**Proyecto Integrador**

**Game Changers**

## **Análisis de Datos de la NBA 🏀**

## **Actividades Realizadas**

**Sprint 1**

**Creación del  repositorio público y compartirlo con el equipo**

Para facilitar la colaboración y el seguimiento del proyecto, se creó un repositorio público en GitHub y se compartió con el equipo.

El repositorio se titula Game-Changers-PF/Proyecto-Final el cual se configuró como público para asegurar que todos los miembros del equipo tengan acceso completo.

Se envió invitación a través de GitHub y una vez aceptada ya se tenía acceso al repositorio.

El repositorio está disponible públicamente en GitHub en el siguiente enlace: https://github.com/Game-Changers-PF/Proyecto-Final

**Identificar los archivos a usar y armar la estructura de carpetas a usar en el repositorio**

Para organizar eficientemente los archivos y recursos del proyecto, hemos definido una estructura de carpetas y hemos identificado los archivos principales a utilizar:

**README.md**: Este archivo contiene la descripción principal del proyecto, incluyendo su propósito, instrucciones de instalación, y otros detalles relevantes.

**.Vscode/**: Esta carpeta contiene configuraciones específicas de Visual Studio Code, como ajustes de configuración y extensiones utilizadas para el desarrollo del proyecto.

**Machine Learning/:** Carpeta que contiene archivos relacionados con el desarrollo de modelos de aprendizaje automático, como scripts de entrenamiento, evaluación de modelos, y otros recursos asociados.

**MachineLearningModel.ipynb:** Archivo Jupyter Notebook que contiene código para actualizar o mejorar el modelo de aprendizaje automático existente.

**Csv\_filtrados/:** almacena los archivos CSV filtrados específicamente para el análisis de datos.

**Csv\_originales/:** contiene los archivos CSV originales sin procesar, utilizados como datos de entrada para el proyecto.

**Diseños/:** contiene el logo del proyecto.

**Limpieza/:** contiene scripts y archivos relacionados con la limpieza de datos, incluyendo procesos ETL (Extract, Transform, Load).

**Myenv/:** contiene el entorno virtual myenv utilizado para gestionar las dependencias y el ambiente de desarrollo específico del proyecto.

**Conexion\_API\_y\_consulta.ipynb:** Archivo Jupyter Notebook que contiene código para la conexión a una API externa y la consulta de datos lo cual nos aportó un valor agregado al archivo brindado por Henry

**Matriz correlacion.png:** Imagen que muestra una matriz de correlación utilizada para visualizar las relaciones entre variables en el análisis de datos.

**Creación base de datos SQL Management Studio**

Creamos la base de datos en SQL y para asegurarnos de que todo el equipo pueda conectarse a la misma configuramos las credenciales de cada miembro para acceder al servidor SQL. Esto implica proporcionar nombres de usuario y contraseñas adecuadas.

**Entidades/tablas que se usaron en la  BD**

Se genero el modelo relacional con todas las tablas y sus relaciones

**Crear las tablas y objetos SQL necesarios**

creamos las estructuras de datos necesarias para almacenar información sobre equipos, jugadores y partidos relacionados con la NBA

**Flujo de  importación de  los datos a la base de datos**

los datos que importamos provienen de fuentes diversas, como archivos CSV descargados de sitios oficiales de la NBA, servicios web que proporcionan actualizaciones en tiempo real de estadísticas de jugadores y partidos, o incluso ingreso manual de datos.

Utilizamos scripts y herramientas para extraer los datos del origen y se realizaron transformaciones básicas para limpiar o estructurar los datos antes de cargarlos en la base.

**Validación de acceso a la  base de datos**

nos aseguramos de que todos los miembros del equipo puedan acceder correctamente a la base de datos de la NBA para realizar sus análisis y consultas.

Definimos qué usuarios específicos tienen acceso a la base de datos y qué tipo de permisos tienen (lectura, escritura, administración, etc.).

Se configuro el firewall y las reglas de red para que solo las personas autorizadas puedan acceder de manera segura y protegida a la información almacenada en esa base de datos.

**Creación método de automatización de ingesta de datos nuevos**

Se importan las librerías necesarias para trabajar con archivos CSV, bases de datos SQL Server, AWS S3, y para cargar variables de entorno desde un archivo .env usando dotenv.

Conectamos con la base de datos SQL Server (connect\_to\_database() usando las credenciales y detalles de conexión del archivo .env. y utilizamos pyodbc para la conexión ODBC.

Leemos el archivo CSV almacenado en un bucket de Amazon S3 especificado (read\_csv\_from\_s3(bucket\_name, key) y con boto3 interactuamos con S3 y pandas para leer el archivo CSV en un DataFrame.

Insertamos los datos del DataFrame en la tabla especificada en la base de datos. Construimos y ejecutamos una consulta SQL de inserción para cada fila del DataFrame y automatizamos el proceso completo de carga de datos, primero leemos el archivo CSV desde S3 (read\_csv\_from\_s3), luego insertamos los datos en la base de datos (insert\_data\_to\_database), se devuelve un diccionario con un estado de código y un mensaje indicando si la operación fue exitosa o si ocurrió algún error.

Definimos los parámetros (bucket\_name, key, table\_name) necesarios para la carga de datos dentro de un bloque que se ejecutara solo cuando se ejecuta el archivo.

