

CONTEXTO DOS DADOS



Mestrado em modelagem matemática da informação
Disciplina visualização de informação
Profa Asla Medeiros e Sá

Relatório sobre Visualização de Informação
EMAp – FGV
Setembro, 2015

Walter_Dominguez
domingwal@gmail.com

Apresentação

O objetivo deste trabalho é entender o conceito de visualização da informação

Este trabalho dividido em 2 partes:

Parte 1: Proposta de visualizações

Parte 2: Anexos

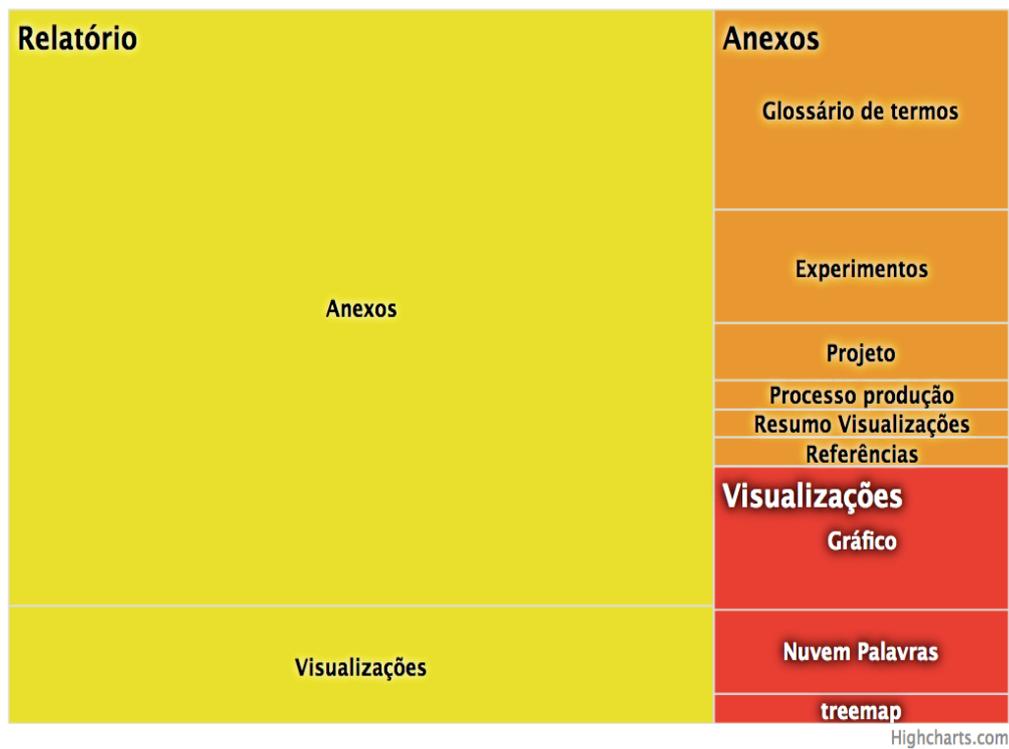
Contém:

Trabalho completo (este documento)

Apresentação do trabalho (incluído neste documento)

Visualizações (incluído neste documento)

Visualização da informação



<http://jsfiddle.net/rpscwjgg/>

Índice

Parte 1: Visualizações

1. Proposta 1 de visualização

Introdução

Objetivo

Requisitos

Papeis

Gráfico

Processo para produzir o gráfico

Prospecção futura

Execução da proposta

2. Proposta 2 de visualização

Introdução

Objetivo

Requisitos

Papeis

Nuvem de palavras

Procedimento para produzir a nuvem

Prospecção futura

3. Proposta 3 de visualização

Introdução

Objetivo

Requisitos

Papeis

Treemap

Procedimento p/ produzir o Treemap

Prospecção futura

Parte 2 Anexos

1. Planejamento do trabalho

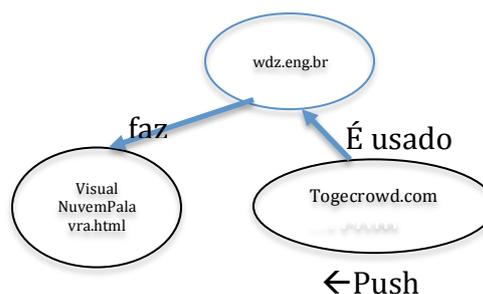
2. Experimentos

3. Glossário de termos

4. Processo p/ produzir visualização

5. Apresentação da visualização escolhida

6. Visualizações



7. Referências

Parte 1: Visualizações

1. Proposta 1 de visualização – Gráfico

Motivação

Este trabalho foi motivado pela necessidade de fazer uma análise da comparação de vários tipos de ferramentas de visualização x grau de satisfação do mercado. Esta pesquisa foi feita pelo site TDMI (site educacional que fala coisas sobre dados).

