



PROCESO DE GESTIÓN DE FORMACIÓN PROFESIONAL INTEGRAL GUÍA DE APRENDIZAJE

1. IDENTIFICACIÓN DE LA GUÍA DE APRENDIZAJE

- **Denominación del programa de formación:** Análisis exploratorio de datos en Python.
- **Código del programa de formación:** 21710112.
- **Nombre del proyecto:** N/A.
- **Fase del proyecto:** N/A.
- **Competencia:** 220501114- Sistematizar datos masivos de acuerdo con métodos de analítica y herramientas tecnológicas.
- **Resultados de aprendizaje:**
 - 220501114-01. Definir el objetivo del análisis a partir del planteamiento de preguntas que permitan conocer los datos recopilados.
 - 220501114-02. Organizar, preparar y visualizar datos que den respuesta a las preguntas definidas para el análisis.
 - 220501114-03. Validar el resultado del análisis de acuerdo con los objetivos definidos para la organización.
- **Duración de la Guía:** 48 horas.

2. PRESENTACIÓN

El SENA le da la bienvenida a la presente guía de aprendizaje, que tiene como fin dar a conocer en detalle la descripción de cada una de las actividades de aprendizaje a desarrollar, los criterios de evaluación y la descripción de las evidencias de aprendizaje que tendrá en el presente curso durante las próximas cuatro semanas, esta guía tiene como objetivo validar la asimilación de conocimientos y el alcance de la competencia establecida para el desarrollo del programa, relacionada con sistematizar datos masivos de acuerdo con métodos de analítica y herramientas tecnológicas.

Desde esta perspectiva, las actividades propuestas están orientadas a comprender las metodologías utilizadas para la ciencia de los datos, de igual manera, el uso de una herramienta metodológica para identificar oportunidades de negocio denominada *Desing Thinking*, que ayuda en el proceso de instalación y el uso de herramientas de *software* necesarios para el análisis, transformación y visualización de los datos, finalmente se aplicará la estadística en análisis univariado, bivariado, y de validación de resultados, que permitan la comparación con los objetivos propuestos o deseados en la organización.

La mayor y mejor disposición para el desarrollo de las actividades ofrecidas mediante esta guía, fortalecerán las habilidades y los conocimientos necesarios para que dentro del contexto productivo pueda contribuir al análisis de datos como mecanismo para la toma de decisiones estratégicas en un contexto empresarial.



3. FORMULACIÓN DE LAS ACTIVIDADES DE APRENDIZAJE

En este apartado se describirán las actividades de aprendizaje para la competencia que plantea el presente programa complementario, de la misma manera que se darán a conocer los tiempos y las evidencias a desarrollar.

En el siguiente cuadro se observa la distribución de tiempo de las evidencias que se realizarán durante las dos (4) semanas planeadas para la presente guía:

Semanas	1	2	3	4
Competencia				
220501114	Infografía: metodologías para el análisis de datos y su relación con los objetivos del negocio. AA1-EV01.	Video: descripción del funcionamiento de la herramienta <i>Python</i> para la manipulación y visualización de la información AA3-EV01. Informe en el que se identifiquen las variables y los componentes estadísticos a partir de una situación planteada. AA3-EV02. Presentación en la que dé a conocer los tipos de gráficos y su función en el análisis de datos. AA3-EV03.		
Horas x semana	8	8	12	4
220501114	Cuestionario sobre elementos básicos de la estadística y el proceso de preparación de datos. AA2-EV01.			Informe de resultados obtenidos del análisis de datos exploratorio realizado al caso de estudio. AA4-EV01..
Horas x semana	4	4		8
Total de horas	12	12	12	12

Como requisito para el desarrollo del presente programa de formación, es importante que usted conozca el espacio de trabajo y poder utilizar las herramientas necesarias en el LMS desde su rol de aprendiz. Además, se le sugiere realizar las siguientes acciones:

- ✓ Actualización de los datos personales.
- ✓ Lea el documento **“Información del programa”**, le permitirá reconocer los objetivos, las actividades y la metodología que se llevará a cabo en el curso.
- ✓ Lea el documento **“Cronograma”**, le permitirá comprender la planeación diseñada para lograr de manera secuencial los objetivos del curso. Tenga en cuenta que entender los pasos y reconocer las actividades a realizar le permitirá tener un estimado del tiempo de dedicación y, así, planear el cumplimiento de los requisitos de acuerdo con su tiempo disponible.
- ✓ Lea el presente documento **“Guía de aprendizaje”**, le orientará en el desarrollo de las actividades.
- ✓ Adicionalmente, realice la presentación en el foro social de acuerdo con las indicaciones de su instructor.

