



**UNIFACS**

UNIVERSIDADE SALVADOR

LAUREATE INTERNATIONAL UNIVERSITIES

□

**UNIVERSIDADE SALVADOR - UNIFACS**  
**MBA GESTÃO DA INFORMAÇÃO E PROJETOS DE BUSINESS**  
**INTELLIGENCE**

**ANDREA DE CARVALHO COSTA**

**MINERAÇÃO DE DADOS**  
**Análise de dados óbitos doença dengue em 2013-2014**  
**Salvador**

Salvador 2015

## Sumário

INTRODUÇÃO.....	3
METODOLOGIA.....	3
OBJETIVO.....	3
PERGUNTAS E RESPOSTAS.....	4
CLASSIFICAÇÃO:.....	6
CLASSIFICAÇÃO 1 – ANO REFERÊNCIA 2013.....	6
<u><a href="#">Tabela1 – ano referência 2013.....</a></u>	<u><a href="#">7</a></u>
UTILIZANDO WEKA – RESPOSTA PARA PERGUNTA 1 E 2.....	7
<u><a href="#">Mineração sexo masculino 2013.....</a></u>	<u><a href="#">8</a></u>
<u><a href="#">Mineração sexo feminino 2013.....</a></u>	<u><a href="#">8</a></u>
<u><a href="#">Mineração – ambos sexos 2013.....</a></u>	<u><a href="#">9</a></u>
<u><a href="#">Mineração – Sexo Feminino.....</a></u>	<u><a href="#">10</a></u>
<u><a href="#">Mineração Sexo Masculino.....</a></u>	<u><a href="#">10</a></u>
UTILIZANDO EXCEL.....	11
<u><a href="#">Gráfico período Janeiro/dezembro -2013 Feminino.....</a></u>	<u><a href="#">11</a></u>
<u><a href="#">Gráfico período Janeiro/dezembro -2013, Masculino.....</a></u>	<u><a href="#">11</a></u>
CLASSIFICAÇÃO 2 – ANO REFERÊNCIA 2014.....	12
<u><a href="#">Tabela2 – ano referência 2014.....</a></u>	<u><a href="#">13</a></u>
UTILIZANDO WEKA.....	13
<u><a href="#">Mineração – ano referência 2014.....</a></u>	<u><a href="#">13</a></u>
<u><a href="#">Mineração sexo feminino 2014.....</a></u>	<u><a href="#">14</a></u>
<u><a href="#">Mineração ambos sexos 2014.....</a></u>	<u><a href="#">14</a></u>
<u><a href="#">Mineração sexo feminino 2014.....</a></u>	<u><a href="#">15</a></u>
<u><a href="#">Mineração sexo masculino 2014.....</a></u>	<u><a href="#">15</a></u>
UTILIZANDO EXCEL.....	16
<u><a href="#">Gráfico período Janeiro/dezembro -2014 Feminino.....</a></u>	<u><a href="#">16</a></u>
<u><a href="#">Gráfico período Janeiro/dezembro -2014 Masculino.....</a></u>	<u><a href="#">16</a></u>
CLASSIFICAÇÃO 3 – NÍVEL DE ESCOLARIDADE, ENTRE HOMENS E MULHERES SALVADOR.....	17
<u><a href="#">Tabela3 – nível escolaridade 2014.....</a></u>	<u><a href="#">17</a></u>
<u><a href="#">Mineração nível escolaridade.....</a></u>	<u><a href="#">18</a></u>
UTILIZANDO WEKA.....	18
<u><a href="#">Mineração nível escolaridade idade.....</a></u>	<u><a href="#">18</a></u>
CLASSIFICAÇÃO 4 – SANEAMENTO BÁSICO RELAÇÃO HOMEM/MULHER.....	19
<u><a href="#">Tabela4 – saneamento básico 2014.....</a></u>	<u><a href="#">19</a></u>
UTILIZANDO WEKA.....	19
<u><a href="#">Mineração saneamento básico sexo feminino 2014.....</a></u>	<u><a href="#">20</a></u>
CLASSIFICAÇÃO 5 - CASOS DE DENGUE NOS 12 DISTRITOS SANITÁRIOS DE SALVADOR.....	20
<u><a href="#">Tabela5 – 12 distritos sanitários 2014.....</a></u>	<u><a href="#">21</a></u>
<u><a href="#">Mineração 12 distritos sanitário.....</a></u>	<u><a href="#">21</a></u>
<u><a href="#">Mineração 12 distritos sanitário.....</a></u>	<u><a href="#">22</a></u>
<u><a href="#">Mineração população 2013.....</a></u>	<u><a href="#">22</a></u>
CONCLUSÃO.....	23
BIBLIOGRAFIA.....	23

## Introdução

Dengue é uma doença infecciosa febril causada pelo Arbovírus do gênero Flavivírus (Sorotipos 1,2,3 e 4), que pode apresentar de forma benigna ou grave (Brasil, 2005), cuja evolução depende de diversos fatores.

A dengue tornou-se endêmica no Brasil e cursa com picos sazonais e ciclos epidêmicos. A transmissão, embora mais intensa no verão, pode ocorrer durante todo o ano, em especial nas regiões de clima quente e úmido.

## Metodologia

Para a análise da situação da dengue em Salvador fez-se um estudo de base de dados secundários a partir dos sistemas de informação disponíveis, a saber: Sistema de Informação em de Saneamento, do Ministério das Cidades; Base de dados do Ministério da Saúde (Datasus), e Censo Demográfico do Instituto Brasileiro de Geografia Estatística, Prefeitura Municipal de Saúde, Centro de Controle de zoonoses. Visando investigar as desigualdades no acesso aos serviços utilizaram-se como unidade de análise os setores censitários, que foram agregados em bairros que define 12 distritos sanitários de Salvador, que são: Centro Histórico, Itapagipe, São Caetano/Valeria, Liberdade, Brotas, Barra / Rio Vermelho, Boca do Rio, Itapoan, Cabula / Beiru, Pau da Lima, Subúrbio Ferroviário e Cajazeiras. A cidade de Salvador é a terceira maior cidade brasileira em população, estimada em torno de 2.902.927, com uma área de 692,819 km<sup>2</sup> (dados do IBGE - 2014).

## Objetivo

Criar uma base de dados com as seguintes informações a verificar quantas mulheres em relações a homens foram a óbitos pela doença em Salvador no ano de 2013 a 2014, distribuídos pela idade, escolaridade, localidade, saneamento básico.

Segundo dados do Sistema de Informação Hospitalar (SIH-SUS) do Ministério da Saúde, nos anos compreendidos entre 2013 e 2014 entre a idade de 01 a 80 anos, ocorreram 96 óbitos no ano de 2013 e em 2014 foram 167 óbitos e internações em residentes no município de Salvador.

O grupo mais atingido no ano de 2013 foi o da faixa etária de 10 a 14 e 20 a 29 anos com 11 internações do sexo feminino, totalizando no ano 44 óbitos, e o segundo mais atingido foi o de 5 a 9 anos com 11 internações do sexo masculino, totalizando 52 óbitos.

Observando a classificação mostra que os dados resultam em um padrão semelhante com grandes danos para as crianças na faixa de 5 a 14 anos.

Para o ano de 2014 temos o grupo mais atingido foi o da faixa etária de 30 a 39 anos com 21 óbitos do sexo feminino, e o segundo mais atingido foi o de 20 a 29 anos com 17 óbitos do sexo masculino.

Observando a classificação mostra que os dados resultam em um padrão semelhante com grandes danos para os jovens na faixa etária de 20 a 39 anos.

## **Perguntas e respostas mineração de dados**

As informações analisadas pelo sistema weka e excel mostram:

1 – Mês, idade que teve mais óbitos do sexo masculino ano 2013?

(Imagem1 e Imagem5)

Resp : Período de março e janeiro;

Resp: Idade de 5 a 9 anos

2 – Mês, idade que teve mais óbitos do sexo feminino ano 2013?

(Imagem2 - Imagem4)

Resp : Período de março e abril;

Resp: Idade 10 a 14 anos

3 – Mês ambos sexos foram a óbitos no ano ano 2013?

(Imagem3)

Resp : Março e Janeiro;

4 – Quantidade de óbitos do sexo feminino e masculino ano 2013 ?

(Imagem6 – Imagem7)

Resp : 44 Feminino e 52 Masculino

5 – Mês que teve mais óbitos do sexo masculino ano 2014?

(Imagem8, Imagem12)

Resp : Período de junho e abril

Resp: Idade 15 a 29 anos

6 – Mês, idade que teve mais óbitos do sexo feminino ano 2014?

(Imagem9 - Imagem11)

Resp : Período de maio e abril

Resp: Idade 20 a 39 anos

7 – Mês ambos sexos foram a óbitos no ano ano 2014?

(Imagem10)

Resp : Maio e Abril

8 – Quantidade de óbitos do sexo feminino e masculino ano 2014 ?

(Imagem13 – Imagem14)

9 – Qual a área que esta mais propícia em ter a doença nos 12 distritos sanitários no ano de 2014?

