

Identificación de consumos de agua atípicos en el municipio de Fusagasugá Cundinamarca.

Presentado por: César Chavarriaga B.

Bases de Datos

Maestría en Analítica de Datos

Universidad Central - Facultad de Ingeniería y Ciencias Básicas

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

El consumo de agua como servicio público es una de las necesidades básicas de cualquier comunidad. Actualmente las tarifas de medición del consumo se componen de un cargo fijo, un básico y un complementario; la unidad de medición de este consumo es el metro cúbico (m³). El municipio de Fusagasugá, una comunidad aledaña de Bogotá, con unas condiciones climatológicas más cálidas y con precipitaciones menos drásticas se destaca como un lugar óptimo para vacacionar; generando en una población flotante que debe ser tenida en cuenta especialmente por sus altos consumos en los lapsos de descanso.

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

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

IDENTIFICACIÓN DE CONSUMOS DE AGUA ATÍPICOS EN EL MUNICIPIO DE FUSAGASUGÁ CUNDINAMARCA.

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

Crear un modelo entidad relación, que permita determinar si el consumo de agua de un cliente de Fusagasugá Cundinamarca se encuentra dentro de las condiciones normales de medición.

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

- Realizar un análisis exploratorio de los datos, mediante el uso de metodologías de estadística descriptiva, con el fin de conocer patrones y tendencias de consumo de agua por parte de los usuarios.
- Identificar y determinar mediante un análisis de correlación de variables, el grado de relación de las variantes implicadas en el consumo de agua de los clientes.
- Definir el reporte resultado para el análisis de los datos atípicos presentados diariamente en el proceso de facturación.
- Determinar la estructura de las vistas donde se alojará la información que se extraerá del reporte.
- Crear procedimientos de almacenado, orientados a la captura y alistamiento de la información de valores atípicos para reportes diarios.

- Implementar trigger para la ejecución de los procedimientos de almacenado y limpieza de tablas temporales.

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

A través del análisis de 11 meses de facturación, se espera identificar el perfil de consumo atípicos y su impacto en las pérdidas de agua, sin embargo, es necesarios contar con los periodos vacacionales debido a que el consumo, aunque alto puede ser real.

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

¿En qué medida los consumos atípicos influyen en el porcentaje de agua no contabilizada en el Municipio de Fusagasugá?

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

La mayor cantidad de perdidas en agua se encuentra en el consumo que se realiza en el predio, y se asocia por el consumo que supera el promedio del mismo uso, estrato, ubicación y barrio.

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

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

Pertenecen a la información historia de la entidad, actualmente esta información no se le realiza ningún tipo de tratamiento de la información aparte de la consulta de la misma.

3.2 . ¿Cuáles son las consideraciones legales o éticas del uso de la información? (Max 100 Palabras) - (Primera entrega)

Con el fin de proteger los datos personales de los suscriptores y con el fin de que los datos no puedan ser atribuidos a un individuo, se realiza un tratamiento de los datos y se reemplazaron los identificadores como nombres, numero de documento etc.

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

Dentro del análisis realizado a los datos, se identifica, que hay varios campos que necesitan ser revisados, tal como la Ubicación en el que no es posible identificar si es rural o urbano y el Campo Barrio, en el cual presenta se presentan vacíos, se considera que el CRM que maneja la entidad debe garantizar la que no existan vacíos en datos relevantes. Se evidencia una cantidad que tiende al aumento lo cual es coherente con el comportamiento de ingreso de información nueva, esto a su vez, se ve reflejado en los metros cúbicos consumidos el cual también va al alza.

4 Diseño del Modelo de Datos del SMBD (Sistema Manejador de Bases de Datos) (Primera entrega) 7

4.1 Características del SMBD (Sistema Manejador de Bases de Datos)

La Tabla de Facturación, se considera la principal, ya que contiene los datos relevantes para el análisis. Cuenta con: 549079 registros, con información del año 2021 desde el mes de enero hasta noviembre.

