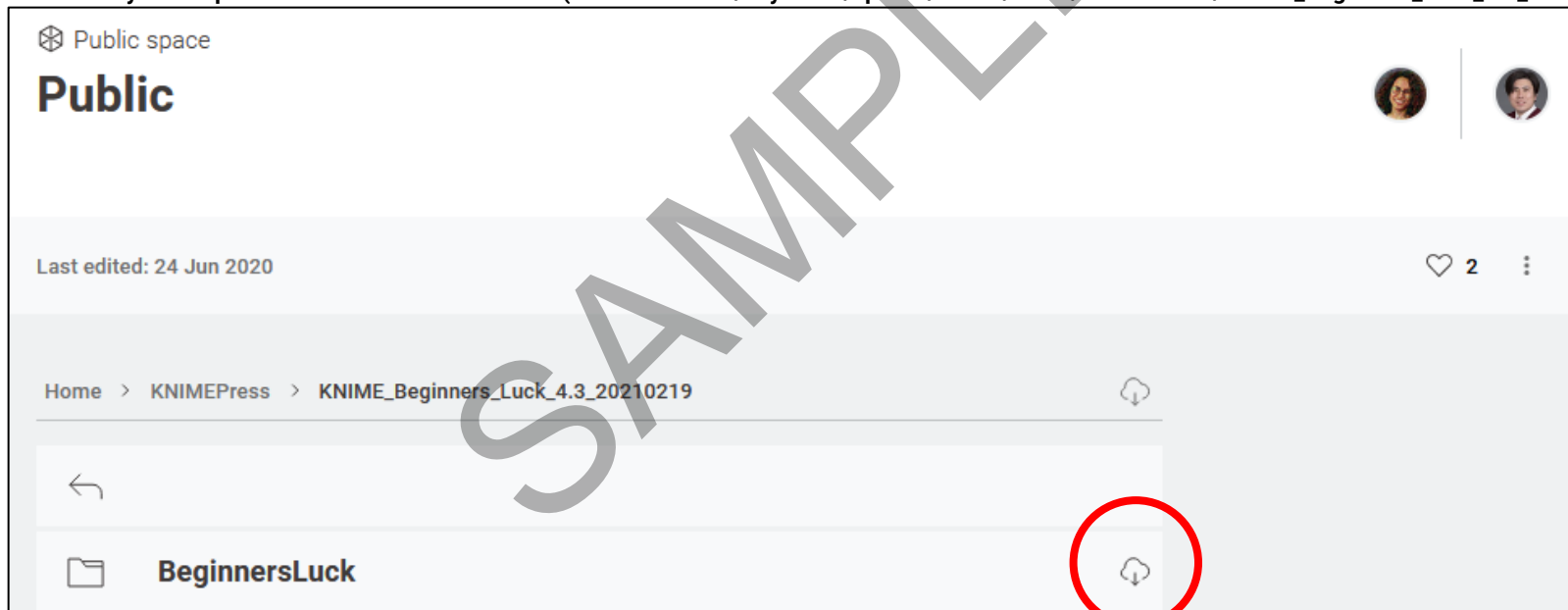
extensiones, incluso las de cheminformatics. Muchos de ellos contienen varios nodos útiles que no están necesariamente restringidos a un dominio en particular.

1.9. Datos y workflows para éste libro

Este libro construye algunos ejemplos y proporciona las soluciones a los ejercicios. Están contenidos en carpeta “BeginnersLuck” descargable desde el [espacio KNIME Hub de uno de los autores](#) de este libro. Para acceder al KNIME Hub, necesitas crear una cuenta con el [KNIME Forum](#). Después de ingresar al KNIME Hub, para descargar los flujos de trabajo, simplemente haga clic en el ícono de la nube.

- Descargue la carpeta completa en su máquina, lo que resultará en un archivo .knar. Luego:
- Haga doble click o importelo al KNIME Explorer haciendo: Select File → Import KNIME Workflow ...

1.26. Workflows y datos para este libro en el KNIME Hub (hub.knime.com/hayasaka/spaces/Public/latest/KNIMEPress/KNIME_Beginners_Luck_4.4_20210802)



Al final de la operación de importación, en el panel KNIME Explorer debería encontrar una carpeta BeginnersLuck que contiene las subcarpetas Chapter2, Chapter3, Chapter4, Chapter5 y Chapter6, cada una con flujos de trabajo y ejercicios que se implementarán en los siguientes capítulos. También debe encontrar una carpeta KBLdata que contenga los datos requeridos

Los datos utilizados para los ejercicios y para los flujos de trabajo demostrativos de este libro fueron generados por el autor o descargados del Repositorio de Aprendizaje Automático de la UCI, un repositorio público de datos. (<http://archive.ics.uci.edu/ml/datasets>). Si el conjunto de datos pertenece al repositorio de la UCI, aquí se proporciona un enlace completo para descargar. Los datos generados por el autor, que no son datos públicos, se encuentran en la carpeta KBLData.

