



Feedback del Módulo 3 Bootcamp Data Analytics

Hola, Sara 😊. Has hecho un gran trabajo esta evaluación y se nota en tu ejercicio. Parece que tu EDA no tiene fin.

A continuación, te señalo las notas de cada parte a tener en cuenta del examen:

Feedback del examen:

- **EDA [Entiende el proceso de EDA]:** Destaca
- **Exploración:** Destaca
- **Uso de histogramas:** Cumple el objetivo
- **Detección de valores únicos en variables categóricas:** Destaca
- **Gestión de nulos:** Cumple el objetivo
- **Visualización [Uso de gráficos adecuados]:** Notable
- **Estadística [Entiende los conceptos de estadística básica]:** Notable
- **Buenas prácticas:** Notable
- **Valoración global:** Notable 👍

COMENTARIOS DEL EXAMEN:

Empecemos por tu repo:

Tienes un Readme fantástico, muy buen trabajo!! Se nota que has dedicado tiempo y cuidado al desarrollarlo, está genial!!! Una pequeña mejora para la estructura de carpetas que has creado es añadir una carpeta para poner dentro todas las imágenes que has creado.

Has hecho un ejercicio muy detallado, incluso en algunos aspectos excesivo. Muy buena exploración inicial en busca de patrones y las conclusiones que has ido obteniendo son adecuadas en casi todos los casos (ahora lo desglosamos). Muy buen contexto del problema y las interpretaciones obtenidas del info y describe muy acertadas. Se nota que has interiorizado los conceptos y que te fijas en los outputs en busca de resultados.

Durante la entrevista has contestado bien a las preguntas y explicas de maravilla el código y las justificaciones que has tomado. En algunos casos te "enredas" un poco por falta de confianza a la hora de tomar decisiones, esto por un lado está genial porque te hace profundizar en el problema buscando la solución correcta, pero por otro lado te hace hacer un sobre análisis y esto supone mucho tiempo dedicado a un problema en concreto (los duplicados por ejemplo). Has hecho una gran exploración de los mismos y finalmente has conseguido tener la confianza para tomar la decisión, es normal al principio esta falta de confianza, pero poco a poco hay que ir trabajando en ganar confianza y tomar decisiones en base a lo que ves en los datos.



Feedback del Módulo 3

Bootcamp Data Analytics

Está genial todo el análisis que has realizado para entender correctamente los datos y ver las transformaciones que tienes que realizar, incluso has hecho una función para unificar los datos de tipo categórico, buen trabajo. En alguna de las transformaciones has hecho un trabajo que no hace falta, por ejemplo en transformar las columnas de mes y año en una sola columna de tipo period, la transformación que has hecho está genial y se nota que te manejas en python y que sabes buscar recursos para conseguir lo que quieres, pero no era necesario realizar esa transformación. Tengo la sensación de que a veces realizas trabajo

de más y eso hace que sientas mayor falta de tiempo o presión para realizar el ejercicio, intenta enfocar en el objetivo y simplificar al máximo y luego si tienes tiempo puedes intentar mejorar alguna cosa que te hayas dejado pendiente.

Muy buena justificación del cambio de los valores negativos de salario a valor absoluto. En este caso has enfocado y profundizado en el problema y tomado una decisión basándote en el análisis que has realizado de una forma fantástica.

En el caso de los valores nulos, has decidido eliminar las columnas de cancelation. Estos valores nulos nos están diciendo que esos clientes no han cancelado sus tarjetas de suscripción y eso es una información valiosa. La mejor opción para tratar estos valores podría haber sido crear una columna nueva que recoja esa información.

Para los nulos de salario has decidido usar KNN, ojo con este método que se puede volver muy pesado cuando tenemos muchos datos. Además, para usarlo correctamente tenemos que pasarle varias columnas que nos "expliquen" la variable que queremos imputar (variables con alta correlación con la que tiene nulos). En este caso, todos los valores nulos de salario pertenecen a la categoría de educación de College, por lo que una imputación más adecuada habría sido hacer la media de las categorías de educación inferior y superior (high school y bachelor) e imputar los valores nulos con esa métrica.

Muy buena práctica el realizar una pequeña exploración antes y después de unir los dos datasets, genial.

En la primera gráfica, has hecho una gran interpretación. En el aspecto visual, el haber creado esa columna nueva uniendo los datos, te ha perjudicado un poco porque hace la gráfica más difícil de entender y añade un día a la fecha que no existe. Aún así, has sabido ver los datos perfectamente, buen trabajo!!

El segundo ejercicio es súper completo, has sacado la correlación, un detalle de los datos, ... y lo has mostrado e interpretado de forma genial (luego hablamos del bonus).



Feedback del Módulo 3 Bootcamp Data Analytics

Al mostrar la distribución de clientes por provincia, el tipo de gráfica y la personalización están genial, sin embargo, los datos que has usado para visualizar no representan la realidad, estás usando el dataframe unido y no tenemos esa cantidad de clientes, tendrías que haber agrupado para contar una única vez cada cliente.

En la gráfica de salarios, puedes ver que el salario de College coincide exactamente con la media de todos los datos y esto es debido a la imputación de valores nulos que has realizado anteriormente.

El resto de gráficas está genial y has hecho muy buenas interpretaciones de los datos, buen trabajo!!

Menudo Bonus te has currado!! Está genial todo lo que has profundizado y las conclusiones que has obtenido. Me encanta ver que usas métodos como la correlación para entender esas relaciones y las diferentes gráficas que has hecho para entender el problema. Has conseguido entender a la perfección el problema y la relación entre esas variables incluyendo otras (tipo de tarjeta) que explicaban el problema.

En resumen, has hecho un gran trabajo y como siempre, te has esforzado un montón. Lo mejor de todo es que lo haces manteniendo una gran actitud y capacidad de aprendizaje. Enhorabuena por tu trabajo Sara, sigue así!! 🥰.