

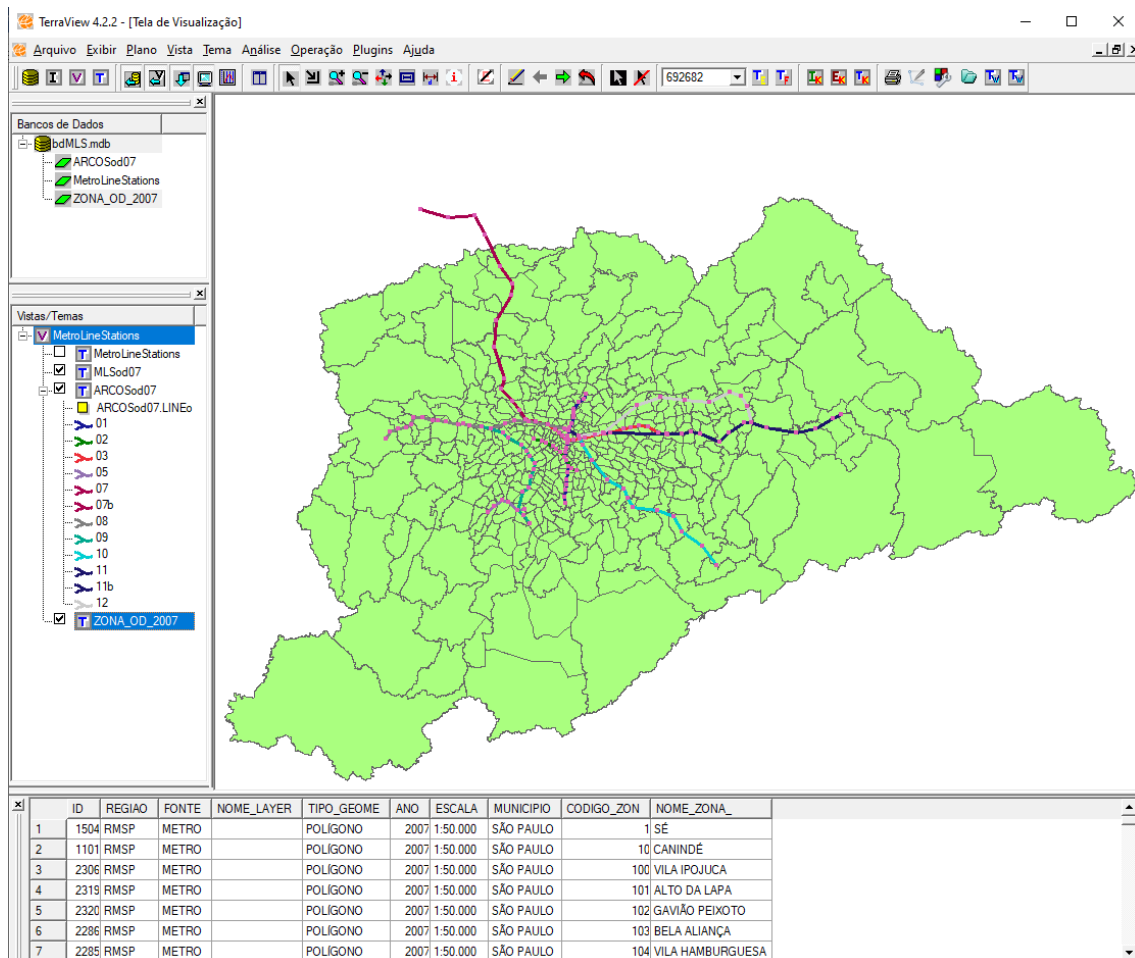
## Exercício 8 – geoprocessamento com software SIG e BD MS-Access

Este exercício pode ser feito como extensão do Aula 7 com Ms-Access, pois requer o uso das estações/linhas em uso antes da Pesquisa OD 2007 do Metrô. Parte dele será feito com software gerenciador de bancos de dados, p.ex., o MS-Access; e o restante no software SIG de preferência. Pode ser feito em Geopackage ou Libre Office Base (com as devidas alterações da sintaxe).

O objetivo do exercício é o de estudar a influência da rede metroferroviária nos tempos das viagens, a trabalho, por transporte coletivo, com destino ao centro da cidade de São Paulo (zonas 1 a 37), a partir das demais regiões. Sugere-se a construção de isócronas de até 30'; mais de 30' a 45', mais de 45' a 60'; mais de 60' a 75'; mais de 75' a 90; mais de 90' a 105'; mais de 105' a 120'; e mais de 120'. Sabe-se que o tempo médio de viagem nesse contexto foi estimado em 74' (próximo de 75').

Em se tratando da Pesquisa OD 2007, há de se ter o mapa do zoneamento criado para tal pesquisa – no caso, os arquivos **ZONA\_OD\_2007.shp** (shx, dbf e prj). O sistema de coordenadas é o da projeção UTM zona 23, hemisfério Sul, sobre o Datum SAD 69. Há de se recordar que o sistema de projeção do arquivo MetroLineStation.csv é o da projeção UTM zona 23, hemisfério Sul, mas o Datum é o WGS 84.

Figura 1 – zonas da pesquisa OD 2007 do Metrô



A Figura 1 ilustra o zoneamento da Pesquisa OD 2007 sob as estações e trechos entre as estações correspondentes. O campo indicado como ID na tabela de dados do zoneamento não tem utilidade e pode ser eliminado, dado que a chave primária desta tabela é **CODIGO\_ZON**.

Os dados agregados, fornecidos pela Prof. Dra. Andreina Nigriello, também funcionária da Companhia do Metropolitano de São Paulo – Metrô, estão no arquivo de planilha eletrônica **coletivo\_picos\_Centro.xls**.

As figuras 2 e 3 mostram extratos de pastas dessa planilha eletrônica, para dois dos intervalos do pico da manhã, respectivamente, 05h30' a 06h30' e 08h30' a 09h30'.

Figura 2 – Número de viagens e tempos médios para o Centro das 05h30' às 06h30'

6	7	Zonas de Destino - Centro									
		1		2		3		4		5	
8	Zonas de Origem	Viagens	Tempo	Viagens	Tempo	Viagens	Tempo	Viagens	Tempo	Viagens	Tempo
9	1	-	.	-	.	-	.	-	.	-	.
10	3	-	.	-	.	-	.	-	.	-	.
11	5	47	45	-	.	-	.	-	.	-	.
12	6	-	.	-	.	-	.	-	.	-	.
13	7	-	.	-	.	-	.	-	.	-	.
14	8	-	.	-	.	-	.	-	.	24	.
15	10	-	.	-	.	-	.	-	.	-	.
16	11	12	60	-	.	7	45	-	.	-	.
17	12	-	.	-	.	-	.	-	.	-	.
18	13	-	.	-	.	-	.	7	50	-	.
19	15	-	.	-	.	-	.	-	.	-	.
20	16	-	.	-	.	-	.	-	.	-	.
21	18	-	.	-	.	-	.	-	.	-	.
22	19	-	.	94	40	-	.	-	.	-	.
23	20	-	.	-	.	-	.	-	.	-	.
24	21	-	.	-	.	-	.	58	60	-	.
25	22	-	.	-	.	-	.	-	.	-	.
26	23	-	.	-	.	-	.	-	.	-	.
27	24	-	.	-	.	-	.	-	.	-	.
28	27	-	.	-	.	-	.	-	.	-	.
29	28	-	.	-	.	-	.	-	.	-	.
30	30	-	.	-	.	-	.	60	55	-	.
31	31	-	.	-	.	-	.	-	.	64	.
32	33	-	.	-	.	-	.	-	.	-	.
33	35	-	.	-	.	-	.	-	.	-	.
34	36	-	.	-	.	-	.	-	.	-	.
35	37	-	.	-	.	-	.	-	.	-	.
36	41	-	.	-	.	84	30	-	.	-	.
37	42	-	.	-	.	-	.	-	.	-	.
38	43	-	.	-	.	-	.	-	.	-	.
39	44	-	.	-	.	-	.	-	.	-	.
40	46	176	100	-	.	-	.	-	.	-	.
41	49	72	45	-	.	-	.	-	.	-	.
42	50	-	.	-	.	-	.	-	.	-	.
43	51	-	.	37	25	-	.	-	.	-	.
44	52	-	.	-	.	-	.	-	.	-	.
45	54	-	.	-	.	-	.	-	.	-	.

