Prueba Técnica Ciencia de datos

Aspirantes al cargo de Ingenieros de Datos

**Objetivo:**

Evaluar las capacidades del postulante, midiendo:

1. Capacidad de análisis
2. Dominio de bases de datos
3. Capacidad de ingesta de datos para ciencia de datos
4. Conocimientos básicos de ingeniería de datos
5. Conocimientos en ETL
   1. Extracción: capacidad de extraer datos de diversos orígenes
   2. Transformación: capacidad de entender el negocio y la lógica de los datos.
   3. Carga: Capacidad de presentar al cliente la información
6. Nivel de Python o C# y JavaScript
7. Conocimientos sobre PowerBi
8. Conocimientos sobre bases de datos relacionales
9. Conocimientos sobre bases de datos no relacionales
10. Conocimientos en georreferenciación

**Método de evaluación de la prueba:**

Se evaluará dominio, capacidad de abstracción y resolución de problemas de ciencias de datos.

**Tiempo recomendado de la prueba:**

8 a 16 horas laborales, contadas 2 horas después de entregada la prueba en horario laboral, o al inicio del día siguiente.

**Conexión a Orígenes de Datos:**

El candidato deberá crear un script o programa en Python o C# que se conecte a dos orígenes de datos, creados por el usuario, utilizando el json adjunto en la última página de este documento.

**Extracción de Datos (ETL):**

El candidato deberá extraer datos de la fuente mencionada en el script o programa, incluyendo en el programa, el archivo json como origen de datos primarios.

Los datos extraídos deben ser transformados de alguna manera, por ejemplo, limpieza de datos, conversión de tipos, agregación, etc.

Se deberá agregar a cada documento coordenadas referentes a la ubicación de la editorial.

**Desarrollo de Software:**

El candidato deberá desarrollar un pequeño programa o script (preferiblemente en un lenguaje como Python o C#) que implemente la lógica de extracción, transformación y carga (ETL) de los datos, almacenando en una base de datos SQL el modelo relacional identificado en el archivo json cumpliendo las formas normales, deberá guardar en una base de datos en MongoDB, el mismo modelo relacional identificado.

**Publicación en Web: 45s\\7809n\\\78{\\9**

Los datos transformados deben ser publicados en una página web accesible, con usuario y contraseña, la contraseña debe ser 45s\7809n\\’78{\9 el usuario debe ser admin, opcional usar otra contraseña.

La página web debe mostrar los datos de manera clara y comprensible, preferiblemente utilizando tecnologías como HTML, CSS y JavaScript, se puede usar Flask de Python, además de mostrar un mapa con la geolocalización de las editoriales y gráficos, preferiblemente interactivos, con los análisis que se consideren relevantes.

**APIs:**

El candidato deberá exponer los datos transformados a través de una API RESTful en FastApi preferiblemente.

La API debe proporcionar endpoints para acceder a los datos de manera eficiente y segura con autenticación básica.

El candidato deberá integrar swagger para conocer la documentación de la API

**Integración con PowerBI:**

El candidato deberá crear un informe en PowerBI que se conecte a la API mencionada anteriormente.

El informe debe incluir visualizaciones significativas de los datos y ser fácilmente comprensible para un usuario final, con los análisis que se consideren relevantes.

**Control de Versiones:**

El desarrollo en su conjunto incluyendo los archivos PowerBi deben estar en un repositorio de git (subversión, Devops, Github o gitlab) a nivel publico

**Escalabilidad:**

Se debe tener en cuenta la escalabilidad del sistema desarrollado, considerando cómo manejaría grandes volúmenes de datos y un mayor número de usuarios.

Esta prueba proporciona una evaluación de las habilidades del ingeniero de datos en los diversos aspectos del que hacer profesional y de las propias habilidades de quien la realice.

