

TABLERO DE CONTROL PARA EL SEGUIMIENTO DE LA ESTRATEGIA

DOCUMENTACIÓN DE USUARIO

Versión 1.0
Noviembre 2018



Contenido

Introducción	3
Arquitectura del Sistema de Integración de Datos.....	5
Proceso de Instalación.....	6
Proceso de Integración de Datos	13
Pentaho Data Integration.....	13
Estructura del Datawarehouse SENAVITAT (MUVH)	14
Datamart obras SENAVITAT.....	14
Proceso ETL	17
Extracción.....	19
Transformación y Carga.....	24
Tabla de Hechos.....	32
Tablero de Control Cuantitativo.....	33
Ejecución Física	33
Desempeño de Constructoras	34
Mapa de Viviendas	35
Nivel Operadores	36
Tablero de Control Cualitativo.....	37
Observaciones Finales.....	38

Introducción

La necesidad de utilización de herramientas de seguimiento de la estrategia se remonta a la popularización del modelo de administración por objetivos por parte de (Drucker, 1954), en contraposición a la administración por procesos, hasta entonces aplicada en la mayoría de las organizaciones, tanto privadas como públicas. Posteriormente, debido que el énfasis colocado en los objetivos financieros provocaba un cortoplacismo sistemático, principalmente por parte de las empresas privadas centradas en las retorno sobre capital, (Kaplan & Norton, 1992) basados en varios estudios de casos, proponen la implementación de un sistema de objetivos integral y balanceado entre el corto, el mediano y el largo plazo. El modelo “Balanced Score Card” o “Cuadro de Mando Integral”, traducido al español, propone la distribución de los objetivos organizacionales en cuatro perspectivas vinculadas de manera causal: (1) Aprendizaje y Crecimiento -> (2) Procesos Internos -> (3) Clientes y Ventas -> (4) Resultados Financieros. La lógica radica en que la inversión en capital humano (Aprendizaje), así como en infraestructura y tecnología (Crecimiento) permite la mejora de los procesos internos, los cuales, a su vez, permiten proporcionar una mayor satisfacción a los clientes, lo que apuntala las ventas y finalmente se traduce en mayor rentabilidad para la compañía. A partir de la publicación de estos hallazgos el esquema de Cuadro de Mando Integral ha ganado gran popularidad entre las organizaciones del sector privado logrando una mayor efectividad en el gerenciamiento y comunicación de la estrategia empresarial (Mukherjee & Pandit, 2009).

Por otro lado, en lo que se refiere al sector público, el nuevo modelo de administración pública, difundido mundialmente a partir de los 1970s, también hace énfasis en la administración por objetivos (Katsamunskaya, 2012). De la mano de organizaciones multilaterales de cooperación y desarrollo, particularmente el Banco Mundial y la CEPAL, se han popularizado en Latinoamérica los esquemas de seguimiento de la estrategia basados en el proceso de producción de bienes públicos (Armijo, 2011). Tal proceso consta de cuatro etapas: (1) Obtención de Recursos; (2) Producción de bienes y servicios; (3) Distribución de bienes y servicios y (4) Evaluación del impacto económico y social de los bienes y servicios públicos. Tal esquema no es muy diferente al de (Kaplan & Norton, 1992): También es un esquema que refleja una causalidad, partiendo de recursos, pasando por procesos internos para la producción y distribución, hasta finalmente producir un impacto económico y social en la población objetivo. La adaptación del Cuadro de Mando Integral para el sector público, por lo tanto, no es muy problemática y resulta compatible con el modelo de CEPAL (BARROS, BATISTA, & RODRÍGUEZ, 2004).

Resulta claro, sin embargo, que cualquiera sea el modelo de seguimiento de la estrategia, siempre estará basado en la definición de un conjunto de indicadores clave o “Key Performance Indicators” (KPI). Estos son los que definen de manera cuantitativa y medible las expresiones reflejadas en los objetivos estratégicos. Por ejemplo, si un objetivo estratégico se establece como “Incrementar la satisfacción de los clientes”, debe tener un indicador asociado que mida de alguna manera dicha satisfacción, para la cual se generalmente se construyen uno o varios índices. Otro ejemplo, de particular interés en el sector público, tiene que ver con la focalización. Generalmente se pretende incrementar el acceso de la población pobre a los servicios públicos, por lo que el porcentaje de pobres atendidos resulta un indicador natural.

En la práctica, a su vez, estos indicadores se obtienen de diferentes fuentes de datos administradas por las organizaciones. La recopilación e integración de los datos de modo a construir indicadores ha dado origen a todo un nuevo campo en la informática, usualmente denominado “Management Information Systems” o “Sistemas de Información Gerencial”. Las mejores prácticas recomiendan la integración de los datos en lo que se conoce como un Datawarehouse (Kimball & Caserta, 2004). El Datawarehouse a su vez consiste en una serie de métricas, a partir de las cuales es posible construir indicadores, asociadas a datos categóricos que permiten su análisis a diferentes niveles o dimensiones. La lógica del Datawarehouse es contener la información en un esquema de datos orientado a la consulta y el análisis, a diferencia de los sistemas de registro o sistemas transaccionales, que se encuentran orientados a minimizar la utilización de recursos (como espacio de disco y memoria), disminuir el riesgo de pérdida de información y maximizar el desempeño (Kimball & Caserta, 2004). Los Datawarehouse, a su vez, son construidos a través de procesos conocidos como ETL (Extracción, Transformación y Carga o *Loading*). Los cuales, a través de herramientas especializadas, logran integrar los datos a partir de diferentes fuentes y orígenes que pueda administrar una organización.

En los sistemas de ETL es usual encontrar altos niveles de normalización en los sistemas de origen, lo cual es eficiente para los procesos transaccionales, pero no muy eficiente para los procesos de consulta y análisis, que requieren un nivel de normalización menor. Las prácticas más comunes apuntan a construir una tabla de hechos, con campos exclusivamente numéricos, que contienen, por un lado, las métricas de interés, mientras que por otro claves foráneas (foreign keys) que apuntan a tablas que contienen dimensiones, es decir, campos, ya sean alfanuméricos o numéricos discretos, que constituyen la base de análisis de las métricas. A veces las dimensiones apuntan a subdimensiones, por lo que la dimensionalidad del esquema se incrementa, pero lejos aún de los niveles encontrados en sistemas transaccionales.

En SENAVITAT (Actualmente Ministerio de Urbanismo Vivienda y Hábitat, MUVH) por el contrario, no se encontraron sistemas transaccionales altamente normalizados, sino que tablas de distinto grado de normalización, algunas de ellas sin ningún grado, por lo que el proceso de ETL tomó caminos poco habituales. Además de ello nos encontramos con datos sobre el mismo objeto (por ejemplo beneficiarios) que se duplicaban en distintas fuentes, respondiendo a aplicativos diferentes que se generaron (por ejemplo) para digitalizar distintos programas. Mucha de la información potencial, a su vez, no logró ser utilizada por la carga incompleta, errónea o falta de carga.

A pesar de todas estas limitaciones, sin embargo, con la información encontrada en SENAVITAT, además de las correcciones realizadas por el equipo de la entidad, altamente comprometido con este esfuerzo, se ha logrado diseñar un ETL que presenta datos suficientes para una primera versión de una herramienta de la estrategia.

En tal sentido, el objetivo del presente componente del proyecto “Cambio Organizacional en el Sector Público Paraguayo: El Caso SENAVITAT”, por lo tanto, fue el de construir un Datawarehouse para la Secretaría Nacional de la Vivienda el Hábitat, el cual permitirá alimentar diferentes herramientas para el seguimiento y análisis de la estrategia, como ser tableros de comando (Scorecards), informes analíticos (dashboards), entre otros, potenciando las herramientas existentes.

Dado que construir a partir de la herramienta en utilización es una práctica que puede reducir en gran medida la resistencia y la necesidad de capacitación, un objetivo específico del proyecto, por tanto, puede constituirse como el mejoramiento de esta herramienta, tanto en el proceso operativo de generación de informes, como en la riqueza de información presentada, sirviendo al mismo tiempo de rendición de cuentas con la sociedad civil. Por otro lado, funcionarios designados se encargan de reportar ciertos indicadores a la Secretaría Técnica de Planificación para el tablero de control presidencial, el cual también se utiliza como herramienta de rendición de cuentas públicas. El diseño de Datawarehouse elaborado para SENAVITAT, por tanto, se concentra en un inicio en el monitoreo de los procesos de producción del ente, en términos de efectividad y eficiencia en la construcción de viviendas y hábitats para la población.

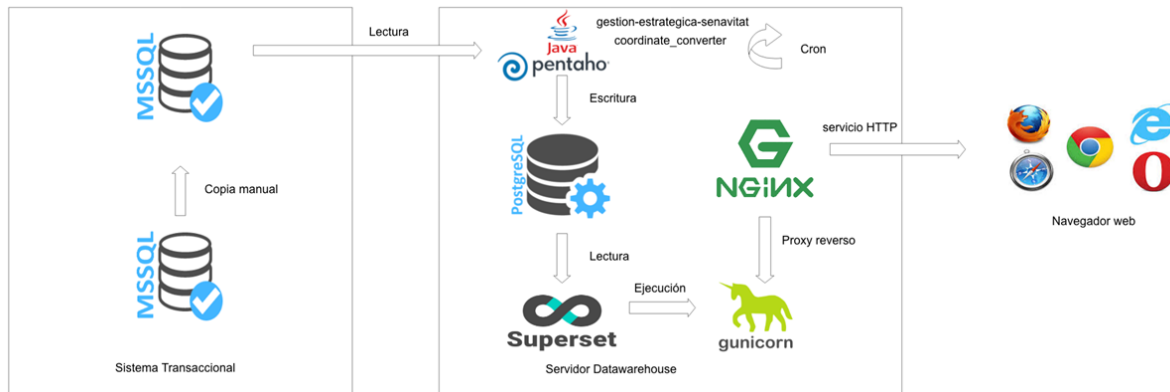
Por otro lado, a pesar de contar con documentos de planificación estratégica y de los resultados de la revisión de la planificación 2015-2018 realizado en un taller organizado por el Instituto Desarrollo, queda claro que la herramienta debe ser lo suficientemente flexible como para adaptarse a nuevos desafíos estratégicos encarados por diferentes administraciones que puedan sucederse en el tiempo. De hecho, unos de los mayores desafíos resaltados por el taller mencionado es la institucionalización del cambio organizacional conseguido por SENAVITAT en la administración 2015-2018, de modo a que nuevas administraciones puedan construir sobre ya elaborado, ahorrando el tiempo y esfuerzo que podría conllevar una reorganización completa para cada cambio de administración.

Finalmente, no podemos ignorar las limitaciones de datos. SENAVITAT cuenta a la fecha con muy pocos de sus procesos efectivamente sistematizados y digitalizados. Si bien se había encarado un proyecto de digitalización del proceso de solicitud, análisis y selección de beneficiarios de soluciones habitacionales, el mismo a la fecha de inicio del proyecto aún no se encontraba implementado. Sin embargo, al centrarnos en esquemas y procedimientos escalables, e incluyendo procedimientos de capacitación para funcionarios de la organización, fue posible empezar con lo existente, enriqueciendo en gran medida las herramientas de seguimiento de la estrategia, al tiempo de dejar sentadas las bases para su ampliación en el futuro.

Arquitectura del Sistema de Integración de Datos

El sistema para la alimentación del tablero o *dashboard* cuantitativo, cuenta con varias componentes o capas. La capa más baja se basa en una copia fiel de las bases de datos de origen, que se administran en el ambiente SQL server de Microsoft®. Estas es consultada por un sistema ETL basado en Hitachi Pentaho PDI community edition®, el cual posee una licencia de software libre y corre sobre una máquina virtual de Java®. Pentaho PDI® se encuentra preparado para integrar informaciones de otras fuentes, como bases de datos de otros sistemas de administración, así como archivos “planos” en diferentes formatos. El resultado del procesamiento del sistema ETL es posteriormente almacenado en un sistema de administración de bases de datos PostgreSQL®, también bajo licencia de software libre. Finalmente se utiliza como herramienta para el diseño y presentación de tableros el Apache Superset®, generador de páginas web basado en el lenguaje de programación Python,

también bajo licencia de software libre. Finalmente, debido a la naturaleza web del aplicativo final, servidores web también son necesario para su despliegue final en navegadores estándar. En este último paso se utilizan los aplicativos Gunicorn© y Nginx©.