TABLA FACTURACIÓN		
VARIABLES	DESCRIPCIÓN	TIPO
ANO	Año de la toma de lectura	Categórica
PERIODO	Periodo (mes) de toma de lectura	Categórica
CICLO	Ciclo, agrupación por zona geografica de inmuebles	Categórica
CODIGOINTERNO	Consecutivo que identifica el inmueble que tiene el servicio	Numérica
TIPO_DOC	Tipo de documento con el que se identifica el suscriptor	Numérica
NUM_DOCUMENTO	Número de documento de identificación del suscriptor	Carácter
USO	identificador del servicio Residencial: 1,2,3,4,5,6; Industrial: 10, Comercial: 11, Oficial: 12	Numérica
ESTRATO	Clasificación del uso residencial	Numérica
UBICACIÓN	Ubicación rural o urbana del predio	Categórica
BARRIO	Barrio de ubicación del predio	Categórica
SECTHIDRAULICO	Codigo de ubicación de la fuente de abastecimiento del servicio	Numérico
FECHAEXPFACTURA	Fecha de expedición de la factura	Fecha
LECTURA ACTUAL	Lectura del Periodo actual (n)(últimos 30 días de facturación)	Numérico
CONSUMO DEL PERIODO EN METROS CÚBICOS (M3)	Metros cúbicos consumidos	Numérico

La Tabla tipo de documento, contiene el descriptivo de los tipos de documentos

TABLA TIPO DOCUMENTO		
VARIABLES	DESCRIPCIÓN	TIPO
TIPO_DOC	Tipo de documento con el que se identifica el suscriptor	Numérico
TIPO_DOC_DESCRIP	Descripción de tipo de documento	Categórica

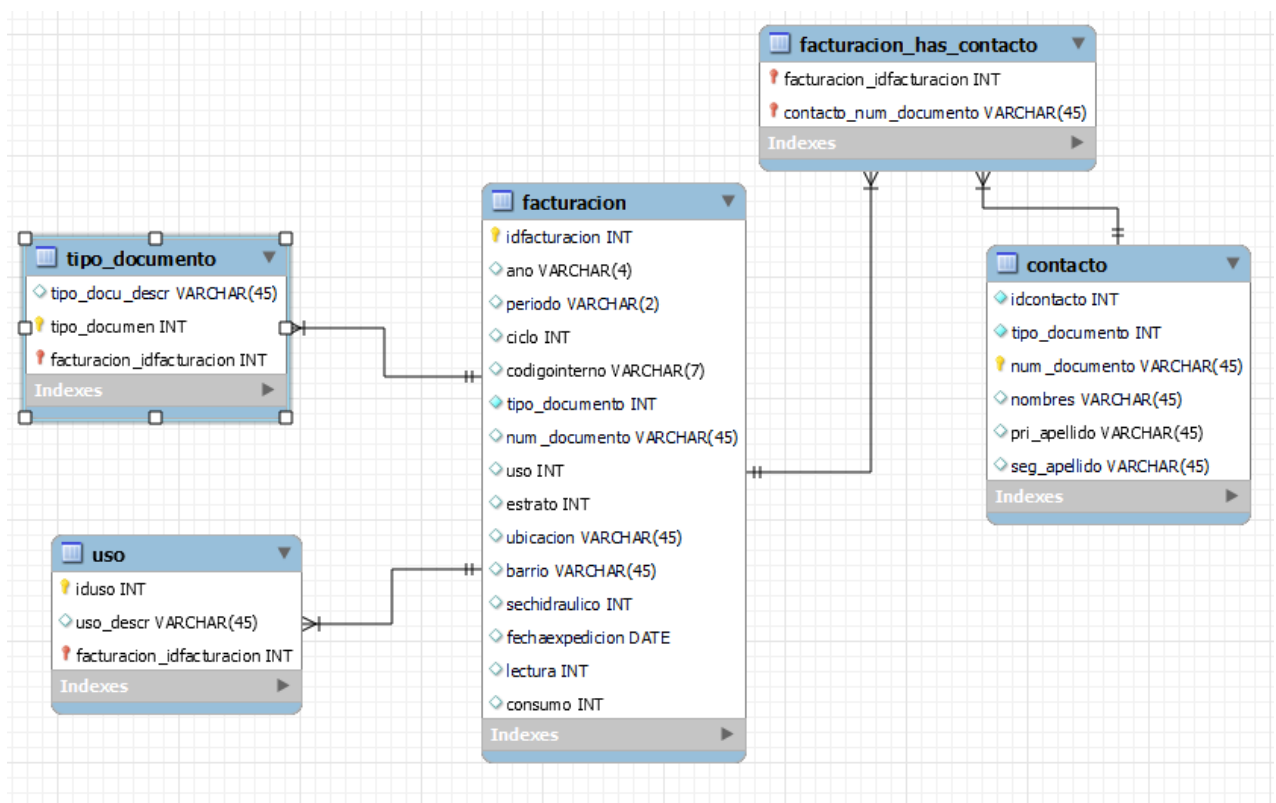
La Tabla Contactos, tiene el detalle de los suscriptores de los inmuebles.

