

Exercício 1 com dados da RMSP – parte 1

Este exercício tem por objetivo ilustrar a verificação e uso dos microdados da amostra de domicílios do Censo 2010 do IBGE, para a RMSP.

(1)

Baixar e descompactar os arquivos abaixo, dos links:

Repositório > Notas de Aula e Exercícios >

dados da RMSP para exercícios > 2012 >

- Setores Censitários e Áreas de Ponderação (sp.zip)
 - APxSetor.dbf
 - 35SEE250GC_SIR.shp/shx/dbf/prj

- Microdados 2010 - Domicílios (Domicilios2010.zip)
 - Domicilios2010.mdb

(2)

Observe que se deve antes criar o BD **Setores2010.mdb** com o TerraView, em seguida importar o shapefile **35SEE250GC_SIR** (.shp, .shx, .dbf e .prj) como plano e tema, conforme indicado na porção inferior da Figura 1.

O TerraView não aceita nome iniciado com numeral, portanto todos os arquivos do shapefile devem ser renomeados, por exemplo, para **SP_Setores**.

O sistema de coordenadas desse plano é o das latitudes e longitudes (em graus decimais) ao Datum SIRGAS 2000.

Observe que se trata do mapa de todos os setores censitários, do Censo 2010, do Estado de São Paulo. Reserve o banco de dados, pois será usado mais tarde.

(3)

Executar o **MS-Access** e abrir o arquivo

Domicilios2010.mdb

Abra a tabela **Amostra_Domicilios_35_RMSP**

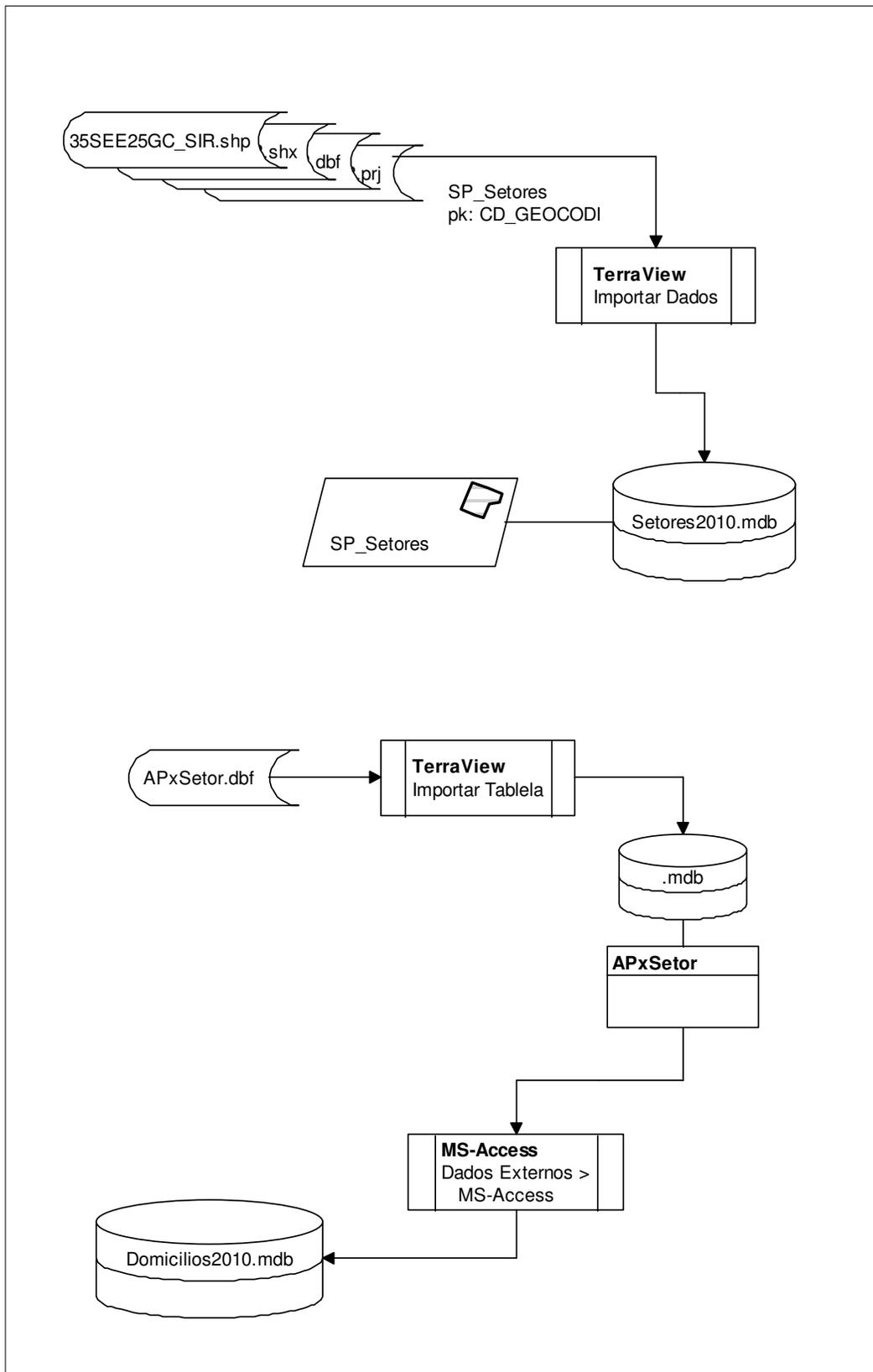
Os códigos V... dos campos originais foram substituídos por códigos mnemônicos