diarioCentro

0530-0630 Centro

0630-0730 Centro

730-830 Centro

0830-0930 Centro

Figura 3 – Número de viagens e tempos médios para o Centro das 08h30' às 09h30'

6	Zonas de Destino - Centro									
	7	1		2		3		4		5
8		Zonas de Origem	Viagens	Tempo	Viagens	Tempo	Viagens	Tempo	Viagens	Tempo
9	1	20	25	-	-	-	-	25	30	-
10	2	-	-	-	-	-	-	-	-	-
11	3	-	-	-	-	-	-	-	-	-
12	4	67	30	-	-	-	-	54	35	-
13	5	-	-	-	-	47	20	58	15	-
14	6	53	20	-	-	-	-	-	-	-
15	7	-	-	36	30	-	-	-	-	90
16	8	-	-	-	-	-	-	-	-	23
17	9	-	-	-	-	-	-	-	-	19
18	10	6	30	-	-	-	-	-	-	-
19	11	-	-	-	-	-	-	-	-	-
20	13	-	-	-	-	-	-	15	45	-
21	15	-	-	36	30	-	-	-	-	-
22	17	-	-	-	-	-	-	-	-	21
23	18	-	-	-	-	-	-	-	-	50
24	19	-	-	-	-	-	-	-	-	-
25	20	-	-	118	40	-	-	-	-	-
26	21	163	50	-	-	-	-	-	-	47
27	22	74	25	-	-	60	15	-	-	60
28	23	-	-	-	-	-	-	-	-	-
29	24	79	20	-	-	-	-	-	-	-
30	25	-	-	113	50	-	-	67	60	-
31	27	76	30	-	-	-	-	-	-	-
32	28	50	18	31	25	31	45	-	-	-
33	29	-	-	43	20	-	-	-	-	31
34	30	60	20	-	-	-	-	-	-	-
35	31	64	20	-	-	-	-	-	-	-
36	32	-	-	-	-	-	-	-	-	-
37	33	4	45	-	-	-	-	4	40	-
38	34	13	15	-	-	18	40	-	-	-
39	35	-	-	-	-	-	-	-	-	-
40	36	-	-	-	-	-	-	-	-	135
41	37	-	-	-	-	-	-	-	-	-
42	39	22	40	-	-	-	-	-	-	-
43	40	46	30	40	45	-	-	-	-	-
44	41	-	-	-	-	-	-	-	-	-
45	43	83	20	-	-	-	-	-	-	-

A primeira planilha, não exibida nas figuras, corresponde ao total das viagens diárias e não serve para o pico da manhã.

Para se ter os totais para o pico da manhã, há necessidade de se somar as viagens e os tempos médios desses quatro intervalos de tempo.

Há de se observar que, as zonas sem registro de origem de viagem, em um dado intervalo de tempo, não tem os correspondentes destinos associados a ela, ou seja, não está listada na coluna Zonas de Origem.

No extrato do intervalo das 05h30' às 06h30', da Figura 2, vê-se que faltam, na coluna Zonas de Origem, as zonas 2, 4, 9, 14, 17, 25, 26, 29, 32, 34, 38 a 40 ...

E, no extrato do intervalo das 08h30' às 09h30', da Figura 3, vê-se que faltam, as zonas 12, 14, 16, 26, 38, 42, ...

A “solução” manual de criar uma planilha para totalização de viagens e recálculo de tempos médios consome tempo e está sujeita a erro.

Uma solução mais eficiente é a de carregar os dados dessas planilhas em um BD, por exemplo, o MS-Access. Isso pode ser feito diretamente do arquivo da planilha, em formato XLS, ou a partir de um arquivo intermediário em formato CSV – *comma separated values*.

De qualquer forma, em cada planilha, apenas a primeira e as duas últimas colunas são relevantes para o cálculo do número de viagens e, sobretudo, do tempo médio para o pico da manhã, conforme ilustrado na Figura 4.

Figura 4 – Totais de viagens e tempos médios para o Centro das 05h30' às 06h30'

