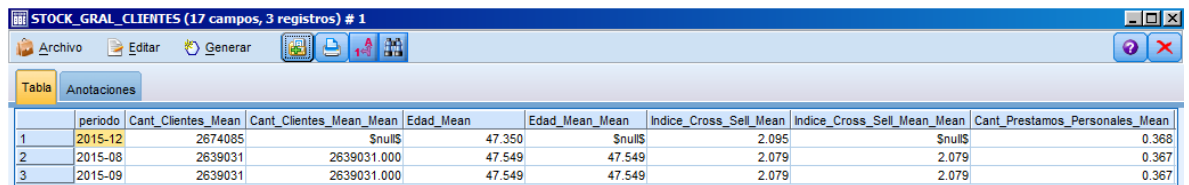


## MANUAL DE OPERACIONES EN SPSS

### Procedimiento sobre archivos fuente

#### A. Tipos de controles en Tablas de Resultados

- a. **General:** En las Tablas de Resultados para verificación se muestra un conjunto de variables para la comparación mes a mes en donde siempre se repite el promedio mensual de clientes en todas éstas tablas. A su vez se obtienen promedios acumulado de dichas variables.



	periodo	Cant_Clientes_Mean	Cant_Clientes_Mean_Mean	Edad_Mean	Edad_Mean_Mean	Indice_Cross_Sell_Mean	Indice_Cross_Sell_Mean_Mean	Cant_Prestamos_Personales_Mean
1	2015-12	2674085	\$null\$	47.350	\$null\$	2.095	\$null\$	0.368
2	2015-08	2639031	2639031.000	47.549	47.549	2.079	2.079	0.367
3	2015-09	2639031	2639031.000	47.549	47.549	2.079	2.079	0.367

**Período:** Cada variable analizada se abre en un registro fila mensual.

**Cant. de Clientes:** Conteo de los registros únicos de los clientes en la tabla de análisis.

**XXXX\_Mean:** Promedio mensual de la variable de la variable seleccionada.

**XXXX\_Mean\_Mean:** Promedio histórico acumulado de los meses anteriores al de análisis, en donde el mes bajo análisis contendrá valor *\$null\$* mientras el resto de los meses repetirá el promedio acumulado en todas las filas.

**XXXX\_Sum:** Conteo del total mensual de valores de la variable seleccionada.

**XXXX\_SDev:** Desvío estándar mensual de valores de la variable seleccionada.

**NOTA 1:** Es importante indicar que en las comparaciones de variables históricas se debe utilizar un criterio para decidir sobre la magnitud de los desvíos basado en la experiencia del analista de datos.

**NOTA 2:** Inicialmente se definió un período “BIS” indicado como AGOSTO2015 que debe ser reemplazado por los períodos a incorporar.

- | TRX_TC (3 campos, 1 registros) # 2 |                       |                            |  |
|------------------------------------|-----------------------|----------------------------|--|
| Archivo                            | Editar                | Generar                    |  |
| Tabla                              | Anotaciones           |                            |  |
| Periodo                            | Clientes_Key_RECUENTO | Importe_Consumo_Total_Mean |  |
| 1 2015-12                          | 922419                | 3148.937                   |  |

- 
- El diagrama de flujo de datos muestra la transformación de un archivo de texto 'VALORACION.txt' en una tabla de datos. El proceso comienza con 'VALORACION.txt', que pasa por un filtro y luego se fusiona con una tabla existente. El resultado se selecciona y se agrega a una nueva tabla. La captura de pantalla de la interfaz de usuario 'INGRESAR MES' muestra la configuración de la entrada de usuario, con los campos 'Almacenamiento' y 'Valores' definidos como 'Cadena' y '2015-12' respectivamente.

La siguiente tabla indica el tipo de control de resultados detallados arriba para cada uno de los módulos en el stream de SPSS:

INDICE	GENERAL	GENERAL Y DETALLE	GENERAL (parametrizable)
STOCK	STOCK_GRAL_CLIENTES	CANC_CTAS_PAQ	INGRESOS
	ALTAS_BAJAS_PROD	RECLAMOS	VALORACION
Base de datos	DOMICILIO		
	EMAILS		
	TELEFONOS		
Usabilidad Canales	RX_ATM		
	TRX_CANALES		
Usabilidad Producto		TRX_TC	
		TRX_TD	

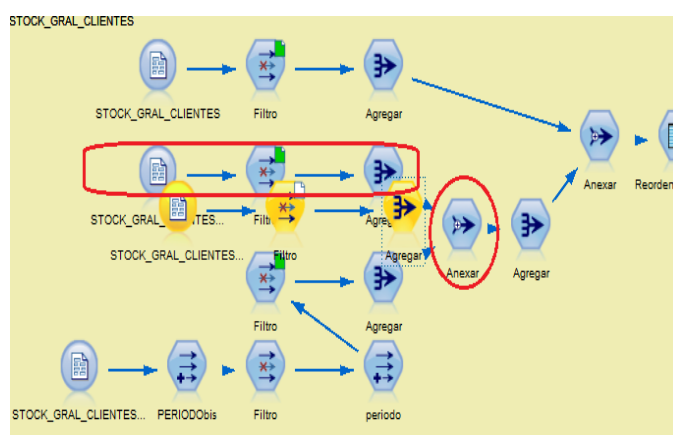
### C. Procedimiento mensual

**General:** Antes de iniciar la corrida de un nuevo mes se debe actualizar cada uno de los módulos en el stream SPSS de comparación de tablas incorporando:

- las nuevas tablas actuales y
- pasando a histórico aquellas tablas del último mes verificado.

1. En todos los módulos se debe copiar y pegar aquellos nodos consecutivos de procesamiento<sup>1</sup> de sólo el último mes anterior al actual.

Por ejemplo para el módulo de Stock:



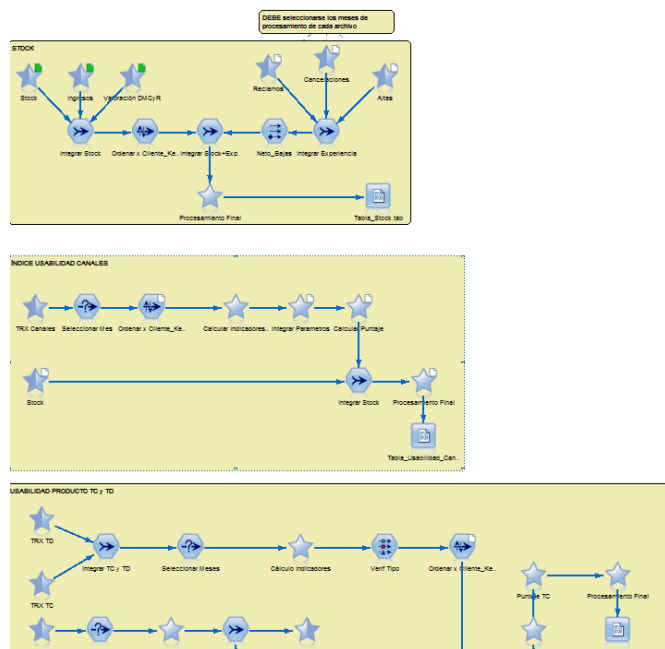
- 1.1 Se copia la rama correspondiente al último mes anterior al actual hasta el nodo "Anexar".
- 1.2 Se pega en el stream
- 1.3 En el nodo "Origen" de la rama pegada se selecciona la ruta a donde debe leer la nueva tabla del mes anterior al actual.
- 1.4 Se conecta la rama al nodo "Anexar", con lo cual deben quedar anexadas todas las ramas de tablas históricas a ese nodo.

2. Con los procesamientos periódicos, conforme se extiendan las tablas de resultados de verificación, se deberán eliminar las ramas más antiguas para evitar distorsiones en el promedio histórico (i.e. los campos **XXXX.Mean\_Mean**). Asimismo, por única vez, en los módulos que correspondan se debe quitar aquellas ramas que contengan un nodo llamado "PERIODObis". Estas ramas leen y procesan una tabla ficticia en un mes histórico (por lo general AGOSTO2015) para poblar de ejemplos iniciales para facilitar la comprensión de los campos en las comparaciones mes a mes.

<sup>1</sup> Llamaremos "RAMA" en un stream SPSS a un conjunto de nodos consecutivos que se inicia en un nodo "Origen", o que termina en un nodo "Resultado" o en un nodo de "Exportación".

