

Examen Ciencia de Datos L1/L2

Observaciones:

- Envía tus respuestas a más tardar 48 horas después de recibir el correo.
- Las respuestas podrán ser enviadas en un PDF, un Jupyter Notebook o con un link a tu repositorio de GitHub.
- Ambas preguntas del examen son prácticas, utiliza Python para resolverlas.
- Tus respuestas deberán contener, en cada paso, una descripción del razonamiento y planteamiento que seguiste, el código con el que generaste tus resultados y tus resultados. Toma en cuenta que a nosotros nos interesa conocer, sobre todo, tu capacidad de comprensión y planteamiento del problema.
- Busca el 80-20 en cada pregunta (el 20% del trabajo que te da 80% del valor en la solución).

Sección A

Datos abiertos de la CDMX

La Agencia Digital de Innovación Pública tiene disponibles datos de las **carpetas de investigación aportados por la PGJ**. La tabla está disponible aquí: <https://datos.cdmx.gob.mx/dataset/carpetas-de-investigacion-fgj-de-la-ciudad-de-mexico/resource/48fcb848-220c-4af0-839b-4fd8ac812c0f>

Utilizando estos datos, responde las siguientes preguntas. Debes de desarrollar el código en Python:

1. ¿Qué pruebas identificarías para asegurar la calidad de estos datos? No es necesario hacerlas, sólo describe la prueba y lo que te dice cada una.
2. Identifica los delitos que van a la alza y a la baja en la CDMX (ten cuidado con los delitos con pocas ocurrencias).
3. ¿Cuál es la alcaldía que más delitos tiene y cuál es la que menos? ¿Por qué crees que sea esto?
4. ¿Existe alguna tendencia estacional en la ocurrencia de delitos (mes, semana, día de la semana, quincenas) en la CDMX? ¿A qué crees que se deba?
5. ¿Cuáles son los delitos que más caracterizan a cada alcaldía? Es decir, delitos que suceden con mayor frecuencia en una alcaldía y con menor frecuencia en las demás.
6. Diseña un indicador que mida el nivel de "inseguridad". Génalo al nivel de desagregación que te parezca más adecuado (ej. manzana, calle, AGEB, etc.). Analiza los resultados ¿Encontraste algún patrón interesante? ¿Qué decisiones se podrían tomar con el indicador?

Sección B

BOPS

Analiza este [caso](#) y responde las siguientes preguntas:

1. ¿Cuántos millones de dólares se ganaron o perdieron a causa del programa? ¿Deberían expandirse a Canadá? Explica tu razonamiento y metodología.

Pista: Existen dos experimentos naturales: Canadá y las zonas que se encuentran lejos de una tienda. Utilízalos.