Proceso de Instalación

Para instalar todo el paquete a partir del sistema operativo en el servidor habilitado en SENAVITAT para el efecto, es necesario contar con los siguientes requisitos:

- Ubuntu > 16.04
- Git
- Conexión a internet
- Terminal

Complidos estos requisitos, la instalación procede a través de la ejecución de los siguientes comandos:

1. Actualizar el listado de paquetes del sistema operativo

- sudo apt update

2. Instalar Java 8

- sudo add-apt-repository ppa:webupd8team/java
- sudo apt-get update
- sudo apt-get install oracle-java8-installer
- sudo apt-get install oracle-java8-set-default

3. Verificar la variable de entorno JAVA_HOME

- env | grep JAVA_HOME

3.a. Si el resultado del comando anterior es vacío

- nano ~/.bashrc

- Añadir la siguiente línea al final del archivo

- export JAVA_HOME=/usr/lib/jvm/java-8-oracle

4. Si se modificó .bashrc, recargar el ambiente

- source .bashrc

5. Instalación de Pentaho Data Integration

- Descargar la última versión de Pentaho Data Integration (8.1 al momento de escribir esta guía)

- <https://sourceforge.net/projects/pentaho/files/Pentaho%208.1/>

- Copiar al sistema y unzipear

- unzip pdi-ce-8.1.0.0-365.zip

6. Instalación de Postgres

- sudo add-apt-repository "deb http://apt.postgresql.org/pub/repos/apt/ \$(lsb_release -sc)-pgdg main"

- wget --quiet -O - https://www.postgresql.org/media/keys/ACCC4CF8.asc | sudo apt-key add –

- sudo apt-get update

- sudo apt-get install postgresql-9.6

7. Cambiar la contraseña del usuario postgres.

- sudo -u postgres psql

- ALTER USER user_name WITH PASSWORD 'new_password';
rm ~/.psql_history

8. crear una nueva base de datos

- sudo -u postgres createdb senavitat_tutorial

9. Instalar el Proyecto Gestión Estratégica Senavitat y configurar variables de entorno

- cd ~/opt

- git clone <https://gitlab.com/instituto-desarrollo/gestion-estrategica-senavitat.git>

- mkdir -p ~/.kettle

- cp gestion-estrategica-senavitat/kettle.properties.example ~/.kettle/ kettle.properties

```
pathSrcTxtTransformations=~/.opt/gestion-estrategica-senavitat/transformations/tables
```

```
pathTgtTables=~/.opt/gestion-estrategica-senavitat/txt/tables
```

```
pathSrcDimTransformations                               ~/.opt/gestion-estrategica-senavitat/transformations/dimensions
```

10. Configurar la base de datos de origen

```
senDbHostname=201.217.51.160
```

```
senDbPort=1433
```

```
senDbName=senavitat
```

```
senDbUsername=Codium
```

```
senDbPass=codium2018
```

- Configuramos la base de datos de destino

```
dwDbHostname=localhost
```

```
dwDbPort=5432
```

```
dwDbName=senavitat
```

```
dwDbUsername=postgres
```

```
dwDbPass=postgres
```

11. Comprobación de la Instalación del ETL

- ./kitchen.sh -file=~/.opt/gestion-estrategica-senavitat/jobs/masterJob.kjb

12. Si falla el comando anterior restaurar a partir de un backup

- pg_restore --clean -d senavitat -U postgres senavitat_dw_21_09_18

13. Instalar el Proyecto Coordinate Converter (Programa en python que convierte las coordenadas GPS de las bases de datos de origen al sistema utilizado por la aplicación de Dashboards)

- pip install --user pipenv

- cd ~/opt

- git clone <https://gitlab.com/instituto-desarrollo/gestion-estrategica-senavitat.git>

- pipenv install

14. Modificar config.yaml para configurar la URL de la base de datos

- pipenv shell

- TARGET_DB='postgres://postgres:postgres@localhost/senavitat' python
coordinate_converter.py

15. Instalar Apache Superset

- # *Install superset*

 - pip install superset

- # *Create an admin user (you will be prompted to set username, first and last name before setting a password)*

 - fabmanager create-admin --app superset

- # *Initialize the database*

 - superset db upgrade

- # *Load some data to play with*

 - superset load_examples

- # *Create default roles and permissions*

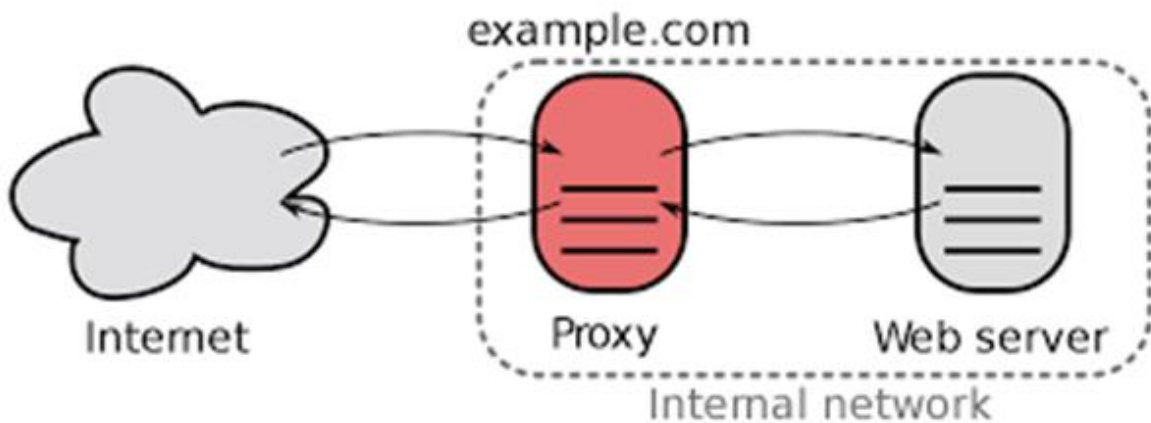
 - superset init

- # *To start a development web server on port 8088, use -p to bind to another port*
superset runserver -d

16. Instalaciones para puesta en producción:

- pip install gunicorn

•`gunicorn \ -w 10 \ -k gevent \ --timeout 120 \ -b 0.0.0.0:8088 \ --limit-request-line 0 \ -limit-request-field_size 0 \ superset:app`



17. Configurar roxy reverso

•`sudo apt install nginx`

17. a. Al cambiar la configuración

•`sudo service nginx restart`

17. b. Para modificar la configuración (crear el archivo si no existe)

•`sudo vim /etc/nginx/conf.d/reporte.senavitat.gov.py.conf`

```
server {
    listen 80;
    server_name reporte.senavitat.gov.py;
    charset utf-8;
    access_log      /var/log/nginx/reporte.senavitat.gov.py.access.log;
    error_log       /var/log/nginx/reporte.senavitat.gov.py.error.log;
    location / {
        ...
    }
}

server {
    location / {
```

```

proxy_pass http://localhost:8088/;

    proxy_set_header    X-Real-IP $remote_addr;

    proxy_set_header    X-Forwarded-For $proxy_add_x_forwarded_for;

    proxy_set_header    X-Forwarded-Proto $scheme;

    proxy_set_header X-Forwarded-Host $http_host;

    add_header Access-Control-Allow-Origin *;

    proxy_set_header    Host $host;

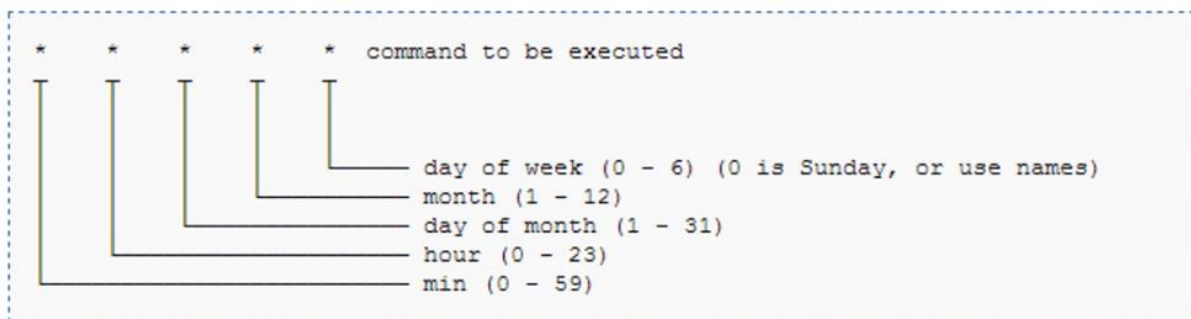
    proxy_buffering off;

    proxy_redirect http://$remote_addr/ http://$remote_addr:1280/;

}
}

```

18. Configurar CRON JObs



18.a. Nuevo job:

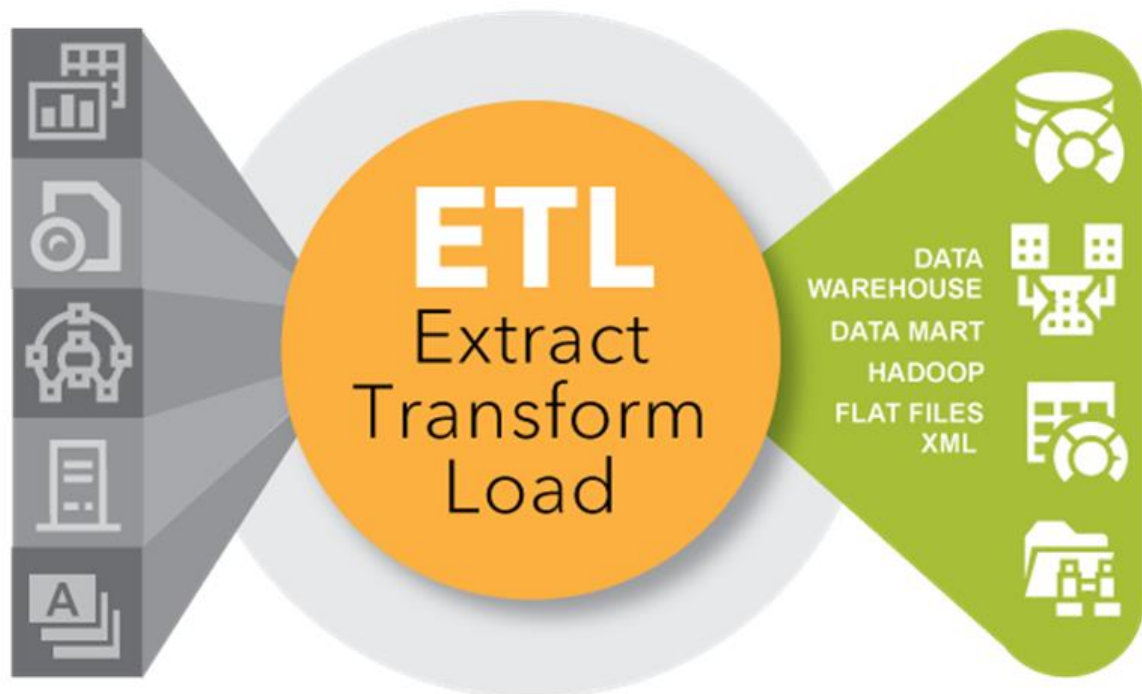
- `crontab -e`

18.b. Listar los Jobs

- `crontab -l`

Proceso de Integración de Datos

Pentaho Data Integration



El proceso de integración de datos, también llamado ETL, utiliza la plataforma pentaho PDI® que permite la obtención de datos de distintas fuentes, su limpieza, corrección, adecuación y transformación, de acuerdo a la lógica de negocios, y finalmente, su publicación en bases de datos orientadas a la consulta, normalmente en datamarts relacionados entre sí por dimensiones comunes, que a su vez, se organizan en un esquema estrella. El esquema estrella consiste en una tabla de hechos central, relacionada con múltiples dimensiones. En la tabla de hechos se utilizan únicamente campos numéricos, campos de código para los registros de hechos, así como campos que contienen las métricas propiamente dichas. Las dimensiones contienen mayores datos que permiten hacer un análisis más detallado de la información.

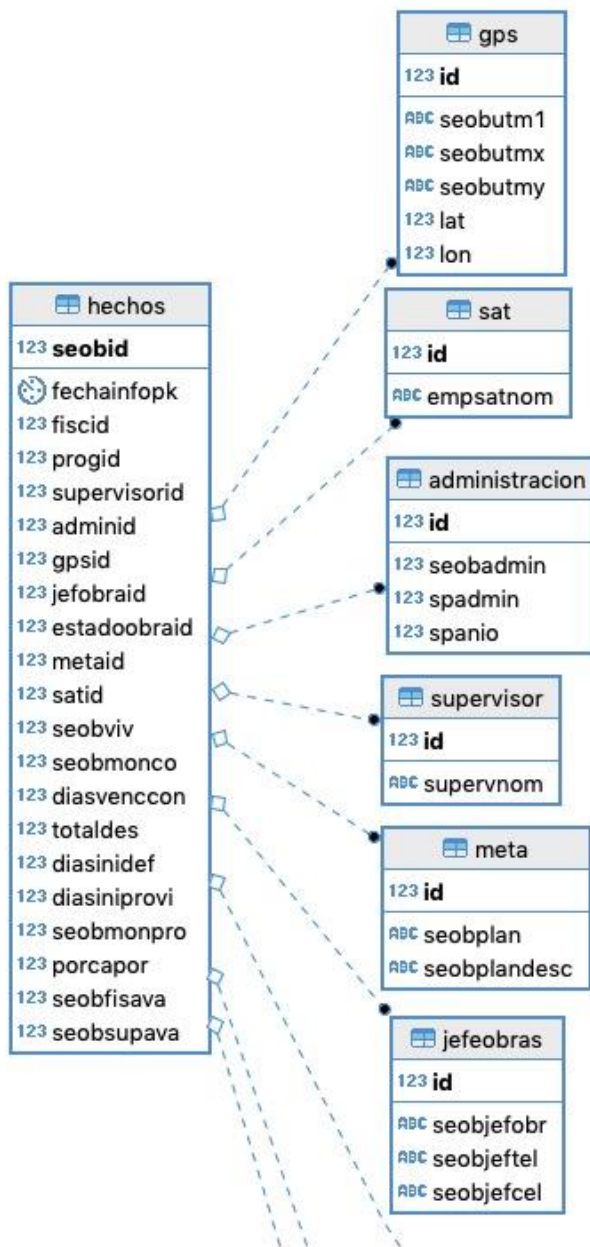
A continuación, se especifica el esquema de datawarehouse instalado en SENAVITAT (hoy MUVH).

