

# Data Project

Tratamiento del Dato

Curso 21-22 - Edición 3ª

# AGENDA

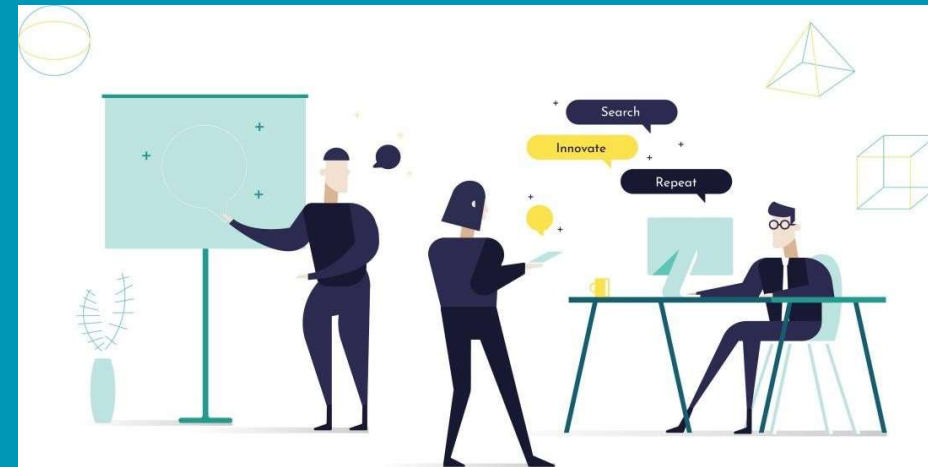
1. Objetivo
2. Metodología
3. Fechas
4. Evaluación
5. Contenido
6. Elementos a Entregar

# OBJETIVO

Un hackaton o data Project es una herramienta formativa que pretende colocar a los alumnos frente a una situación real dentro de un entorno de colaboración y confianza. Las decisiones erróneas servirán de aprendizaje al grupo..

En estas sesiones se presentará un caso de negocio para todos y por equipos deberá aportarse la mejor solución a nivel técnico y de negocio.

Este modelo permite al alumno colaborar con un equipo cercano a la realidad y hacer una presentación similar a lo que haría en una empresa.



**Objetivo...**  
**APRENDER**

# METODOLOGÍA

## Equipos:

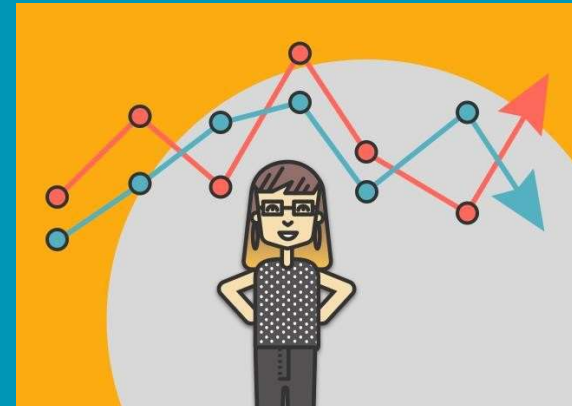
Los proyectos se presentarán en equipos de 5 personas, estos equipos serán compensados entre gente cercana a negocio y gente con una visión más técnica.

## Contenido:

El caso de uso se os presentará 2 semanas antes de la defensa del mismo, durante este tiempo habrá diferentes fechas de corte que tendréis que manejar

## Consultas de Negocio:

Se fijarán unas fechas donde podréis pregunta a los profesores dudas de negocio, que no tendrán penalización y que serán públicas al resto de compañeros. Esto permitirá aclarar dudas sobre el alcance del caso de uso.



## Preguntas

### Consultas Técnicas:

Durante la fase de desarrollo, los equipos podrán preguntar a los profesores para evitar bloqueos o errores. Estas preguntas, serán recibidas por los profesores que os dirán la penalización en nota en caso de contestarla.

**Equipo1:** Roberto, cómo ingesto un csv en hdfs?

**Roberto:** Enlace a web externa (coste 0,2)

Referencia de apuntes (coste 0,3)

Respuesta completa (coste 0,4)

**Equipo1:** Queremos la referencia a apuntes



# FECHAS

Fechas para el primer Data Project:

Publicación:	<b>13 – Enero</b>
Dudas Negocio (límite):	<b>20 – Enero</b>
Dudas Técnicas (límite):	<b>29 – Enero</b>
Implementación:	<b>2 – Febrero</b>
Entrega de Material:	<b>2 - Febrero</b>



# EVALUACIÓN

La solución aportada por cada uno de los equipos se evaluará en dos aspectos teniendo en cuenta los siguientes puntos:



## NEGOCIO

- Originalidad
- Presentación
- Viabilidad de negocio
- Adecuación al requerimiento



## TECNOLOGÍA

- Diseño de Arquitectura
- Idoneidad de las piezas
- Implementación Técnica
- Justificación de las decisiones

$$\text{Nota Final} = (\text{Nota Negocio} + \text{Nota de Tecnología}) / 2 - \text{Penalizaciones}$$

# USE CASE 1

Meta ha contactado con EDEM para la realización de un producto de cara a su nuevo metaverso. Para ello necesita informar a sus usuarios mediante proximidad de los mismos, usando coordenadas GPS.

La idea es que un usuario del metaverso reciba una alerta si un amigo se acerca en su radio de interacción.

Para ello han abierto una RFP pública en la que tendréis que presentar vuestro producto y que aporta esta dentro del metaverso de meta.





## USE CASE 2

**Zurich quiere validar su modelo de negocio de actividad de seguros personalizados. Para ello quiere diseñar un modelo que cuantifique la actividad física (aeróbica) de una persona en un día.**

**Para ello Zurich ha solicitado a EDEM un piloto en el que se vea cuánto ha caminado cada persona en kms y le asigne una puntuación o prima en relación a la media.**



**ZURICH®**

# ELEMENTOS A ENTREGAR

Los componentes del equipo tendrán que presentar los siguientes elementos como parte de la presentación:

- **Presentación:** Harán una presentación de 30 min explicando su solución
- **Diseño de Arquitectura:** Diagrama que explique las diferentes piezas que han utilizado
- **Justificación de piezas:** Justificación de las decisiones que han llevado a elegir unas herramientas frente a otras
- **Solución:** Deberán entregar una máquina virtual / Docker / Github, entorno web, donde se pueda ejecutar la solución
- **Vídeo:** Cada equipo deberá entregar un vídeo corto (<5 min) explicativo del Caso en youtube
- **Origen de Datos:** El equipo debe entregar una lista de los datasets explorados y razones por las que han sido incluidos o excluidos.
- **Modelos de Datos:** Diseño del modelo, junto con los atributos de las mismas, definiciones etc

# RECURSOS

Para la generación de datos ambos equipos partiréis de un github con un generador de datos aleatorios.

<https://github.com/a10pepo/datagenerator>