TABLA CONTACTO		
VARIABLES	DESCRIPCIÓN	TIPO
ID_CONTACTO	Consecutivo de contacto	Numérico
TIPO_DOC	Tipo de documento con el que se identifica el suscriptor	Numérico
NOMBRES	Nombres del propietario del servicio	Carácter
PRIMER APELLIDO	Primer apellido del propietario del servicio	Carácter
SEGUNDO APELLIDO	Segundo apellido del propietario del servicio	Carácter

La Tabla Uso, tiene el descriptivo del identificador de los códigos de estratificación

TABLA USO		
VARIABLES	DESCRIPCIÓN	TIPO
ID_USO	identificador del servicio Residencial: 1,2,3,4,5,6; Industrial: 10, Comercial: 11, Oficial: 12	Numérico
UDO_DESCRIPCION	Descripción del código de uso	Categórica

4.2 Diagrama modelo de datos (Primera entrega)



4.3 Imágenes de la Base de Datos (Primera entrega)

Facturación

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
1	ANO	PERIODO	CICLO	CODIGOINTERNO	TIPO_DOC	NUM_DOCUMENTO	USO	ESTRATO	UBICACION	BARRIO	SECTHIDRAULICO	FECHAEXPFACTURA	LECTURA	CONSUMO	
2	2021	01	2	1207636	2	28991781	1	4	URBANO	MANILA	704	4/4/2019	228	6	
3	2021	01	2	1263845	2	20568317	1	4	URBANO	SANTA ANA	704	3/11/2020	2063	6	
4	2021	01	2	1060565	2	19244569	1	4	URBANO	MANDALAY	703	8/27/2019	130	9	
5	2021	01	2	1186352	2	20391062	1	4	URBANO	COOTRAN	703	8/6/2020	30	9	
6	2021	01	2	1100288	2	991100288	1	4	URBANO	BALMORA	704		418	8	
7	2021	01	2	1404391	2	991404391	1	4	URBANO	PIEDRAGR	704	3/21/2019	1190	6	
8	2021	01	2	1153154	2	991153154	1	3	URBANO	BALMORA	704	3/6/2020	434	15	
9	2021	01	2	1075704	2	991075704	1	3	URBANO	BALMORA	704		828	12	
10	2021	01	2	1073220	2	991073220	1	4	URBANO	MANILA	704	6/20/2016	1494	5	
11	2021	01	2	1065333	2	991065333	1	4	URBANO	PIEDRAGR	704		808	17	
12	2021	01	2	1064757	2	991064757	1	4	URBANO	PIEDRAGR	704		329	9	
13	2021	01	2	1084615	2	991084615	1	4	URBANO	BALMORA	704	2/6/2017	2771	29	
14	2021	01	2	1072131	2	991072131	1	4	URBANO	VILLA CLAY	704	2/1/2017	404	3	
15	2021	01	2	1066257	2	991066257	1	4	URBANO	PIEDRAGR	704	7/2/2019	892	17	
16	2021	01	2	1073568	2	991073568	1	4	URBANO	MANILA	704		1623	5	
17	2021	01	2	2462461	2	992462461	1	4	URBANO	SANTA ANA	704	3/6/2020	584	12	
18	2021	01	2	2423155	2	992423155	1	3	URBANO	POTOSI	703	5/6/2019	535	7	
19	2021	01	2	1053099	2	991053099	3	3	URBANO	LUXEMBURGO	703	8/18/2016	726	37	