Llamamos a la función principal y mostramos la respuesta que indica si la operación fue exitosa o si hubo algún error.

**SPRINT 2**

1) **Creación del archivo .pbix en Power BI**

Creamos un archivo .pbix en Power BI como nuestro centro de análisis para datos estadísticos detallados de jugadores en la NBA.

2) **Importación e integración de datos**

Importamos datos detallados de juegos, rendimiento de jugadores, estadísticas de equipos, porcentajes de tiros, rebotes y otras métricas clave desde fuentes confiables como una API de la NBA que se consiguió, bases de datos y CSV de juegos y salarios.

3) **Limpieza y preparación de datos**

Aplicamos procesos de limpieza y transformación para asegurar que los datos estén consistentes y libres de errores. Esto incluye la corrección de formatos, la eliminación de datos duplicados y la estandarización de nombres y categorías para facilitar el análisis.

4) **Creación de medidas/columnas/relaciones**

Se crearon medidas y cálculos que permiten evaluar el rendimiento de los jugadores en diferentes aspectos del juego, como puntos anotados por partido, porcentaje de tiros de campo, efectividad en tiros de tres puntos, promedio de rebotes por juego, asistencias, robos de balón, y bloqueos, entre otros.

5) **Diseño de reportes y gráficos a usar**

Diseñamos reportes interactivos y visualizaciones gráficas que muestran de manera clara y comprensible las estadísticas y el rendimiento de los jugadores. Estas visualizaciones incluyen gráficos de barras, gráficos de líneas, araña, tarjetas y tablas dinámicas que permiten a los usuarios explorar los datos desde diferentes perspectivas.

6) **Análisis general del tablero**

Creamos paneles de control personalizados en Power BI que permiten a gerentes, técnicos y otros interesados en la compra de jugadores acceder fácilmente a las estadísticas clave y a un ranking de los 10 mejores jugadores. Estos paneles pueden incluir resúmenes de salarios, comparativas entre jugadores, análisis histórico de desempeño y recomendaciones basadas en datos.

7) **Pruebas de calidad del reporte**

Realizamos pruebas exhaustivas para validar la precisión de los datos y la funcionalidad de los reportes. Aseguramos que todas las métricas calculadas sean precisas y reflejen correctamente el desempeño real de los jugadores.

8) **Entrega y soporte continuo**

Entregamos el archivo .pbix finalizado junto con documentación detallada sobre su uso y mantenimiento. Proporcionamos soporte continuo para responder preguntas y realizar ajustes según las necesidades específicas de los usuarios.

6) **Vistas del tablero**

Creamos vistas específicas dentro de nuestros paneles en Power BI que permiten a gerentes, técnicos y otros interesados en la compra de jugadores acceder fácilmente a las estadísticas clave. Estos paneles pueden incluir resúmenes ejecutivos, comparativas entre jugadores, análisis histórico de desempeño y recomendaciones basadas en datos.

7) **Pruebas de calidad del reporte**

Realizamos pruebas para validar la precisión de los datos y la funcionalidad de los reportes en Power BI. Aseguramos que todas las métricas calculadas sean precisas y reflejen correctamente el desempeño real de los jugadores.

8) **Resultados principales y líneas futuras de análisis**

Los principales hallazgos obtenidos a partir del análisis de datos en Power BI, destacando las tendencias más significativas y los insights clave sobre el rendimiento de los jugadores en la NBA. Además, identificamos posibles áreas para futuros análisis, como la integración de datos en tiempo real, modelos predictivos de desempeño o análisis comparativos con ligas internacionales.

**Dashboard en Versión Móvil 📊📱**

Hemos creado un dashboard móvil que te permite ver estadísticas detalladas de jugadores de la NBA desde tu teléfono o tableta. Este dashboard te proporciona toda la información que necesitas sobre el rendimiento de los jugadores, estés donde estés.

Decidimos qué estadísticas eran más importantes para los fanáticos de la NBA, como puntos por partido, asistencias, rebotes y porcentaje de tiros.

Pensamos en cómo organizar esta información de manera que fuera fácil de leer en una pantalla pequeña.

Hicimos bocetos para ver cómo se verían las estadísticas en el teléfono, asegurándonos de que todo fuera claro y accesible.

Diseñamos la disposición de los gráficos y tablas para que se adaptaran bien a diferentes tamaños de pantalla.

Se ajusto el dashboard para que fuera fácil de usar con el tacto. Los botones y controles están diseñados para ser tocados con precisión, y los filtros permiten seleccionar jugadores o estadísticas específicas sin complicaciones.

Incorporamos opciones para comparar estadísticas de diferentes jugadores con solo unos toques.

Probamos el dashboard en varios teléfonos y tabletas para asegurarnos de que se viera y funcionara bien en todos ellos.

Publicamos el dashboard en Power BI, asegurándonos de que estuviera disponible en la aplicación móvil para que cualquier usuario autorizado pudiera acceder a él.

Ahora tienes acceso a un dashboard móvil que te muestra estadísticas detalladas de jugadores de la NBA. Esta herramienta está diseñada para que puedas ver fácilmente la información que necesitas sobre el rendimiento de los jugadores, estés en casa o en movimiento. Todo está pensado para que sea fácil de usar y accesible desde tu teléfono o tableta.