

# Capítulo 4

## Proceso de desarrollo del sistema de BI

En este capítulo se engloban todos los aspectos relacionados al proceso de desarrollo del sistema de BI.

### 4.1. Planificación

A continuación se presentan las distintas fases definidas para el proyecto, un cronograma de trabajo planteado y los hitos destacados.

#### 4.1.1. Fases

El plan de trabajo se enmarca dentro del enfoque presentado por Kimball en The Data Warehouse Toolkit [32] llamado Ciclo de Vida Dimensional del Negocio.

Es importante destacar que, aunque se hayan seguido los lineamientos presentados, en la fase de *Análisis y Especificación de Requerimientos* se agrega una etapa de *Prototipado* con los siguientes objetivos:

1. Aprender rápido y poder incluir las modificaciones necesarias en la fase de diseño. Esto nos permite aportar sobre nuevas funcionalidades e incluso hacer descartes cuando las propuestas no son viables.
2. Realizar pruebas que verifiquen y determinen como va a reaccionar los usuarios ante diferentes situaciones como la carga inicial, la actualización o las distintas consultas de las herramientas de análisis.
3. Testear la herramienta antes de comenzar el proceso de desarrollo. Los prototipos se analizan y se evalúan para detectar errores y posibles mejoras, confirmar que cuenta con las características deseadas y asegurar que se puede adaptar los requerimientos de análisis.

En dicha etapa se presentan a los especialistas de la IM distintos prototipos de análisis y uso basados en los datos fuentes y los requerimientos de usuario que se identificaron.

#### 4.1.2. Cronograma e hitos relevantes

A continuación, se muestra en fig. 4.1 el cronograma definido para la realización del proyecto:

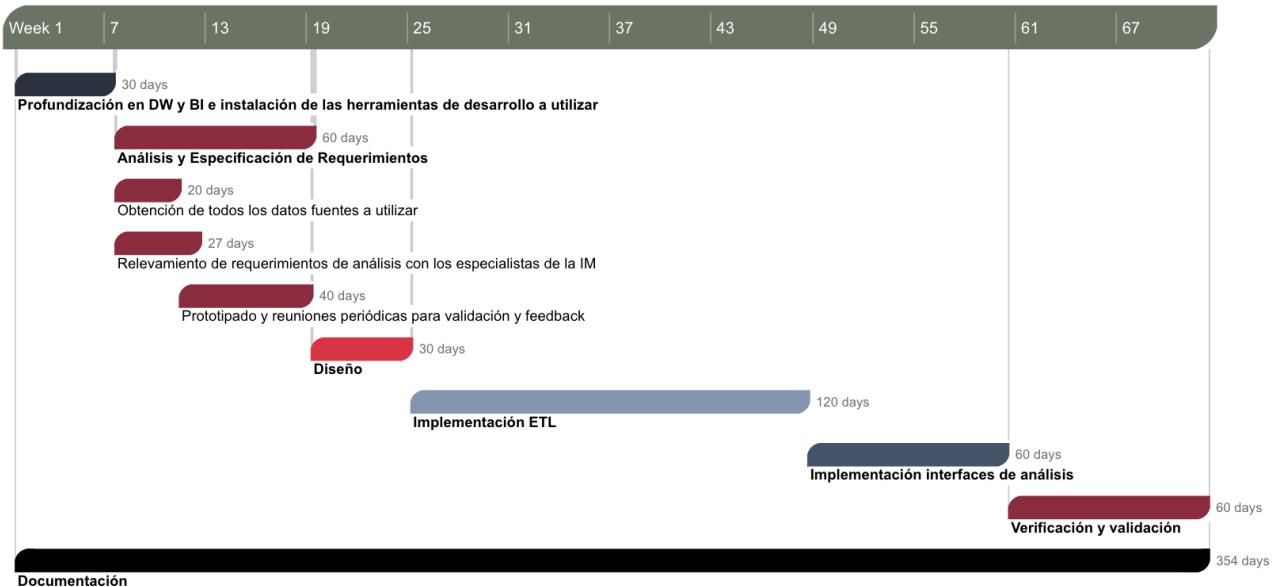


Figura 4.1: Planificación

Dentro de la planificación presentada se pueden destacar los distintos hitos dentro del plan:

1. Requerimientos de análisis y uso especificados.
2. Diseño conceptual, lógico y físico completo.
3. Implementación de requerimientos de análisis.
4. Implementación de requerimientos de uso.
5. Despliegue de la solución en ambiente productivo.
6. Documentación autocontenida.

## **4.2. Análisis**

Dentro del alcance del proyecto se encuentra la etapa de Análisis. En esta etapa se trabaja con los especialistas del Servicio de Calidad de Aire para poder comprender y plasmar formalmente las funcionalidades y requisitos que se espera del sistema de BI.

La etapa de Especificación de Requerimientos se realiza mediante un enfoque combinado, utilizando paralelamente el enfoque orientado en las fuentes y el enfoque orientado al usuario, y luego integrando los resultados de ambas etapas para lograr requerimientos que sean de utilidad para los usuarios y que puedan ser realizados con los datos disponibles.

Como base para la identificación de requerimientos, se visualizan las principales actividades contaminantes, como la quema de combustible para el transporte y la industria; el uso de leña tanto en el área comercial como residencial para la calefacción y cocción de alimentos, la generación de energía eléctrica a partir de centrales térmicas y quema de residuos a cielo abierto, de las cuales se necesita realizar un análisis, y las fuentes de datos que se disponen.

### **4.2.1. Enfoque orientado en las fuentes**

Como primer paso, se investigan todas las fuentes de datos internas, abiertas al público o no, con las que se dispone desde el Servicio de Calidad de Aire, desde bases de datos disponibles hasta archivos con datos que puedan ser de utilidad.

Luego, se investigan fuentes de datos externas que puedan ser cruzadas con las obtenidas del Servicio de Calidad de Aire para darle un mayor valor agregado al análisis y lograr visualizar las mediciones de contaminación junto con otros indicadores, para identificar relaciones entre los distintos tipos de datos. Para dicha búsqueda se utiliza el Catálogo de Datos Abiertos de la Agencia de Gobierno Electrónico y Sociedad de la Información y del Conocimiento (AGESIC) [2] donde se encuentra una gran variedad de datos y tipos de datos.

#### **4.2.1.1. Barrios de Montevideo**

Shapefile [29] de polígonos con los barrios de Montevideo. Además se tiene una descripción del área ocupada, nombre número y código.

La información de los barrios de Montevideo es mantenida en el Servicio de Geomática [6]. Dichos datos son provistos por el Instituto Nacional de Estadística (INE) [5].

En fig. 4.2 y fig. 4.3 se puede visualizar una sección de los datos:

#### **4.2.1.2. Centros Comunales Zonales (CCZ) de Montevideo**

Shapefile de polígonos con los CCZ de Montevideo. Además se informa el nombre de cada CCZ.

La información de los CCZ de Montevideo es mantenida en el Servicio de Geomática.

En fig. 4.4 y fig. 4.5 se puede visualizar una sección de los datos:

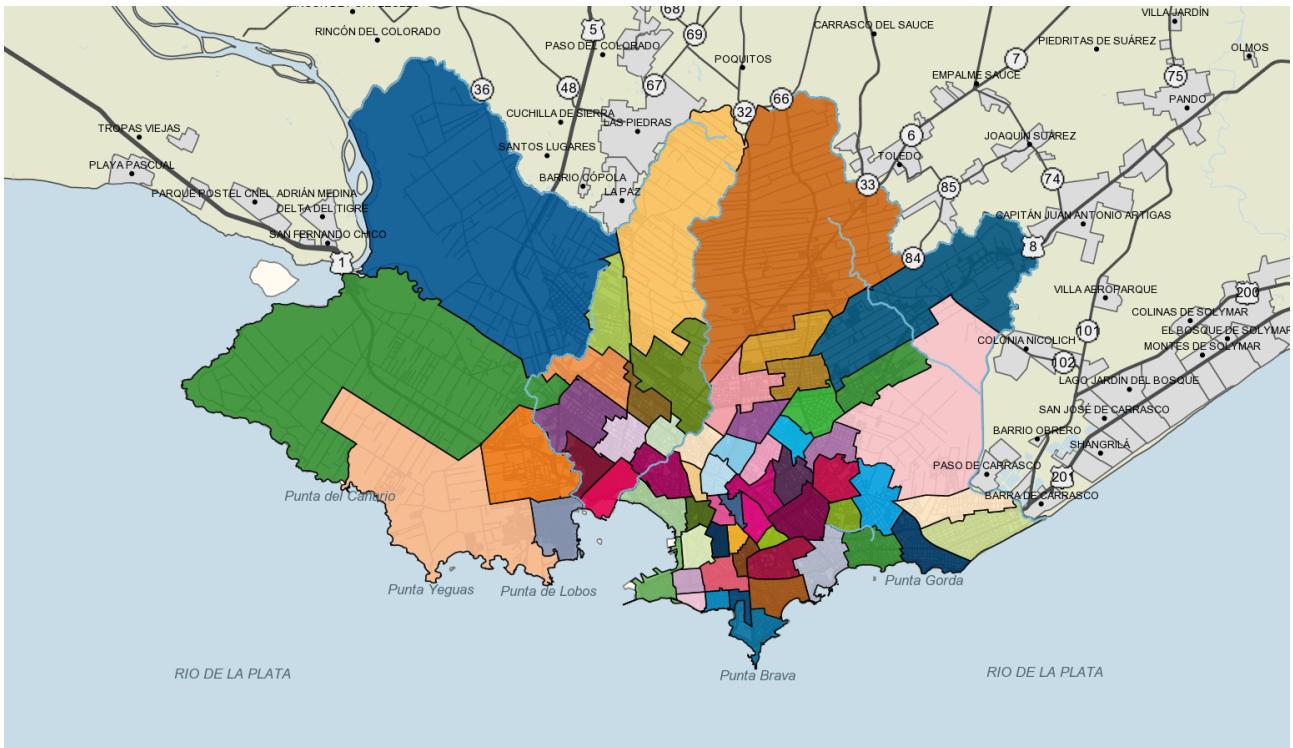


Figura 4.2: Mapa Barrios de Montevideo

AREA_KM,N,1►NOMBBARR,C,25	NROB►CDB/
2.111 CIUDAD VIEJA	1 CJ
1.297 CENTRO	2 CE
0.694 BARRIO SUR	3 SU
2.279 CORDON	4 CD
0.798 PALERMO	5 PL
0.766 PARQUE RODO	6 PR
2.73 PUNTA CARRETAS	7 PC
4.155 BUCEO	9 BU
3.375 PQUE. BATLLE, V. DOLORES	10 PB
3.501 MALVIN	11 MA
1.714 MALVIN NORTE	12 MN
3.816 PUNTA GORDA	13 PG
6.168 CARRASCO	14 CR
4.643 CARRASCO NORTE	15 CN
32.162 BA ñADOS DE CARRASCO	16 BC
3.126 MARO ñAS, PARQUE GUARANI	17 MO
2.69 FLOR DE MARO ñAS	18 FL

Figura 4.3: Datos Barrios de Montevideo

#### 4.2.1.3. Descripción de campos de medidas históricas

La siguiente fuente de datos sirve de meta-information [11] para la descripción de los archivos de medidas históricas de calidad de aire.

Se describen el formato y significado de cada campo de los archivos de datos.

En la fig. 4.6 se puede visualizar una sección conteniendo los datos:

En la misma fuente de datos se especifican los archivos de medidas históricas que se disponibilizan, como se puede visualizar en fig. 4.7 y fig. 4.8:

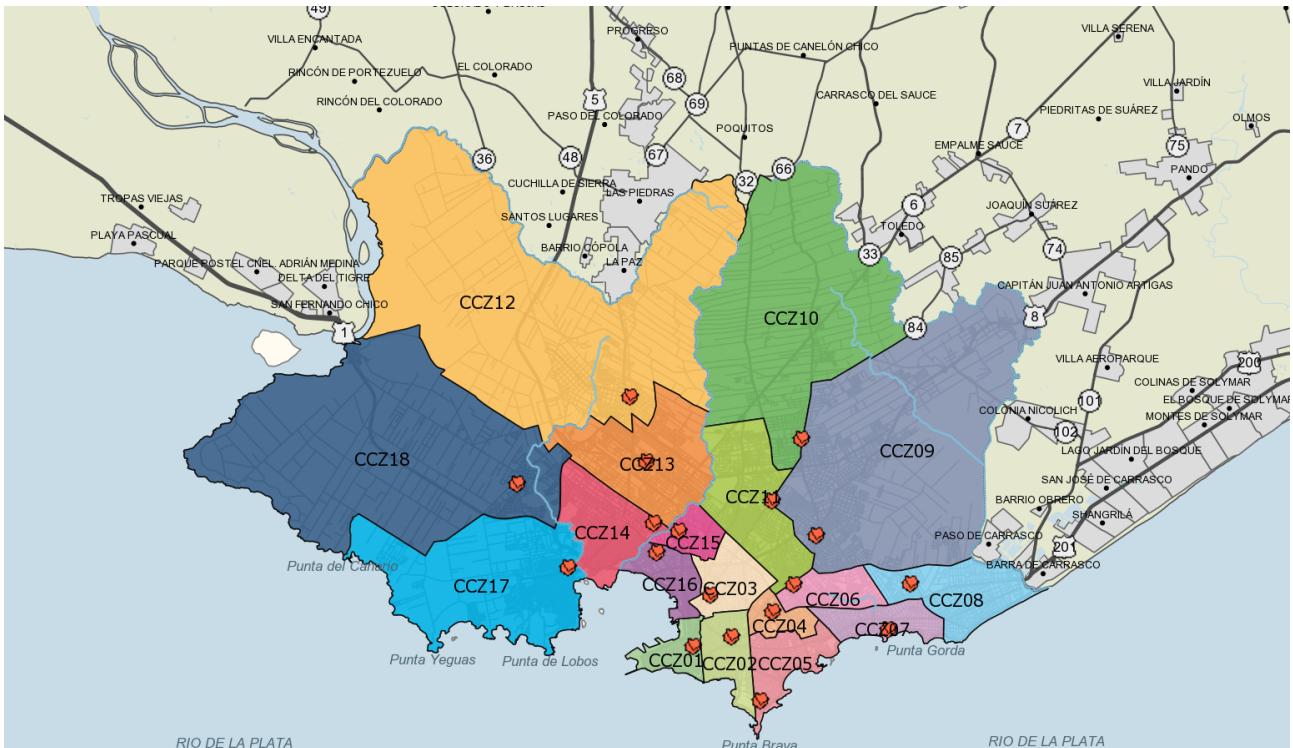


Figura 4.4: Mapa CCZ de Montevideo

GID,N,19,0	ZONA_LEGAL,C,5
8647694	CCZ12
8647695	CCZ10
8647696	CCZ17
8647697	CCZ14
8647698	CCZ09
8647699	CCZ08
8647700	CCZ07
8647701	CCZ13
8647702	CCZ11
8647703	CCZ15
8647704	CCZ02
8647705	CCZ03
8647706	CCZ04
8647707	CCZ05
8647708	CCZ06
8647709	CCZ18
8647710	CCZ01
8647711	CCZ16

Figura 4.5: Datos CCZ de Montevideo

#### 4.2.1.4. Medidas históricas calidad de aire

Estas medidas van desde el 2005 hasta el 2018 [19] y para cada una se detallé que contaminante se mide, las unidades, el valor y el método de medición. Así como otra información de interés como se explica en el archivo de descripción de la metadata.