Objetivo

Interpretar a pesquisa realizada pelo site TDMI, das ferramentas de visualização interativa de dados do mercado, usadas com sucesso pelas empresas.

Requisitos

Da visualização: Dar subsídios de conhecimento em visualização para análise de gráfico comparativo.

Dos dados fornecidos: Foram obtidos de questionários com executivos de TI e negócios, VP's de BI/DW.

A análise de dados de várias organizações (378 respondedores) foi feita através da análise visual auto-dirigida, e progresso em direção a objetivos de negócios (tais como gestão de decisão), e recomendar as melhores práticas. Os questionários foram realizados em 2015.

Marketing, vendas, pessoal de fornecedores de tecnologia, bem como estudantes foram excluídos.

Papeis

Iniciador – Asla Medeiros e Sá

Coordenador – Walter Dominguez

Jornalista, Designer, Cientista de dados, Comunicador, Cientista cognitivo - Walter

Público

Criador dos dados: TDWI (Educação onsite)

Consumidor dos dados: Analista de negócios/BI, Arquiteto de Dados, Gestor do conhecimento

Análise da Visualização

Gráfico a ser analisado:

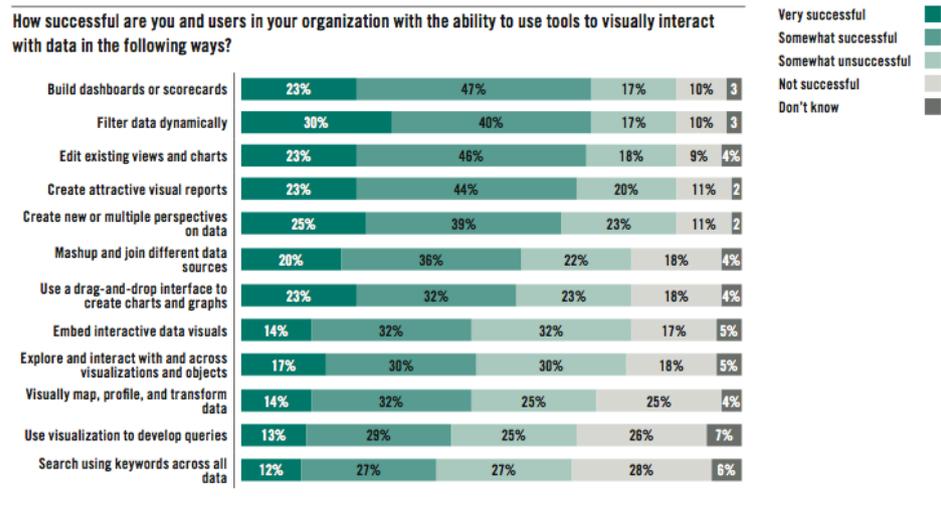
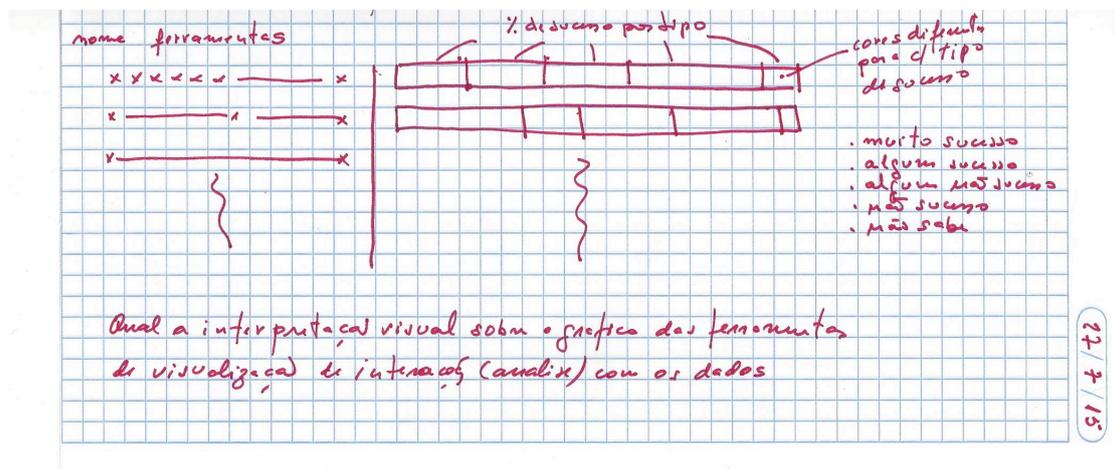


Figure 5. Based on answers from 378 respondents who could select only one answer per row.