3.1. Actividad de aprendizaje de la competencia 220501114 - Sistematizar datos masivos de acuerdo con métodos de analítica y herramientas tecnológicas.



Las actividades planteadas para esta competencia permiten fortalecer en el aprendiz la capacidad de identificar, analizar y seleccionar datos de acuerdo con las necesidades y los requerimientos organizacionales, así como el reconocimiento de metodologías y herramientas necesarias para llevar a cabo dicho proceso.

3.1.1 Actividad de aprendizaje AA1 - Comprender las metodológicas para el análisis de datos y la identificación de los objetivos de negocio.

Esta actividad de aprendizaje posibilita que se logre identificar las metodologías necesarias para llevar a cabo un proyecto de analítica de datos según los objetivos y necesidades de determinado negocio.

Duración: 8 horas.

Materiales de formación a consultar: para desarrollar esta actividad adecuadamente, se recomienda revisar los contenidos presentados en el componente de formación: “**Análisis de datos: el futuro de las organizaciones.**”

Evidencias: a continuación, se describen las acciones y la correspondiente evidencia que conforma la actividad de aprendizaje:

❖ Evidencia AA1-EV01. Infografía: metodologías para el análisis de datos y su relación con los objetivos del negocio.

El objetivo de esta actividad es que a partir del análisis de la metodología ASUM-DM que se utiliza en proyectos de ciencia de los datos, y la herramienta *Desing Thinking* para el entendimiento del negocio, se realice una infografía en la que se describa el uso de cada una de estas técnicas.

Tenga en cuenta que en una infografía puede usarse la combinación de imágenes y textos, los cuales se organizan con el objetivo de explicar de manera concreta un tema, empleando apoyo visual.

En ella debe ser evidente lo siguiente:

- Los pasos más importantes para implementar la metodología ASUM-DM, resaltando las características más importantes de dicha metodología.
- Defina claramente los elementos a tener en cuenta cuando se utiliza *Desing Thinking*, plantee ejemplos.

Lineamientos generales para la entrega de la evidencia:

- **Producto por entregar:** infografía en la que se describan los elementos fundamentales de la metodología ASUM-DM.
- Debe tener las siguientes partes:
 - ✓ Título.
 - ✓ Texto.
 - ✓ Imágenes, iconos y otros elementos gráficos a color que apoyen con claridad la temática tratada.



- ✓ Desarrollo.
- ✓ Fuentes.
- Se deben usar imágenes, iconos o gráficos No puede presentar errores ortográficos y de puntuación. En el siguiente enlace podrá encontrar ejemplos de infografías de comparación que le servirán como guía para el desarrollo de la actividad planteada: https://www.canva.com/es_co/plantillas/?query=infograf%C3%ADas-comparacion
- **Extensión:** mínimo 2 páginas, máximo 3. Debe presentarse en hoja tamaño carta y siguiendo los lineamientos del numeral 1. Estilo, del instructivo APA, séptima edición (disponible en el Sistema de Bibliotecas del SENA, en el enlace <https://biblioteca.sena.edu.co/images/PDF/InstructivoAPA.pdf>)
- **Formato:** PDF.
- Para hacer el envío de la evidencia remítase al área de la actividad correspondiente y acceda al espacio: **Infografía: metodologías para el análisis de datos y su relación con los objetivos del negocio. AA1-EV01.**

3.1.2 Actividad de aprendizaje AA2 - Entender los conceptos básicos de la estadística.

El objetivo de esta actividad es comprender cada uno de los procesos necesarios para la preparación de datos, además de los componentes de la estadística como medio fundamental para el análisis exploratorio de datos, esta actividad de aprendizaje será evaluada mediante un cuestionario en línea, directamente en la plataforma.