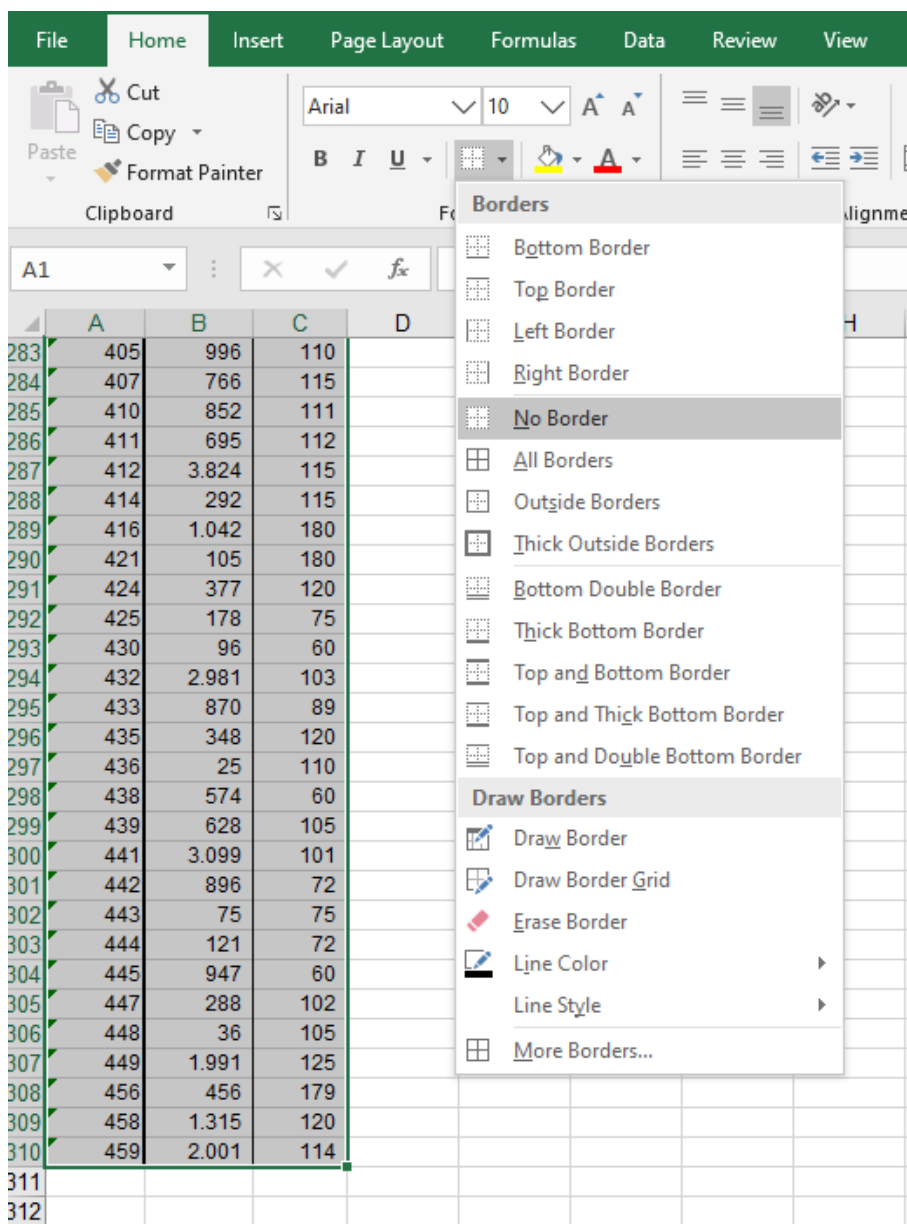
	A	BS	BT	BU	BV	BW	BX	BY	BZ
1	Número de Viagens								
2	Região Metropolitana								
3	Companhia do Metrô								
4	Pesquisa Origem e Destino								
5									
6									
7	Zonas de Origem	5	36		37		Total Centro		
8		Tempo	Viagens	Tempo	Viagens	Tempo	Viagens	Tempo	
35	37	.	-	.	-	.	17	30	
36	41	.	-	.	-	.	231	33	
37	42	.	-	.	-	.	26	135	
38	43	90	-	.	-	.	83	90	
39	44	.	-	.	-	.	62	60	
40	46	.	-	.	-	.	176	100	
41	49	.	-	.	-	.	72	45	
42	50	.	-	.	-	.	49	40	
43	51	.	-	.	-	.	84	36	
44	52	.	-	.	-	.	42	50	
45	54	.	-	.	-	.	46	40	
46	55	.	-	.	-	.	65	25	
47	57	60	-	.	-	.	127	60	
48	59	.	-	.	-	.	169	40	
49	61	.	-	.	-	.	197	43	
50	62	.	-	.	-	.	28	30	
51	63	.	-	.	-	.	3	60	
52	64	.	-	.	-	.	331	33	
53	65	.	-	.	-	.	37	60	
54	68	.	-	.	-	.	59	70	
55	69	.	-	.	-	.	56	40	
56	71	.	-	.	-	.	41	60	
57	73	.	-	.	-	.	47	40	
58	76	.	-	.	-	.	99	35	
59	78	.	-	.	-	.	34	15	
60	80	.	-	.	-	.	36	50	
61	82	.	-	.	-	.	91	120	
62	83	.	-	.	-	.	72	20	
63	84	.	-	.	-	.	25	38	
64	86	.	-	.	-	.	142	28	
65	91	.	-	.	-	.	5	35	
66	92	.	-	.	-	.	35	45	
67	96	.	-	.	-	.	58	40	
68	97	.	-	.	-	.	18	80	
69	101	.	78	30	-	.	78	30	
70	102	.	-	.	-	.	8	45	
71	104	.	-	.	-	.	61	99	

Também as linhas relevantes para sua inserção em uma tabela de banco de dados são apenas aquelas que contêm dados. Não há totais em branco (viagens ou tempo zerados) e isso favorece tal inserção.

Há de se eliminar não apenas as linhas e colunas desnecessárias (ou inconvenientes), como também se preocupar com a “limpeza” de cada planilha. Qualquer formatação de célula (veja Figura 5), por exemplo, as linhas usadas para facilitar a leitura dos dados introduz ruído nos dados e resulta em erros durante o carregamento dos dados no BD.

Remover as oito primeiras linhas (cabecalho e nomes dos campos) e também a última (totais das viagens às zonas (1 a 37) do Centro). Remover da 2ª coluna (B) até a antepenúltima. Deixar apenas a coluna A e as duas últimas, renumeradas como B e C após terem sido eliminadas as intermediárias. Remover a formatação das células. Nas quatro planilhas que serão carregadas no BD.

Figura 5 – Limpeza das planilhas – em destaque a remoção da formatação



Em cada planilha, inserir uma linha no topo com os novos nomes das colunas: Origem, Viagens e Tempo.

O carregamento dos dados, seja de um único intervalo, seja de um número qualquer de intervalos, deve ser feito numa única tabela, p.ex. PMCentro. Neste caso, são quatro planilhas a carregar: 0530-0630 Centro, 0630-0730 Centro, 0730-0830 Centro e 0839-0930 Centro.

Criar um novo BD e importar a 1ª planilha numa nova tabela – figuras 6 a 14. Acrescentar os dados das demais planilhas à essa tabela, tendo o cuidado de não esquecer nenhuma e de não carregar mais de uma vez uma planilha.

Figura 6 – Importação de dados de arquivo de planilhas (.xls / .xlsx) no MS-Access