En la fig. 4.9 se puede visualizar una sección de uno de los archivos fuente conteniendo los

Descripcion de campos de los archivos historicos	
POLLUTANT_ID:	Identificacion de contaminante. Estos identificadores se describen en el archivo contaminantes.csv
POLLUTANT_AVERAGING:	Periodo de toma de muestra en horas
DATE:	Fecha y hora que fue tomado el dato
POLLUTANT_VALUE:	Valor del parametro
POLLUTANT_UNIT:	Unidades en que esta expresado el contaminante
STATION_ID:	Denominacion de la zona que caracteriza la estacion.
X:	Localizacion. Coordenada X de la estacion (UTM- Zona 21)
Y:	Localizacion. Coordenada Y de la estacion (UTM- Zona 21)
ID_ESTACIONES:	Identificacion de la estacion segun consta en el archivo estaciones.csv
METHOD_ID:	Identifica el metodo con el cual se midio el Parametro. Los metodos se describen en el archivo metodos.csv

Figura 4.6: Descripción medidas históricas calidad de aire

Las estaciones automaticas con datos horarios se presentan los datos anuales.		
CV_pm10_14.csv	Ciudad vieja	2014
CV_pm2_15.csv	Ciudad vieja	2015
CV_pm2_16.csv	Ciudad vieja	2016
CV_pm2_17.csv	Ciudad vieja	2017
CV_pm2_18.csv	Ciudad vieja	2018
TRC_pm10_15.csv	Tres Cruces P	2014-2015
TRC_pm10_16.csv	Tres Cruces P	2016
TRC_pm10_17.csv	Tres Cruces P	2017
TRC_pm10_18.csv	Tres Cruces P	2018
TRC_NO2_15.csv	Tres Cruces N	2015
TRC_NO2_16.csv	Tres Cruces N	2016
TRC_NO2_17.csv	Tres Cruces N	2017
TRC_NO2_18.csv	Tres Cruces N	2018
CM_pm2_17.csv	Curva de Ma	2017
CM_pm2_18.csv	Curva de Ma	2018
CM_NO2_17.csv	Curva de Ma	2017
CM_NO2_18.csv	Curva de Ma	2018
CM_O3_17.csv	Curva de Ma	2017
CM_O3_18.csv	Curva de Ma	2018
CO_pm2_18.csv	Colon PM2.5	2018
CO_NO2_18.csv	Colon NO2	2018
CO_O3_18.csv	Colon O3	2018

Figura 4.7: Archivos de medidas históricas calidad de aire

Las Estaciones manuales con datos cada 24/48 horas presentan en archivos por parametro		
hn.csv	Humo Negro	2004-2018
PM2.csv	Material part	2007-2013
PM10.csv	Material part	2007-2018
pts.csv	Particulado t	2004-2018
SO2.csv	SO2 manual	2012-2018

Figura 4.8: Archivos de medidas históricas calidad de aire

datos:

A	B	C	D	E	F	G	H	I	J	K
	pollutant_id	pollutant_av date		pollutant_va	pollutant_un	station_id	X	Y	ID_estacion	method_id
1378	PM2		1 1/18 0:00	5	ug/m3	UYMVD_E8	570970	6149046	Colon	UYMVD_NO2
1379	PM2		1 1/18 1:00	5	ug/m3	UYMVD_E8	570970	6149046	Colon	UYMVD_NO2
1380	PM2		1 1/18 2:00	5	ug/m3	UYMVD_E8	570970	6149046	Colon	UYMVD_NO2
1381	PM2		1 1/18 3:00	5	ug/m3	UYMVD_E8	570970	6149046	Colon	UYMVD_NO2
1382	PM2		1 1/18 4:00	5	ug/m3	UYMVD_E8	570970	6149046	Colon	UYMVD_NO2
1383	PM2		1 1/18 5:00	5	ug/m3	UYMVD_E8	570970	6149046	Colon	UYMVD_NO2
1384	PM2		1 1/18 6:00	11	ug/m3	UYMVD_E8	570970	6149046	Colon	UYMVD_NO2

Figura 4.9: Medidas históricas calidad de aire

#### 4.2.1.5. Medidas históricas de sensores de la plataforma IoT - PM10

Este conjunto contiene las medidas históricas de los sensores de calidad del aire de la plataforma de Internet of Things (IoT) [14]. Los valores en tiempo real de estos sensores se pueden acceder a través de los servicios abiertos de la plataforma.

En la fig. 4.10 se puede visualizar una sección de los datos:

_id	id_node	sensor	timestamp	valor_v
1	AQM60	PM10	2017-03...	59138
2	AQM60	PM10	2017-03...	57787
3	AQM60	PM10	2017-03...	54746
4	AQM60	PM10	2017-03...	55875
5	AQM60	PM10	2017-03...	57302

Figura 4.10: Medidas históricas de sensores de la plataforma IoT - PM10

#### 4.2.1.6. Lista de Identificación de contaminantes

Descripción de los tipos de contaminantes que se miden [18].

En la fig. 4.11 se puede visualizar una sección de los datos:

_id	CONTA...	EXPLICACION
1	PM2	Material particulado menor de 2.5 micras de diam...
2	PM10	Material particulado menor de 10 micras de diam...
3	HN	Humo Negro (black smoke)
4	PTS	Material particulado total
5	SO2	Dioxido de azufre
6	NO2	Dioxido de nitrogeno
7	O3	Ozono

Figura 4.11: Lista de Identificación de contaminantes

#### 4.2.1.7. Métodos de medición

Descripción de los métodos de medición utilizados [20].

En la fig. 4.12 se puede visualizar una sección de los datos:

_id	METODO	EXPLICACION
1	UYMVD_PM2_b	Medida de material particulado menor de 2.5 micras de diametro (PM2.5) utilizando Nefelometria...
2	UYMVD_PM2_D	Medida de material particulado menor de 2.5 micras de diametro (PM2.5) utilizando Un equipo ...
3	UYMVD_PM10_2	Medida de material particulado menor de 10 micras de diametro (PM10) utilizando Nefelometria...
4	UYMVD_PM10_D	Medida de material particulado menor de 10 micras de diametro (PM10) utilizando equipo dicot...
5	UYMVD_PM10_H	Medida de PM10 utilizando un equipo de alto volumen para recoleccion de muestra
6	UYMVD PTS	Medida de material particulado total (PTS) utilizando un equipo de alto volumen para recoleccio...
7	UYMVD_SO2_T	Medida de dióxido de azufre (SO2) en tren de monitoreo y posterior analisis con cromatografia i...
8	UYMVD_NO2_2	Medida de dióxido de nitrógeno (NO2), semiconductor sensible al gas
9	UYMVD_HN	Medida de Humo Negro (black smoke) utilizando reflectometria como principio de deteccion
10	UYMVD_O3	Medida de ozono (O3) semiconductor sensible al gas

Figura 4.12: Métodos de medición

#### 4.2.1.8. Estaciones de medición

Lista las estaciones de medición [13].

En la fig. 4.13 se puede visualizar una sección de los datos:

_id	ESTACION	X	Y	EXPLICACION
1	Ciudad Vieja 3	572796	6137122	E1
2	Ciudad Vieja 1	572831	6136774	E1
3	Ciudad Vieja 2	572452	6137044	E1
4	Centro	574385	6137072	E2
5	Tres Cruces	576888	6138953	E5

Figura 4.13: Estaciones de medición

#### 4.2.1.9. ANCAP

Datos mediciones y datos meteorológicos históricos de la estación de ANCAP [10]. Estos datos fueron provistos por ANCAP para el uso exclusivo que se enmarca dentro del Proyecto de Grado.

En la fig. 4.14 se puede visualizar una sección de los datos:

#### 4.2.1.10. Volumen vehicular

El sistema de conteo vehicular que utiliza el Centro de Gestión de Movilidad (CGM) [7] para la medición de los volúmenes vehiculares que circulan por las arterias que se encuentran centrali-

	ANCAP1 CO Station 1	ANCAP1 NO Station 1	ANCAP1 NOx Station 1	ANCAP1 NO2 Station 1	ANCAP1 PM 2.5 Station 1	ANCAP1 SO2 Station 1	ANCAP1 TRS Station 1	ANCAP1 Global Sun rad 1 W/m2	ANCAP1 Relative Humidity 1 %	ANCAP1 Temperature Extern 1 °C	ANCAP1 Temp Internal 1 °C	ANCAP1 Wind Direction 1 degrees	ANCAP1 Wind 1 m/s
	mg/m3	microg/m3	microg/m3	microg/m3	microg/m3	microg/m3	microg/m3						
12/31/16 22:00	<b>2.00</b>				0.00	1.43	0	64	26.1	19.6	80	2.0	
12/31/16 23:00	<b>1.92</b>				0.01	1.26	0	66	25.6	19.6	62	1.4	
1/1/17 0:00	<b>1.77</b>				0.01	1.00	0	69	25.3	19.4	69	1.0	
1/1/17 1:00	<b>1.40</b>				0.61	1.77	0	73	25.6	19.4	46	1.3	
1/1/17 2:00	<b>0.76</b>				0.09	1.01	0	78	25.6	19.4	36	1.6	
1/1/17 3:00	<b>0.67</b>				0.12	0.98	0	80	25.6	19.3	13	1.1	

Figura 4.14: ANCAP

zadas, utiliza analíticas de video para la generación de los datos necesarios para la elaboración de planes y coordinaciones semafóricas. Los detectores miden el total de vehículos y su clasificación por largo de los mismos, por sentido de circulación, pudiendo discriminarse por carril dentro de cada sentido [8] [9].