Feche a tabela e visualize sua estrutura com a opção **Modo Design**

Verifique os nomes dos campos, seus tipos e sobretudo suas descrições.

Observe que a tabela possui a chave primária **Controle**

Figura 1 – Importação de arquivos e conversão de dados (em .dbf)



Feche o Modo Design e Abra novamente a tabela.

Observe as colunas e seus conteúdos:

UF = 35	para todas as linhas (Estado de SP)
CodMun	com 5 dígitos (se concatenado com UF, corresponde ao código IBGE do Município)
AP	Área de Ponderação com 13 dígitos, os primeiros 7 correspondem ao código IBGE do Município
Controle	Chave primária da tabela – identifica o Domicílio
PesoS	fator de ponderação ou de expansão da amostra. Os três primeiros dígitos representam a parte inteira, os demais 13 a parte fracionária. Devemos criar nova coluna numérica (Double) com base nesta, para poder realizar cálculos com os dados do arquivo: Peso = Val(Esquerda(Peso,3)&”.”&Direita(Peso,13))
SitDom	indica se o domicílio está em área Urbana (1) ou Rural (2).

Demais campos serão vistos oportunamente, quando necessário.

Cada linha da tabela dos microdados **Amostra_Domicilios_35_RMSP** corresponde a um domicílio da amostra. O fator de expansão da amostra para gerar estimativas para o universo, a RMSP, é o Peso (PesoS).

Na dimensão espacial, apesar dos dados representarem domicílios individuais da amostra, não se sabe seus endereços, tampouco a quais setores censitários eles pertencem, apenas suas áreas de ponderação (AP) ou municípios (UF+Codmun).

(4)

Importe o arquivo **APxSetor.dbf** como tabela para o banco de dados **Domicilios2010.mdb**

As versões mais recentes do MS-Office não leem arquivos **.dbf** (DBase), portanto é necessário recorrer a algum outro software que faça a conversão para um formato que possa ser importado, por exemplo o **.xlsx** (MS-Excel), pelo uso do software LibreOffice.

Uma forma encontrada em sala de aula foi a de importar o arquivo APxSetor.dbf como Tabela Externa num banco de dados qualquer do TerraView. Em seguida, abrir o mesmo banco de dados (**.mdb**) com o MS-Access importar a tabela APxSetor para o banco de dados **Domicilios2010.mdb**.

A tabela indica as correspondências entre áreas de ponderação, setores censitários, distritos e municípios – verifique suas colunas.

Observe as colunas e seus conteúdos:

AREAP	Área de Ponderação, com 13 dígitos, como a AP da tabela Amostra_Domicilios_35_RMSP
SETOR	Setor Censitário, com 15 dígitos, os nove primeiros correspondem ao Distrito(coluna seguinte)
DISTRITO	
MUNICIPIO	

Elimine outras colunas eventualmente criadas pelo MS-Access.

Qual a relação entre as **áreas de ponderação (AP)** e os **setores censitários** ou **distritos** (a tabela não tem colunas para identificar o Setor Censitário ou mesmo o Distrito de um Domicílio, somente o Município e a Área de Ponderação)?

Para responder a questão acima haverá necessidade de se executar algumas queries, esquematizadas na Figuras 2 (versão simples) ou, alternativamente, na Figura 3 (versão elaborada).