## Procedimiento sobre Modelo de Índices

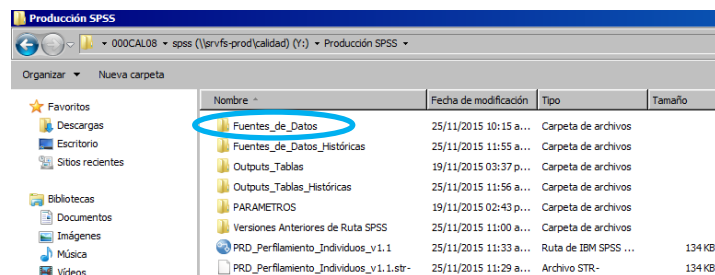
### Descripción de la Ruta en SPSS Modeler – Visión General



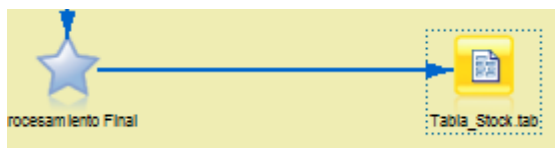
### Procedimiento para el Procesamiento de Índices

#### 3. Generar Tabla de Stock de Clientes

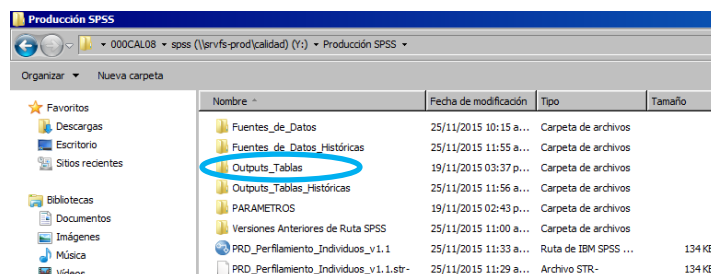
- Guardar los archivos planos enviados por COA, llamados “STOCK\_GRAL\_CLIENTES.txt”, “VALORACION.txt”, “RECLAMOS.txt”, “CANCELACIONES.txt”, “ALTAS\_BAJAS\_PROD.txt” y **archivos fuente de Ingresos**, en la carpeta **Producción SPSS -> Fuentes\_de\_Datos -> Stock**.



- Ejecutar el proceso desde el último nodo de exportación,



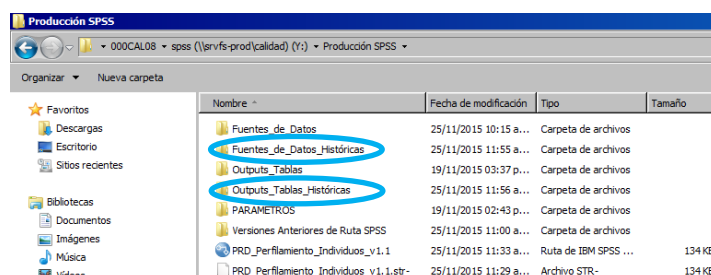
que guardará el resultado con el nombre “Tabla\_Stock.tab” en la carpeta “Outputs\_Tablas”.



#### 4. Realizar copias históricas de archivos fuente y output “Tabla\_Stock”

- c. Una vez ejecutada la Tabla Stock, copiar las fuentes de datos utilizadas (ver listado Anexo A) en la carpeta “Fuentes\_de\_Datos\_Históricas” agregando en el nombre del archivo el mes y el periodo de la extracción.
- d. Copiar la “Tabla\_Stock.tab” procesada en la carpeta “Outputs\_Tablas\_Históricas”, agregando en el nombre del archivo el mes y el periodo de la extracción.

De esta manera, en el periodo siguiente, cuando se envíen y corran la nueva tabla Stock de ese periodo, se puedan pisar los archivos ya que se habrá guardado una copia histórica de cada uno de ellos.



#### 5. Ejecutar Índices

- e. La Tabla Stock se utilizará ahora como un input más, en los supernodos Stock en el procesamiento del resto de los índices.

- f. Ejecutar el proceso desde el último nodo de exportación, que guardará el resultado de cada uno de los índices en la ruta de “Outputs\_Tablas” (ver listado Anexo B)
- g. Repetir paso 2, para realizar copias de archivos fuente y output de cada uno de los índices para guardar la historia.

#### ANEXO A - Archivos fuente de los índices

ÍNDICE	ARCHIVOS FUENTE
STOCK	STOCK_GRAL_CLIENTES
	VALORACION
	RECLAMOS
	CANC_CTAS_PAQ
	ALTAS_BAJAS_PROD
USABILIDAD CANALES	TRX_CANALES
	TRX_ATM
	Tabla_Stock (Procesado)
USABILIDAD PRODUCTOS	TRX_TC
	TRX_TC_RUBRO
	TRX_TD
	TRX_TD_RUBRO
	Tabla_Stock (Procesado)
BASE DE DATOS	EMAILS
	TELEFONOS
	DOMICLIOS
	Tabla_Stock (Procesado)
INTERÉS	STOCK_PLASTICOS_TC
	STOCK_PLASTICOS_TD
	Tabla_Stock (Procesado)
EXPERIENCIA	RECLAMOS_TIPOL
	CANC_CTAS_PAQ_MOTIVOS

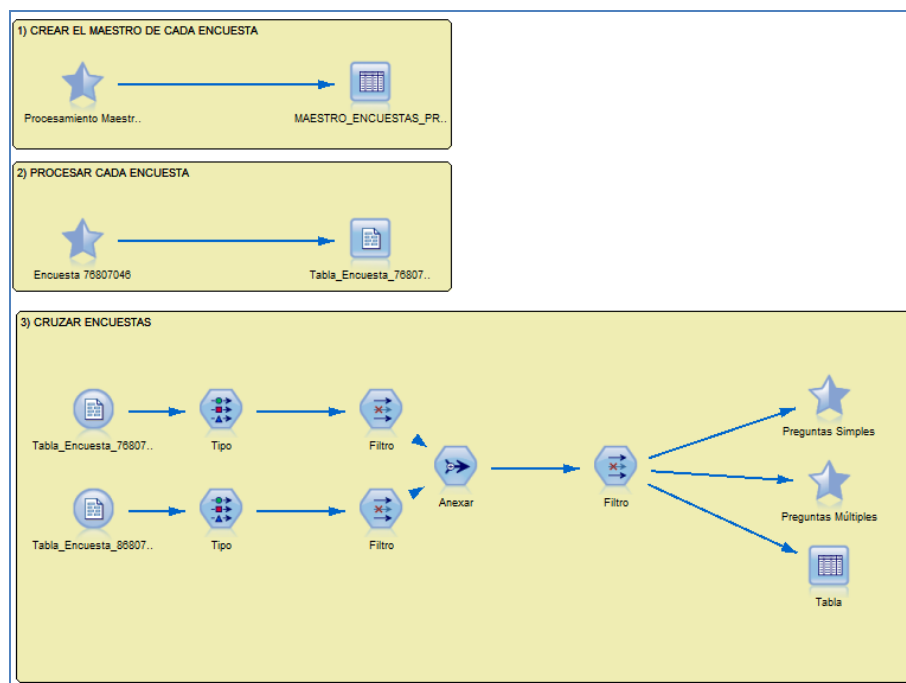
#### ANEXO B - Outputs de Procesamiento

ÍNDICE	OUTPUT PROCESAMIENTO
STOCK	Tabla_Stock
USABILIDAD CANALES	Tabla_Usabilidad_Canales
USABILIDAD PRODUCTOS	Tabla_Usabilidad_Productos
	Tabla Detalle Usabilidad Rubros TC
	Tabla Detalle Usabilidad Rubros TD

<b>BASE DE DATOS</b>	Tabla Base de Datos
<b>INTERÉS</b>	Tabla_Interés
<b>EXPERIENCIA</b>	

## Procedimiento sobre Tratamiento de Encuestas

### Descripción de la Ruta en SPSS Modeler – Visión General



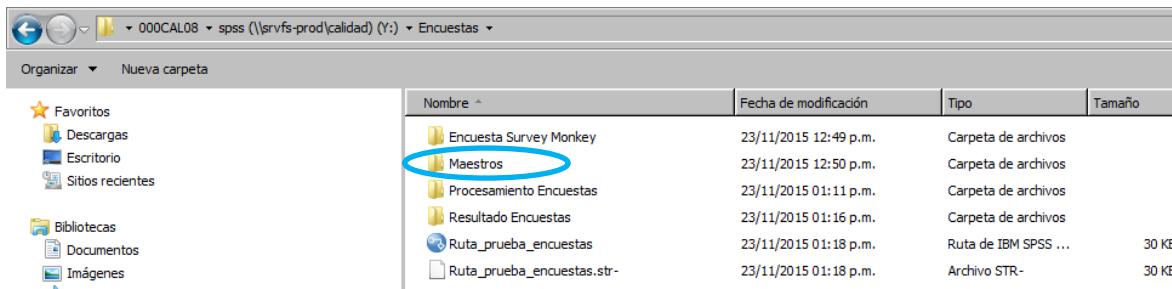
### Descripción de la Ruta en SPSS Modeler – Visión Detallada

El procedimiento para el tratamiento de las encuestas realizadas a través de Survey Monkey consiste en 3 pasos fundamentales:

#### 1) Crear el Maestro de cada encuesta entrante

El objetivo del Maestro de Encuestas consiste en registrar las preguntas y respuestas que el área está generando, y realizar agrupaciones lógicas que permitan agilizar el proceso de búsqueda y análisis de encuestas de distinto tipo.

