



**INSTITUTO PERUANO DE ADMINISTRACIÓN DE EMPRESAS
IPAE**

**CARRERA PROFESIONAL DE DESARROLLO DE SISTEMAS
DE LA INFORMACIÓN CON MENCIÓN EN INTELIGENCIA DE
NEGOCIOS**

**PROYECTO DE ANALYTICS PARA EL AREA DE DESARROLLO Y
DISEÑO DE JUEGO DE LA EMPRESA CANVARITECH S.A.C**

**Trabajo de aplicación profesional para obtener el título de Desarrollo de
Sistemas de la información con mención en Inteligencia de Negocios**

**COVEÑAS SULLON, WALTER ARTURO
(0000-0001-6614-593X)**

**GARCÍA VARILLAS, ANDERSON ALONSO
(0000-0002-7890-6326)**

**ZETA ALCOSER, ALONZO AMOR
(0000-0003-4619-0308)**

Piura - Perú

2022

*Dedicada a los padres de cada integrante y a nuestros profesores que nos acompañaron
durante toda nuestra carrera.*

Índice General

Resumen Ejecutivo	10
Introducción	11
Capítulo I: Modelo de Negocio	13
Nominación de la empresa	13
Historia y Rubro de la Empresa	13
Misión de la Empresa	13
Visión de la Empresa	13
Organigrama de la Empresa	14
Capítulo II: Mapa Estratégico del Negocio	15
Descripción del Problema	15
Formulación del Problema	15
<i>Problema General</i>	15
<i>Problemas Específicos</i>	15
Antecedentes.....	15
<i>Antecedentes internacionales</i>	15
<i>Antecedentes nacionales</i>	16
Métricas y OKRs del área	16
Objetivos.....	17
<i>Objetivo General</i>	17

<i>Objetivos Específicos</i>	17
Limitaciones y Alcance.....	18
Justificación	18
<i>Justificación Teórica</i>	18
<i>Justificación Práctica</i>	19
<i>Justificación Metodológica</i>	19
Stakeholders del Proyecto	19
Funciones en el Proyecto.....	19
Hardware y Software empleado en el Proyecto	20
Capítulo III: Marco Teórico	21
Metodología Kimball	23
Capítulo IV: Método para el Análisis de Datos	24
Fuente de datos	24
Modelamiento multidimensional	25
<i>Modelo de una base de datos</i>	25
<i>Definiendo el modelo</i>	27
Instrumento modelado	27
Proceso ETL.....	52
<i>Extracción de Datos</i>	52
<i>Transformación de Datos</i>	52
<i>Carga de Datos</i>	52

Capítulo V: Elaboración y presentación de sistema de reportería	54
Capítulo VI: Aplicación técnica de analítica.....	62
Conclusiones	64
Recomendaciones	65
Referencias bibliograficas.....	66
Anexos.....	67

Índice de Tablas

Tabla 1. Resultados de Encuesta de Satisfacción	70
--	----

Índice de Figuras

Figura 1. <i>Organización de la empresa</i>	14
Figura 2. <i>Matriz del proceso Kimball implementado</i>	23
Figura 3. <i>Diagrama de Base de Datos Transaccional</i>	25
Figura 4. <i>Diagrama de Base de Datos del Data Mart</i>	26
Figura 5. <i>Opción para restaurar backup de una base de datos</i>	28
Figura 6. <i>Opción para ver el diagrama de la base de datos transaccional</i>	29
Figura 7. <i>Query para el llenado de la dimensión Player</i>	30
Figura 8. <i>Query para el llenado de la dimensión Level</i>	30
Figura 9. <i>Query para el llenado de la dimensión Time</i>	31
Figura 10. <i>Query para el llenado de la dimensión Ubication</i>	31
Figura 11. <i>Query para el llenado de la tabla Sessions</i>	32
Figura 12. <i>Instalación de extensión Microsoft Analysis Services Projects</i>	33
Figura 13. <i>Creación de Proyecto en Visual Studio para Analysis Services</i>	34
Figura 14. <i>Opción para crear conexión entre las bases de datos</i>	35
Figura 15. <i>Opción para elegir el tipo de administrador de conexión</i>	36
Figura 16. <i>Flujo de proceso ETL</i>	37
Figura 17. <i>Tarea para el limpiado de registros</i>	38
Figura 18. <i>Query para el limpiado de datos</i>	39
Figura 19. <i>Flujo para las tareas de llenado de datos</i>	40

Figura 20. <i>Origen OLE DB para el llenado de datos de Player</i>	41
Figura 21. <i>Vinculación de la consulta y las columnas de la tabla</i>	42
Figura 22. <i>Ejecución de proceso ETL</i>	43
Figura 23. <i>Elección de Origen de Datos en Power BI</i>	44
Figura 24. <i>Conexión con base de datos en Power BI</i>	45
Figura 25. <i>Visualizaciones y Campos en Power BI</i>	46
Figura 26. <i>Tabla para Análisis de Canasta de Mercado</i>	47
Figura 27. <i>Datos de la tabla Market_Basket_Analysis</i>	47
Figura 28. <i>Gráfico MBA para las combinaciones de niveles</i>	48
Figura 29. <i>Código Python para implementar librerías</i>	49
Figura 30. <i>Código Python para transformación de datos</i>	50
Figura 31. <i>Código Python para obtener datos con regla de soporte</i>	50
Figura 32. <i>Código Python para generar gráfica de diagrama de redes</i>	51
Figura 33. <i>Opción para ejecutar código Python</i>	51
Figura 34. <i>Carga de Datos en Visual Studio</i>	53
Figura 35. <i>Reporte sobre Niveles con más minutos jugados</i>	55
Figura 36. <i>Filtros del dashboard de los niveles con más minutos jugados</i>	56
Figura 37. <i>Resumen del dashboard de niveles más jugados</i>	57
Figura 38. <i>Gráficas del dashboard de niveles más jugados</i>	57
Figura 39. <i>Reporte de Niveles más escogidos</i>	58
Figura 40. <i>Filtros del dashboard de los niveles más escogidos</i>	59
Figura 41. <i>Resumen del dashboard de los niveles más elegidos</i>	60

Figura 42. <i>Gráficos del dashboard de los niveles más escogidos.....</i>	61
Figura 43. <i>Reporte de Análisis de Canasta de Mercado.....</i>	63
Figura 44. <i>Encuesta de satisfacción para el área de diseño de videojuegos.....</i>	67
Figura 45. <i>Resumen de primera pregunta de la encuesta de satisfacción</i>	68
Figura 46. <i>Resumen de segunda pregunta de la encuesta de satisfacción</i>	68
Figura 47. <i>Resumen de tercera pregunta de la encuesta de satisfacción</i>	69
Figura 48. <i>Resumen de cuarta pregunta de la encuesta de satisfacción.....</i>	69
Figura 49. <i>Resumen de quinta pregunta de la encuesta de satisfacción.....</i>	70

Resumen Ejecutivo

En el presente proyecto se ha elaborado un sistema de analítica para el área de diseño y desarrollo de videojuegos. Para ello se ha elaborado un Data Mart mediante un proceso ETL que se obtuvo gracias a la base de datos de un videojuego proporcionada por la empresa canvaritech S.A.C. La elaboración de este proyecto ha sido desarrollada bajo la metodología Kimball debido a que el sistema está orientado a un área específica de la empresa. Se hicieron dashboards con métricas usadas por el área y un Análisis de Canasta de Mercado adaptado a la necesidad del área por saber las combinaciones de niveles que más se juegan. Basándose en la metodología Kimball se pudo desarrollar el sistema de analítica que permitió generar un Data Mart siguiendo un proceso ETL. Este sistema realiza reportes que ayuden al decisor a saber qué nivel del juego tiene más minutos jugados, cuál nivel se escoge más y las combinaciones de 3 niveles para la elaboración de evento dentro del videojuego aplicando un Análisis de Canasta de Mercado.

Introducción

En la elaboración del presente proyecto se investigó sobre las actividades, funciones, objetivos e importancia de un área de diseño de videojuego. Todo videojuego tiene que tener un diseño el cuál permita crear la experiencia para el jugador, usuario que juega, en su interacción, diversión y estética dentro del videojuego.

Esta investigación previa ha permitido tener mejor contexto y conocimiento para el desarrollo del proyecto. Este proyecto ayuda a la empresa Canvaritech S.A.C en obtener indicadores que permitan saber si la experiencia o diseño de su juego está teniendo buenos resultado o no. Los objetivos han sido predecir cuáles son las combinaciones de 3 niveles diferentes qué el jugador elige más, desarrollar un proceso ETL usando la base de datos del videojuego para el sistema de analítica e identificar los niveles más jugados dentro del videojuego. Estos objetivos ayudan en el cumplimiento del objetivo general Desarrollar un sistema de analítica para apoyar al área de diseño de juego.

En el informe del proyecto se podrá conocer más sobre la empresa en el capítulo de modelo de negocio. En el capítulo Mapa Estratégico de la Empresa se puede conocer más sobre los problemas, objetivos, antecedentes, limitación y alcance del proyecto, la justificación, stakeholders, funciones y el hardware y software empleado para la realización del proyecto. En el capítulo de marco teórico. En el capítulo de Método para el análisis de datos se describe sobre la base de datos origen y la que se ha planteado para el modelo multidimensional, así como el proceso técnico realizado para la generación del sistema. En el capítulo Proceso ETL se detalla sobre las etapas de extracción, transformación y carga de datos. En el capítulo de Elaboración y presentación de sistema de reportería se describe sobre la elaboración de los reportes que la empresa visualizará en el sistema. El capítulo de Aplicación Técnica de Analítica se detalla sobre la aplicación realizada para un Análisis de Canasta de Mercado que

permite obtener las combinaciones de los niveles más jugados. Finalmente, los últimos capítulos son conclusiones y recomendaciones que se han obtenido luego de haber terminado la realización del proyecto.

Capítulo I. Modelo De Negocio

1.1. Nominación de la empresa

La empresa Canvaritech S.A.C con RUC 20606439629.

1.2. Historia y Rubro de la Empresa

La empresa Canvaritech fue formalizada y registrada en Piura el 01 de septiembre del 2020. Sin embargo, la idea comenzó en agosto del 2019 al haber ganado un hackathon organizado por Hub UDEP (incubadora de la Universidad Privada de Piura) y Plan Internacional.

La empresa desarrolla productos y servicios de aprendizaje en emprendimiento de negocios en base a la gamificación, edutainment y los videojuegos. Actualmente ofrece talleres virtuales y han desarrollado un videojuego de mesa educativo sobre emprendimiento de negocios.

1.3. Misión de la Empresa

Ofrecer entretenimiento y educación en nuestras marcas para niños y jóvenes.

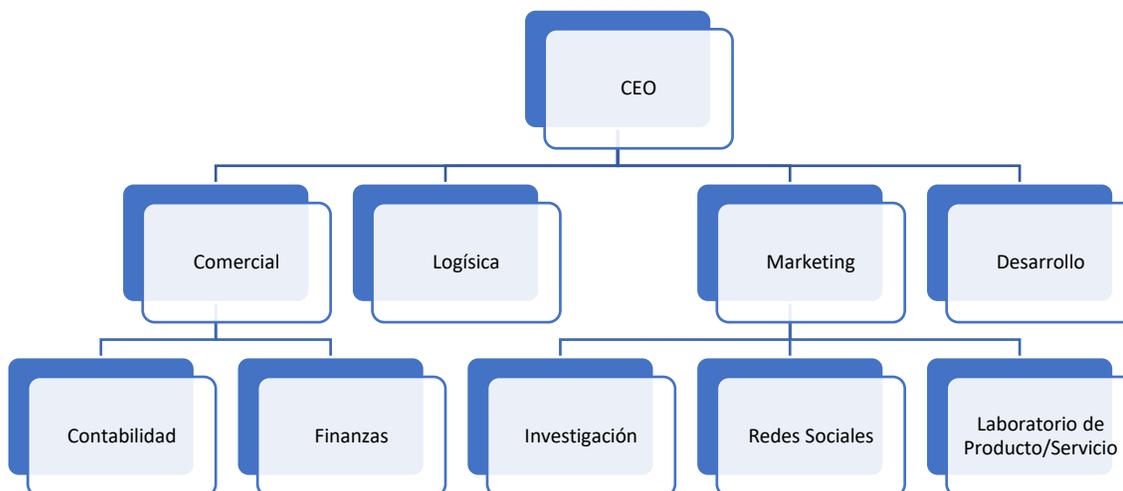
1.4. Visión de la Empresa

Todos los niños y jóvenes de Latinoamérica tengan opciones de aprendizaje entretenido.

1.5. Organigrama de la Empresa

Figura 1.

Organización de la empresa



Nota. Mapa jerárquico de la organización otorgado por la empresa canvaritech S.A.C

Capítulo II. Mapa Estratégico Del Negocio

2.1. Descripción del Problema

El área de desarrollo la empresa CANVARITECH S.A.C quiere analizar el diseño del videojuego educativo que ha sido desarrollado para poder saber que niveles son los más jugados y conocer más a sus jugadores o usuarios.

2.2. Formulación del Problema

2.2.1. Problema General

¿Cómo se puede analizar el diseño del juego?

2.2.2. Problemas Específicos

- ¿Qué niveles se pueden juntar para colocarlos en eventos dentro del videojuego?
- ¿Cómo se puede extraer información de la base de datos para analizar el videojuego?
- ¿Cuáles son los niveles más jugados?

2.3. Antecedentes

2.3.1. Antecedentes internacionales

En 2017 Nintendo, una de las empresas de videojuegos más grande del mundo, lanzaba su nueva consola con grandes expectativas de sus usuarios. Detrás de este lanzamiento, hubo todo un proyecto en el diseño y desarrollo. Nintendo creó su servicio My Nintendo, donde los usuarios de todos los juegos distribuidos por la empresa podían registrarse. Con este servicio se les ofrecía ofertas y promociones. Pero el verdadero objetivo era recolectar data para analizar las preferencias y comportamiento de los jugadores. Así, tras 4 años de desarrollo. Nintendo lanzo una consola que tendría aceptación. Esta consola es Nintendo Switch, una de las consolas más adquiridas junto a Xbox y PlayStation. (Deusto Blog, 2021)

Uno de los videojuegos móviles con más descargas es Candy Crush. Su compañía desarrolladora, King, usa Análisis e Inteligencia de Negocio para mejorar su videojuego en cada versión nueva. Uno de los productores de Candy Crush comentó: “Los usuarios del videojuego estaban abandonando masivamente el nivel 65, sin saber las razones. Con 725 niveles en total dentro del videojuego, para la saga de Candy Crush esta tendencia era un problema. King, la empresa, empleó análisis de datos para revelar que la mayoría de las personas abandonaban este nivel debido a un elemento en particular de ese nivel que no dejaba a los jugadores pasar al siguiente nivel. Entonces el elemento fue eliminado y la retención de usuarios fue creciendo de nuevo”. (Equipo Indicative, 2021)

2.3.2. Antecedentes nacionales

Hasta la actualidad la mayoría de videojuegos peruanos lanzados que han tenido éxito y se obtiene información. No han usado un sistema de BI o BA propio. Solo han usado servicios externos de analítica de datos de videojuegos de Google, Unity o Unreal.

