Proyecto de Bases de datos

Cindy Tatiana Pinzón, Andrés Felipe Gutierrez

Facultad de Ingenieria y Ciencias Básicas
Universidad Central
Maestría en Analítica de Datos
Curso de Bases de Datos
Bogotá, Colombia
cpinzont2@ucentral.edu.co, agutierrezf2@ucentral.edu.co

October 8, 2022

Contents

L	Inti	roducción (Max 250 Palabras) - (<i>Primera entrega</i>)	3
2		racterísticas del proyecto de investigación (Max 500 Palabras) Primera entrega) Titulo del proyecto de investigación (Max 100 Palabras) - (Primera	3
	2.2	entrega)	3 3
	2.3	Alcance (Max 200 Palabras) - (Primera entrega)	4
	2.4	Pregunta de investigación (Max 100 Palabras) - (Primera entrega) .	4
	2.5	Hipotesis (Max 100 Palabras) - (<i>Primera entrega</i>)	4
3	Ref	dexiones sobre el origen de datos e información (Max 400	
3		dexiones sobre el origen de datos e información (Max 400 abras) - (<i>Primera entrega</i>)	5
3		$abras) - (Primera\ entrega)$	5
3	Pala	abras) - (Primera entrega)	5
3	Pala	abras) - (<i>Primera entrega</i>) ¿Cual es el origen de los datos e información ? (Max 100 Palabras)	
3	Pala 3.1	abras) - (Primera entrega) ¿Cual es el origen de los datos e información ? (Max 100 Palabras) - (Primera entrega)	
3	Pala 3.1	Abras) - (Primera entrega) ¿Cual es el origen de los datos e información ? (Max 100 Palabras) - (Primera entrega)	5
3	Pala 3.1 3.2	abras) - (Primera entrega) ¿Cual es el origen de los datos e información? (Max 100 Palabras) - (Primera entrega)	5
3	Pala 3.1 3.2	Abras) - (Primera entrega) ¿Cual es el origen de los datos e información ? (Max 100 Palabras) - (Primera entrega)	5
3	Pala 3.1 3.2	Abras) - (Primera entrega) ¿Cual es el origen de los datos e información ? (Max 100 Palabras) - (Primera entrega) ¿Cuales son las consideraciones legales o eticas del uso de la información? (Max 100 Palabras) - (Primera entrega) ¿Cuales son los retos de la información y los datos que utilizara en la base de datos en terminos de la calidad y la consolidación?	E.5

4	Diseño del Modelo de Datos del SMBD (Sistema Manejador de	
	Bases de Datos)(Primera entrega)	6
	4.1 Características del SMBD (Sistema Manejador de Bases de Datos)	
	para el proyecto (<i>Primera entrega</i>)	6
	4.2 Diagrama modelo de datos (<i>Primera entrega</i>)	6
	4.3 Imágenes de la Base de Datos (<i>Primera entrega</i>)	7
	4.4 Código SQL - lenguaje de definición de datos (DDL) (Primera	
	entrega)	7
	4.5 Código SQL - Manipulación de datos (DML) (Primera entrega)	9
	4.6 Código SQL + Resultados: Vistas (<i>Primera entrega</i>)	11
	4.7 Código SQL + Resultados: Triggers (<i>Primera entrega</i>)	12
	4.8 Código SQL + Resultados: Funciones (<i>Primera entrega</i>)	12
	4.9 Código SQL + Resultados: procedimientos almacenados (<i>Primera</i>	
	entrega)	12
5	Bases de Datos No-SQL (Segunda entrega)	14
	5.1 Diagrama Bases de Datos No-SQL (Segunda entrega)	14
	5.2 SMBD utilizado para la Base de Datos No-SQL (Segunda entrega)	14
6	Aplicación de ETL (Extract, Transform, Load) y Bodega de	
	Datos (Tercera entrega)	15
	6.1 Ejemplo de aplicación de ETL y Bodega de Datos (Tercera entrega)	15
7	Lecciones aprendidas (Tercera entrega)	16
8	Bibliografía	17