#### A) Abrir carpeta “Encuestas”, dentro del compartido SPSS de área Calidad:





Dentro de la carpeta “Maestros”, encontraremos la planilla Excel “Maestro para encuestas” que posee las siguientes Macros en la solapa “BASE SURVEY MONKEY”:

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q
1																	
2																	
3																	
4																	
5																	
6																	
7																	
8																	
9																	
10	Identificación del encuestado	Identificac	Fecha de	Fecha de	Dirección	Dirección	Nombre	Apellido	Datos per	<b>ÁREA DE COPIADO DE PREGUNTAS</b> (a partir de celda J10)							
11																	
12	4334292981	76807046	#####	#####	200.71.232.209												
13	4334290267	76807046	#####	#####	200.71.232.209												
14	4333416594	76807046	#####	#####	200.71.232.209												
15																	
16																	

- Copiar y pegar los encabezados de la encuesta a procesar en la fila 10, comenzando en la columna A de esta planilla, y respetando que las preguntas de la encuesta sean siempre copiadas a partir de la celda J10 (área de preguntas).
- Hacer click en las macros 1° y 2° para completar las solapas de “MAESTRO DE ENCUESTAS” y “MAESTRO DE PREGUNTAS”

**A.1) En la solapa “MAESTRO DE ENCUESTAS”,** completar los campos marcados en amarillo.

	A	B	C	D	E	F	G	H	I	J	K	L
1	GRUPO_ENCUESTA	TIPO_ENCUESTA	OBJETO_ENCUESTA	SOLICITADOPOR	DISTRIBUCION_MUESTRA	FECHA_INICIO	FECHA_CIERRE	COD_ENCUESTA	OBSERVACIONES			
2	ClientLab	Express	MacroConectados	RRII	Todos	11/16/2015	11/16/2015	76807046				
3												

GRUPO\_ENCUESTA y TIPO\_ENCUESTA: Clasificación de primer y segundo nivel que agrupa a la encuesta.

OBJETO\_ENCUESTA: Objeto que está evaluando la encuesta.

SOLICITADOPOR: Área solicitante de la encuesta.

DISTRIBUCION\_MUESTRA: Muestra estudiada por la encuesta. Ejemplo: mujeres mayores de 50 años.

OBSERVACIONES: Campo opcional.

**NOTA:** Es importante definir un listado de variables por cada atributo marcado en amarillo, con el objetivo de realizar análisis cruzados por alguna de las variables descriptas.

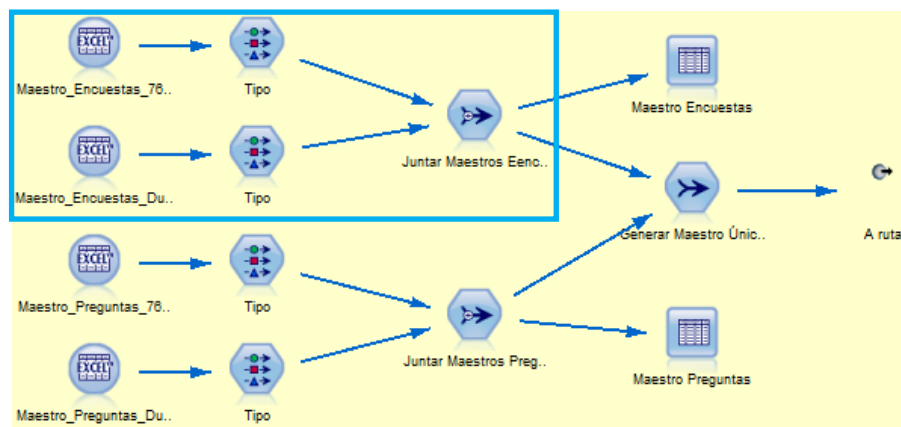
**A.2) En la solapa generada “MAESTRO DE PREGUNTAS”,** completar los campos marcados en amarillo.

	A	B	C	D	E	F	G	H
1	COD_ENCUESTA	TEMÁTICA	SUBTEMÁTICA	TIPO DE PREGUNTA	COD_PREGUNTA	PREGUNTA	COD_RESPUESTA	RESPUESTA
2	76807046	Conocimiento Cliente	Área de Trabajo	Pregunta Selección Múltiple	q001	¿Usted o alguien de su familia trabaja en alguna de las siguientes actividades?	q001_001	Marketing o publicidad
3	76807046	Conocimiento Cliente	Área de Trabajo	Pregunta Selección Múltiple	q001		q001_002	Relaciones públicas
4	76807046	Conocimiento Cliente	Área de Trabajo	Pregunta Selección Múltiple	q001		q001_003	Publicidad
5	76807046	Conocimiento Cliente	Área de Trabajo	Pregunta Selección Múltiple	q001		q001_004	TV, radio o Internet
6	76807046	Conocimiento Cliente	Área de Trabajo	Pregunta Selección Múltiple	q001		q001_005	Bancos o servicios financieros
7	76807046	Conocimiento Cliente	Área de Trabajo	Pregunta Selección Múltiple	q001		q001_006	Ninguna de las anteriores
8	76807046	Conocimiento Cliente	Digitalización	Pregunta Selección Múltiple	q002	¿Qué compañía/s de celular posee contratadas?	q002_001	Movistar
9	76807046	Conocimiento Cliente	Digitalización	Pregunta Selección Múltiple	q002		q002_002	Claro
10	76807046	Conocimiento Cliente	Digitalización	Pregunta Selección Múltiple	q002		q002_003	Personal
11	76807046	Conocimiento Cliente	Digitalización	Pregunta Selección Múltiple	q002		q002_004	Nextel
12	76807046	Conocimiento Cliente	Digitalización	Pregunta Selección Múltiple	q002		q002_other	Otro (especificar)
13	76807046	Conocimiento Cliente	Digitalización	Pregunta Selección Simple	q003	Aproximadamente ¿Con qué frecuencia utiliza el servicio de Internet en su celular?	q003	Respuesta
14	76807046	Usabilidad Producto	Promoción TC	Pregunta Selección Simple	q004	Considerando todos los aspectos de esta promoción. Usted diría que...	q004	Respuesta
15	76807046	Usabilidad Producto	Promoción TC	Pregunta Selección Simple	q005	Si este beneficio se encontrara disponible, Usted considera que...	q005	Respuesta
16	76807046	Usabilidad Producto	Promoción TC	Pregunta Selección Múltiple	q006	Por favor, indique cuál es su opinión acerca de cada aspecto del beneficio	q006_001	La duración de la promoción
17	76807046	Usabilidad Producto	Promoción TC	Pregunta Selección Múltiple	q006		q006_002	La cantidad de beneficios

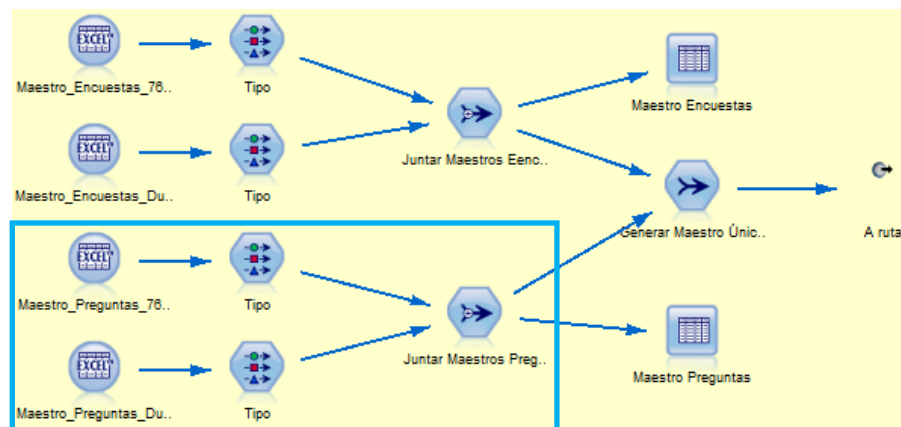
TEMÁTICA y SUBTEMÁTICA: Clasificación de primer y segundo nivel del concepto que se desea analizar en cada pregunta.

**A1.3) Ir a la solapa “BASE SURVEY MONKEY” y hacer click en la Macro 3, “Generar Archivos de Maestros”.** Se crearán en la carpeta “Maestros” 2 planillas nuevas con el código de la encuesta procesada.