2.4. Métricas y OKRs del área

Dentro del área se manejan métricas relacionadas al diseño y desarrollo del videojuego. Estas métricas son:

- Cantidad de errores reportados al mes
- Cantidad de minutos jugados al mes
- Cantidad de minutos jugados por nivel al mes
- Cantidad de usuarios recurrentes al mes
- Cantidad de usuarios nuevos al mes
- Cantidad de usuarios inactivos al mes
- Valoración del juego

La empresa no trabaja con KPIs sino usa OKRs (Objetivos y resultados claves). Esto se centra más en medir los resultados para saber si un objetivo se ha complicado. Cada área hace OKRs por trimestre. Pero también hay OKRs anuales. Para este proyecto se considerará los anuales. Estos son los OKRs del área desarrollo y diseño:

- Objetivo: Retener usuarios activos
 - Resultado clave: Tener un churn rate o tasa de cancelación en 10%.
 - Resultado clave: Tener una tasa de usuarios activos en 90%
- Objetivo: Tener estrategia para temporadas o campañas dentro del juego
 - Resultado clave: Tener 3 combinaciones de niveles más jugados
 - Resultado clave: Saber los 2 niveles menos jugados

2.5. Objetivos

2.5.1. Objetivo General

Desarrollar un sistema de analítica para apoyar al área de diseño de juego.

2.5.2. Objetivos Específicos

Predecir cuáles son las combinaciones de 3 niveles diferentes que el jugador elige más.

Desarrollar un proceso ETL usando la base de datos del videojuego para el sistema de analítica.

Identificar los niveles más jugados dentro del videojuego.

2.6. Limitaciones y Alcance

La limitación del sistema es:

- El sistema funciona solo para analizar videojuegos que tengan niveles.
- El sistema se puede adaptar si se desea hacer un análisis con datos masivos.

Los alcances del proyecto son:

- El sistema se podrá usar de manera online y offline.
- El sistema usará gráficos para mostrar el análisis de forma entendible.
- El sistema servirá para el área de diseño de juego.

2.7. Justificación

2.7.1. Justificación Teórica

El análisis e inteligencia de negocios permite a las empresas dar valor a sus datos que obtienen de sus usuarios o clientes. La base de datos que se suele usar en un servicio o producto, es de tipo transaccional. Es decir, para insertar, modificar, actualizar, leer o eliminar datos. Pero esos datos pueden dar más información y ayudar a la empresa a tomar decisiones. Para ello se debe hacer un proceso de ETL para poder obtener un Data Mart o Data Warehouse, un conjunto de Data Mart. En un Data Mart se guardará una base de datos de dimensiones y hechos. Cada Data Mart es estructurado en base a las medidas que se quieren obtener y son estas medidas las que ayudarán al negocio a tomar decisiones para mejorar su servicio.

2.7.2. Justificación Práctica

El punto más alto de un videojuego suele ser entre el primer y segundo mes de lanzamiento. Después de estos meses suele haber una caída de usuarios activos. Por eso es necesario analizar la experiencia del juego en esos meses claves para en los próximos meses sacar más novedades de acuerdo a los usuarios y preferencias dentro del videojuego. Es por eso que un sistema de analítica permitirá filtrar todos los datos obtenidos los primeros meses para poder medir la experiencia con métricas que les servirá al área de diseño de juego.

2.7.3. Justificación Metodológica

La metodología que se usará para lograr desarrollar el proyecto será Kimball. Esta metodología es específicamente para desarrollar sistemas con un ciclo de vida dimensional. Estas dimensiones serán del videojuego como nivel, jugador ubicación y tiempo. Esta metodología permite centrarse en el negocio, la infraestructura dónde se obtendrán los datos y dar una solución enfocada en un área del negocio. En este caso enfocado al área de Diseño de Juego o Game Design.

2.8. Stakeholders del Proyecto

Los interesados de este proyecto es el encargado del diseño del videojuego principalmente. Pero también los jugadores. Aunque sea otro departamento o área, Marketing también le interesará esos datos para conocer más a los usuarios a quién hará las campañas. Y el área comercial también para saber si el videojuego está teniendo usuarios activos que generen rentabilidad.

2.9. Funciones en el Proyecto

Las funciones de este proyecto serán las siguientes:

- Desarrollador en jefe

Su función principal es hacer la actualización de datos.

- Diseñador de Juego en jefe

Su función es diseñar el videojuego teniendo en cuenta al usuario y el objetivo de la empresa.

2.10. Hardware y Software empleado en el Proyecto

Los softwares usados principalmente en el proyecto son:

- SQL Server 2018
- Visual Studio Community 2019
- Power BI
- Python 3.8
- Entorno Conda para Python
- Pycharm Community 2020.3.3

El hardware usado para este proyecto es:

- Laptop de sistema operativo Windows 10, 4GB de RAM y disco duro de 1TB.

Capítulo III. Marco Teórico

La Inteligencia de Negocios en la industria de los videojuegos no es nuevo. Desde hace dos décadas se iniciaba a usarla de forma básica. Inicialmente se usaba investigaciones de mercado y producto tradicionales como encuestas, focus group, los beta testings. Pero esto era enfocado en muestras poblacionales.

Con el auge de las tiendas comerciales como Steam, Origin, Play Station Store, Play Store, entre otros. Las empresas desarrolladoras o las distribuidoras empezaron a almacenar datos de sus jugadores para conocerlos más. Estos datos empezaron a cobrar valor para las áreas de marketing y producción de los videojuegos. Pero también se empezó a brindar parte de esta información a las áreas de Diseño de Videojuegos.

Para entender mejor el beneficio que puede dar el análisis de negocio e inteligencia de negocio a esta área. Se debe entender qué es lo que hace dicha área. El diseñador de videojuegos es el encargado de realizar el diseño del videojuego, es decir elaborar un sistema que integre las funcionalidades del videojuego, las reglas, la experiencia que se quiere dar, la narrativa y el sonido. Esto teniendo en cuenta la diversión y el tipo de jugador o usuario a quién va dirigido. Es importante que tenga un equilibrio entre estos factores.

Pero esto no solo termina una vez lanzado el videojuego. Pues actualmente hay muchos videojuegos que son servicios. Pues dentro de un período determinado se ofrece más opciones, funcionalidades o variedades al jugador. La mayoría de estas son de pago, aunque también puede ser gratis. Esto depende del modelo de negocio. Es por esto que los diseñadores de videojuegos necesitan saber cómo es el comportamiento de los jugadores dentro del videojuego.

Uno de las técnicas usadas por negocios de venta de productos es el Análisis de Canasta de Mercado. El cuál permite obtener combinaciones de productos que frecuente son

comprados. Se podría definir como “Una de las técnicas claves usadas por grandes compañías de retail que quieren saber asociaciones entre sus productos... Permite identificar relaciones entre productos que los clientes compran.” (Susan Li, 2017). Si bien es una técnica usada en la industria retail, también se puede adaptar a otras industrias como la de videojuegos. Aunque su aplicación depende del diseño del videojuego.

Esta técnica trabaja con una regla de asociación las cuáles tiene dos factores: Antecedente y consecuente. El antecedente es un ítem que se ha encontrado en las ventas o transacciones de un sistema. El antecedente podría tener un consecuente, que es el ítem que se adquirió o eligió luego del antecedente. Ambos forman una combinación.

Para hacer el análisis de las combinaciones se usa un algoritmo en el cual participan tres elementos bases: El soporte (support), la confianza(confidence) y el ascenso(lift). El soporte es calculado por el número de ventas o transacciones de una combinación, dividida entre el total de ventas o transacciones de los ítems que conforman la combinación. Su resultado es el porcentaje de soporte de dicha combinación. La confianza es calculada por el número de ventas o transacciones de una combinación entre el total de ventas o transacciones del ítem más popular en la combinación. Su resultado es el porcentaje de confianza de dicha combinación. El ascenso es calculado entre la división del porcentaje de soporte y el porcentaje de confianza de una combinación. Su resultado es un número que permite saber si la combinación es frecuente o no. Si su valor es 1 o más significa que hay mayor probabilidad que se hagan ventas o transacciones por la combinación. En caso sea menor que 1, no valdría la pena usar la combinación debido a que no es muy frecuente dicha combinación.

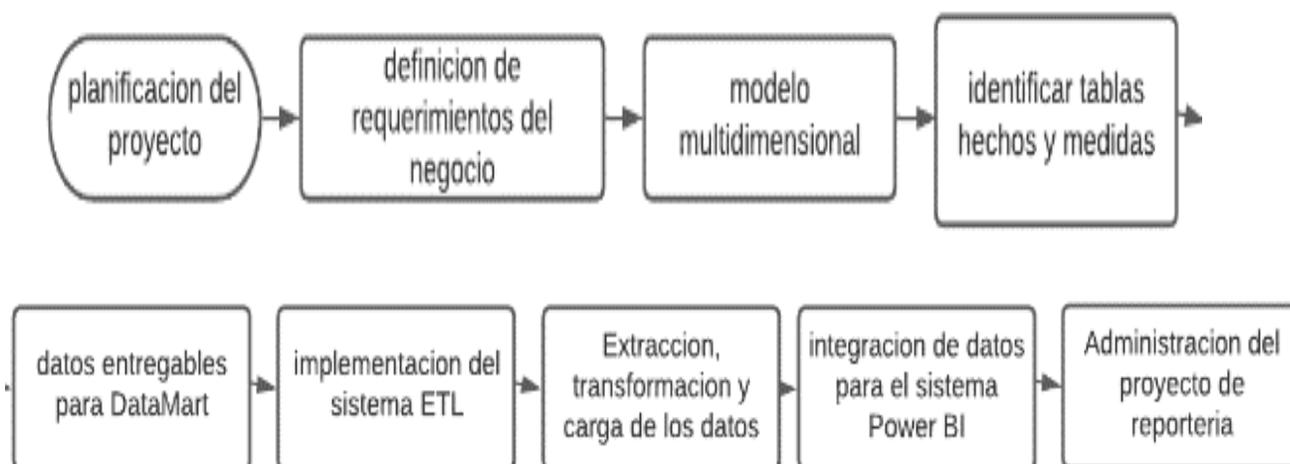
3.1. Metodología Kimball:

La metodología usada sería el proceso Kimball, Esta metodología permite centrarse en el negocio, la infraestructura dónde se obtendrán los datos y dar una solución enfocada en un área del negocio es clave entender qué es lo que deben hacer nuestros usuarios en su trabajo y cuál es el ciclo de vida del dato el uso de técnicas y diseños flexibles. Es probable que nuevos requerimientos surjan, nuevos módulos o sistemas deban ser incorporados e incluso puede ocurrir cambio de gerentes.

A continuación, nuestra matriz del proceso Kimball implementado:

Figura 2.

Matriz del proceso Kimball implementado



Nota. El grafico representa nuestra matriz del proceso Kimball implementado

Capítulo IV. Método Para El Análisis De Datos

En el análisis cualitativo se intentará trabajar de la mano con el análisis cuantitativo. Pues se trabajará con medidas y posteriormente se tratará esa información con modelos estadísticos que permitirán interpretar los datos para una mejor toma de decisiones en base a predicciones.

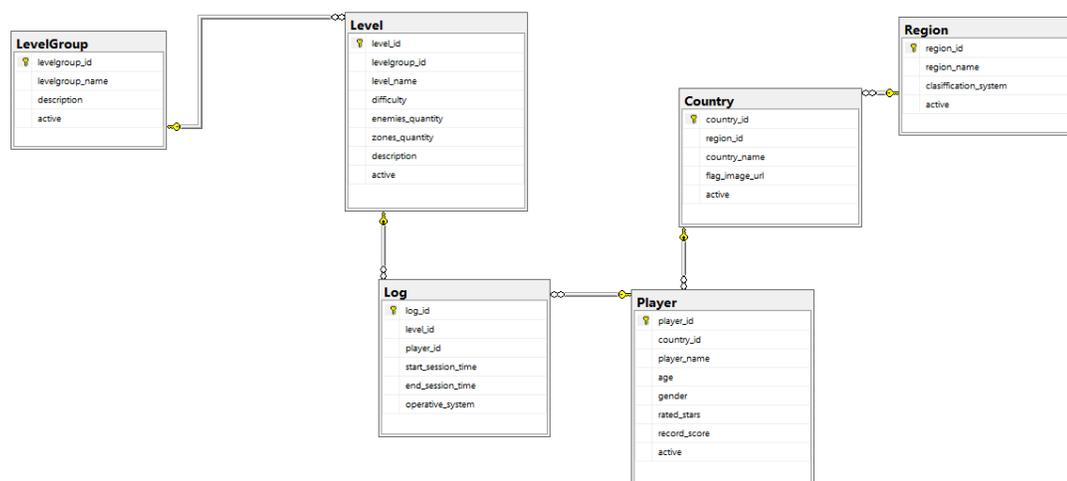
Es por ello que los métodos que se usarán en este proyecto son:

- Análisis cualitativo
- Análisis cuantitativo
- Análisis estadístico
- Análisis predictivo

4.1. Fuente de datos

La fuente de la base de datos de origen está almacenada en Amazon Web Service, la cual tiene levantada una base de datos SQL Server. El videojuego tiene una base de datos local en SQLite la cuál sincroniza la data con el servicio de Amazon.

Esta base de datos tiene 6 tablas las cuales son “LevelGroup”, “Level”, “Log”, “Player”, “Country” y “Region”. En la tabla “LevelGroup” se guarda los rubros de negocio que estarán disponibles dentro del videojuego como por ejemplo el rubro de Tecnología, Tradicionales, Deportes, entre otros. La tabla “Level” guardará cada nivel o negocio dentro del videojuego como por ejemplo Tienda de Repostería, Club de Futbol, Empresa de Soporte Técnico, entre otros. La tabla “Player” guardará la información del jugador, la tabla “Region” guarda los continentes donde está disponible el videojuego, la tabla “Country” guarda los países donde está disponible el videojuego y la tabla “Log” guarda un historial breve de sesiones del jugador. El esquema del diseño de esta base de datos es la siguiente:

Figura 3.*Diagrama de Base de Datos Transaccional*

Nota. El gráfico representa al modelo de la base de datos transaccional que maneja la empresa.