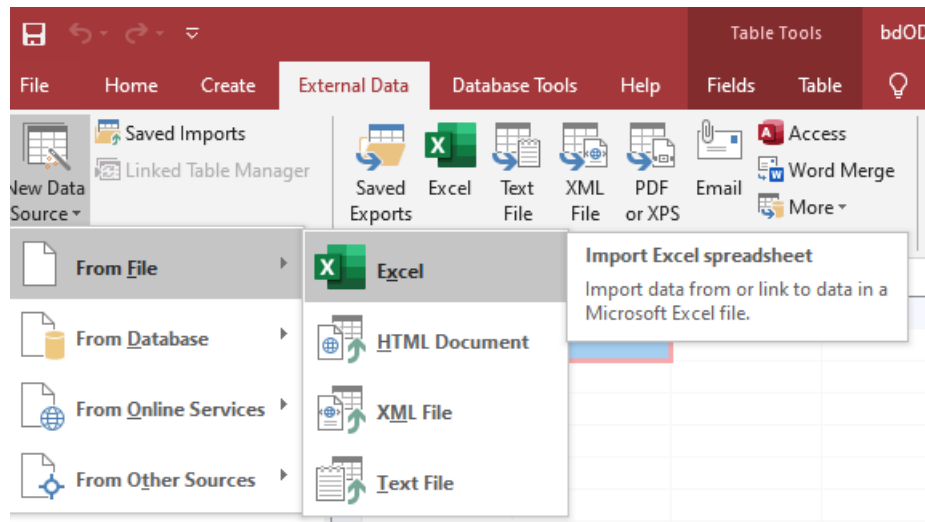


Figura 7 – Importação da primeira planilha como uma nova tabela

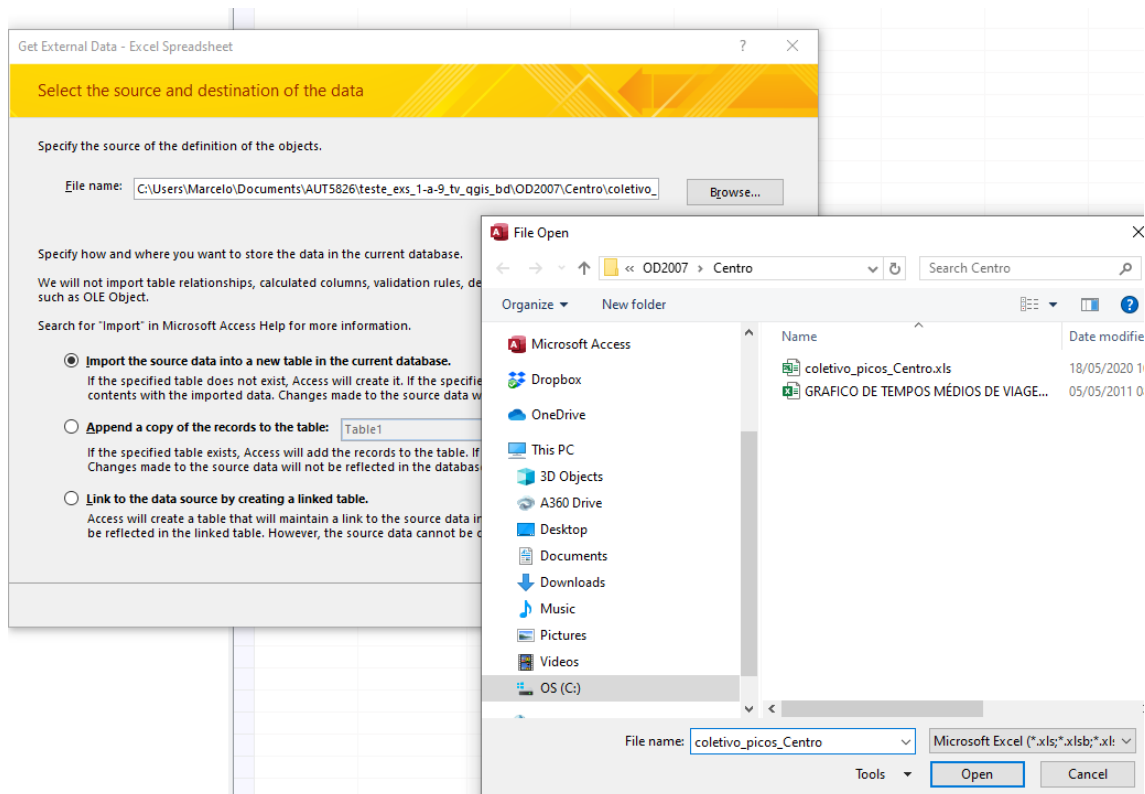


Figura 8 – A primeira planilha a importar é a 0530-0630 Centro

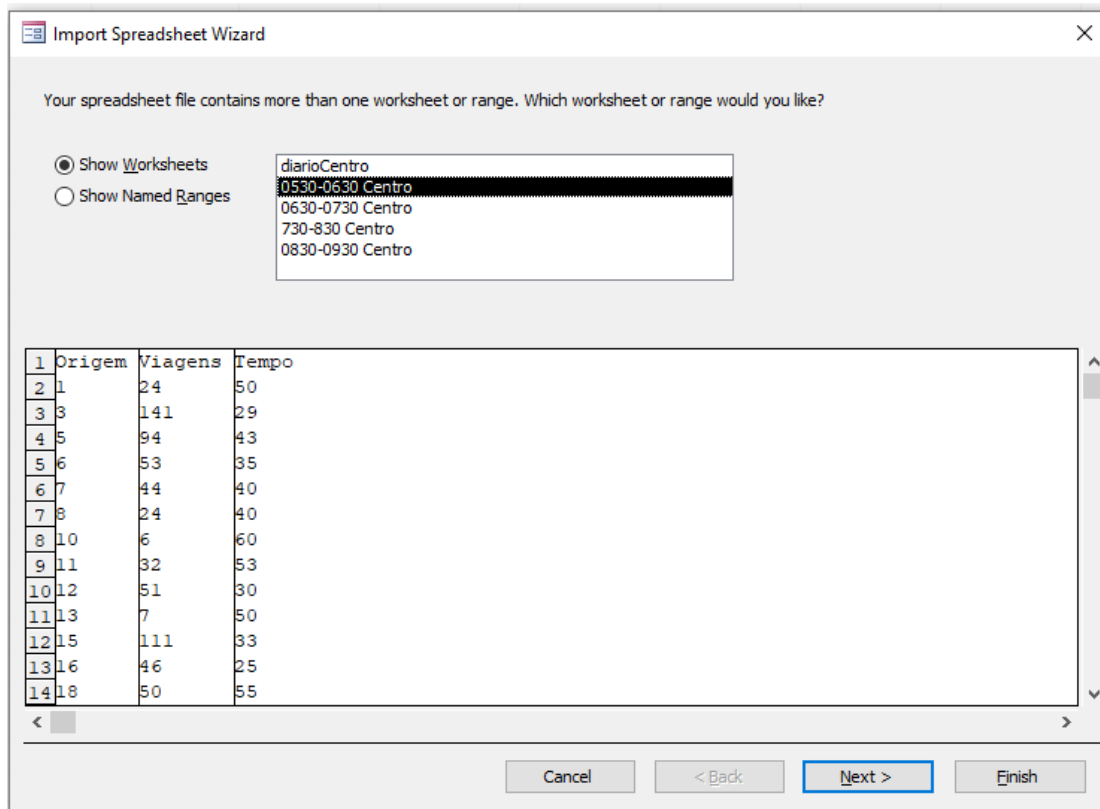
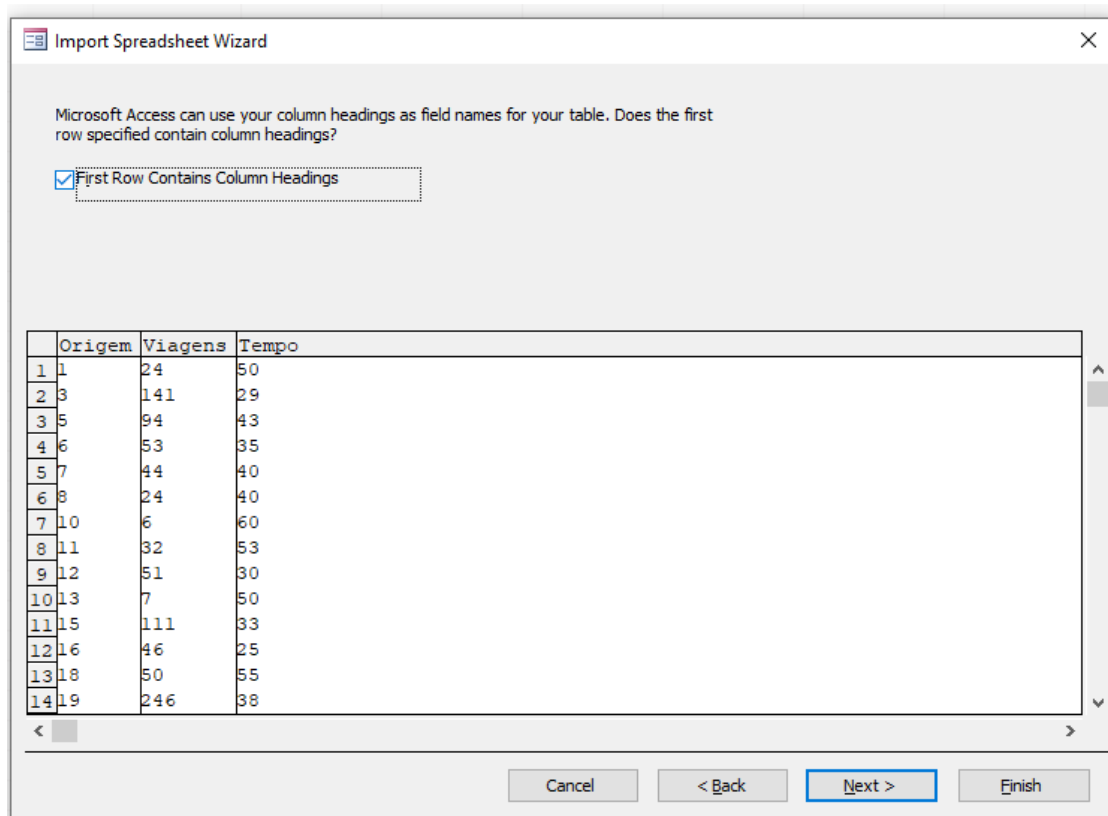


Figura 9 – Primeira linha deve conter os nomes dos campos



É importante em todas as planilhas os nomes das colunas sejam os mesmos.