Elabore a seguinte query (**Consulta1**):

```
Select AREAP, DISTRITO
      from APxSETOR
      group by AREAP, DISTRITO;
```

A consulta retorna os distintos pares (AREAP, DISTRITO). Observe que coluna SETOR foi desconsiderada, por isso o número de linhas da consulta deve ser menor do que a da tabela APxSETOR.

Figura 2 – Queries para verificar a relação entre AREAP e DISTRITO (versão simples)

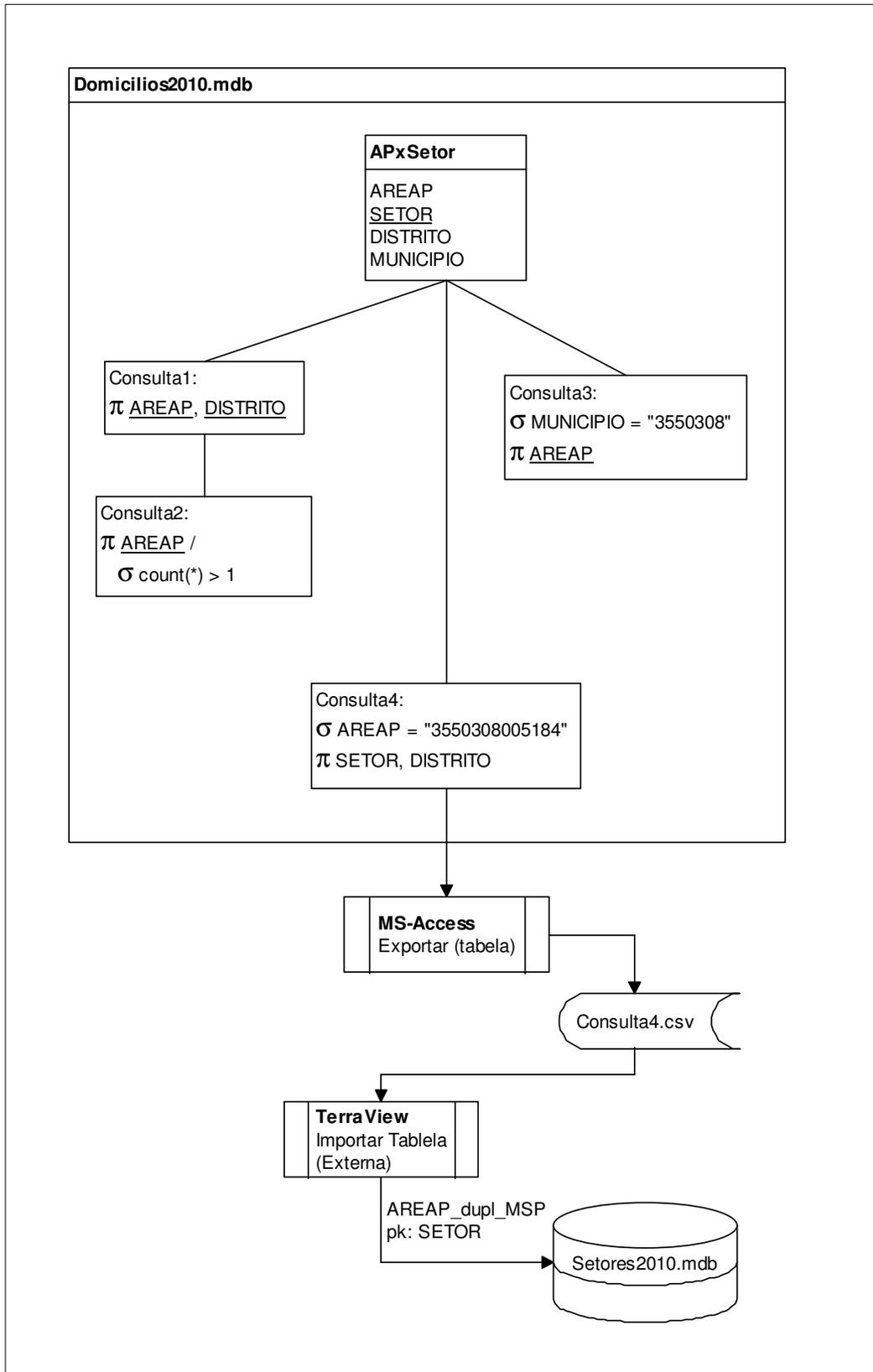
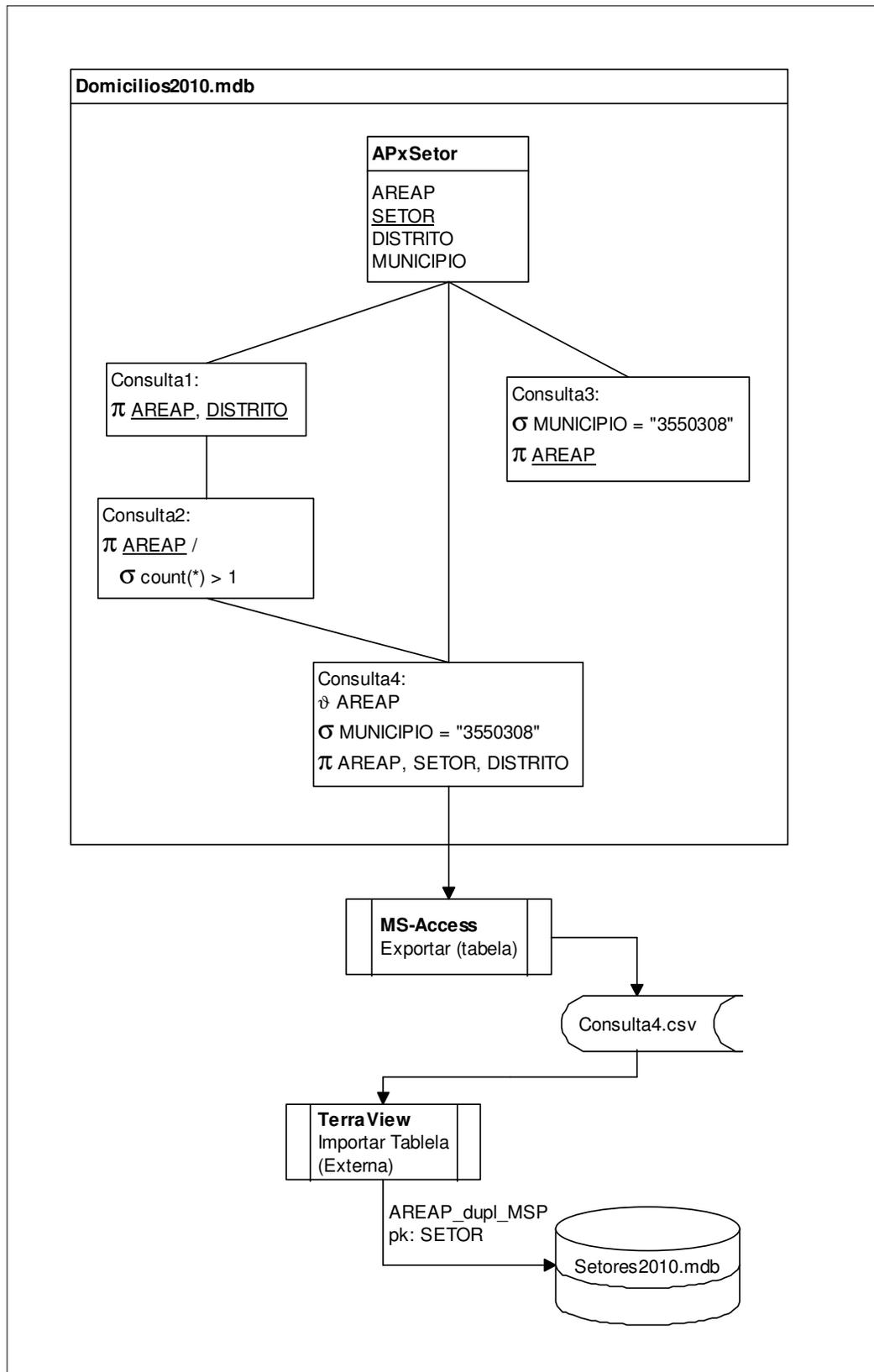