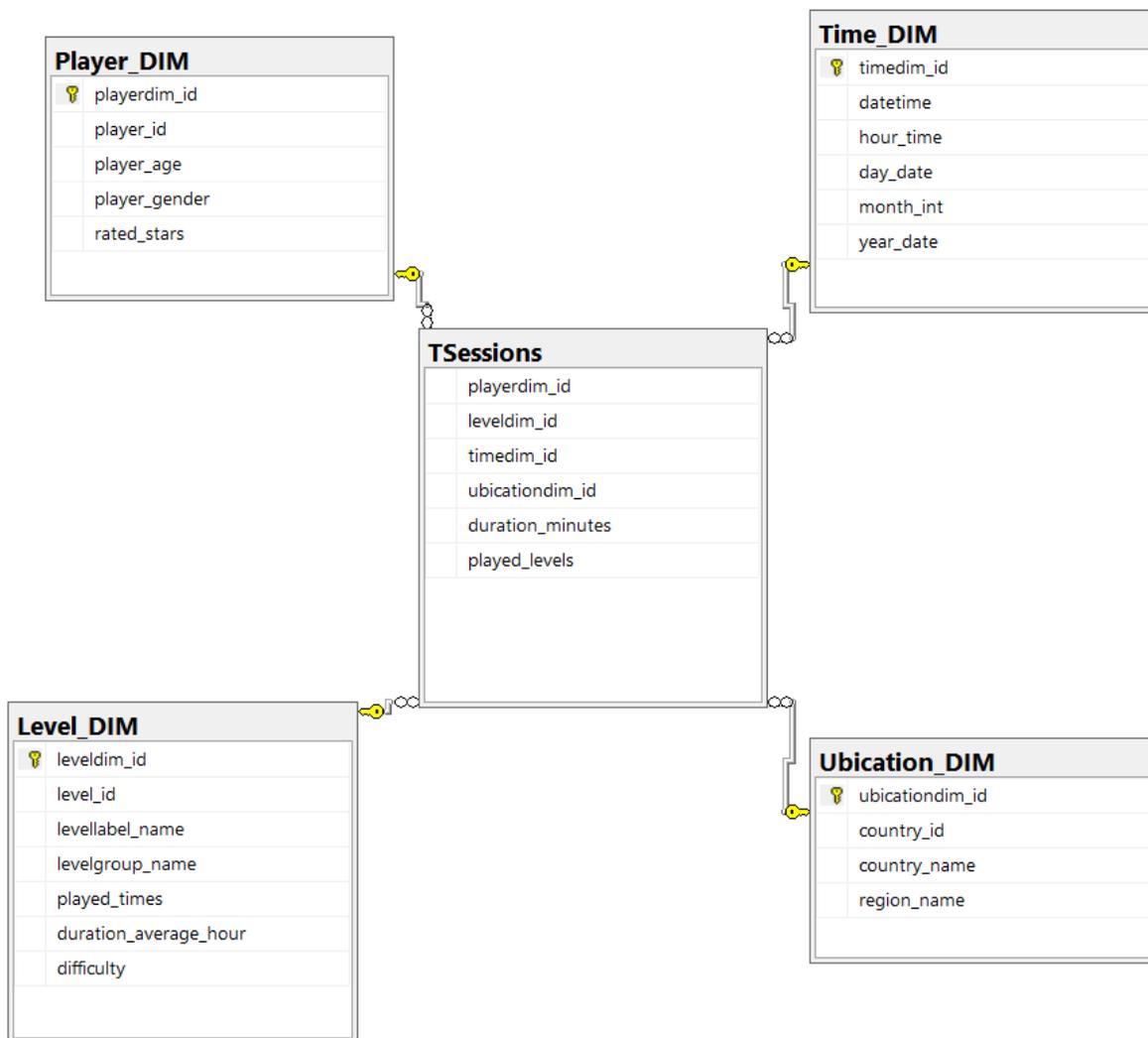
4.2. Modelamiento multidimensional

4.2.1. Modelo de una base de datos

Se ha trabajado con el modelo estrella para el modelamiento multidimensional debido a que permite posicionar como enfoque principal a una tabla hecho, que esta vez se ha llamado “TSessions”. Esto porque en la tabla de origen hay una tabla Log que permite guardar un historial de actividad del jugador. Entonces este modelo se ajusta más al modelado de la base de datos origen. El modelamiento que se ha desarrollado es el siguiente:

Figura 4.

Diagrama de Base de Datos del Data Mart



Nota. El gráfico representa al modelo de la base de datos dimensional creada para el proyecto.

4.2.2. Definiendo el modelo

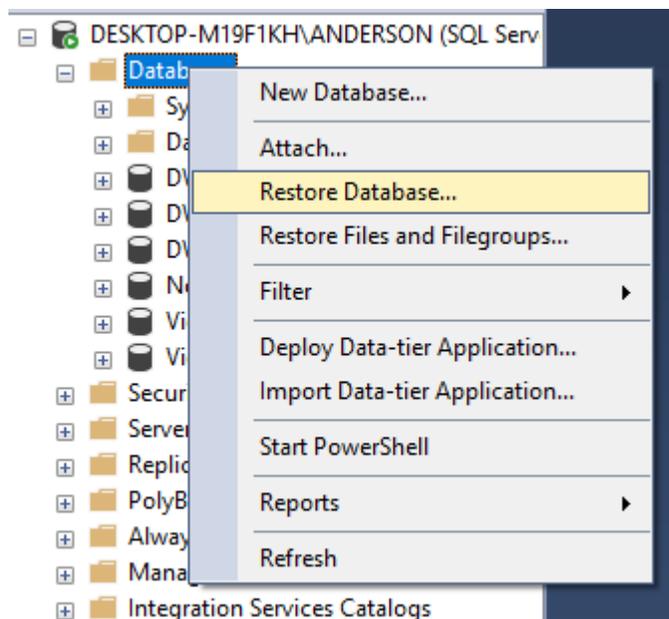
Se ha creado 4 dimensiones “Player_DIM” que guarda los datos de un jugador, “Level_DIM” guarda los datos del nivel, “Ubication_DIM” guarda los datos de los países disponibles y “Time_DIM” que guarda los datos de fecha y hora de las sesiones. La tabla principal será “TSessions” para guardar las medidas y las llaves foráneas de cada dimensión. Las medidas que se obtendrán son la duración de una sesión en minutos y las veces que se ha jugado un nivel.

4.3. Instrumento modelado

La fuente de origen se restauró con su data de agosto del 2021 hasta diciembre del mismo año. Para alojar esta base de datos se usó el gestor SQL Server 2018. Para crear una base de datos a partir de un backup. Se debe dar click derecho en “Database” y elegir la opción “Restaurar Base de Datos...”. Luego se elige el archivo de backup y se da click en restaurar.

Figura 5.

Opción para restaurar backup de una base de datos

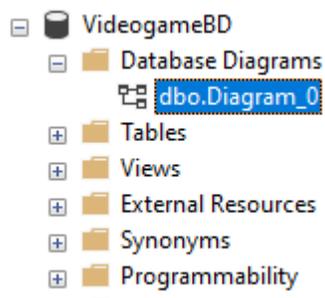


Nota. El gráfico ha sido obtenido con una captura de pantalla obtenida durante el proceso para restaurar la base de datos.

Se puede observar de manera general la estructura de la base de datos de la fuente de origen con el diagrama. Para ver el diagrama se despliega “Diagramas de Base de Datos” y se da doble click izquierdo.

Figura 6.

Opción para ver el diagrama de la base de datos transaccional



Nota. El gráfico ha sido obtenido con una captura de pantalla en SQL Server.

Se creó una nueva base de datos que servirá para alojar el Data Mart. Antes se tuvo que modelar sus dimensiones y tabla central de Sesiones, tomando en cuenta el diagrama estrella.

Antes de empezar el proceso de ETL se probó desde SQL Server. Para eso se crearon “queries” es decir consultas para la base de datos. Se crearon 4 consultas para el llenado de cada dimensión y una más para la tabla de sesiones.

Para obtener el registro para la dimensión de jugadores se selecciona los campos del ID del jugador, la edad, el enero y la valoración de estrellas. Todos estos datos se encuentran en la tabla Player de la base de datos transaccional.

Figura 7.

Query para el llenado de la dimensión Player

```
----- Consulta para llenar Player DIM
select player_id,age as player_age, gender as player_gender, rated_stars from Player
```

Nota. El gráfico se ha obtenido en una captura de pantalla de la consulta SQL para obtener los datos para la dimensión Player.

Para obtener el registro para la dimensión de los niveles se selecciona los campos del ID del nivel, el nombre del nivel, el grupo de nivel, el contador de veces jugados, el promedio de horas jugadas y la dificultad del nivel. Todos estos datos se obtienen de las tablas Log, Level y LevelGroup. Para unificar estos datos de distintas tablas se usa inner join como sentencia de SQL. Al final se ordenan con group by teniendo en cuenta el ID del nivel, el nombre, la dificultad y el nombre del grupo de niveles. Esto es para que no se repitan datos.

Figura 8.

Query para el llenado de la dimensión Level

```
-- CONSULTA DEFINITIVA PARA NIVEL DIM
select lv.level_id,lv.level_name as levellabel_name, lvg.levelgroup_name, count(lv.level_id) as played_times,
AVG(DATEDIFF(hour,lg.start_session_time,lg.end_session_time)) as duration_average_hour, lv.difficulty from Log lg
inner join Level lv on lv.level_id =lg.level_id
inner join LevelGroup lvg on lv.levelgroup_id = lvg.levelgroup_id
group by lv.level_id, lv.level_name, lv.difficulty, lvg.levelgroup_name
```

Nota. El gráfico se ha obtenido en una captura de pantalla de la consulta SQL para obtener los datos para la dimensión Level.

Para obtener el registro para la dimensión de tiempo se selecciona los campos del de fecha y hora de inicio de una sesión, la hora en número, el día de la fecha en número, el mes de la fecha en número y el año. Todos estos datos se encuentran en la tabla Log de la base de datos transaccional.

Figura 9.

Query para el llenado de la dimensión Time

```
-----Consulta para llenar dimension Time
select distinct lg.start_session_time as 'datetime',
DATEPART(HOUR,lg.start_session_time) as hour_time,
DAY(lg.start_session_time) as day_time,
MONTH(lg.start_session_time) as month_time,
YEAR(lg.start_session_time) as year_time
from Log lg
```

Nota. El gráfico se ha obtenido en una captura de pantalla de la consulta SQL para obtener los datos para la dimensión Time.

Para obtener el registro para la dimensión de ubicación se selecciona los campos del ID del país, el nombre del país, la región o continente del país. Todos estos datos se encuentran en las tablas Country y Region de la base de datos transaccional. Se usa inner join para interceptar los datos de estas dos tablas.

Figura 10.

Query para el llenado de la dimensión Ubication

```
-----Consulta para dimension de ubicación
select distinct c.country_id, c.country_name, r.region_name from Country c
inner join Region r on r.region_id = c.region_id
```

Nota. El gráfico se ha obtenido en una captura de pantalla de la consulta SQL para obtener los datos para la dimensión Ubication.

Para obtener el registro para la tabla hechos o sessions se selecciona los campos del ID de cada dimensión es decir de Player, Level, Time y Ubication. También se seleccionan los campos del contador de veces jugados de un nivel y la duración en minutos jugados. Los IDs se encuentran en las tablas dimensionales y los campos de las medidas se obtienen de las tablas Log y Player de la base de datos transaccional.

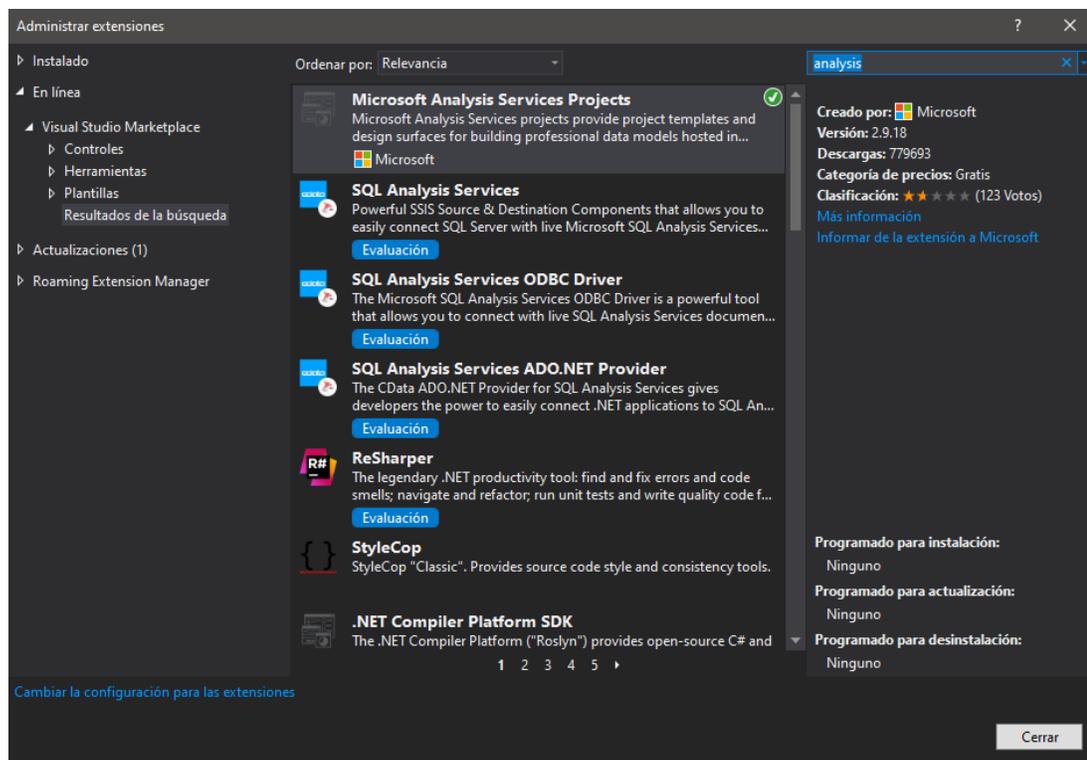
Figura 11.