Figura 10 – Configuração da coluna Origem

You can specify information about each of the fields you are importing. Select fields in the area below. You can then modify field information in the 'Field Options' area.

Field Options

Field Name:  Data Type:

Indexed:    Do not import field (Skip)

	Origem	Viagens	Tempo
1	1	24	50
2	3	141	29
3	5	94	43
4	6	53	35
5	7	44	40
6	8	24	40
7	10	6	60
8	11	32	53
9	12	51	30
10	13	7	50
11	15	111	33
12	16	46	25
13	18	50	55
14	19	246	38

<

O tipo de dado da coluna Origem deve ser compatível com o tipo de dado da coluna correspondente no *layer* das zonas OD 2007 no software SIG escolhido. Na Figura 10, está como Short Text, pois é assim que o software Terra View exibe a coluna ZONA\_COD após ela ser indicada como coluna de ligação de dados. Se, por outro lado, o software SIG enxergar a coluna ZONA\_COD como Integer, então deve-se selecionar o tipo de dados como Integer..



Figura 11 – Configuração da coluna Viagens

You can specify information about each of the fields you are importing. Select fields in the area below. You can then modify field information in the 'Field Options' area.

Field Options

Field Name:  Data Type:

Indexed:   Do not import field (Skip)

	Origem	Viagens	Tempo
1	1	24	50
2	3	141	29
3	5	94	43
4	6	53	35
5	7	44	40
6	8	24	40
7	10	6	60
8	11	32	53
9	12	51	30
10	13	7	50
11	15	111	33
12	16	46	25
13	18	50	55
14	19	246	38

Buttons: Cancel, < Back, Next >, Finish

As colunas Viagens e Tempo devem ser configuradas como de valores numéricos. No caso ilustrado na Figura 11, elas aparecem inicialmente como Double (um exagero), e podem ser alteradas para Single. Se não fizer agora isso pode ser corrigido mais tarde, clicando-se com o botão direito na tabela e selecionando Design View.

Figura 12 – Sem chave primária

Microsoft Access recommends that you define a primary key for your new table. A primary key is used to uniquely identify each record in your table. It allows you to retrieve data more quickly.

Let Access add primary key.

Choose my own primary key. Origem

No primary key.

	Origem	Viagens	Tempo
1	1	24	50
2	3	141	29
3	5	94	43
4	6	53	35
5	7	44	40
6	8	24	40
7	10	6	60
8	11	32	53
9	12	51	30
10	13	7	50
11	15	111	33
12	16	46	25
13	18	50	55
14	19	246	38

Cancel < Back Next > Finish

A tabela não possui chave primária, caso contrário não seria possível acrescentar os dados das demais planilhas.

Figura 13 – Nomear a tabela como PMCentro

That's all the information the wizard needs to import your data.

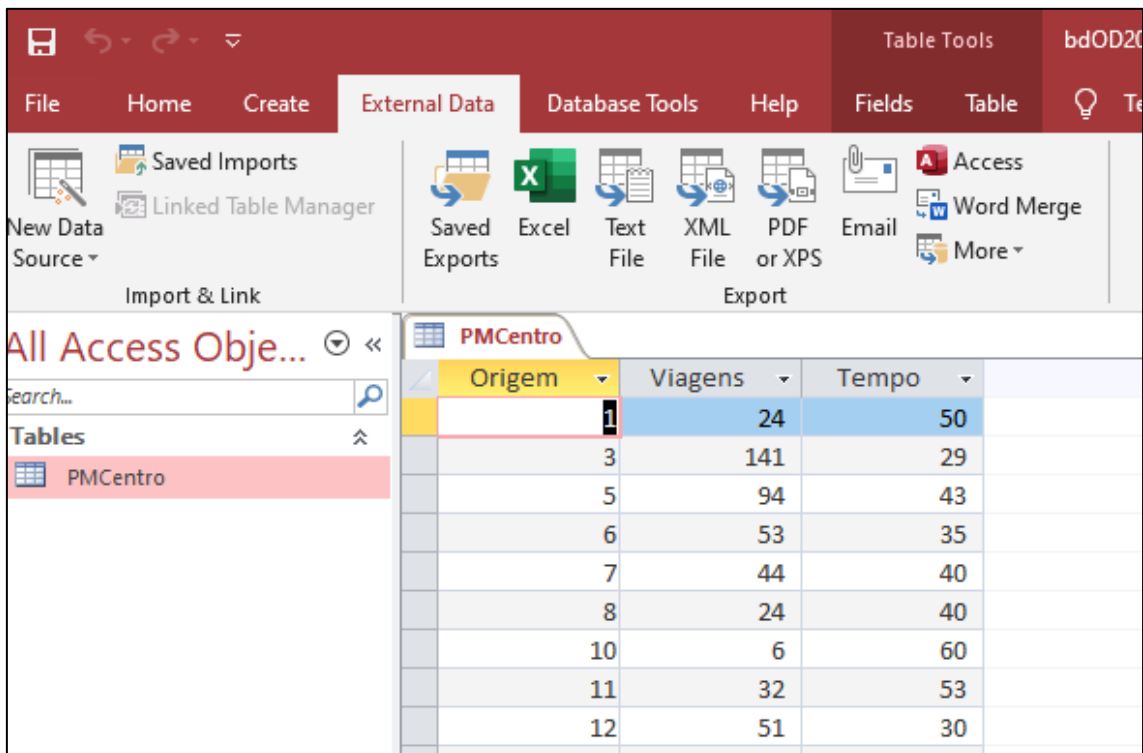
Import to Table:  
PMCentro

I would like a wizard to analyze my table after importing the data.

Cancel < Back Next > Finish

A tabela deve ser renomeada para não ficar com o nome da primeira planilha carregada.