Estructura del Datawarehouse SENAVITAT (MUVH)

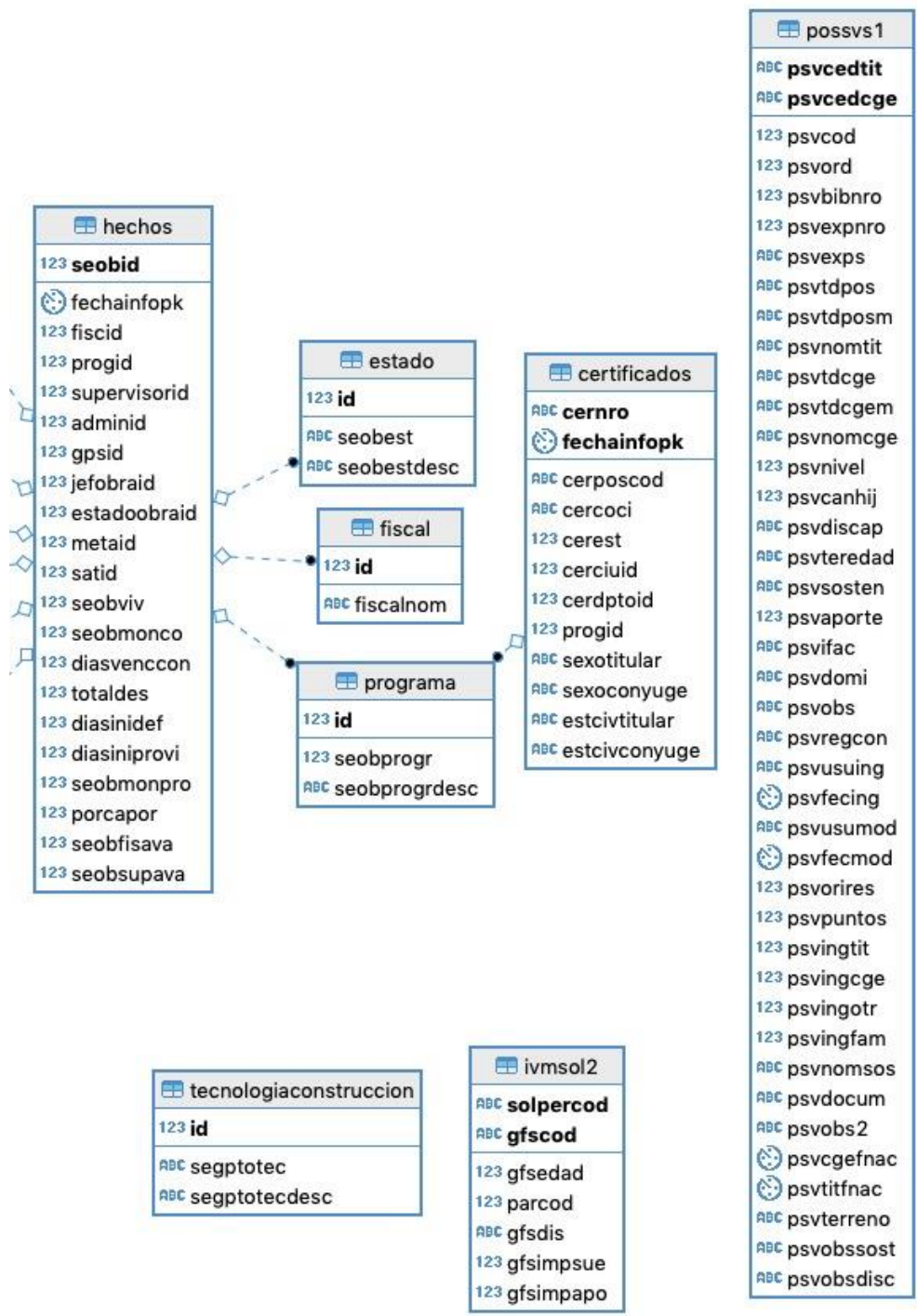
Las mejores prácticas en relación a la estructura de un Datawarehouse orientado a la consulta y análisis de la información para el soporte de toma de decisiones en los procesos de negocio, sugiere la organización de la información de datamarts, los cuales consisten en un esquema datos donde una tabla central, contiene las métricas de negocios y claves sencillas, numéricas, generadas usualmente por el proceso ETL, que apunten a tablas que contengan las dimensiones de análisis, con datos numéricos y alfanuméricos en forma de atributos.

Ejemplo de métricas, en el contexto del caso de SENAVITAT, están dadas por la cantidad de proyectos, la cantidad de viviendas en cada proyecto, el costo monetario de cada proyecto y el costo unitario de cada vivienda. También es posible el cómputo de métricas relacionadas a la eficiencia en términos de tiempos, por ejemplo la cantidad de días transcurridos desde el inicio del proyecto, la cantidad de días de atraso de un proyecto, si es que ya se ha dado la fecha de vencimiento de contrato, o la cantidad de días transcurridos desde una etapa o hito de la obra hasta otra siguiente. Finalmente, otra métrica interesante es el avance porcentual de la obra, definido en términos operativos por documentos internos de SENAVITAT. Debido a que los procesos constructivos son raramente secuenciales (por ejemplo: una vivienda y luego la siguiente) y generalmente se ejecutan en paralelo (por ejemplo: los cimientos, para todas las viviendas) no hay que confundir el avance porcentual de un proyecto con la cantidad de viviendas construidas. Por dar un ejemplo, un avance del 10% podría indicar que los materiales, equipos y mano de obra ya se encuentran en el sitio, el 20% podría indicar la culminación de cimientos y obras complementarias, etc.; lo cual de vuelta se aclara es de modo a graficar el concepto, ya que los criterios de esta evaluación dependen de procesos internos del ente, que pueden estar definidos para cada programa o incluso para cada proyecto.

Datamart obras SENAVITAT (MUVH)



En tal sentido se tomara la decisión, de manera conjunta con el equipo de contrapartida en SENAVITAT, conformado por la Dirección General de Tecnología y la Dirección General de Planificación Habitacional, de empezar con un diseño mínimo, escalable, tomando como punto de partida el centro de la estrategia que tiene que ver con el control de las obras. El datamart de seguimiento de obras, por lo tanto, fue diseñado como punto de partida.



El mismo coloca las métricas en dos tablas de hechos, la primera denominada hechos, contiene el estado de las métricas a la fecha del último informe, la segunda, hechosnap contiene las métricas de acuerdo a valores históricos de informes anteriores, con lo cual es posible darles un seguimiento de su evolución a lo largo del tiempo.

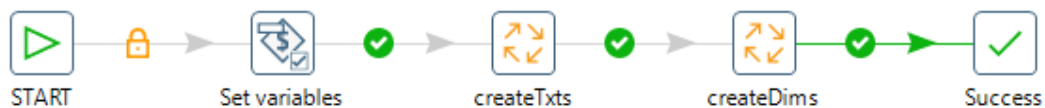
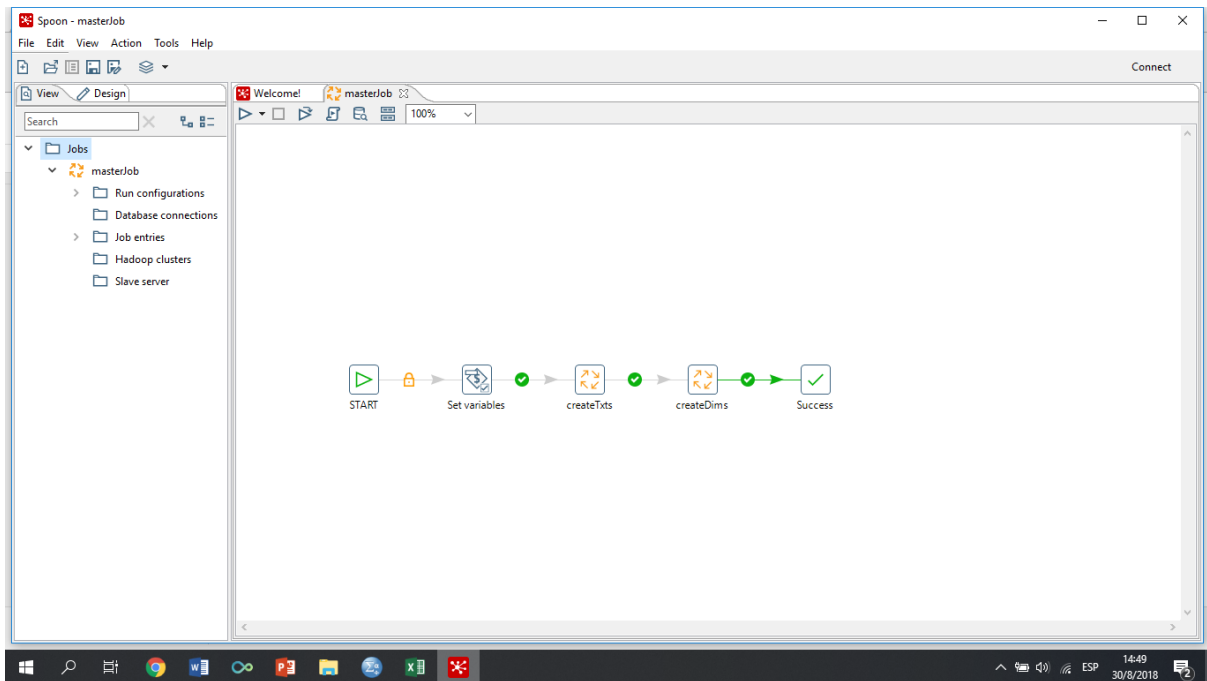
Datawarehouse SENAIVAT					
Tablas	HECHOS, HECHOSSNAP (HISTORICO)				
Campo	Tipo Dato	Funcion	Apunta a:	Descripción	Origen (tabla SEGOBRA salvo otra especific.)
seobid	INTEGER	Clave Primaria		Codigo de la Obra	SEGOBRA.SEOBID
fechainfopk	DATE	Clave Primaria		Fecha de Proceso ETL	Fecha de sistema
fiscid	INTEGER	Clave Dimension	fiscal.id	Codigo del Fiscal	FiscalCod
progid	INTEGER	Clave Dimension	programa.id	Codigo del Programa	SEOBProgr
supervisorid	INTEGER	Clave Dimension	supervisor.id	Codigo del Supervisor	SupervCod
adminid	INTEGER	Clave Dimension	administracion.id	Codigo de tipo de administracion	SEOBAdmin
gpsid	INTEGER	Clave Dimension	gps.id	Codigo de Ubicación GPS	SEOBtm1, SEOBtmX, SEOBtmY
jefobraid	INTEGER	Clave Dimension	jefeobras.id	Codigo Jefe de obra	SEOBJefObr
estadoobraid	INTEGER	Clave Dimension	estado.id	Codigo de Estado de la obra	SEOBEst
metaid	INTEGER	Clave Dimension	meta.id	Codigo de Meta de Planificación	SEOBPlan
satid	INTEGER	Clave Dimension	sat.id	Codigo de SAT/Empresa	EmpSatCod
seobviv	INTEGER	Metrica		Cantidad de Viviendas	segobra.seobviv
seobmonco	BIGINT	Metrica		Monto total del Contrato	segobra.SEOBMonCo
diasvenccon	INTEGER	Metrica		Dias desde el venc. Del contrato	segobra.SEOBPlaEO-FechaProceso
totaldes	BIGINT	Metrica		Total desembolsado	segobra.SEOBMonAn+segobra.SEOBDetS
diasinidef	INTEGER	Metrica		Dias Inicio hasta recep. definitiva	seobgra.SEOBReDe-segobra.SEOBAct
diasiniprovi	INTEGER	Metrica		Dias inicio hasta recep. Provisoria	SEOBRePr-segobra.SEOBAct
seobmonpro	BIGINT	Metrica		Monto Total Proyecto FONAVIS	segobra.SEOBMonPro
porcapor	DOUBLE	Metrica		Aporte % beneficiario	segobra.SEOBApoBen/SEOBMonPro
seobfisava	DOUBLE	Metrica		Avance % S/ Fiscal	segobra.SEOBFisAva
seobsupava	DOUBLE	Metrica		Avance % S/ Supervisor	segobra.SEOBSupAva

Las tablas de hechos y hechossnap se asocian con las tablas que contienen las dimensiones a través de las claves foráneas o claves de dimensión, bajo la lógica general hechos.dimensionid=dimension.id. Por ejemplo, para rescatar un atributo de la dimensión programa, la sentencia join deberá incluir el comando “on progid=programa.id”.

El cuadro superior, lista el conjunto de dimensiones en el datamart de obras, tanto en la versión actual “hechos” como la versión histórica “hechossnap”.