Query para el llenado de la tabla Sessions

```
-----Consulta para llenar dimension hechos
select dp.playerdim_id, dl.leveldim_id, dt.timedim_id, du.ubicaciondim_id, tlg.duration_minutes, COUNT(dl.level_id) as played_levels
from (select *, DATEDIFF(minute, start_session_time, end_session_time) as duration_minutes from Log) tlg
INNER JOIN [VideogameBD-DM].dbo.Player_DIM dp ON tlg.player_id = dp.player_id
INNER JOIN [VideogameBD-DM].dbo.Level_DIM dl ON tlg.level_id = dl.level_id
INNER JOIN [VideogameBD-DM].dbo.Time_DIM dt ON tlg.start_session_time = dt.datetime
INNER JOIN Player tp ON tp.player_id = dp.player_id
INNER JOIN [VideogameBD-DM].dbo.Ubication_DIM du ON tp.country_id = du.country_id
GROUP BY tlg.log_id, tlg.duration_minutes, dp.playerdim_id, dl.leveldim_id, dt.timedim_id, du.ubicaciondim_id
```

Nota. El gráfico se ha obtenido en una captura de pantalla de la consulta SQL para obtener los datos para la tabla Sessions.

Para iniciar el proceso ETL se usó el programa Visual Studio 2019. Se tuvo que instalar la extensión de “Microsoft Analysis Services Projects” de Microsoft.

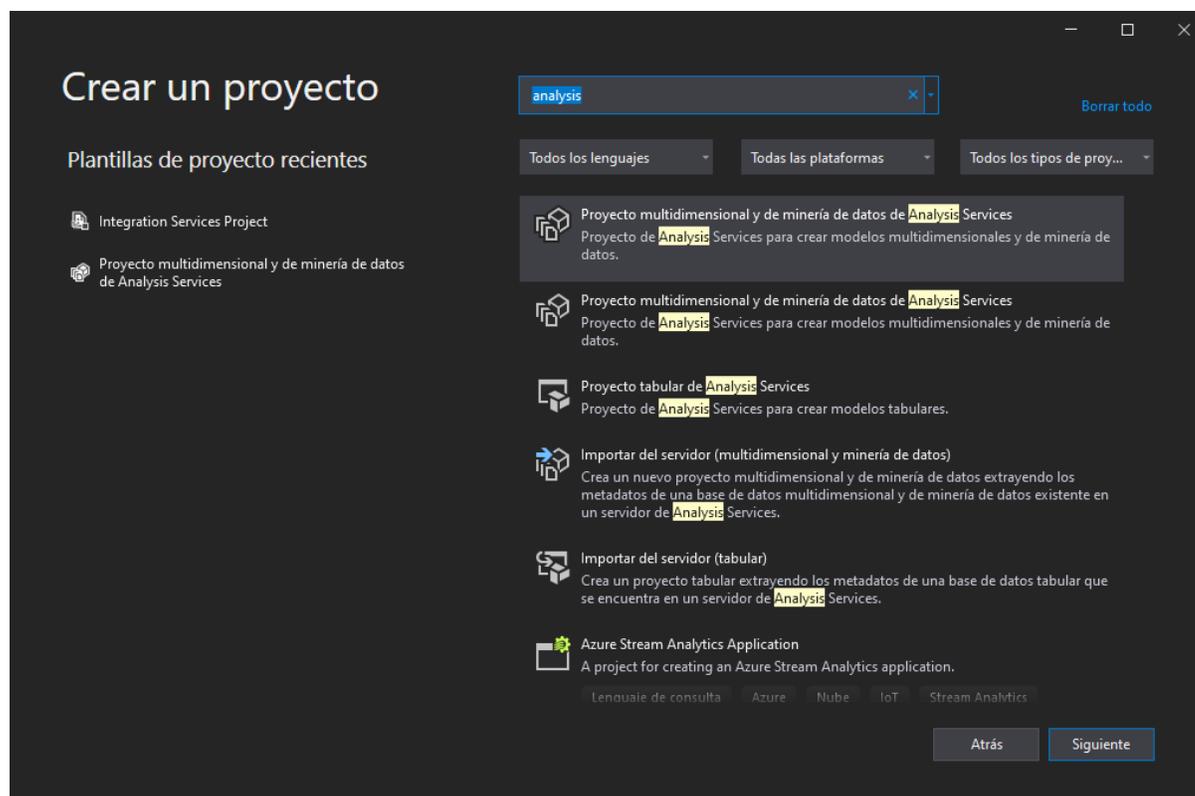
Figura 12.*Instalación de extensión Microsoft Analysis Services Projects*

Nota. El gráfico se ha obtenido en una captura de pantalla al instalar la extensión a usar desde Visual Studio.

Luego se creó un nuevo proyecto de “Proyecto Multidimensional y de Minería de Datos de Analysis Services”.

Figura 13.

Creación de Proyecto en Visual Studio para Analysis Services

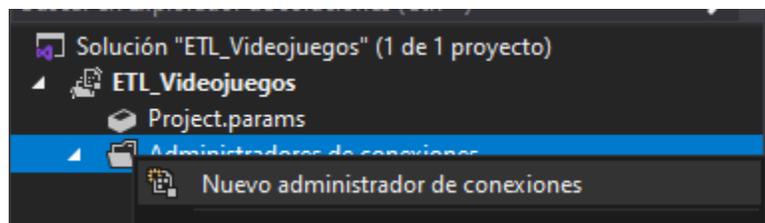


Nota. El gráfico se ha obtenido en una captura de pantalla al crear un proyecto multidimensional y minería de datos en Visual Studio.

Se genera la conexión con la base de datos se da click derecho en “Administraciones de Conexiones”. Y se elige la opción “Nuevo Administrador de Conexiones”.

Figura 14.

Opción para crear conexión entre las bases de datos

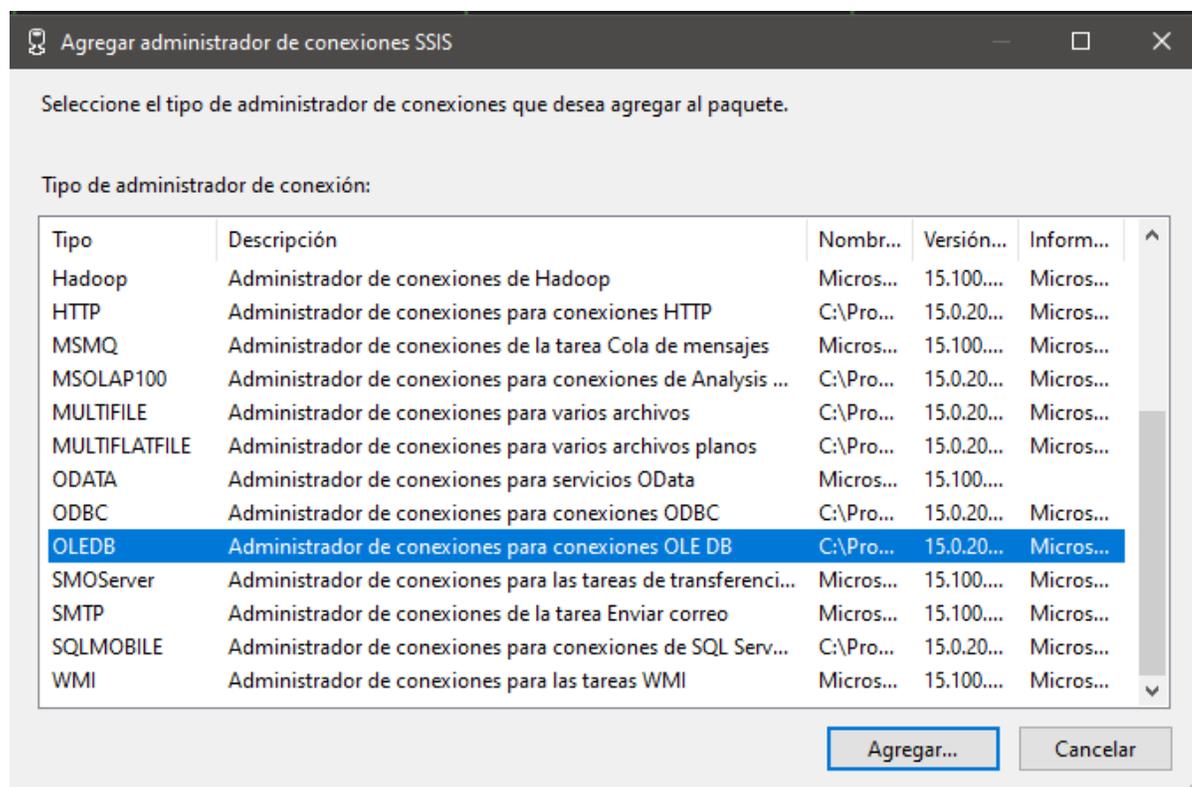


Nota. El gráfico se ha obtenido en una captura de pantalla al crear una nueva conexión con desde Visual Studio con la base de datos alojada en SQL Server.

Luego se eligió OLEDB como tipo de base de datos debido a que el gestor donde están las bases de datos es en SQL Server. Luego se añade la conexión de la base de datos y se elige las bases de datos de la fuente de origen y del data mart.

Figura 15.

Opción para elegir el tipo de administrador de conexión

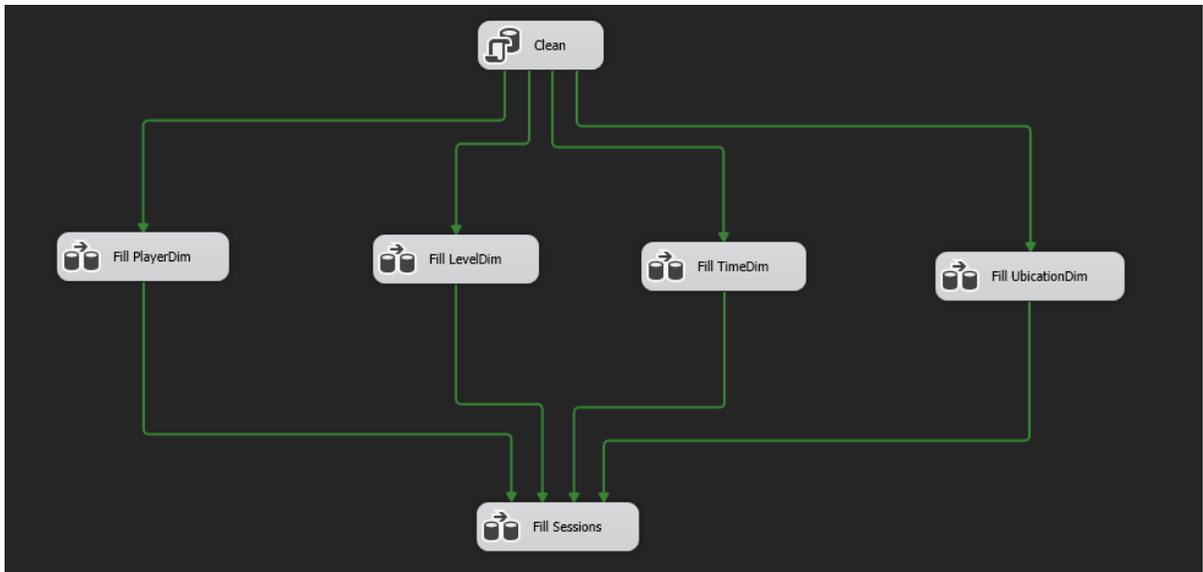


Nota. El gráfico se ha obtenido en una captura de pantalla de los tipos de administradores de conexión.

Luego se añaden las tareas que se ejecutarán en el proceso ETL. Primero se limpiarán los datos del data mart para empezar a llenar las dimensiones. Y luego de esto ya se puede llenar la tabla de sesiones.

Figura 16.

Flujo de proceso ETL

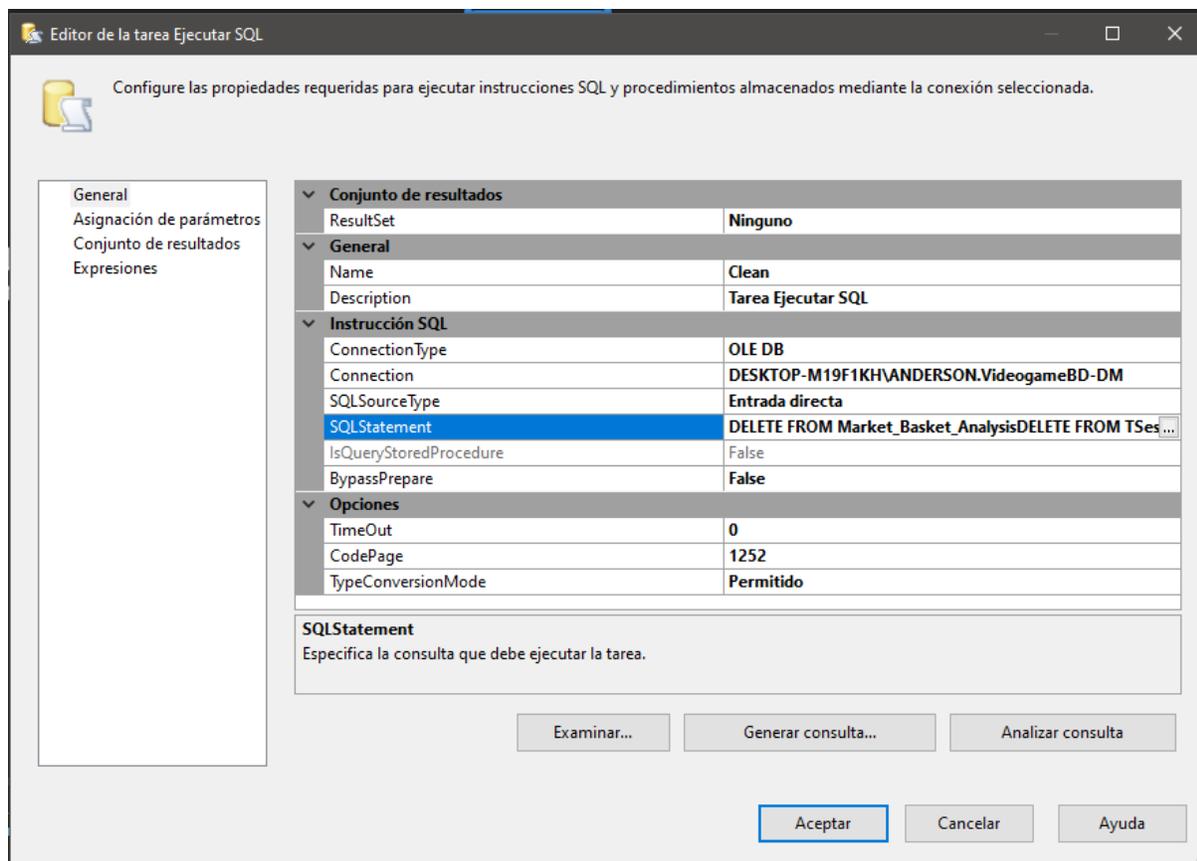


Nota. El gráfico se ha obtenido en una captura de pantalla del flujo en Visual Studio del proceso ETL.