(Imagem18, Imagem19, Imagem20)

Resp : Cabula/Beiru, Barra/Rio Vermelho e Boca do Rio

## Classificação:

1. Períodos disponíveis – janeiro de 2013 a dezembro de 2014,  
Seleções disponíveis,
2. Município 29020 Salvador,
3. Região de Saúde (CIR)29020 Salvador,
4. Macrorregião de Saúde - todas categorias,
5. Divisão administ estadual - 2901-Salvador,
6. Microrregião IBGE – 29021 – Salvador,
7. Região Metropolitana - 29010 Salvador,
8. Estabelecimento – Todas as categorias,
9. Caráter atendimento - Todas as categorias,
- 10.Regime - Todas as categorias
- 11.Capítulo CID-10 – Algumas doenças infecciosas e parasitarias,
- 12.Lista Morb CID-10 –Dengue [dengue clássico, Febre hemorrágica devida ao vírus da dengue],
- 13.Faixa Etária de 1 a 80 anos,
- 14.Sexo Masculino e Feminino.

## Classificação 1 – ano referência 2013

Mês	Idade 1 a 4			Idade 5 a 9			Idade 10 a 14			Idade 15 a 19			Geral
	Masc	Fem	Total	Masc	Fem	Total	Masc	Fem	Total	Masc	Fem	Tot	
Janeiro	2	0	2	1	0	1	2	2	4	0	0	0	7
Fevereiro	2	0	2	1	0	1	0	1	1	0	1	1	5
março	1	1	2	1	0	1	0	2	2	1	0	1	6
abril	1	0	1	0	0	0	0	3	3	1	1	2	6
maio	1	0	1	0	0	0	1	1	2	0	0	0	3
junho	1	0	1	0	0	0	0	0	0	2	1	3	4
julho	0	0	0	1	0	1	2	0	2	1	1	2	5
agosto	0	1	1	3	0	3	0	0	0	0	0	0	4
setembro	1	1	2	2	0	2	1	0	1	0	0	0	5
outubro	0	0	0	1	0	1	0	0	0	0	1	1	2
novembro	0	0	0	1	0	1	0	2	2	1	0	1	4
dezembro	1	0	1	0	1	1	0	0	0	1	2	3	5
Total	10	3	13	11	1	12	6	11	17	7	7	14	56

Mês	Idade 20 a 29			Idade 30 a 39			Idade 40 a 49			Idade 50 a 59			Geral
	Masc	Fem	Total	Masc	Fem	Total	Masc	Fem	Total	Masc	Fem	Tot	
Janeiro	1	1	2	0	0	0	2	0	2	1	0	1	5
Fevereiro	0	1	1	0	0	0	0	0	0	0	0	0	1
março	0	1	1	2	1	3	0	0	0	1	0	1	5
abril	1	0	1	0	3	3	1	0	1	0	0	0	5
maio	0	1	1	0	0	0	1	1	2	0	1	1	4
junho	1	1	2	1	0	1	1	0	1	0	0	0	4
julho	1	1	2	0	0	0	0	0	0	1	1	2	4
agosto	0	2	2	0	0	0	0	0	0	0	0	0	2
setembro	0	2	2	0	1	1	0	0	0	0	0	0	3
outubro	0	0	0	0	0	0	0	0	0	0	0	0	0
novembro	0	0	0	0	0	0	0	0	0	0	0	0	0
dezembro	0	1	1	0	0	0	0	0	0	1	0	1	2
Total	4	11	15	3	5	8	5	1	6	4	2	6	35

Mês	Idade 60 a 69			Idade 70 a 79			Idade 80 anos			Geral
	Masc	Fem	Total	Masc	Fem	Total	Masc	Fem	Total	
Janeiro	0	0	0	0	1	1	0	0	0	1
Fevereiro	0	0	0	0	1	1	1	0	1	2
março	0	0	0	0	0	0	1	0	1	1
abril	0	0	0	0	0	0	0	0	0	0
maio	0	0	0	0	0	0	0	0	0	0
junho	0	0	0	0	0	0	0	0	0	0
julho	0	0	0	0	0	0	0	0	0	0
agosto	0	0	0	0	0	0	0	0	0	0
setembro	0	0	0	0	0	0	0	0	0	0
outubro	0	0	0	0	0	0	0	0	0	0
novembro	0	0	0	0	0	0	0	0	0	0
Dezembro	0	0	0	0	1	1	0	0	0	1
Total	0	0	0	0	3	3	2	0	2	5

Tabela1 – ano referência 2013

Fonte : <http://tabnet.datasus.gov.br/cgi/tabcgi.exe?sih/cnv/niba.def>

## Utilizando weka – Resposta para pergunta 1 e 2

Utilizando o software para minerar os dados weka, com o arquivo doença-dengue.arff, foi visualizado que o mês onde teve mais óbitos do sexo masculino foi o período de março e janeiro de 2013;

The screenshot shows the Weka Explorer interface with the 'Classify' tab selected. The classifier is 'LeastMedSq -S 4 -G 0'. The 'Test options' section shows 'Cross-validation' with 'Folds' set to 10 and 'Percentage split' at 66%. The 'Classifier output' pane displays a 'Linear Regression Model' for 'Total\_Masc ='. The model coefficients are listed as follows:

```

0 * idade=70-79,80,30-39,20-29,50-59,1-80,40-49,10-14,15-19,1-4,5-9 +
0 * idade=80,30-39,20-29,50-59,1-80,40-49,10-14,15-19,1-4,5-9 +
0 * idade=30-39,20-29,50-59,1-80,40-49,10-14,15-19,1-4,5-9 +
0 * idade=20-29,50-59,1-80,40-49,10-14,15-19,1-4,5-9 +
0 * idade=50-59,1-80,40-49,10-14,15-19,1-4,5-9 +
0 * idade=1-80,40-49,10-14,15-19,1-4,5-9 +
0 * idade=40-49,10-14,15-19,1-4,5-9 +
0 * Mes=Fevereiro,Setembro,Abril,Junho,Julho,Marco,Janeiro +
0 * Mes=Setembro,Abril,Junho,Julho,Marco,Janeiro +
0 * Mes=Abril,Junho,Julho,Marco,Janeiro +
0 * Mes=Junho,Julho,Marco,Janeiro +
0 * Mes=Julho,Marco,Janeiro +
0 * Mes=Marco,Janeiro +
0

```

Below the model coefficients, the 'Summary' section provides the following statistics:

Correlation coefficient	0
Mean absolute error	0.3939
Root mean squared error	0.7487
Relative absolute error	72.9533 %
Root relative squared error	116.9376 %
Total Number of Instances	132

The 'Time taken to build model' is 1.08 seconds. The status bar at the bottom shows 'OK' and the system clock is 21:19 on 21/02/2015.

Mineração sexo masculino 2013

Para o sexo feminino o período de março e abril foram os que tiveram maior índice de óbitos em 2013;

The screenshot shows the Weka Explorer interface with the 'Classify' tab selected. The classifier is 'LeastMedSq -S 4 -G 0'. The 'Test options' section shows 'Cross-validation' with 'Folds' set to 10 and 'Percentage split' at 66%. The 'Classifier output' pane displays a 'Linear Regression Model' for 'Total\_Fem ='. The model coefficients are listed as follows:

```

0 * idade=80,40-49,5-9,50-59,70-79,1-4,1-80,30-39,15-19,20-29,10-14 +
0 * idade=40-49,5-9,50-59,70-79,1-4,1-80,30-39,15-19,20-29,10-14 +
0 * idade=5-9,50-59,70-79,1-4,1-80,30-39,15-19,20-29,10-14 +
0 * idade=50-59,70-79,1-4,1-80,30-39,15-19,20-29,10-14 +
0 * idade=70-79,1-4,1-80,30-39,15-19,20-29,10-14 +
0 * idade=1-4,1-80,30-39,15-19,20-29,10-14 +
0 * idade=1-80,30-39,15-19,20-29,10-14 +
0 * idade=30-39,15-19,20-29,10-14 +
0 * Mes=Julho,Setembro,Janeiro,Fevereiro,Maio,Dezembro,Marco,Abril +
0 * Mes=Setembro,Janeiro,Fevereiro,Maio,Dezembro,Marco,Abril +
0 * Mes=Janeiro,Fevereiro,Maio,Dezembro,Marco,Abril +
0 * Mes=Fevereiro,Maio,Dezembro,Marco,Abril +
0 * Mes=Maio,Dezembro,Marco,Abril +
0 * Mes=Dezembro,Marco,Abril +
0 * Mes=Marco,Abril +
0

