

Data Project 2

Tratamiento del Dato

Curso 20-21 - Edición 2ª

Fecha 14/01/2021

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:

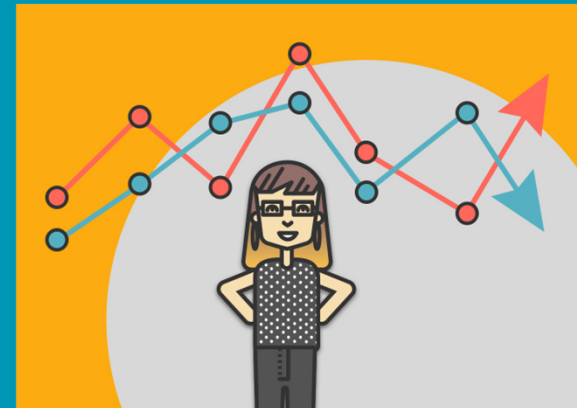
Los proyectos se presentarán en equipos de 5 personas, estos equipos serán compensados entre gente cercana a negocio y gente con una visión más técnica.

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 segundo Data Project:

Publicación:	14 – Enero
Dudas Negocio (límite):	21 – Enero
Dudas Técnicas (límite):	26 – Enero
Implementación:	5 – Febrero
Entrega de Material:	5 – Febrero



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

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

CONTENIDO



El tiempo ha pasado e QualityLive ha seguido creciendo y habéis pasado el primer corte, pero queda lo más difícil, vender la solución real.

QualityLive desea abrir una ronda de RFP a diferentes empresas, pero antes, como líderes de su equipo de IT os han pedido que hagáis una estimación con los siguientes puntos:

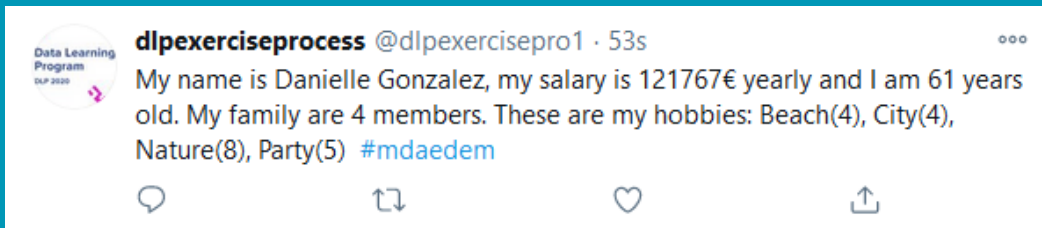
- Roles Necesarios (Tecnologías y Perfiles ejemplo LinkedIn)
- Coste desarrollo (MD)
- Tiempo necesario para puesta en marcha
- Coste de Mantenimiento
- Desarrollo MVP

REQUERIMIENTO

QualityLive apuesta por las redes sociales para la implementación de su caso de uso por lo que os ha pedido lo siguiente como parte de la estimación y del MVP:

- 1) Leer la información de la cuenta de twitter y responder acorde a esta información:
- 2) Debéis leer los mensajes con el #mdaedem

dlpexerciseprocess
@dlpexercisepro1



- 3) Una vez leídos los tweets debéis cruzar la información entre ambos, es decir, encontrar el mejor match de las opciones disponibles y responder al tweet, en el menor tiempo posible.
- 4) En la respuesta podéis incluir tanta información como consideréis (fotos ciudad, Wikipedia, enlaces...), pero como mínimo el Code de la vivienda que le ofrecéis.

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, entorno web, donde se pueda ejecutar la solución
- **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.
- **Github & Youtube**