



## **Algoritmo de agrupamiento no supervisado *K-Means* con Python**

**Servicio Nacional de Aprendizaje - SENA**

Nivel de formación: **Complementaria**

## 01 Presentación

Estudia Algoritmo de Agrupamiento No Supervisado *K-Means* con Python y, como egresado SENA, desarrollarás la habilidad de determinar las **características y criterios de agrupamiento de datos** según los **objetivos del análisis**. Además, dominarás la aplicación de algoritmos no supervisados para **agrupar y validar datos** en función de los **objetivos definidos**. Así, estarás preparado para **desempeñarte** en el **sector productivo**, donde aprovecharás al máximo los datos para tomar decisiones y prever comportamientos futuros mediante análisis estadísticos y técnicas de analítica de datos.

Este curso tendrá una duración de **48 horas** y se impartirá en **modalidad 100% virtual**. Para **inscribirte**, debes contar con un computador o *tablet* con **acceso a internet**.

¡Súmate a esta propuesta de formación y haz parte de los miles de colombianos que le apuestan al cambio!

Inscríbete en [www.senasofiaplus.edu.co](http://www.senasofiaplus.edu.co)



**Código**  
21710111



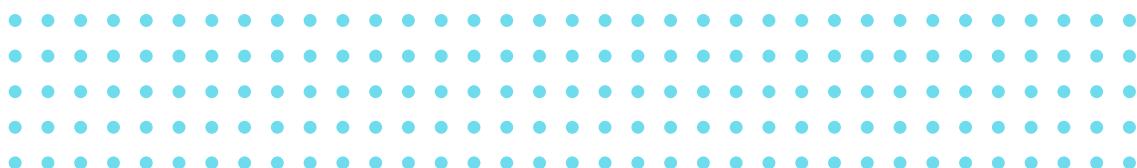
**Horas**  
48



**Duración**  
1 mes



**Modalidad**  
Virtual



## 02 Justificación del programa

En la actualidad, existe una alta inversión en *Big Data* e inteligencia artificial por parte de las organizaciones, y es por ello que es muy importante entender cómo en el mundo cada vez se producen y almacenan más y más datos en todos los procesos desarrollados al interior de las organizaciones. Gracias a los procesos de transformación digital y tendencias tecnológicas como la industria 4.0, las empresas, además de mantenerse en el mercado, buscan ser cada vez más competitivas y deben integrar a su recurso humano profesionales con capacidades de análisis y aprovechamiento de datos masivos.

En este aspecto, el análisis de los datos es una herramienta muy poderosa, que permite a las gerencias de las organizaciones tener acceso a información de valor que se convierte en el insumo más importante para la toma de decisiones. Si bien es una realidad la cantidad de datos y las distintas fuentes que permiten recopilar y generar *Datasets* o conjuntos de datos, estos pueden ser tan diversos, dispersos o de gran tamaño y valor que requieren ser organizados o agrupados por características similares o por los requisitos propios del análisis a desarrollar en la organización.

En este sentido, el SENA ofrece el programa de formación complementaria en algoritmo de agrupamiento no supervisado *K-Means* con Python, con el objetivo de brindar una formación que permita adquirir habilidades y destrezas en el manejo y selección del algoritmo de agrupación adecuado según el tipo de análisis o tipo de trabajo a realizar con ellos, utilizando para este proceso el lenguaje de programación **Python**, de tal manera que los aprendices en esta área contribuyan al sector productivo encontrando ventajas en el aprovechamiento de sus datos para la toma de decisiones y la predicción de comportamientos futuros, basados en análisis estadístico y la aplicación de técnicas propias del campo de la analítica de datos.

## 03 Competencia a desarrollar

220501114 - Sistematizar datos masivos de acuerdo con métodos de analítica y herramientas tecnológicas.

## 04 Perfil de ingreso

Para la realización de este programa, es necesario que el aprendiz tenga un adecuado manejo de las siguientes herramientas informáticas y de comunicación:

Correo electrónico, chats, Messenger, procesadores de texto, software de presentación, Internet, navegadores, otros sistemas y herramientas tecnológicas para la formación virtual.

Cumplir con los requisitos para formación complementaria del SENA. Estar registrado en SofiaPlus.

## 05 Estrategia metodológica

Centrada en la construcción de autonomía para garantizar la calidad de la formación en el marco de la formación por competencias, el aprendizaje por proyectos y el uso de técnicas didácticas activas que estimulan el pensamiento para la resolución de problemas simulados y reales; soportada en la utilización de las tecnologías de la información y la comunicación, integradas en ambientes virtuales de aprendizaje que, en todo caso, recrean el contexto productivo y vinculan al aprendiz con la realidad cotidiana y el desarrollo de las competencias.

Igualmente, debe estimular de manera permanente la autocrítica y la reflexión del aprendiz sobre el quehacer y los resultados de aprendizaje que logra a través de la vinculación activa de las cuatro fuentes de información para la construcción de conocimiento:

- El instructor - Tutor.
- El entorno.
- Las TIC.
- El trabajo colaborativo.