Duración: 8 horas.

Materiales de formación a consultar: para desarrollar esta actividad adecuadamente, se recomienda revisar los contenidos presentados en el componente de formación: “**Análisis de datos: el futuro de las organizaciones**”

Evidencias: a continuación, se describen las acciones y las correspondientes evidencias que conforman la actividad de aprendizaje:

- ❖ **Evidencia AA2-EV01. Cuestionario sobre elementos básicos de la estadística y el proceso de preparación de datos.**

Este cuestionario tiene como fin valorar los conocimientos que se tienen respecto a los elementos de la estadística, el proceso de preparación de datos, y las técnicas a emplear para su normalización, conocimientos que resultan fundamentales para el óptimo desarrollo de un análisis de datos.

Lineamientos generales para la entrega de la evidencia:

- **Producto a entregar:** cuestionario resuelto en su totalidad en la plataforma.
- **Extensión:** el cuestionario dispuesto en la plataforma establecerá el tiempo de desarrollo, de acuerdo con el tipo y número de preguntas establecidas.
- Para realizar la prueba, acceda al **Cuestionario sobre elementos básicos de la estadística y el proceso de preparación de datos AA2-EV01.**

3.1.3 Actividad de aprendizaje AA3 - Implementar las herramientas y librerías necesarias para el análisis de los datos



El objetivo de la actividad aquí propuesta es el reconocimiento de los elementos necesarios para la implementación de las herramientas y librerías basadas en *Python*, las cuales son fundamentales para el análisis y procesamiento de los datos en una organización.

Duración: 24 horas.

Materiales de formación a consultar: para desarrollar esta actividad adecuadamente, se recomienda revisar los contenidos presentados en el componente de formación: “**Recursos y herramientas para el análisis efectivo de datos.**”

Evidencias: a continuación, se describen las acciones y las correspondientes evidencias que conforman la actividad de aprendizaje:

- ❖ **Evidencia AA3-EV01. Video: descripción del funcionamiento de la herramienta *Python* para la manipulación y visualización de la información.**

El objetivo de esta actividad es que el estudiante comprenda en detalle cómo funciona la herramienta *Python* para la manipulación y visualización de la información y describa por medio de un video, el paso a paso requerido para desplegar su entorno de trabajo, de igual manera, que establezca los requerimientos mínimos del sistema, para que el funcionamiento de la herramienta sea el óptimo, en el componente formativo encontrará diversas ayudas multimediales que le ayudarán con la puesta en marcha y los primeros pasos en el entorno de desarrollo.

Elementos para tener en cuenta para la elaboración del video:

- El video debe mostrar la herramienta Anaconda y Jupiter Lab instaladas correctamente.
- El video debe mostrar la instalación de las librerías explicadas en el componente formativo.
- El video debe mostrar los ejemplos que se realizan en el componente formativo.

Lineamientos para la entrega del video:

- **Producto por entregar:** video narrado en el que se explique la herramienta anaconda y Júpiter LAB, el video debe mostrar que estas herramientas estén instaladas correctamente, debe explicar el uso de esta herramienta, con una introducción de máximos 2 minutos, posteriormente el video debe exponer diferentes ejemplos.
- **Extensión:** máximo diez (10) minutos.
- **Formato:** vídeo MP4 o en formato que permita su visualización de manera sencilla en cualquier reproductor de video.
- Para hacer el envío de la evidencia remítase al área de la actividad correspondiente y acceda al espacio: **Video: descripción del funcionamiento de la herramienta *Python* para la manipulación y visualización de la información. AA3-EV01.**

- ❖ **Evidencia AA3-EV02. Informe en el que se identifiquen las variables y los componentes estadísticos a partir de una situación planteada.**



El objetivo de esta evidencia de aprendizaje es que, a partir del análisis de un caso puntual, se identifiquen las variables que se pueden presentar en el proceso de recolección de datos, además de establecer el proceso que se debe realizar en caso de que las variables presenten errores, no contengan información, o se encuentren en un formato distinto.