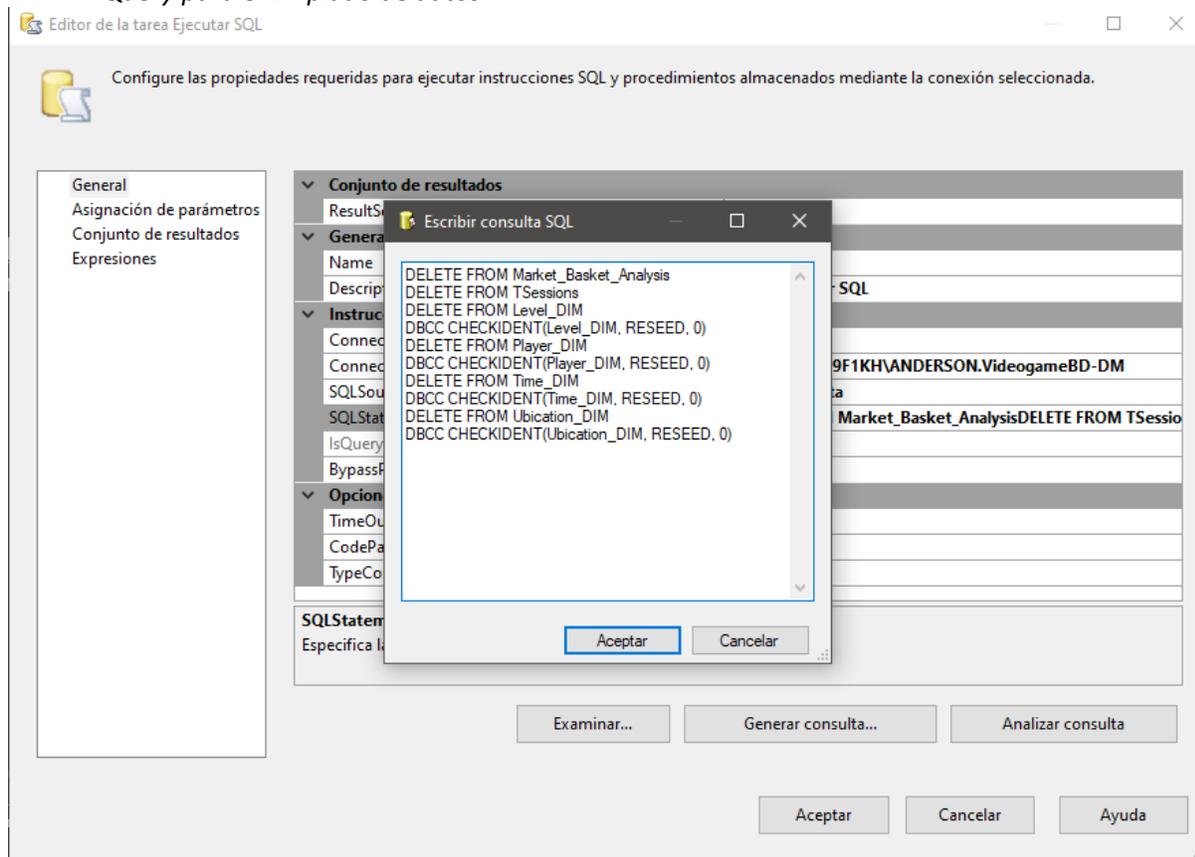
Para la tarea de limpiar datos se elige la conexión del data mart y se inserta la sentencia SQL para que borre los registros y se reinicie las llaves identity. Es decir, para que, con los nuevos datos, las llaves primarias no continúen el conteo anterior sino vuelva a iniciar desde 1.

Figura 17.

Tarea para el limpiado de registros



Nota. El gráfico se ha obtenido en una captura de pantalla del editor de tarea.

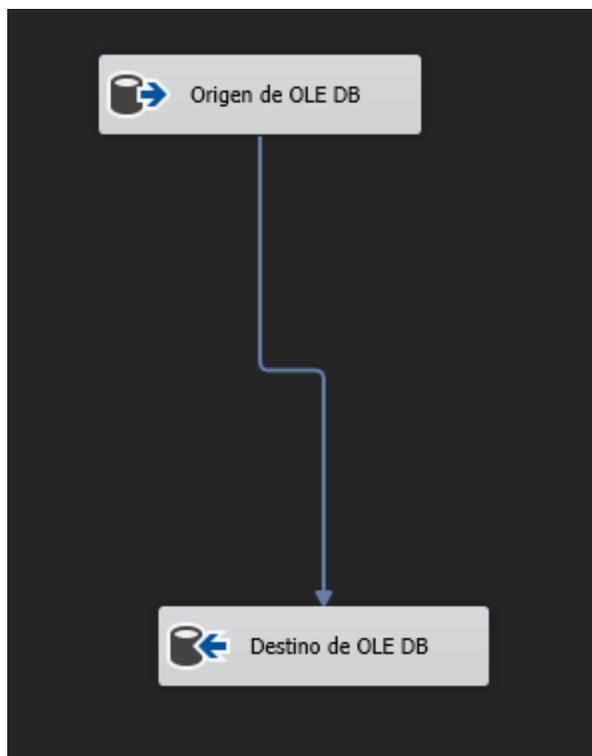
Figura 18.*Query para el limpiado de datos*

Nota. El gráfico se ha obtenido en una captura de pantalla del editor de tarea luego de colocar la sentencia para eliminar datos de las tablas de dimensiones y sesiones.

Para las tareas de llenado de cada dimensión se elige la fuente de origen OLE DB y la fuente de destino OLE DB.

Figura 19.

Flujo para las tareas de llenado de datos

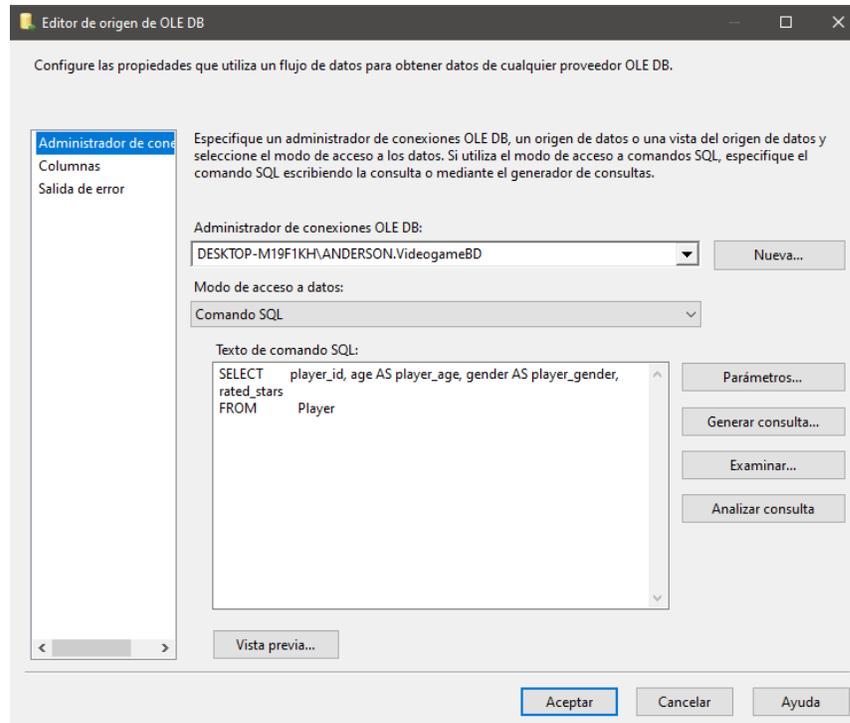


Nota. El gráfico se ha obtenido en una captura de pantalla del flujo para el llenado de datos que se usará para el llenado de datos de las dimensiones y tabla sesión.

En el destino se elige la base de datos y se agrega la sentencia SQL que previamente se ha validado en SQL Server. Para corroborar que todo estuvo bien en la vinculación de la consulta y los campos de la base de datos se puede verificar seleccionando la opción de Analizar Consulta.

Figura 20.

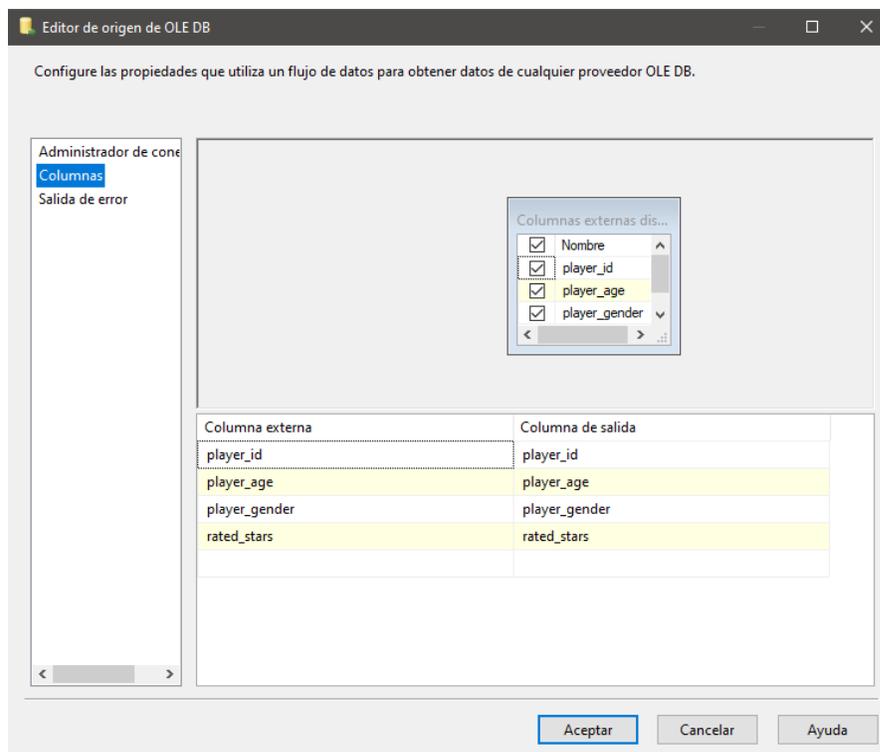
Origen OLE DB para el llenado de datos de Player



Nota. El gráfico se ha obtenido en una captura de pantalla del editor de origen donde está la consulta para el llenado de datos.

Figura 21.

Vinculación de la consulta y las columnas de la tabla

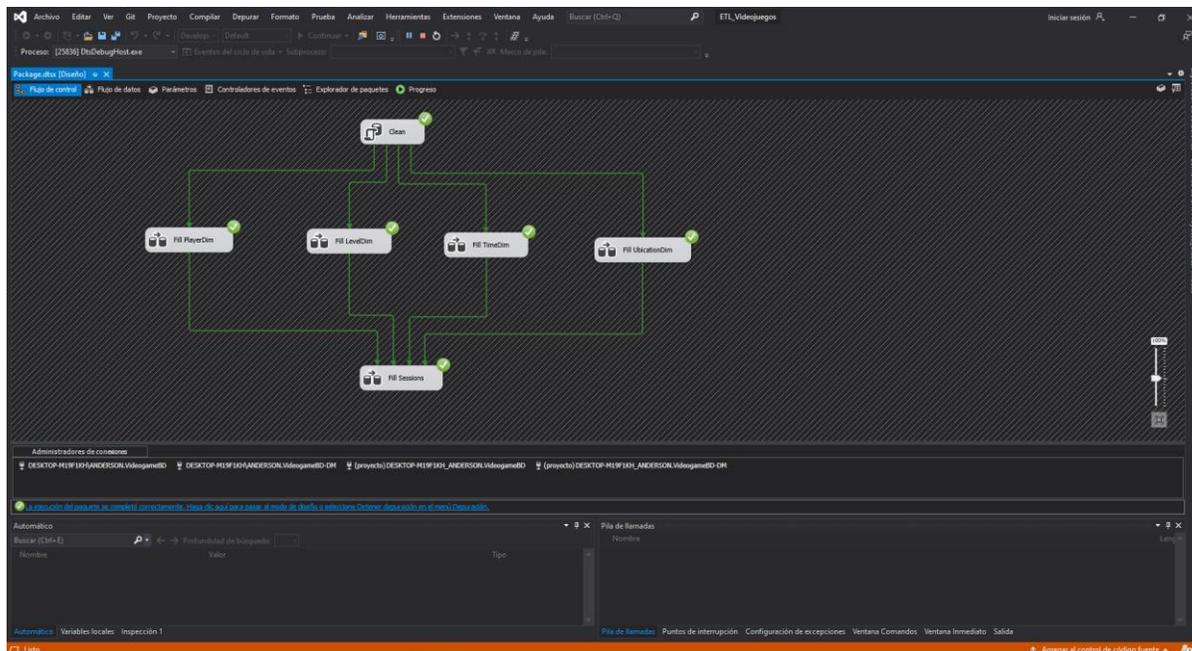


Nota. El gráfico se ha obtenido en una captura de pantalla del editor de origen luego de hacer la consulta y verificar que se hayan creado las columnas necesarias.

Este proceso se hace para todas las dimensiones y hechos. Luego se ejecuta el proceso para comprobar que no hay errores y funcione.

Figura 22.