```

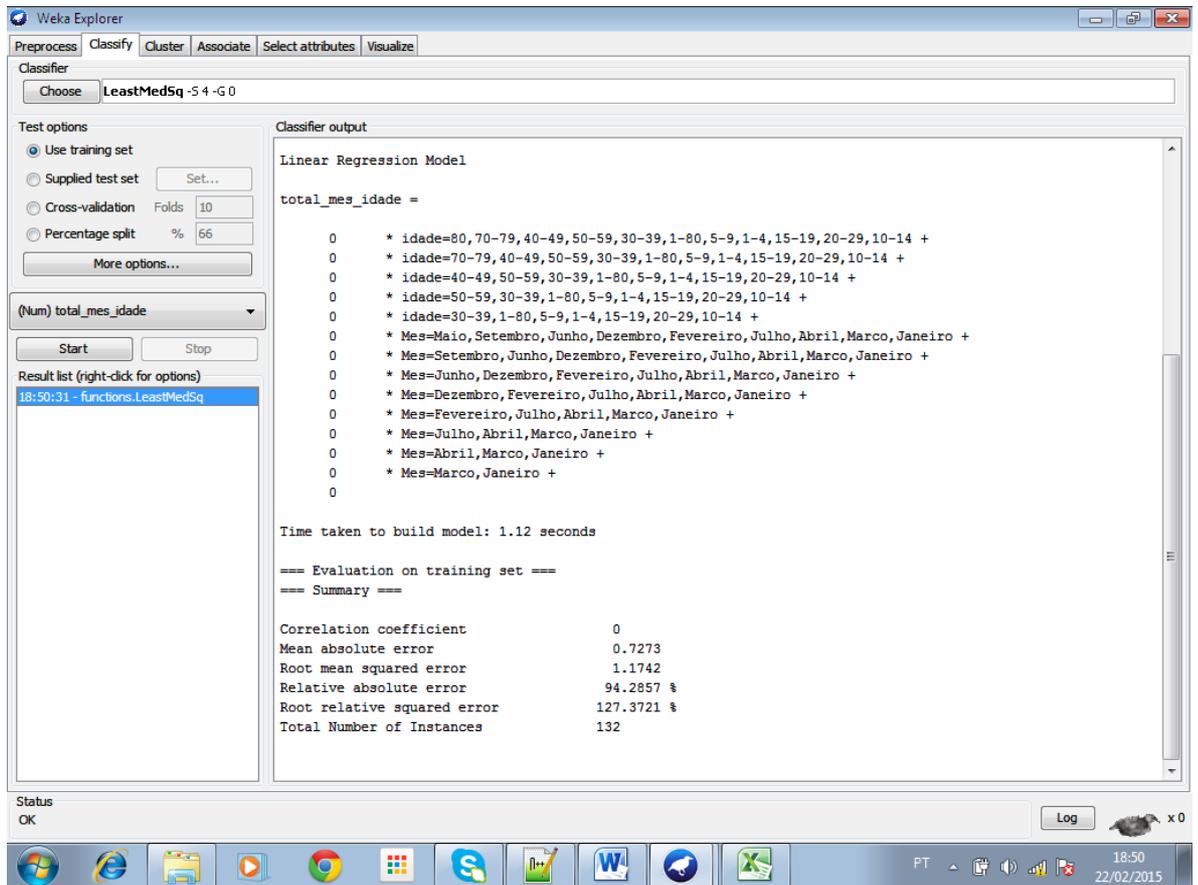
Below the model coefficients, the 'Summary' section provides the following statistics:

Correlation coefficient	0
Mean absolute error	0.3333
Root mean squared error	0.7177
Relative absolute error	66.6921 %
Root relative squared error	111.4569 %
Total Number of Instances	132

The 'Time taken to build model' is 1.21 seconds. The status bar at the bottom shows 'OK' and the system clock is 21:20 on 21/02/2015.

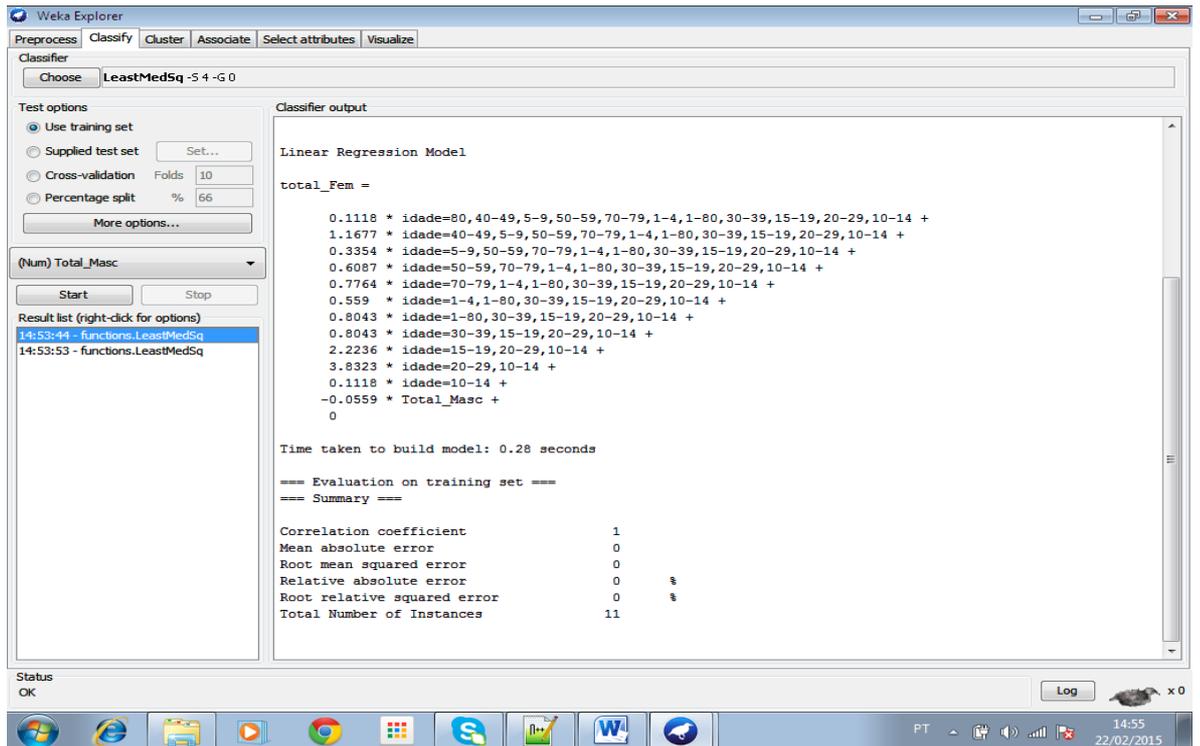
Mineração sexo feminino 2013

Para ambos sexos os meses que obtiveram maior índice de óbitos por dengue foram os períodos de março e janeiro 2013.



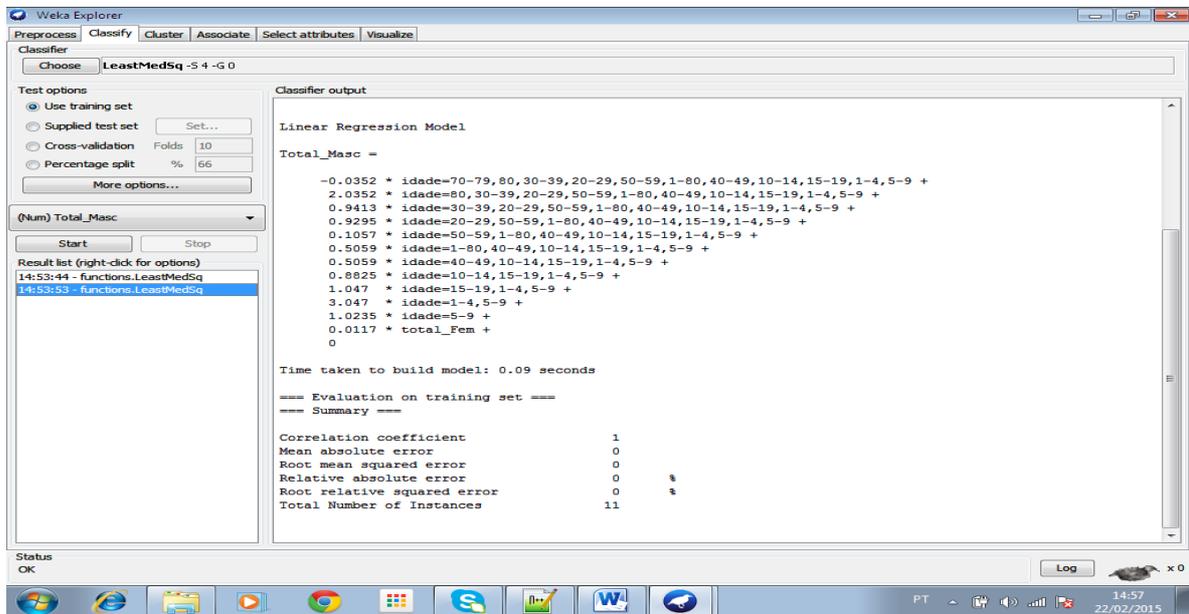
### Mineração – ambos sexos 2013

Para o sexo feminino a idade que teve o maior número de óbitos foi o de 10 a 14 anos (doença-dengue\_qtd.arff)



### Mineração - Sexo Feminino

Para o sexo masculino a idade que teve o maior número de óbitos foi o de 5 a 9 anos (doença-dengue\_qtd.arff)



### Mineração Sexo Masculino

## Utilizando excel

O gráfico mostra o período de janeiro a dezembro as informações dos casos de óbitos da dengue em relação a idade e sexo feminino.

