Bases de Datos Avanzadas

Manuel Flores Quiñones A01206898

David Sebastián Ramírez Navarro A01206423

José Ramón Romero Chávez A01700318

Alejandro Salmón Félix Díaz A01201854

1. **Detalle de actividades**

Para la segunda entrega del proyecto será necesario el hacer el diseño de nuestra data warehouse, para esto necesitamos diseñar los cubos de cada integrante que puedan satisfacer los requerimientos para nuestras consultas individuales.

Durante el proceso de desarrollo y planeación de los Data Marts debemos analizar con cuidado las preguntas que se plantearon de forma individual. Estas preguntas brindarán las bases para la información requerida tanto en las dimensiones y en los hechos, no solo definirá la información que tendrán estas tablas sino que también definirán la cantidad de tablas necesarias.

Al terminar el análisis, se pasa a la fase de construcción y modelado del modelo estrella para el proyecto, dependiendo de las preguntas individuales se modela con un diagrama de entidad relación las tabla de hechos y las dimensiones que se conectan con la misma.

Una vez que tengamos bien definidos nuestros DataMarts es momento de pasar a la parte práctica de la entrega. La implementación del proyecto en la base de datos que representara nuestra data warehouse.

**2. Responsabilidades y Roles**

El proyecto será dividido en diferentes fases:

1. Planeación y documentación de las actividades.
2. Diseño de los cubos de los integrantes del equipo.
3. Chequeo de los cubos y garantizar la información necesaria para las consultas.
4. Implementación de las tablas hechos y dimensiones de cada integrante
5. Diseño y codificación del modelo ETL.
6. Creación de vistas materializadas necesarias para obtener la información.
7. Recuperar la información de las vistas e introducirlas en las tablas individuales de hechos y dimensiones.
8. Diseñar las consultas que contesten las preguntas establecidas de forma individual.
9. Chequeo de resultados de consultas.

Estas actividades serán repartidas entre los cuatro integrantes del equipo de una forma equitativa y cada integrante tiene la responsabilidad de terminar con sus Cubos.

El equipo dividirá las tareas en roles y estos roles serán asignados entre los integrantes.

2.2 Roles

Cada integrante tiene la responsabilidad de completar las tareas asignadas en su rol en el proyecto. Cada integrante en la primer entrega tenía asignada una confederación (en el caso de dos integrantes tenían dos cada uno, uno en la instancia slp y otra en qro). En esta ocasión, basándonos en las necesidades que deben de ser cumplidas en las preguntas individuales será necesario establecer un rol a cada integrante del equipo.

Ramón será el encargado de hacer el cubo que cumpla las necesidades de cumplir cualquier consulta relacionada con la temporada.

David será el responsable de hacer el cubo que cumpla con las necesidades de cumplir cualquier consulta relacionada con los equipos en cualquier parte del mundo.

Salmón será el responsable de hacer el cubo que cumpla con las necesidades de cumplir cualquier consulta relacionada con las confederaciones de los equipos.

Manuel será el responsable de hacer el cubo que cumpla con las necesidades de cumplir cualquier consulta relacionada con las ligas en la que juegan los equipos.

A parte del diseño e implementación de los cubos, cada integrante tiene la tarea de hacer (e implementar) las consultas que contesten las preguntas establecidas de forma individual que estén relacionadas con su rol en la organización.

**3. Distribución del tiempo**

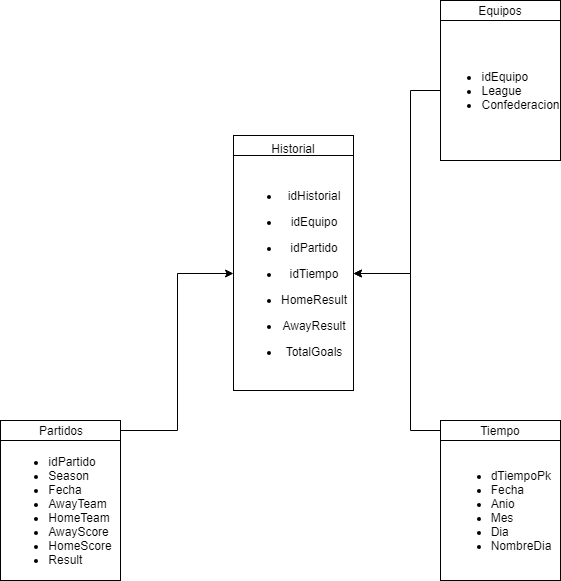
|  |  |  |
| --- | --- | --- |
| Fecha | Tiempo invertido | Tarea a realizar |
| 14/10/17 | 60 minutos | Planeacion y definicion de las preguntas (5 por integrante) |
| 16/10/17 | 30 minutos | Chequeo de preguntas individuales con el profesor |
| 16/10/17 | 15 minutos | Corrección de preguntas erróneas o mal planteadas |
| 18/10/17 | 40 minutos | Diseño de los cubos individuales con hechos y dimensiones. |
| 20/10/17 | 120 minutos | Documentación básica del proyecto, diseño a detalle de los cubos individuales con sus llaves primarias y atributos. |
| 23/10/17 | 30 minutos | Chequeo del diseño de los cubos con el profesor y realizar correcciones |
| 24/10/17 | 60 minutos | Realizar pruebas en la base de datos del equipo con vistas materializadas y simulacro de consultas para contestar las preguntas individuales |
| 25/10/17 | 180 minutos | Creación de cubos en la base de datos del equipo, creación de las vistas e implementación de las consultas |