El sistema genera un registro cada 5 minutos, las medidas que se explican a continuación corresponden a la suma de todos los vehículos de los 5 minutos anteriores a la fecha y hora del registro. También se publican datos particulares del sensor que permiten conocer su ubicación y entender el sentido de circulación que se está midiendo.

1. volumen: Cantidad de vehículos que se contabilizó en los 5 minutos previos en un carril dado.
2. volumen hora: Estimación de la cantidad de vehículos/hora que representan los 5 minutos previos medidos.

En la fig. 4.15 se puede visualizar una sección de los datos:

id_detector	id_carril	fecha	hora	dsc_avenida	dsc_int_anterior	dsc_int_siguiente	latitud	longitud	volumen	volumen_hora
1	3	11/1/18	0:00:00	Sarmiento	Requena	Rambla	-34.915369	-56.168358	1	12
1	3	11/1/18	0:05:00	Sarmiento	Requena	Rambla	-34.915369	-56.168358	1	12
1	3	11/1/18	0:10:00	Sarmiento	Requena	Rambla	-34.915369	-56.168358	0	0
1	3	11/1/18	0:15:00	Sarmiento	Requena	Rambla	-34.915369	-56.168358	0	0
1	3	11/1/18	0:20:00	Sarmiento	Requena	Rambla	-34.915369	-56.168358	1	12
1	3	11/1/18	0:25:00	Sarmiento	Requena	Rambla	-34.915369	-56.168358	0	0
1	3	11/1/18	0:30:00	Sarmiento	Requena	Rambla	-34.915369	-56.168358	0	0
1	3	11/1/18	0:35:00	Sarmiento	Requena	Rambla	-34.915369	-56.168358	0	0
1	3	11/1/18	0:40:00	Sarmiento	Requena	Rambla	-34.915369	-56.168358	0	0
1	3	11/1/18	0:45:00	Sarmiento	Requena	Rambla	-34.915369	-56.168358	0	0
1	3	11/1/18	0:50:00	Sarmiento	Requena	Rambla	-34.915369	-56.168358	0	0
1	3	11/1/18	0:55:00	Sarmiento	Requena	Rambla	-34.915369	-56.168358	0	0
1	3	11/1/18	1:00:00	Sarmiento	Requena	Rambla	-34.915369	-56.168358	0	0

Figura 4.15: Volumen Vehicular

#### 4.2.1.11. Vías de tránsito

Shapefile de líneas con las vías de tránsito de Montevideo [21]. Además se tiene una descripción de la estructura y contenidos de las vías.

La información de vías de tránsito es mantenida en el Servicio de Geomática, por la Unidad de Nomenclatura y Numeración.

En la fig. 4.16 se puede visualizar la fuente de datos:



Figura 4.16: Vías de tránsito

#### 4.2.1.12. Encuesta continua de hogares

Microdatos, cuestionarios, diccionario de variables y notas de la Encuesta Continua de Hogares [12].

La encuesta continua de hogares es realizada por INE.

En la fig. 4.17 se puede visualizar la fuente de datos:

<u>_id</u>	<b>DESCRIPCIÓN DE LA VARIABLE</b>	<b>ECH 2017</b>
4	HOGAR. IDENTIFICACIÓN	
5	AÑO DE LA ENCUESTA	ANIO
6	IDENTIFICACIÓN DEL CUESTIONARIO	NUMERO
7	DEPARTAMENTO	DPTO
8	NOMBRE DEL DEPARTAMENTO	NOMDPTO
9	SECCIÓN CENSAL	SECC
10	SEGMENTO CENSAL	SEGM
11	LOCALIDAD AGRUPADA	LOCAGR
12	NOMBRE DE LA LOCALIDAD AGRUPADA	NOM_LOCAGR
13	CENTRO COMUNAL ZONAL	CCZ
14	BARRIO	BARRIO
15	NOMBRE DEL BARRIO	NOMBARRIO
16	MES DE LA ENCUESTA	MES
17	ESTRATO	ESTRED13

Figura 4.17: Encuesta continua de hogares

#### 4.2.1.13. Esquema de las fuentes de datos

En la fig. 4.18 se presenta un modelo conceptual de las fuentes de datos:

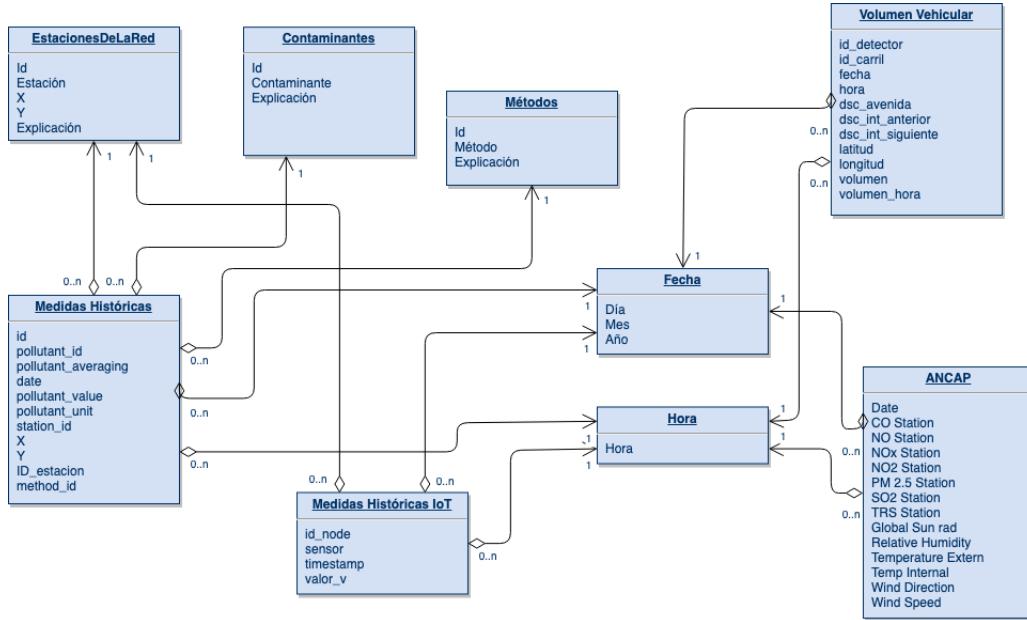


Figura 4.18: Esquema conceptual de las fuentes

#### 4.2.2. Enfoque orientado al usuario

Para identificar las necesidades que los usuarios del sistema de BI quieren o pueden atacar, se realizan sucesivas reuniones presenciales donde se discuten y ponen en la mesa ideas o solicitudes de análisis de datos o situaciones puntuales que han surgido en el Servicio de Calidad de Aire en los últimos años.

Dentro de los puntos que se lograron identificar se encuentran:

1. Relacionar la combustión realizada en la quema de leña o carbón al hacer asados con los días festivos, por ejemplo, un 1ro de mayo.
2. Investigar en qué medida varía la emisión de contaminantes en los días que la Central Batlle se encuentra encendida generando energía.
3. Investigar la relación entre distintos contaminantes, esto es, cuánto influye en la emisión de un contaminante el aumento en la emisión de otro. Por ejemplo, cuánto influye un aumento de emisión de PM10 en la presencia de Humo Negro.

#### 4.2.3. Estudio de factibilidad

Con las fuentes externas e internas disponibles se realizan distintos prototipos [45] con cruza-mientos entre las fuentes de datos internas y externas para identificar, junto a los especialistas de la IM, cuáles son los relacionamientos más relevantes que pueden convertirse en requerimientos para un posterior análisis y desarrollo de los mismos.

Los prototipos fueron presentadas en las herramientas de análisis de datos, viendo distintos ejemplos reales de cómo se visualizan los datos ingresados en el sistema.

Esta forma de trabajo permite obtener feedback de los usuarios en etapas tempranas del desarrollo del sistema, lo que brinda la posibilidad de realizar cambios o ajustar tempranamente los requerimientos.