Figura 3 – Queries para verificar a relação entre AREAP e DISTRITO (versão elaborada)



Para verificar se há casos áreas de ponderação com setores em distritos distintos é necessária a query complementar (**Consulta2**):

```
Select AREAP
from Consulta1
group by AREAP having count(*) > 1;
```

Se o resultado for uma tabela vazia, toda Área de Ponderação está contida num único Distrito, para todos os municípios da RMSP, e os microdados podem ser agregados também por Distrito (não apenas por Município).

Senão, pode-se verificar para quais municípios da RMSP é possível agregar os microdados.

A query **Consulta2** retorna os seguintes resultados (reorganizados para melhor análise):

Município	AREAP			
3509205	3509205003002			
3510609	3510609005005	3510609005006		
3513009	3513009003010			
3515103	3515103003003			
3515707	3515707003003	3515707003005	3515707003006	3515707003007
	3515707003008			
3518800	3518800005018	3518800005022	3518800005024	3518800005028
	3518800005030			
3526209	3526209003001			
3528502	3528502003001	3528502003002		
3530607	3530607005002	3530607005005	3530607005006	3530607005007
	3530607005011	3530607005014	3530607005018	3530607005022
3539806	3539806003004			
3543303	3543303003001	3543303003002	3543303003005	3543303003006
3545001	3545001001001			
3547809	3547809005013	3547809005014	3547809005017	
3548708	3548708005017			
3550308	3550308005184			

Há 15 municípios com casos de áreas de ponderação em mais de um Distrito. Nesses municípios, à primeira vista, não seria possível agregar os microdados por seus distritos.

Nos demais 24 municípios da RMSP, por não haver casos como os verificados acima, ou por haver um único Distrito (a sede), essa agregação pode ser realizada. Dada sua complexidade (extensão territorial e heterogeneidade), seria desejável que o Município de São Paulo estivesse neste grupo!

O código IBGE do Município de São Paulo é 3550308, e está indicado na última linha da tabela acima.

Identifica-se um único caso (AREAP de número **3550308005184**) em 310 (conforme indica a query **Consulta3** abaixo), portanto merece investigação mais detalhada.

```
Select AREAP
      from APxSETOR
      where MUNICIPIO = "3550308"
      group by AREAP;
```

Pode-se buscar na tabela APxSETOR quais os setores correspondentes à AREAP de número **3550308005184** (Conforme a query **Consulta4** abaixo), exportá-los no formato **.csv (.txt)** para em seguida importá-los como Tabela Externa no BD **Setores2010.mdb** com o TerraView, confrontá-los com o Tema **SP_Setores**, e verificar se de fato há um erro na associação de um ou mais desses setores com a Área de Ponderação indicada, conforme esquematizado na Figura 2.

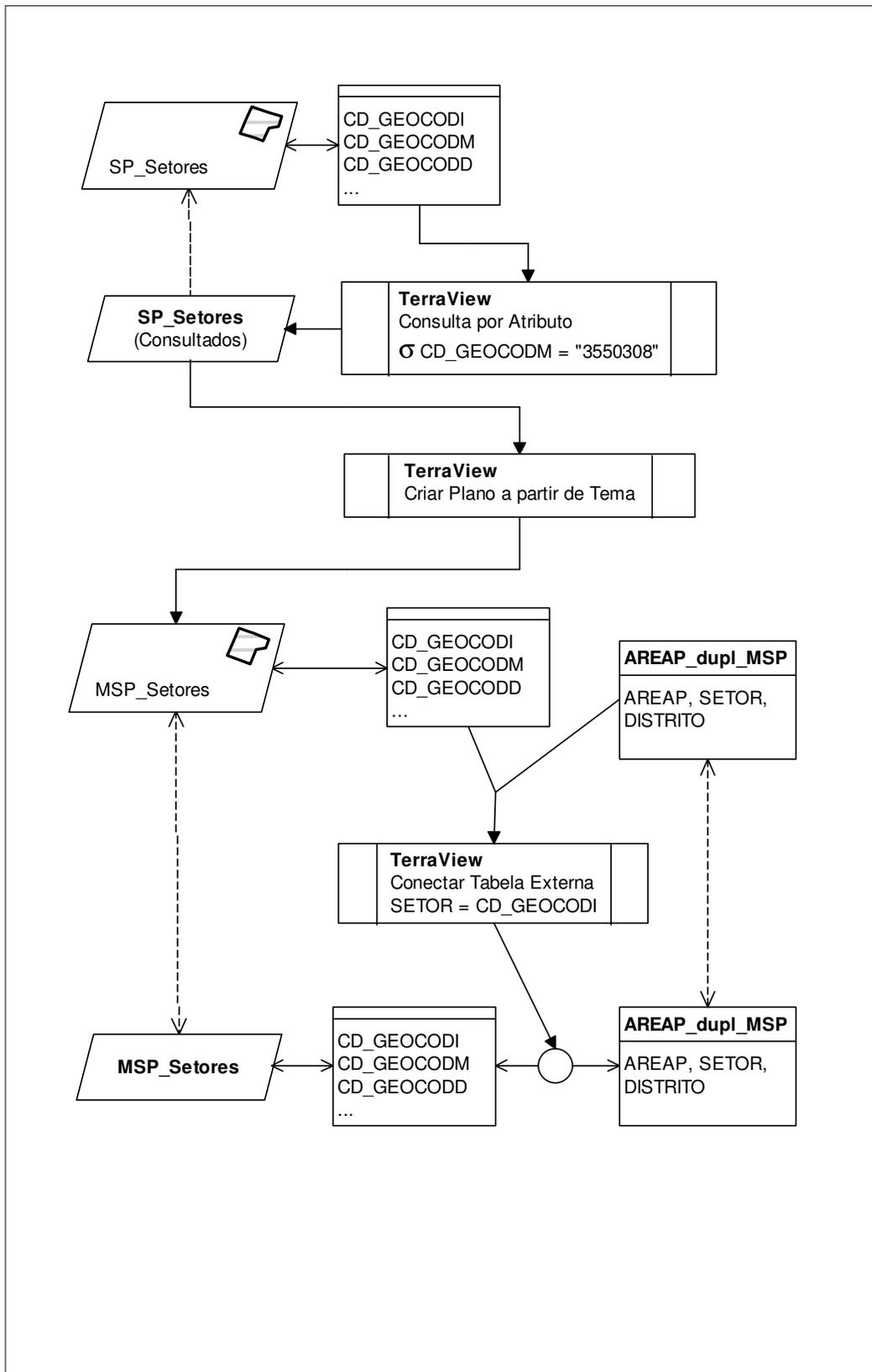
```
Select SETOR, DISTRITO
      from APxSETOR
      where AREAP = "3550308005184";
```