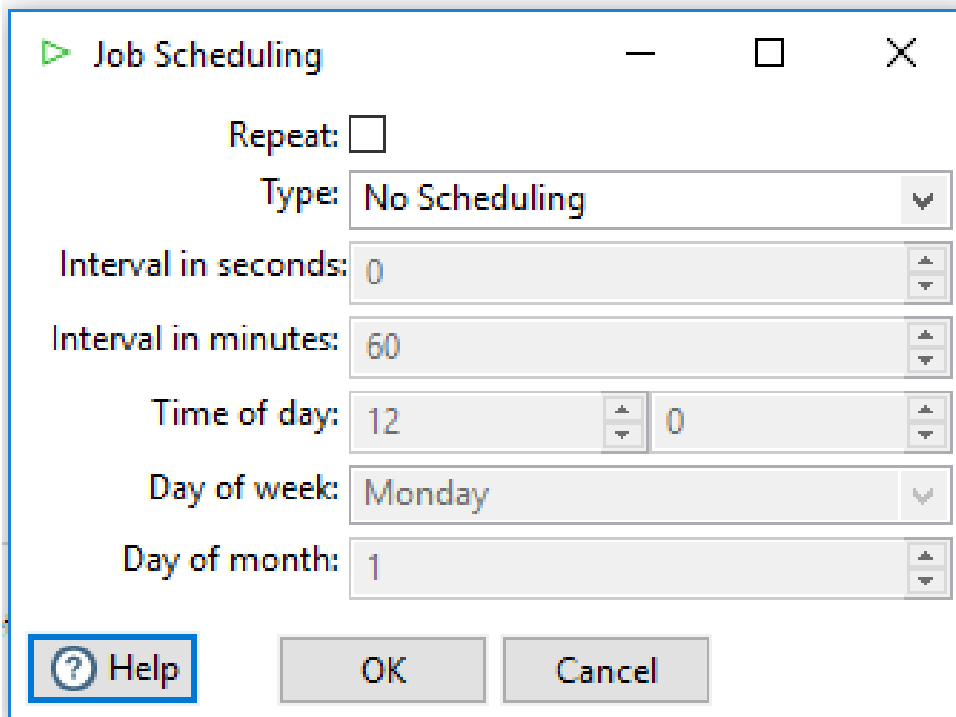
Proceso ETL

El proceso de ETL en general se encuentra codificado en el archivo *masterJob.kjb* el contiene las instrucciones en un formato xml interpretable por tanto por la aplicacion de diseño de PDI, denominada Spoon, así como las aplicaciones de producción como Pan. Al abrir masterJob en Spoon se puede observar la siguiente pantalla:



La pestaña activa presenta el flujo de datos diseñado para el ETL, mientras que en el panel izquierdo se presentan diferentes herramientas para modificar dicho flujo. El proceso masterJob contiene la siguiente secuencia de pasos:

1. Start: Es un paso standar y requerido para todos los jobs y permite definir la forma en que dicho proceso se va a correr. Al hacer click en start aparece una ventana que permite calendarizar el proceso, si es que se va a repetir, de lo contrario el proceso se corre a demanda del usuario en el Spoon. Otras herramientas de PDI permiten delegar a aplicaciones del sistema operativo el corrido de los procesos, por ejemplo el Pan, pasando los procesos ETL a ser parte del conjunto de procesos corridos en batch por el sistema.



2. Set Variables:

Este paso permite definir variables de modo a estandarizar los procesos, por ejemplo, los directorios de los archivos de resultados intermedios, las cadenas de conexión a diferentes sistemas de administración de base de datos, la información necesaria para procesos de consulta y alimentación de servicios web, entre otros.

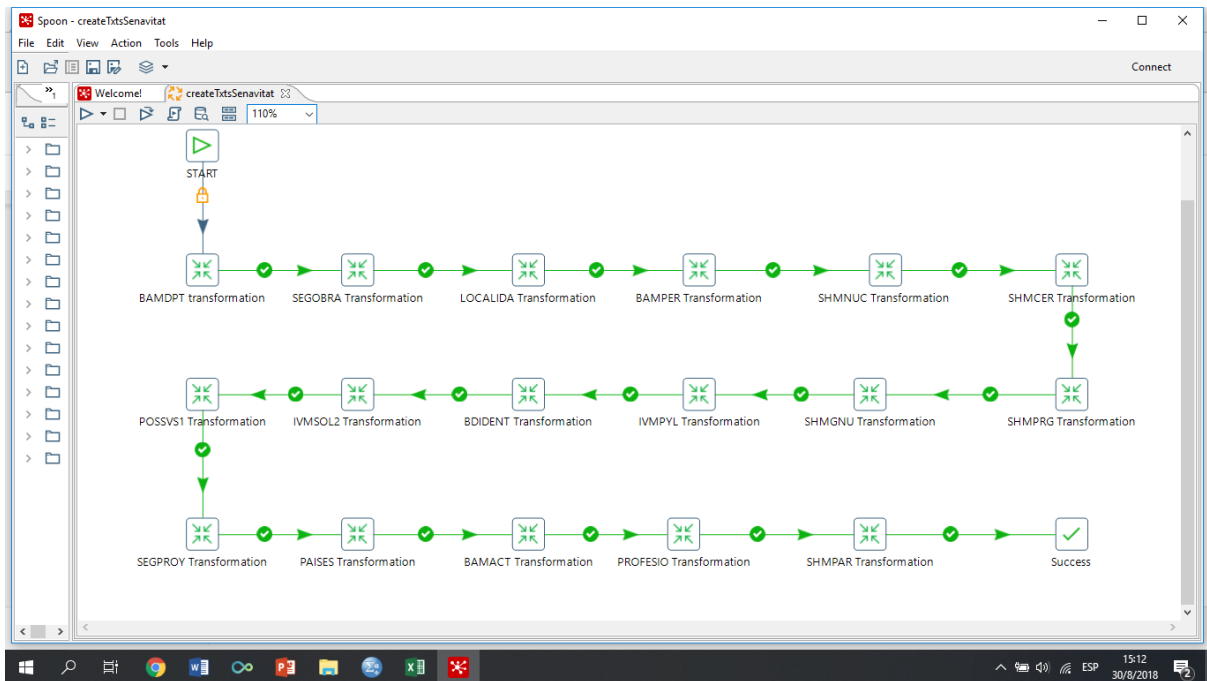
3. El siguiente paso ejecuta el archivo job **createTxts.kjb**, que se encarga del proceso de extracción propiamente dicha y que será analizado en mayor detalle posteriormente. En este apartado basta decir que se conecta a las bases de datos producción de SENAVITAT y captura una copia actualizada de las tablas que el sistema va a utilizar para producir la información integrada. Dichas copias son guardadas como archivos planos en un repositorio, de modo a no requerir de vuelta una conexión a la base de datos si parte del proceso falla o se trunca, además de dejar registros intermedios para control.

4. El siguiente paso ejecuta el archivo job **createDims.kjb** que se encarga de depurar, conformar y cargar los datos en un esquema integrado orientado a la consulta, del cual posteriormente se alimentan las aplicaciones cliente orientadas al seguimiento y análisis de la estrategia.

5. El último paso, **Success**, simplemente verifica que los pasos anteriores se han corrido sin inconvenientes.

Extracción

El archivo **createTxtsSenavitat.kjb** se encarga de conectarse a las bases de datos de SENAVITAT de modo a obtener las tablas necesarias para el proceso de integración de datos y guardarlas en el repositorio para su posterior procesamiento.

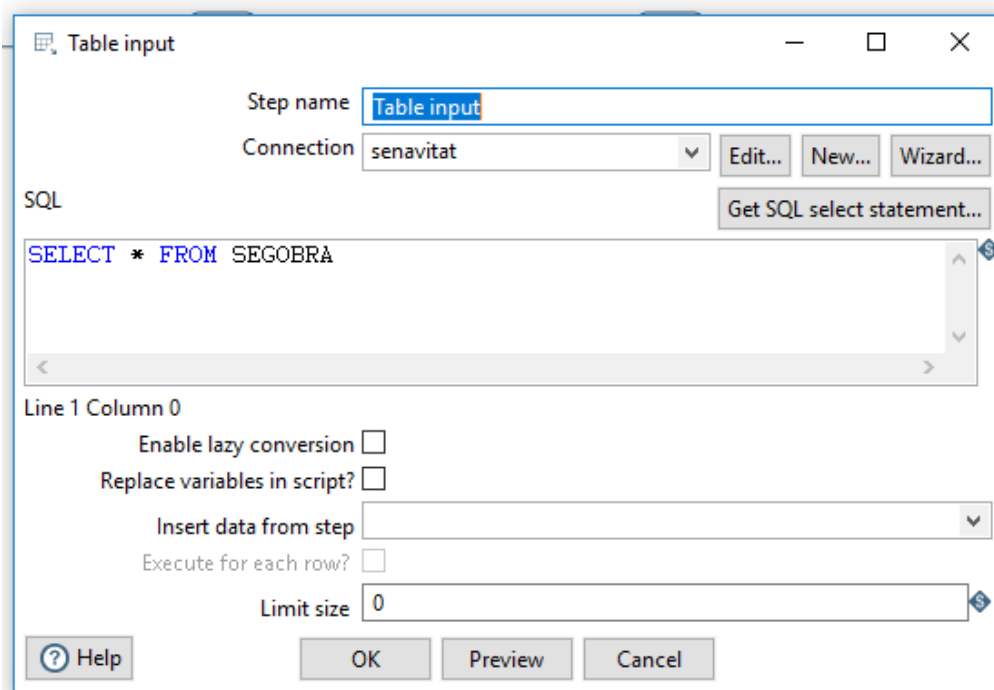


La estructura del proceso es bastante simple ya que se encuentra estructurada en diferentes niveles de abstracción. A nivel de job su único trabajo es llamar secuencialmente a diferentes archivos de transformación (.ktr), denominados según la tabla a la cual apuntan, de modo que estas transformaciones realicen la extracción correspondiente. Esta forma de trabajar, a pesar de ser algo lenta (ya que la aplicación también permite hacerlo en paralelo), facilitará la identificación de cualquier problema, a medida que transcurra, en el paso correspondiente, además de permitir la incorporación de alertas al administrador en una iteración siguiente del diseño.

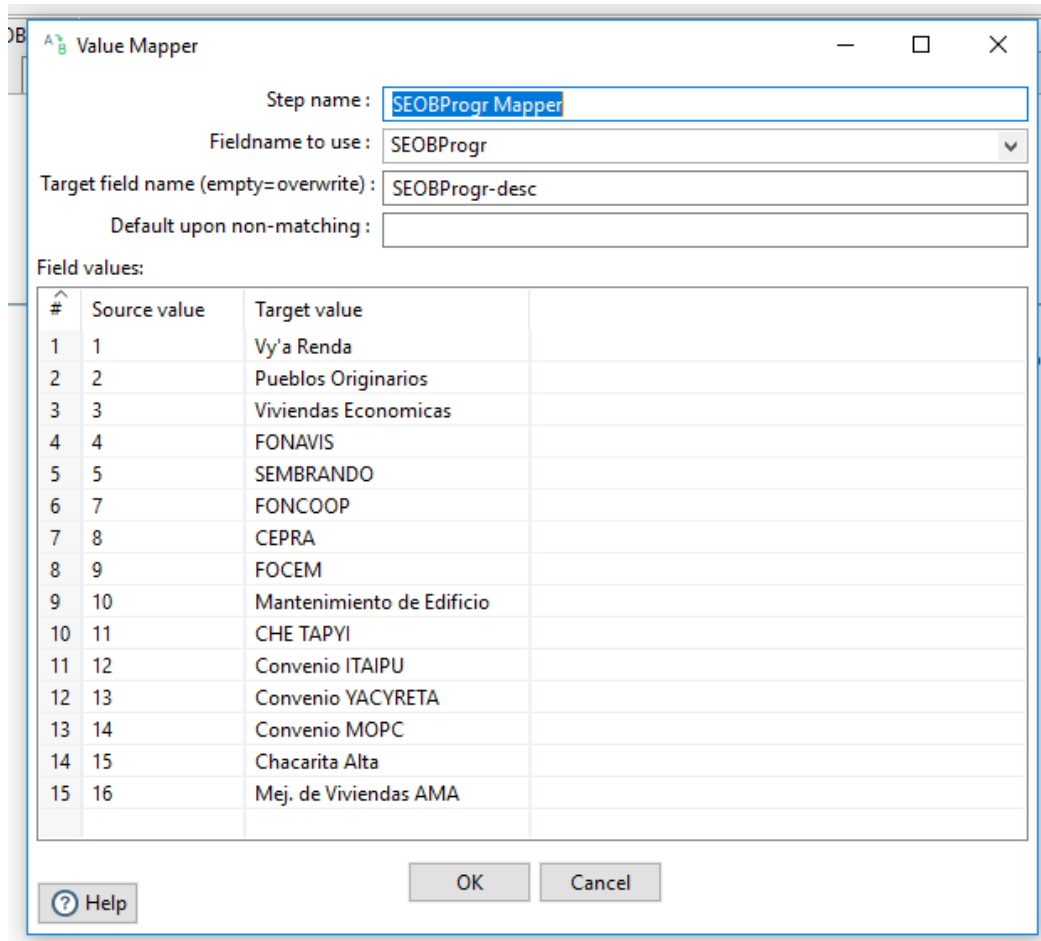


A modo de ejemplo vemos la extracción de la tabla SEGOBRA, que consiste de una serie de pasos:

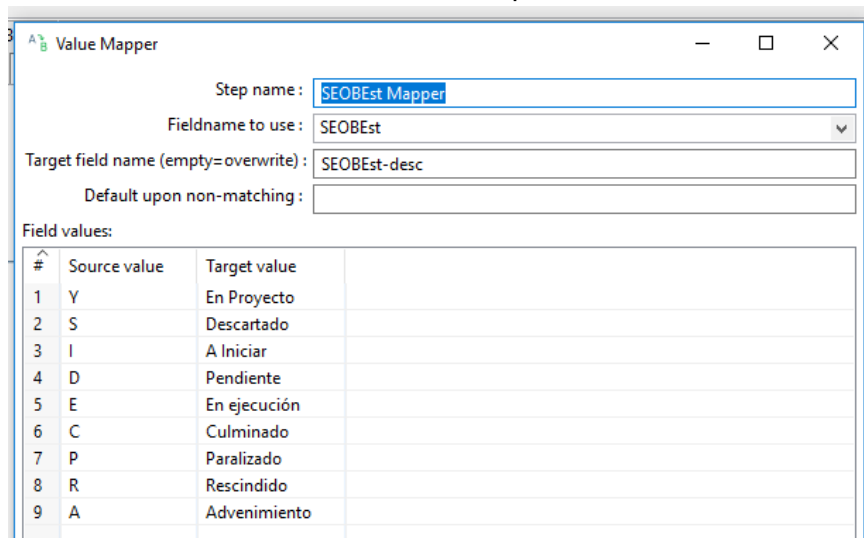
1. Table Input: Este paso, utilizando una conexión llamada *senvitat* que apunta a la base de datos de producción de dicho ente. El paso ejecuta una sencilla sentencia SQL para obtener el conjunto completo de registros de dicha tabla, con todos sus campos, sin mucha más complejidad.



