

**DISEÑO DE GUÍA DE ENSEÑANZA EN ANALÍTICA DE DATOS PARA ESTUDIANTES DEL  
SECTOR SALUD.**

**DANIELA JULIANA TABORDA BETANCUR**

**UNIVERSIDAD TECNOLÓGICA DE PEREIRA  
INGENIERIA DE SISTEMAS Y COMPUTACION  
PEREIRA  
2018**

**DISEÑO DE GUÍA DE ENSEÑANZA EN ANALÍTICA DE DATOS PARA ESTUDIANTES DEL  
SECTOR SALUD.**

**DANIELA JULIANA TABORDA BETANCUR**

**MONOGRAFÍA PARA OPTAR AL TÍTULO DE INGENIERO DE SISTEMAS Y COMPUTACIÓN**

**ASESOR**

**JUAN DE JESUS VELOZA MORA**

**UNIVERSIDAD TECNOLÓGICA DE PEREIRA  
INGENIERIA DE SISTEMAS Y COMPUTACION**

**PEREIRA**

**2018**

## **TABLA DE CONTENIDO**

### **INTRODUCCION**

#### **1. GENERALIDADES**

##### **1.1. PLANTEAMIENTO DEL PROBLEMA**

#### **2. OBJETIVOS**

##### **2.1. OBJETIVOS GENERAL**

##### **2.2. OBJETIVOS ESPECIFICOS**

#### **3. JUSTIFICACION**

#### **4. MARCO TEORICO**

##### **4.1. ¿QUE ES ANÁLISIS DE DATOS?**

##### **4.2. HISTORIA Y EVOLUCION DEL ANALISIS DE DATOS**

###### **4.2.1. Áreas del análisis de datos**

###### **4.2.2. Análisis descriptivo**

###### **4.2.3. Análisis exploratorio**

###### **4.2.4. Análisis confirmatorio y expectativo**

###### **4.2.5. Investigación con variables nominales**

###### **4.2.6. Análisis de datos cuantificados**

###### **4.2.7. Análisis de datos no cuantificados**

###### **4.2.8. Análisis de datos no estructurados**

##### **4.3. MANUAL DE INSTALACIÓN WEKA.**

#### **5. EJEMPLOS DE COMO UTILIZAR EL ANALISIS DE DATOS CON LA HERRAMIENTA WEKA**

##### **5.1. Información que contiene las diferentes consultas sobre los indicadores de lactancia materna**

##### **5.2. Gráfica Mortalidad por desnutrición en menores de cinco años periodos 2014, 2015,2016 y 2017 en el Departamento del Putumayo.**

#### **6. CONCLUSIONES**

#### **7. BIBLIOGRAFIA**

## INTRODUCCION

Los volúmenes de información que se manejan en la actualidad han llevado a repensar el modelo educativo en su formación base para los actuales profesionales. La formación base impartida en los currículos de primeros semestres en carreras profesionales requieren hoy en día de las competencias básicas en analítica de datos como una herramienta fundamental en el desempeño profesional. Este documento pretende hacer una aproximación a los elementos que se requieren para alcanzar estas competencias en la formación de profesionales de la salud.

Esta guía va a hacer una Unidad Conceptual en el cual se desarrollaran los conceptos básicos del Enfoque del Marco Lógico, que permite contextualizar la metodología que se plantea y conocer los criterios centrales para su utilización.

## **1. GENERALIDADES**

### **1.1. PLANTEAMIENTO DEL PROBLEMA**

Ante el volumen de información que se manejan en la actualidad profesiones como las orientadas a la ciencias de la salud deben observar las herramientas que ofrecen campos como el machine learning, la inteligencia artificial, la minería de datos que les permita el acompañamiento de un experto virtual en el aprendizaje, diagnóstico, toma de decisiones las cuales enfrentan en su labor profesional. Elementos de estudio donde un software que le permite buscar temas específicos sobre salud en bibliotecas virtuales que contienen la mejor información, algoritmos de aprendizaje automático que le permiten apoyar el diagnóstico y la toma de decisiones en la mayoría de enfermedades existentes, son herramientas que se encuentran disponibles para hacer posible esta labor. Como pregunta del proceso de investigación se pretende mostrar hasta donde el conocimiento de estas herramientas puede abrir nuevos procesos de aprendizaje en los estudiantes que aspiran ser profesionales de este sector. Esta guía pretende servir de motivador en estos procesos.

## **2. OBJETIVOS**

### **2.1. OBJETIVOS GENERAL**

Diseñar desde una perspectiva práctica una guía de enseñanza en analítica de datos para estudiantes del sector salud.

### **2.2. OBJETIVOS ESPECIFICOS**

- Observar el estado del arte en la enseñanza de analítica de datos.
- Establecer las competencias transversales en este campo para la formación profesional.
- Diseñar una guía metodológica que permita el desarrollo de estas competencias.

### **3. JUSTIFICACION**

En la actualidad, el análisis de datos se utiliza en muchas industrias, independientemente de la rama a utilizar, nos da las bases para tomar o no una decisión o cerciorarnos si una hipótesis es cierta o no.

Actualmente, muchas industrias usan el análisis de datos para sacar conclusiones y decidir acciones a implementar. Cabe mencionar que la ciencia también usa el análisis de datos para comprobar o descartar teorías o modelos existentes.

## **4. MARCO TEORICO**

### **4.1. ¿QUE ES ANÁLISIS DE DATOS?**

El análisis de datos con lleva un proceso organizado de generar conocimiento, es por ello que basada en la lectura de varios documentos señalados en las referencias, se pretende esquematizar y sintetizar como se inicia y bajo qué enfoques se puede realizar una investigación científica, haciendo parte de una metodología que abarca desde el método científico utilizado, siguiendo con la escogencia de la población y muestra a estudiar, las técnicas de recolección de datos y la representación y análisis de esos datos, de manera que cumplan con dos requisitos de toda investigación como es que los resultados obtenidos sean Válidos y Confiables.

En los primeros esquemas se sintetizara como se inicia la generación del conocimiento científico, bajo qué postura se genera en análisis de datos y de que tipos de análisis se habla de acuerdo a diferentes tipos de enfoques, posteriormente se hará énfasis en la parte de la recolección y análisis de los datos a través de distintas técnicas que se disponen para la misma.

El análisis de datos es un proceso consistente en inspeccionar, limpiar y transformar datos con el objetivo de resaltar información útil, lo que sugiere conclusiones, y apoyo en la toma de decisiones. El análisis de datos tiene múltiples facetas y enfoques, que abarca diversas técnicas en una variedad de nombres, en diferentes negocios, la ciencia, y los dominios de las ciencias sociales.

El propósito del análisis es resumir las observaciones llevadas a cabo de forma tal que proporcionen respuesta a las interrogantes de la investigación. La interpretación, más que una operación distinta, es un aspecto especial del análisis su objetivo es "buscar un significado más amplio a las respuestas mediante su trabazón con otros conocimientos disponibles (Selltiz, 1970) que permitan la definición y clarificación de los conceptos y las relaciones entre éstos y los hechos materia de la investigación.

## **4.2. HISTORIA Y EVOLUCION DEL ANALISIS DE DATOS**

### **4.2.1. Áreas del análisis de datos**

Al hablar de la historia del análisis de datos tengo que hablar también de:

### **4.2.2. Análisis descriptivo**

Ayudará a observar el comportamiento de la muestra en estudio, a través de tablas, gráficos. Los resultados recogidos en la muestra se resumen en una matriz de datos  $N \times M$ , en la cual  $N$  es el número de unidades de análisis utilizadas (número de casos) y  $M$  es el número de características de dichas unidades, unidades de las que tenemos información.

### **4.2.3. Análisis exploratorio**

Pretende partir de un conocimiento profundo y creciente de los datos para, trabajando inductivamente, llegar a un modelo ajustado de los datos.

### **4.2.4. Análisis confirmatorio y expectativo**

La mayor parte de las técnicas tradicionales de análisis estadístico de los datos tienen un carácter deductivo confirmatorio. De todas las técnicas de análisis confirmatorio la más útil para el sociólogo es aquella que parte del análisis de variables, entre las que cabe distinguir: nominales y de intervalo o de razón

### **4.2.5. Investigación con variables nominales**

En la investigación con variables nominales se utilizan técnicas como el análisis de varianza o factorial.

Los datos a partir de los cuales el investigador inicia el análisis, son diferentes según el nivel de elaboración realizado, el cual depende de la naturaleza del problema de investigación y consecuentemente, del tipo de investigación; también de las técnicas y procedimientos seguidos en elaboración.