Caso houvessem mais casos de AREAP em mais de um DISTRITO do Município de São Paulo, a query **Consulta4** pode ser melhorada, conforme indicado na Figura 3, e escrita como segue.

```
Select AREAP, SETOR, DISTRITO
      from APxSETOR inner join Consulta2 on
      APxSETOR.AREAP = Consulta2.AREAP
      where MUNICIPIO = "3550308";
```

Em ambos casos, a query **Consulta4** é importada como Tabela Externa, com novo nome, por exemplo **AREAP_dupl_MSP**.

Figura 4 – Operações com o TerraView para análise das associações entre a(s) Área(s) de Ponderação e os Setores Censitários



(5)

O plano SP_Setores contém os polígonos dos setores censitários do Censo 2010 para todo o Estado de São Paulo.

Dado que se está analisando o caso particular do Município de São Paulo, sugere-se a criação de novo plano, MSP_Setores com apenas os setores do Município de São Paulo, por uma Consulta por Atributos, conforme indicado esquematicamente na Figura 4.

A tabela externa AREAP_dupl_MSP deve ser conectada a esse novo plano/tema MSP_Setores, pelos campos SETOR e CD_GEOCODI, respectivamente, conforme esquematizado na Figura 4.

Cabe observar que a grande maioria dos campos da tabela externa, visíveis, uma vez conectada ao tema MSP_Setores, apresenta vazios. Isso porque a conexão foi feita para apenas uma das áreas de ponderação, a que possivelmente está com erro.

Pode-se visualizar os setores correspondentes, pode-se executar uma Consulta por Atributos, com a condição:

SETOR is not null.

Uma análise visual irá indicar se há erro. O Setor Censitário de número “355030841000104”, corresponde à Área de Ponderação “3550308005184” e ao Distrito “355030841” (Jaguapé), enquanto aos seus confrontantes, correspondem a Áreas de Ponderação ou Distrito opostos:

Setor	Área de Ponderação	Distrito
355030841000046	3550308005108	355030841 (Jaguapé)
355030841000047	3550308005108	355030841 (Jaguapé)
355030841000104	3550308005184	355030841 (Jaguapé)
355030867000104	3550308005184	355030867 (Rio Pequeno)
355030867000068	3550308005184	355030867 (Rio Pequeno)

Conclusão, a Área de Ponderação do Setor Censitário número “355030841000104” está incorreta na tabela APxSETOR, e deve ser alterada de “3550308005184” para “3550308005108”, no BD **Domicilios2010.mdb**.