

Data Science Challenge

Data & Analytics

Bienvenido al Challenge de Data Science de Mercado Libre. Agradecemos tu interés por formar parte de nuestro equipo, y nos gustaría que te enfrentarás a los desafíos con los que trabajamos en nuestro día a día. Por medio de esta prueba nos gustaría conocer cómo te desenvuelves en las diferentes etapas de la solución de un problema de negocio, desde el entendimiento del problema, el planteamiento de preguntas, la obtención y exploración de la información, el planteamiento y creación de la solución analítica y la presentación de tus resultados. ¡Te deseamos mucha suerte!

1. Instrucciones Generales

El objetivo de esta prueba es conocer la manera en que abordarías algunos casos de negocio, desde el entendimiento del problema hasta la creación de una solución analítica, y por qué no, las propuestas que se te ocurran!

No es necesario que tengas conocimiento particular de la industria en la que nos encontramos, pues lo que queremos evaluar es la forma en la que te desenvuelves en cada fase del ciclo de Data Science.

Te presentaremos unos casos de negocio y podrás escoger **tres** en los que desees trabajar. Finalmente debes realizar una presentación de 20 minutos en la que nos guíes a través del problema, la forma en que lo abordaste, tus análisis, presentación de resultados y conclusiones.

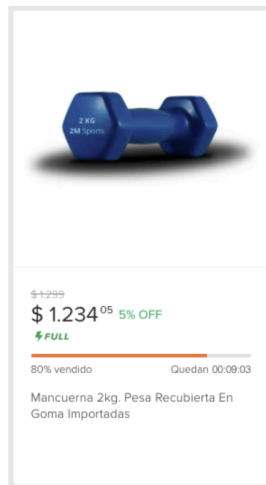
Adicionalmente, todos tus desarrollos deben estar contenidos en un **repositorio público de GitHub bajo tu usuario**. Este repositorio debe poder ser clonado y utilizado para replicar tu código de extracción de información, análisis exploratorios y modelos. Es deseable que utilices **Python en Jupyter Notebooks** por presentación y organización.

2.Casos de Negocio

A continuación te presentamos algunos problemas de negocio con los que nos enfrentamos día a día. **Debes escoger tres casos** de negocio y plantear una solución analítica que te permita generar valor tanto a la unidad de negocio como a Mercado Libre en general.

1. **Optimización de stock:** Mercado Libre ofrece a los sellers del Marketplace la posibilidad de almacenar productos en sus bodegas, teniendo en cuenta que el espacio es limitado, se debe sugerir para cada ítem la cantidad óptima de inventario que debe ser almacenado. Para esto contamos con la información de cantidad de inventario inicial, las ventas y el inventario actual. ¿Cómo abordarías el problema de sugerir primero si se debe o no comprar inventario del ítem y segundo las unidades que se deberían comprar?
2. **Generación de insights para ofertas relámpago:** En conjunto con el desafío te compartimos un archivo llamado ofertas_relampago.csv el cual posee información de los resultados de ofertas del tipo relámpago para un periodo de tiempo y un país determinado.

Estas ofertas en mercadolibre se pueden ver de la siguiente manera:



Es decir, son ofertas que tienen una duración definida de algunas horas y un porcentaje de unidades (stock) comprometidas.

El objetivo de este desafío es hacer un **EDA** sobre estos datos buscando insights sobre este tipo de ofertas.

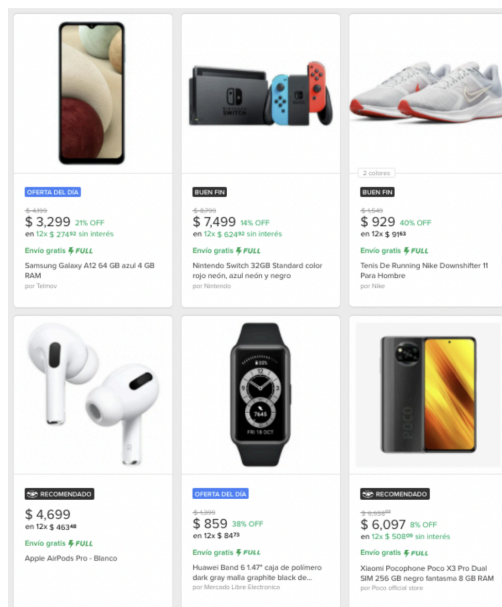
Las columnas del dataset son autoexplicativas pero puedes preguntarnos cualquier duda.

3. **Serie de tiempo:** Pronosticar las ventas de un producto o de una categoría es un desafío recurrente para cualquier ecommerce. En este caso, el desafío va a ser pronosticar la cantidad de unidades diarias que van a vender 3 categorías distintas de MELI. El dataset “series.csv” tiene las ventas en unidades diarias de 3 categorías que poseen un id único y su fecha de venta.

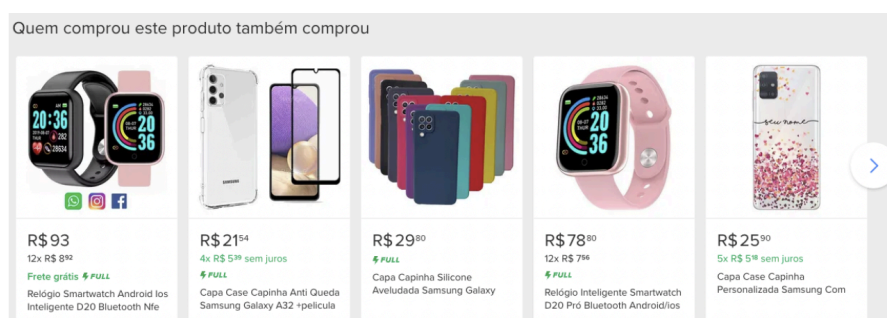
El objetivo de este desafío es construir un modelo de forecast que permita estimar las ventas de 3 semanas a nivel diario utilizando la historia de ventas de la categoría. Es decir, predecir las ventas de los siguientes 21 días. Las métricas y la medición de la performance del forecast son un punto clave de este desafío.