De acuerdo a estas consideraciones, los datos que se utilizan en el análisis pueden ser:

- Datos cuantificados
- Datos no cuantificados
- Datos no estructurados

#### **4.2.6. Análisis de datos cuantificados**

Algunos tipos de estudios, por su naturaleza, aportan datos elaborados, es decir, cuantificados.

- El tratamiento estadístico de los datos permite un análisis adecuado que puede tener diversos alcances, los cuales dependen de los objetivos de la investigación y de las hipótesis formuladas.
- Determinar lo que es típico en el grupo estudiado. (Se utiliza algunas de las medidas de tendencia central, según el caso)
- Indicar si existen variaciones entre los sujetos del grupo, señalando de qué tipo y magnitud son. (Se utiliza alguna de las medidas de variabilidad; cada una proporciona datos sobre un aspecto diferente)
- Mostrar la forma cómo están distribuidos los individuos con respecto a la variable que se mide. (Se utiliza el desarrollo de una curva de distribución).
- Mostrar la relación existente entre dos o más variables. (Se aplica el coeficiente de variabilidad).

#### **4.2.7. Análisis de datos no cuantificados**

No todos los aspectos del material recogido pueden ser categorizados y, consecuentemente, cuantificados, debido, en algunos casos, a la falta de precisión en la definición de las categorías, lo que dificulta el análisis de los resultados.

Por este motivo, se recomienda considerar que cada categoría propuesta comprenda un amplio margen de criterios para las respuestas. De todos modos, los datos sin elaborar, "pueden ser utilizados en el análisis e interpretación sin tener en cuenta si han sido o no cuantificados en todos los aspectos", pues cumplen una función importante:

- Ayudan a entender el significado de las categorías
- Aclaran la naturaleza de las relaciones entre las variables determinadas estadísticamente
- Permiten orientar al investigador a formular nuevas hipótesis para futuras investigaciones

#### **4.2.8. Análisis de datos no estructurados**

El material no estructurado es el que proviene, por ejemplo, de observaciones o entrevistas no estructuradas, en las cuales se recoge mucho material, a veces valioso, pero sin ninguna pauta que permita alguna forma de organización y menos de clasificación. En algunos casos, los estudios de nivel exploratorio, que no se inician con hipótesis, cubren aspectos diversos, los cuales conducen al acopio de datos en cantidad excesiva y no estructurado.

### 4.3. MANUAL DE INSTALACIÓN WEKA.

Para realizar la descarga de la herramienta weka debemos ingresar al siguiente link <https://www.cs.waikato.ac.nz/ml/weka/> en este se podrá encontrar el software para para diferentes sistemas operativos como Windows, Mac OSX y Linux.

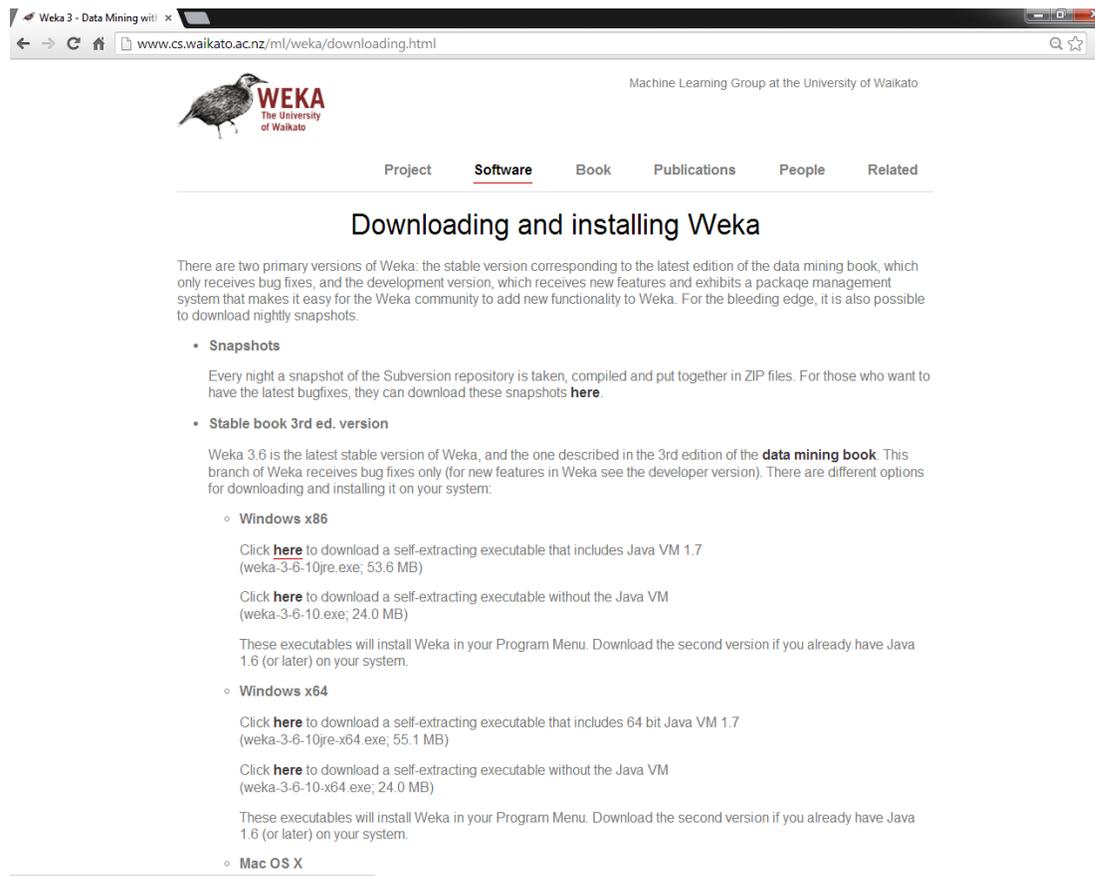


IMAGEN 1. PASO 1

Después de seleccionar el sistema operativo y la arquitectura (En caso de Windows x86 y x64) encontrar la siguiente página. En la parte inferior izquierda podrá observar el porcentaje de descarga de la herramienta.

