

Programador Junior en Machine Learning I



**Argentina
programa
4.0**



**Ministerio de
Desarrollo Productivo
Argentina**

**Secretaría de
Economía del Conocimiento**

Semana 4

- + **Introducción al análisis de datos**
 - + ¿Qué es el análisis de datos?
 - + Etapas del análisis de datos: recolección, procesamiento, análisis y visualización
 - + Tipos de análisis de datos: descriptivo, exploratorio, inferencial y predictivo
- + **Introducción a Python para el análisis de datos**



Introducción al análisis de datos

¿Qué es el análisis de datos?

El análisis de datos es el proceso de examinar, transformar y modelar datos con el fin de extraer información útil, generar conclusiones y apoyar la toma de decisiones. El análisis de datos se puede aplicar a diversos campos, como la ciencia, la ingeniería, la medicina, las finanzas, el marketing, la educación y otros. El análisis de datos implica el uso de técnicas estadísticas, matemáticas, computacionales y de visualización para explorar y comprender los datos. El análisis de datos se puede clasificar en diferentes tipos según el objetivo y el método que se emplee, como el análisis descriptivo, el análisis inferencial, el análisis predictivo y el análisis prescriptivo.

Etapas del análisis de datos: recolección, procesamiento, análisis y visualización

El análisis de datos es un proceso que consiste en transformar los datos en bruto en información útil para la toma de decisiones. Para ello, se siguen las siguientes etapas:

Recopilación de datos: se obtienen los datos de las fuentes disponibles, como bases de datos, archivos, encuestas, etc. Se debe asegurar que los datos sean de calidad, es decir, completos, precisos, relevantes y actualizados.

Preparación de datos: se limpian y organizan los datos, eliminando los errores, las inconsistencias, los duplicados y los valores faltantes. Se seleccionan los datos que se van a analizar y se transforman en un formato adecuado para el análisis.

Análisis de datos: se aplican técnicas estadísticas y matemáticas para explorar, describir, modelar y predecir los datos. Se pueden utilizar herramientas como tablas dinámicas, gráficos, software especializado, etc. Se busca identificar patrones, tendencias, relaciones y anomalías en los datos.

Acción: se interpretan los resultados del análisis y se extraen conclusiones y recomendaciones. Se comunican los hallazgos a las personas o entidades interesadas, utilizando medios como informes, presentaciones, dashboards, etc. Se toman decisiones basadas en la evidencia y se evalúan sus impactos.

Tipos de análisis de datos: descriptivo, exploratorio, inferencial y predictivo

Los datos son valiosos para mejorar procesos, productos y servicios. Para aprovecharlos al máximo, se utilizan cuatro tipos de análisis: descriptivo, exploratorio, inferencial y predictivo.

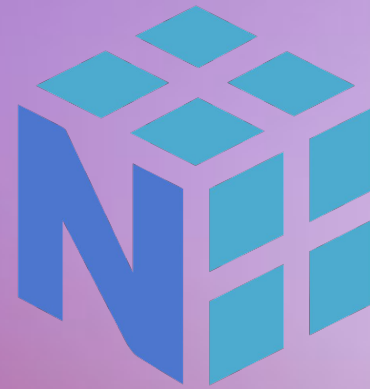
El **análisis descriptivo** resume los datos mediante tablas, gráficos y medidas estadísticas. No busca explicar causas o consecuencias, solo describe lo que ha ocurrido o está ocurriendo.

El **análisis exploratorio** descubre patrones, tendencias y relaciones ocultas sin partir de hipótesis previas. Genera nuevas preguntas e hipótesis explorando los datos.

El **análisis inferencial** prueba o refuta hipótesis sobre una población a partir de una muestra. Estima la confianza o significación de los resultados e infiere lo que puede ocurrir en casos no observados.

El **análisis predictivo** utiliza datos históricos o actuales para predecir el futuro. Construye modelos predictivos mediante técnicas estadísticas o algoritmos de aprendizaje automático.

Introducción a Python para el análisis de datos



NumPy

matplotlib

