**Preguntas de Entrevistas Técnicas**

**Plataforma para practicar challenge técnicos de DATA:**

* [Code.golf](https://code.golf/ideas)
* [Kaggle](https://www.kaggle.com/)
* [Ejemplos para Power BI](https://learn.microsoft.com/es-es/power-bi/create-reports/sample-datasets)
* [GeeksforGeeks](https://www.geeksforgeeks.org/)
* [Pramp](https://www.pramp.com/dev/uc-data-science)
* [Coderbyte](https://coderbyte.com/challenges?utm_campaign=NewHomepage)

**PREGUNTAS ESPAÑOL:**

**SQL**

* Describa la diferencia entre claves primarias y foráneas en una base de datos SQL.
* Explique la diferencia entre la cláusula WHERE y la cláusula HAVING.
* Explique que es una función de agregación y mencione al menos 5 de las más comunes. (Count, AVG, Max, Min, STDDEV, Variance)
* Nombre los diferentes objetos de una base de datos (pista: 6 son permanentes y 1 temporal)
* ¿Qué es una subconsulta? Explique de manera breve su sintaxis.
* Describa y explique los diferentes tipos de JOINS para SQL.

**Python**

* Escriba una función que permita calcular la raíz cuadrada de un número.
* Escriba una función que dada una lista de números devuelva los outliers.
* Nombrar las diferencias entre una Tupla y una Lista.
* Especificar la diferencia entre variables locales y globales.
* ¿Python distingue entre mayúsculas y minúsculas?
* ¿Qué son los iteradores de Python? Ejemplifique.
* ¿Qué es una variable categórica en Pandas?
* ¿Qué es una serie en pandas?
* ¿Qué es un index en Pandas? ¿Cómo podría acceder a él? ¿Puede haber más de uno?.
* ¿Qué es la vectorización de operaciones en Pandas?
* ¿Qué librerías de visualización para Python conocen? ¿Y para graficar mapas geográficos? ¿Y para realizar Dashboards?
* ¿A qué se refieren con el nivel en los lenguajes de programación? ¿A qué nivel pertenece Python y por que?
* ¿Qué alcance tienen las variables en las funciones de Python?
* ¿Qué tipos nativos de estructuras de datos de Python conocen? ¿Cuáles son inmutables?
* ¿Qué es NumPy? ¿Cuál es la estructura de datos que utiliza y en qué se diferencia de una lista?
* ¿Para qué se usan break, continue y pass en python?
* ¿Cuál es la diferencia entre listas y tuplas?
* ¿Cuál es la diferencia entre listas y sets?
* ¿Cómo encontrar el modo de una lista sin usar ninguna función matemática o biblioteca incorporada?

**Data Analysis**

* ¿Cuál es la mejor forma para determinar las características más importantes en un modelo?
* ¿Cómo podría usar los datos de un GPS de un vehículo para determinar la calidad del conductor?
* ¿Cuáles serían buenas métricas para un producto e-commerce?(MercadoLibre - Amazon)¿Y para un producto de suscripción? (Netflix-Spotify-Etc)
* ¿Cuáles serían buenas métricas de éxito para un producto de consumo que depende en gran medida del compromiso y la interacción? (Snapchat, Pinterest, Facebook, etc.) ¿Un producto de mensajería? (GroupMe, Hangouts, Snapchat, etc.)
* ¿Cuáles serían buenas métricas para productos que ofrecen ventas dentro de aplicaciones? (Ej: Aplicaciones de Juegos)
* Supongamos que eres dueño de Netflix. ¿Cuáles serían las métricas que te interesa utilizar para definir en qué películas invertir?
* ¿Qué es la complejidad algorítmica? ¿Con qué símbolo o letra se representa la notación de la cota superior?

**Database**

* Explicar el esquema de tipo Estrella.
* Explicar el modelo de tipo Copo de Nieve.
* ¿Qué tipo de relaciones en el esquema debemos evitar?
* ¿Qué formatos de archivos para almacenar datos conoce?
* ¿A partir de qué tipo de archivo o de dato se puede generar un DataFrame nativamente en pandas? Nombra algunos.

**Machine learning**

* ¿Qué técnicas de machine learning has implementado?
* ¿Qué es el sobreajuste? ¿De qué forma evitar el sobreajuste en los modelos?
* ¿Cómo has evitado que tus modelos arrojan sesgos en los datos?
* ¿De qué manera determinan que un conjunto de datos es de buena calidad?
* ¿Qué es un diccionario y cuáles son las características importantes de los diccionarios?
* ¿Qué es la normalización y la desnormalización en una base de datos?
* ¿Qué nuevas tecnologías te gustaría incorporar a corto plazo? ¿Por qué?

**PREGUNTAS INGLÉS**

* What did you learn this week?
* What excites or interests you about data?
* What is a recent technical challenge you experienced and how did you solve it?
* Write an SQL query that makes recommendations using the pages that your friends liked. Assume you have two tables: a two-column table of users and their friends, and a two-column table of users and the pages they liked. It should not recommend pages you already like.
* Explain different python functions and what levels of speed each of them have.
* What’s the difference between Compiler Vs Interpreter
* Basic sql problems like to join
* What is the CAP Theorem?
* how to find mode of a list without using any in-built maths function or library
* Can you describe your workflow when you create a data analysis?