1 Introducción (Max 250 Palabras) - (Primera entrega)

La gestión de información empresarial se vuelve cada vez más importante, sin embargo, la sobrecarga de datos está generando estrés y confusión en las organizaciones, como lo demuestra el estudio de Open Text, el cual deja visto que el 79% de los españoles experimenta efectos de la "sobrecarga de información". El entendimiento, análisis y fácil acceso a la información parte de un trabajo arduo, que involucra conocimientos técnicos, experiencia tanto en datos como en operación de negocio y solo es posible si se cuenta con el tiempo necesario para la medición y evaluación de cada indicador. Al investigar y evaluar diferentes soluciones para integración, análisis y visualización de data, se encuentran diferentes alternativas que van desde las negociaciones con proveedores expertos en materia que solo entreguen este tipo de servicios bajo contratos de un tiempo estipulado y con curva de aprendiza je en implementación, hasta el contrato directo del capital humano dedicado 100% a este proceso, lo que conlleva a la responsabilidad de tener una línea especializada para lograr enfocar el proceso a la necesidad de la operación; sin embargo no se encuentran servicios de información capaces de desarrollar análisis y visualizaciones de forma rápida y eficaz, con la confidencialidad de la data necesaria y que puede ser, esporádica, ocasional o de solo una vez. Es por esto por lo que InfoExpress se crea con la finalidad de cubrir estas necesidades.

2 Características del proyecto de investigación (Max 500 Palabras) - (*Primera entrega*)

2.1 Titulo del proyecto de investigación (Max 100 Palabras) - (Primera entrega)

InfoExpress, Servicio de análisis, unificación y visualización de información

2.2 Objetivo general (Max 100 Palabras) - (Primera entrega)

Construir las diferentes herramientas tecnológicas necesarias para transaccionar los datos de un servicio de análisis, unificación y visualización de datos, con el fin de apoyar a las empresas a sus análisis de resultados y entendimiento de la evolución del negocio de forma esporádica mediante técnicas de transformación y correlación de información.

2.2.1 Objetivos especificos (Max 100 Palabras) - (Primera entrega)

Entender la necesidad a nivel de experiencia de usuario y necesidad operacional
del servicio en cuanto a criterios necesarios para generar solicitudes de
informes, con el fin de realizar un esquema de solicitud amigable para el
usuario y con la información necesaria para dar respuesta a la solicitud,
mediante la experticia y comparación con el mercado del ofrecimiento del
servicio.

- Diseñar y programar la estructura de datos y el enlace de este a la pagina web y/o apps del servicio para la recolección y visualización de solicitudes, mediante diagrama de base datos.
- Diseñar y programar la estructura de las vistas del estado de las solicitudes y el historial de ellas para la recepción, entendimiento y seguimiento tanto del analista a cargo de la solicitud como del usuario que sube la solicitud, mediante un esquema amigable a ambos tipos de usuario.

2.3 Alcance (Max 200 Palabras) - (Primera entrega)

Se desea lograr un alcance aplicativo al desarrollar este proyecto de investigación, en cual se radicará la aplicación de los métodos y técnicas de diseño y programación de bases de datos, a través de los contenidos impartidos en la asignatura "Bases de datos", perteneciente a la maestría de Analítica de datos de la Universidad Central.

2.4 Pregunta de investigación (Max 100 Palabras) - (Primera entrega)

¿Cuál es la forma mas optima de transaccionar los datos de un servicio de análisis, unificación y visualización de información de manera online?

2.5 Hipotesis (Max 100 Palabras) - (Primera entrega)

La forma mas optima de transaccionar los datos de un servicio de análisis, unificación y visualización de información de manera online es a través del desarrollo de una base de datos relacional con vistas en donde los diferentes tipos de usuarios logren visualizar toda la información de cada solicitud.

3 Reflexiones sobre el origen de datos e información (Max 400 Palabras) - (*Primera entrega*)

3.1 ¿Cual es el origen de los datos e información ? (Max 100 Palabras) - (*Primera entrega*)

La base de datos que se plantea crear almacenará datos de origen de transacción tomados de aplicativo web y/o apps que el servicio requiera.

3.2 ¿Cuales son las consideraciones legales o eticas del uso de la información? (Max 100 Palabras) - (*Primera entrega*)

La base de datos pensada para Info Express, requiere no solo un diseño amigable para el usuario sino además la garantía de la privacidad y total resguardo de la información consignada en cada solicitud, puesto que cada requerimiento puede contener información sensible de las empresas usuarias del servicio tales como ventas, información de empleados, información de desempeño de planes estratégicos, etc. Info Express es un servicio que busca apalancar aquellos análisis esporádicos que hoy algunas empresas no logran en su día a día, por ello debe claridad y veracidad en las cifras entregadas, comprometiéndose a no tener alteraciones en la información recibida.