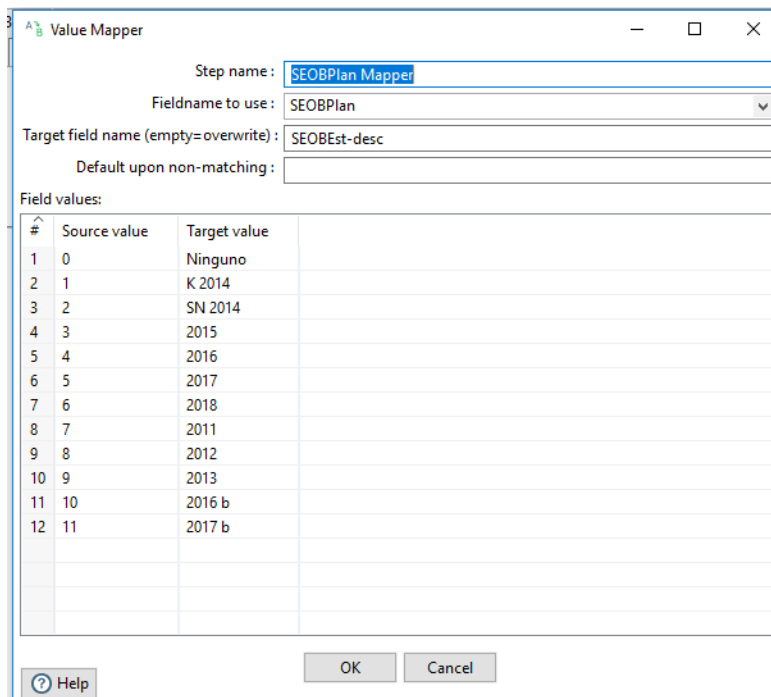
2. Value Mapper: El siguiente paso, de tipo Value Mapper, asigna un campo de tipo cadena a cada ocurrencia del campo numérico entero SEOBprog, en la tabla de origen, el cual contiene los códigos de los programas a los cuales pertenece cada obra, dado que al momento de la elaboración del proceso de integración de datos, no se contaba con una tabla en las bases de datos de SENAVITAT que contenga esta información. En términos de su actualización, dado un nuevo programa, esto implica adicionar un nuevo registro al grid del paso Value Mapper, o si en el futuro se elabora una tabla con dicha información, la sustitución de este paso por otro paso de extracción para esta última.



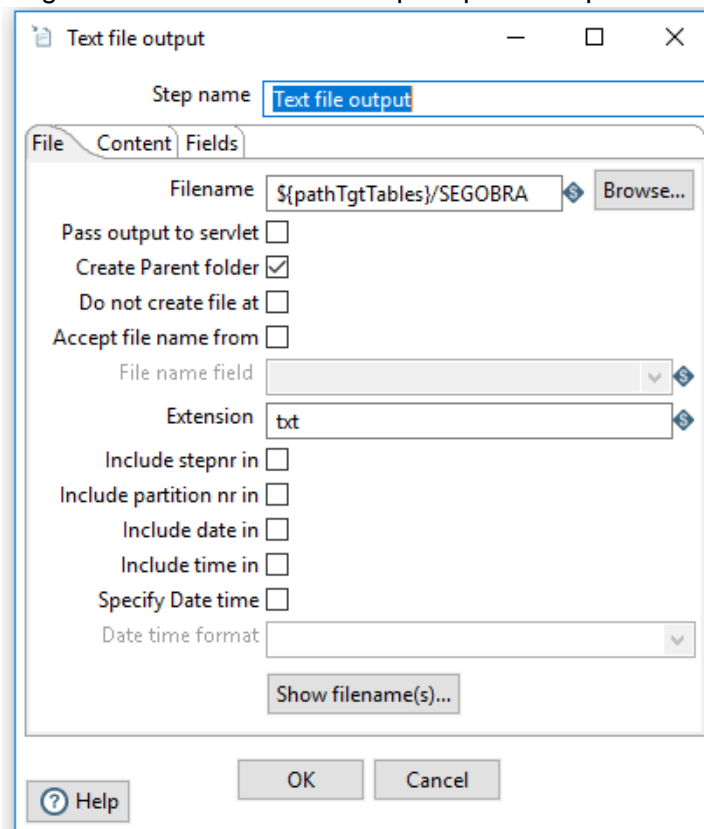
- Value Mapper: El siguiente paso funciona de manera muy similar al anterior, solo que en este caso mapea el campo SEOBEst, o estado de la obra, el cual tampoco contaba con una tabla definida al momento del presente diseño.



- Value Mapper: De nuevo la misma lógica para el campo SEOBplan, según se ve en el snapshot.



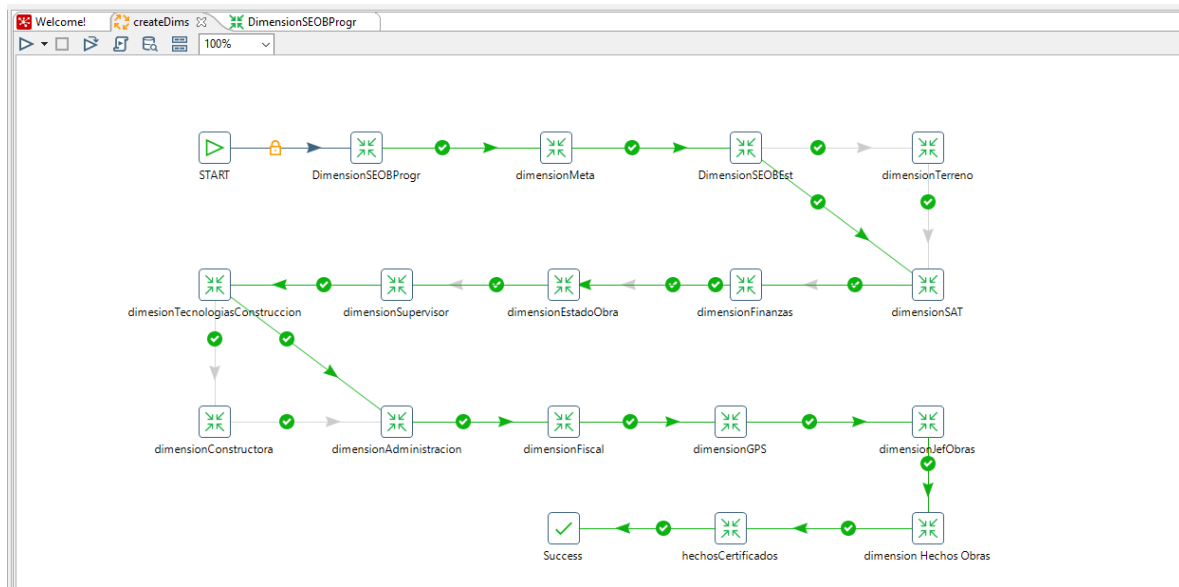
5. Text Input: Finalmente, la transformacion que apunta a la tabla de seguimiento de obras (SEGOBRA) en las bases de datos de SENAVITAT, luego de las depuraciones mencionadas, se guarda un archivo de texto para posterior procesamiento.



Las demás transformaciones que se refieren a otras tablas de SENAVITAT, proceden de manera muy similar.

Transformación y Carga de Dimensiones

El archivo createDims.kjb contiene el job destinado a crear las dimensiones y las métricas utilizadas en el datamart relacionado a los proyectos y obras de SENAVITAT. A partir del paso START, se observa con claridad que cada transformación se corresponde con la elaboración de una dimensión en particular, hasta en las dos últimas se crea la tabla de hechos, la cual se conecta a las dimensiones previamente elaboradas para conformar un esquema orientado a la consulta, comúnmente denominado esquema estrella.

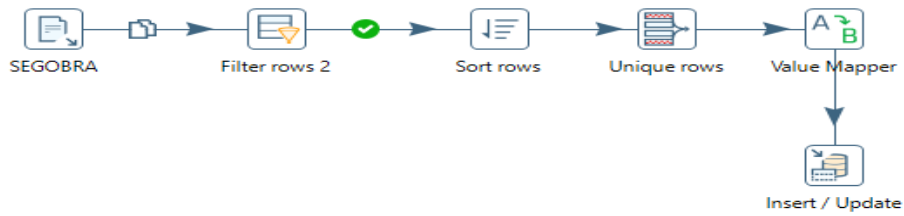


Pasamos ahora a examinar la creación de algunos de los componentes principales. Los demás pueden inferirse de estos ejemplos ya que se sigue la misma lógica.

Dimension Programa

La creación de la mayoría de las dimensiones, por parte del proceso ETL, consiste en normalizar distintas tablas de seguimiento de la base transaccional de SENAVITAT, debido a la estructura no normalizada encontrada en los sistemas de la entidad. Este es un enfoque bastante contrastante con la mayoría de las implementaciones de un Datawarehouse, debido a que lo usual es encontrar sistemas sumamente normalizados, de modo a optimizar el desempeño transaccional, minimizar recursos en disco y memoria, etc. De igual manera, un mínimo de normalización en el Datawarehouse es deseable y recomendado por las mejores prácticas.

De esta manera, a partir de la tabla seguimiento de obras, SEGOBRA, se normaliza la dimensión Programa.



Para ello se observa que el primer paso, tipo TEXT INPUT, consiste en recuperar del repositorio de archivos temporales la tabla de seguimiento de obras o SEGOBRA de SENAVITAT, de la cual se toma exclusivamente el campo de número de proyecto, o SEOBProg. El segundo paso FILTER ROWS consiste en el filtrado de casos nulos.

Filter rows

Step name:

Send 'true' data to step:

Send 'false' data to step:

The condition:

Buttons: Help, OK, Cancel

Sort rows

Step name:

Sort directory:

TMP-file prefix:

Sort size (rows in memory):

Free memory threshold (in %):

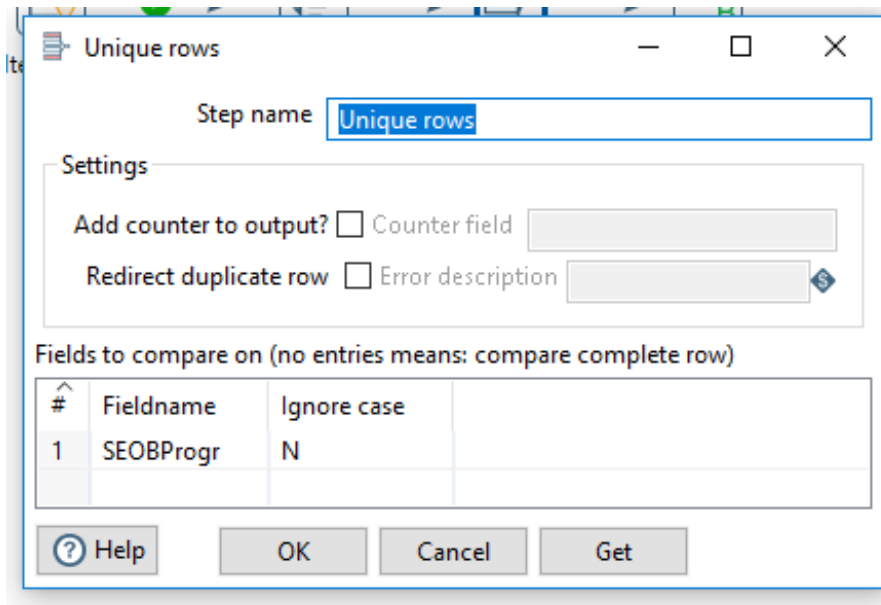
Compress TMP Files?