Figura 14 – Tabela criada com os dados da primeira planilha

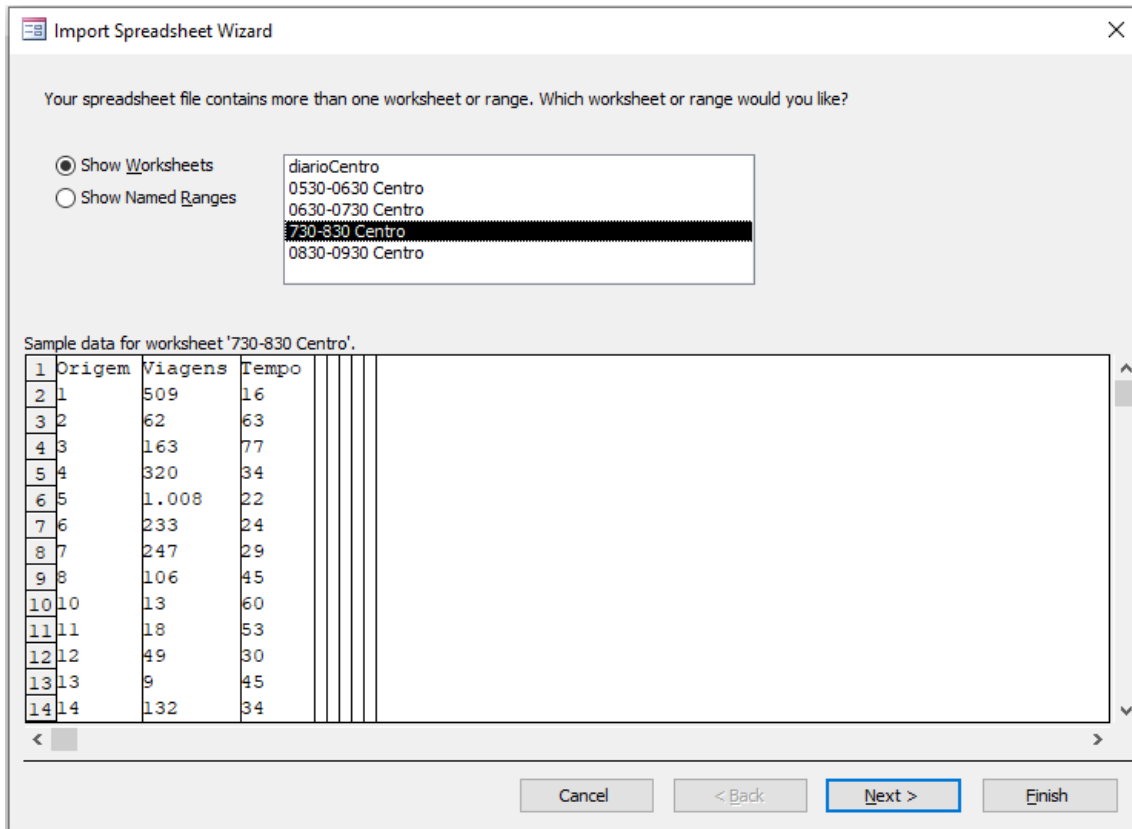


The screenshot shows the Microsoft Access interface with the 'External Data' ribbon selected. The ribbon contains options for 'Import & Link' (New Data Source, Saved Imports, Linked Table Manager) and 'Export' (Saved Exports, Excel, Text File, XML File, PDF or XPS, Email, Access, Word Merge, More). The 'Table Tools' ribbon is also visible, showing 'Fields' and 'Table' tabs. The 'All Access Objects' task pane on the left shows the 'Tables' section with 'PMCentro' selected. The main window displays the 'PMCentro' table with the following data:

Origem	Viagens	Tempo
1	24	50
3	141	29
5	94	43
6	53	35
7	44	40
8	24	40
10	6	60
11	32	53
12	51	30

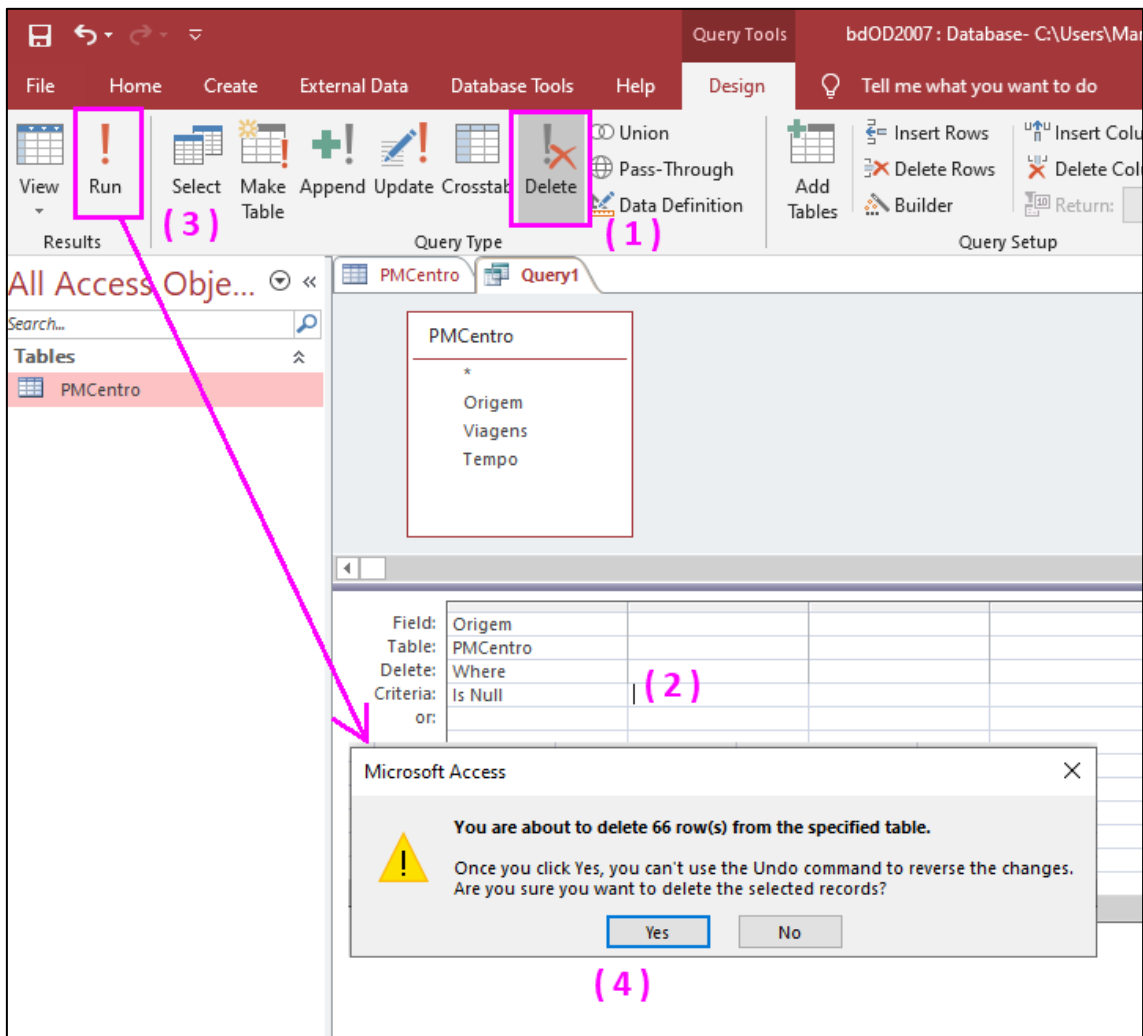
Após a criação da tabela PMCentro, com os dados da primeira planilha, deve se acrescentar os dados das demais a ela. Em cada caso, ao invés de selecionar uma nova tabela, indicar que se deseja acrescentar os dados à tabela existente (veja a segunda opção – *append a copy of the records to the table* - da Figura 7).

Figura 15 – Sinal de haver ruído numa das planilhas



Pode ocorrer de, em algum momento ocorrer algum erro ou sinalização de haver algo errado com uma das planilhas. A Figura 15 ilustra um exemplo disso. Parece haver resquício de formatação em colunas sem dados. Se isso acontecer, interrompa a importação dos dados, abra a planilha e (neste caso exemplo) remova colunas à direita das colunas com dados, de modo a eliminar o problema.