3.3 ¿Cuales son los retos de la información y los datos que utilizara en la base de datos en terminos de la calidad y la consolidación? (Max 100 Palabras) - (*Primera entrega*)

Dentro de infoExpress uno de los retos más grandes en términos de consolidación de datos es los diferentes formatos en los cuales puede compartir la información el usuario del servicio, ya que puede venir desde varias sabanas de datos en Excel hasta bancos de datos en otras aplicaciones. Otro reto enorme en términos de calidad es la capacidad de almacenamiento de cada una de las transacciones de solicitudes, dejando un registro de cada paso llevado y sin reemplazo de información en pasos anteriores.

3.4 ¿Que espera de la utilización de un sistema de Bases de Datos para su proyecto? (Max 100 Palabras) - (Primera entrega)

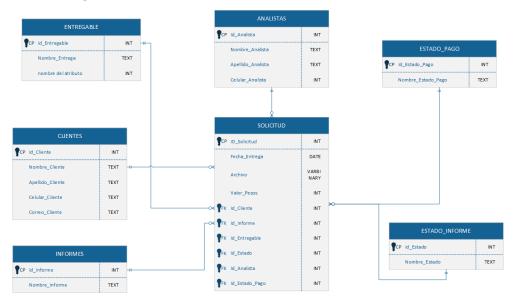
Se espera que con la utilización de un sistema de base de datos InfoExpress logre unificar todas las solicitudes y respuestas dadas del servicio, así mismo entregue la información necesaria para detectar y prever las cargas a nivel numérico y de tipo de solicitud de la herramienta. Por otra parte, con la finalidad de apalancar el entendimiento y la correcta ejecución de la operación detrás del sistema, se espera lograr el entendimiento de los tiempos de respuesta y flujos de estas.

4 Diseño del Modelo de Datos del SMBD (Sistema Manejador de Bases de Datos)(Primera entrega)

4.1 Características del SMBD (Sistema Manejador de Bases de Datos) para el proyecto (*Primera entrega*)

Para el proyecto se eligió Microsoft SQL Server dado que la obtencion de la información se encuentra en servidores de Windows, que facilita la intergración con este sistema operativo, por otra parte otras herramientas permite manipular la información desde SQL Server asi como la carga de información en archivos con extension xlsx.

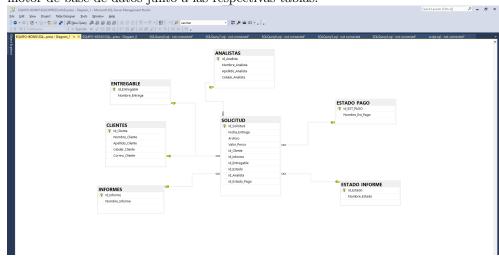
4.2 Diagrama modelo de datos (Primera entrega)



4.3 Imágenes de la Base de Datos (Primera entrega)

□ InfoExpress
□ Database Diagrams
□ Tables
□ System Tables
□ FileTables
□ External Tables
□ Graph Tables
□ Image:
□ Ima

En la imagen anterior se puede apreciar el árbol de carpetas que maneja el motor de base de datos junto a las respectivas tablas.



Al crear la base de datos el administrador SQL Management Studio permite visualizar un modelo de base de datos de todo el esquema creado como se vio en el modelo relacional.