Only pass unique rows? (verifies keys only)

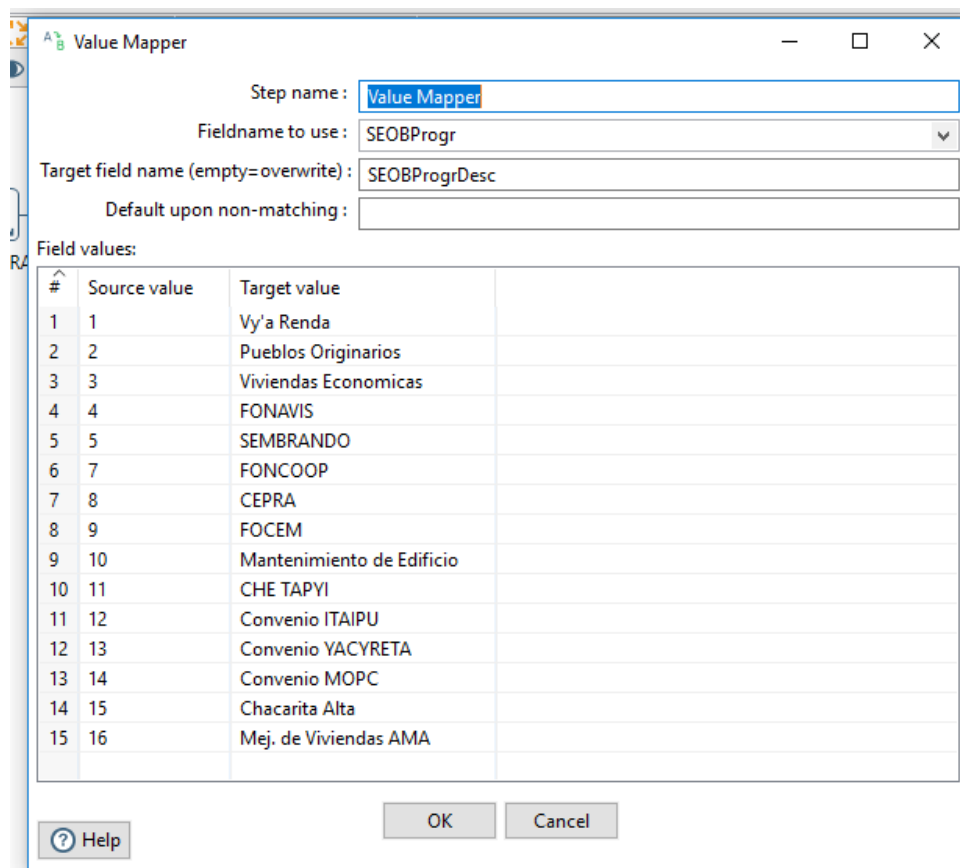
Fields:

#	Fieldname	Ascending	Case sensitive compare?	Sort based on current locale?	Collator Strength	Presorted?
1	SEOBProg	Y	N	N	0	N

Buttons: Help, OK, Cancel, Get Fields



Mientras que posteriormente, se ordenan los datos de forma ascendente en el paso SORT ROWS, para luego de-duplicar las ocurrencias en el paso UNIQUE ROWS. A continuación en el paso VALUE MAPPER, se asigna a cada instancia de SEOBProgr la cadena que describe el programa al cual corresponde.



Finalmente, una vez asignados los nombres de proyectos, en forma de cadena o STRING, a cada uno de los códigos de programas, dato que no figura en las tablas de SENAVITAT sino

que se maneja a nivel de aplicativos, se colocan en la tabla creada en el esquema relacional mediante un paso INSERT UPDATE.

Step name:

Connection:

Target schema:

Target table:

Commit size:

Don't perform any updates:

The key(s) to look up the value(s):

#	Table field	Comparator	Stream field1	Stream field2
1	SEOBProgr	=	SEOBProgr	

Update fields:

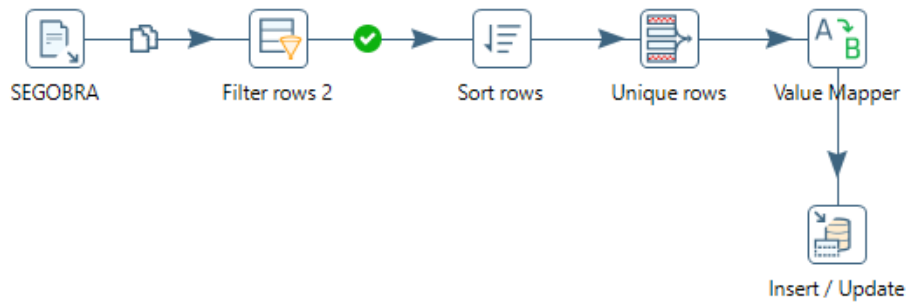
#	Table field	Stream field	Update
1	SEOBProgr	SEOBProgr	Y
2	SEOBProgrDesc	SEOBProgrDesc	Y

El esquema relacional funciona se apunta en la conexión DW, y el nombre de la tabla se define como simplemente PROGRAMA, que pasa a ser la dimensión.

Nótese que con este procedimiento es posible detectar la posibilidad de programas nuevas, no asignados a cadenas o STRINGS, de tal manera a facilitar el mantenimiento del sistema.

Dimension Meta

De manera muy similar se procede a la creación de la siguiente dimensión. La dimensión meta se corresponde con un periodo de tiempo creado por SENAVITAT asignado como metas de programas y proyectos. La información sobre esta dimensión se obtiene mediante el siguiente flujo: Por un lado se sube a memoria la información de la tabla SEGOBRA a partir de un paso text input.



Posteriormente se filtra por el campo SEOBPlan, los valores nulos e inválidos.

Filter rows

Step name: Filter rows 2

Send 'true' data to step: Sort rows

Send 'false' data to step:

The condition:

SEOBPlan IS NOT NULL

AND NOT

SEOBPlan REGEXP [\s+]

Help OK Cancel

Sort rows

Step name: Sort rows

Sort directory: %%java.io.tmpdir%%

TMP-file prefix: out

Sort size (rows in memory): 1000000

Free memory threshold (in %):

Compress TMP Files?

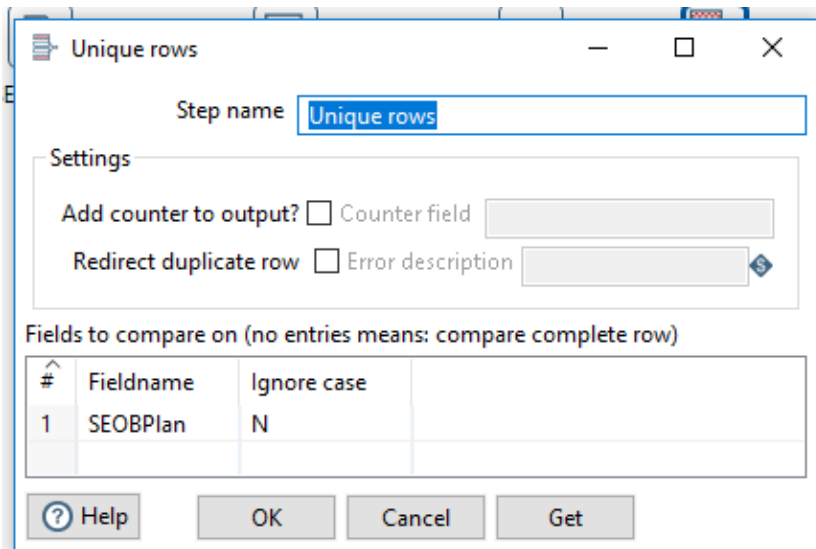
Only pass unique rows? (verifies keys only)

Fields:

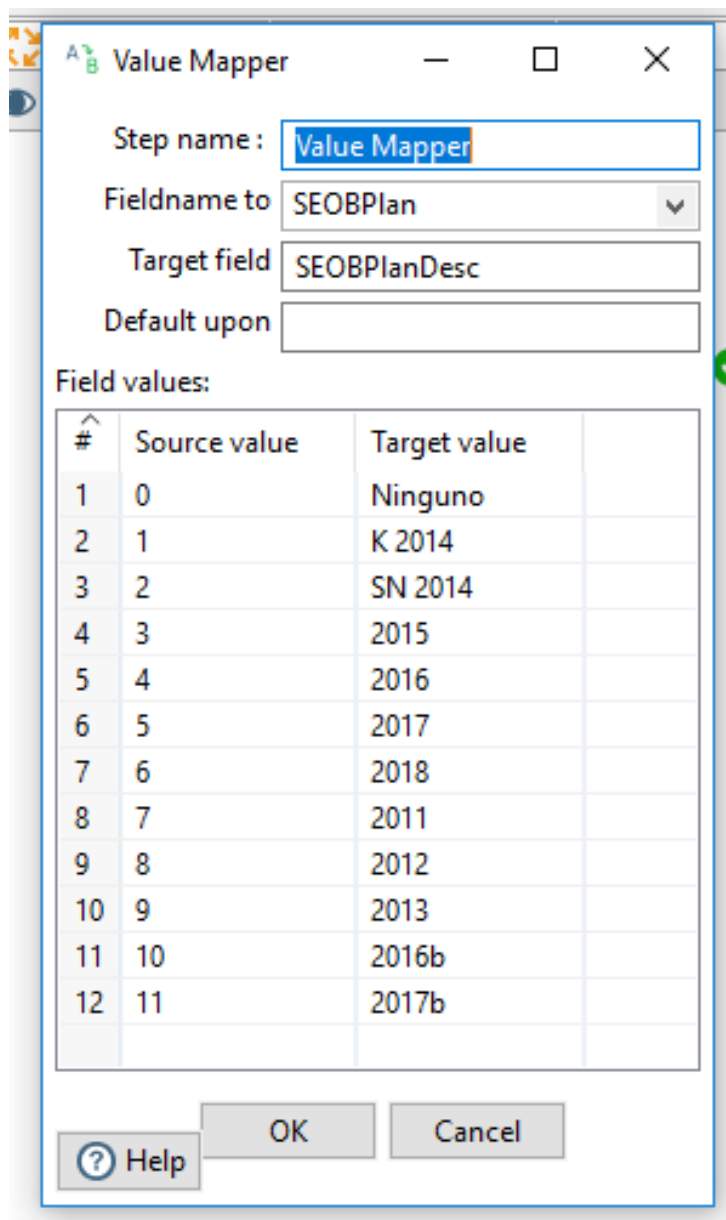
#	Fieldname	Ascending	Case sensitive compare?	Sort based on current locale?	Collator Strength	Presorted?
1	SEOBPlan	Y	N	N	0	N

Help OK Cancel Get Fields

Luego se ordena el conjunto de filas por el campo SEOBPlan en forma ascendente.



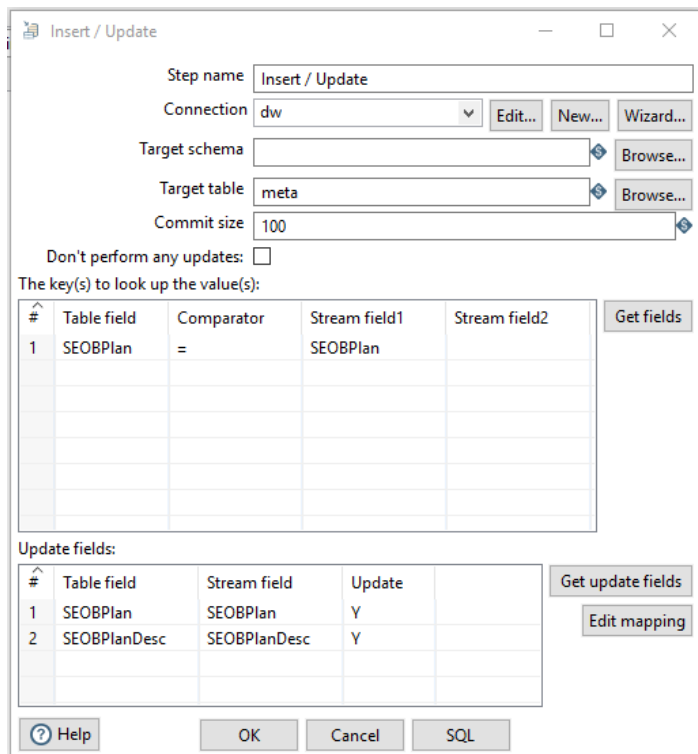
A continuación se hallan los valores únicos para el campo mencionado, de modo a normalizar la información.



Finalmente, a los valores normalizados, únicos y ordenados el campo, se le asignan etiquetas descriptivas mediante un paso de tipo ValueMapper.

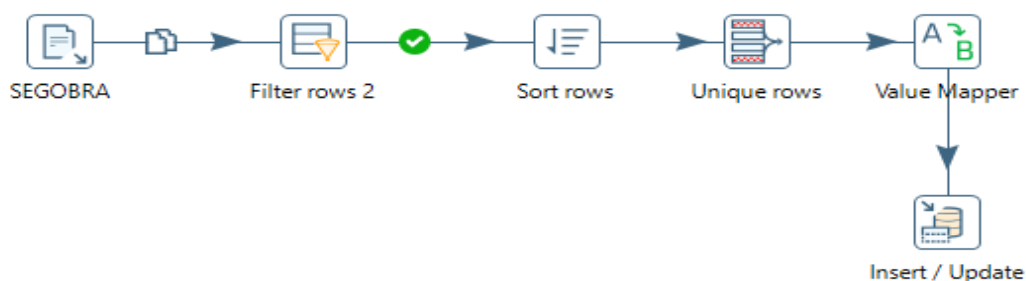
De esta manera, en el momento que se asigne un nuevo plan, el administrador del sistema podrá visualizar rápidamente la ausencia de la etiqueta correspondiente y proceder a colocarla.