A continuación se describen los distintos cruzamientos que se realizaron pero que no fueron tomados en cuenta para un posterior desarrollo:

Cruzamientos potenciales descartados	
Cruzamiento	Motivo
Evolución en el tiempo de la contaminación del aire relacionándolo con la densidad de vehículos en las calles y los tipos de caminería.	Aunque el estudio de en qué vías presentan una mayor concentración según un mayor volumen vehicular es muy prometedor, sucede que alrededores de las cámaras vehiculares y de las estaciones de monitoreo de la red, los tipos de caminería son iguales, imposibilitando un análisis por tipo de caminería, por ejemplo, logrando una comparación entre pavimento y caminos de tierra.

A su vez se descartan las siguientes fuentes de datos:

1. Vías de tránsito.
2. Encuesta continua de hogares.

#### 4.2.4. Especificación de requerimientos funcionales

Basados en ambos enfoques y en el estudio de factibilidad se presentan los requerimientos funcionales acordados con los especialistas de la IM según las fuentes de datos disponibles y las necesidades identificadas. Primeramente se especifican los requerimientos funcionales de análisis:

Requerimientos Funcionales de Análisis		
ID	Requerimiento Funcional	Descripción
RFA1	RFA1 - Relación Volumen Vehicular y Contaminación	Analizar la evolución en el tiempo de la contaminación del aire relacionándolo con la densidad de vehículos en las calles.
RFA2	RFA2 - Relación Actividad Industrial y Contaminación	Analizar la evolución en el tiempo de la contaminación del aire que generan las industrias a su alrededor teniendo en cuenta su nivel de actividad.
RFA3	RFA3 - Relación Días Especiales y Contaminación	Analizar la evolución en el tiempo de la contaminación del aire generados por la combustión, específicamente la quema de leña generada por los asados, en días especiales como los festivos, feriados, fechas patria y días con eventos destacables.

A continuación se describen los requerimientos funcionales de uso, éstos son, las herramientas a utilizar por los especialistas de la IM para el manejo de la plataforma.

Requerimientos Funcionales de Uso		
ID	Requerimiento Funcional	Descripción
RFU1	RFU1 - Aplicación web de carga y actualización	Permitir acceso efectivo a usuarios finales mediante la implementación de un aplicativo para la realización de la carga y actualización del DW que resulte amigable y no requiera de amplio conocimiento técnico.
RFU2	RFU2 - Generación de reporte de BI de sección en Informe Anual	Automatizar la realización de un reporte de BI de la sección Resultados de Red de Monitoreo del Informe Anual, donde detallan las características de cada estación y se muestra un resumen de los resultados expresados utilizando el ICAire.
RFU3	RFU3 - Generación de un dashboard para análisis	Generar un dashboard con gráficas e indicadores predefinidos para un posterior análisis

#### 4.2.4.1. RFA1 - Relación Volumen Vehicular y Contaminación

*Analizar la evolución en el tiempo de la contaminación del aire relacionándolo con la densidad de vehículos en las calles.*

El siguiente requerimiento intenta realizar un análisis sobre uno de los principales focos de contaminación de aire como es el transporte. El objetivo del mismo es poder visualizar si a mayor densidad de autos existe una mayor contaminación.

El mismo surge del cruzamiento de una fuente de datos interna del Servicio de Calidad de Aire y una fuente de datos externa obtenida del catálogo de datos abiertos de AGESIC.

Para relacionar el volumen vehicular en las calles de Montevideo con la contaminación, es necesario investigar en paralelo, con datos horarios, la emisión de los contaminantes PM2, PM10 o PTS, junto con el volumen de vehículos que se encuentran alrededor de las estaciones de la red en las que se realizan las mediciones de dichos contaminantes.

Para tomar datos de casos reales para realizar la investigación, se define junto a los especialistas del Servicio de Calidad de Aire un rango de distancia entre una estación de la red y las cámaras de conteo vehicular. El valor definido es *1km* de radio pero es modificable al momento de la carga.

En la etapa de carga de datos de contaminación y volumen vehicular se filtrarán y no se relacionarán los datos obtenidos de estaciones de monitoreo y cámaras de conteo vehicular que se encuentren a un radio de más de *1km*.

Analizando las fuentes de datos con las que se cuenta en el momento de la elaboración del informe, se puede observar que:

Para el conteo vehicular:

1. Se obtuvieron datos del conteo vehicular del 2017 y 2018, por lo tanto se podrá relacionar con las mediciones de contaminación para ese período.
2. Aunque se obtuvo un número considerable de cámaras de conteo vehicular, se podrán relacionar sólo las que se encuentren en un radio de *1km* con alguna estación de la red.

Para las mediciones de contaminación:

1. Existen datos de PM2, PM10 y PTS semanales en varias estaciones de monitoreo, como Ciudad Vieja y Portones, pero no horarios o por minutos como es requerido.
2. Existen datos horarios de PM10 en Ciudad Vieja y Tres Cruces, pero no se encuentran en el mismo período de medición del conteo vehicular.

Por lo tanto, al momento de la elaboración de este informe será posible relacionar:

1. Mediciones de contaminación de PM10, obtenidas en la estación de la red que se encuentra en Tres Cruces para el período 2017.
2. Datos de conteo vehicular de las cámaras que se encuentran a menos de *1km* de la estación de Tres Cruces para el período 2017. Se puede visualizar los datos obtenidos en fig. 4.19.

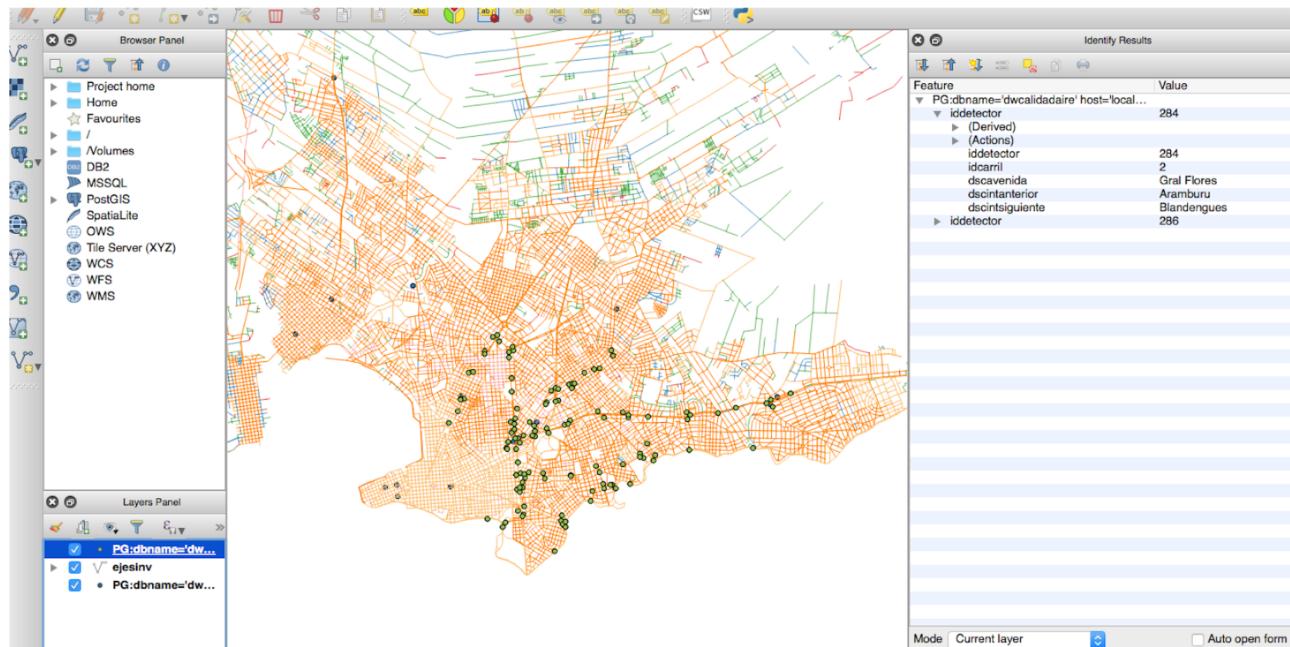


Figura 4.19: Radio de *1km* alrededor de las estaciones de la red y cámaras vehiculares

En el prototipo desarrollado para mostrar la relación entre ambos datos, se realizó una carga inicial manual y se generó un cubo de ejemplo para ser visualizado en Saiku, como se puede ver en fig. 4.20, que lograra mostrar básicamente la relación entre ambos datos.

