

# **Reto del Datatón 2025**

## **Primera Fase**

*Tu nuevo hogar en México: recomendador inteligente de zonas*

### **Introducción**

El reto parte del siguiente escenario: un extranjero llega a la Ciudad de México con la intención de comprar o rentar vivienda, pero se enfrenta a un problema clave, ya que no conoce las características, ventajas ni desventajas de las distintas zonas de la ciudad. Para atender esta necesidad, el desafío consiste en construir una aplicación, página web o dashboard que le otorgue información oportuna y le de recomendaciones acerca de las mejores zonas para vivir. Dichas recomendaciones deben ser personalizadas de acuerdo con las prioridades y necesidades del usuario.

La solución que entregue la aplicación debe emplear datos abiertos y las fuentes de información que ustedes consideren más relevantes. Esto queda a juicio de cada equipo. Algunos ejemplos incluyen: indicadores socioeconómicos, niveles de seguridad, opciones de transporte, disponibilidad de servicios (por ejemplo: escuelas, mercados, parques, hospitales), niveles de contaminación u otros aspectos que a ustedes se les ocurran.

Su aplicación deberá combinar múltiples variables y entregar a los usuarios un sistema de recomendaciones claro y confiable. Cada equipo decidirá el formato que utilice su aplicación para entregar la información. Algunas alternativas incluyen mapas y listados. De esta manera, el usuario, sin importar su nivel de conocimiento previo sobre la ciudad, podrá recibir sugerencias concretas y bien fundamentadas que le ayuden a tomar una buena decisión acerca de dónde establecerse.

Lo que aprendas aquí te servirá para entrevistas de trabajo y proyectos reales. Muchos empleadores buscan justo estas habilidades: análisis de datos, programación y comunicación de resultados.

## Objetivo

Diseñar y desarrollar una aplicación interactiva que:

1. Reciba como entrada las características y prioridades del usuario (e.g. presupuesto, transporte, estilo de vida, seguridad, etc.). Cada equipo deberá decidir qué entradas incluir y en qué formato.
2. Ejecute un algoritmo utilizando dichas entradas y las fuentes de información que el equipo decida utilizar. El algoritmo procesará dichos datos para entregar una comparación o recomendación de distintas zonas de la ciudad.
3. Entregue una recomendación personalizada de 3 a 5 zonas ideales, acompañada de información relevante y fácil de interpretar. Cada equipo decidirá el formato y manera en que la aplicación entregará la información.

## Talleres

Para apoyarlos en el inicio del reto recuerden que habrá una serie de talleres durante la semana del 8-12 de septiembre a la hora de la comida para no afectar sus clases. La asistencia a los talleres es altamente sugerida, aunque faltar no elimina a los equipos. Sugerimos que al menos uno o dos miembros por equipo acudan para aprovechar este recurso y reconocer el apoyo y trabajo de los alumnos que prepararon este recurso para apoyarlos.

Taller	Lugar y hora	Tutor(a)
Uso de datos	Lu 8-sep, 2-4pm, RH-313	Nuria Alas
Estructurar el problema	Ma 9-sep, 2-4pm, Epic	Jana Boltvinik
Visualización de datos	Mie 10-sep, 2-4pm, RH-313	Fabián Orduña
Crear mapas	Jue 11-sep, 2.30-5pm, RH-301	Harry Gomar
PPTs efectivas	Vie 12-sep, 2-4pm, RH-301	Samari García

## Datos y fuentes sugeridas

Cada uno de los equipos deberá decidir las fuentes de información a utilizar. Se sugiere que los equipos empiecen con un diseño simple utilizando pocas fuentes de información y una vez que su aplicación ya funciones agreguen fuentes adicionales si el tiempo se los permite. Algunas fuentes de información que pudieran resultarles de utilidad incluyen:

- **INEGI.** Datos socioeconómicos, demográficos y comerciales. Tengan en consideración que los equipos que pasen a la segunda fase seguramente tendrán que utilizar datos de INEGI, así que sería útil irse familiarizando con ellos.
- **Datos Abiertos de la CDMX.** Pueden encontrar información de seguridad, transporte, servicios públicos y movilidad.
- **Portales inmobiliarios.** Pueden encontrar precios de compra y renta de vivienda, ya sea a partir de datasets descargados o mediante *web scraping*.
- **Google Places API / OpenStreetMap.** Información sobre puntos de interés como hospitales, escuelas, parques y supermercados.
- **SEMOVI.** Datos sobre rutas, tráfico y transporte público.
- **SEDEMA.** Datos de contaminación.