Como paso final, se procede a cargar la dimensión en el esquema de DW.



Dimensión Estado

Otra dimensión fundamental para el análisis de la estrategia en SENAVITAT, resulta ser el estado de la obra, es decir se encuentra aún planificada, en ejecución, culminada, paralizada, etc. Tal dimensión se obtiene mediante la normalización del campo **Seobest** de la tabla segobra, de manera muy similar a los otros procedimientos.



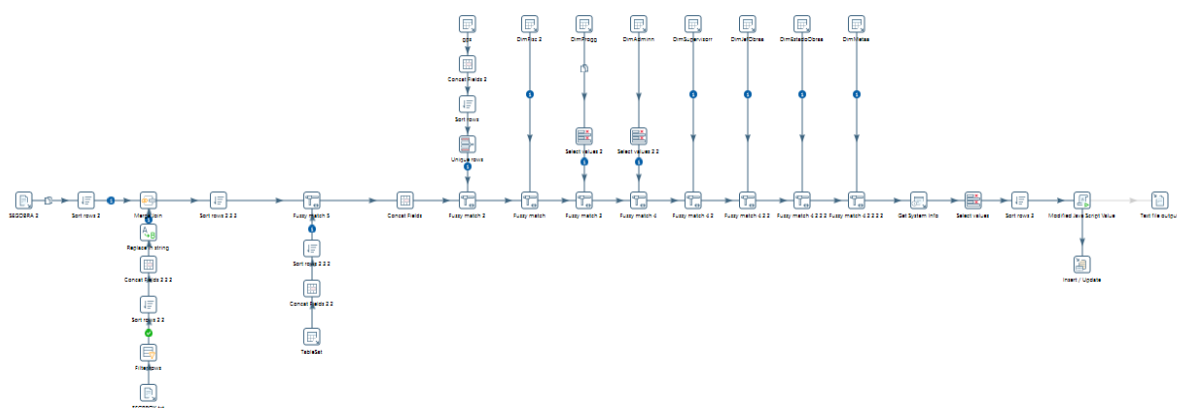
Otras dimensiones, cuyas fuentes de datos, pueden visualizarse la tabla que resume el esquema del de ETL se obtienen de manera muy similar, mediante procesos orientados a la normalización de la información, deduplicación, limpieza y agregación de etiquetas alfanumérica con información institucional.

Tabla de Hechos

El proceso de construcción de la construcción del datamart para seguimiento de obras consiste en relacionar las tablas de hechos con las diferentes dimensiones. Dicho procesos se encuentra en la transformación denominada dimPipeline.ktr. Básicamente la lógica es, a partir de la tabla que contiene los hechos o métricas, en este caso la tabla de seguimiento de obras, o SEGOBRA, se asigna una clave externa para cada dimensión y se las vincula. Para ello se utilizan reglas de lógica exacta (o difusa).

Para mejor entendimiento consideremos este ejemplo: La dimensión programa contiene información sobre los diferentes programas del ente (FONAVIS, Sembrando Oportunidades, Che Tapy, etc.). En la tabla SEGOBRA existe un campo se vincula con el programa específico, pero no existía una tabla de “detalle de programa”, la cual se crea a nivel de ETL con las informaciones obtenidas del relevamiento. El proceso dimPipeline, en tal sentido, relaciona línea por línea los valores de código de programa en SEGOBRA con los valores de código de programa en la tabla-dimensión programa. De tal manera, si corremos una sentencia de consulta en código sql: “SELECT * FROM HECHOS JOIN PROGRAMA ON HECHOS.PROGID=PROGRAMA.ID”, obtendremos la información, a nivel de proyectos, conjuntamente con los detalles del programa.

La misma lógica se aplica para las demás dimensiones construidas con el ETL a partir de datos estructurados en el sistema de SENAVITAT, sometido a normalización y complementado con información no sistematizada. Las dimensiones que no necesitaron de esta transformación, debido a que ya contenían tablas detalle, se vinculan de forma directa.



Hechos Históricos

Al final del proceso dimPipeline.ktr, se generan dos tablas de resultado: una es la tabla hechos, que contiene las métricas a la fecha del informe y que permite tener información del estado actual de la implementación de la estrategia. La otra es la tabla hechossnap que contiene la información acumulada, y permite realizar un seguimiento histórico.

Tablero de Control Cuantitativo

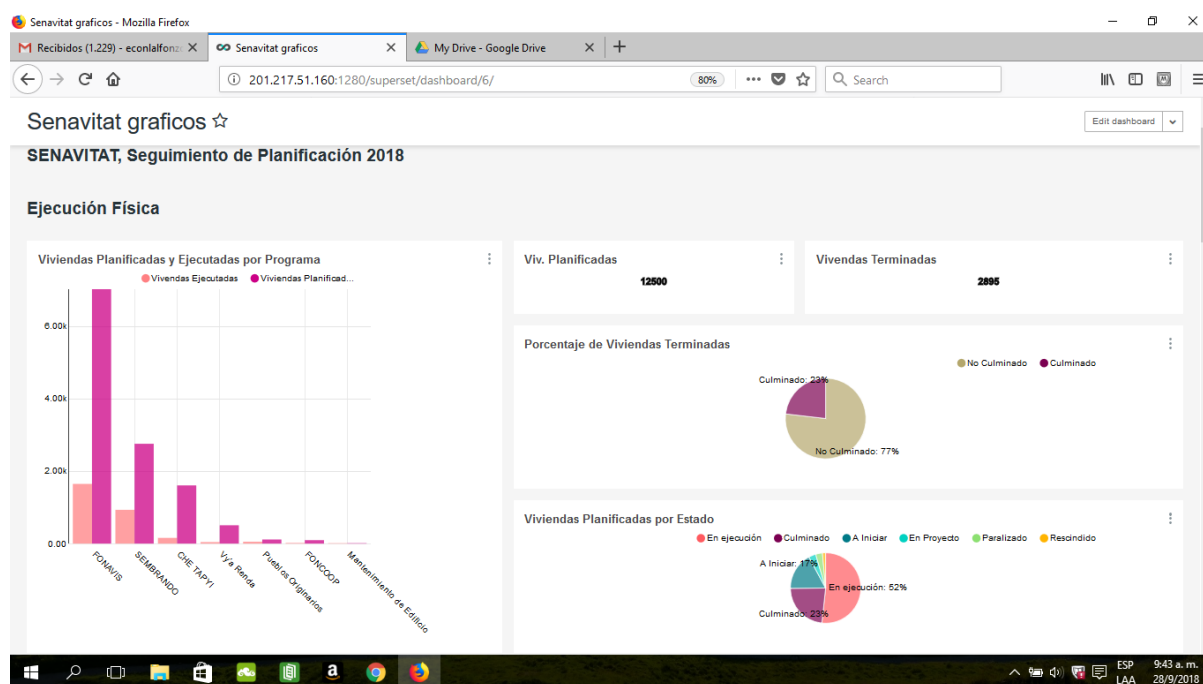
El tablero de Control para SENAVITAT (Hoy MUVH) fue concebido originalmente para el seguimiento integral de la estrategia, incluyendo todos los procesos y perspectivas, sin embargo, debido a limitaciones de sistemas y a prioridades de las autoridades del ente en el momento de llevar a cabo el convenio de colaboración, hemos acotado la primera versión al monitoreo de los procesos de producción, dentro de las perspectivas beneficiarios y procesos internos de la estrategia.

De ninguna manera esta versión debe considerarse como la definitiva, sino más bien como un paso inicial hacia la incorporación de los demás aspectos estratégicos en el monitoreo en el futuro. Tarea que deberá tener en cuenta, además, los diferentes cambios a las estructuras de los sistemas transaccionales que pueden darse como resultado del proceso de mejora continua.

De todas maneras, iniciamos este primer esfuerzo con una serie de elementos informativos, agrupados en varias secciones:

Ejecución Física

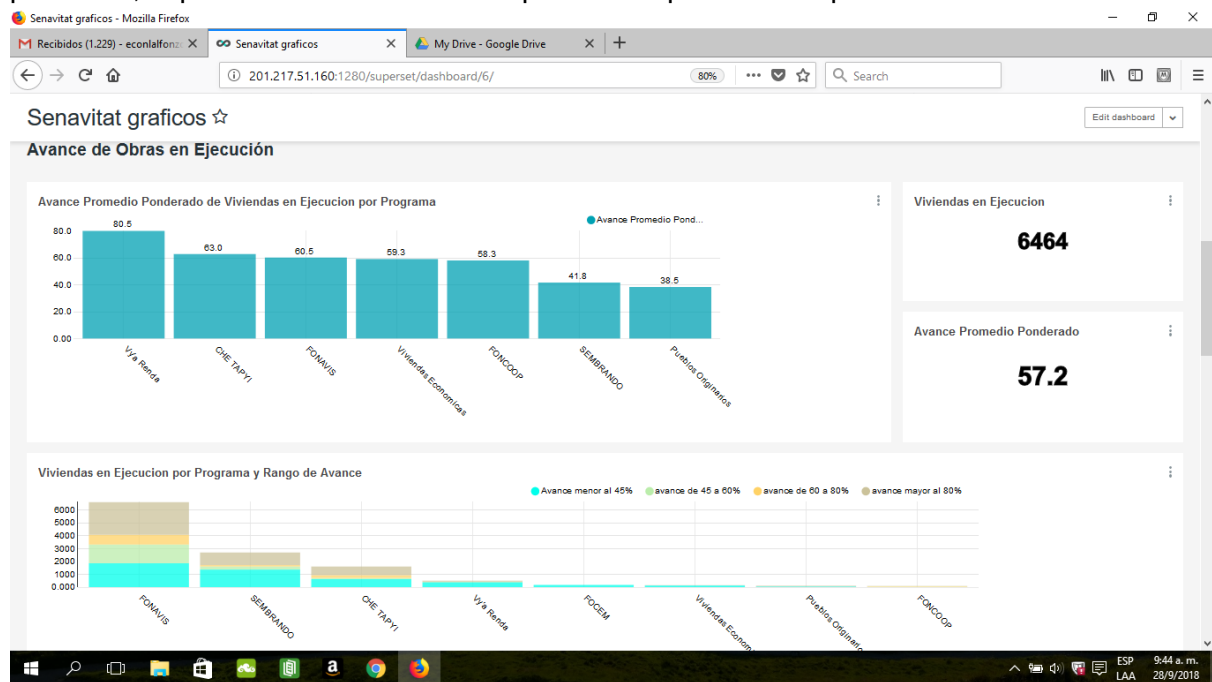
En esta sección se monitorea principalmente el avance físico de los diferentes proyectos y obras gestionadas por SENAVITAT en sus diferentes programas. En particular se cuantifican datos como la cantidad de viviendas planificadas, la cantidad de viviendas terminadas, el porcentaje de viviendas terminadas, la distribución de las viviendas planificadas según estado y según programa.



Se presentan además datos referentes a la ejecución promedio ponderada de los diferentes proyectos, agrupados por programa y rango de avance. Debido a que la construcción de

viviendas y componentes de hábitat; como calles, veredas, plazas, centros comunitarios, etc; gestionados por SENAVITAT, no necesariamente se realiza de manera secuencial, sino generalmente de manera simultánea y en paralelo, el grado de avance de los proyectos no se puede disgregar, por ejemplo, en cantidad de viviendas finalizadas, etc. Los criterios del porcentaje de avance son definidos por el ente, y se entiende que supera el 0% cuando ya ha habido un desembolso y las empresas constructoras tienen maquinarias y personal en el sitio, y alcanza un 100% con el proceso de entrega de viviendas culminadas.

Por otro lado, al agrupar los datos de ejecución promedio por dimensiones de análisis más abstractas, por ejemplo, por programas, no es muy apropiado utilizar un promedio simple, debido a que las obras tienen una gran variabilidad en tamaño y complejidad. Para dar una solución práctica a este problema, se utiliza como criterio de ponderación la cantidad de viviendas de un determinado proyecto. Por tanto, cuando en el tablero hablamos de ejecución promedio, implícitamente hablamos de promedios ponderados por cantidad de viviendas.