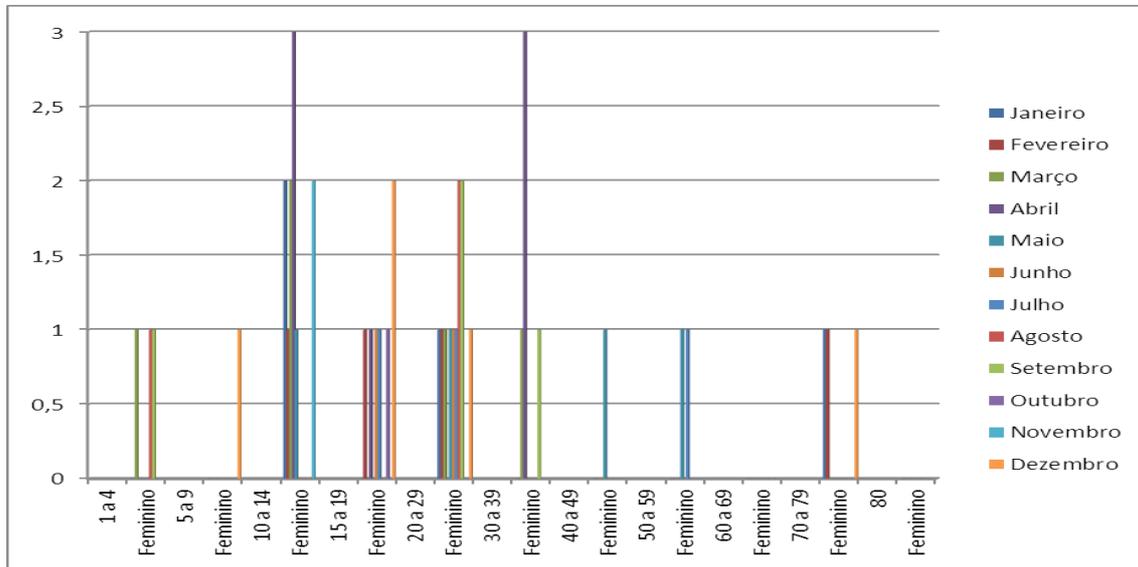


Gráfico período Janeiro/dezembro -2013 Feminino

O gráfico mostra o período de janeiro a dezembro as informações dos casos de óbitos da dengue em relação a idade e sexo masculino.

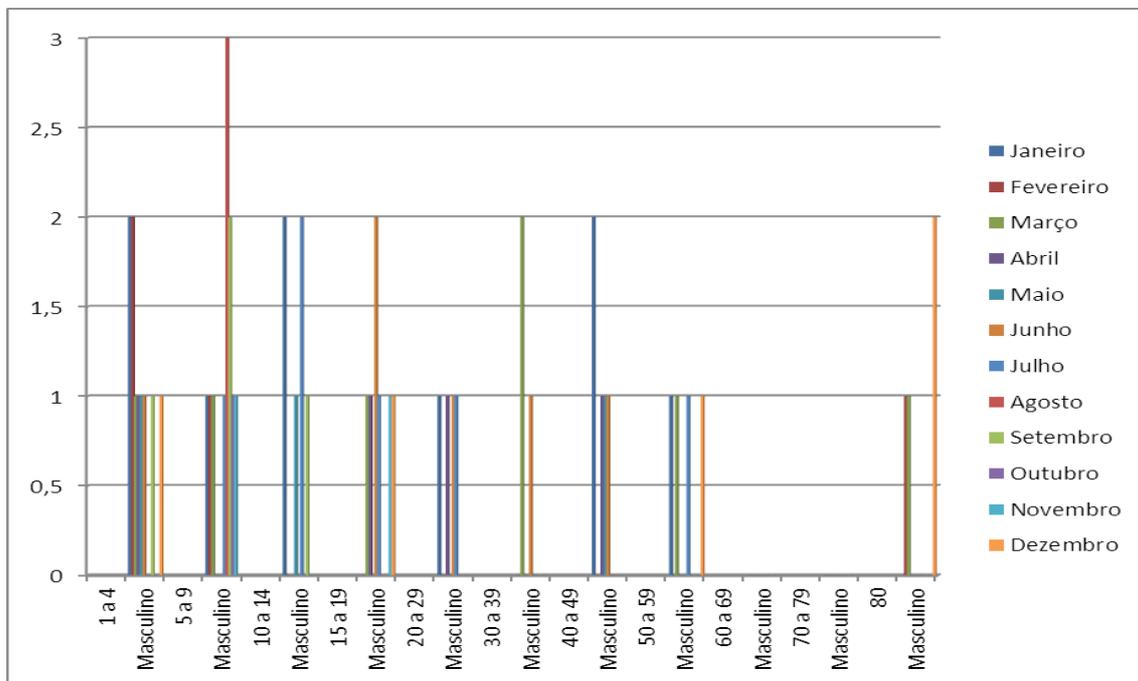


Gráfico período Janeiro/dezembro -2013, Masculino

## Classificação 2 – ano referência 2014

Mês	Idade 1 a 4			Idade 5 a 9			Idade 10 a 14			Idade 15 a 19			Geral
	Masc	Fem	Total	Masc	Fem	Total	Masc	Fem	Total	Masc	Fem	Tot	
Janeiro	0	1	1	1	0	1	1	0	1	0	0	0	3
Fevereiro	0	0	0	0	1	1	1	0	1	4	0	4	6
Março	0	1	1	2	0	2	0	3	3	0	1	1	7
Abril	1	2	3	2	1	3	3	0	3	5	4	9	18
Maio	0	1	1	0	1	1	2	2	4	1	0	1	7
Junho	0	1	1	1	1	2	3	1	4	1	2	3	10
Julho	0	0	0	1	0	1	0	0	0	0	1	1	2
Agosto	0	0	0	1	1	2	1	1	2	1	1	2	6
Setembro	0	3	3	0	1	1	0	0	0	1	0	1	5
Outubro	0	0	0	0	0	0	0	0	0	1	0	1	1
Novembro	0	0	0	0	2	2	0	0	0	0	0	0	2
Dezembro	0	0	0	1	0	1	0	0	0	1	0	1	2
Total	1	9	10	9	8	17	11	7	18	15	9	24	

Mês	Idade 20 a 29			Idade 30 a 39			Idade 40 a 49			Idade 50 a 59			Geral
	Masc	Fem	Total	Masc	Fem	Total	Masc	Fem	Total	Masc	Fem	Tot	
Janeiro	0	1	1	0	2	2	0	0	0	0	1	1	4
Fevereiro	0	0	0	2	0	2	0	0	0	0	0	0	2
março	2	0	2	0	1	1	0	0	0	1	0	1	4
abril	2	3	5	1	8	9	1	0	1	1	1	2	17
maio	3	3	6	2	5	7	2	2	4	0	1	1	18
junho	1	1	2	3	1	4	0	1	1	0	1	1	8
julho	1	2	3	2	0	2	1	2	3	1	0	1	9
agosto	1	3	4	2	0	2	0	2	2	0	2	2	10
setembro	1	0	1	2	2	4	2	0	3	0	1	1	8
outubro	3	3	6	0	2	2	4	0	4	0	0	0	12
novembro	0	0	0	0	0	0	0	0	0	0	0	0	0
dezembro	3	0	3	0	0	0	0	0	0	0	2	2	5
Total	17	16	33	14	21	35	10	7	18	3	9	12	

Mês	Idade 60 a 69			Idade 70 a 79			Idade 80 anos		
	Masc	Fem	Total	Masc	Fem	Total	Masc	Fem	Total
Janeiro	0	0	0	0	0	0	0	0	0
Fevereiro	0	1	1	0	0	0	0	0	0
março	0	2	2	0	0	0	0	0	0
abril	0	3	3	0	0	0	0	1	1
maio	0	0	0	0	0	0	0	0	0
junho	1	0	1	1	0	1	0	0	0

julho	0	0	0	0	0	0	0	0	0
agosto	0	0	0	1	0	1	0	0	0
setembro	0	0	0	0	0	0	0	0	0
outubro	0	0	0	0	0	0	0	0	0
novembro	0	0	0	0	0	0	0	0	0
dezembro	0	0	0	0	0	0	0	0	0
Total	1	6	7	2	0	2	0	1	1

Tabela2 – ano referência 2014

## Utilizando weka

Utilizando o software para minerar os dados weka, com o arquivo doença-dengue\_2014.arff, foi visualizado que o mês onde teve mais óbitos do sexo masculino foi o período de junho e abril de 2014;

The screenshot shows the Weka Explorer interface with the 'Classify' tab selected. The classifier chosen is 'LeastMedSq -S 4 -G 0'. The 'Classifier output' pane displays the following linear regression model:

```

total_ano_masculino =
  1.5816 * Mes=Janeiro, Dezembro, Marco, Julho, Setembro, Agosto, Fevereiro, Outubro, Maio, Junho, Abril +
  3.0966 * Mes=Dezembro, Marco, Julho, Setembro, Agosto, Fevereiro, Outubro, Maio, Junho, Abril +
 -0.4507 * Mes=Marco, Julho, Setembro, Agosto, Fevereiro, Outubro, Maio, Junho, Abril +
  1.0966 * Mes=Julho, Setembro, Agosto, Fevereiro, Outubro, Maio, Junho, Abril +
 -0.1932 * Mes=Setembro, Agosto, Fevereiro, Outubro, Maio, Junho, Abril +
  0.7103 * Mes=Agosto, Fevereiro, Outubro, Maio, Junho, Abril +
  0.7726 * Mes=Fevereiro, Outubro, Maio, Junho, Abril +
  0.5816 * Mes=Outubro, Maio, Junho, Abril +
  0.9055 * Mes=Maio, Junho, Abril +
  1.6438 * Mes=Junho, Abril +
  3.3905 * Mes=Abril +
  0.0644 * total_mes_ano +
  0.0322 * total_ano_feminino +
 -0.1932
  
```

The evaluation summary shows a correlation coefficient of 1, a mean absolute error of 0, and a total number of instances of 12.