Figura 16 – Query para eliminar linhas vazias



Outro problema desse tipo é aparecerem, após o carregamento dos dados, linhas vazias na tabela. Isso pode ser corrigido nas planilhas, pela remoção de um punhado de linhas após a última linha de dados, e/ou no BD com a eliminação das linhas vazias.

A sequência de queries a rodar sobre esta única tabela deve funcionar, independentemente de quantas planilhas tiveram seus dados acrescentados a ela. Não deve importar se as zonas de origem das viagens se repetem, pois seus valores são diferentes em cada intervalo em que foram registradas viagens originando delas.

Deseja-se os tempos médios de cada zona Origem das viagens para o Centro (zonas 1 a 37) no pico da manhã. Trata-se da média ponderada dos tempos médios em cada intervalo com viagens registradas para cada zona Origem, o que demanda duas queries.

#### Query 1:

```
select Origem, sum(Viagens) as TV, sum(Viagens * Tempo) as VT
from PMCentro
group by Origem ;
```

e

**Query 2:**

```
select Origem, VT/TV as Tempo  
from Query1 ;
```

ou, em Design Mode:

**Query 1:**

Field:	Origem	TV:Viagens	VT: [Viagens]*[Tempo]
Table:	PMCentro	PMCentro	PMCentro
Total:	Group By	Sum	Sum
Sort:			
Show:	[ V ]	[ V ]	[ V ]
Criteria:			
Or:			

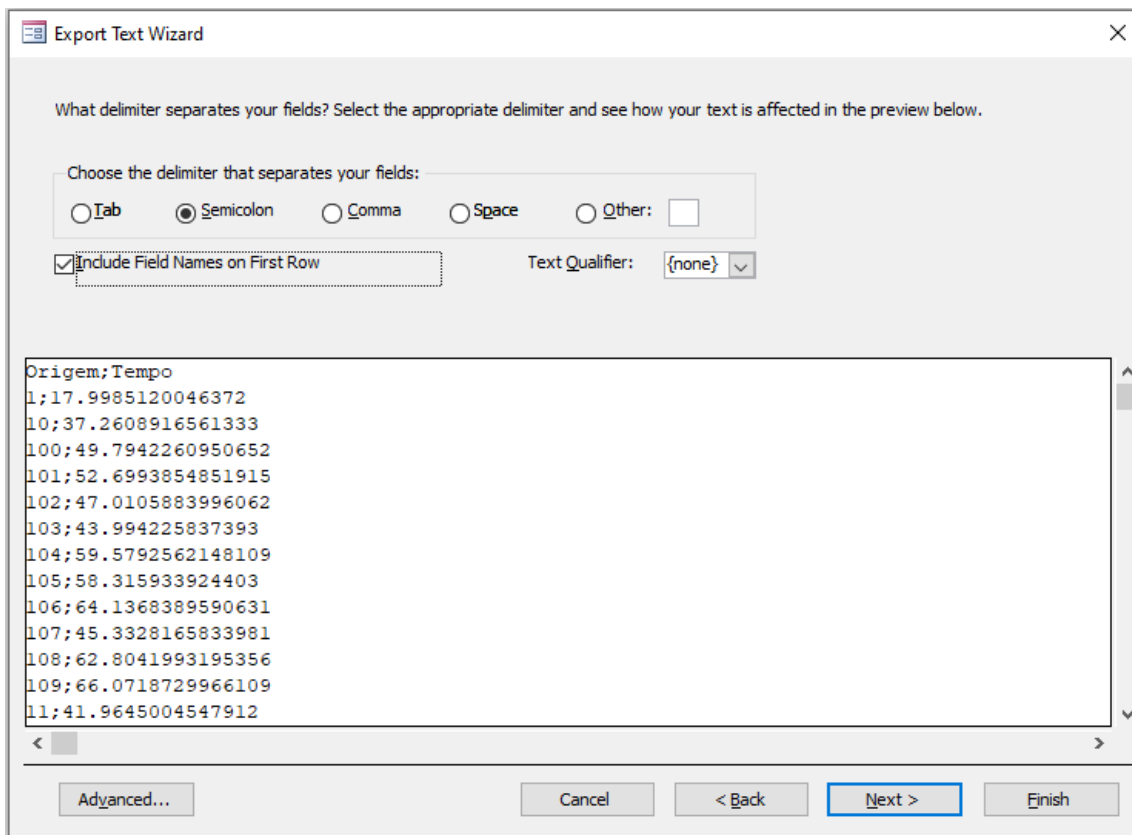
e

**Query2:**

Field:	Origem	Tempo:[VT]/[TV]
Table:	Query1	Query1
Sort:		
Show:	[ V ]	[ V ]
Criteria:		
Or:		

O resultado da [Query2](#) pode ser exportado em formato CSV para ser carregado no software SIG escolhido e conectado ao *layer* das zonas da Pesquisa OD 2007, conforme ilustra a Figura 17.

Figura 17 – exportação da Query 2 como PMCentro.csv



A Figura 17 ilustra exemplo de exportação em formato (.txt / .csv) para software SIG. Deve-se incluir os nomes das colunas na primeira linha e indicar como separador o ponto e vírgula (;). No exemplo, os dados estão sendo produzidos para serem carregados no BD do software SIG Terra View 4. Por isso o separador decimal é o ponto (.) e os dados da coluna Origem, definidos como do tipo texto, não possuem delimitadores ( Text Qualifier: {none} )

Após a importação desses dados e a sua vinculação com as zonas da Pesquisa OD 2007, o mapa pode ser “colorido” com base nos tempos médios de viagens de transporte coletivo ao Centro na escala definida inicialmente para as isócronas (até 30’, 30-45’, 45-60’, 60-75’, 75-90’, 90-120’ e mais de 120’) – conforme ilustrado na Figura 18.

Figura 18 – Isócronas de tempo médio ao Centro no pico da manhã, sobre as zonas da pesquisa OD 2007 do Metrô – toda a RMSP