**A1.4) Una vez creados los Maestros de Encuesta con la macro, incorporarlos en SPSS en nodos origen, para anexar cada Maestro de Encuesta, como se muestra abajo:**

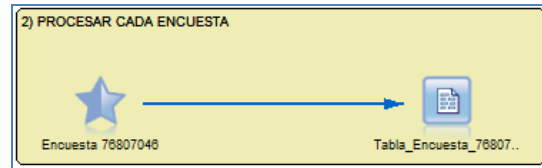


**A1.5) Una vez creados los Maestros de Preguntas con la macro, incorporarlos en SPSS en nodos origen, para anexar cada Maestro de Preguntas, como se muestra abajo:**

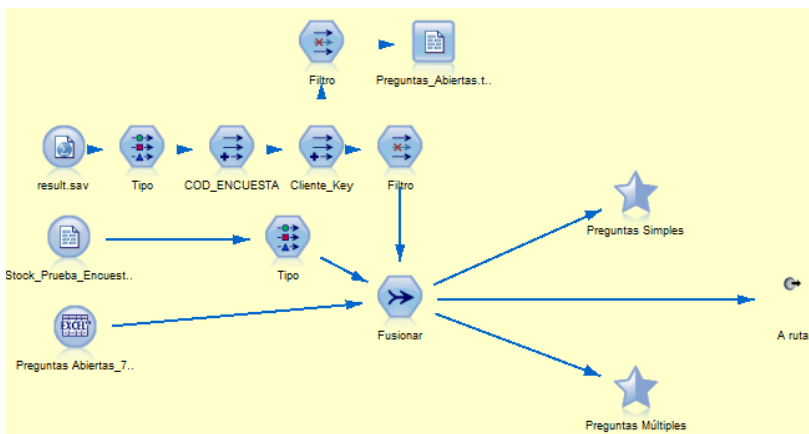


**A1.6) Para manejar fácilmente un gran volumen de columnas, integraremos (Fusionar) en SPSS los campos de cada encuesta y sus preguntas en una única tabla maestra.**

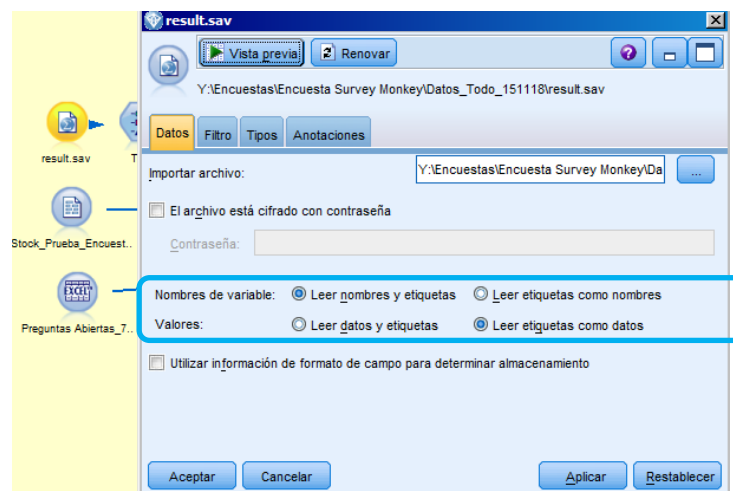
2) Procesar los resultados de cada encuesta entrante individualmente



**Detalle del Supernodo “Encuesta XXXX”:**

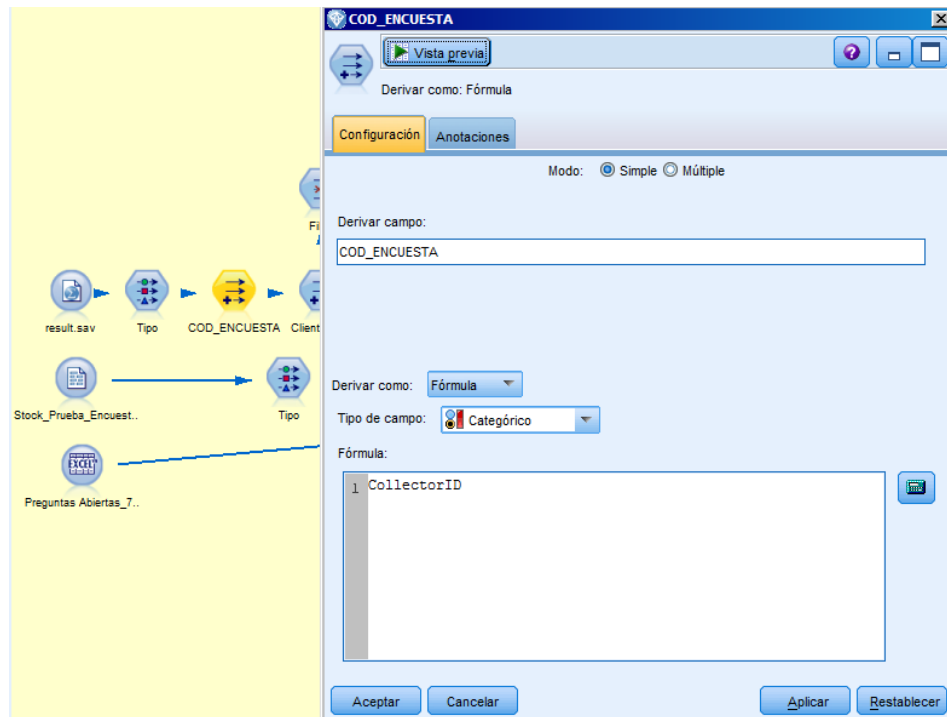


A) Importar los resultados de Survey Monkey con formato .sav; respetando las siguientes propiedades del Nodo de Origen “Archivo Statistics”.



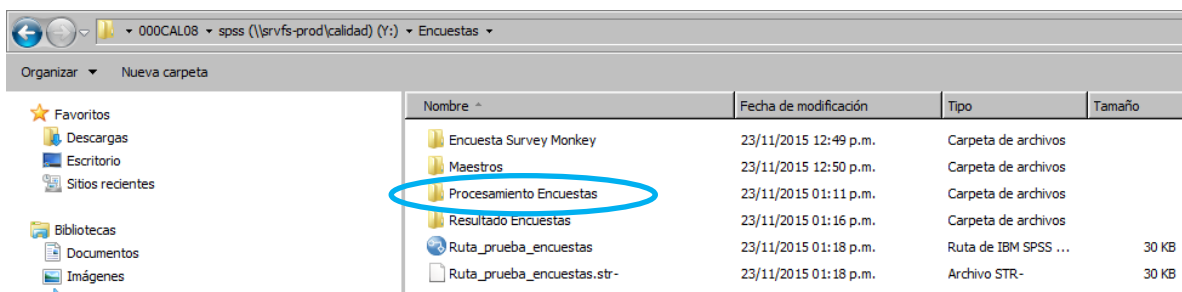
- Leer el tipo de las variables (Nodo Tipo).

- Cambiar a través de un Nodo Derivar el tipo de la variable CollectorID a CATEGÓRICO para utilizar en los gráficos posteriores y renombrarlo con “COD\_ENCUESTA”.

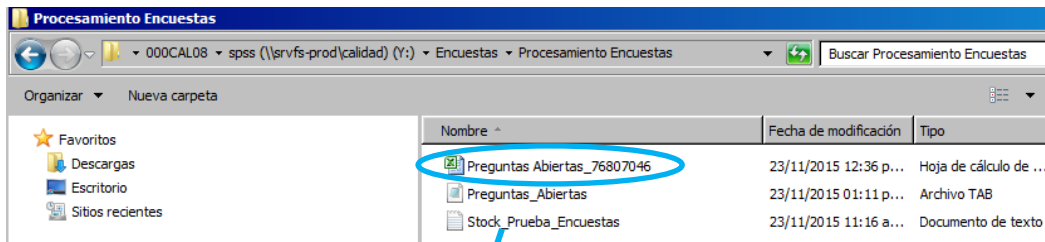


- En otro Nodo Derivar, renombrar la variable RespondentID como Cliente\_Key, y filtrar los campos que no sean necesarios.

A) Filtrar el Cliente\_Key, COD\_ENCUESTA y las respuestas de las preguntas abiertas del archivo fuente .sav, para exportarlas y codificar sus respuestas manualmente. Exportar en archivo plano, en la carpeta “Procesamiento Encuestas”:



B) Renombrar el Archivo de Preguntas Abiertas con un guión bajo y el Código de Encuesta. Integrar como una fuente más el archivo con las preguntas ya codificadas, fusionándolas por Cliente\_Key.

