

# Datathon: guía del caso

## Instrucciones

¡Bienvenidos al Datathon! ¡Esperamos que esta jornada os sirva para pasarlo bien, aprender y acabéis sorprendiéndonos con vuestras propuestas!

Primero lo más importante.

El Datathon es el evento que tendrá lugar el próximo **miércoles 27 de octubre**. Ese día llegaréis allí y comenzaréis a trabajar en equipo con el objetivo de presentar vuestra solución del caso frente a un jurado experto, a partir de las 16 horas, con un máximo de 13 minutos por equipo (3' para proyectar la presentación y disponerse, **7' para presentar vuestra propuesta** y 3' para responder preguntas de los evaluadores).

Como es un desafío complejo, recibiréis los requisitos y las pautas con anticipación. También sabréis con antelación con qué compañeros compartiréis equipo, y de esa manera podréis organizaros, repartiros trabajo de investigación y diseñar una estrategia que tendrá como objetivo aprovechar al 100% el día del evento para trabajar en vuestra solución del caso y generar una presentación **impactante e innovadora**.

A partir de recibir estos requisitos y conocer a vuestros equipos, sois libres de planificar cómo resolverlo, pero no queremos equipos que tengan su caso resuelto antes del evento, ya que no podrán ser valorados en la dinámica del Datathon por las personas de que estarán ayudando y evaluando. Claro que podéis diseñar estrategias, hacer investigaciones, llevar cosas adelantadas, etc., pero **el día del evento es el reservado para ejecutar vuestra solución**.

¡Ahora sí! ¡Vamos al caso!

## El desafío

En este desafío los equipos deberán responder a una **RFI (Solicitud de Información)** por parte de una cadena de supermercados ficticia llamada **SuperCenter**.

Esta RFI, que simula un caso que una empresa como Capgemini podría recibir típicamente, describe la necesidad que el cliente trasmite a diferentes proveedores (en este caso vosotros, los equipos), quienes deberán competir para conseguir el proyecto.

Las consignas están claramente representadas en el documento RFI, aunque en este documento os damos algunos tips adicionales.

Es importante que cumpláis con cada paso:

1. **Enviar las preguntas y dudas** que podáis tener sobre lo que el cliente necesita en la fecha que se designa y utilizando la template provista. Las preguntas serán respondidas por el “cliente”, consolidadas y enviadas a todos los equipos por igual, para que todos tengan la misma información.
2. Participar (máximo 2 por grupo) de la **videollamada de kick off**. (será grabado y puesto a disposición de todos posteriormente). Vosotros decidís quienes participan de la videollamada.
3. **Participar como equipo completo** el día del Datathon como ya se ha mencionado arriba.
4. **Entregar lo solicitado** por el cliente en tiempo y forma.

## Evaluación

La solución como vais a ver en los otros documentos comprende un desarrollo tecnológico para resolver un problema de negocio del “cliente”.

La resolución de un problema de esta naturaleza la dejamos abierta al ingenio de los equipos, es decir que **no hay un solo método de resolución ni una única estrategia ganadora**.

Valoraremos por igual:

1. **Soluciones técnicas** que utilicen los datos provistos y puedan presentar una demo de sistemas de recomendación de productos utilizando alguna herramienta de programación (por ejemplo, programando en Python).
2. **Explicar conceptualmente la solución** sin necesidad de recurrir a programación (aunque sí tenéis que ser capaces de explicar cómo funciona el sistema, qué datos le alimentan, qué resultados produce el algoritmo o el software).
3. Hacer **uso de los datos provistos** en generar estadísticas, visualizaciones y todo tipo de análisis, aunque esto resultará un elemento complementario ya que necesariamente debe estar centrada la presentación en la solución propuesta.

En definitiva, todo dependerá de las habilidades de vuestros compañeros de equipo. Si entre las personas de su equipo alguna o varias poseen conocimientos de programación, os podéis animar a algo más técnico. Por el contrario, si no los hubiera, se puede optar por presentar un caso más enfocado a *negocio*.