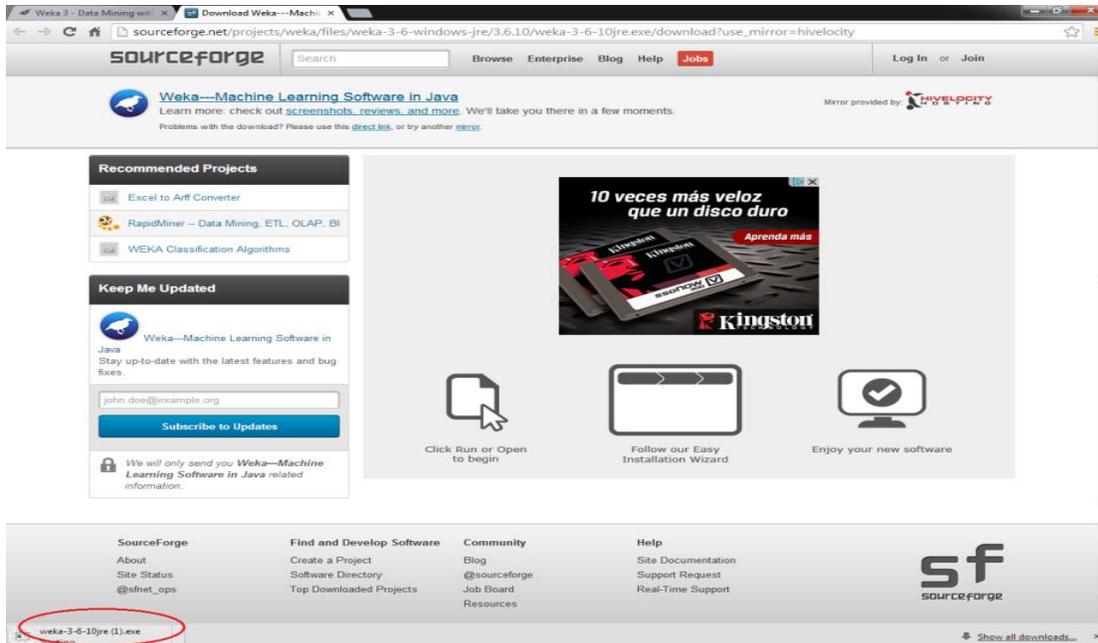


IMAGEN 2. PASO 2

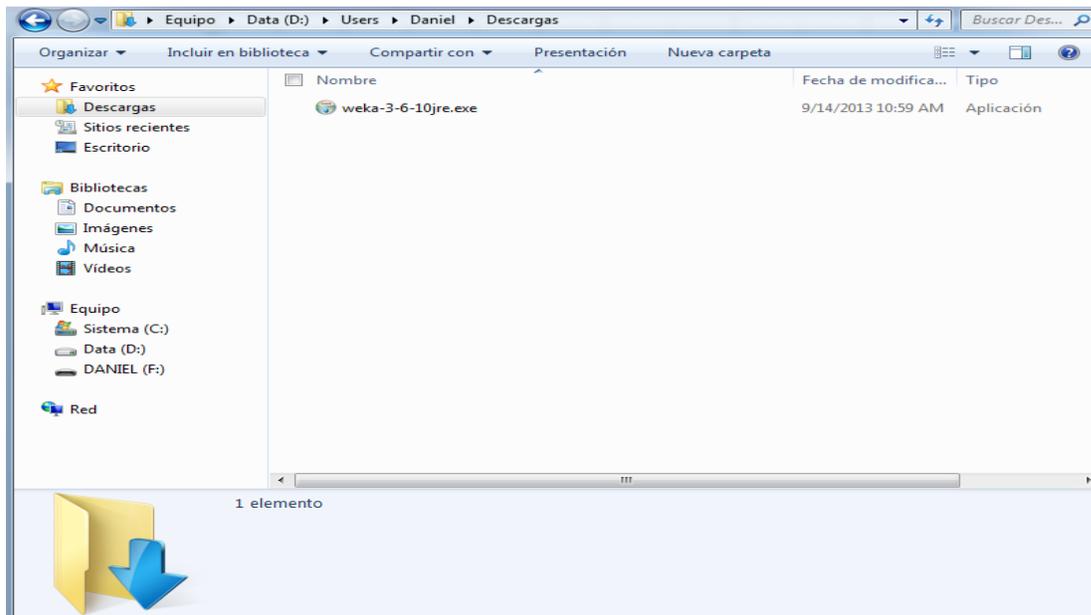


IMAGEN 3. PASO 3: En la siguiente ventana podrá visualizar el ejecutable de la herramienta descargado.

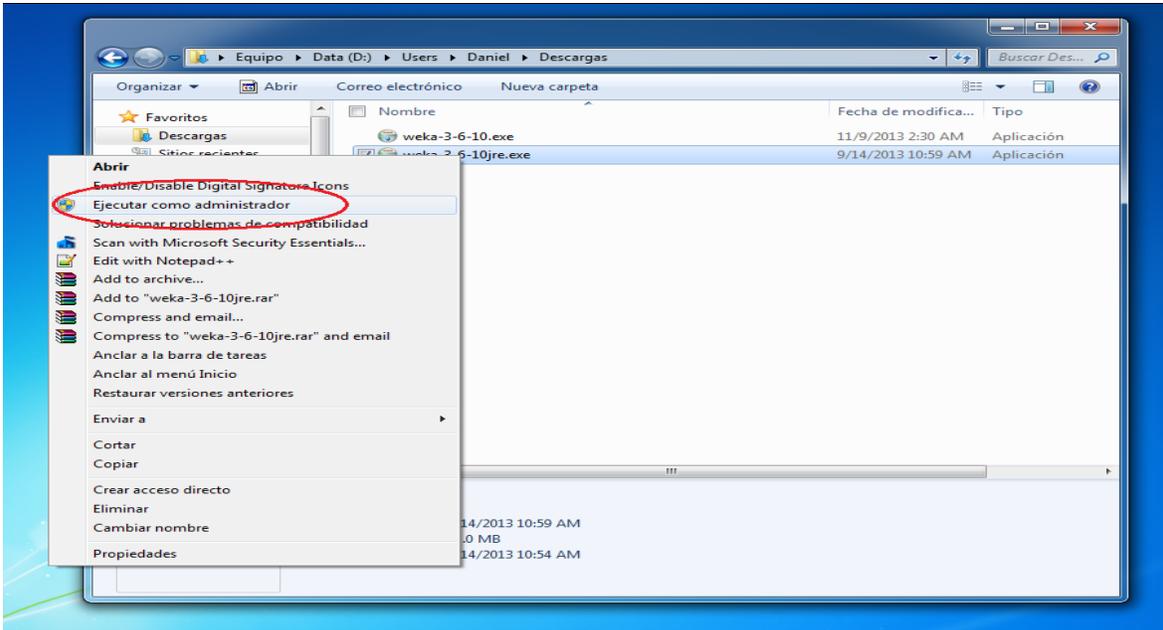


IMAGEN 4. PASO 4: Hacer clic derecho sobre el instalador, seguido de un doble clic en la opción Ejecutar como administrador.

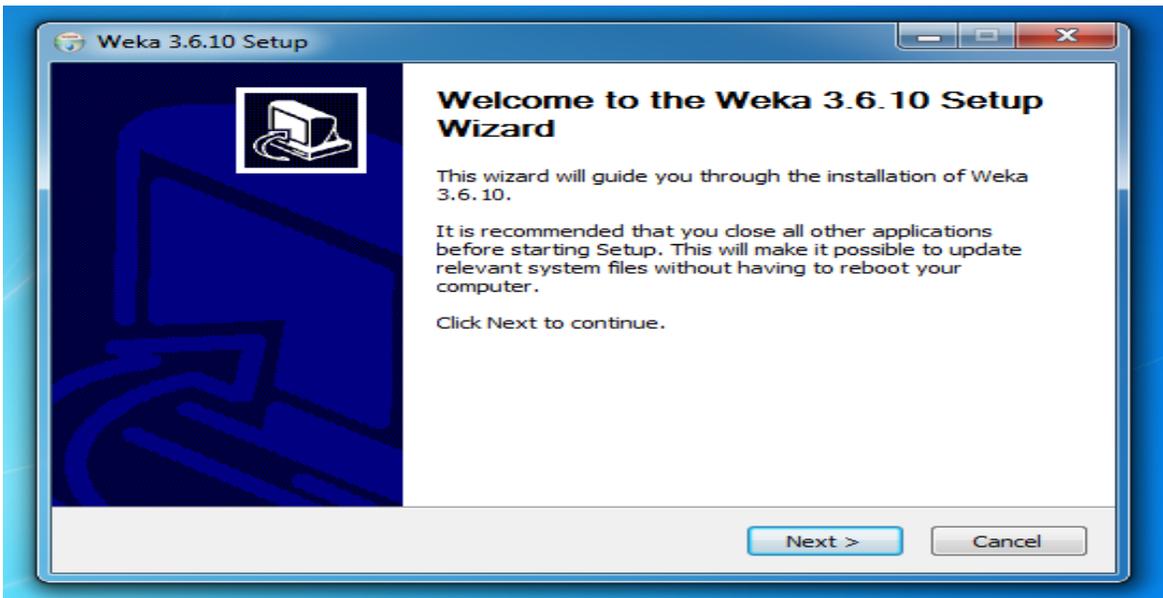


IMAGEN 5. PASO 5: Hacer clic en la opción *Next* para iniciar la instalación.

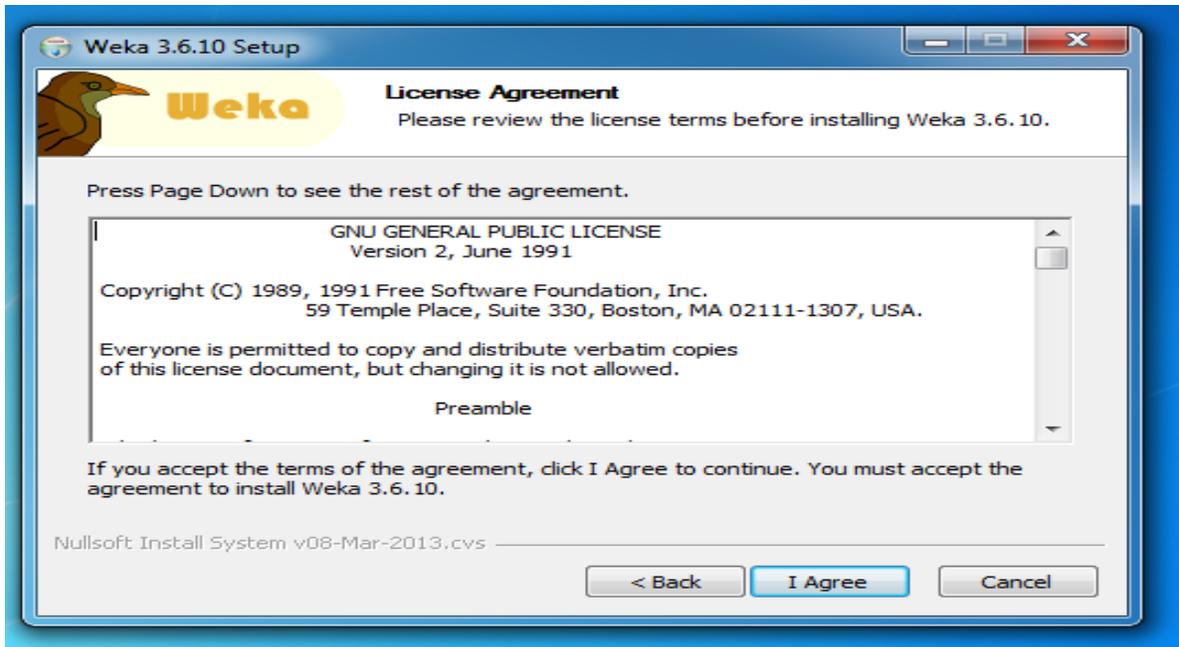


IMAGEN 6. PASO 6: Hacer clic en la opción *I Agree* para aceptar los términos de instalación.