Mineração – ano referência 2014

Para o sexo feminino o período de maio e abril onde tiveram o maior índice de óbitos dengue em 2014;

The screenshot shows the Weka Explorer interface with the 'Classify' tab selected. The classifier chosen is 'LeastMedSq -S 4 -G 0'. The 'Test options' section is set to 'Use training set' with 'Folds' set to 10 and 'Percentage split' at 66%. The 'Classifier output' pane displays the following text:

```

Linear Regression Model
total_ano_feminino =
  0.8211 * Mes=Novembro, Fevereiro, Julho, Outubro, Janeiro, Setembro, Março, Junho, Agosto, Maio, Abril +
 -1.0297 * Mes=Fevereiro, Julho, Outubro, Janeiro, Setembro, Março, Junho, Agosto, Maio, Abril +
  2.6848 * Mes=Julho, Outubro, Janeiro, Setembro, Março, Junho, Agosto, Maio, Abril +
 -0.3284 * Mes=Outubro, Janeiro, Setembro, Março, Junho, Agosto, Maio, Abril +
  0.9853 * Mes=Janeiro, Setembro, Março, Junho, Agosto, Maio, Abril +
  1.1034 * Mes=Setembro, Março, Junho, Agosto, Maio, Abril +
  1.2841 * Mes=Março, Junho, Agosto, Maio, Abril +
 -0.1052 * Mes=Junho, Agosto, Maio, Abril +
  1.4172 * Mes=Agosto, Maio, Abril +
  3.7882 * Mes=Maio, Abril +
  6.4154 * Mes=Abril +
  0.1198 * total_mes_ano +
  0.0444 * total_ano_masculino +
  0.9392

Time taken to build model: 0.25 seconds

=== Evaluation on training set ===
=== Summary ===

Correlation coefficient      1
Mean absolute error        0
Root mean squared error    0
Relative absolute error    0 %
Root relative squared error 0 %
Total Number of Instances  12

```

The 'Result list' on the left shows two entries for 'functions.LeastMedSq' at 18:24:55 and 18:43:22, with the second entry selected.

#### Mineração sexo feminino 2014

Para ambos sexos os meses que obtiveram maior índice de óbitos por dengue foram os períodos de maio e abril 2014.

The screenshot shows the Weka Explorer interface with the 'Classify' tab selected. The classifier chosen is 'LeastMedSq -S 4 -G 0'. The 'Test options' section is set to 'Use training set' with 'Folds' set to 10 and 'Percentage split' at 66%. The 'Classifier output' pane displays the following text:

```

Linear Regression Model
total_mes_ano =
  3.507 * Mes=Janeiro, Dezembro, Fevereiro, Julho, Março, Setembro, Outubro, Agosto, Junho, Maio, Abril +
  0.2107 * Mes=Dezembro, Fevereiro, Julho, Março, Setembro, Outubro, Agosto, Junho, Maio, Abril +
  0.4871 * Mes=Fevereiro, Julho, Março, Setembro, Outubro, Agosto, Junho, Maio, Abril +
  2.2764 * Mes=Julho, Março, Setembro, Outubro, Agosto, Junho, Maio, Abril +
 -0.7236 * Mes=Março, Setembro, Outubro, Agosto, Junho, Maio, Abril +
  2.0702 * Mes=Setembro, Outubro, Agosto, Junho, Maio, Abril +
  0.1405 * Mes=Outubro, Agosto, Junho, Maio, Abril +
  1.623 * Mes=Agosto, Junho, Maio, Abril +
  1.3009 * Mes=Junho, Maio, Abril +
  5.2963 * Mes=Maio, Abril +
  6.8477 * Mes=Abril +
  0.2565 * total_ano_masculino +
  0.3267 * total_ano_feminino +
  1.3466

Time taken to build model: 0.11 seconds

=== Evaluation on training set ===
=== Summary ===

Correlation coefficient      1
Mean absolute error        0
Root mean squared error    0
Relative absolute error    0 %
Root relative squared error 0 %
Total Number of Instances  12

```

The 'Result list' on the left shows one entry for 'functions.LeastMedSq' at 18:45:57, which is selected.

#### Mineração ambos sexos 2014

Em relação ao sexo feminino a idade entre 20 a 39 anos foram os que teve o maior número de óbitos (doença-dengue\_qtd\_2014.arff)

The screenshot shows the Weka Explorer interface with the 'Classify' tab selected. The classifier is 'LeastMedSq -S 4 -G 0'. The 'Test options' are set to 'Use training set' with 'Folds' set to 10 and 'Percentage split' set to 66. The 'Classifier output' pane displays the following Linear Regression Model:

```

Linear Regression Model
total_Fem =
  1.1713 * idade=80,60-69,40-49,10-14,5-9,1-80,1-4,15-19,50-59,20-29,30-39 +
  4.9143 * idade=60-69,40-49,10-14,5-9,1-80,1-4,15-19,50-59,20-29,30-39 +
  0.2291 * idade=40-49,10-14,5-9,1-80,1-4,15-19,50-59,20-29,30-39 +
  -0.0857 * idade=10-14,5-9,1-80,1-4,15-19,50-59,20-29,30-39 +
  1.1713 * idade=5-9,1-80,1-4,15-19,50-59,20-29,30-39 +
  0.8426 * idade=1-80,1-4,15-19,50-59,20-29,30-39 +
  0.8426 * idade=1-4,15-19,50-59,20-29,30-39 +
  -1.1992 * idade=15-19,50-59,20-29,30-39 +
  1.0279 * idade=50-59,20-29,30-39 +
  5.8008 * idade=20-29,30-39 +
  5.257 * idade=30-39 +
  0.0857 * Total_Masc +
  -0.1713

Time taken to build model: 0.08 seconds

=== Evaluation on training set ===
=== Summary ===

Correlation coefficient      1
Mean absolute error         0
Root mean squared error     0
Relative absolute error     0 %
Root relative squared error 0 %
Total Number of Instances  11
  
```

The 'Result list' on the left shows two entries: '19:00:08 - functions.LeastMedSq' and '19:00:48 - functions.LeastMedSq', with the latter selected.

#### Mineração sexo feminino 2014

Em relação ao sexo masculino a idade que teve o maior número de óbitos foi o de 15 a 29 anos (doença-dengue\_qtd\_2014.arff)

The screenshot shows the Weka Explorer interface with the 'Classify' tab selected. The classifier is 'LeastMedSq -S 4 -G 0'. The 'Test options' are set to 'Use training set' with 'Folds' set to 10 and 'Percentage split' set to 66. The 'Classifier output' pane displays the following Linear Regression Model:

```

Linear Regression Model
Total_Masc =
  0.7251 * idade=60-69,1-4,70-79,50-59,1-80,5-9,40-49,10-14,30-39,15-19,20-29 +
  -0.1649 * idade=1-4,70-79,50-59,1-80,5-9,40-49,10-14,30-39,15-19,20-29 +
  1.4948 * idade=70-79,50-59,1-80,5-9,40-49,10-14,30-39,15-19,20-29 +
  0.5052 * idade=50-59,1-80,5-9,40-49,10-14,30-39,15-19,20-29 +
  3.0275 * idade=1-80,5-9,40-49,10-14,30-39,15-19,20-29 +
  3.0275 * idade=5-9,40-49,10-14,30-39,15-19,20-29 +
  1.055 * idade=40-49,10-14,30-39,15-19,20-29 +
  1 * idade=10-14,30-39,15-19,20-29 +
  2.2304 * idade=30-39,15-19,20-29 +
  1.6597 * idade=15-19,20-29 +
  1.6152 * idade=20-29 +
  0.055 * total_Fem +
  -0.055

Time taken to build model: 0.1 seconds

=== Evaluation on training set ===
=== Summary ===

Correlation coefficient      1
Mean absolute error         0
Root mean squared error     0
Relative absolute error     0 %
Root relative squared error 0 %
Total Number of Instances  11
  
```

The 'Result list' on the left shows two entries: '19:00:08 - functions.LeastMedSq' and '19:04:48 - functions.LeastMedSq', with the latter selected.

#### Mineração sexo masculino 2014

## Utilizando excel

O gráfico mostra o período de janeiro a dezembro de 2014 as informações dos casos de óbitos da dengue em relação a idade e sexo feminino.

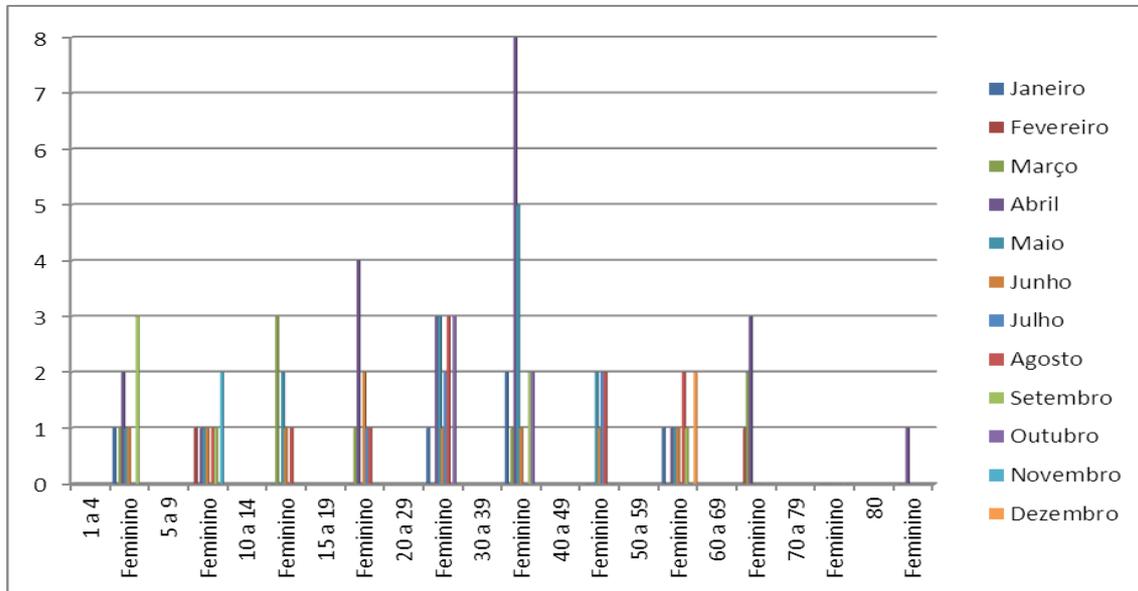


Gráfico período Janeiro/dezembro -2014 Feminino

O gráfico mostra o período de janeiro a dezembro de 2014 as informações dos casos de óbitos da dengue em relação a idade e sexo masculino.