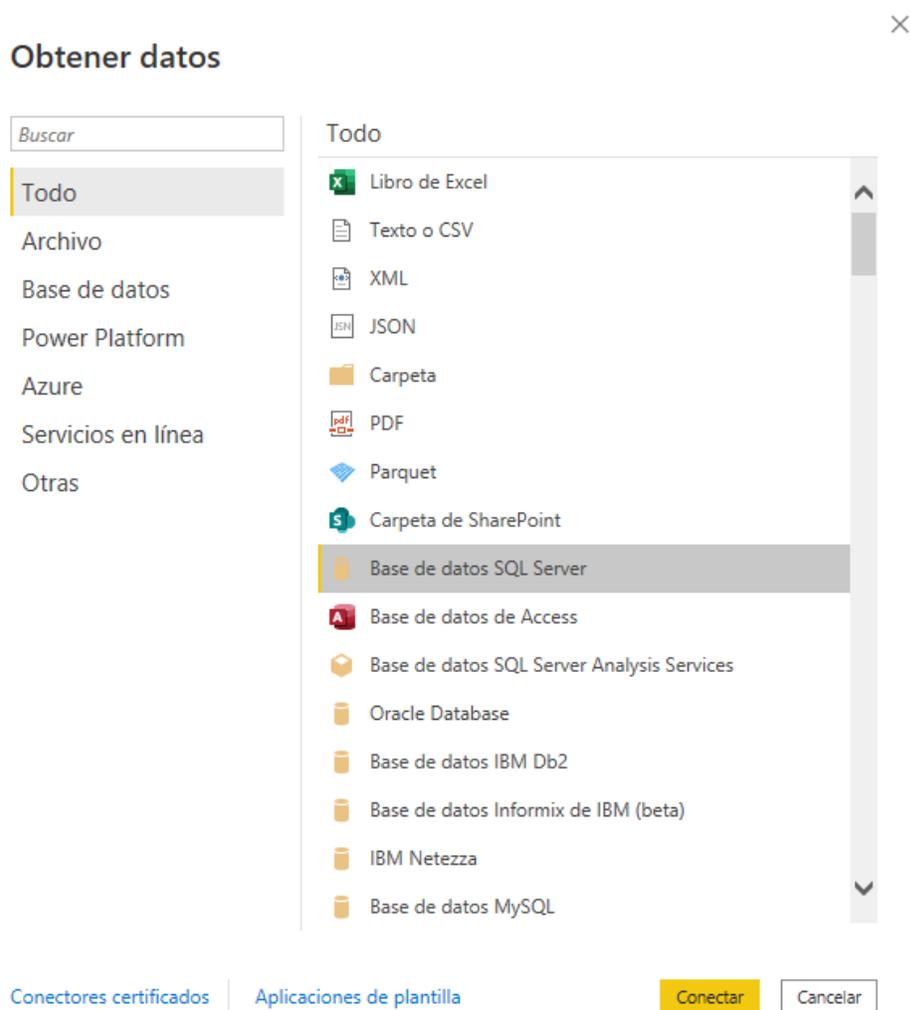
Ejecución de proceso ETL



Nota. El gráfico se ha obtenido en una captura de pantalla de la ejecución del proceso ETL y validar el correcto proceso en Visual Studio.

Se puede observar que todas las tareas tienen los íconos de check verde. Esto indica que el proceso se ha ejecutado correctamente.

Para la realización del sistema de analítica y reportería se ha usado el programa Power BI. Primero se debe crear un nuevo proyecto y se debe elegir la conexión con la base de datos de solo el data mart. La cual ya se encuentra poblada después del proceso ETL.

Figura 23.*Elección de Origen de Datos en Power BI*

Nota. El gráfico se ha obtenido en una captura de pantalla en Power BI para elegir el tipo de base de datos de origen para hacer el sistema de reportería.

Figura 24.

Conexión con base de datos en Power BI



Base de datos SQL Server ✕

Servidor ⓘ

Base de datos (opcional)

Modo Conectividad de datos ⓘ
 Importar
 DirectQuery

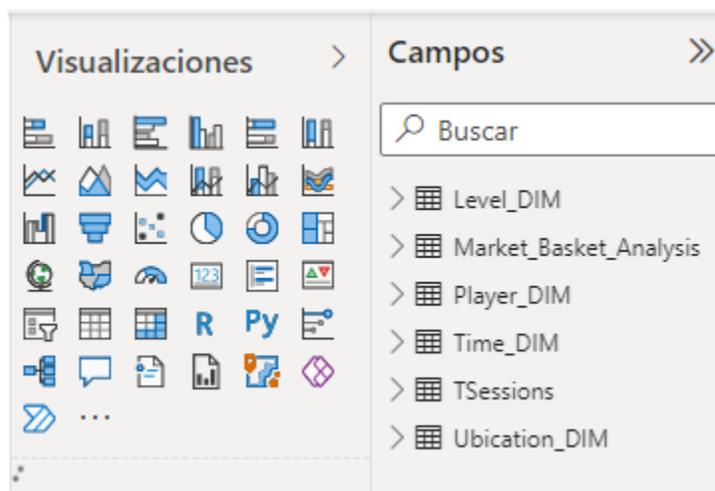
▸ Opciones avanzadas

Nota. El gráfico se ha obtenido en una captura de pantalla de la asignación de la base de datos alojada en SQL Server.

Para los reportes solo se debe arrastrar los elementos que se desean usar y vincularlas con los campos de las dimensiones y las medidas de la tabla Sesiones. Para las gráficas se usan elementos visuales como gráfico de barras y de líneas.

Figura 25.

Visualizaciones y Campos en Power BI

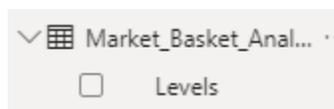


Nota. El gráfico se ha obtenido en una captura de pantalla de las tablas creadas de la base de datos multidimensional importada en Power BI desde SQL Server.

Para generar el Análisis de Canasta de Mercado se usará el objeto visual para Python. Pero previamente se generará una nueva tabla para filtrar los datos según los niveles jugados cada día.

Figura 26.

Tabla para Análisis de Canasta de Mercado



Nota. El gráfico se ha obtenido en una captura de pantalla de la tabla creada exclusivamente para el Análisis de Canasta de Mercado.

Figura 27.

Datos de la tabla Market_Basket_Analysis

Levels
Cake Shop
Cake Shop,Cake Shop,Ed Tech,FinTech,Market,Football,Foo
Cake Shop,Cake Shop,FinTech,Football,Football,Football,Ed T
Cake Shop,Cake Shop,FinTech,Market,Market,Market,Foot
Cake Shop,Cake Shop,Football,Ed Tech,Cake Shop,FinTech,F
Cake Shop,Cake Shop,Football,FinTech,Cake Shop,Cake Sho
Cake Shop,Ed Tech,FinTech,Market,Cake Shop,Cake Shop,C
Cake Shop,Ed Tech,Football,Cake Shop,FinTech,FinTech,Cak
Cake Shop,Ed Tech,Market,Football,Ed Tech,FinTech,Footba
Cake Shop,Ed Tech,Market,Market,Cake Shop,Football,Ed T
Cake Shop,Football,Cake Shop,FinTech,Football,FinTech,Fin
Cake Shop,Football,Cake Shop,Football,Ed Tech,Cake Shop,C
Cake Shop,Football,Ed Tech,Cake Shop,FinTech,FinTech,Fin
Cake Shop,Football,Ed Tech,Cake Shop,FinTech,Football,Fin
Cake Shop,Football,FinTech,FinTech,Market,FinTech,FinTeci
Cake Shop,Football,FinTech,Market,FinTech,Market,Cake Sh
Cake Shop,Football,Football,Cake Shop,Cake Shop,Football,F
Cake Shop,Football,Football,Cake Shop,Market,FinTech,Mar
Cake Shop,Football,Market,Cake Shop
Cake Shop,Football,Market,Cake Shop,Ed Tech,Cake Shop,E
Cake Shop,Football,Market,Football,Cake Shop,Football,Fin
Cake Shop,Market
Cake Shop,Market,Ed Tech,Cake Shop,FinTech,Market,Ed T
Cake Shop,Market,FinTech,Cake Shop,FinTech,Ed Tech
Cake Shop,Market,FinTech,Football,Football,Cake Shop,Ed T
Cake Shop,Market,Football,FinTech,Ed Tech,Market,Footba
Cake Shop,Market,Football,Market,Football,FinTech,Market
Ed Tech

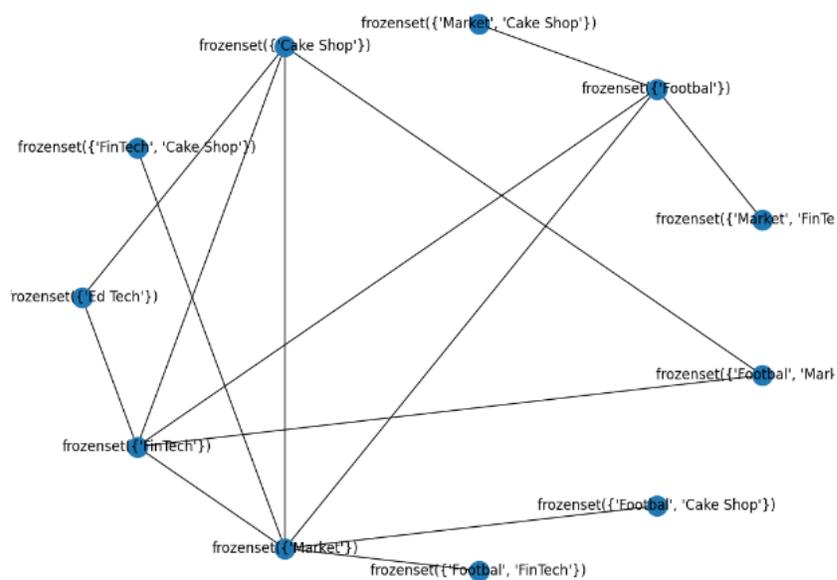
Nota. El gráfico se ha obtenido en una captura de pantalla de las combinaciones de los niveles que se están guardando en la tabla para el Análisis de Canasta de Mercado.

Luego se volverá a la página de reportería y se añadirá el código para hacer un diagrama de redes y obtener las combinaciones de los niveles que se juegan más.

Figura 28.

Gráfico MBA para las combinaciones de niveles

Levels



Nota. El gráfico se ha obtenido en una captura de pantalla del diagrama de redes que muestran las combinaciones obtenidas.

Primero se importan las librerías necesarias a usar.

Figura 29.

Código Python para implementar librerías

```
import matplotlib.pyplot as plt
import pandas as pd
from mlxtend.frequent_patterns import apriori
from mlxtend.frequent_patterns import association_rules
from mlxtend.preprocessing import TransactionEncoder
import numpy as np
from mlxtend.frequent_patterns import association_rules
from mlxtend.frequent_patterns import apriori
import networkx as nx
from matplotlib.pyplot import figure
```

Nota. El gráfico se ha obtenido en una captura de pantalla de las librerías usadas para el script de Python que ejecuta el análisis de canasta de mercado.

Luego se transforman los datos.

Figura 30.

Código Python para transformación de datos

```
mydata = dataset.Levels
# Aplicamos una transformación para los registros (divisor por ,)
transacciones = mydata.apply(lambda t: t.split(','))
transacciones = list(transacciones)
codificador = TransactionEncoder().fit(transacciones)
# Transformado los datos originales
transformado = codificador.transform(transacciones)
tranf_ejemplo = pd.DataFrame(transformado, columns = codificador.columns_)
.....
```

Nota. El gráfico muestra que los datos de los niveles se guardan en una nueva variable para hacer la transformación de datos usando el método fit() y luego el método transform(). Una vez transformado se guarda en una variable llamada tranf_ejemplo.

Luego se crea filtran los datos para obtener solo las combinaciones cuyo soporte sea mayor o igual de 0.5. Esta regla fue establecida por la propia empresa.

Figura 31.

Código Python para obtener datos con regla de soporte

```
frecuencia_itemset = apriori(tranf_ejemplo, min_support = 0.001, max_len = 3, use_colnames=True)
reglas_1=association_rules(frecuencia_itemset, metric = 'lift', min_threshold=1).sort_values('lift', ascending=False).reset_index(drop=True)
# Mostramos algunas reglas
data_top = reglas_1[(reglas_1['support'] >= 0.5)]
.....
```

Nota. El gráfico muestra las reglas usadas para el filtrado de la data.

Finalmente se crea la gráfica.

Figura 32.

Código Python para generar gráfica de diagrama de redes

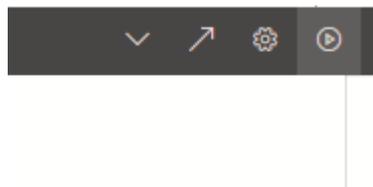
```
# Graficamos
G = nx.Graph()
G = nx.from_pandas_edgelist(data_top, 'antecedents', 'consequents')
figure(figsize=(10, 8))
nx.draw_shell(G, with_labels=True)
plt.show()
```

Nota. El gráfico muestra cómo se genera el gráfico de diagrama de redes usando la librería panda, networkx y matplotlib.

Para poder generar la gráfica se debe ejecutar el script en el ícono del triángulo de Ejecutar.

Figura 33.

Opción para ejecutar código Python



4.4. Proceso ETL

4.4.1. Extracción de Datos

Para realizar la extracción se exportó en formato de backup la base de datos alojada en Amazon y se restauró en un gestor SQL Server local. Se le puso de nombre "VideogameBD".

4.4.2. Transformación de Datos

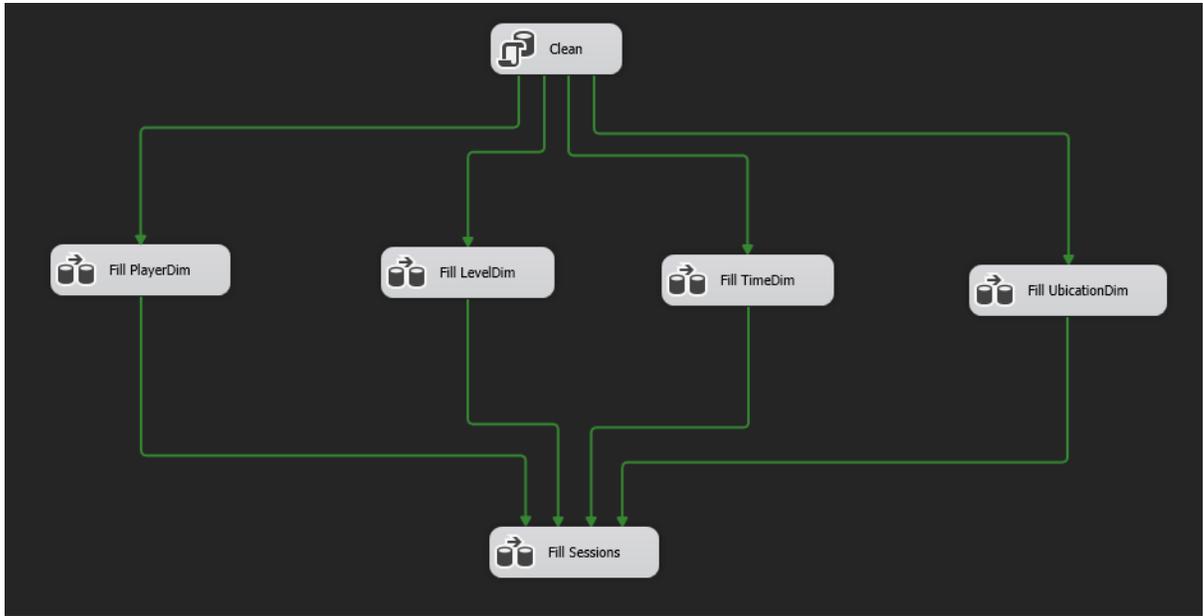
En el proceso de transformación, se limpiaron los datos de la fuente de origen. También se estandarizaron a nivela de base de datos algunos formatos como los formatos de los campos para que coincidan entre las bases de datos transaccional y del data mart, verificar que no haya espacios en los registros y tampoco errores ortográficos. Esto se hizo para que no pueda haber errores posteriores. Se hizo un control de la ortografía de los campos, principalmente de los nombres de niveles, grupos de niveles, nombre de países y nombre de regiones. Se comprobó que los datos estén de manera uniforme.