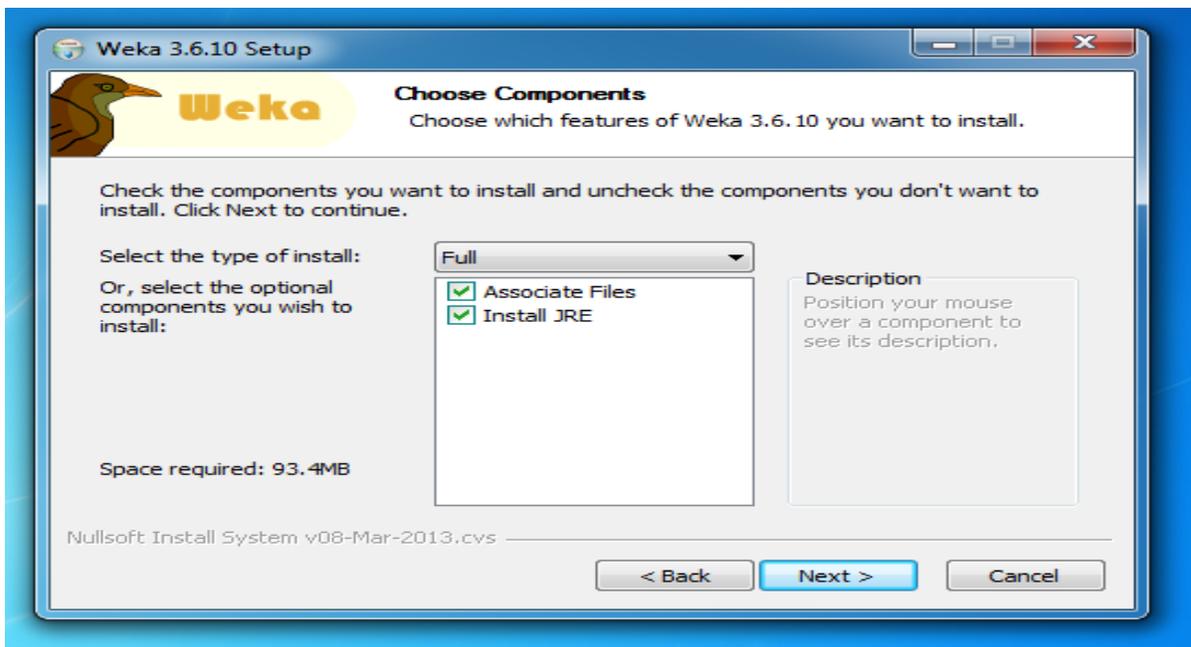


IMAGEN 7. PASO 7: Hacer clic en la opción *Next* para continuar con la instalación.

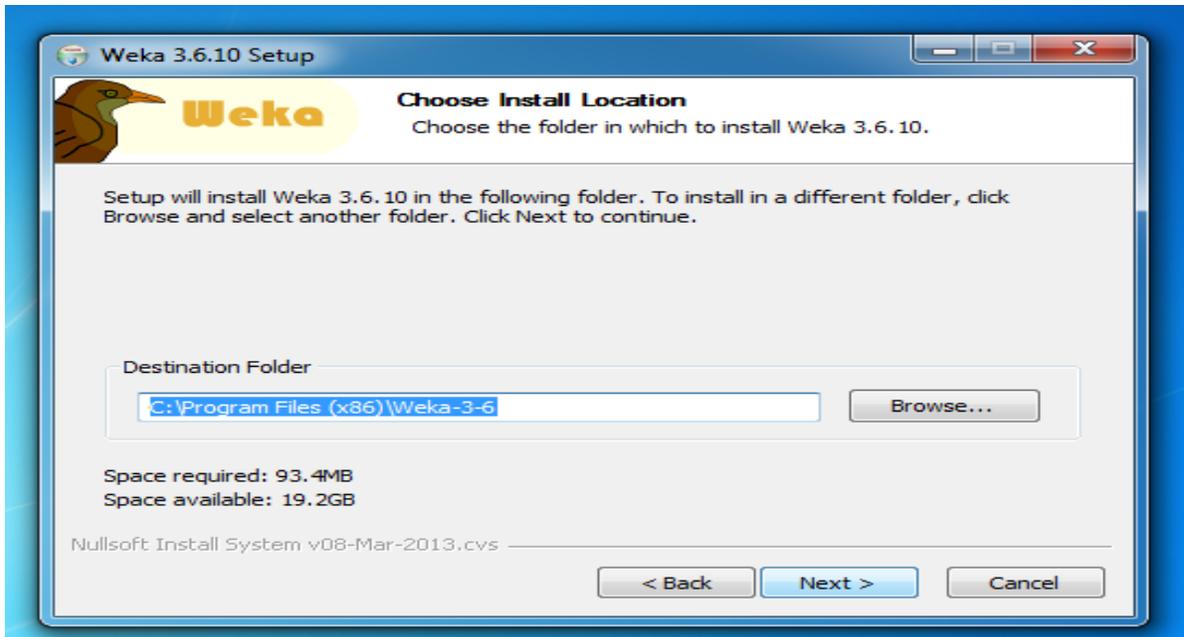


IMAGEN 8. PASO 8: Hacer clic en la opción *Next* para continuar con la instalación.

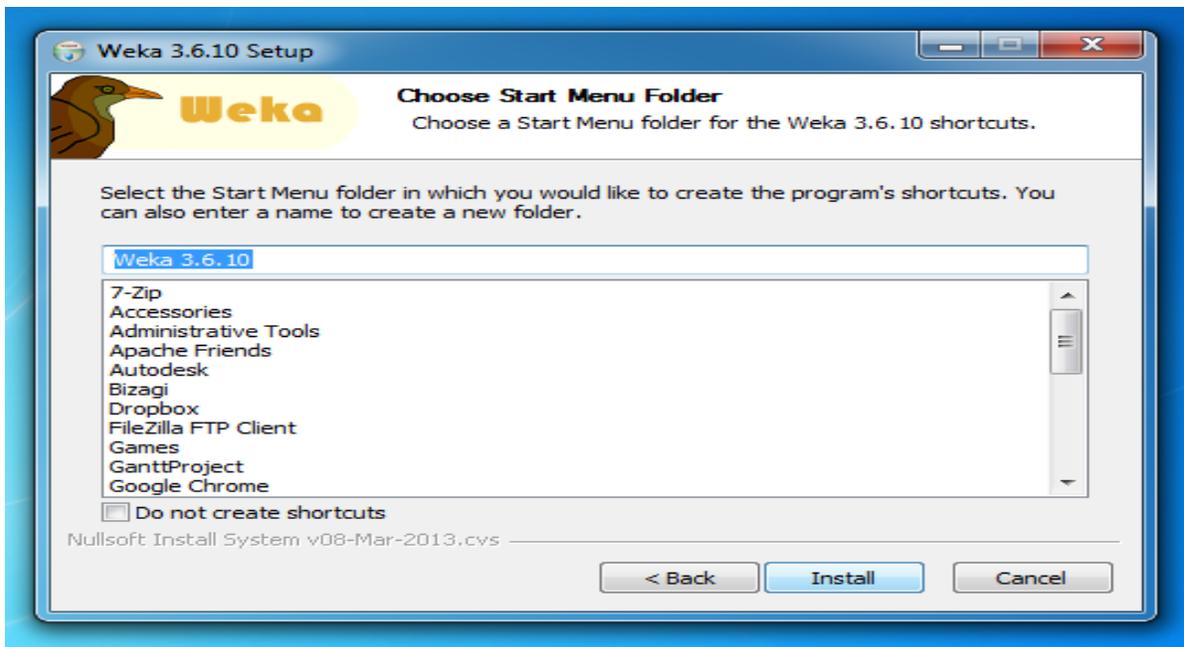


IMAGEN 9. PASO 9: Hacer clic en la opción *Install* para continuar con la instalación.