Por otro lado, se logrará comprender cómo realizar un análisis univariado y multivariado implementando los elementos propios de la estadística mediante las herramientas de análisis de datos de *Python*.

Para realizar esta actividad debe tener en cuenta lo dispuesto en el componente formativo, y leer con atención el caso anexo propuesto.

Para la elaboración del informe tenga en cuenta los siguientes aspectos:

- Debe tener configurado correctamente el entorno de desarrollo Anaconda y Jupyter LAB.
- Descargar de los anexos la data para el análisis del caso propuesto.
- No tomar pantallazos completos, solo de la zona requerida para la explicación, así evita tener muchas hojas en el informe.
- Explicar el funcionamiento de cada pantallazo.
- En anexos se encuentra el documento “Caso_Propuesto”, en el cual se explican los requerimientos para realizar el informe, de igual manera se provee de la data para el análisis del caso “Data_Caso_Propuesto”.

Lineamientos para la entrega del informe:

- **Producto por entregar:** informe detallando el paso a paso a través de pantallazos en los códigos utilizados para realizar las diferentes acciones como, por ejemplo: lectura de archivos, visualización de los datos ordenamientos de datos, selección de variables, métodos de corrección de variables, aplicación de los métodos estadísticos a los datos, y demás funciones que se consideren necesarias para complementar el informe.
- Incluir portada con los datos del aprendiz.
- **Extensión:** máximo diez (10) páginas.
- **Formato:** PDF.
- Para hacer el envío de la evidencia remítase al área de la actividad correspondiente y acceda al espacio: **Informe en el que se identifiquen las variables y los componentes estadísticos a partir de una situación planteada. AA3-EV02.**

- ❖ **Evidencia AA3-EV03. Presentación en la que dé a conocer los tipos de gráficos y su función en el análisis de datos.**

Esta evidencia de aprendizaje busca que se reconozcan los tipos de gráficos y su aplicación para el análisis de datos.

Elementos para tener en cuenta para la elaboración de la presentación:

- Descargue las imágenes que se generan directamente en la plataforma para que realice su presentación.
- Instale las librerías necesarias para la creación de gráficos.
- Explique el uso de cada gráfica y los resultados obtenidos.



- Copie el segmento de código utilizado para la generación de la gráfica.

Lineamientos para la entrega de la presentación:

- **Producto por entregar:** presentación en donde se explique la funcionalidad de cada uno de los gráficos que se pueden utilizar para el análisis de datos.
- Incluir Portada con los datos del aprendiz.
- **Extensión:** máximo diez (10) diapositivas
- **Formato:** PPT
- Para hacer el envío de la evidencia remítase al área de la actividad correspondiente y acceda al espacio **Presentación en la que dé a conocer los tipos de gráficos y su función en el análisis de datos. AA3-EV03.**

3.1.4 Actividad de aprendizaje AA4 - Establecer los resultados obtenidos frente al análisis exploratorios de datos.

El objetivo de esta actividad de aprendizaje es comprender los pasos necesarios para realizar las validaciones de los datos y las valoraciones respectivas en la aprobación o desaprobación de las hipótesis planteadas, se identificará como aplicar la estadística para el análisis de varianzas y significancias, se reconocerá la mejor ruta para encontrar respuestas a las preguntas que pueden surgir al iniciar la exploración de los datos y tomar decisiones asertivas que favorezcan a determinada organización.

Duración: 8 horas.

Materiales de formación a consultar: para desarrollar esta actividad adecuadamente, se recomienda revisar los contenidos presentados en el componente de formación: **“Recursos y herramientas para el análisis efectivo de datos”**.

Evidencias: a continuación, se describen las acciones y la correspondiente evidencia que conforma la actividad de aprendizaje:

- ❖ **Evidencia AA4-EV01. Informe de resultados obtenidos del análisis de datos exploratorio realizado al caso de estudio.**

Esta evidencia busca que entienda cómo los resultados obtenidos a partir de análisis de varianzas y diversas pruebas que se verán en el componente formativo. La toma de decisiones a partir del análisis de datos sirve en la verificación respecto a las preguntas o hipótesis planteadas del caso de estudio que se viene desarrollando, se busca que se identifiquen los elementos para validar el modelo utilizado con respecto a los objetivos que se pudiesen haber planteado.