Algunos puntos interesantes para estudiar serían las estaciones de monitoreo en el Palacio Legislativo y en Portones de Carrasco pero el conteo vehicular se tienen datos sólo de 2017 y necesitaríamos datos horarios de los contaminantes en dichas estaciones de monitoreo. UTE tiene estos datos horarios pero no se logró obtener los datos al momento de realizar el informe. En la estación de ANCAP se tienen datos horarios pero no se tiene conteo vehicular alrededor.

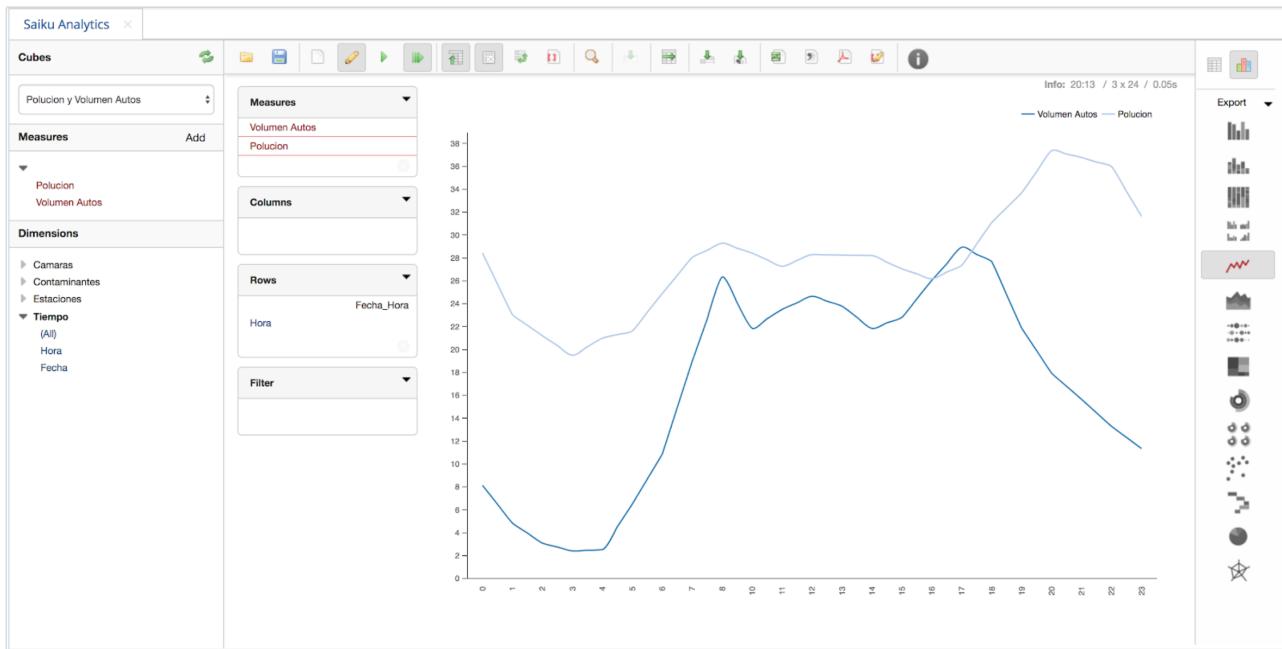


Figura 4.20: Prototipo - Relación Contaminación - Volumen Vehicular

#### 4.2.4.2. RFA2 - Relación Actividad Industrial y Contaminación

Analizar la evolución en el tiempo de la contaminación del aire que generan las industrias a su alrededor teniendo en cuenta su nivel de actividad.

El requerimiento pretende realizar un análisis sobre otro de los principales focos de contaminación de aire como son los procesos industriales. El objetivo es poder entender si a mayor actividad se encuentra una mayor contaminación en el aire.

El mismo surge del cruzamiento de una fuente de datos interna del Servicio de Calidad de Aire y fuentes de datos externas de industrias obtenidas.

Para relacionar la actividad industrial con la contaminación, es necesario investigar en paralelo, con datos diarios, la emisión de los contaminantes PM2, PM10 o PTS, junto con el nivel de actividad de las industrias que se encuentran alrededor de las estaciones de la red en las que se realizan las mediciones de dichos contaminantes.

Por lo tanto, para tomar datos de casos reales para realizar la investigación, se define junto a los especialistas del Servicio de Calidad de Aire un rango de distancia entre una estación de la red y una industria. El valor definido es 3km de radio pero es modificable al momento de la carga.

En la etapa de carga de datos de contaminación y nivel de actividad se filtrarán y no se relacionarán los datos obtenidos de estaciones de monitoreo e industrias que se encuentren a un radio de más de 3km.

Analizando las fuentes de datos con las que se cuenta en el momento de la elaboración del informe, se puede observar que:

Para el nivel de actividad:

1. Dentro de las industrias a investigar, de la cual se logró conseguir datos actuales, se

encuentra la Central Batlle mediante la información provista por Administración del Mercado Eléctrico (ADME) [1].

2. Se obtuvieron datos de nivel de actividad desde 2015 por lo tanto se podrá relacionar con las mediciones de contaminación para ese período.
3. Aunque existen un número considerable de estaciones de la red, se podrán relacionar sólo las que se encuentren en un radio de  $3km$  con la Central Batlle. Para el momento que se realizó el informe, la estación de la red en Ciudad Vieja es la única estación en el rango de la cual se disponían datos para cruzar.

Para las mediciones de contaminación:

1. Existen datos de múltiples contaminantes con periodicidad diaria y semanal.

Por lo tanto, al momento de la elaboración de este informe será posible relacionar:

1. Mediciones de múltiples contaminantes obtenidas en la estación de la red que se encuentra en Ciudad Vieja para el período desde 2005 hasta la fecha.
2. Datos de la actividad industrial de la Central Batlle desde 2015. Se puede visualizar los datos obtenidos en fig. 4.21.

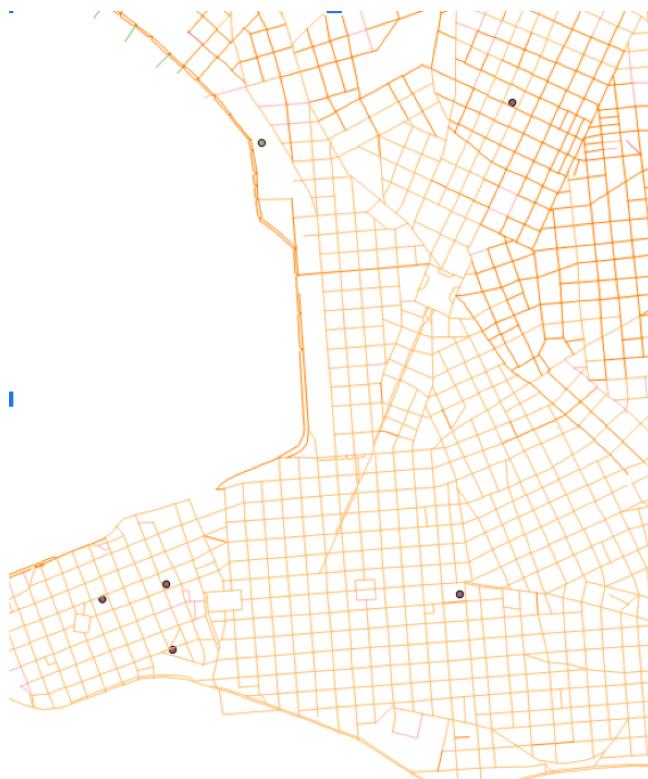


Figura 4.21: Radio de  $3km$  alrededor de la Central Batlle

En el prototipo desarrollado para mostrar la relación entre ambos datos, se realizó una carga inicial manual y se generó un cubo de ejemplo para ser visualizado en Saiku, como se puede ver en fig. 4.22, fig. 4.23 y fig. 4.24, que lograra mostrar básicamente la relación entre ambos datos.