O que se deseja:



Pergunta: Qual a interpretação visual sobre o gráfico das ferramentas de visualização de interações (análise) com os dados?

Variáveis de dados: 1 categoria nominal (nome da ferramenta) 1 quantitativa (Grau de sucesso).

Variáveis visual: Comprimento da barra, cores na barra

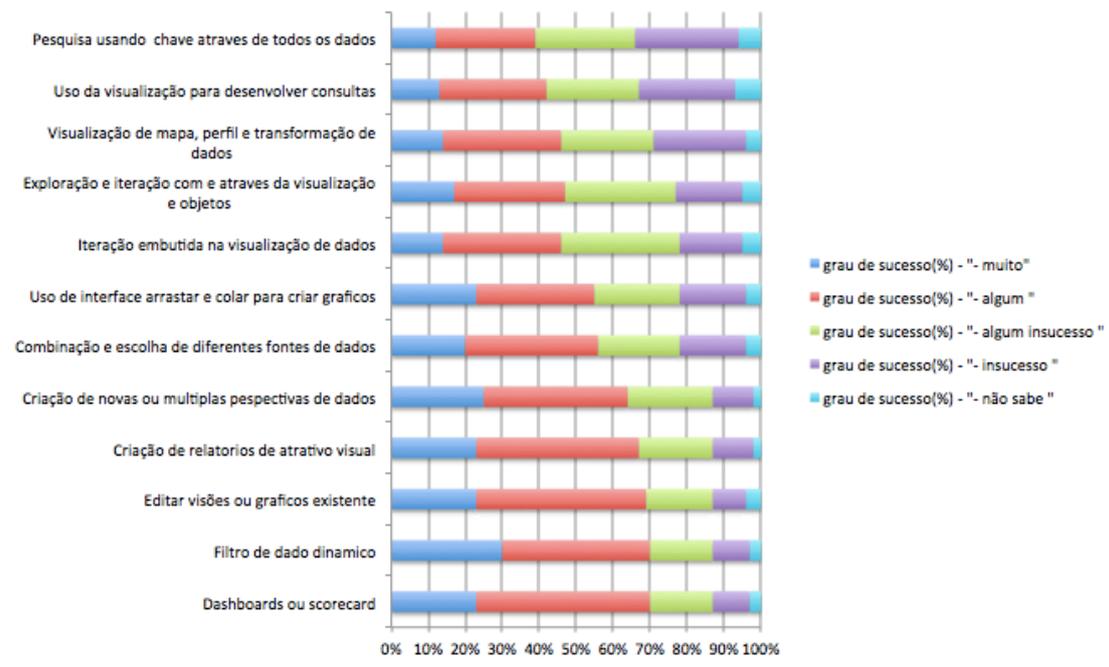
Método: Comparação de categorias.

Iteração: nenhuma

Descrição: ajuda a mostrar a gama de valores quantitativos.

Ele apresenta um barra que distribui o grau de sucesso observado pelas empresas (portanto, a posição inicial não é o ponto zero).
 Usando tais cartas permite identificar a diversidade das medições dentro de uma categoria e visualizar parte do todo em todas as categorias.
 Uso das cores: diferenciar percentuais

Ferramenta de visualização (Excel)



Fonte

Gráfico convertido para planilha do Excel:

ferramentas	grau de sucesso(%) -				
	" muito "	" algum "	algum insucesso "	insucesso "	" não sabe "
Dashboards ou scorecard	23	47	17	10	3
Filtro de dado dinâmico	30	40	17	10	3
Editar visões ou gráficos existente	23	46	18	9	4
Criação de relatórios de atrativo visual	23	44	20	11	2
Criação de novas ou múltiplas perspectivas de dados	25	39	23	11	2
Combinação e escolha de diferentes fontes de dados	20	36	22	18	4
Uso de interface arrastar e colar para criar gráficos	23	32	23	18	4
Iteração embutida na visualização de dados	14	32	32	17	5
Exploração e iteração com e através da visualização e objetos	17	30	30	18	5
Visualização de mapa, perfil e transformação de dados	14	32	25	25	4
Uso da visualização para desenvolver consultas	13	29	25	26	7
Pesquisa usando chave através de todos os dados	12	27	27	28	6