Documento en JSON para evaluación técnica  
  
{

{

"id": 1,

"titulo": "El señor de los anillos",

"autor\_nombre": "J.R.R. Tolkien",

"autor\_nacionalidad": "Inglaterra",

"autor\_fecha\_nacimiento": "1892-01-03",

"autor\_genero": "Fantasía",

"nombre\_editorial": "Minotauro",

"ubicacion\_editorial": "España",

"isbn": "9788445072515",

"precio": 20.50,

"cantidad\_stock": 100

},

{

"id": 2,

"titulo": "Cien años de soledad",

"autor\_nombre": "Gabriel García Márquez",

"autor\_nacionalidad": "Colombia",

"autor\_fecha\_nacimiento": "1927-03-06",

"autor\_genero": "Realismo mágico",

"nombre\_editorial": "Diana",

"ubicacion\_editorial": "Colombia",

"isbn": "9788423349802",

"precio": 18.75,

"cantidad\_stock": 75

},

{

"id": 3,

"titulo": "1984",

"autor\_nombre": "George Orwell",

"autor\_nacionalidad": "Reino Unido",

"autor\_fecha\_nacimiento": "1903-06-25",

"autor\_genero": "Ciencia ficción",

"nombre\_editorial": "Secker & Warburg",

"ubicacion\_editorial": "Reino Unido",

"isbn": "978-0451524935",

"precio": 15.99,

"cantidad\_stock": 50

},

{

"id": 6,

"titulo": "Moby Dick",

"autor\_nombre": "Herman Melville",

"autor\_nacionalidad": "Estados Unidos",

"autor\_fecha\_nacimiento": "1819-08-01",

"autor\_genero": "Ficción de aventuras",

"nombre\_editorial": "Harper & Brothers",

"ubicacion\_editorial": "Estados Unidos",

"isbn": "9780142000083",

"precio": 22.99,

"cantidad\_stock": 80

},

{

"id": 7,

"titulo": "Orgullo y prejuicio",

"autor\_nombre": "Jane Austen",

"autor\_nacionalidad": "Reino Unido",

"autor\_fecha\_nacimiento": "1775-12-16",

"autor\_genero": "Novela romántica",

"nombre\_editorial": "Thomas Egerton",

"ubicacion\_editorial": "Reino Unido",

"isbn": "9780141439518",

"precio": 17.50,

"cantidad\_stock": 90

},

{

"id": 8,

"titulo": "El gran Gatsby",

"autor\_nombre": "F. Scott Fitzgerald",

"autor\_nacionalidad": "Estados Unidos",

"autor\_fecha\_nacimiento": "1896-09-24",

"autor\_genero": "Novela contemporánea",

"nombre\_editorial": "Charles Scribner's Sons",

"ubicacion\_editorial": "Estados Unidos",

"isbn": "9780743273565",

"precio": 19.95,

"cantidad\_stock": 60

},

{

"id": 9,

"titulo": "Don Quijote de la Mancha",

"autor\_nombre": "Miguel de Cervantes",

"autor\_nacionalidad": "España",

"autor\_fecha\_nacimiento": "1547-09-29",

"autor\_genero": "Novela picaresca",

"nombre\_editorial": "Francisco de Robles",

"ubicacion\_editorial": "España",

"isbn": "9788424119656",

"precio": 16.00,

"cantidad\_stock": 70

},

{

"id": 10,

"titulo": "Matar un ruiseñor",

"autor\_nombre": "Harper Lee",

"autor\_nacionalidad": "Estados Unidos",

"autor\_fecha\_nacimiento": "1926-04-28",

"autor\_genero": "Novela dramática",

"nombre\_editorial": "J. B. Lippincott & Co.",

"ubicacion\_editorial": "Estados Unidos",

"isbn": "9780061120084",

"precio": 21.99,

"cantidad\_stock": 85

},

{

"id": 11,

"titulo": "Crimen y castigo",

"autor\_nombre": "Fyodor Dostoevsky",

"autor\_nacionalidad": "Rusia",

"autor\_fecha\_nacimiento": "1821-11-11",

"autor\_genero": "Novela psicológica",

"nombre\_editorial": "The Russian Messenger",

"ubicacion\_editorial": "Rusia",

"isbn": "9780486415871",

"precio": 18.50,

"cantidad\_stock": 95

}

]

}