**4. Especificaciones del proceso de negocio, indicadores y dimensiones**

David -> equipo

* Equipo con mejor diferencia de goles en la história
* El equipo con peor diferencia de goles de manera histórica
* Equipo con mejor diferencia de goles en la última en la historia
* El equipo con peor diferencia de goles en la historia
* Equipo con mejor promedio de goles en la historia

Modelo de Cubos:



Dimensiones:

* D\_Equipos (**PK\_TEAM**, TEAM, LEAGUE, CONFEDERACION)
* D\_Partidos (**pk\_partido**, CONFEDERACION, LEAGUE, SEASON, FECHA, HOMETEAM, AWAYTEAM, HOMESCORE, AWAYSCORE, GOLESPARTIDO)
* D\_Tiempo (**pk\_tiempo**, Año, Mes, Día)

Hechos:

* H\_Historial (**pk\_historial, pk\_partido, pk\_hometeam, pk\_awayteam,**  HomeResult, AwayResult, TotalGoals)

Manuel ->liga

* La liga con más goles
* La liga con menos goles
* La liga con más goles en casa
* La liga con menos goles en casa
* La liga con más goles en contra
* La liga con menos goles en contra

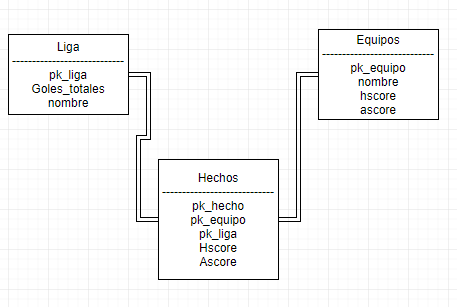
Modelo Cubo:

Dimensiones:

* Equipos (**pk\_equipo**, nombre, hscore, ascore)
* Liga(**pk\_liga**, goles\_totales, nombre)

Hechos:

* Hecho (**pk\_hecho, pk\_equipo, pk\_liga,** hscore, ascore)



Salmón: confederaciones

* Confederación con más goles en la última temporada
* Confederación con menos goles en la última temporada
* Confederación con más goles en el histórico
* Confederación con menos goles en el histórico
* Confederación con más cantidad de partidos por temporada

