**CIENTIFICO DE DATOS**

**SISTECREDITO SAS**

**Proceso de Selección**

**Nombre Completo: DIEGO FERNANDO OSORIO DIAZ**

**Numero Identificación: 1010238755**

**Observaciones:**

* Use este documento para incluir la mayor cantidad de la información textual y gráfica que responda a las preguntas.
* Es posible que requiera utilizar herramientas externas. Incluya los enlaces de los artefactos externos, o los archivos adjuntos correspondientes dentro del correo. Se recomienda que pueda respaldar los distintos hallazgos con su debido soporte.
* Los siguientes casos, corresponden a la mezcla de información cierta y ficticia, construidos para fines de medir ciertos conocimientos y habilidades requeridos para el cargo en mención. Los datos que se proveen para el ejercicio no representan de ninguna forma a los clientes de Sistecredito SAS.
* No hay una extensión mínima o máxima.
* Guarde este documento en formato PDF, cuyo nombre debe tener la siguiente estructura: “<primer\_nombre>\_<primer\_apellido>\_casos”. Por ejemplo: ricardo\_sanchez\_casos.
* Para cualquier otro archivo, utilice la siguiente estructura: “<primer\_ nombre>\_<primer\_ apellido>\_<alias\_archivo>”. Por ejemplo: ricardo\_sanchez\_depuracion\_datos.
* Envíe todos los documentos y soportes en un único correo, con el asunto: “Nombre del aspirante\_CientificoDatos”, a los siguientes correos: [lmendoza@sistecredito.com](mailto:lmendoza@sistecredito.com), [yguzman@sistecredito.com](mailto:yguzman@sistecredito.com) . En el correo, explique brevemente el contenido de los archivos adicionales, si llega a tenerlos.
* Si tiene alguna duda o solicitud al respecto, por favor comunicarse al correo: [lmendoza@sistecredito.com](mailto:lmendoza@sistecredito.com).
* **Plazo límite para enviar la prueba: miércoles 06 de diciembre 2023, 06:00 pm**

# PRUEBA: CIENTÍFICO DE DATOS

1. DATOS:

Este conjunto de datos presenta una compilación de información relevante para analizar las ventas de productos en tiendas de Ecuador. Los datos de entrenamiento se componen de series temporales que registran las ventas de distintas familias de productos, promociones activas para esa fecha, y la identificación de tiendas y productos, entre otras. Adicionalmente, se incluyen archivos que complementan con metadatos que pueden ser cruciales para el desarrollo de modelos predictivos robustos y confiables, tales como información exógena que influye sobre las ventas en las tiendas, por ejemplo, el precio del petróleo para una fecha determinada y un calendario detallado de festividades y eventos.

1. OBJETIVOS:

El reto del candidato consiste en usar los datos para entender las dinámicas en las diferentes familias de productos al realizar un análisis de estos. Además, se debe construir un modelo predictivo que estime las ventas futuras, donde se espera que demuestre habilidades en el manejo, limpieza y análisis de series temporales, así como en la aplicación de métodos estadísticos y de machine learning. Es clave que el candidato exhiba competencias para entender cómo influencian las variables externas a las tendencias de ventas y cómo adaptar sus modelos a estas dinámicas. Más allá de buscar la mejor solución de manera inmediata, se valorará el razonamiento crítico del candidato para proponer mejoras continuas del modelo, identificando falencias en los aspectos técnicos, prácticos o teóricos que puedan afectar los resultados.

1. RESULTADOS:

El candidato finalizará esta prueba entregando un informe y una presentación que explique el problema, insights obtenidos de los datos, las metodologías aplicadas, el porqué de la selección de modelos o algoritmos aplicados, la relevancia de las variables escogidas y el análisis de los resultados obtenidos. Además, deberá identificar y proponer estrategias para mejorar el modelo, discutiendo las limitaciones encontradas y cómo éstas pudieron influir en los resultados. La presentación deberá ser clara y comprensible, capaz de traducir términos técnicos a un contexto de negocio, permitiendo entender cómo los insights obtenidos pueden ser aplicados para optimizar las decisiones comerciales. Se espera que el candidato muestre la habilidad de reflexionar sobre su trabajo, manifestando una comprensión profunda de la ciencia de datos no solo como una disciplina técnica, sino también como una herramienta para la toma de decisiones estratégicas.