TIP: Dividir el dataset en entrenamiento, testing y validación correctamente es muy importante en problemas de forecasting!

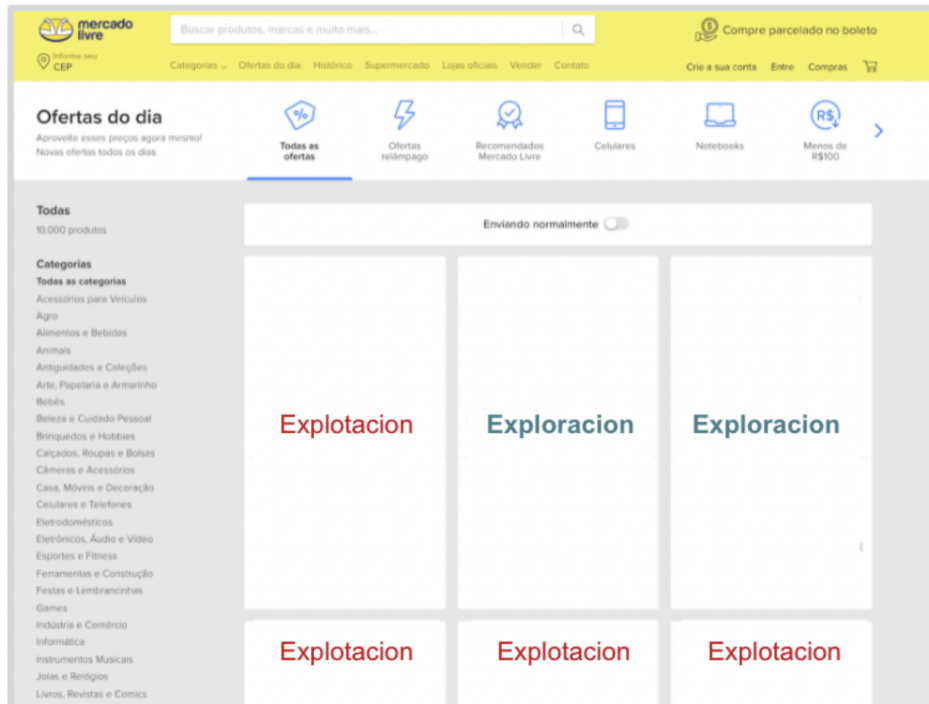
4. Exploración VS Explotación: Durante las campañas comerciales como el black friday o el cyber monday los listados de publicaciones como el que muestra esta captura



buscan exponer a los usuarios variedades de publicaciones permitiendo el discovery de buenas ofertas o productos. Otro ejemplo de un listado como estos es la sección de ofertas o por ejemplo los carruseles de recomendaciones en las páginas de producto.



Un problema interesante para estos listados es el de tener un buen mix entre exploración y explotación en las recomendaciones o en el ordenamiento. Donde por explotación nos referimos a colocar en las primeras posiciones productos que sabemos funcionan bien en estas posiciones, mientras que explorar refiere a colocar productos que todavía no sabemos cómo van a performar pero son buenos candidatos. El desafío es diseñar un sistema que mediante la interacción en producción con los usuarios nos permita definir en cuantas posiciones del ordenamiento queremos explorar y en cuantas explotar. La imagen de abajo refleja este problema.



Para este challenge te proponemos hacer un **diseño y propuesta de experimentación en producción** y te dejamos algunas preguntas a modo de guía:

- ¿Que tipo de enfoque utilizar para definir en que % de posiciones explorar y en que % explotar?
- ¿Qué métrica usarías para medir este experimento?
- ¿Cómo aprovecharías la interacción y el feedback de los usuarios?

4. Puntos Clave de Presentación

Una vez planteado y resuelto tus casos de negocio debes realizar la socialización de tus resultados. Algunos puntos a tener en cuenta son:

- *¿Qué problema se quiso resolver con el desarrollo de la solución?*
- *¿Qué información buscaste y utilizaste para el desarrollo de la solución? Incluye las hipótesis que te llevaron a descargar dicha información.*
- *Muestra los insights más relevantes que hayas obtenido de tu análisis exploratorio. ¿Cómo estos te guiaron a la hora de plantear la solución?*
- *Danos un contexto de la solución que escogiste, y cómo la abordaste. ¿Qué metodologías aplicaste? ¿Qué métricas de evaluación utilizaste?*
- *¿Cuál es tu solución final? ¿Cómo se comporta? ¿Cómo soluciona o ayuda a solucionar el problema de negocio?*
- *Concluye y plantea próximos pasos que harías después de entregar esta solución.*

Es importante que estos análisis estén respaldados por cifras, y que estas puedan ser consultadas y replicadas en los notebooks del repositorio de Github.

5. Información Adicional

Si encuentras algún problema o dudas, no lo pienses dos veces y contáctanos por medio de correo electrónico o los canales que te hayamos proporcionado cuando te enviamos esta prueba.