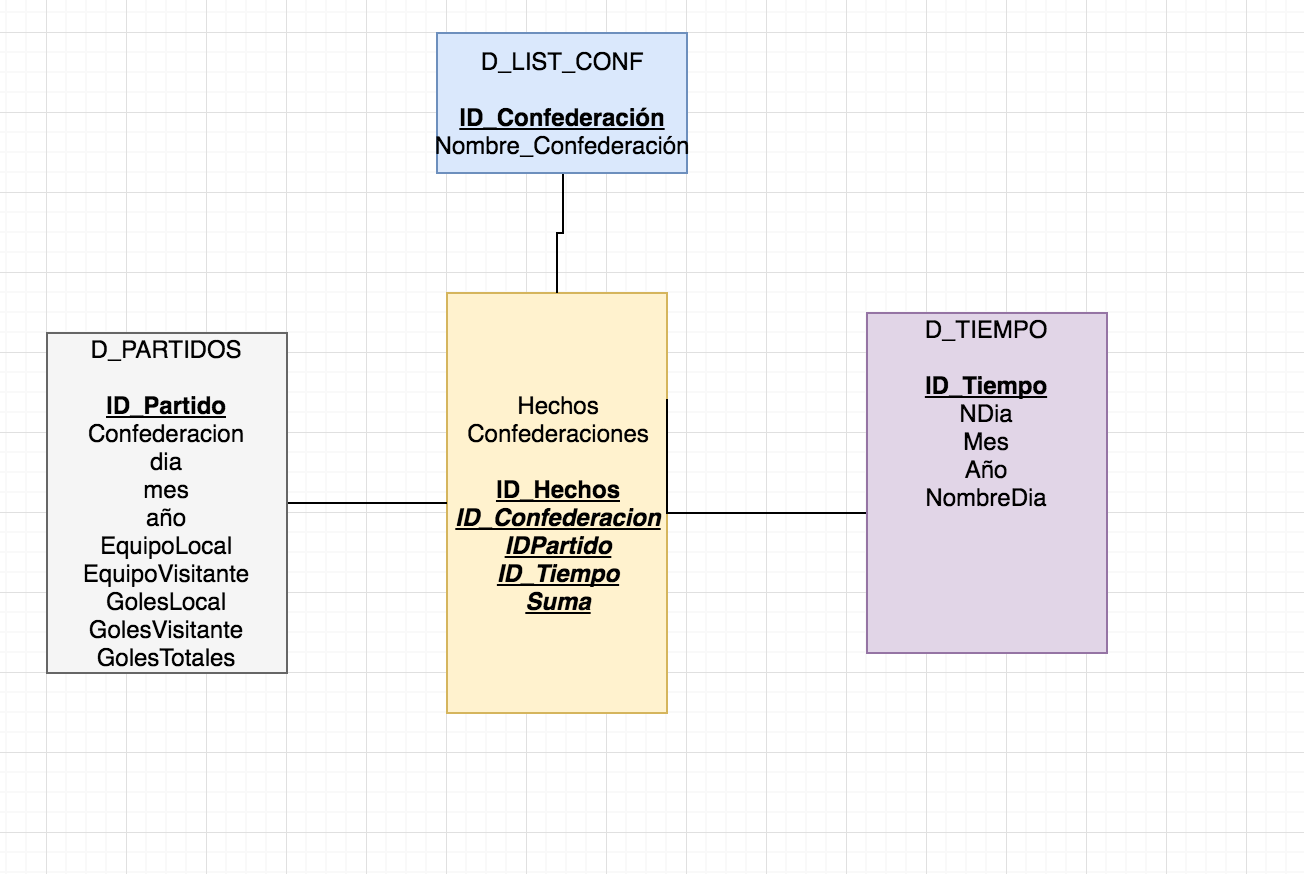
Modelos de Cubo:

Tiempo(**pk\_tiempo**, dia, mes, año);

Confederaciones(**pk\_confederaciones**, nombre)

Partidos (**pk\_partido**, local,visitante, goleslocal, golesvisitante, sumagoles, Confederacion)

Hechos(**pk\_hechos**,***pk\_partido, pk\_tiempo****,* ***pkConfederacion***, goles acumulados);