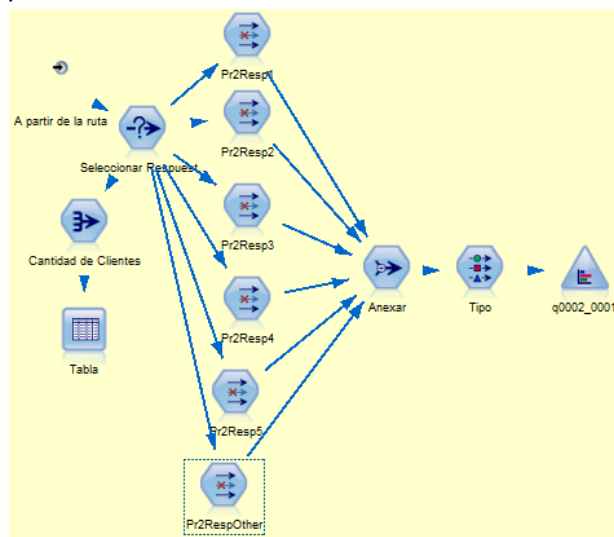


C) Integrar (fusionar) con el archivo de Stock de Clientes.

**NOTA:** Definir el periodo con información de clientes a analizar ya que el mismo puede ser distinto al periodo de encuestas.

Ejemplo: tomar el último archivo de Stock de Clientes si se desea comparar encuestas de distintos periodos.

- D) **Supernodo “Preguntas Simples”:** Realizar gráficos de distribución para las variables correspondientes a preguntas simples.
- E) **Supernodo “Preguntas Múltiples”:** Por cada pregunta múltiple, debemos generar un stream compuesto por:



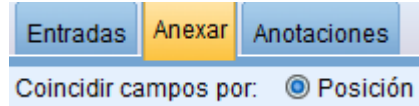
- Nodo Seleccionar: seleccionar todas las columnas / respuestas que pertenecen a una pregunta del tipo múltiple.
- Nodos Filtro: seleccionar en forma individual cada respuesta en un nodo.

Filtro			Anotaciones		
			Campos: 149 de entrada, 40 filtrados, 0 cambiados de nombre, 109 de salida		
Campo	Filtro	Campo			
q0002_0001	→	q0002_0001			
q0002_0002	→	q0002_0002			
q0002_0003	→	q0002_0003			
q0002_0004	→	q0002_0004			
q0002_0005	→	q0002_0005			
q0002_other	→	q0002_other			

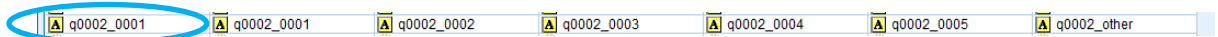
El objetivo es obtener "tablas temporales" donde para la pregunta en análisis, quede 1 columna con todas las respuestas.

- c. Nodo Anexar: luego cada una de las respuestas que se encuentran en un nodo de filtro se anexan como tablas unas debajo de las otras.

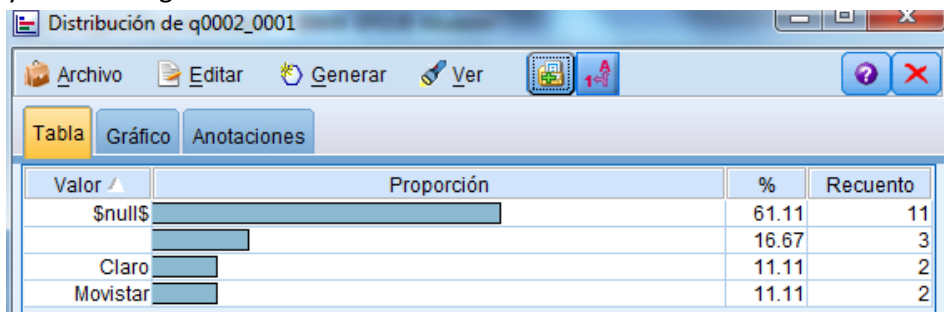
Para ello debe hacerse coincidir los campos por Posición



De esta manera cada respuesta quedará en una única columna que llevará el nombre de la primera respuesta y podrá ser explotada como una respuesta simple



y realizar un gráfico de distribución



**NOTA:** el % y los recuentos siempre cuentan la cantidad de respuestas y no de encuestados que han respondido (en el ejemplo son 3 clientes, donde 1 de ellos respondió Claro y Movistar)

- d. Nodo Agregar: mediante esta agregación a nivel de cliente, se identifica la cantidad de clientes que han respondido a la pregunta múltiple en análisis.

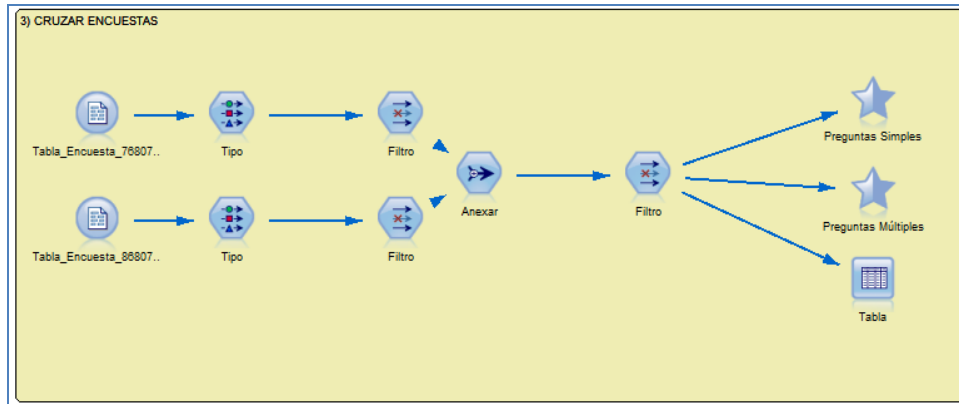
Agregados básicos

Campos agregados:

Campo	Suma	Media	Mín.	Máx.	Desv. est.	Mediana	Recuento	Varianza	Primer cuartil	Tercer cuartil
Cliente_Key	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

- F) Exportar la "Tabla\_Encuesta\_X.tab" que contiene ya el Stock de Clientes fusionado y las Preguntas Abiertas ya codificadas, como archivo plano separado por tabulaciones, colocando en X el Código de la encuesta o "CollectorID".

### 3) Cruzar las encuestas deseadas



A) **Tomar del Maestro las preguntas a comparar, en función de las distintas variables de filtro.** (Por ejemplo, Tema de la pregunta, Tipo de encuesta, Temática. Etc.)

B) **Importar con nodos origen las Encuestas procesadas individualmente y filtrar solo aquellas preguntas a comparar seleccionadas en el Maestro en el paso anterior.**

**NOTA:** Recordar que cada una de las encuestas traerá la variable de Stock de Clientes procesada anteriormente.

Dependiendo del análisis, se requerirá seleccionar una foto de Stock o se genere un nuevo Stock con el cruce de las encuestas. Ejemplo: Comparar cómo respondieron en Encuesta B, clientes de la Encuesta A.

C) **Anexarlas (deben poseer la misma cantidad de columnas).** Seleccionar “Coincidir campos por Posición” y las columnas resultantes tendrán el nombre de la primera Tabla\_Encuesta.

**Anexar**

Vista previa

Añadir 2 conjuntos de datos

Entradas Anexar Anotaciones

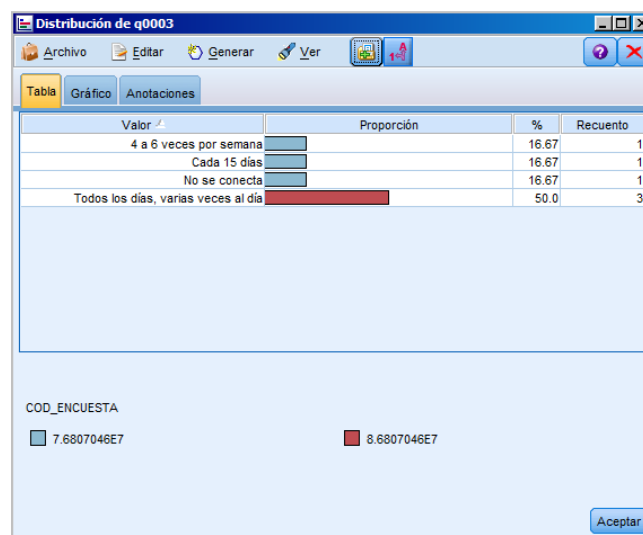
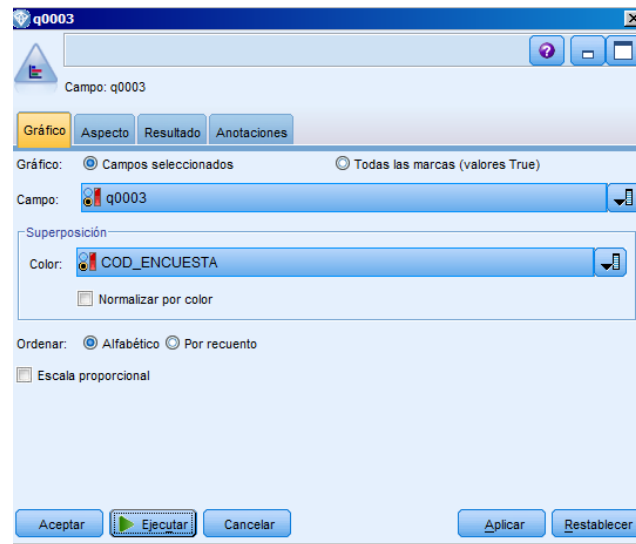
Coincidir campos por: ☒ Posición ☐ Nombre

Vista previa de estructura y coincidencias de campos

Campo de resultado	1[Tabla_Encuesta_7680704...	2[Tabla_Encuesta_8680704...
Cliente_Key	Cliente_Key	Cliente_Key
CODIGO_CLIENTE	CODIGO_CLIENTE	CODIGO_CLIENTE
periodo	periodo	periodo
Fecha	Fecha	Fecha
fecha_key	fecha_key	fecha_key
Antig_Cliente	Antig_Cliente	Antig_Cliente
Antig_Cliente_Prod_Vig	Antig_Cliente_Prod_Vig	Antig_Cliente_Prod_Vig
Sexo	Sexo	Sexo
Fecha_Nacimiento	Fecha_Nacimiento	Fecha_Nacimiento
Edad	Edad	Edad
Estado_Civil	Estado_Civil	Estado_Civil

D) Filtrar campos que no sean necesarios.

