

El objetivo de este proyecto:

Adalab nos ha vuelto a contratar, en concreto quieren seguir ampliando su contenido y evaluar como es el perfil de una analista de datos. Para eso nos han pasado una serie de datos recogidos de una encuesta realizada por Kaggle. En esta encuesta se recopila información sobre un montón de encuestados que trabajan como analistas de datos en todo el mundo.

Entre las preguntas que tenían que contestar los encuestados tenemos información sobre que perfil tienen (analistas de datos, científicos de datos, ingeniero de datos, product owner etc.) que herramientas de visualización usan, cuantos años de experiencia tienen como analistas de datos, cuál es su nivel de formación, etc.

Basándonos en toda esta información deberemos decirle a Adalab como se encuentra ahora mismo el mercado del análisis de datos, que herramientas son las que más se usan o que herramientas serían interesantes para incluir dentro de la formación de Adalab, etc.

Día 1	Día 2	Día 3	Día 4	Día 5	Día 6	Día 7
<div>SPRINT PLANNING</div> <div>- Definir working agreements del equipo de desarrollo</div> <div>- Crear el tablero del Proyecto en GitHub</div>	<div>TRABAJAMOS EN EL PROYECTO</div> <div>- Daily (unos 5-10 min)</div> <div>- Cargar los datos necesarios para llevar a cabo el proyecto.</div> <div>- Hacer una exploración de los datos que tenemos en cada dataframe:<div>- ¿Qué información tenemos en los disintos ficheros?</div><div>- ¿Cuáles son los nombres de las columnas?</div><div>- ¿Cuántas filas y columnas tenemos en cada uno de los dataframes?</div><div>- ¿Cuáles son los tipos de los datos de cada una de las columnas?</div><div>- etc.</div></div>	<div>TRABAJAMOS EN EL PROYECTO: Una vez que sepamos que información tenemos contenida en nuestros datos (realizado el día anterior) deberemos:</div> <div>- Daily (unos 5-10 min)</div> <div>- Pensar algunas de las preguntas que queremos contestar para solucionar el problema que nos ha planteado el cliente (Adalab). NOTA: estas preguntas podrán ir cambiando a lo largo del proyecto, pero necesitamos definir las para poder focalizar nuestro análisis</div> <div>- Una vez definidas las preguntas, establecer que columnas son las necesarias y eliminar aquellas que no necesitaremos.</div>	<div>TRABAJAMOS EN EL PROYECTO</div> <div>- Daily (unos 5-10 min)</div> <div>- Unir todos los conjuntos de datos con las columnas que seleccionamos el día anterior para tener toda la información necesaria en un único fichero</div> <div>- Explorar este nuevo dataframe creado, ¿cuántos nulos tenemos?, ¿qué valores únicos tenemos en cada columna?, ¿tenemos que cambiar el nombre de las columnas?, ¿qué tipo de datos tenemos en cada columna?. Estas son algunas de las ideas, pero podréis explorar el dataset todo lo que queráis y con todo lo que se os ocurra</div> <div>- Recordad guardar este dataframe creado para seguir trabajando con él en el futuro.</div>	<div>TRABAJAMOS EN EL PROYECTO</div> <div>- Daily (unos 5-10 min)</div> <div>- Hasta el día de hoy hemos estado explorando los datos y quedándonos solo con las columnas y los datos que nos interesaban.</div> <div>- Hoy hemos aprendido a hacer filtros en Pandas, el objetivo de la tarea de hoy será que hagáis una serie de filtros basandoos las preguntas que definisteis al inicio del sprint que nos ayuden a ir conociendo más nuestros datos en profundidad.