Lo más importante es generar una presentación que transmita confianza, sea creíble y que impacte al “cliente” con su nivel de innovación y creatividad.

Si llegado el caso, algunos equipos están virtualmente empatados (según los criterios anteriores), tratándose de un Datathon, **se priorizará aquel que mejor tratamiento de los datos haya conseguido** (aunque no es una condición suficiente para ganar).

## Cuentas de Google Cloud

Como veréis en la RFI, la solución debe estar basada en la plataforma cloud de Google (Google Cloud Platform).

Para daros la oportunidad de conocer e incluso utilizar la plataforma GCP como parte del desafío, os daremos también acceso a cada uno de los participantes a una cuenta de equipo de GCP donde dispondréis de un crédito asignado para poder gastar en recursos si así lo quisiérais.

No es necesaria su utilización como parte de la solución, pero sí esperamos al menos un nivel de investigación y una propuesta que incluya componentes de Google en la ideación de la misma.

Las instrucciones para utilizar dichas cuentas serán provistas separadamente.



## Otras recomendaciones

- ✓ El desafío comienza **desde el momento en que estáis leyendo esta documentación** y concluye el día del evento.
- ✓ Como se mencionó anteriormente, esto quiere decir que sí se puede comenzar a resolverlo desde este momento, aunque debéis estar preparados para **trabajar el día del evento**, donde además podrían presentarse variaciones al caso. Si venís con un caso muy cerrado el día del evento, corréis el riesgo de que se actualice alguna consigna o dato y sea más difícil adaptarse. (¡es parte del juego!)
- ✓ No hay una recomendación acerca de los canales que podéis utilizar para organizaros como equipo. Es algo que decidís vosotros (grupo de Whatsapp, teams, email, etc.)
- ✓ La presentación al jurado la puede hacer un miembro, varios o todos, pero sí o sí tenéis que estar **todos presentes frente al jurado como equipo y todos preparados para responder a las preguntas** que se les pueden llegar a hacer.
- ✓ Los datos que os daremos son una herramienta que podéis utilizar para hacer toda clase de deducciones acerca de los productos, frecuencias, asociaciones, etc... Todo dato que necesitéis y no se haya provisto, es susceptible de ser “asumido” por el equipo. Esto es, si el equipo precisa más detalles financieros o de productos que no se han dado, se acepta que esos datos puedan ser inventados por el equipo, siempre y cuando tengan sentido con el resto del caso y siempre y cuando su invención sea un soporte para la propuesta y no un modo de obtener una ventaja.
- ✓ La RFI menciona en la página 3, sección 3.2, realizar una “demostración efectiva del funcionamiento de la solución”. Esto quiere decir que la presentación que hagáis en la defensa tiene que enseñar cómo, a través de unos *inputs*, el sistema genera recomendaciones de *otros productos similares o complementarios*, cuál es la lógica del sistema que permite generar esas recomendaciones y como mínimo una ilustración de su funcionamiento.
- ✓ En la sección 4.1 de la RFI se menciona que la solución debe estar basada en “Métodos analíticos avanzados, matemáticos y estadísticos”. Los sistemas de recomendación de productos son soluciones que se basan en los datos obtenidos y aplican métodos matemáticos /estadísticos para generalizar y extraer patrones. Como parte del reto esperamos que los equipos se instruyan acerca de qué clase de algoritmos son los que se pueden utilizar y sean capaces de transmitirlo brevemente en la presentación.
- ✓ Las cuentas de Google Cloud están para ser exploradas, aún si no las usáis efectivamente. Es importante que presentéis algún componente de Google Cloud como parte de la solución, y para eso podréis investigar qué sistemas de recomendación propone Google en sus diversas documentaciones disponibles en internet.
- ✓ En la sección de planificación es importante plantear al “cliente” de qué manera abordar el proyecto con mínimos riesgos. ¿Realizar una prueba de concepto antes que hacer el proyecto completo? ¿Hacer un experimento con ciertos clientes?, etc...