## PM2\_5

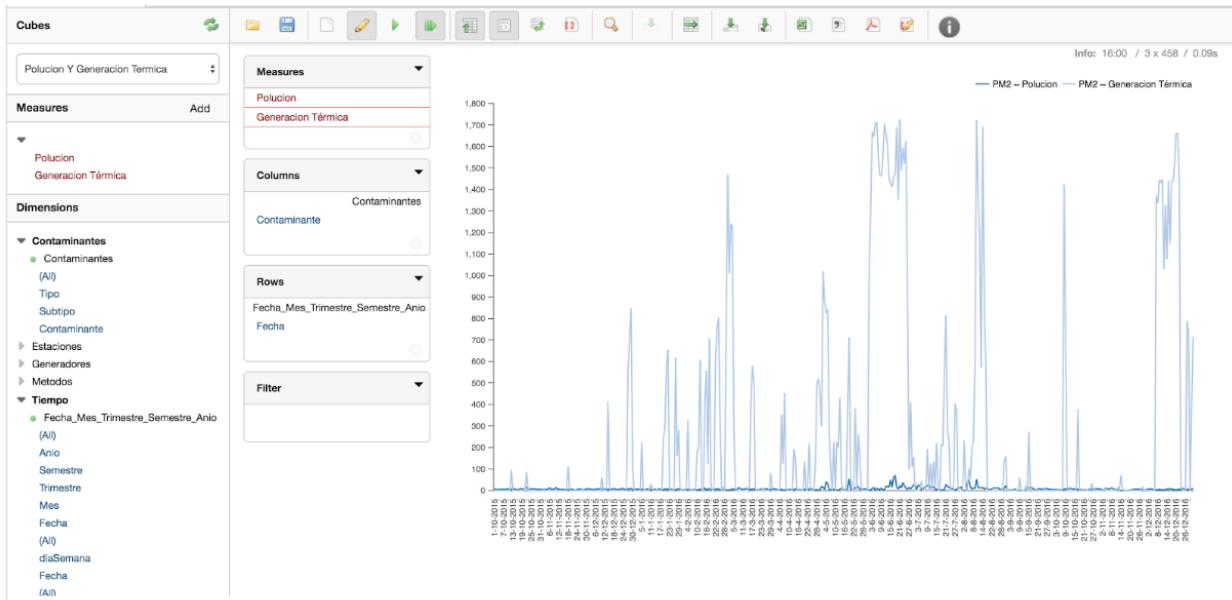


Figura 4.22: Prototipo - Relación Contaminación en PM2.5 - Actividad Industrial

## Humo Negro

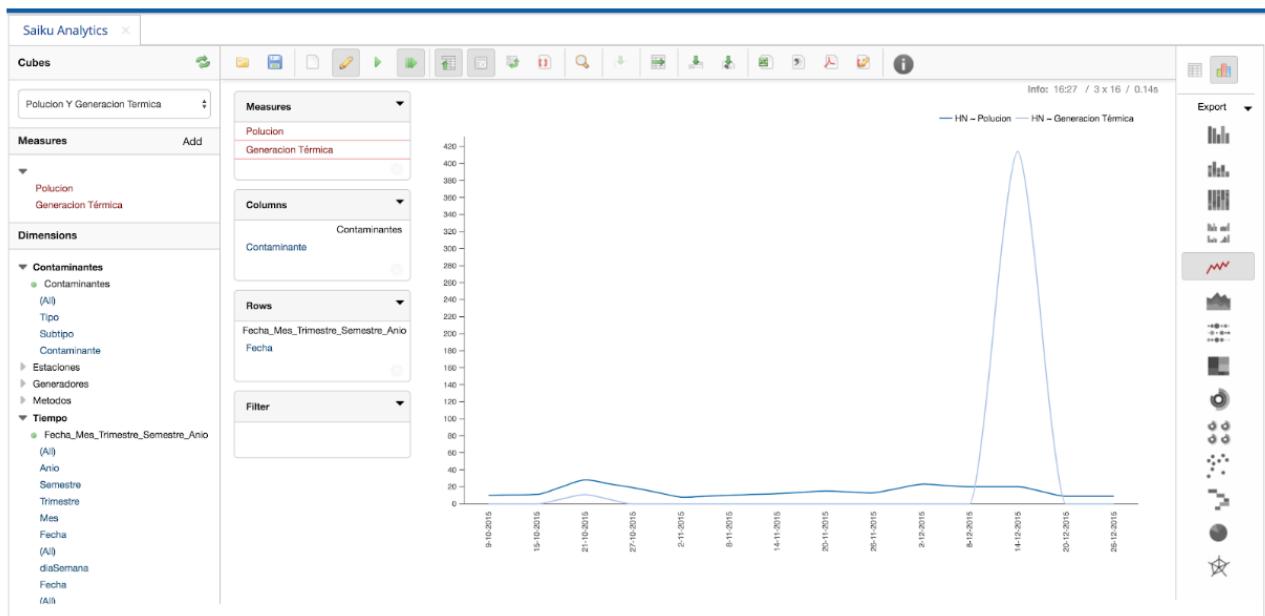


Figura 4.23: Prototipo - Relación Contaminación en Humo Negro - Actividad Industrial

### 4.2.4.3. RFA3 - Relación Días Especiales y Contaminación

*Analizar la evolución en el tiempo de la contaminación del aire generados por la combustión, específicamente la quema de leña generada por los asados, en días especiales como los festivos, feriados, fechas patria y días con eventos destacables.*

## SO2

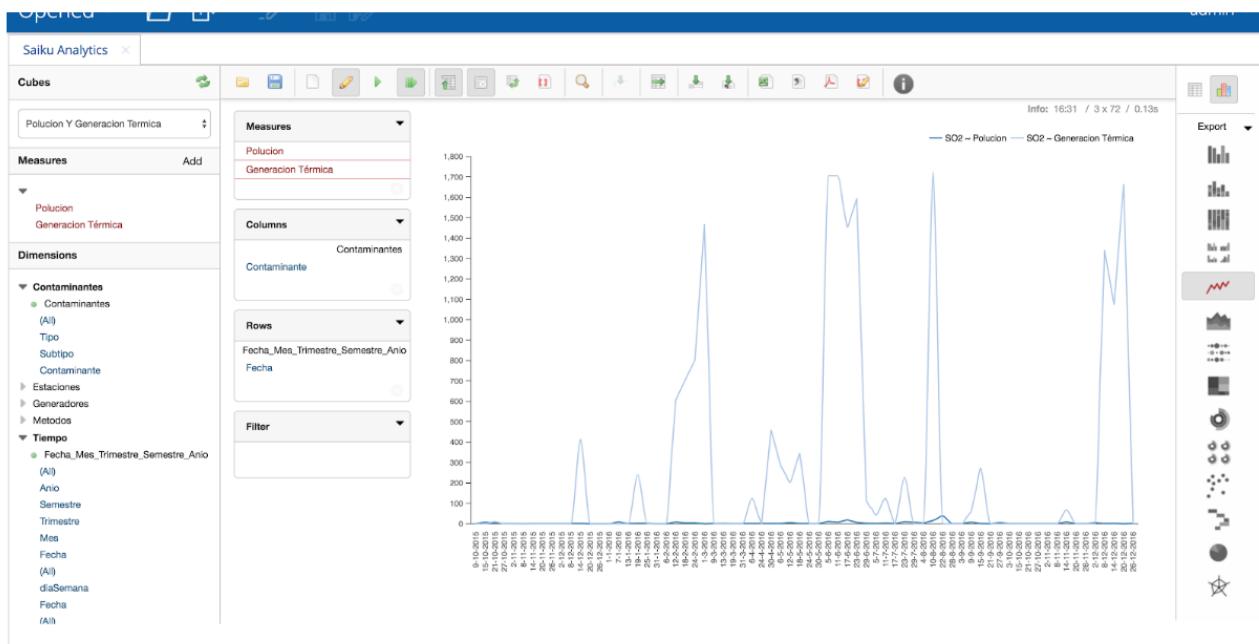


Figura 4.24: Prototipo - Relación Contaminación en SO2 - Actividad Industrial

El requerimiento pretende realizar un análisis sobre otro de los principales focos de contaminación de aire como son los procesos de quema de leña por distintos motivos, en especial, los asados. El objetivo es poder entender si en los días que se produce una mayor quema de leña se encuentra una mayor contaminación en el aire.

El mismo surge del cruzamiento de una fuente de datos interna del Servicio de Calidad de Aire y fuentes de datos externas para obtener los días festivos, feriados y fechas patria, y del ingreso manual de fechas con eventos destacables.