E) **Supernodo “Preguntas Simples”**: Realizar gráficos de distribución para las variables correspondientes a preguntas simples, realizando una superposición por COD\_ENCUESTA, para visualizar los resultados por encuesta.



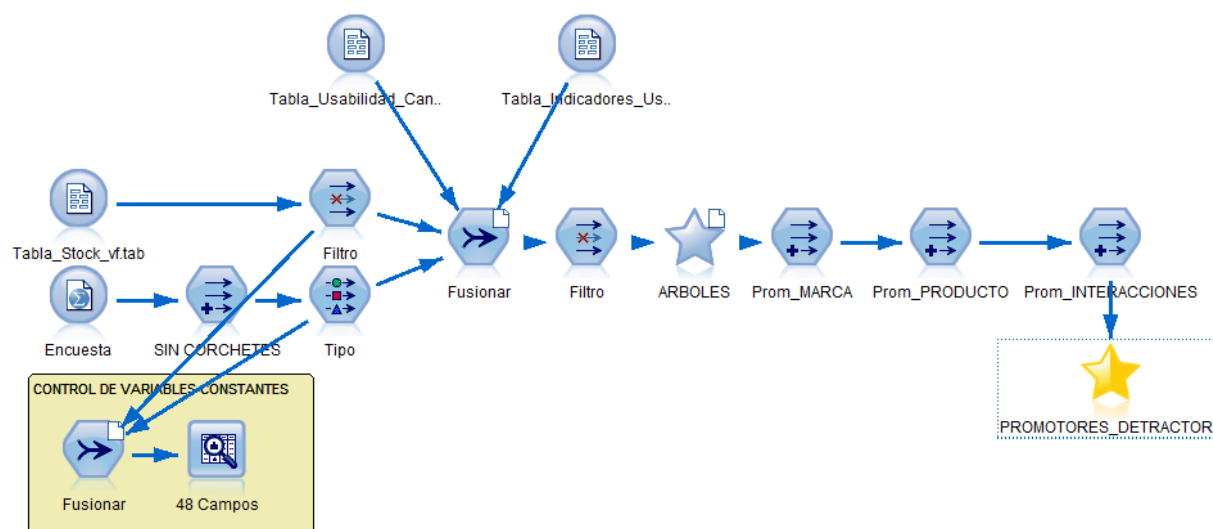
F) **Supernodo “Preguntas Múltiples”**: Por cada respuesta asociada a preguntas múltiples que se desea comparar, se debe realizar el mismo procedimiento especificado en el punto 2, en miras de transformar la misma en una pregunta simple que puede ser analizada mediante barras de distribución.



## Procedimiento sobre Modelo de ÍNDICES EXPERIENCIA CLIENTES

### Descripción de la Ruta en SPSS Modeler – Visión General

El siguiente modelo para el procesamiento de encuestas integra las respuestas de experiencia de las variables Marcas, Productos e Interacciones con las variables de perfilamiento e índices de usabilidad (productos y canales), con el objetivo de obtener un índice que se extrapole a toda la población de clientes individuos, mediante el uso de técnicas de árbol de decisión.



Desde el punto de vista técnico para un modelo de árbol de decisión se debe poder identificar correlación entre determinados campos del perfilamiento, otros índices y las respuestas de las encuestas de experiencia. A tal efecto se propone la siguiente selección de campos, los cuales el modelo en SPSS en definitiva determinará cuáles de ellos serán predictores:

DATOS DEMOGRAFICOS	DATOS DE RELACIÓN CON EL BANCO	MÉTRICAS DE COMPORTAMIENTO
Division	Antig_Cliente	Indicador_Cuenta
Region	Antig_Cliente_Prod_Vig	Indicador_Paquete
Estado_Civil	Rango_Antiguedad	Indicador_PF
Sexo	Segmento_Comercial	Indicador_PP
INGRESO_FINAL	SubSegmento_Comercial	Indicador_Reclamos_Ingresados
Edad	Indice_Cross_Sell	Indicador_Reclamos_Reinicio
		Indicador_Reclamos_Resueltos
		Indicador_TC
		Indicador_TD

		Neto_Altas
		Neto_Bajas
		Posicion_Activa
		Posicion_Pasiva
		Saldo_Promedio_Paquete_3
		Saldo_Promedio_TC_3
		Saldo_promedio_vista_3
		Valor_Promedio

**NOTA:** en caso de querer incorporar nuevas variables, por considerarse que pueden impactar en los predictores, se debe sacar los filtros aplicados a los archivos de origen.

### LA ENCUESTA DE OPINION - Conceptos generales y formatos

La encuesta detalla las respuestas de opinión sobre la MARCA, el PRODUCTO y la INTERACCION brindados por el banco, a través de 6 preguntas generales.

Los resultados registrados son volcados en archivo plano con formato SPSS Statistics (result.sav) según:

	CollectorIm	RespondentID	CollectorID	StartDate	EndDate	IPAddress	EmailAddress	FirstName	LastName	CustomData	q0001_0001	q0001_0002	q0001_0003	q0001_0004	q0001_0005	q0001_0006	q0002
1	Web Link 1	4489530251.0...	8034134...	02/04/2016 18:44:53	02/04/20...	200.71.232...					1.000	2.000	3.000	4.000	5.000	5.000	[5320076]
2	Web Link 1	4489530935.0...	8034134...	02/04/2016 18:45:26	02/04/20...	200.71.232...					2.000	2.000	2.000	2.000	2.000	2.000	[8351473]
3	Web Link 1	4489531868.0...	8034134...	02/04/2016 18:45:44	02/04/20...	200.71.232...					5.000	5.000	5.000	5.000	5.000	5.000	[6848937]
4	Web Link 1	4489532741.0...	8034134...	02/04/2016 18:46:05	02/04/20...	200.71.232...					1.000	2.000	4.000	4.000	5.000	5.000	[8121354]
5	Web Link 1	4489533569.0...	8034134...	02/04/2016 18:46:24	02/04/20...	200.71.232...					2.000	2.000	2.000	2.000	2.000	2.000	[5341503]
6	Web Link 1	4489535009.0...	8034134...	02/04/2016 18:46:42	02/04/20...	200.71.232...					5.000	5.000	5.000	5.000	5.000	5.000	[7463455]

donde para este modelo sólo se utilizarán los campos de respuestas con el contenido de respuesta codificada de 1 a 5 que se corresponden con:

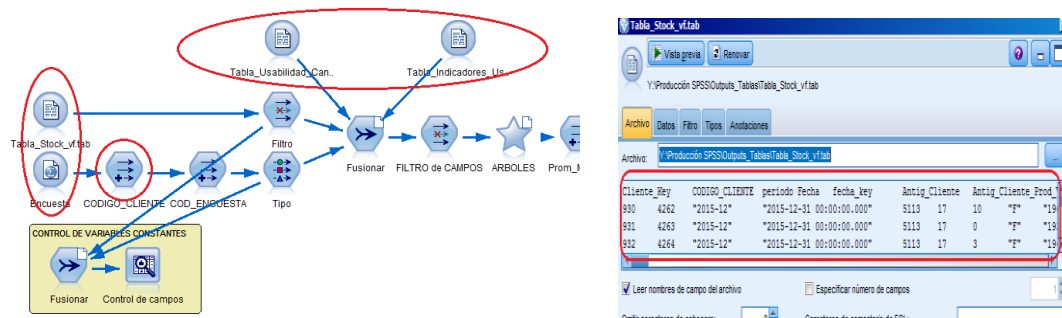
1. Totalmente en desacuerdo
2. Algo en desacuerdo
3. Ni de acuerdo ni en desacuerdo
4. Algo de acuerdo
5. Totalmente de acuerdo

y con la última columna que indica el CODIGO\_CLIENTE.