Para el desarrollo de esta actividad debió haber comprendido cada uno de los momentos anteriores en donde se realiza el análisis del negocio, pasando por la preparación, transformación y visualización de los datos, al final en este componente se deben implantar técnicas estadísticas que validen las hipótesis planteadas.

Elementos para tener en cuenta para la elaboración del informe:

- Continuar trabajando con base en los datos del caso de estudio planteado en AA3.



- Aplicar las técnicas y pruebas estadísticas para la validación de hipótesis.

Lineamientos para la entrega del Informe:

- **Producto por entregar:** informe ejecutivo donde presente los objetivos del negocio del caso de estudio planteado, y en el que se explique cómo realizó la preparación de los datos, incluyendo las gráficas que generó el análisis, y finalmente se explique si los resultados obtenidos tienen relación con respecto a los objetivos del negocio.
- **Extensión:** máximo diez (10) páginas.
- **Formato:** PDF.
- Para hacer el envío de la evidencia remítase al área de la actividad correspondiente y acceda al espacio: **Informe de resultados obtenidos del análisis de datos exploratorio realizado al caso de estudio. AA4-EV01.**

4. ACTIVIDADES DE EVALUACIÓN

Evidencias de Aprendizaje	Criterios de Evaluación	Técnicas e Instrumentos de Evaluación
<p>Evidencia de desempeño:</p> <p>Infografía: metodologías para el análisis de datos y su relación con los objetivos del negocio. AA1-EV01.</p>	<p>Recopila información acerca de la situación actual de la empresa o negocio objeto de estudio de acuerdo con los requerimientos técnicos.</p> <p>Identifica los objetivos de negocio de la empresa u organización del objeto de análisis según planteamiento de preguntas.</p> <p>Identifica fuentes de datos de acuerdo con la técnica seleccionada.</p> <p>Prepara los datos para su análisis en busca de errores y datos faltantes según estructura de datos.</p> <p>Define las preguntas objeto de análisis de acuerdo con el objeto de estudio.</p> <p>Determina la población, muestra, variables y demás componentes estadísticos requeridos para el análisis de los datos según procesos estadísticos.</p> <p>Selecciona un tipo de grafico que permita visualizar los datos y variables analizadas.</p>	<p>Lista de chequeo:</p> <p>IE-AA1-EV01</p>



Evidencias de Aprendizaje	Criterios de Evaluación	Técnicas e Instrumentos de Evaluación
<p>Evidencia de conocimiento:</p> <p>Cuestionario sobre elementos básicos de la estadística y el proceso de preparación de datos. AA2-EV01.</p>	<p>Recopila información acerca de la situación actual de la empresa o negocio objeto de estudio de acuerdo con los requerimientos técnicos.</p> <p>Identifica los objetivos de negocio de la empresa u organización del objeto de análisis según planteamiento de preguntas.</p> <p>Identifica fuentes de datos de acuerdo con la técnica seleccionada.</p> <p>Prepara los datos para su análisis en busca de errores y datos faltantes según estructura de datos.</p> <p>Define las preguntas objeto de análisis de acuerdo con el objeto de estudio.</p> <p>Determina la población, muestra, variables y demás componentes estadísticos requeridos para el análisis de los datos según procesos estadísticos.</p> <p>Selecciona un tipo de grafico que permita visualizar los datos y variables analizadas.</p>	<p>Cuestionario</p> <p>IE-AA2-EV01</p>
<p>Evidencia de desempeño:</p> <p>Video: descripción del funcionamiento de la herramienta <i>Python</i> para la manipulación y visualización de la información. AA3-EV01.</p>	<p>Define el método para hacer el análisis exploratorio de los datos de acuerdo con el tipo de análisis de datos.</p> <p>Prepara los datos recopilados de diferentes fuentes ordenando e indexando los datos según ordenamiento de los datos.</p>	<p>Lista de chequeo:</p> <p>IE-AA3-EV01</p>
<p>Evidencia de desempeño:</p> <p>Informe en el que se identifiquen las variables y los componentes estadísticos a partir de una situación planteada. AA3-EV02.</p>	<p>Prepara el dataset o base de datos a ser analizada de acuerdo con la estructura de los datos.</p> <p>Determina las variables objeto de análisis según planteamiento de preguntas.</p> <p>Selecciona una técnica de modelado de acuerdo con los requerimientos solicitados.</p>	<p>Lista de chequeo:</p> <p>IE-AA3-EV02</p>