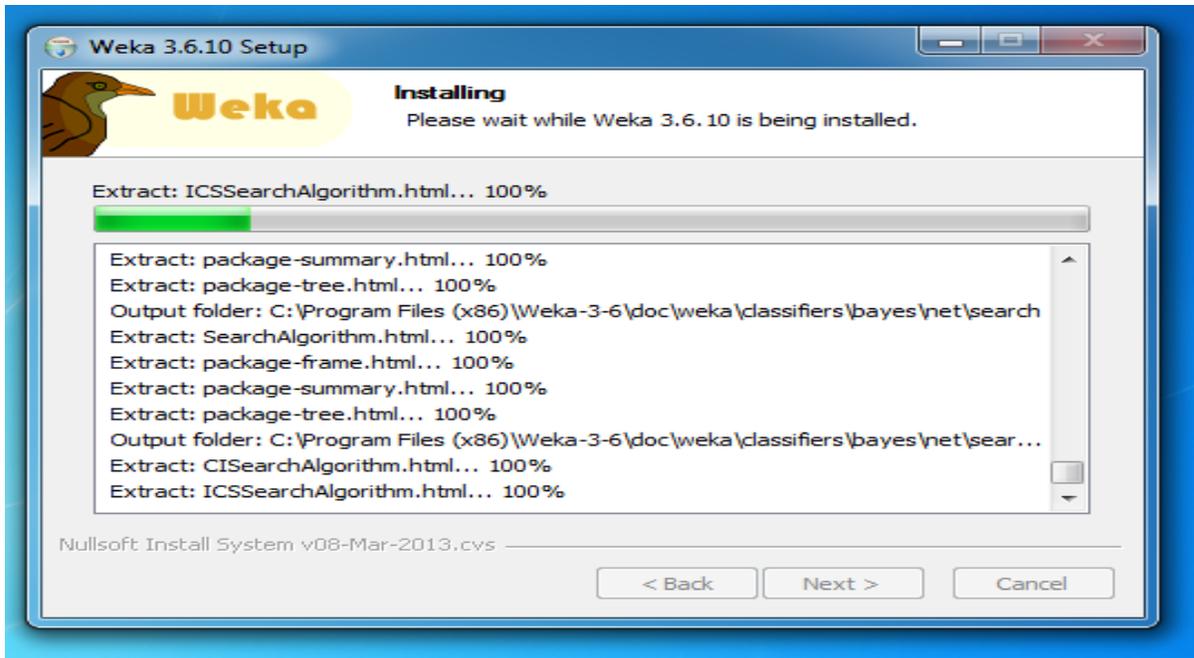


IMAGEN 10. PASO 10: Después de realizar la opción *Install* podrá observar la siguiente ventana, esta indica el porcentaje de instalación.



IMAGEN 11. PASO 11: Hacer clic en la opción *Instalar* para realizar la instalación de los complementos de Java



IMAGEN 12. PASO 12: La siguiente ventana muestra el proceso de configuración de Java.

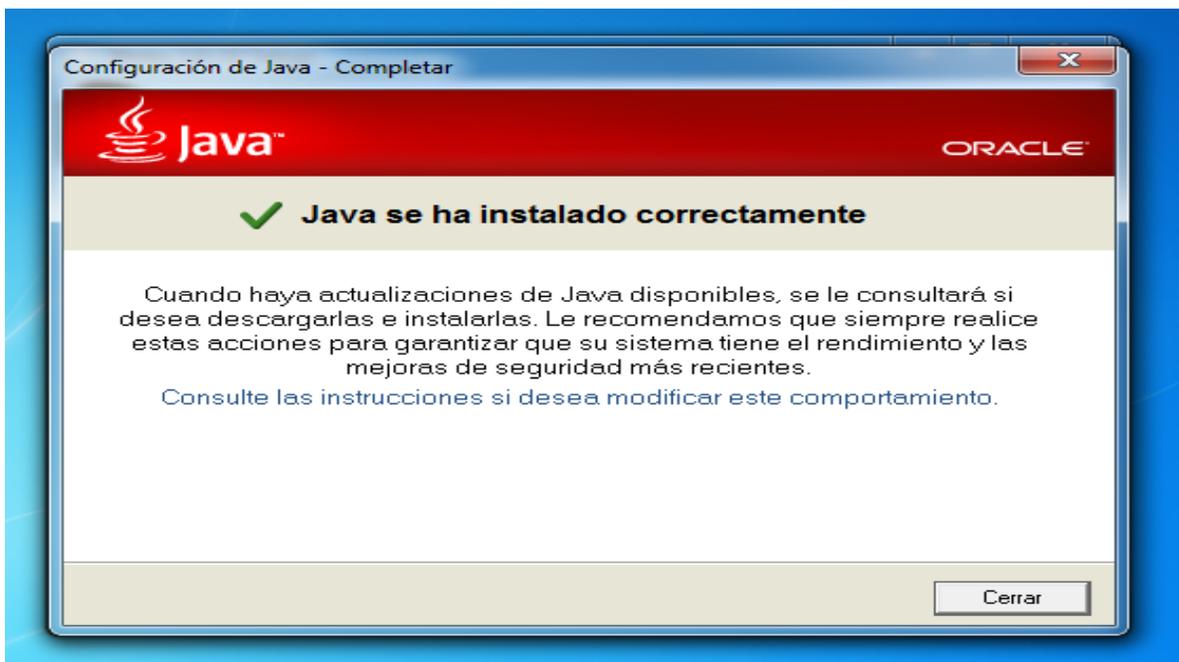


IMAGEN13. PASO 13: Hacer clic en la opción *Cerrar* para terminar con la configuración de Java.

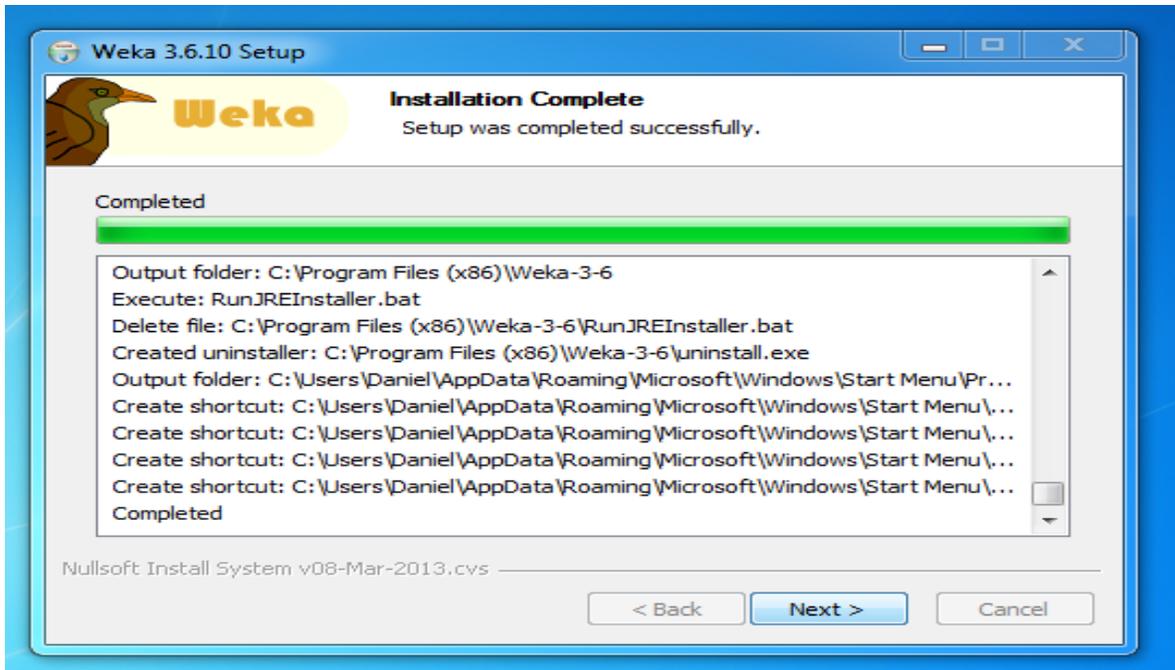


IMAGEN 14. PASO 14: La siguiente venta visualiza el porcentaje de instalación de la herramienta Weka.