4.4.3. Carga de Datos

Se ha trabajado con el modelo estrella para el modelamiento multidimensional debido a que permite posicionar como enfoque principal a una tabla hecho, que esta vez se ha llamado "TSessions". Esto porque en la tabla de origen hay una tabla Log que permite guardar un historial de actividad del jugador. Entonces este modelo se ajusta con propiedad "Identity", es decir que su valor es incrementativo de manera automática. De esta manera cuando se vuelva a cargar los datos, las llaves identity se restauren desde 0. Luego de comprobar que las sentencias están correctas, se creó un proyecto de Servicio de Análisis en Visual Studio para generar el proceso ETL de carga. Al ejecutar el proyecto se pudo hacer la carga de la base de datos del Data Mart, llamada "VideogameBD-DM", de forma correcta.

Figura 34.

Carga de Datos en Visual Studio



Nota. El gráfico se ha obtenido en una captura de pantalla del flujo del proceso ETL usado para el proyecto en Visual Studio.

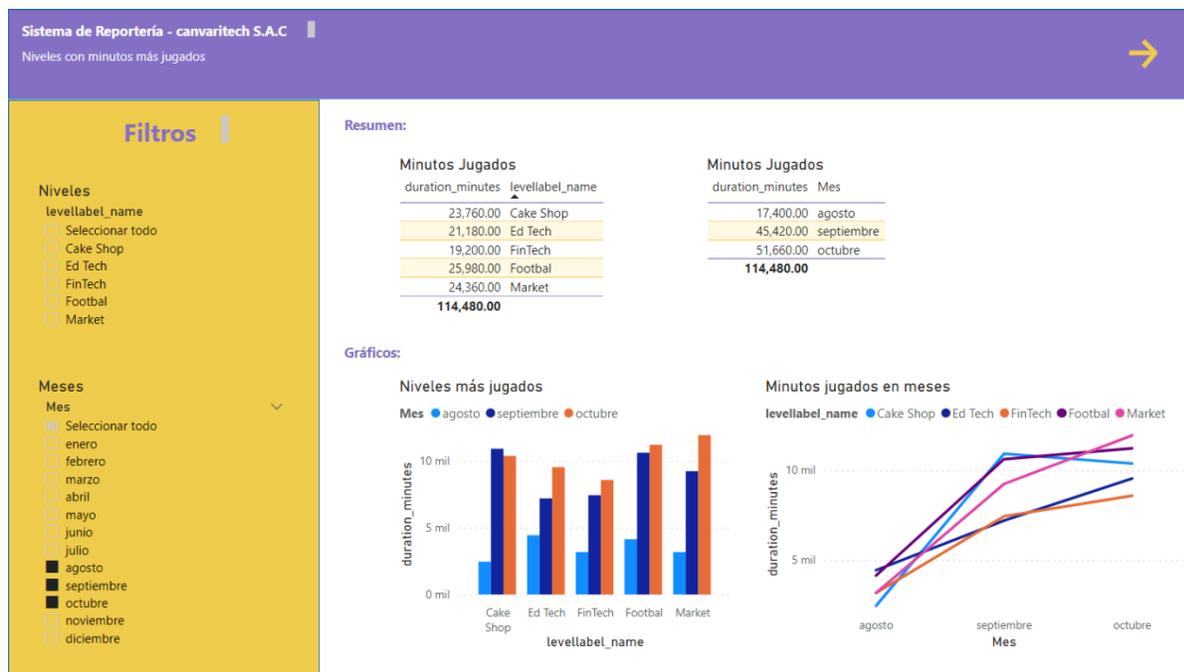
Capítulo V. Elaboración y Presentación de Sistema de Reportería

Para la elaboración del sistema de reportería se ha usado el programa de Microsoft, Power BI. Se vinculó la base de datos del Data Mart a un nuevo proyecto en este programa. Los reportes que se han sacado son sobre los niveles más jugados en base a los minutos jugados y sobre los niveles que más veces se eligió para jugar teniendo en cuenta su grado de dificultad.

El primer reporte permite mostrar los niveles más jugados mensualmente y que se pueda filtrar por mes o nivel. Por ejemplo, se puede ver que en los 3 primeros meses el nivel "Football" es el que más minutos se ha jugado. Con este reporte ayuda al objetivo de saber los niveles más jugados.

Figura 35.

Reporte sobre Niveles con más minutos jugados



Nota. El gráfico se ha obtenido en una captura de pantalla del dashboard para saber los niveles con más minutos jugados en Power BI. Se muestran los diferentes filtros y gráficas.

Figura 36.

Filtros del dashboard de los niveles con más minutos jugados



Nota. El gráfico se ha obtenido en una captura de pantalla de los filtros del dashboard para saber los niveles con más minutos jugados en Power BI.

En este primer dashboard se puede filtrar la data por niveles y el mes. Esto permite que se pueda ver cuántas veces se ha jugado cada nivel en un mes específico o un rango de meses. Así como ver durante cada mes o rango de meses que niveles han sido los más jugados.

Figura 37.

Resumen del dashboard de niveles más jugados

Resumen:

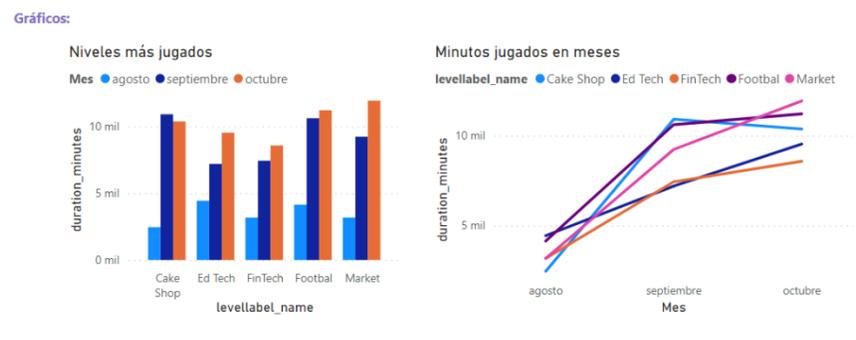
Minutos Jugados		Minutos Jugados	
duration_minutes	levellabel_name	duration_minutes	Mes
23,760.00	Cake Shop	17,400.00	agosto
21,180.00	Ed Tech	45,420.00	septiembre
19,200.00	FinTech	51,660.00	octubre
25,980.00	Footbal		
24,360.00	Market		
114,480.00			

Nota. El gráfico se ha obtenido en una captura de pantalla del resumen del dashboard para saber los niveles con más minutos jugados en Power BI.

En el dashboard se puede ver un resumen cuantitativo de los minutos jugados por nivel y por mes.

Figura 38.

Gráficas del dashboard de niveles más jugados

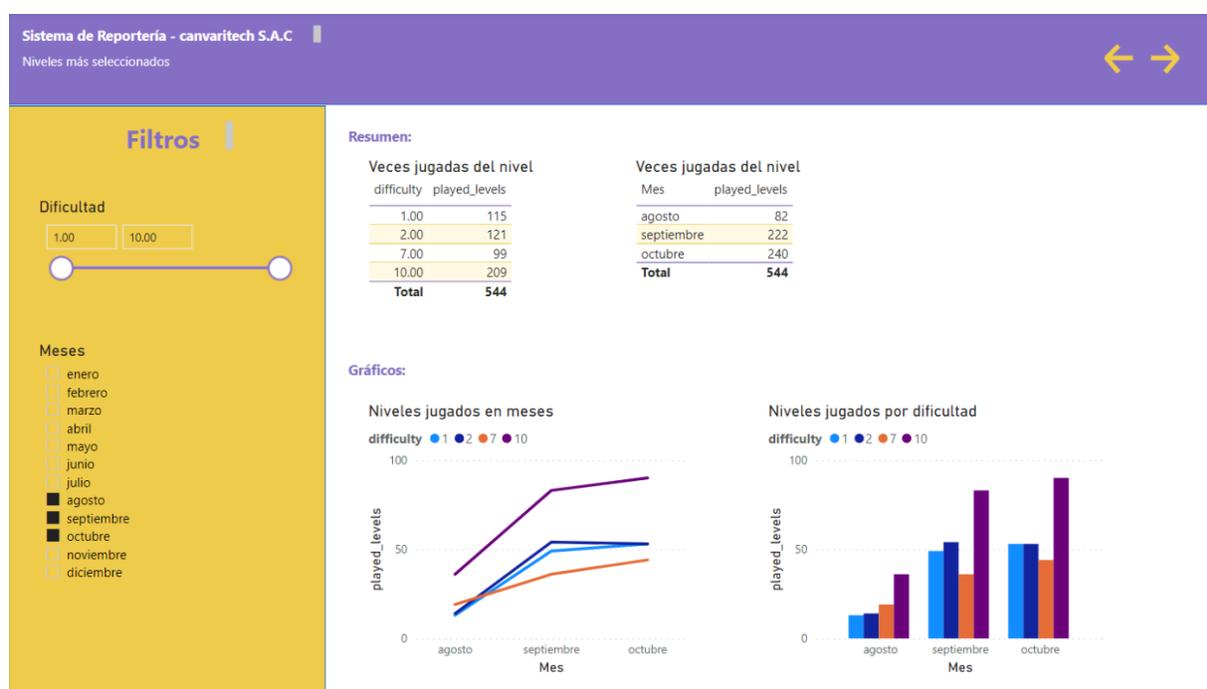


Nota. El gráfico se ha obtenido en una captura de pantalla de los gráficos del dashboard para saber los niveles con más minutos jugados en Power BI.

En el dashboard también se tienen dos gráficos. En la primera gráfica que muestran los niveles más jugados por mes, de esta manera se puede ver el crecimiento de cada nivel mediante las barras. En el segundo gráfico se ve una línea de regresión para ver los minutos jugados de cada nivel durante los meses, de esta manera se puede ir viendo si hay un pico de subida o bajada.

Figura 39.

Reporte de Niveles más escogidos



Nota. El gráfico se ha obtenido en una captura de pantalla del dashboard para saber las veces jugadas de un nivel según el grado de dificultad en Power BI. Se muestran los diferentes filtros y gráficas.

El segundo reporte se puede ver la cantidad de veces que un nivel es elegido mensualmente y se puede filtrar por la dificultad del nivel y los meses. Por ejemplo, el decisor puede ver que los niveles que más se escogen para jugar son los que tiene una dificultad máxima.

Figura 40.

Filtros del dashboard de los niveles más escogidos



Nota. El gráfico se ha obtenido en una captura de pantalla de los filtros del dashboard para saber los niveles más elegidos en Power BI.

En este segundo dashboard se puede filtrar la data por el número de dificultad y meses. Esto permite que se pueda ver si la dificultad puede afectar los minutos jugados.

Figura 41.*Resumen del dashboard de los niveles más elegidos*

Resumen:

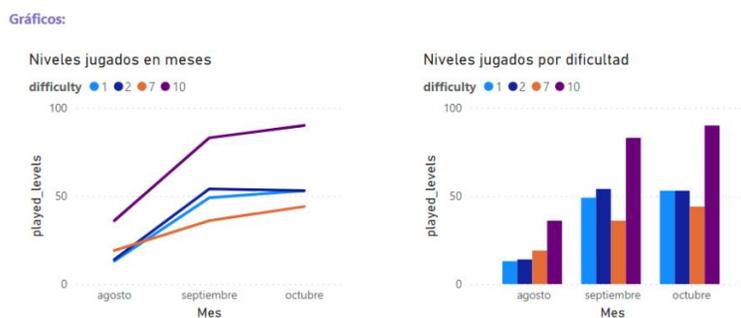
Veces jugadas del nivel		Veces jugadas del nivel	
difficulty	played_levels	Mes	played_levels
1.00	115	agosto	82
2.00	121	septiembre	222
7.00	99	octubre	240
10.00	209	Total	544
Total	544		

Nota. El gráfico se ha obtenido en una captura de pantalla del resumen del dashboard para saber los niveles más elegidos en Power BI.

En el dashboard se puede ver un resumen cuantitativo de las veces jugadas de un nivel, es decir cuántas veces se escogió un nivel para ser jugado. Esto se muestra por cada dificultad que tienen los niveles y por meses.

Figura 42.

Gráficos del dashboard de los niveles más escogidos



Nota. El gráfico se ha obtenido en una captura de pantalla de los gráficos del dashboard para saber los niveles más elegidos en Power BI.

En el dashboard también se tienen dos gráficos. En la primera gráfica muestran las veces jugadas de cada dificultad de nivel durante los meses, de esta manera se puede ver si el grado de dificultad de un nivel afecta en un crecimiento o no. En el segundo gráfico se ve las veces jugadas de cada dificultad de nivel de cada mes, de esta manera se puede ver en cada mes que grado de dificultad fue el más o menos jugado.