4.4 Código SQL - lenguaje de definición de datos (DDL) $(Primera\ entrega)$

USE [InfoExpress]
GO

```
--Creación de las tabla [ANALISTAS]
CREATE TABLE [dbo].[ANALISTAS](
[Id_Analista] int IDENTITY(1,1) NOT NULL PRIMARY KEY,
[Nombre_Analista] text NULL,
[Apellido_Analista] text NULL,
[Celular_Analista] text NULL,
);
--Creación de las tabla [CLIENTES]
CREATE TABLE [dbo].[CLIENTES](
[Id_Cliente] [int] IDENTITY(1,1) NOT NULL,
[Nombre_Cliente] text NULL,
[Apellido_Cliente] text NULL,
[Celular_Cliente] text NULL,
[Correo_Cliente] text NULL,
);
--Creación de las tabla [ENTREGABLE]
CREATE TABLE [dbo]. [ENTREGABLE] (
[Id_Entregable] int IDENTITY(1,1) NOT NULL,
[Nombre_Entrega] text NULL
);
--Creación de las tabla [ESTADO_INFORME]
CREATE TABLE [dbo].[ESTADO_INFORME](
[Id_Estado] [int] IDENTITY(1,1) NOT NULL,
[Nombre_Estado] [text] NULL
);
--Creación de las tabla [ESTADO_PAGO]
CREATE TABLE [dbo].[ESTADO_PAGO](
[Id_EST_PAGO] [int] IDENTITY(1,1) NOT NULL,
[Nombre_Est_Pago] [text] NULL
);
--Creación de las tabla [INFORMES]
CREATE TABLE [dbo].[INFORMES](
[Id_Informe] [int] IDENTITY(1,1) NOT NULL,
[Nombre_Informe] [text] NULL
);
--Creación de las tabla [SOLICITUD]
CREATE TABLE [dbo].[SOLICITUD](
[Id_Solicitud] [int] IDENTITY(1,1) NOT NULL,
[Fecha_Entrega] [date] NULL,
```

```
[Archivo] [varchar] (500) NULL,
[Valor_Pesos] [int] NULL,
[Id_Cliente] [int] NOT NULL,
[Id_Informe] [int] NOT NULL,
[Id_Entregable] [int] NOT NULL,
[Id_Estado] [int] NOT NULL,
[Id_Analista] [int] NOT NULL,
[id_Estado_Pago] [int] NOT NULL,
--Creación de llaves foraneas mediante constraints
CONSTRAINT [FK_SOLICITUD_ANALISTAS] KEY
([Id_Analista]) REFERENCES [ANALISTAS] ([Id_Analista]),
CONSTRAINT [FK_SOLICITUD_ENTREGABLE] KEY
([Id_Entregable]) REFERENCES [ENTREGABLES] ([Id_Entregable]),
CONSTRAINT [FK_SOLICITUD_CLIENTES] KEY
([Id_Cliente]) REFERENCES [CLIENTES] ([Id_Cliente]),
CONSTRAINT [FK_SOLICITUD_ESTADO_INFORME] KEY
([Id_Estado]) REFERENCES [ESTADO_INFORME] ([Id_Estado]),
CONSTRAINT [FK_SOLICITUD_ESTADO_PAGO] KEY
([id_Estado_Pago]) REFERENCES [ESTADO_PAGO] ([id_Estado_Pago]),
CONSTRAINT [FK_SOLICITUD_INFORMES] KEY
([Id_Informe]) REFERENCES [INFORMES] ([Id_Informe])
);
DROP TABLE [Archivo_Clientes]
DROP TABLE [Archivo_Analistas]
     Código SQL - Manipulación de datos (DML) (Primera
4.5
      entrega)
----Inserción de datos apartir de archivos de .xlsx
INSERT INTO (CLIENTES)
 [Nombre_Cliente]
,[Apellido_Cliente]
,[Celular_Cliente]
,[Correo_Cliente]
SELECT
       [Nombre_Cliente]
```

```
,[Apellido_Cliente]
      ,[Celular_Cliente]
      ,[Correo_Cliente]
 FROM [InfoExpress].[dbo].[Archivo_Clientes];
  INSERT INTO (ANALISTAS)
      , [Nombre_Analista]
      ,[Apellido_Analista]
      ,[Celular_Analista]
SELECT
      , [Nombre_Analista]
      ,[Apellido_Analista]
      ,[Celular_Analista]
 FROM [InfoExpress].[dbo].[Archivo_Clientes];
  --Insercion de datos establecidos
INSERT INTO [ENTREGABLE]
([Id_Entregable], [Nombre_Entrega])
VALUES (1, 'EXCEL'),
       (2, 'POWER BI'),
   (3, 'PDF'),
   (4, 'POWER POINT')
 INSERT INTO [dbo].[ESTADO_INFORME]
 ([Id_Estado], [Nombre_Estado])
 VALUES (1, 'EN ESPERA PARA ASIGNAR ANALISTA'),
        (2, 'EN PROCESO'),
(3, 'EN PAUSA'),
(4, 'TERMINADO')
INSERT INTO [INFORMES]
([Id_Informe], [Nombre_Informe])
VALUES (1, 'Distribucion'),
       (2, 'Negociaciones'),
   (3, 'Inventarios'),
   (4, 'Participación'),
   (5, 'Ventas'),
   (6, 'Vencimientos'),
   (7, 'Cobertura')
   (8, 'Chequeo Precios'),
   (9, 'Actividad Comercial');
```