### Descripción de la Ruta en SPSS Modeler – Visión Detallada

El procedimiento para la corrida del stream SPSS consta de los siguientes ítems que se deberán ejecutar en el orden aquí abajo indicado:

- 1) Asegurarse de que se dispongan las últimas versiones de las tablas con las variables predictoras (al momento de la preparación de este documento, se consta con **Tabla\_Stock**, **Tabla\_Usabilidad\_Canales** y **Tabla\_Indicadores\_Usabilidad\_TC\_TD**) y disponibles en sus correspondientes carpetas de Producción. Asimismo se deberá disponer del file **result.sav** proveniente de Survey Monkey en formato SPSS Statistics. Una manera sencilla de asegurarse la lectura de las 4 tablas es editando los nodos y verificar en cada uno que se leen los campos:



**NOTA:** el presente proceso, sólo contempla 1 única encuesta a la vez. Sin embargo, se recomienda que en la medida que se vayan generando nuevas encuestas, se vayan acumulando en un único archivo o se anexas (nodo Anexar de la función Operaciones con Registros), con el objetivo de obtener mejores predictores, ante la mayor cantidad de casos.

- 2) Se sigue un nodo de definición del campo CODIGO\_CLIENTE ya que en las exportaciones en formato SPSS proveniente desde Survey Monkey no se incorpora ese título de campo y habitualmente se muestran los códigos en último campo del archivo, el cuál variará su nombre en función de la cantidad de preguntas que posea la encuesta.
- 3) Se dispone de una rama auxiliar recuadrada bajo el título "CONTROL DE VARIABLES CONSTANTES" que ejecuta un proceso de auditoría de datos para el control de qué variables son constantes, las cuáles no pueden ser utilizadas en los procesos árbol de decisión justamente porque sobre ellas no hay nada que decidir ya que siempre toman un mismo valor<sup>2</sup>. Esta rutina corre un nodo de control de distribución de variables para analizar la inexistencia de valores o de valores constantes:

<sup>2</sup> Puede suceder que determinados campos de algunas de las tablas del perfilamiento sean valores constantes (ej.: Neto de Bajas, posee un valor único equivalente a 0, porque no existen clientes encuestados con bajas de producto). También puede suceder que para esos clientes existan campos sin valores (i.e. celdas "\$null\$" en los campos)

**Auditoría de datos de [48 campos]**

Archivo Editar Generar

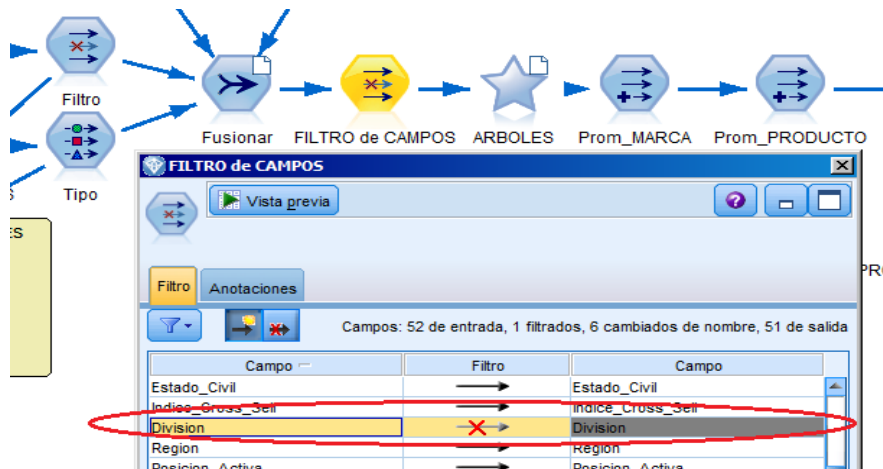
Auditoría Calidad Anotaciones

Campo	Gráfico	Medida	Mín.	Máx.	...	...	Válido
Indice_Cross_Sell		Continuo	1	5	...	...	25
Division		Nominal	--	--	...	1	25
Region		Nominal	--	--	...	1	25
Posicion_Activa		Continuo	0.000	88158.160	...	...	25
Saldo_promedio_vista_3		Continuo	0.000	188366.690	...	...	25
Saldo_Promedio_Paquete_3		Continuo	0.000	0.000	...	...	25

Aquí a la izquierda se observan 6 variables. 3 de ellas: **División**, **Región** y **Saldo\_Promedio\_Paquete\_3** no podrían participar como entrada en un proceso de árbol de decisión. Y los restantes 3 campos:

**Indice\_Cross\_Sell**, **Posición\_Activa** y **Saldo\_Promedio\_vista\_3** que sí podrían incorporarse al proceso

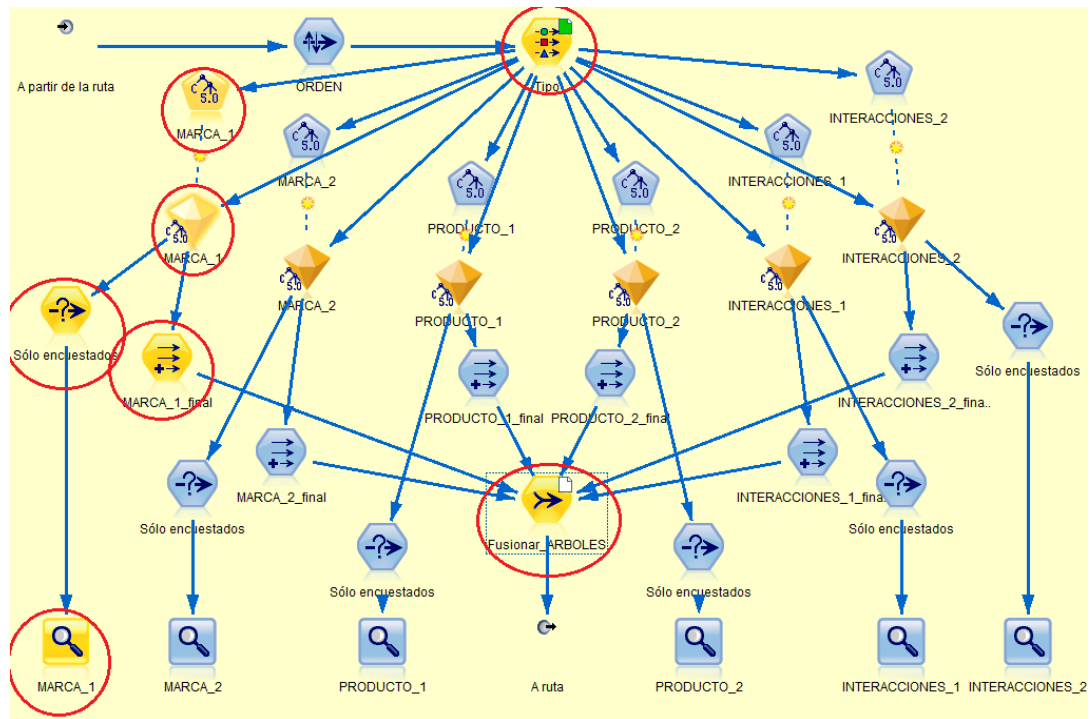
- 4) Luego del control de arriba, se deberá filtrar en el nodo FILTRO de CAMPOS aquellos campos que sean constantes o sin valores (i.e. "\$null\$") identificados en el paso anterior:



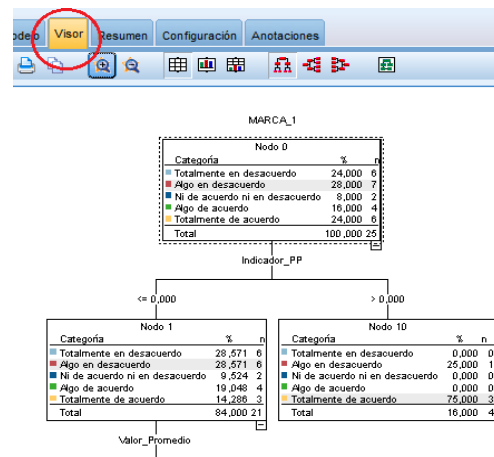
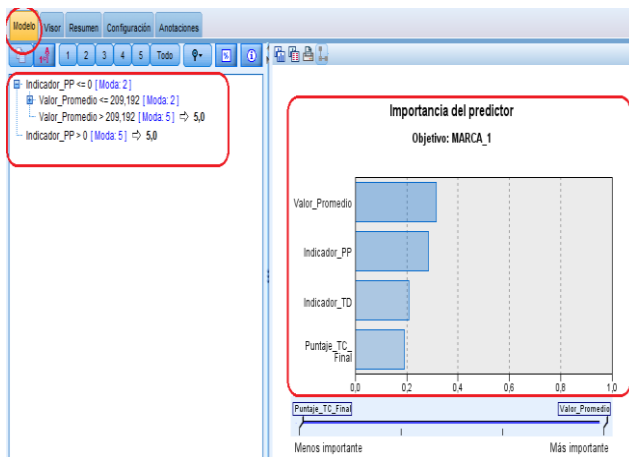
Por ejemplo aquí a la izquierda, en la edición del nodo FILTRO de CAMPOS se indica quitar el campo "Division" para los procesos que siguen