Procedimento para produzir o gráfico:

- Escolher o gráfico e base de dados do que se deseja obter
- Transformar Gráfico -> Tabela (procedimento manual)
- Analisar tabela identificando os componentes de visualização (proc manual)

Melhorar gráfico procurando novos ângulos para melhorar a visualização
(proc analítico no computador)

Prospecção futura:

- Colocar número percentual nas barras coloridas.
- Descrição e uso de cada ferramenta
- Automatizar a coleta da informação.

2. Proposta 2 de visualização – Nuvem de palavras (ver em <http://wdz.eng.br/VisualNuvemPalavras.xml.html>)

Motivação

Proposta para dar subsidio na identificação de um tópico, através da visualização de estrutura textual (dados nominais).

Será mostrada a partir de um texto no formato XML as palavras com maior frequência variando seu tamanho e tonalidade de cor .

A abordagem do trabalho é através de melhorias contínuas.

Foram feitos vários experimentos com o webservice <http://tagcrowd.com>

Objetivo

Mostrar as palavras mais relevantes de um tópico.

Requisitos

Dar subsidios para identificar o nome de um tópico.

Fonte:

O arquivo contém um tópico e esta no format xml.

Contem: texto puro com tag's xml, html, javascript, svg e outros scripts que possam surgir posteriormente.

Visualização:

Linguagem do texto: Português

Numero máximo de palavras para mostrar? : 40 (pode variar)

Frequencia minima? : 5 (pode variar)

Mostra frequencia junta as palavras?: yes (pode ser não)

Agrupa somente palavras similares? (Sómente em Inglês): sim (pode ser não)

Converte para letra minuscula? : original

Cor das palavras: uma cor (azul) com duas tonalidades

Não mostrar as seguinte palavras:

Papeis

Iniciador – Asla Medeiros e Sá

Coordenador – Walter Dominguez

Jornalista, Designer, Cientista de dados, Comunicador, Cientista cognitivo -

Walter

Público

Criador dos dados: Conteudista (VisualNuvemPalavras.xml)

Consumidor dos dados: Conteudista e Gestor do conhecimento

Análise da Visualização

Pergunta

O nome do tópico "Visualização de Nuvem de Palavras de um tópico" esta coerente com o conteúdo deste tópico?

nuvem (7) palavras (10) **tópico** (26) visualizacao (7)
created at TagCrowd.com

Componentes de sintaxe gráfica da informação

Variáveis de dados: 1 categoria nominal (palavras)

1 quantitativa (frequência da palavra).

Variáveis visual: Tamanho da letra, tonalidade (claro/escuro) cor (azul)

Método: Comparação da frequência das palavras.

Iteração: Texto fonte, linguagem do texto, número máximo de palavras para mostrar, frequência mínima, mostrar frequência sim ou não, converter para letra minúscula ou original, não mostrar as seguintes palavras.

Cor: Azul favorece o exercício intelectual e tranquiliza.

Descrição: Cada palavra tem seu tamanho regido pela relevância em determinado corpus de texto. Contagem simples das ocorrências de determinada palavra no texto. Uma palavra citada 27 vezes vai ter um tamanho proporcionalmente maior do que uma palavra citada 15 vezes

Processo para obtenção:

Fonte: VisualNuvemPalavras.xml

Seleção:

Palavras não selecionadas da fonte:

alt aparecao blank blog bold br caracteristicas class cloud
componentes desc div entao exemplos font font-weight height
href http id img informacao justify li linguagem links nao ordem
org px quantidade sao sintaxe site span src strong style table
tag tambem tarciziosilva tbody td text-align tipo title titulo topicos
tr tras ul vao web width wikipedia wp wp-content wp-image www

Ferramenta: <http://tagcrowd.com>

Passar os seguinte parâmetros:

Choose your text source:

Paste Text Web Page URL Upload File

Paste text to be visualized:
plain text, 500 kilobyte max

Visualize!

Options:

Language of text:
Ignore common words in this language Portuguese

Maximum number of words to show?
25 - 100 is a good range 50

Minimum frequency?
Don't show infrequent words 1

Show frequencies?
Show word count next to each word no yes

Group similar words? (English only)
eg: learn, learned, learning -> learn no yes

Convert to lowercase?
eg: PhD -> phd, FBI -> fbi, Rio -> rio lowercase original

Don't show these words:
Exclude unwanted words.