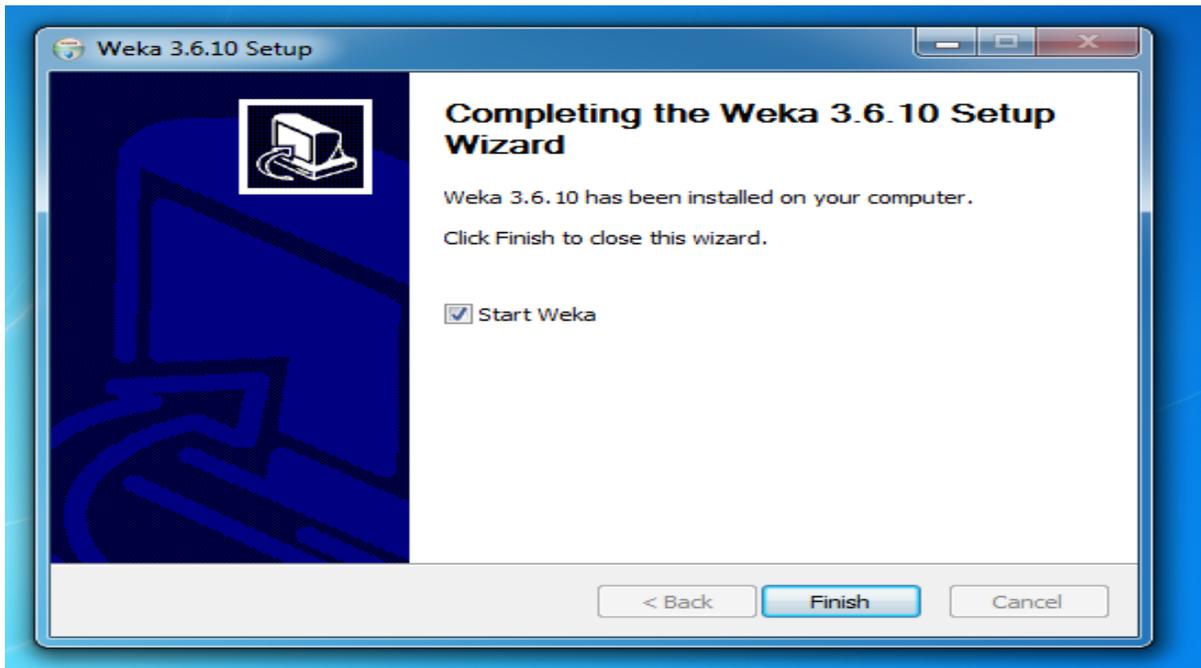


IMAGEN 15. PASO 15: Hacer clic en la opción *Finish* para concluir con la instalación.

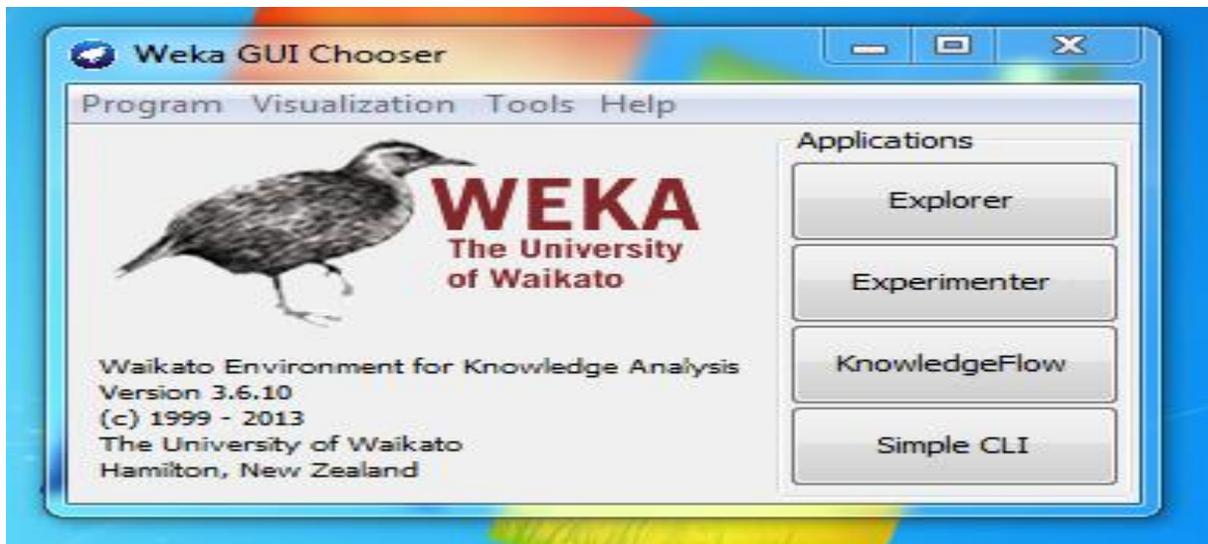


IMAGEN 16. PASO 16: Finalmente, una vez realizada la instalación en su totalidad, la siguiente ventana se desplegará. Esto indica que el proceso de instalación se completó con éxito.

#### 4.3.1 Exportar archivos para weka

En este caso vamos a utilizar weka en formato CSV donde se puede exportar fácilmente desde Microsoft Excel, por lo que una vez que puede obtener sus datos en Excel, puede convertirlos fácilmente en formato CSV.

De hecho, es mucho más probable que esté en formato de valores separados por comas (CSV). Este es un formato simple donde los datos se presentan en una tabla de filas y columnas y se usa una coma para separar los valores en una fila. Las comillas también se pueden usar para rodear valores, especialmente si los datos contienen cadenas de texto con espacios.

Tenemos nuestro archivo donde podemos observar que tiene al final el .xlsx  
Como lo muestra la imagen

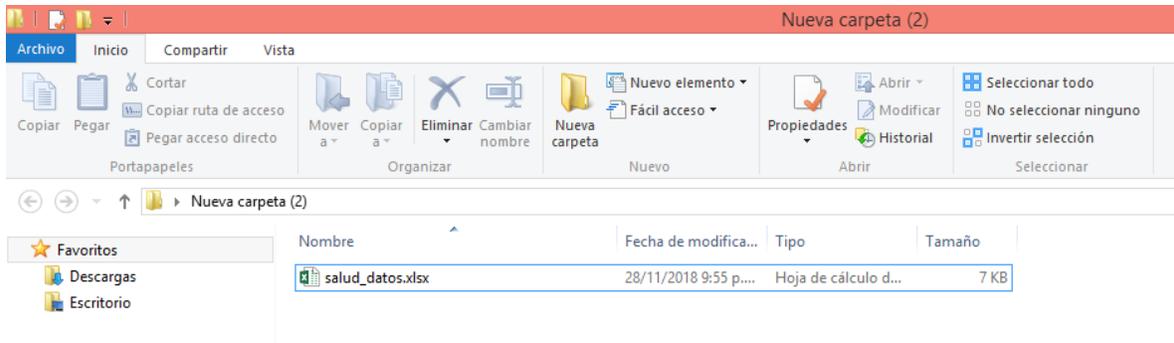


IMAGEN 17. Tabla con formato.xlsx.