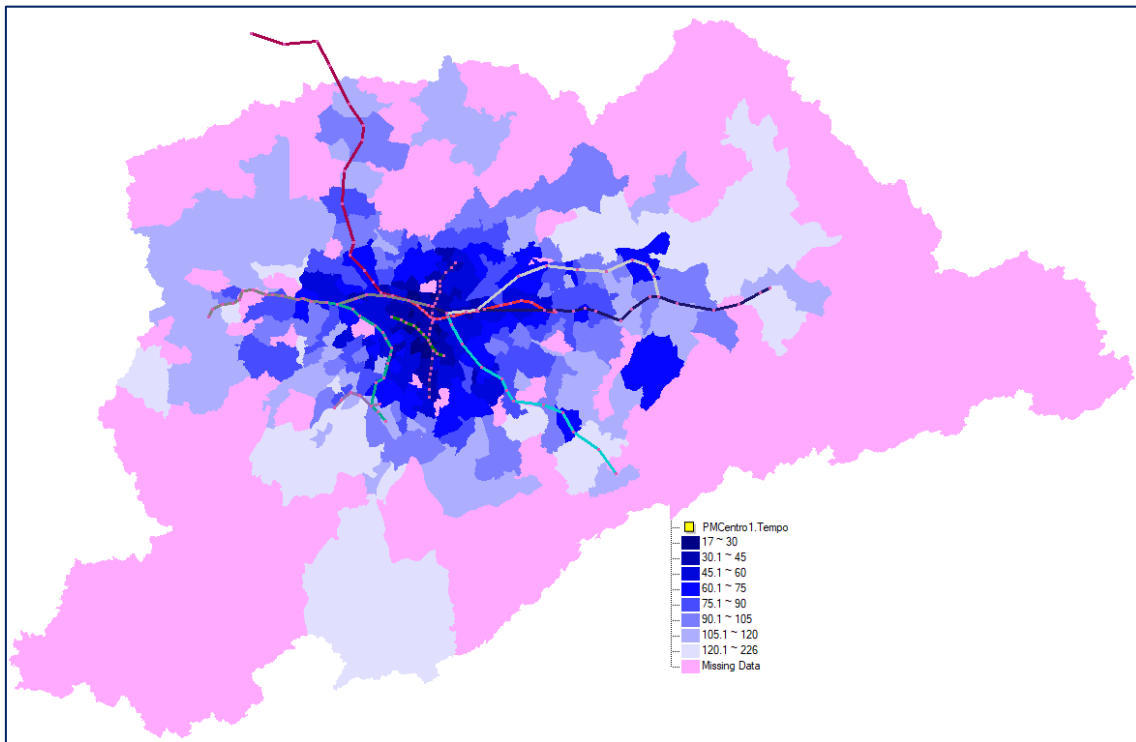
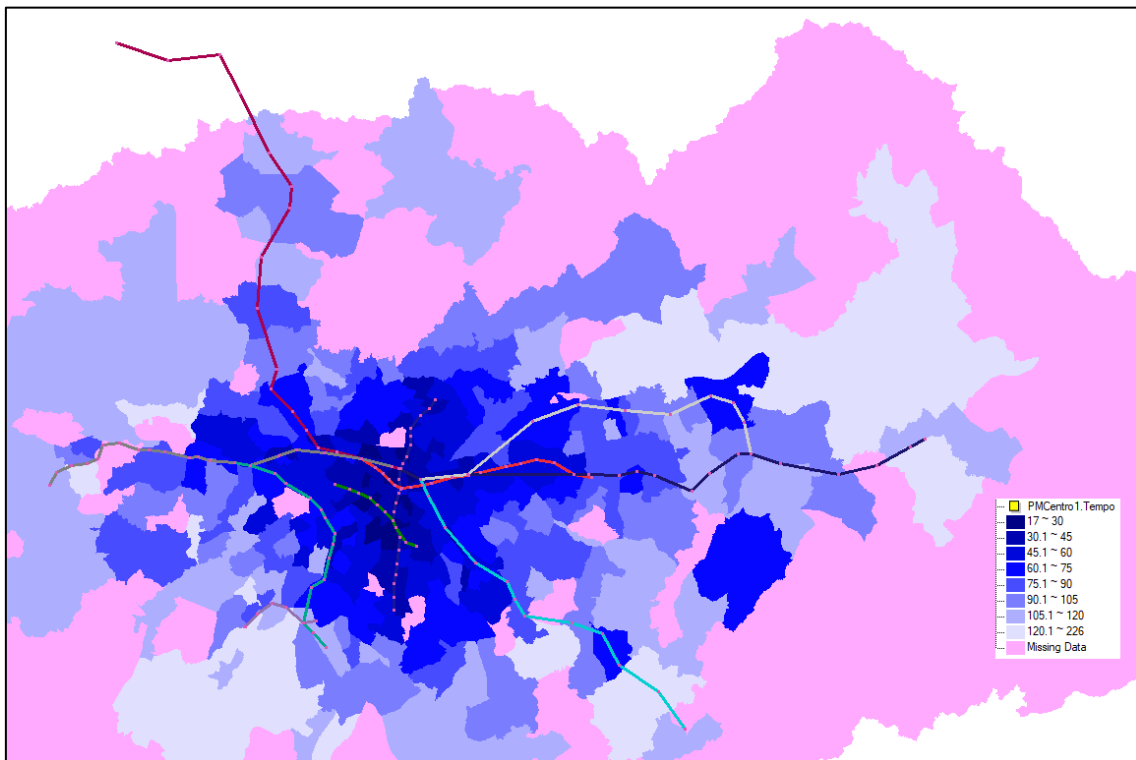


Figura 19 – Isócronas de tempo médio ao Centro no pico da manhã, sobre as zonas da pesquisa OD 2007 do Metrô – zoom nas linhas do sistema metroferroviário



Pode-se analisar visualmente o efeito das estações (pontos de acesso à rede metroferroviária), com o auxílio dos segmentos representativos das linhas (Figura 19).



Cabe observar que, em 2007, não havia integração tarifária nos moldes que existe atualmente, em especial entre os modos ônibus e metrô; ou entre ônibus e trem, mas, também entre linhas operadas pelo Metrô e pela CPTM. Especificamente, isso ocorria no caso da linha 05 Lilás, muitas pessoas andavam a pé até a estação Santo Amaro ou Largo 13 por não poderem pagar duas tarifas, apenas a da CPTM a partir da estação Santo Amaro; ou apenas do ônibus, a partir da estação Largo 13. Isso foi alterado no início de 2012 (integração entre EMTU, SPTrans e CPTM), e em meados de 2013 (integração entre a linha 05 Lilás do Metrô e a EMTU).

Com essas e outras integrações, pós OD 2007, aliadas à própria expansão da rede, houve em 2012 - 2013, um aumento de mais de um milhão de passageiros por dia – o equivalente na época à população da cidade de Campinas, ou ao total de passageiros do metrô do Rio de Janeiro.

**É possível visualizar isso no mapa com as estações, linhas e as isócronas?**

### **O que se pode fazer em termos de geoprocessamento**

Sugestões:

- (1) Criar centroides para as zonas OD, com o aproveitamento dos tempos médios de viagem para o Centro no período do Pico da Manhã;
- (2) Criar diagrama de Voronoi a partir das estações operacionais em 31.12.2006 (veja Figura 20)

Tratam-se de polígonos centrados nas estações. Qualquer ponto dentro de um polígono está mais próximo (especialmente) da estação correspondente, do que de quaisquer outras;

Como a área de estudo é a RMSP, a extensão espacial do diagrama a ser gerado deve ser igual a dela, neste caso a extensão espacial do zoneamento da OD 2007.

Há de se observar que podem haver sobreposições nas coordenadas dos pontos representativos de estações servidas por mais de uma linha, e que disso podem resultar erros de processamento geométrico; Também de pontos muito próximos, pelo fato da base MetroLineStations.csv indicar pares (estação, linha) e não propriamente as estações; Há de se reduzir essa base para apenas as estações antes de se gerar o diagrama;

- (3) Coletar dados das estações em cada polígono correspondente;
- (4) Criar sequência de buffers individuais, a partir de cada estação operacional em 31.12.2006 – um a cada quilômetro (km);
- (5) Fazer a interseção espacial entre os diagramas de Voronoi e as sequências de buffers dessas estações. Capturar os dados de ambos no novo *layer*. Eliminar as redundâncias (colunas repetidas);
- (6) A partir da interseção dos buffers de 1km, 2km, ... de cada estação; com os polígonos de Voronoi dessas mesmas estações, coletar os dados dos centroides das zonas da Pesquisa OD 2007 que se encontram dentro delas: os tempos mínimo, máximo, etc., além do número de pontos coletados;

(7) Esses dados podem ser exportados em formato csv e analisados com o uso de software de planilha eletrônica e/ou banco de dados. Tal tabela conteria: estação; intervalo de distância (0-1km, 1-2km, ...); menor tempo médio de viagem; maior tempo médio de viagem; número de centroides coletados, etc.

Figura 20 – Diagrama de Voronoi das estações/linhas ativas em 31.12.2006