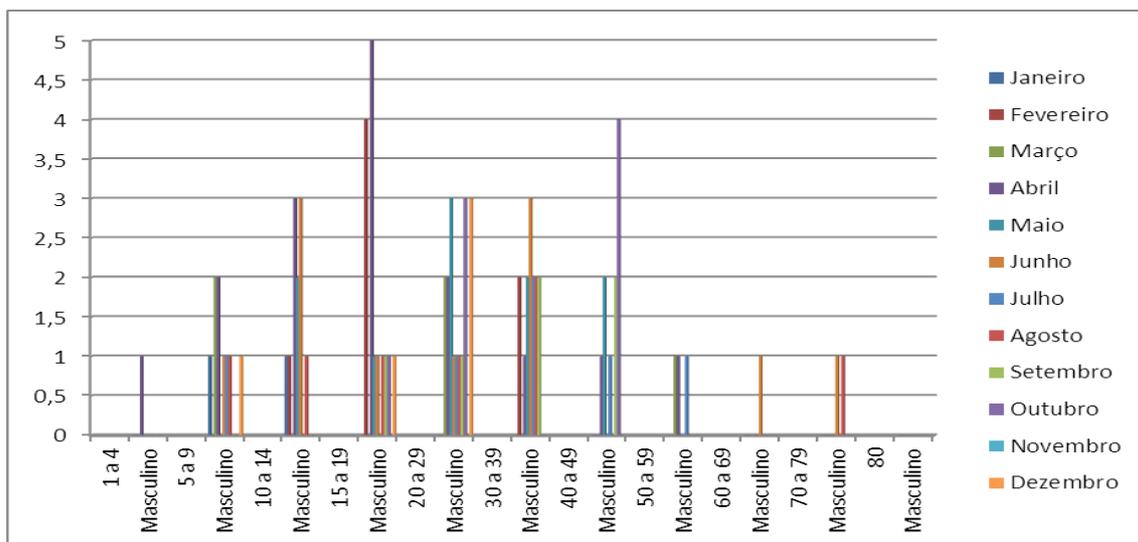


Gráfico período Janeiro/dezembro -2014 Masculino

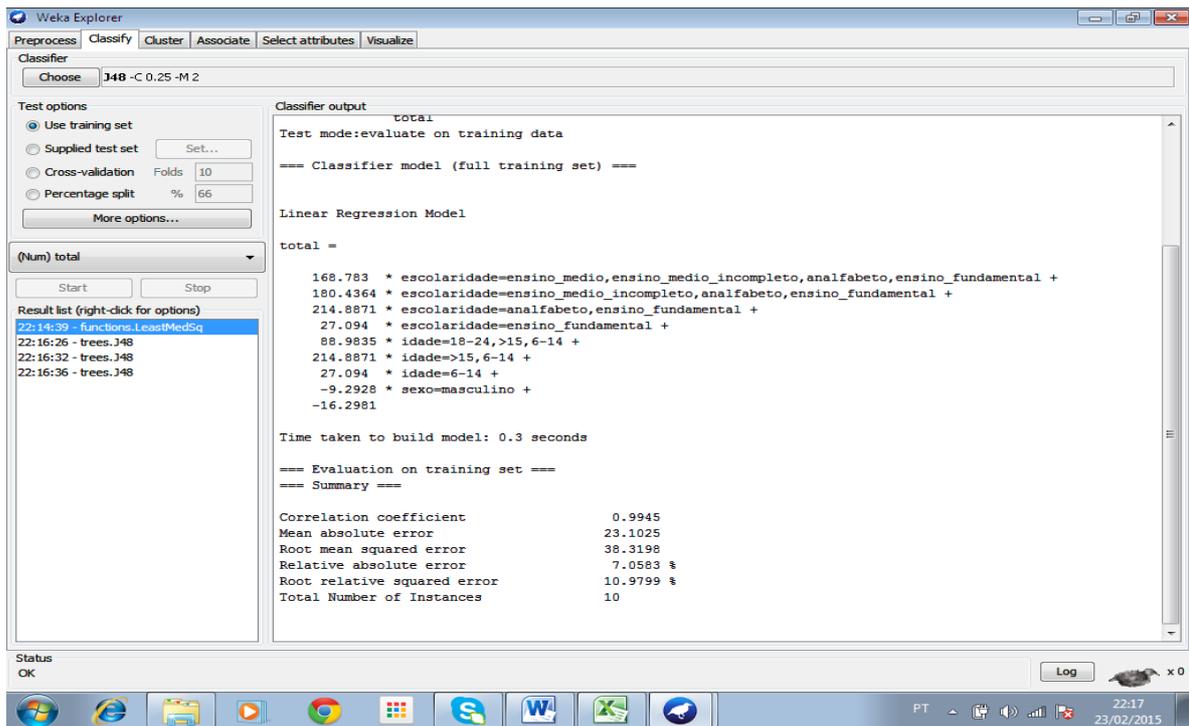
### Classificação 3 – Nível de escolaridade, entre homens e mulheres Salvador

Com base nas informações prestadas entre óbitos, o próximo nível é verificar a escolaridade de cada indivíduo em relação a idade.

Total de homens analfabetos com 15 anos ou mais de idade	857.795
Total de mulheres analfabetas com 15 anos ou mais de idade	836.270
Total de homens entre 6 e 14 anos de idade que frequentavam ensino fundamental	910.703
Total de mulheres entre 6 e 14 anos de idade que frequentavam ensino fundamental	891.738
Total de homens entre 15 e 17 anos de idade que frequentavam ensino médio	121.405
Total de mulheres entre 15 e 17 anos de idade que frequentavam ensino médio	174.272
Total de homens entre 18 e 24 anos de idade que frequentavam ensino superior	55.642
Total de mulheres entre 18 e 24 anos de idade que frequentavam ensino superior	80.436
Total de homens entre 18 e 24 anos de idade que não haviam concluído o ensino médio e não estavam frequentando a escola	412.612
Total de mulheres entre 18 e 24 anos de idade que não haviam concluído o ensino médio e não estavam frequentando a escola	309.031

Tabela3 – nível escolaridade 2014

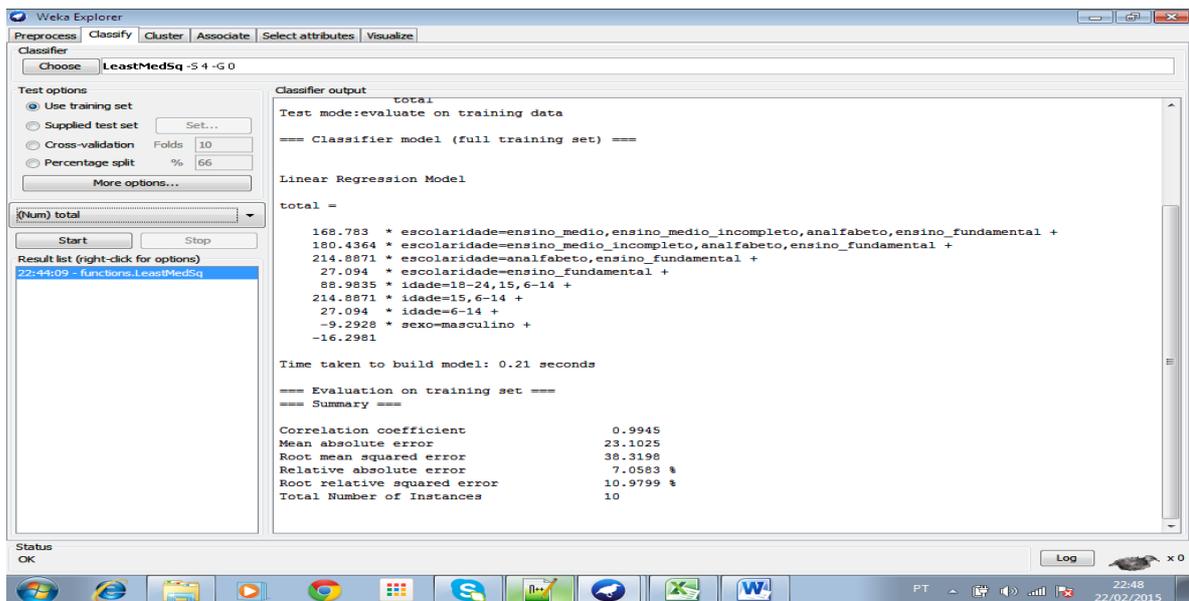
Fonte: IBGE, Censo Demográfico 2012. <br /> [www.ibge.gov.br/apps/snig/v1/](http://www.ibge.gov.br/apps/snig/v1/) - Para obter as informações do Sistema Nacional de Informação de Gênero.