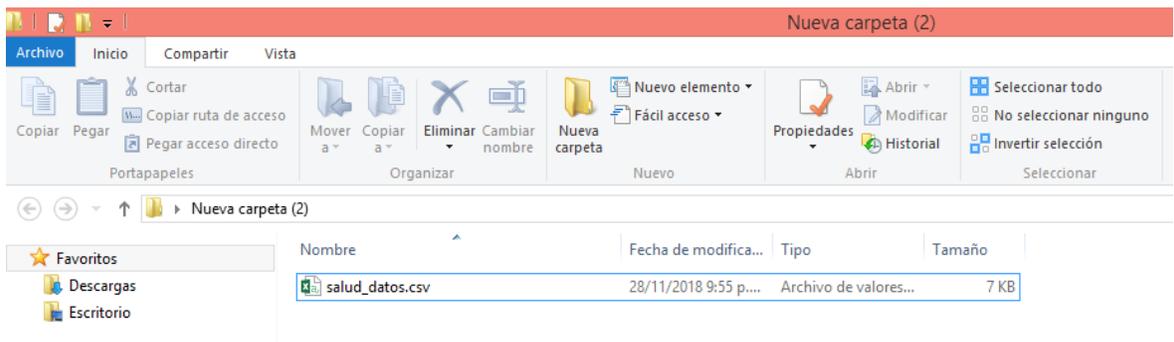


IMAGEN 18. Se pasa el formato .xlsx por .csv

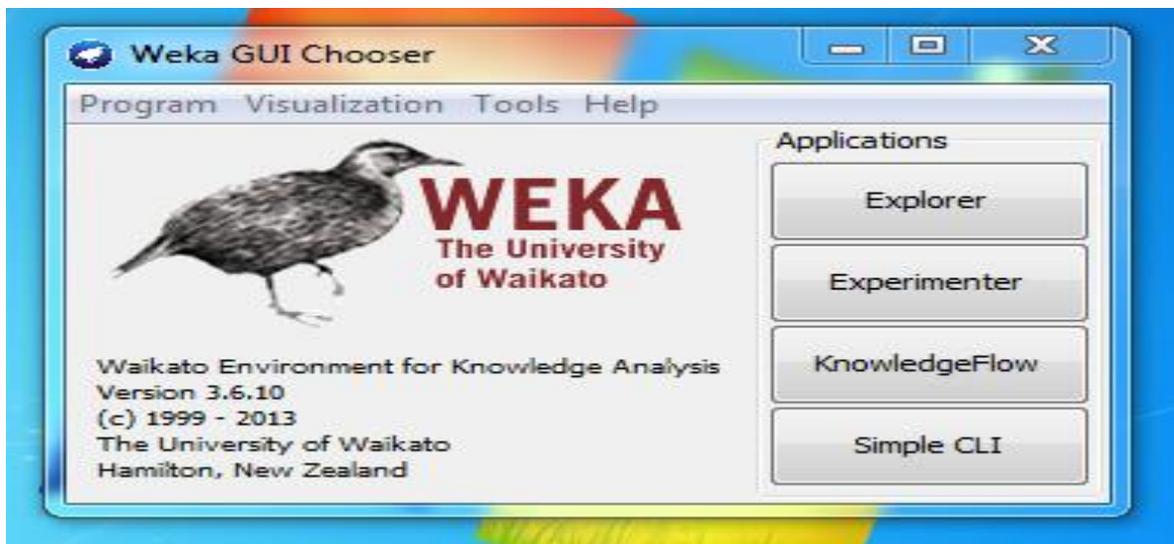


IMAGEN 19. Iniciamos el selector weka le damos en la opción Explorer

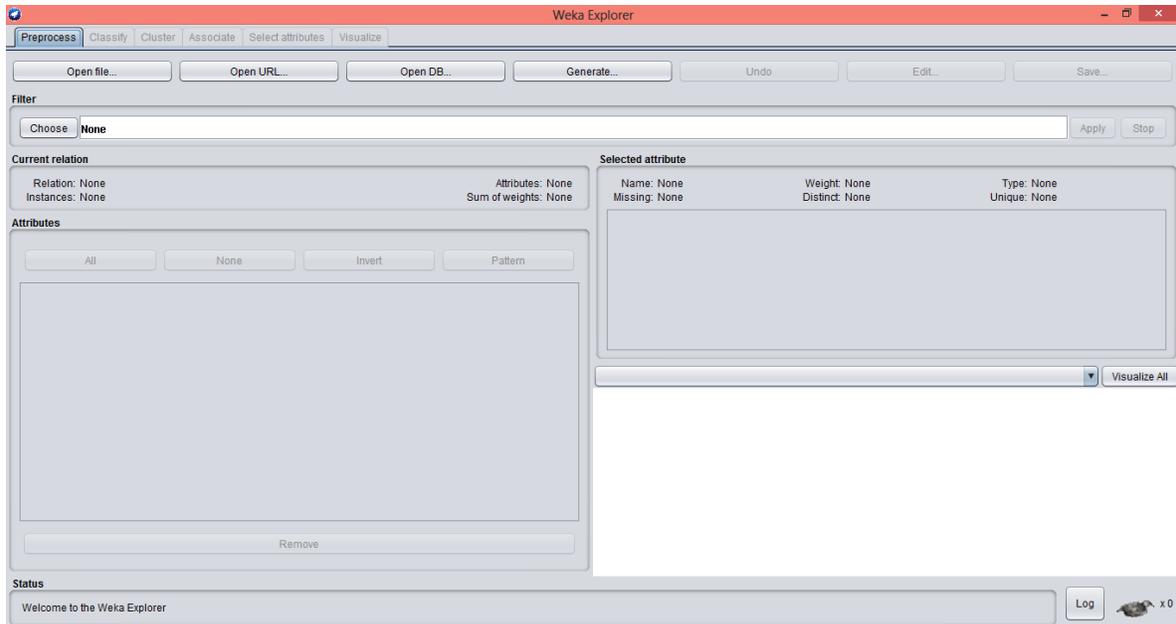


IMAGEN 20. Nos muestra la siguiente ventana damos en open file y nos muestra la siguiente ventana

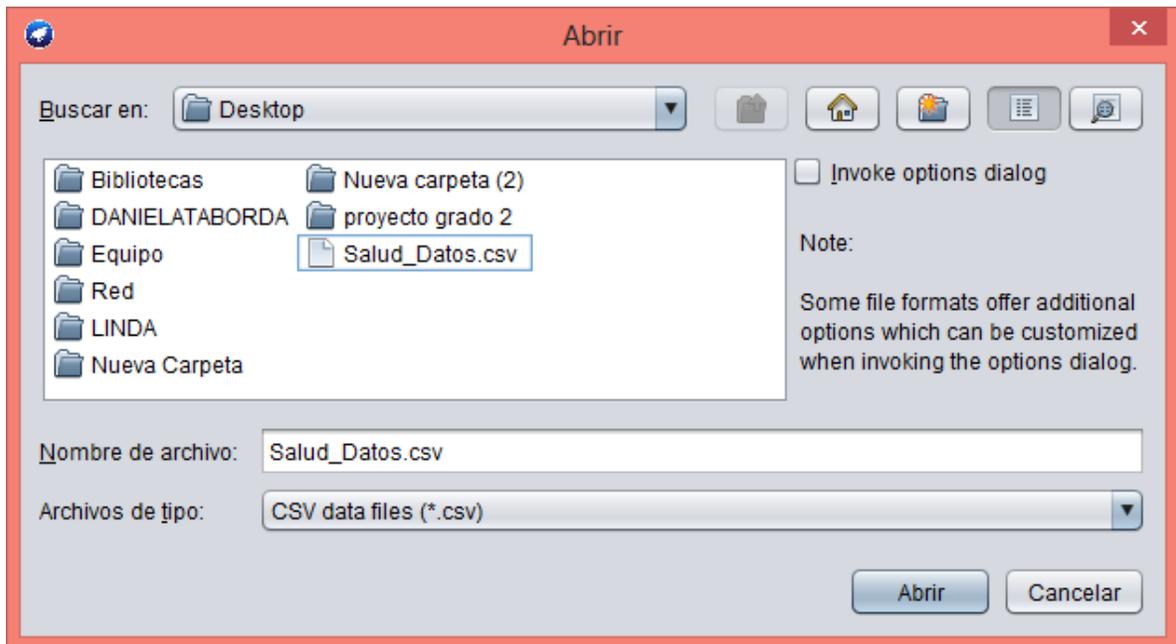


IMAGEN 21. Seleccionamos donde aparece archivo de tipos debemos de cambiarlo el formato a csv para que nos aparezca el documento y le damos abrir.