Datos del repositorio de aprendizaje automático de la UCI:

- Adult.data: <http://archive.ics.uci.edu/ml/datasets/Adult>
- Iris data: <http://archive.ics.uci.edu/ml/datasets/Iris>
- Yellow-small.data (Balloons) <http://archive.ics.uci.edu/ml/datasets/Balloons>
- Wine data: <http://archive.ics.uci.edu/ml/datasets/Wine>

1.10. Ejercicios

Ejercicio 1

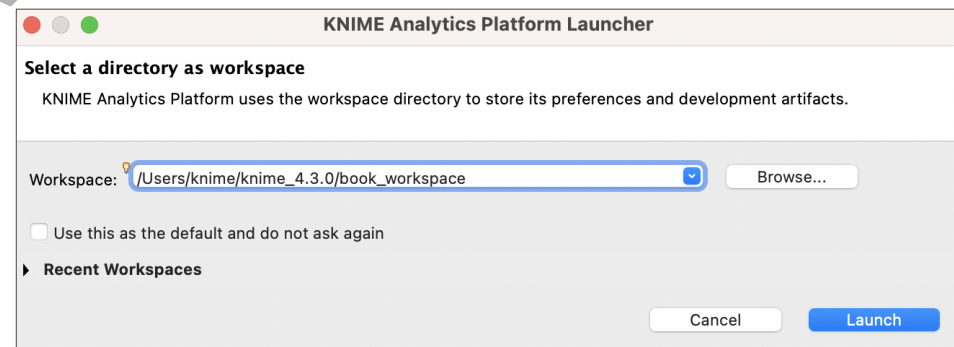
Cree su propio espacio de trabajo y asígnele el nombre "book_workspace". Utilizará este espacio de trabajo para los próximos ejercicios y flujos de trabajo.

Solución al ejercicio 1

- Lanzar KNIME
- En la ventana del Workspace Launcher, haga clic en "Browse"
- Seleccione la ruta para su nuevo workspace
- Haga clic en "OK"

Para mantener esto como su workspace predeterminado, habilite la opción en la esquina inferior izquierda.

1.27. Ejercicio 1: Crear el workspace "book_workspace"



Ejercicio 2

Instale las siguientes extensiones:

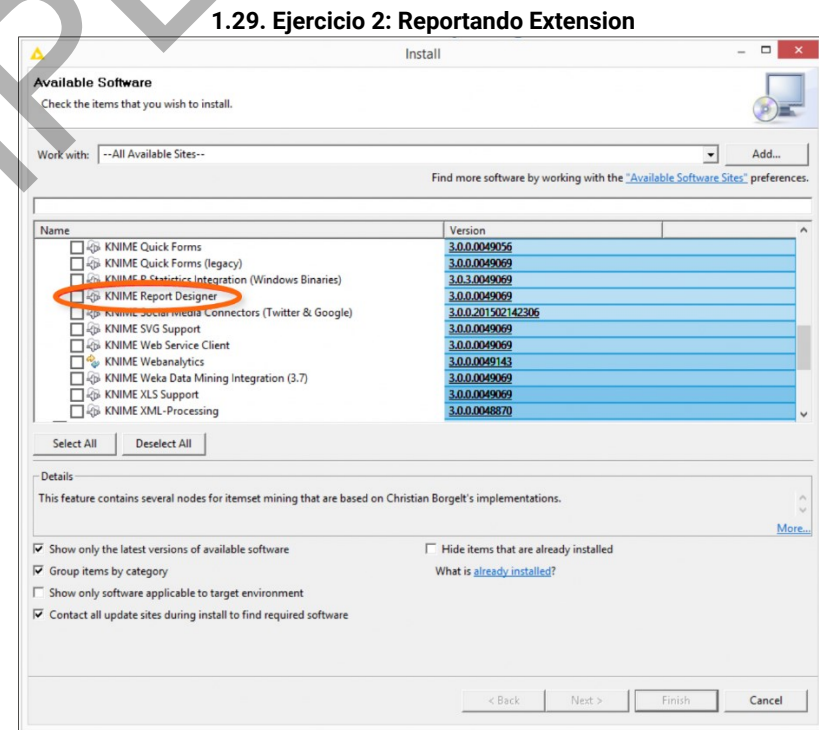
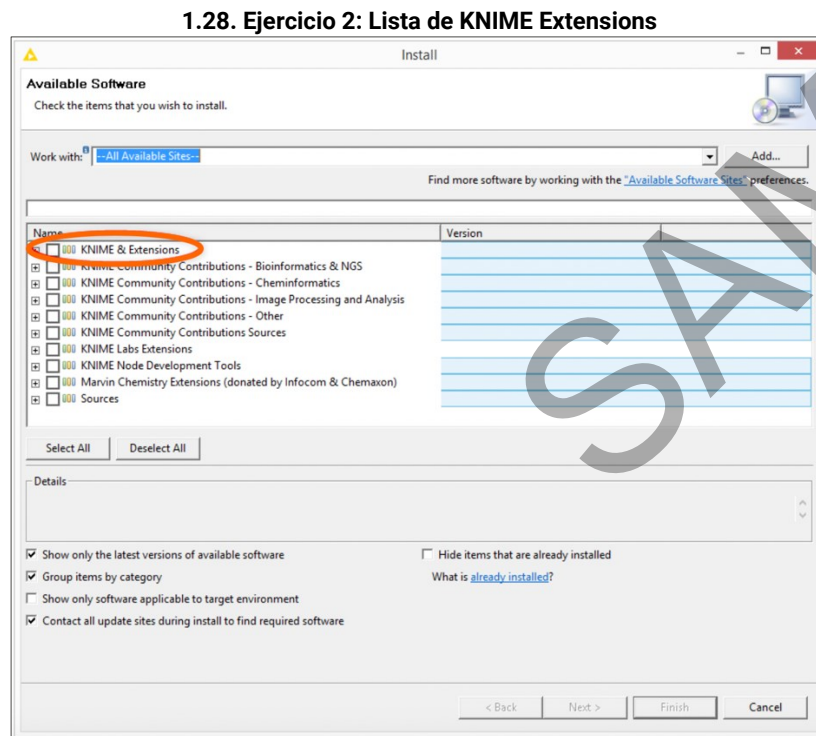
- KNIME Database
- KNIME Javascript Views
- KNIME Report Designer