**NOTA:** en caso de no filtrarse, cuando se ejecute el proceso total, traerá un error asociado a variables constantes.

- 5) Luego del filtro de campos para procesamiento sigue un análisis de efectividad de la bondad del modelo con las variables seleccionadas. Para ello se explicará una sola de las 6 ramas asociadas a los 6 procedimientos de árbol de decisión. Se tiene el siguiente supernodo "ARBOLES" con las 6 ramas de árboles definidos:

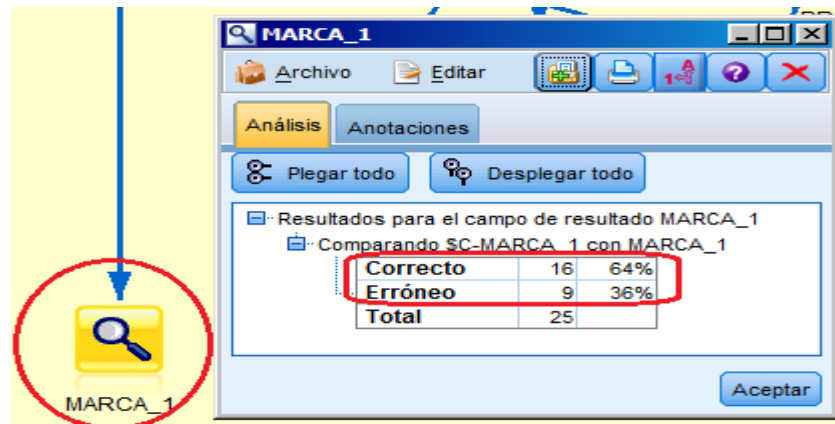


- I. En el caso del nodo MARCA\_1 de parametrización de árbol de decisión C5.0 es el generador del árbol y está definido por defecto en forma automática. El nugget que sigue es el resultado y muestra el par de gráficos siguientes según se seleccione solapas “Modelo” o “Visor”. En “Modelo” se define la regla de decisión específica para MARCA\_1 y también los predictores con sus rangos de correlación, y en “Visor” la misma regla en forma gráfica, típicamente de árbol:



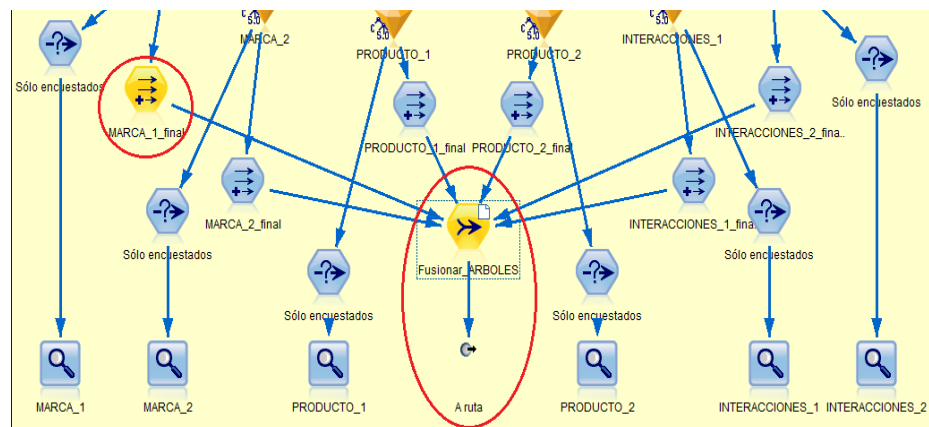
- II. El siguiente control de eficacia del árbol de decisión C5.0 se ejecuta desde el nodo Análisis “MARCA\_1” el cual cuenta los casos en donde la

estimación del campo MARCA\_1 que calcula el árbol C5.0 coincide con la verdadero registro de los clientes encuestados:



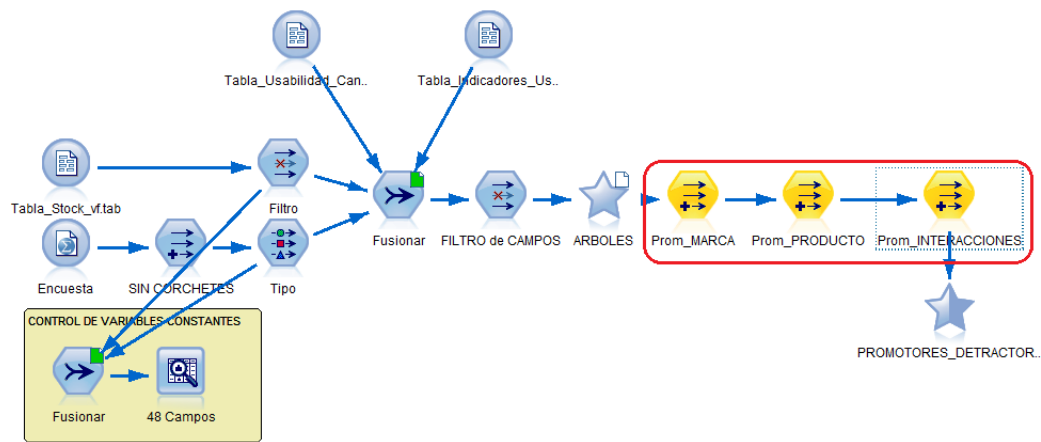
En esta rama el resultado de estimación es relativamente bueno con el 64% de los casos son estimados para MARCA\_1.

- III. A continuación el nuevo campo creado por el proceso C5.0 extiende, en el nodo “MARCA\_1\_final” la estimación de MARCA\_1 a los restantes clientes no encuestados según la regla de decisión determinada en el proceso, y se incorpora en el nodo fusión<sup>3</sup>. Este nodo “Fusionar ARBOLES” funde las 6 estimaciones extendidas a toda la población.



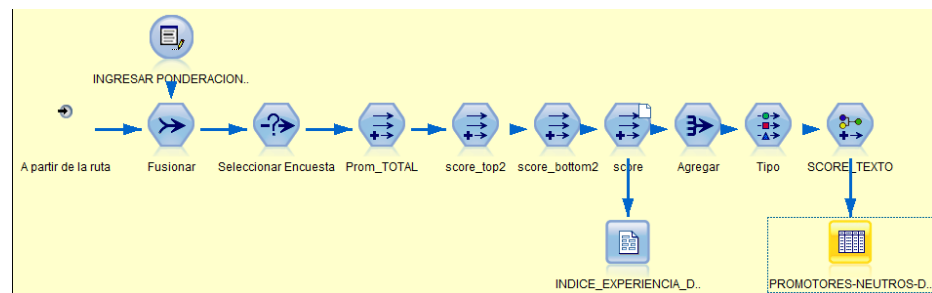
- 6) Los siguientes nodos que siguen al supernodo “ARBOLES” corresponden al cálculo de índice de experiencia:

<sup>3</sup> Es importante tener en cuenta que las reales respuestas de MARCA\_1 dadas por los clientes encuestados NO son reemplazados por las estimaciones del proceso de árbol de decisión C5.0.



7) El supernodo “PROMOTORES\_DETRACTORES” ejecuta los cálculos del índice:

- I. Inicialmente se incorpora la opción de ponderar qué aspectos influenciar en el índice, ya sea apalancar las respuestas de MARCA, las de PRODUCTO y/o las de INTERACCION. Como siempre, los coeficientes de ponderación elegidos deben sumar 1; por defecto se muestran ponderadores iguales a 0.33<sup>4</sup>.



- II. Se sigue un nodo selección llamado “Seleccionar encuesta” para indicar la encuesta sobre la cual se desea calcular el índice EXPERIENCIA\_CLIENTE. Por defecto, en este nodo no hay ninguna encuesta seleccionada.
- III. A continuación se siguen nodos para el cálculo del campo “score”, una tabla agregación que cuenta la cantidad de casos de score=1, score=0 y score=-1 los cuales se etiquetan como clientes PROMOTORES, NEUTROS y DETRACTORES y finalmente se exporta la tabla “INDICE\_EXPERIENCIA\_DETALLE” con el agregado de los nuevos 6 campos estimados de MARCA, PRODUCTO e INTERACCION, y los “score\_TOP”, “score\_BOTTOM” y “score” finales.

<sup>4</sup> Valores de ponderación iguales entre sí (y siempre sumando 1) no produce ninguna influencia y los promedios ponderados resultantes son simples promedios.