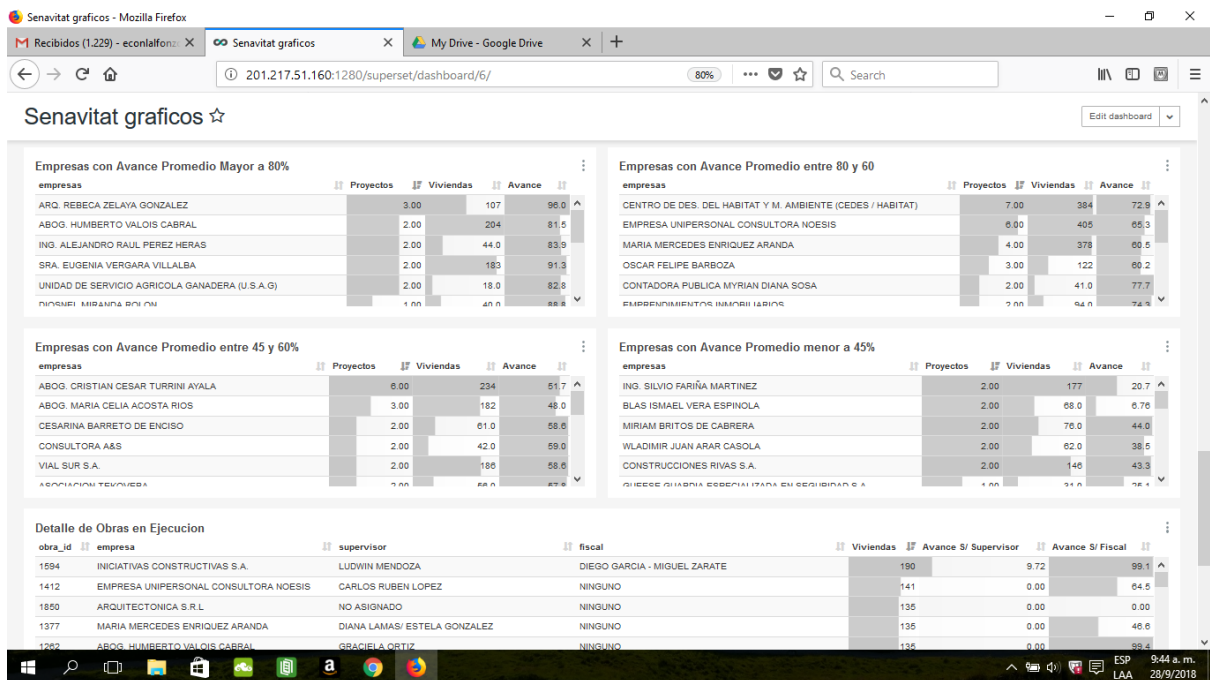
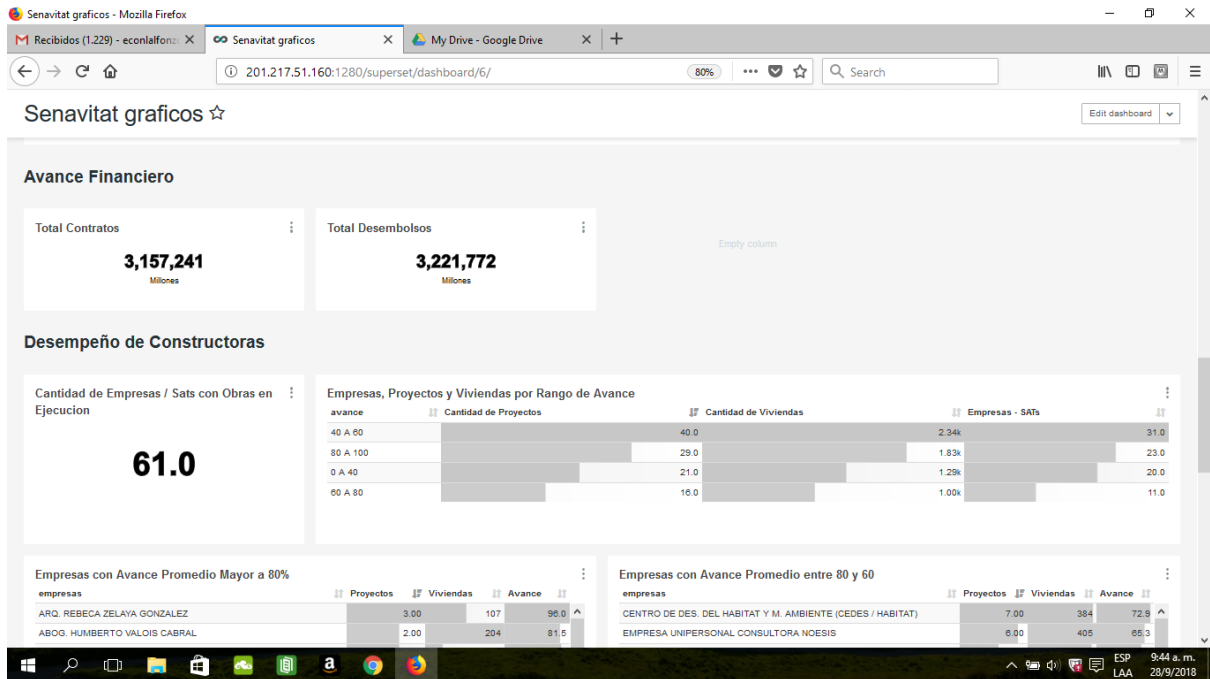


Desempeño de Constructoras

Además de esta información más agregada, resulta fundamental para el seguimiento de la estrategia de producción de SENAVITAT, poder seguir en detalle el desempeño de las constructoras. Esta a su vez, se encuentran monitoreadas por los niveles medios de la institución, por lo que este instrumento permite que la estrategia descienda a nivel operador.

Según se observa, se parte de la cantidad total de constructoras que tienen obras en ejecución. Posteriormente se visualizan métricas resumen por rango de avance, en porcentaje, que incluyen la cantidad de proyectos administrados, la cantidad de viviendas y la cantidad de empresas, ya sean constructoras o Servicios de Asistencia Técnicas, dentro de tal rango de avance.

Posteriormente se agrupan las empresas por rango de avance, entrando en detalle a nivel individual, con las métricas anteriores, además de la métrica de avance en sí misma. Concluyendo la sección, con el detalle de las obras ejecutadas.



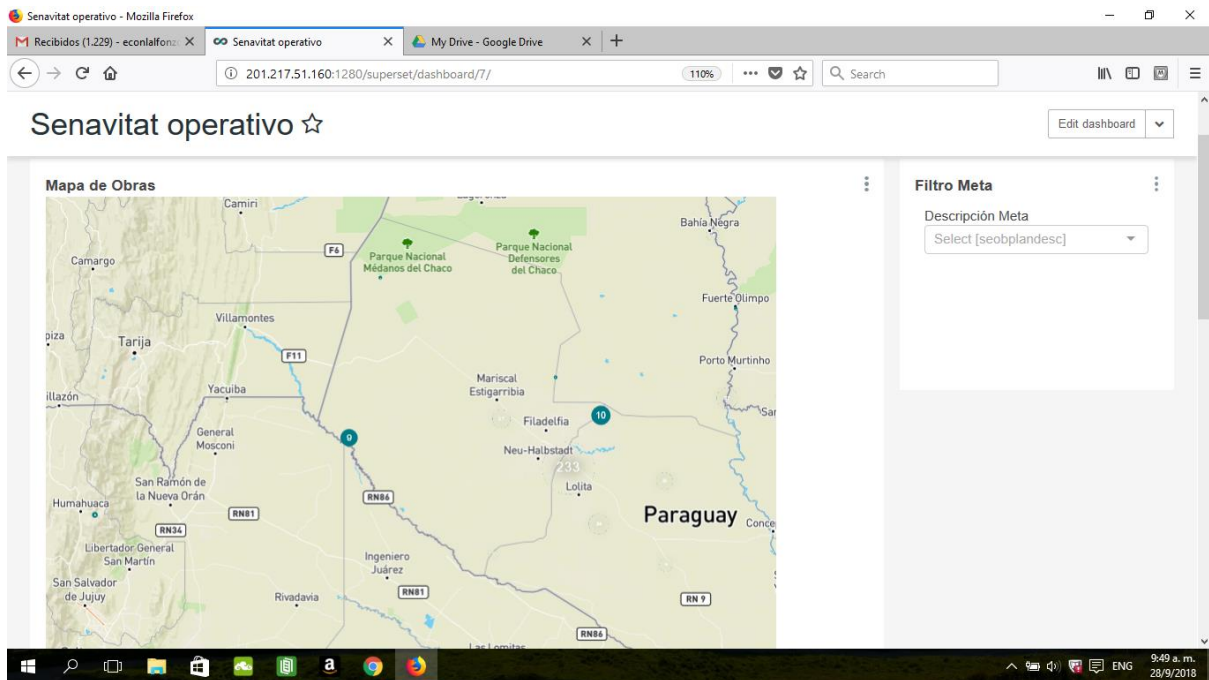
Mapa de Viviendas

El tablero también permite visualizar la información de manera geográfica. Un primer modelo presentado a la administración del ente permitió identificar algunas pocas obras que no se encontraban en la ubicación esperada debido, posiblemente, a errores en la carga de los datos georreferenciados.

La vista geográfica también permite el despliegue de algunas métricas adicionales, configurables por el usuario. Cuenta con etiquetas, colores y tamaño del círculo, que pueden

asociarse, por ejemplo, a los nombres del proyecto, el estado de la obra y la cantidad de viviendas.

Cabe destacar que esta visualización no pretende reemplazar el “mapa de viviendas” disponible en la página web, orientado a la ciudadanía, sino facilitar el control interno.



Nivel Operadores

A nivel de la gestión de los operadores, representados en SENAVIDAT por los supervisores y los fiscales de obra, encontramos importante incluir métricas que incluyan la cantidad de fiscales supervisados (en el caso de los primeros), la cantidad de obras; de modo a dimensionar el volumen de trabajo de cada operador; así mismo incluimos indicadores sobre atraso de las obras, particularmente el atraso máximo, el avance físico promedio ponderado, y el avance financiero, en términos de desembolso.

Senavitat operativo ☆ Edit dashboard

Fiscales **Supervisores** Programas

Obras por Supervisor

Show 10 entries Search:

Nombre del Supervisor	Cantidad de Fiscales	Cantidad de Obras	Atraso Máximo	Porcentaje de Avance	Desembolso
GERMAN FLEITAS - LUIS SCHEMBORI	3	3	-3006	100	7695831800
JULIO ENCINA - CARMEN TAPIOLAS	1	1	-2474	80	619091400
VICTORIANO FIGUEREDO	3	3	-2474	46	12252697256
ANTONIO BRIZUELA SERVIN	8	10	-2349	43	30557129530
CARMEN TAPIOLAS	3	3	-2284	77	2367314864
VICTORIANO FIGUEREDO - JORGE LOPEZ	1	1	-2264	35	1946998820
CARMEN TAPIOLAS - ANTONIO BRIZUELA	1	1	-2255	62	1535925799
FRANCISCO BAEZ - LAURA COLMAN	2	9	-2208	60	4960476920
FRANCISCO BAEZ	15	30	-2134	60	28666719478
DOMINGO CACERES	1	1	-2100	51	501227520

Previous 1 2 3 4 5 12 Next

Senavitat operativo ☆ Edit dashboard

Fiscales **Supervisores** Programas

Obras por Fiscal

Show 10 entries Search:

Nombre del Fiscal	Cantidad de Obras	Cantidad de Viviendas	Atraso Máximo	Porcentaje de Avance	Desembolso
HEBE LOPEZ	1	47	-3006	100	1843048647
LUIS LEIVA	6	103	-2632	80	5710961039
ELADIO FRANCO NAVONI	3	123	-2474	68	7253134842
GUSTAVO MARTINEZ	2	12	-2474	100	3053444415
DARIO GONZALEZ - RENE VILLABA	1	100	-2422	100	11138337927
AURORA VERA	3	51	-2421	100	3825301420
FRANCISCO FRACHI	2	124	-2399	85	5935686083
FRANCISCO FRACHI - CARMEN AIRALDI	1	30	-2399	100	1507697925
LILIAN ROLON - MIGUELA SANABRIA	1	100	-2397	100	10030940288
JUAN CARLOS SALAZAR - MANUEL IBARROLA	1	23	-2395	100	853297263

Previous 1 2 3 4 5 16 Next

Tablero de Control Cualitativo

En términos de un tablero de control cualitativo, para el seguimiento de la estrategia, sugerimos la adopción de los criterios y procedimientos establecidos por el MECIP (Modelo Estándar de Control Interno del Paraguay) de modo a evitar una duplicación de trabajo e información redundante, así como para lograr una sinergia con el área de control interno de

la entidad. En tal sentido, la información existente en planillas electrónicas puede utilizarse para alimentar el tablero a través de procedimientos ETL, o directamente desde la herramienta Apache Superset ®.

Entendemos, por tanto, que el diagnóstico del sistema de control interno (formulario 6) puede servir para tal efecto, ya que incluye una serie de indicadores que cubren todos los aspectos necesarios para la mejora continua organizacional, conjuntamente con la dirección estratégica basada en el monitoreo cuantitativo. En particular el diagnóstico del sistema de control interno incluye indicadores sobre: acuerdos y compromisos éticos; desarrollo del talento humano; protocolos de buen gobierno; planes y programas; modelo de gestión por procesos; estructura organizacional; contexto estratégico del riesgo; identificación, análisis y valoración de los riesgos; administración de los riesgos; políticas de operación y procedimiento; controles y transparencia; indicadores; manual de operación; información interna y externa; sistemas de información; comunicación pública e institucional; rendición de cuentas; autoevaluación del control, la gestión y el sistema de control interno, auditoría interna, plan de mejoramiento institucional, funcional e individual. Informaciones que ya se utilizan en la gestión de la entidad, por ello la mejoría observada en la calificación de cumplimiento del MECIP.

Observaciones Finales

De vuelta cabe recalcar que el objetivo de esta primera versión es la de, sobre todo, capacitar a los recursos humanos de la entidad en el uso de herramientas para el seguimiento de la estrategia, ofreciendo un producto mínimo viable, además de varias de capacitación in situ. Queda claro que el dominio de las herramientas introducidas a la gestión del ente requiere de mayor capacitación y práctica. En tal sentido la selección de los componentes que hemos realizado garantiza la disponibilidad de materiales de apoyo, tutoriales y manuales de usuario para cada aplicación, además de la disponibilidad gratuita de actualizaciones, al tratarse, en todos los casos de software libre de versión comunitaria.