### Entradas posibles

La aplicación deberá solicitar al usuario proporcionar una lista de entradas para realizar la búsqueda de vivienda. A continuación se enlistan algunos ejemplos de información a solicitar. Los equipos pueden usar elementos de estos ejemplos u otros que se les ocurran. No es obligatorio incluir todos los elementos de la siguiente lista, de hecho se sugiere empezar con pocos elementos. Cada equipo deberá decidir qué información solicitar y el formato para recopilar dicha información.

- Presupuesto disponible (rango en pesos mexicanos).
- Características demográficas del usuario (por ejemplo: pareja, con/sin hijos, edades, etc.)
- Distancia máxima a trabajo, escuela o lugar de interés.
- Nivel de prioridad que le dan a la seguridad, acceso a transporte público, contaminación (alta, media o baja).
- Servicios clave requeridos (hospitales, escuelas, supermercados, áreas verdes).
- Estilo de vida preferido (por ejemplo: familiar, cultural, vida nocturna).
- Preferencia de densidad (zonas activas vs. zonas residenciales).

### Salidas esperadas

Los resultados que reciba el usuario deben ser claros, fáciles de entender y útiles. Para clasificar las zonas, deberán crear un puntaje que combine la información que es relevante para los usuarios de acuerdo a las entradas del paso anterior. Cada equipo podrá decidir cómo la aplicación entregará los resultados y la información que recibirán los usuarios. Algunos ejemplos:

- Lista de zonas recomendadas con un puntaje comparativo.

- Mapa interactivo que muestre dichas zonas.
- Fichas resumidas con las características principales de la(s) zona(s) mas recomendada(s).
- Rango estimado de precios de vivienda y características relevantes para el usuario de acuerdo a las entradas en la(s) zona(s) sugerida(s).

### **Requisitos técnicos mínimos**

- Algoritmo que calcule una puntuación y categorice las zonas según los datos y entradas de los usuarios.
- Interfaz amigable (web o móvil).
- Uso de al menos tres fuentes de datos distintas.
- Visualización geoespacial (mapa con marcadores o zonas coloreadas).

### **Evaluación de la primera fase**

La semana del 29 de septiembre al 3 de octubre se llevará a cabo la final de la primera fase. Se le asignará a cada equipo un horario entre 2.30 y 4.00pm para mostrar su solución propuesta. Esta evaluación consistirá en:

- 1) Presentación de **máximo 10 minutos** frente a un grupo de jueces. La presentación deberá explicar el desarrollo de su producto y los aspectos técnicos relevantes, incluyendo: (i) ¿cómo decidieron las entradas?, (ii) ¿cómo construyeron el puntaje para clasificar zonas?, (iii) cualquier otro aspecto técnico relevante. Los jueces serán miembros de la facultad de diversos departamentos académicos.
- 2) Mostrar que su aplicación, página web o dashboard funciona en tiempo real. Los jueces les solicitarán una búsqueda distinta y su solución debe funcionar y entregar resultados.
- 3) Resolver preguntas que tengan los jueces.

Criterios de evaluación:

- Calidad técnica en el método usado para clasificar las zonas, justificación de las entradas utilizadas.
- Calidad en el uso de datos: integración, limpieza y generación de información relevante.
- Experiencia de usuario: diseño visual, claridad, originalidad / creatividad y facilidad de navegación.
- Presentación del proyecto: narrativa, claridad y calidad de exposición.
- Ejercicio donde se muestre la interactividad de la plataforma en tiempo real.