Tipo_documento

TIPO_DOC	TIPO_DOC_DESCRIP
1	TARJETA DE IDENTIDAD
2	CEDULA CIUDADANIA
3	NIT
4	CEDULA EXTRANJERIA
5	PASAPORTE
6	SIN IDENTIFICAR

Contacto

A	B	C	D	E
TIPO_DOCUMENTO	NUM_DOCUMENTO	NOMBRES	PRI_APELLIDO	SEG_APELLIDO
2	28991781	ANA CLEOFE	RODRIGUEZ	VDA DE LUNA
2	20568317	AUDREY HERLINDA	TORRES	PARDO
2	19244569	JOSE BERNABE	SANCHEZ	LEGUIZAMON
2	20391062	NYDIA MARLENE	MARTINEZ	BARACALDO
2	991100288	PATRICIA	GIRALDO	
2	991404391	ORFA MERCED	CORIANO	
2	991153154	JAIR	HERNANDEZ	
2	991075704	MILICIADES	IRANF	PINILIA

USO

ID_USO	USO_DESCRIPCION
1	Residencial
10	Industrial
11	Comercial
12	Oficial

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

DDL for bdproyecto.contacto

```
1 CREATE TABLE `contacto` (  
2   `idcontacto` int NOT NULL,  
3   `tipo_documento` int DEFAULT NULL,  
4   `num_documento` varchar(45) DEFAULT NULL,  
5   `nombres` varchar(45) DEFAULT NULL,  
6   `pri_apellido` varchar(45) DEFAULT NULL,  
7   `seg_apellido` varchar(45) DEFAULT NULL,  
8   PRIMARY KEY (`idcontacto`)  
9 ) ENGINE=InnoDB DEFAULT CHARSET=utf8mb4 COLLATE=utf8mb4_0900_ai_ci
```

DDL for bdproyecto.facturacion

```
1 CREATE TABLE `facturacion` (  
2   `id_factura` int NOT NULL,  
3   `facturacioncol` varchar(45) DEFAULT NULL,  
4   `ano` varchar(4) DEFAULT NULL,  
5   `periodo` varchar(2) DEFAULT NULL,  
6   `ciclo` int DEFAULT NULL,  
7   `codigointerno` varchar(7) DEFAULT NULL,  
8   `tipo_doc` int DEFAULT NULL,  
9   `num_documento` varchar(45) DEFAULT NULL,  
10  `uso` varchar(1) DEFAULT NULL,  
11  `estrato` varchar(2) DEFAULT NULL,  
12  `ubicacion` varchar(45) DEFAULT NULL,  
13  `barrio` varchar(45) DEFAULT NULL,  
14  `sectorhidraulico` varchar(3) DEFAULT NULL,  
15  `fechalectura` date DEFAULT NULL,  
16  `lectura` int DEFAULT NULL,  
17  `consumo` int DEFAULT NULL,  
18  PRIMARY KEY (`id_factura`)  
19 ) ENGINE=InnoDB DEFAULT CHARSET=utf8mb4 COLLATE=utf8mb4_0900_ai_ci
```

DDL for bdproyecto.tipo_documento

```
1  CREATE TABLE `tipo_documento` (  
2    `tipo_documento` int NOT NULL,  
3    `tipo_docu_des` varchar(45) DEFAULT NULL,  
4    PRIMARY KEY (`tipo_documento`)  
5  ) ENGINE=InnoDB DEFAULT CHARSET=utf8mb4 COLLATE=utf8mb4_0900_ai_ci COMMENT=''
```

DDL for bdproyecto.usos

```
1  CREATE TABLE `usos` (  
2    `id_uso` int NOT NULL,  
3    `uso_descrip` varchar(45) DEFAULT NULL,  
4    PRIMARY KEY (`id_uso`)  
5  ) ENGINE=InnoDB DEFAULT CHARSET=utf8mb4 COLLATE=utf8mb4_0900_ai_ci
```