Para relacionar los días especiales con la contaminación, es necesario investigar en paralelo, con datos horarios y diarios, la emisión de los contaminantes PM2, PM10 o PTS, junto con los días especiales. Cabe destacar que son relevantes los datos obtenidos de todas las estaciones de la red.

Por lo tanto, para tomar datos de casos reales para realizar la investigación, se define junto a los especialistas del Servicio de Calidad de Aire los días que serán tomados en cuenta para la realización del análisis.

Analizando las fuentes de datos con las que se cuenta en el momento de la elaboración del informe, se puede observar que:

Para los días especiales:

1. Se obtuvieron datos de los días festivos del Ministerio de Turismo (MINTUR) [22].
2. Se obtienen e ingresan los días que hay partidos de fútbol relevantes (Selección Uruguaya de Fútbol, partidos de los equipos grandes).

Para las mediciones de contaminación:

1. Existen datos de múltiples contaminantes con periodicidad diaria y semanal.

Por lo tanto, al momento de la elaboración de este informe será posible relacionar:

1. Mediciones de múltiples contaminantes obtenidos en las estaciones de la red de monitoreo para el período desde 2005 hasta la fecha.
2. Días especiales tomados de las fuentes externas y de las ingresadas manualmente.

En el prototipo desarrollado para mostrar la relación entre ambos datos, se realizó una carga inicial manual y se generó un cubo de ejemplo para ser visualizado en Saiku, como se puede ver en fig. 4.25, que lograra mostrar básicamente la relación entre ambos datos.

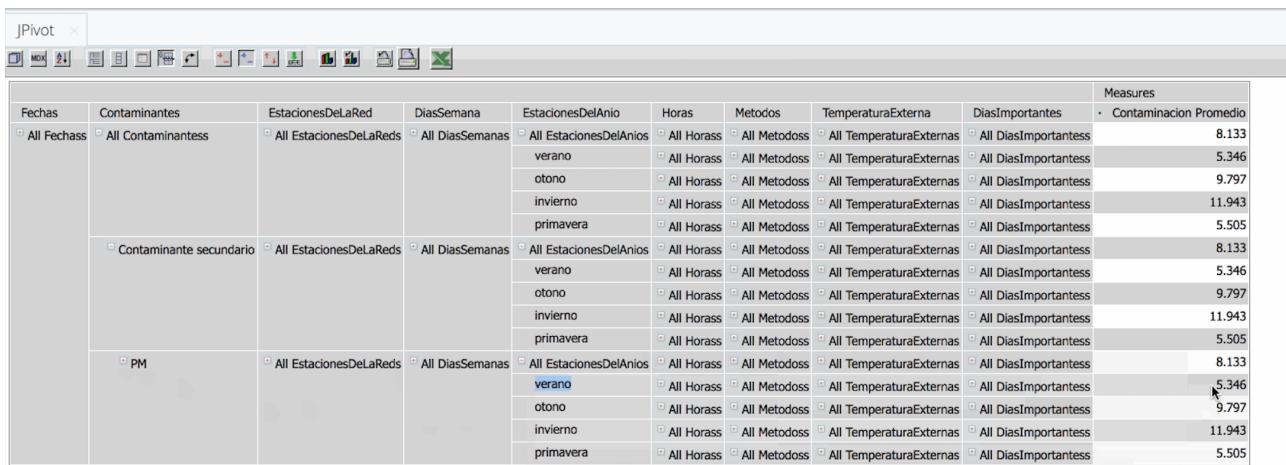


Figura 4.25: Prototipo - Relación Contaminación en PM2.5 - Días Importantes

#### 4.2.4.4. RFU1 - Aplicación web de carga y actualización

*Permitir acceso efectivo a usuarios finales mediante la implementación de un aplicativo para la realización de la carga y actualización del DW que resulte amigable y no requiera de amplio conocimiento técnico.*

Para promover el uso del sistema de BI es necesario brindar herramientas simples e intuitivas para las tareas que son indispensables de realizar periódicamente para su gestión, como lo son la carga y actualización de los datos desde las fuentes. Para cumplir dicho motivo se disponibiliza una aplicación web fácilmente accesible.

#### 4.2.4.5. RFU2 - Generación de reporte de BI de sección en Informe Anual

*Automatizar la realización de un reporte de BI de la sección Resultados de Red de Monitoreo del Informe Anual, donde detallan las características de cada estación y se muestra un resumen de los resultados expresados utilizando el ICAire.*

Actualmente, los especialistas del servicio ECCA, generan manualmente la información, estadísticas y reportes de la sección Resultados de Red de Monitoreo del Informe Anual como se visualiza en fig. 4.26. Se espera la posibilidad de generar dicho reporte de BI de la sección a demanda, desde aplicativo web a implementar, para con los datos que se encuentren en el DW para el año indicado.

## 2.2 Estación 1: Ciudad Vieja

Estación de Base

### Academia Uruguay

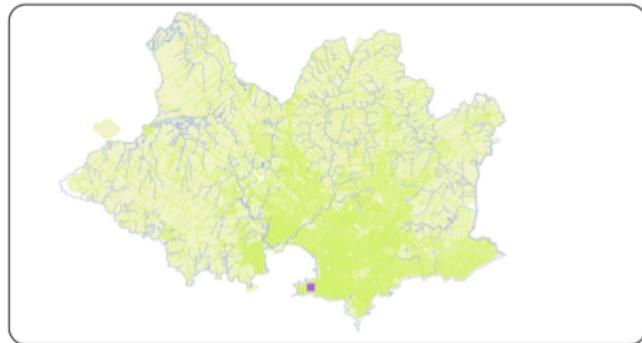
UTM21 S  
SIRGAS 2000  
X=572795 Altura sobre Nivel del mar 20m  
Y=6137106 Elevación desde la calzada 9 m

### Sede de Asociación de Empleados Bancarios de Uruguay

UTM21 S  
SIRGAS 2000  
X=572831 Altura sobre Nivel del mar 14m  
Y=6136758 Elevación desde la calzada 3 m

Parámetro	Unidades	Método de medida	Periodo evaluado
-----------	----------	------------------	------------------

Material Particulado PM2.5	PM10	ug/m <sup>3</sup>	Light Scattering	Horaria
Humo Negro	HN	ug/m <sup>3</sup>	Reflectometría	24 horas cada 6 días
Dióxido de Azufre	SO <sub>2</sub>	ug/m <sup>3</sup>	Cromatografía Iónica	24 horas cada 6 días



Mapa 2.2 Ciudad Vieja

	Promedio (ug/m <sup>3</sup> )	IC (ug/m <sup>3</sup> )	Número de Muestras	Máximo (ug/m <sup>3</sup> )	Mínimo (ug/m <sup>3</sup> )	Percentil 95 (ug/m <sup>3</sup> )
PM2.5	13	(12-15)	348	67	3	26
SO <sub>2</sub>	5	(3-6)	15	21	1	20
Humo Negro	14	(11-17)	30	49	1	24

Tabla 2.2 Resultados Ciudad Vieja

Figura 4.26: Sección Resultados de Red de Monitoreo - Informe Anual

#### 4.2.4.6. RFU3 - Generación de un dashboard para análisis

*Generar un dashboard con gráficas e indicadores pre-definidos para un posterior análisis*

Con el objetivo de brindar herramientas de BI visuales para descubrir insights que logren retratar fielmente el estado de la calidad de aire en Montevideo mediante un acceso eficaz y eficiente a la información, se genera un Dashboard con gráficas e indicadores pre-establecidos para poder ser consultado de manera rápida e intuitiva.

#### 4.2.5. Especificación de requerimientos no funcionales

Basados en ambos enfoques y en el estudio de factibilidad se presentan los requerimientos no funcionales acordados con los especialistas de la IM que deberá cumplir el sistema.

1. Implementar la solución utilizando la suite de Pentaho Community Edition (CE).
2. Funcionar en arquitecturas de varios niveles.
3. Mantener una relación adecuada con Bases de Datos (BDs) Fuentes:
  - a) Acceso a BDs heterogéneas y multiplataforma.

- b)* Independiente de los Sistemas de Producción.
4. Proveer documentación técnica detallada del sistema, su instalación, configuración, inicialización y monitoreo para su posterior gestión, actualización y soporte por parte del Servicio de Calidad de Aire.
  5. Proveer documentación funcional del sistema, un manual de usuario y tutoriales web para su posterior utilización por parte del Servicio de Calidad de Aire.