Capítulo VI. Aplicación Técnica de Analítica

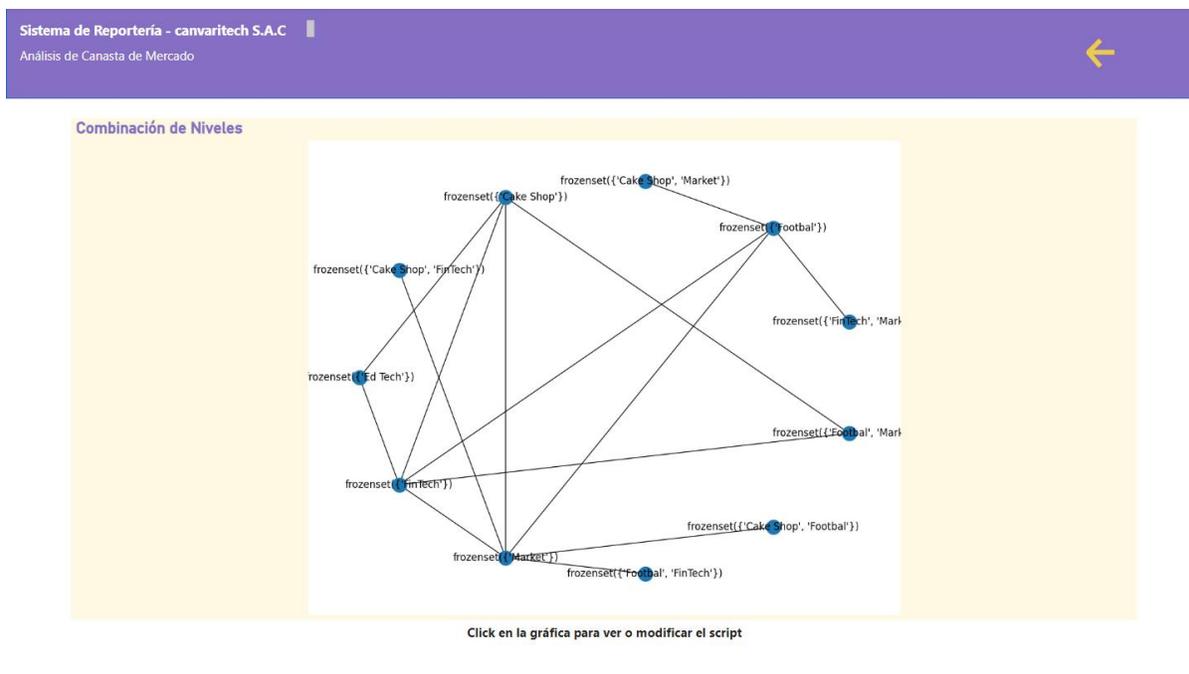
Guardar los datos y solo poder filtrarlos no permite sacar el mayor potencial para el beneficio de la empresa, pues con esos datos se puede obtener información que permita predecir y ayude a tomar decisiones a la empresa. La analítica permite ayudar a las empresas en ese sentido. Hay diferentes tipos de analítica, la que se ha usado en este proyecto es la analítica predictiva.

La analítica predictiva permite hacer predicciones, tendencias y previsiones de manera precisa a partir de un conjunto de datos, en este caso el data mart generado.

La aplicación de la analítica se ha realizado en Power BI de Microsoft con un script de Python para poder hacer un Análisis de Canasta de Mercado o MBA por sus siglas en inglés, Market Basket Analysis. Un MBA permite predecir las combinaciones de un producto o servicio a partir de un primer consumo. Por ejemplo, los clientes luego de comprar un producto A compran un producto D y producto F. Entonces la empresa puede generar una oferta o combo de estos productos. Para este proyecto, según los objetivos específicos, se desea saber las combinaciones de 3 niveles para que ayude a la empresa a crear eventos dentro del videojuego y combinar diferentes niveles.

Estos eventos donde se combinan varios niveles son muy frecuentes en videojuegos para generar variedad. En este caso el beneficio es que no solo se agruparan niveles al azar sino se agruparan niveles luego de un aplicar MBA.

El decisor puede ver en un diagrama de redes generado por el script de Python desde Power BI. Este diagrama permite ver el siguiente nivel o niveles y sus relaciones entre ellos mismos dentro del videojuego.

Figura 43.*Reporte de Análisis de Canasta de Mercado*

Nota. El gráfico se ha obtenido en una captura de pantalla del diagrama de redes que muestran las combinaciones obtenidas.

Conclusiones

Al finalizar el proyecto, la empresa tendrá un sistema de analítica para que el área de diseño de videojuegos pueda tomar decisiones en base a datos reales del videojuego. Los objetivos han sido cumplidos y se resolvieron los problemas planteados. Con este sistema la empresa puede saber los niveles más jugados en base a minutos, tener un proceso ETL en base a su base de datos del videojuego y saber las combinaciones de 3 niveles que les ayude a crear eventos dentro del videojuego.

El sistema permite escalar a otro videojuego que tenga un diseño por niveles. Además de poder soportar un volumen de datos grande que le sirva a la empresa en el tiempo.

Se pueda concluir que el proyecto ha cumplido con su objetivo y ha solucionado el problema de la empresa respecto a su área de diseño de videojuegos.

Recomendaciones

Durante el proyecto se pudo observar que la base de datos necesitó cambiar formatos para su correcto uso en la generación del data mart por lo que se recomienda modificar su base de datos montada en un servicio cloud en proveedor de servicios en la nube, con la estandarización de sus tablas hechas en este proyecto.

También se recomienda que el área tenga un micro proceso para la actualización de la base de datos multidimensional. Esto permitirá que no haya mucho volumen de datos a cargar ni tiempo de espera porque con el tiempo la base de datos transaccional va aumentando sus datos. Una buena práctica es tener un rango de tiempo establecido, por ejemplo, en una hora de madrugada una vez al mes para que se carguen los datos mientras no se tiene un uso concurrente en la base de datos transaccional.

Otra recomendación es usar una base de datos en la nube con el proveedor Azure de Microsoft para automatizar la carga y actualización de datos. La base de datos se podría ir replicando en otra alojada en Azure y dese Visual Studio se use esa base de datos para hacer el proceso ETL. La base de datos dimensional también se alojaría en Azure para poder hacer la conexión con Power BI de manera directa. Esto hará que todo el proceso del sistema desde la extracción de datos hasta mostrar los dashboards sea más automático. (Microsoft, y otros, Microsoft Corp, 2022), (Microsoft, AngelosP, ghogen, v-rajagt-zz, & hyoshioka0128, Microsoft Corp, 2022)

REFERENCIAS BIBLIOGRAFICAS

Chawla, N. V. (Abril de 2011). *ResearchGate GmbH*. Researchgate.

https://www.researchgate.net/publication/220183340_Market_basket_analysis_wth_net_works

Deepnote. (s.f.). *Deepnote*. Deepnote. [https://deepnote.com/@code-along-tutorials/Market-](https://deepnote.com/@code-along-tutorials/Market-Basket-Analysis-in-Python-An-Implementation-with-Online-Retail-Data-6231620b-cba3-4935-bde8-8ce1490868bf)

[Basket-Analysis-in-Python-An-Implementation-with-Online-Retail-Data-6231620b-cba3-4935-bde8-8ce1490868bf](https://deepnote.com/@code-along-tutorials/Market-Basket-Analysis-in-Python-An-Implementation-with-Online-Retail-Data-6231620b-cba3-4935-bde8-8ce1490868bf)

Deusto Blog. (2021). *Deusto Blog*. Deusto Blog. [https://blogs.deusto.es/master-informatica/el-bi-](https://blogs.deusto.es/master-informatica/el-bi-de-la-gran-n/)

[de-la-gran-n/](https://blogs.deusto.es/master-informatica/el-bi-de-la-gran-n/)

Equipo Indicative. (2021). *Indicative*. Indicative. [https://www.indicative.com/resource/gaming-](https://www.indicative.com/resource/gaming-analytics/)

[analytics/](https://www.indicative.com/resource/gaming-analytics/)

Microsoft, AngelosP, ghogen, v-rajagt-zz, & hyoshioka0128. (29 de Abril de 2022). *Microsoft*

Corp. Microsoft: <https://learn.microsoft.com/es-es/visualstudio/azure/azure-sql-database-add-connected-service?view=vs-2022>

Microsoft, davidiseminger, chetankt2305, TimShererWithAquent, mihart, DCtheGeek, . . . v-

[kents. \(21 de Setiembre de 2022\). *Microsoft Corp*. https://learn.microsoft.com/es-es/power-bi/connect-data/service-azure-sql-database-with-direct-connect](https://learn.microsoft.com/es-es/power-bi/connect-data/service-azure-sql-database-with-direct-connect)

Programador Clic . (2022). *Programador Clic* . Programador Clic.

<https://programmerclick.com/article/1274797018/>

Susan Li. (24 de Setiembre de 2017). *Towards Data Science*. Towards Data Science Web Site.

<https://towardsdatascience.com/a-gentle-introduction-on-market-basket-analysis-association-rules-fa4b986a40ce>

ANEXOS

Figura 44.

Encuesta de satisfacción para el área de diseño de videojuegos

Encuesta de Satisfacción

Esta encuesta está dirigida al equipo de diseño de juego para poder evaluar la satisfacción respecto al sistema de analítica realizado para el área.

[Iniciar sesión en Google](#) para guardar lo que llevas hecho. [Más información](#)

***Obligatorio**

Facilidad para usar el sistema de reportería y analítica *

1 2 3 4 5

Muy difícil Muy fácil

¿Pudiste saber los niveles más jugados del videojuego? *

Sí

No

¿Pudiste saber las mejores combinaciones de 3 niveles? *

Sí

No

Nivel en el que el sistema cumplió su objetivo *

1 2 3 4 5

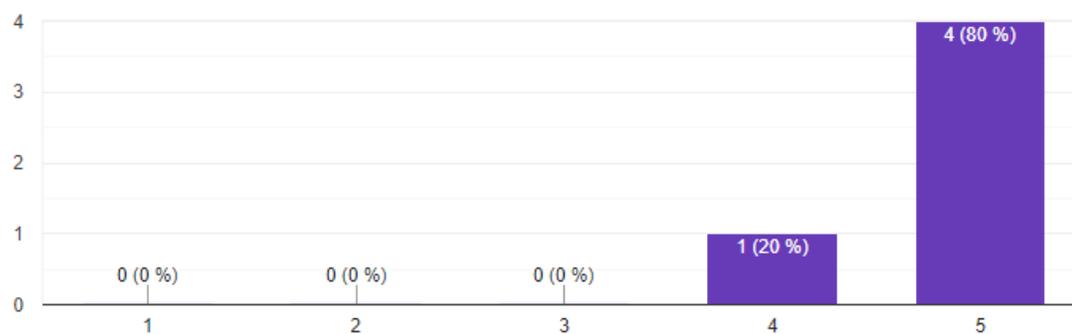
No cumplió en nada Cumplió en todo

Figura 45.

Resumen de primera pregunta de la encuesta de satisfacción

Facilidad para usar el sistema de reportería y analítica

5 respuestas

**Figura 46.**

Resumen de segunda pregunta de la encuesta de satisfacción

¿Pudiste saber los niveles más jugados del videojuego?

5 respuestas

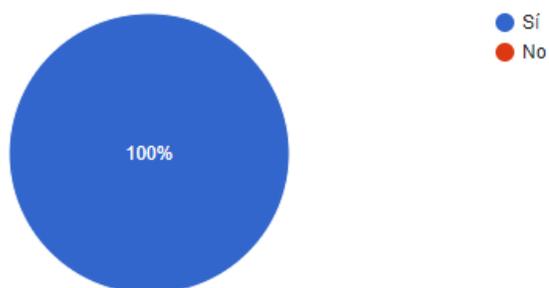
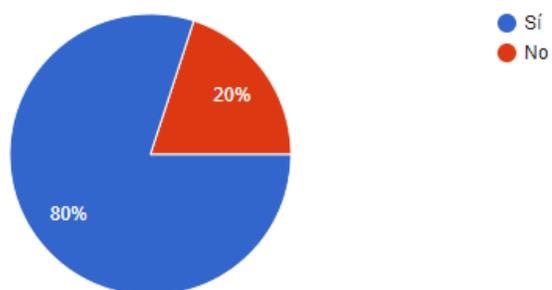


Figura 47.

Resumen de tercera pregunta de la encuesta de satisfacción

¿Pudiste saber las mejores combinaciones de 3 niveles?

5 respuestas

**Figura 48.**

Resumen de cuarta pregunta de la encuesta de satisfacción

Nivel en el que el sistema cumplió su objetivo

5 respuestas

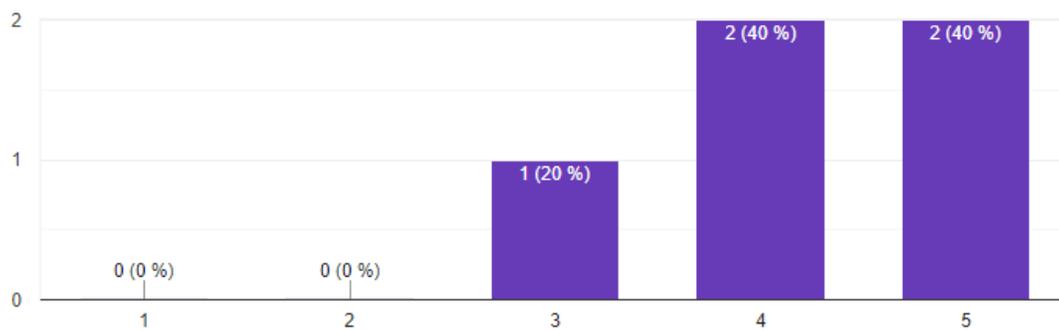
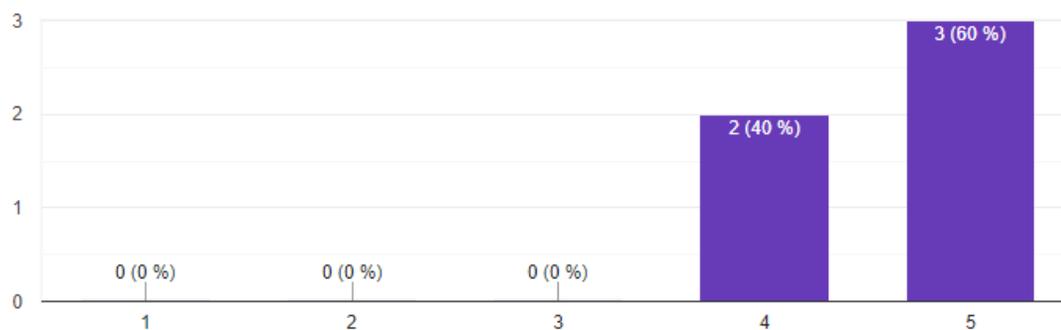


Figura 49.

Resumen de quinta pregunta de la encuesta de satisfacción

Nivel en el que usarás este sistema

5 respuestas

**Tabla 1.**

Resultados de Encuesta de Satisfacción

Ítem	Resultado en base de 1 a 10	Conclusión
Fácil uso del sistema	8	Cumplido
Entendimiento de reportes	9	Cumplido
Cumplimiento de objetivo	4	Cumplido
Frecuencia de uso	5	Cumplido