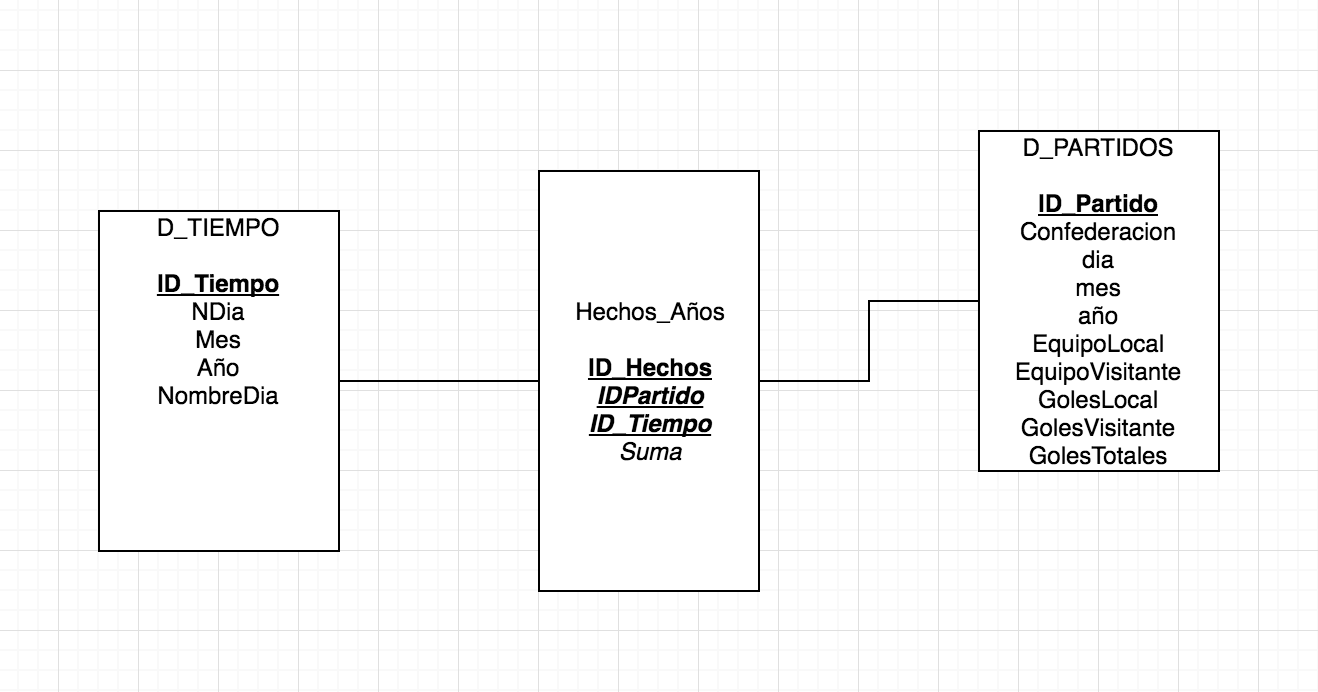


Ramón - Años

1. El año en la que más goles ha habido en total
2. El año con menos goles que ha habido en total
3. El año con menos partidos en total
4. El año con más partidos en total
5. El año con más empates

Modelo Cubo:

* D\_Partidos (**pk\_partido**, CONFEDERACION, LEAGUE, SEASON, FECHA, HOMETEAM, AWAYTEAM, HOMESCORE, AWAYSCORE, GOLESPARTIDO)
* D\_Tiempo (**pk\_tiempo**, Año, Mes, Día, NombreDia)
* H\_Historial (**pk\_historial, pk\_partido, pk\_partido,** suma)



**Describe procesos de la organización seleccionada, identificando los indicadores que determinan el desempeño del proceso y las dimensiones por las que deben analizarse para tomar decisiones que impacten en la mejora continua del proceso**

Reporte del aprendizaje obtenido en la fase.

David: Aprendí toda la teoría de lo que es un data warehouse, me di cuenta de cuáles son sus utilidades y por que son necesarias a la hora de manejar bases de datos muy grandes con mucha historia, me interesó mucho que con el historial de datos puedes sacar tendencias y conclusiones según el tiempo y el historial de datos.

Sólo diría que hace falta una clase de laboratorio el cual se dedique a la generación de scripts y consultas avanzadas ya que no tengo mucha práctica con eso.

Manuel: En esta entrega aprendí que es muy práctico y de mucha utilidad utilizar una data warehouse por que te brinda la posibilidad de analizar cualquier tipo de dato en alguna empresa. Esto es algo muy importante y muy solicitado en la industria ya que es una forma de que las empresas o socios afiliados puedan analizar el desempeño de una empresa y tomar acciones/decisiones en base al análisis.

Ramón: Se me hizo interesante la idea conceptual que existe en el paso de las simples bases de datos transaccionales al modelaje y creación de los modelos de cubos. Además de que aprendí que este es el proceso ideal que finaliza en la minería de datos. Y que en este proceso no se hace de un dia para otro, es necesario tener idea de lo que en realidad se pide y es necesario hacer, tener una idea conceptual lamentablemente no siempre es suficiente haciendo durante la práctica, ya que si no se practica lo suficiente la curva de aprendizaje de una nueva tecnología puede llegar a ser muy tardada.

Salmón: Después de trabajar con SQL, vi que las herramientas como Vistas Materializadas, Procedimientos, Triggers, entre otras cosas. Es fundamental en cualquier sistema que desarrolle estar consciente de que una base de datos eficiente y bien diseñada facilita de una manera impresionante el manejo de toda la información, transacciones e indicadores que puede existir con el usuario a través de alguna aplicación.

En General el equipo tuvo buen interacción y comunicación, sin embargo, todavía existe áreas de oportunidad que buscaremos desarrollar haciendo uso de una mejor organización en tareas y tiempos con respecto a cada una de las fases del proyecto