Solucion al Ejercicio 2

Del Top Menu, seleccione **"File" → "Install KNIME Extensions"**

Seleccione las extensiones requeridas (KNIME Database,...)

Haga Click en **"Next"** y siga instrucciones



Ejercicio 3

Busque todos los "Row Filter" nodes en el Node Repository.

En el panel "Node Description", ¿puede ud explicar la diferencia entre un "Row Filter", a "Reference Row Filter", and a "Nominal Value Row Filter"?

Muestre el efecto de esos nodos en las siguientes tablas

Tabla Original

Position	name	team
1	The Black Rose	4
2	Cynthia	4
3	Tinkerbell	4
4	Mother	4
5	Augusta	3
6	The Seven Seas	3

Tabla de Referencia

Ranking	scores
1	22
3	14
4	10

Solucion al Ejercicio 3

Row Filter

Este nodo permite el filtrado de filas de acuerdo con ciertos criterios. Puede incluir o excluir: ciertos rangos (por número de fila), filas con un determinado ID de fila y filas con un cierto valor en una columna seleccionable (atributo). En el siguiente ejemplo, usamos el siguiente criterio de filtro: team > 3

Original

Position	name	team
1	The Black Rose	4
2	Cynthia	4
3	Tinkerbell	4
4	Mother	4
5	Augusta	3
6	The Seven Seas	3

Filtrada

Position	name	team
1	The Black Rose	4
2	Cynthia	4
3	Tinkerbell	4
4	Mother	4

Reference Row Filter

Este nodo tiene dos tablas de entrada. La primera tabla de entrada, conectada al puerto inferior, se toma como tabla de referencia; la segunda tabla de entrada, conectada al puerto superior, es la tabla a filtrar. Tienes que elegir la columna de referencia en la tabla de referencia y la columna de filtrado en la segunda tabla. Todas las filas con un valor en la columna de filtrado que también existe en la columna de referencia se mantienen, si se selecciona la opción "incluir"; se eliminan si se selecciona la opción "excluir"

Tabla de Referencia

Ranking	scores
1	22
3	14
4	10

Tabla Filtrada

Position	name	team
1	The Black Rose	4
2	Cynthia	4
3	Tinkerbell	4
4	Mother	4
5	Augusta	3
6	The Seven Seas	3

Tabla resultante

Position	name	team
1	The Black Rose	4
3	Tinkerbell	4
4	Mother	4

En el ejemplo de arriba, estamos usando el "Ranking" como la columna de referencia en la table de referencia y la "Position" como la columna de filtrado en la table filtrada. Hemos esgojido de incluir las filas comunes

Nominal Value Row Filter

Filtra las filas según el valor seleccionado de un atributo nominal. Se puede seleccionar una columna nominal y uno o más valores nominales de este atributo como criterio de filtro. Las filas que tienen estos valores nominales en la columna seleccionada se incluyen en los datos de salida. Básicamente es un filtro de fila aplicado a una columna con valores nominales. Las columnas nominales son columnas de cadena y los valores nominales son los valores que contienen.

En el siguiente ejemplo, usamos "name" como columna nominal y "name = Cynthia" como criterio de filtrado.

Original table

Position	name	team
1	The Black Rose	4
2	Cynthia	4
3	Tinkerbell	4
4	Mother	4
5	Augusta	3
6	The Seven Seas	3

Filtered table

Position	name	team
2	Cynthia	4