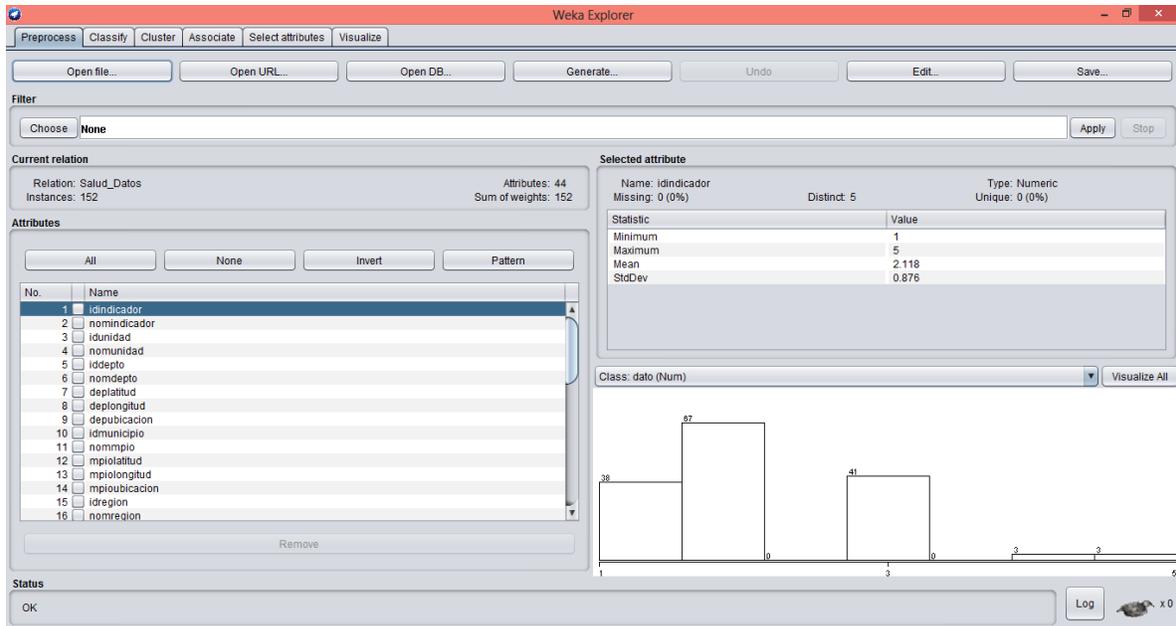


IMAGEN 22. Después de darle abrir nos muestra todo los atributos.

## 5. EJEMPLOS DE COMO UTILIZAR EL ANALISIS DE DATOS CON LA HERRAMIENTA WEKA

### 5.1. Información que contiene las diferentes consultas sobre los indicadores de lactancia materna

En este ejemplo se hace un análisis de como la duración de lactancia materna exclusiva en menores de 3 años donde se está evaluando:

El identificador único del indicador de salud (idindicador), Nombre del indicador de salud (nomindicador), Identificador de la unidad (id unidad) entre otros.

Link del ejemplo:

<https://www.datos.gov.co/Salud-y-Proteccion-Social/Salud-Datos/5kgh-si6m/data>

En este ejemplo lo mostramos en weka y esta herramienta nos muestra el análisis gráficamente.





TABLA 3.

### 5.2. Gráfica Mortalidad por desnutrición en menores de cinco años periodos 2014, 2015,2016 y 2017 en el Departamento del Putumayo.

Ejemplo se analiza un conjunto de datos contiene el número de casos de mortalidad por desnutrición en niños menores de cinco años, en el departamento del Putumayo durante los años 2014, 2015 ,2016 y hasta 31 de marzo de 2017.

Link del ejemplo:

<https://www.datos.gov.co/Salud-y-Protecci-n-Social/Mortalidad-por-desnutrici-n-en-menores-de-cinco-a-/8vbb-ztj3/data>

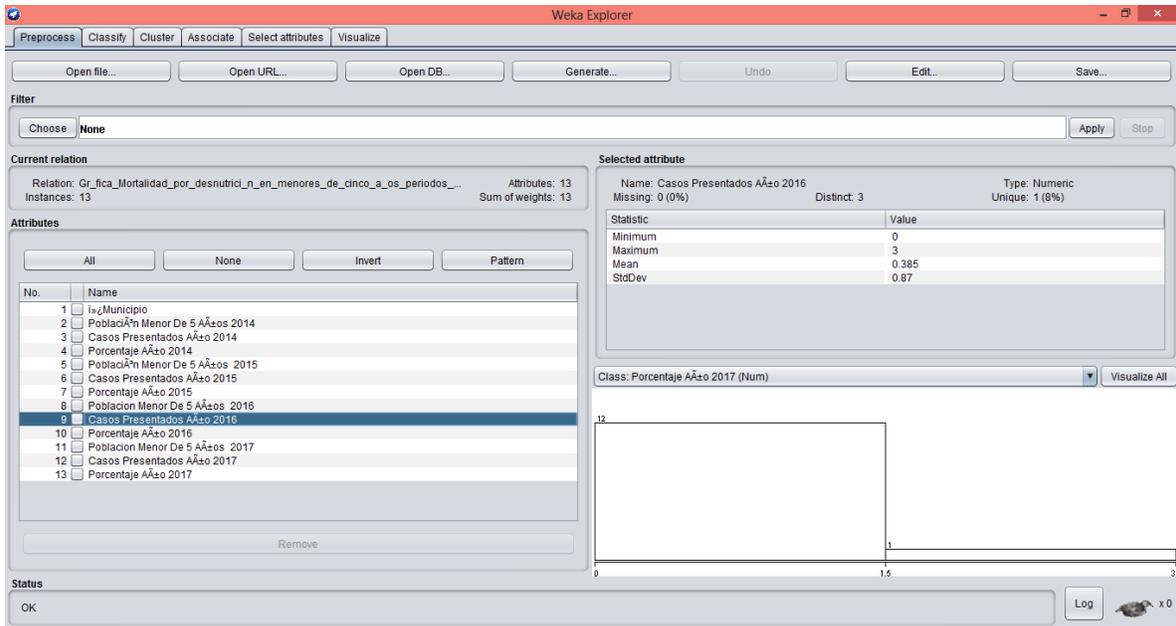


TABLA 1.

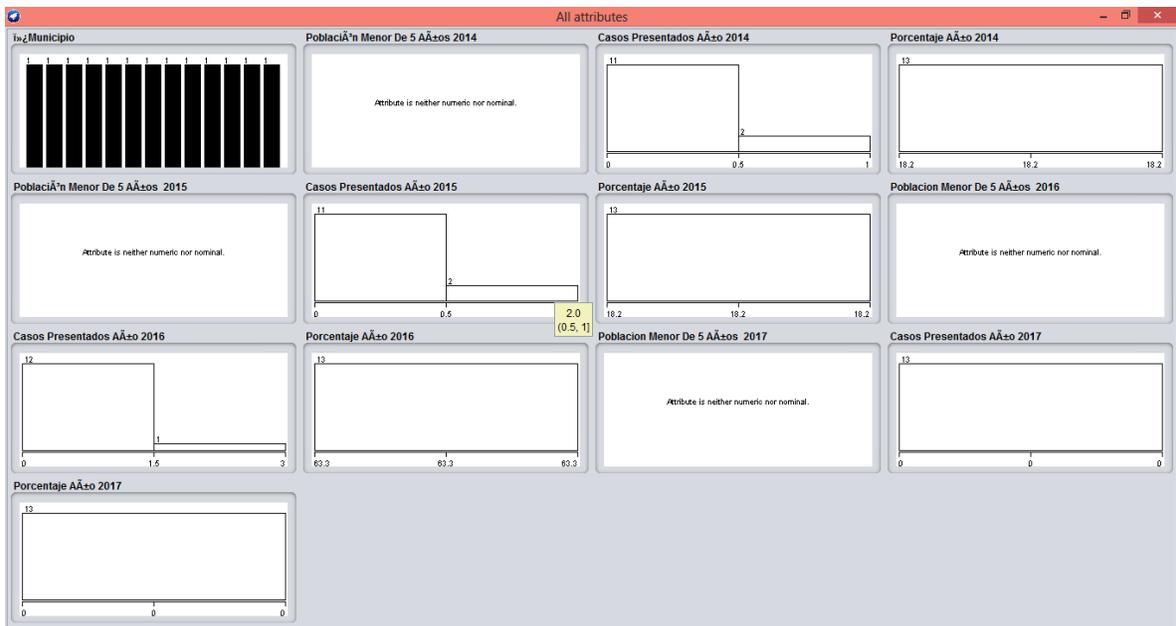


TABLA 2.

## 6. CONCLUSIONES

- Se puede observar que en analítica de datos ayuda a que no se escape ningún detalle en el momento de tomar una decisión.
- La herramienta weka facilita el análisis de datos debido a que muestra de una manera general la recolección de datos y el análisis de los mismos.
- Se puede observar que el uso de herramientas tecnológicas ayudan al diagnóstico temprano de algún tipo de enfermedad.

## 7. BIBLIOGRAFIA

- <https://www.datos.gov.co/>
- <https://www.datos.gov.co/Salud-y-Protecci-n-Social/Mortalidad-por-desnutrici-n-en-menores-de-cinco-a-/8vbb-ztj3/data>
- <https://www.datos.gov.co/Salud-y-Protecci-n-Social/Salud-Datos/5kgh-si6m/data>
- <https://www.informee.co/analitica-sector-salud/>
- <https://ticnegocios.camaravalencia.com/servicios/tendencias/los-efectos-positivos-de-la-analitica-de-datos-en-la-empresa/>
- <https://searchdatacenter.techtarget.com/es/guia/Principios-de-la-analitica-de-datos-una-guia-esencial>