4.6 Código SQL + Resultados: Vistas (Primera entrega)

CREATE VIEW [dbo].[VW_SOLICITUDES_Resumen] AS

```
SELECT [Id_Solicitud]
      , [Fecha_Entrega]
      ,[Archivo]
      , [Valor_Pesos]
      ,CONCAT([Nombre_Cliente],' ',[Apellido_Cliente]) AS [Cliente]
      , [Nombre_Informe]
  , [Nombre_Entrega]
      , [Nombre_Estado]
      ,CONCAT([Nombre_Analista], ' ',[Apellido_Analista]) AS [Analista]
      , [Nombre_Est_Pago]
 FROM [SOLICITUD]
 INNER JOIN [ANALISTAS] ON [SOLICITUD].[Id_Analista] = [ANALISTAS].[Id_Analista]
 INNER JOIN [ENTREGABLE] ON [SOLICITUD].[Id_Entregable] = [ENTREGABLE].[Id_Entregable]
 INNER JOIN [ESTADO_INFORME] ON [ESTADO_INFORME].[Id_Estado] = [SOLICITUD].[Id_Estado]
 INNER JOIN [CLIENTES] ON [CLIENTES].[Id_Cliente] = [SOLICITUD].[ID_Cliente]
 INNER JOIN [INFORMES] ON [INFORMES].[Id_Informe] = [SOLICITUD].[Id_Informe]
 INNER JOIN [ESTADO_PAGO] ON [ESTADO_PAGO].[Id_EST_PAGO] = [SOLICITUD].[id_Estado_Pago]
```

	ld_Solicitud	Fecha_Entrega	Archivo	Valor_Pesos	Cliente	Nombre_Informe	Nombre_Entrega	Nombre_Estado	Analista	Nombre_Est_Pago
1	3	2020-10-07	Nisi.ppt	550931973	Melli Daunay	Ventas	POWER POINT	TERMINADO	Beverlie Tumioto	CON ABONO
2	4	2020-06-06	NULL	822878	Hedwig Baffin	Cobertura	EXCEL	EN PROCESO	Sara-ann Rayworth	CON ABONO
3	5	2021-07-15	Fringilla Phoncus mpeg	8163015	Nicolas Linneman	Cobertura	EXCEL	TERMINADO	Jacquenette Stronack	SIN ABONO
4	6	2020-08-22	Praesent Lectus Vestibulum į peg	1236470	Erv Fouracre	Vencimientos	POWER BI	EN PAUSA	Wynne Brigham	SIN ABONO
5	7	2021-12-11	Gravida.mpeg	7321348	Carrissa Askwith	Actividad Comercial	POWER POINT	TERMINADO	Randy Mallon	PAGO TOTAL
6	8	2020-09-21	QuisqueUt xls	7138601	Keir Dollar	Ventas	POWER POINT	TERMINADO	Kellby Dowker	PAGO TOTAL
7	9	2021-04-05	Viverra Dapibus, ppt	783262	Faustine Hindenburg	Ventas	PDF	EN ESPERA PARA ASIGNAR ANALISTA	Lek Petchell	PAGO TOTAL
8	10	2020-10-01	Primisln.avi	1870853	Terrye Gleasane	Negociaciones	PDF	TERMINADO	Diana Whisby	PAGO TOTAL
9	11	2020-06-26	Ante.mp3	8506307	Leonelle Kohrsen	Actividad Comercial	PDF	EN PAUSA	Josiah Grevel	SIN ABONO
10	12	2021-12-18	Ut.tiff	9861205	Fayth Loades	Distribucion	POWER BI	EN PAUSA	Laurette Stolberg	SIN ABONO
11	13	2022-05-12	NULL	6254826	Brandi Tremeer	Negociaciones	POWER POINT	EN PAUSA	Claus Kinchington	SIN ABONO
12	14	2022-08-16	Eu.avi	1465564	Bondy Garmey	Distribucion	EXCEL	EN PAUSA	Josiah McMeanma	SIN ABONO
13	15	2021-07-11	NULL	3704788	Luciana Pavelka	Negociaciones	EXCEL	EN ESPERA PARA ASIGNAR ANALISTA	Gib Threadgold	PAGO TOTAL
14	16	2021-02-26	NULL	8350092	Luise Peatheyjohns	Actividad Comercial	EXCEL	EN ESPERA PARA ASIGNAR ANALISTA	Reynold Chesterton	PAGO TOTAL
15	17	2022-04-13	Donec Diam.txt	5309322	Hansiain Shedd	Actividad Comercial	POWER POINT	EN PAUSA	Pamell Jarrett	SIN ABONO
16	18	2021-07-05	NULL	1159760	Nollie Ayrton	Cobertura	EXCEL	TERMINADO	Casar Tailby	SIN ABONO
17	19	2021-03-31	NULL	4674108	Shirlene Madden	Cobertura	POWER POINT	EN PROCESO	Haleigh Setchell	CON ABONO
18	20	2021-02-21	NULL	4779289	Cynthia Bolsover	Cobertura	EXCEL	EN PROCESO	Trina Muckian	PAGO TOTAL
19	21	2020-09-25	Et Ultrices xls	4963397	Phebe Sansam	Actividad Comercial	POWER POINT	EN PROCESO	George Doucette	PAGO TOTAL
20	22	2020-05-30	NULL	540857	Nap De Pero	Negociaciones	PDF	TERMINADO	Jacquenette Stronack	CON ABONO