Mineração nível escolaridade

## Utilizando weka

O software mostra que entre a idade de 06 a 14 anos o grau de estudo está no nível fundamental, entre o sexo masculino.



Mineração nível escolaridade idade

## Classificação 4 – Saneamento básico relação homem/mulher

Em Salvador, os bairros do Subúrbio Ferroviário, do Cabula, Pau da Lima, São Marcos, Nordeste de Amaralina e Beiru/Tancredo Neves são os que mais apresentam casos de doenças relacionadas à irregularidade da oferta de saneamento básico, segundo destaca o médico infectologista Robson Reis, que é membro da Secretaria Municipal de Saúde e professor de Urgências e Emergências Médicas da Escola Bahiana de Medicina e Saúde Pública.

Segundo o site ( <http://www.tratabrasil.org.br/datafiles/estudos/ranking/release-ranking-2014.pdf> ) mostra que Salvador está na posição 34º em saneamento básico.

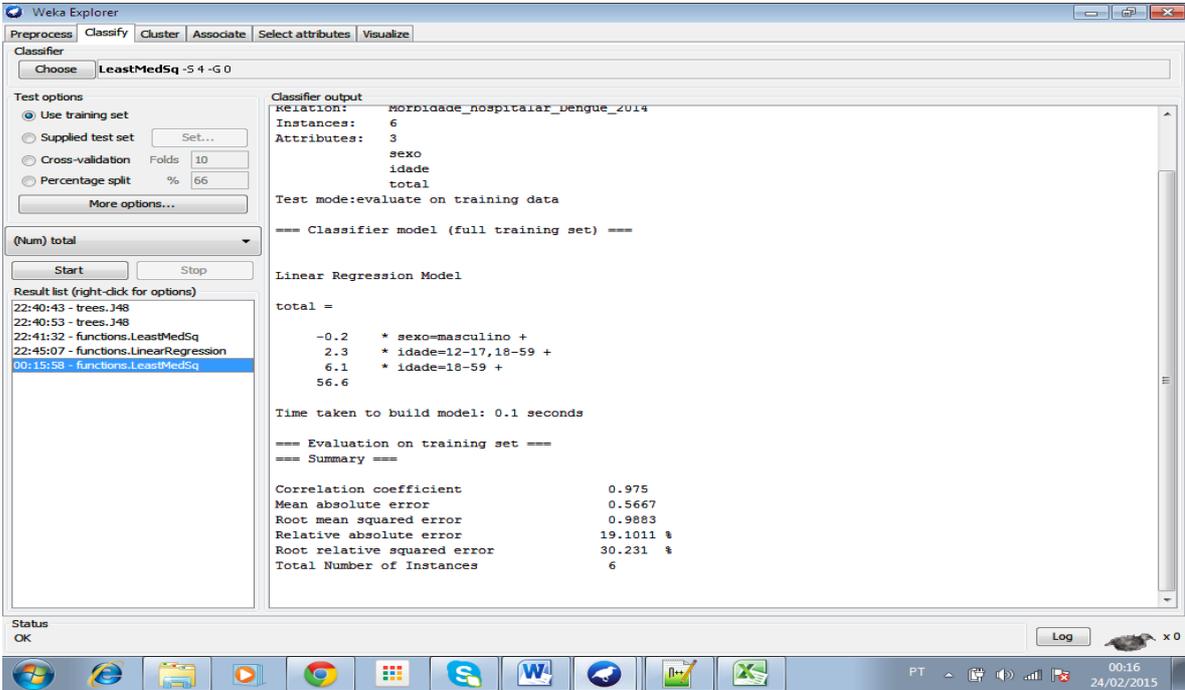
Fonte: <http://www.correio24horas.com.br/detalhe/noticia/agenda-bahia-saneamento-basico-e-saude-da-populacao-andam-de-maos-dadas/?cHash=a33706dfadbed62f7a4517847cfd4317>

Taxa de adequação do saneamento entre mulheres de 0 a 11 anos de idade residentes em domicílios particulares permanentes urbanos	56,4
Taxa de adequação do saneamento entre homens de 0 a 11 anos de idade residentes em domicílios particulares permanentes urbanos	56,6
Taxa de adequação do saneamento entre mulheres de 12 a 17 anos de idade residentes em domicílios particulares permanentes urbanos	58,7
Taxa de adequação do saneamento entre homens de 12 a 17 anos idade residente em domicílios particulares permanentes urbanos	57,4
Taxa de adequação do saneamento entre mulheres de 18 a 59 anos de idade residentes em domicílios particulares permanentes urbanos	64,8
Taxa de adequação do saneamento entre homens de 18 a 59 anos de idade residentes em domicílios particulares permanentes urbanos	63,1

Tabela4 – saneamento básico 2014

## Utilizando weka

As informações mostram que a idade entre 18 e 59 do sexo feminino tem o melhor saneamento.



The screenshot shows the Weka Explorer interface with the 'Classify' tab selected. The classifier chosen is 'LeastMedSq'. The 'Classifier output' pane displays the following information:

```
Relation: Morbiidade_hospitalar_dengue_2014
Instances: 6
Attributes: 3
  sexo
  idade
  total

Test mode: evaluate on training data

=== Classifier model (full training set) ===

Linear Regression Model

total =
-0.2 * sexo=masculino +
2.3 * idade=12-17,18-59 +
6.1 * idade=18-59 +
56.6

Time taken to build model: 0.1 seconds

=== Evaluation on training set ===
=== Summary ===
Correlation coefficient          0.975
Mean absolute error             0.5667
Root mean squared error         0.9883
Relative absolute error         19.1011 %
Root relative squared error     30.231 %
Total Number of Instances      6
```

Mineração saneamento básico sexo feminino 2014

## Classificação 5 - Casos de dengue nos 12 distritos sanitários de Salvador

Em 2014, até a SE 32 (09/08/2014), foram notificados 5.619 casos de Dengue, comparando-se o mesmo período com o ano de 2013 (N= 1867) percebe-se um incremento de 201% no número de casos. Entretanto, quando comparamos o ano de 2014 com os anos 2011 e 2012 há uma redução de 22% o que demonstra que o ano de 2013 apresentou casos abaixo do esperado.

Dos casos notificados no ano vigente, 1.537 (27%) foram classificados como Dengue/Dengue Clássico, 13 casos (0,23%) foram classificados como Dengue Grave, 42 (0,75%) como Dengue com sinais de alarme, 1.591 (28,3%) inconclusivos, 559 (9,95%) em investigação e 1877 (33,4%) foram descartados.

Os Distritos Sanitários de Cabula/Beiru, Barra/Rio Vermelho e Boca do

Rio apresentaram o maior número de casos notificados e também as maiores taxas de incidência.

Distrito Sanitário	Índice de Infestação	Índices Suspeitos	População
Subúrbio Ferroviário	3.56	56.17	347157
Brotas	8.91	222.72	219106
Cajazeiras	1.86	59.51	600356
Pau da Lima	5.28	122.61	120000
Itapoan	10.47	212.47	289687
Cabula/Beiru	21.95	289.05	415851
Itapagipe	2.26	70.84	175045
Boca do Rio	13.35	536.92	136147
Centro Histórico	2.52	178.18	77451
Barra/Rio vermelho	18.58	278.42	385270
Liberdade	7.01	198.58	103377
São Caetano/Valeria	4.24	83.61	277482

Tabela5 – 12 distritos sanitários 2014

The screenshot shows the Weka Explorer interface with the following details:

- Classifier:** LeastMedSq -S 4 -G 0
- Test options:**
  - Use training set (selected)
  - Supplied test set: Set...
  - Cross-validation: Folds 10
  - Percentage split: % 66
- Classifier output:**

```

Linear regression model
indice_infestacao =

1.6072 * distrito=Itapagipe,Centro_Historico,Suburbio_Ferroviario,Sao_Caetano-Valeria,Pau_da_Lima,Liber
0.3262 * distrito=Centro_Historico,Suburbio_Ferroviario,Sao_Caetano-Valeria,Pau_da_Lima,Liberdade,Brotas
0.5056 * distrito=Suburbio_Ferroviario,Sao_Caetano-Valeria,Pau_da_Lima,Liberdade,Brotas,Itapoan,Boca_dc
0.8263 * distrito=Sao_Caetano-Valeria,Pau_da_Lima,Liberdade,Brotas,Itapoan,Boca_do_Rio,Barra-Rio_Vermel
1.4169 * distrito=Pau_da_Lima,Liberdade,Brotas,Itapoan,Boca_do_Rio,Barra-Rio_Vermelho,Cabula-Beiru +
1.6252 * distrito=Liberdade,Brotas,Itapoan,Boca_do_Rio,Barra-Rio_Vermelho,Cabula-Beiru +
1.5167 * distrito=Brotas,Itapoan,Boca_do_Rio,Barra-Rio_Vermelho,Cabula-Beiru +
1.3765 * distrito=Itapoan,Boca_do_Rio,Barra-Rio_Vermelho,Cabula-Beiru +
2.671 * distrito=Boca_do_Rio,Barra-Rio_Vermelho,Cabula-Beiru +
5.0299 * distrito=Barra-Rio_Vermelho,Cabula-Beiru +
3.2602 * distrito=Cabula-Beiru +
0.002 * inc_suspeitos +
0 * populacao +
0.004

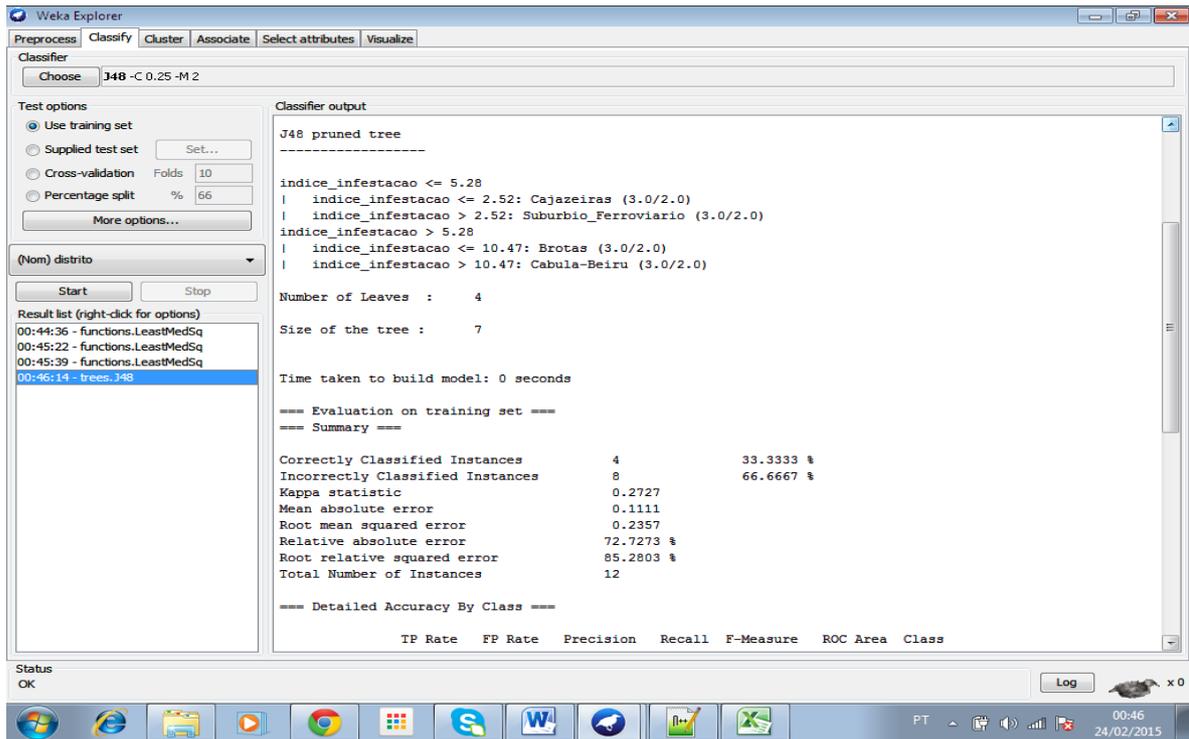
Time taken to build model: 0.77 seconds

=== Evaluation on training set ===
=== Summary ===

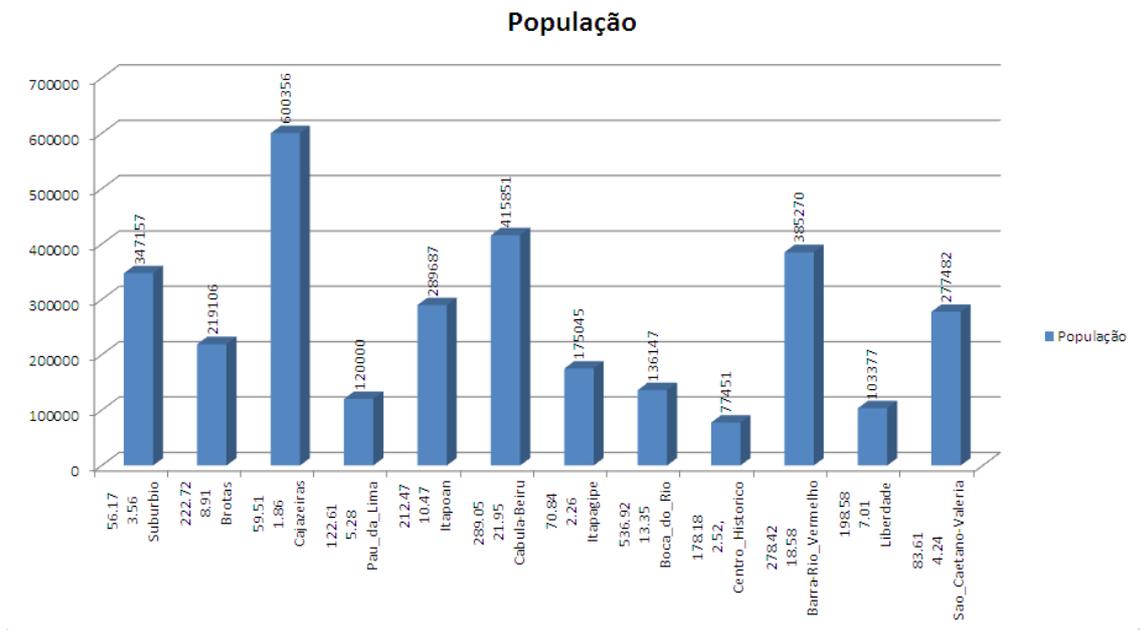
Correlation coefficient      1
Mean absolute error        0
Root mean squared error    0
Relative absolute error    0 %
Root relative squared error 0 %
Total Number of Instances  12

```

Mineração 12 distritos sanitário



Mineração 12 distritos sanitário



Mineração população 2013

## Conclusão

A evolução da doença demonstrou tendência ascendente do número de casos e do risco de adoecer por dengue até 2014. Durante o período analisado, a maior concentração de casos de dengue ocorreu no sexo feminino e com a idade entre 20 a 39 anos, com exceção do ano de 2013, onde houve predomínio de casos no sexo masculino. Na análise por ciclo de vida, até o ano de 2014, houve predomínio do risco de adoecer entre adultos na faixa etária de 25 a 39 anos, após esse período, as crianças e adolescentes passaram a ser os mais acometidos pela doença (Tabela1, Tabela2).

A distribuição da ocorrência da doença por distrito sanitário demonstrou que o Centro Histórico, Itapagipe e Itapoan apresentaram os maiores coeficientes de incidência por Dengue.

## Bibliografia

Dados do IBGE População de Salvador – Acessado em 15/02/2015

<http://cidades.ibge.gov.br/xtras/perfil.php?codmun=292740>

Coordenação de População e Indicadores Sociais: Estimativas da população residente com data de referência -Acessado em 15/02/2015

[http://www.ibge.gov.br/home/estatistica/populacao/estimativa2014/estimativa\\_tcu.shtm](http://www.ibge.gov.br/home/estatistica/populacao/estimativa2014/estimativa_tcu.shtm)

Área territorial brasileira - Acessado em 15/02/2015

<http://www.ibge.gov.br/home/geociencias/areaterritorial/area.php?nome=%>

Doenças transmissíveis, endemias, epidemias e pandemia

<http://books.scielo.org/id/8pmmmy> - Acessado em 15/02/2015

Fonte: IBGE, Censo Demográfico - Informações do Censo Demográfico 2010 - Acessado em 20/02/2015

<http://www.censo2010.ibge.gov.br>

<http://www.pnud.org.br/atlas/ranking/Ranking-IDHM-Municipios-2010.aspx>

Censo Demográfico 2010: Características Urbanísticas do Entorno dos Domicílios - Acessado em 20/02/2015

[http://www.ibge.gov.br/estadosat/temas.php?sigla=ba&tema=censodemog2010\\_snig](http://www.ibge.gov.br/estadosat/temas.php?sigla=ba&tema=censodemog2010_snig)

Boletim da dengue em 09/08/2012 - Acessado em 20/02/2015

[http://www.saude.salvador.ba.gov.br/dengue/arquivo/Boletim\\_dengue\\_09\\_08\\_2012.pdf](http://www.saude.salvador.ba.gov.br/dengue/arquivo/Boletim_dengue_09_08_2012.pdf)

Luta contra a doença dengue - Acessado em 21/02/2015

<http://g1.globo.com/luta-contr-a-dengue/noticia/2011/03/secretaria-de-saude-confirma-dois-casos-de-dengue-4-em-salvador.html>

Casos de óbitos em Salvador - Acessado em 22/02/2015

<http://www.combateadengue.com.br/dengue-mata-duas-pessoas-e-infestacao-avanca-na-capital-baiana/#ixzz3QchOko58>

<http://www.combateadengue.com.br/dengue-mata-duas-pessoas-e-infestacao-avanca-na-capital-baiana/>

Saneamento básico - Acessado em 22/02/2015

<http://www.tratabrasil.org.br/datafiles/estudos/ranking/release-ranking-2014.pdf>

<http://www.tratabrasil.org.br/datafiles/estudos/ranking/relatorio-completo-2014.pdf>

Distrito sanitário de salvador – Acessado em 25/02/2015

<http://www.saude.salvador.ba.gov.br/dengue/arquivo/Informe-Epidemiologico-sobre-dengue-atualizado.pdf>