



Programación 2

Tecnicatura Universitaria en Inteligencia Artificial

2024

TRABAJO PRÁCTICO FINAL DE PANDAS

1. Análisis de Datos de la línea 144 utilizando Pandas

1.1. Objetivo

El objetivo de este trabajo práctico es que los estudiantes adquieran experiencia en la manipulación y análisis de datos utilizando la librería pandas en Python, aplicando conceptos de programación y estadísticas a un conjunto de datos de la línea 144 utilizando Pandas.

1.2. Conjunto de Datos

La línea 144 brinda atención, contención y asesoramiento a personas en situación de violencia de género. Esta línea fue creada en 2013 para cumplir con la ley N° 26.485 (más información en: <https://www.argentina.gob.ar/capital-humano/generos/linea-144>).

Se dispone de archivos de datos sobre las llamadas recibidas por este número. Cada registro representa una llamada e incluye información como la fecha, la provincia, el género de la persona que llama, su edad, su nacionalidad, entre otros datos. Los mismos pueden consultarse y descargarse en el sitio <https://www.datos.gob.ar/dataset/generos-base-datos-linea-144>. En el mismo sitio pueden consultarse los campos con sus tipos de datos y descripción.

En este trabajo práctico se analizarán los datos correspondientes a los años 2020 al 2022 (no incluiremos los datos del año 2023 por no estar completos).

1.3. Descripción del Trabajo

1. **Carga de Datos:** Los estudiantes deben cargar el conjunto de datos en un `DataFrame` de pandas.
2. **Limpieza de Datos:** Los estudiantes deben realizar una limpieza inicial de los datos, incluyendo un análisis de valores faltantes y la identificación de posibles errores.
3. **Análisis Descriptivo:** Los estudiantes deben realizar un análisis descriptivo de los datos, calculando estadísticas resumen (media, mediana, desviación estándar, etc.) para las variables relevantes. También deben visualizar los datos utilizando boxplot. ¿Existen valores atípicos?

4. **Análisis Temporal:** Los estudiantes deben realizar un análisis temporal de los datos, calculando la evolución de las llamadas a lo largo del tiempo. ¿Hubo aumento de llamadas? ¿Hubo un cambio en las edades de quienes denuncian?
5. **Preguntas de Investigación:** Plantear al menos tres preguntas de investigación relacionadas con los datos y utilizar pandas para responderlas. Por ejemplo, ¿cuál fue la edad promedio de los denunciantes de una provincia y un año específicos? ¿Cuál fue la provincia desde donde se realizaron más llamadas?

1.4. Entrega

Los estudiantes deben entregar el código producido, las visualizaciones y las respuestas a las preguntas de investigación en formato `.ipynb`. **Explicar de forma apropiada todos los resultados.**

1.5. Evaluación

El trabajo práctico se evaluará en función de la calidad del código, la limpieza y el análisis de datos, la presentación de resultados y la capacidad de los estudiantes para responder a las preguntas de investigación.

Fecha de entrega martes 13/08/2024 a las 23.59 (horario servidor FCEIA).