</div>	<div>TRABAJAMOS EN EL PROYECTO</div> <div>- Daily (unos 5-10 min)</div> <div>- Empezar a realizar algunas gráficas básicas que nos ayuden a identificar posibles erratas en los datos, extraer algunas conclusiones de forma más clara que nos ayuden a tomar las decisiones correctas para la limpieza de datos que tendremos que hacer en el siguiente sprint</div>	<div>SPRINT REVIEW Y RETRO</div> <div>- Sprint Review: 5 min. ¿qué es lo que hemos hecho? ¿qué es lo que vamos a hacer? ¿Hemos tenido algun bloqueo?</div> <div>- Retro: una cosa a destacar y a mejorar de cada una de las compañeras y de nosotras mismas.</div>
Día 8	Día 9	Día 10	Día 11	Día 12	Día 13	Día 14
<div>SPRINT PLANNING</div> <div>- Revisar working agreements del equipo de desarrollo</div> <div>- Revisar User Stories</div>	<div>TRABAJAMOS EN EL PROYECTO</div> <div>- Daily (unos 5-10 min)</div> <div>- En la clase invertida hemos aprendido el apply que nos puede ayudar a empezar a limpiar las columnas que hemos elegido para resolver el problema del cliente.</div> <div>- Deberéis empezar a limpiar el dataframe creando funciones y aplicándolas a las columnas que correspondan</div>	<div>TRABAJAMOS EN EL PROYECTO</div> <div>- Daily (unos 5-10 min)</div> <div>- Limpieza del dataframe para llegar a tener los datos preparados para realizar las gráficas que necesitamos para contestar las preguntas del cliente</div>	<div>TRABAJAMOS EN EL PROYECTO</div> <div>- Daily (unos 5-10 min)</div> <div>- Limpieza del dataframe para llegar a tener los datos preparados para realizar las gráficas que necesitamos para contestar las preguntas del cliente</div>	<div>Día de Evaluación no hay hora de proyecto</div>	<div>TRABAJAMOS EN EL PROYECTO</div> <div>- Daily (unos 5-10 min)</div> <div>- Limpieza del dataframe para llegar a tener los datos preparados para realizar las gráficas que necesitamos para contestar las preguntas del cliente</div>	<div>SPRINT REVIEW Y RETRO</div> <div>- Sprint Review: 5 min. ¿qué es lo que hemos hecho? ¿qué es lo que vamos a hacer? ¿Hemos tenido algun bloqueo?</div> <div>- Retro: una cosa a destacar y a mejorar de cada una de las compañeras y de nosotras mismas.</div>
Día 15	Día 16	Día 17	Día 18	Día 19	Día 20	Día 21
<div>SPRINT PLANNING</div> <div>- Revisar working agreements del equipo de desarrollo</div> <div>- Revisar User Stories</div>	<div>TRABAJAMOS EN EL PROYECTO</div> <div>- Daily (unos 5-10 min)</div> <div>- Limpieza del dataframe para llegar a tener los datos preparados para realizar las gráficas que necesitamos para contestar las preguntas del cliente</div>	<div>TRABAJAMOS EN EL PROYECTO</div> <div>- Daily (unos 5-10 min)</div> <div>- Realización de gráficas para contestar el requerimiento del cliente</div>	<div>TRABAJAMOS EN EL PROYECTO</div> <div>- Daily (unos 5-10 min)</div> <div>- Realización de gráficas para contestar el requerimiento del cliente</div>	<div>TRABAJAMOS EN EL PROYECTO</div> <div>- Daily (unos 5-10 min)</div> <div>- Realización de gráficas para contestar el requerimiento del cliente</div>	<div>TRABAJAMOS EN EL PROYECTO</div> <div>- Daily (unos 5-10 min)</div> <div>- Dejar el repo organizado para la entrega del proyecto</div> <div>- Revisión de las gráficas que hemos realizado (evitar tener gráficas con contenido en inglés y español, que las leyendas se lean, que los colores entre gráficas son consistentes, etc)</div>	<div>TRABAJAMOS EN EL PROYECTO</div> <div>- Daily (unos 5-10 min)</div> <div>- BONUS: introducir todo nuestro código en Funciones y Clases</div>