Evidencias de Aprendizaje	Criterios de Evaluación	Técnicas e Instrumentos de Evaluación
Evidencia de conocimiento: Presentación en la que dé a conocer los tipos de gráficos y su función en el análisis de datos. AA3-EV03.	Genera el modelo de análisis de acuerdo con la técnica de modelado de datos utilizada. Prepara la infraestructura tecnológica a ser utilizada en el modelo propuesto de acuerdo con las herramientas para análisis y visualización de datos. Genera el tipo de gráfico que permite visualizar resultados del análisis.	Lista de chequeo: IE-AA3-EV03
Evidencia de desempeño: Informe de resultados obtenidos del análisis de datos exploratorio realizado al caso de estudio. AA4-EV01.	Evalúa el modelo de análisis construido según evacuación de resultados. Comprueba el resultado del modelo vs los objetivos de la organización de acuerdo con el análisis aplicado. Verifica si el resultado del modelo de análisis da respuesta a las preguntas de análisis planteadas para la empresa u organización.	Rúbrica: IE-AA4-EV01

5. GLOSARIO DE TÉRMINOS

ANOVA: análisis de varianza.

Anaconda: herramienta Utilizada para la instalación de Jupyter LAB que permite la ejecución del código *Python*.

Datos: se utiliza para representar una medida o una descripción, y que en grandes volúmenes se usan para el análisis y generación de nueva información.

Gráficos: visualización de información agrupada que permite tener mejores análisis de la información.

Jupyter LAB: libros que permiten al usuario ejecutar código de *Python* de manera sencilla en bloques de código y a través de una interfaz gráfica.

Procesamiento de datos: lecturas de archivos, transformación y visualización de los datos requeridos para analizar o generar nueva información.

Python: lenguaje de programación de alto nivel que se utiliza entre otras cosas en el análisis y procesamiento de información de grandes volúmenes de datos.



6. REFERENTES BIBLIOGRÁFICOS

Alonso, J. Quintero, L. (2021). Guía de buenas prácticas para la mitigación del riesgo de modelo de analítica. Editorial Universidad Icesi. DOI: <https://doi.org/10.18046/EUI/bda.g.1>

Stanford University. (2021) *Explore the Stanford D.School*. <https://dschool.stanford.edu/>

Walpole, R., Myers, R., Myers, S., Keying, L. (2012). Probabilidad y estadística para ingenieros y ciencias. PEARSON.

7. CONTROL DEL DOCUMENTO

	Nombre	Cargo	Dependencia	Fecha
Autor (es)	Ronald Alexander Vacca Ascanio	Experto temático	Regional Distrito Capital. Centro De Diseño Metrología	Marzo 2022
	Jeimy Lorena Romero Perilla	Diseñador Instruccional	Regional Norte de Santander. Centro de la industria, la empresa y los servicios	Marzo 2022
	María Inés Machado López	Diseñador Instruccional	Regional Norte de Santander. Centro de la industria, la empresa y los servicios	Marzo 2022
	Carolina Coca Salazar	Asesora Metodológica	Regional Distrito Capital. Centro de Diseño y Metrología	Marzo 2022
	Rafael Neftalí Lizcano Reyes	Responsable Equipo Diseño Curricular	Regional Santander. Centro Industrial del Diseño y la Manufactura	Abril 2022
	Jhon Jairo Rodríguez Pérez	Corrector de estilo	Regional Distrito Capital. Centro de Diseño y Metrología	Abril 2022

8. CONTROL DE CAMBIOS (diligenciar únicamente si realiza ajustes a la guía)

	Nombre	Cargo	Dependencia	Fecha	Razón del Cambio
Autor (es)					