CREATE VIEW VW_SOLICITUDES_Resumen_Pagos AS

```
SELECT [Id_Solicitud]
    ,[Valor_Pesos]
    ,[Nombre_Est_Pago]
FROM [SOLICITUD]
INNER JOIN [ESTADO_PAGO] ON [ESTADO_PAGO].[Id_EST_PAGO] = [SOLICITUD].[id_Estado_Pago]
```

1 3 550931973 CON ABONO 2 4 822878 CON ABONO 3 5 8163015 SIN ABONO 4 6 1236470 SIN ABONO 5 7 7321348 PAGO TOTAL 6 8 7138601 PAGO TOTAL 7 9 783262 PAGO TOTAL	
3 5 8163015 SIN ABONO 4 6 1236470 SIN ABONO 5 7 7321348 PAGO TOTAL 6 8 7138601 PAGO TOTAL	
4 6 1236470 SIN ABONO 5 7 7321348 PAGO TOTAL 6 8 7138601 PAGO TOTAL	
5 7 7321348 PAGO TOTAL 6 8 7138601 PAGO TOTAL	
6 8 7138601 PAGO TOTAL	
7 9 783262 PAGO TOTAL	
8 10 1870853 PAGO TOTAL	
9 11 8506307 SIN ABONO	
10 12 9861205 SIN ABONO	
11 13 6254826 SIN ABONO	
12 14 1465564 SIN ABONO	
13 15 3704788 PAGO TOTAL	
14 16 8350092 PAGO TOTAL	
15 17 5309322 SIN ABONO	
16 18 1159760 SIN ABONO	
17 19 4674108 CON ABONO	
18 20 4779289 PAGO TOTAL	
19 21 4963397 PAGO TOTAL	
20 22 540857 CON ABONO	

- 4.7 Código SQL + Resultados: Triggers (Primera entrega)
- 4.8 Código SQL + Resultados: Funciones (Primera entrega)
- 4.9 Código SQL + Resultados: procedimientos almacenados ($Primera\ entrega$)

CREATE PROCEDURE sp_CONSULTA_FECHA @FechaIni DATE,

@FechaFin DATE AS SELECT [Id_Solicitud] ,[Fecha_Entrega] ,[Archivo] ,[Valor_Pesos] ,[Cliente] ,[Nombre_Informe] ,[Nombre_Entrega] ,[Nombre_Estado] ,[Analista] ,[Nombre_Est_Pago] FROM [VW_SOLICITUDES_Resumen] WHERE [Fecha_Entrega] BETWEEN @FechaIni AND @FechaFin ---- test EXEC sp_CONSULTA_FECHA '2020-10-07', '2021-10-08';

- 5 Bases de Datos No-SQL (Segunda entrega)
- 5.1 Diagrama Bases de Datos No-SQL (Segunda entrega)
- 5.2 SMBD utilizado para la Base de Datos No-SQL (Segunda $_{entrega)}$

- 6 Aplicación de ETL (Extract, Transform, Load) y Bodega de Datos (Tercera entrega)
- 6.1 Ejemplo de aplicación de ETL y Bodega de Datos (Tercera entrega)

7 Lecciones aprendidas (Tercera entrega)

8 Bibliografía