Procedimento para usar Gerador de nuvem de palavras:
<http://tagcrowd.com>

1. Escolher Paste text ou WebPage url ou upload file.
2. Se Past text então copiar texto fonte e colar em: Paste text to be visualized:
Se não colocar nome da url ou nome do arquivo
3. Em Options, preencher:
 - Language of text: Portuguese
 - Maximum number of words to show? : 40
 - Minimum frequency? : 5
 - Show frequencies?: yes
 - Group similar words? (English only): yes
 - Convert to lowercase? : original
 - Don't show these words: Deixar em branco
4. Clicar em VISUALIZE!
5. Enquanto objetivo não for atendido:
 - Repetir passo 2 variando os parametros e inclua palavras que não deseja aparecer (copiar item anterior) em Don't show these words:
 - Volte de novo ao passo 3

Proposta 3 de visualização - Treemap

Motivação

Visualizar os principais tópicos deste relatório mostrando o volume de informação de cada um dos tópicos até o 2º nível.

Objetivo

Visualizar na estrutura hierárquica o tipo Treemap dos tópicos do índice e seus tamanhos proporcional a quantidade de páginas.

Requisitos

Dar subsídios para mostrar o tamanho dos tópicos diferenciando-os pelas cores e suas tonalidades.

Usar o Highcharts.com e alterar parâmetros/código para fazer visualização desejada.

Fonte

Nome: nome do tópico

Relacionamento: identificador do tópico filho

Valor: proporcional ao número de páginas do tópico

Visualização

Nome do tópico

Cor do tópico

Quantidade de páginas proporcional

Papeis

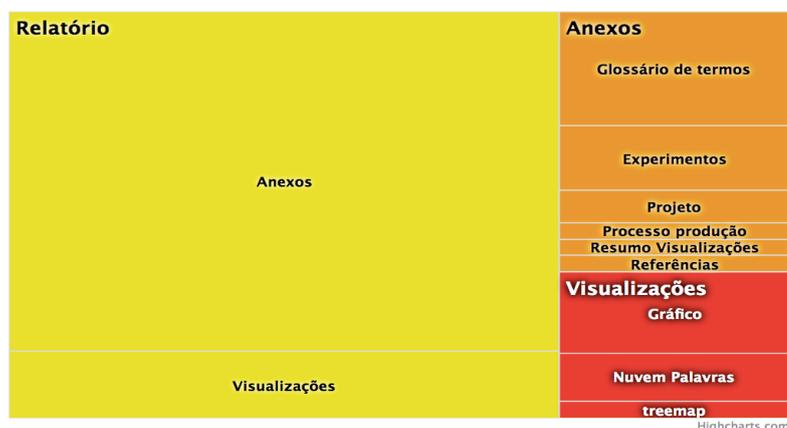
Iniciador – Asla Medeiros e Sá

Coordenador – Walter Dominguez

Jornalista, Designer, Cientista de dados, Comunicador, Cientista cognitivo – Walter

Análise da Visualização

Visualização da informação



Highcharts.com

Componentes de sintaxe gráfica da informação

Variáveis de dados: 1 categoria nominal (nome do tópico e subtópico)
1 quantitativa (proporcional quantidade de páginas).

Variáveis visual: Tamanho do tópico, tonalidade (claro/escuro) cor (azul)

Método: Comparação dos tamanhos dos tópicos em número de páginas.

Iteração: não tem.

Descrição: Cada tópico tem seu tamanho regido pelo tamanho em quantidade proporcional de páginas. Cada tópico tem uma cor e cada subtópico tem tonalidades diferentes referente a cor do tópico.

Processo para obtenção

Fonte:

Entra em <http://jsfiddle.net/rpscwjgg/> e alterar dados em data no javascript:

```
data: [{
  id: 'A',
  name: 'Visualizações',
  color: "#EC2500"
}, {
  id: 'B',
  name: 'Relatório',
  color: "#ECE100"
}, {
  id: 'O',
  name: 'Anexos',
  color: "#EC9800"
},
{
  name: 'Gráfico',
  parent: 'A',
  value: 5
}, {
  name: 'Nuvem Palavras',
  parent: 'A',
  value: 3
}, {
  name: 'treemap',
  parent: 'A',
  value: 1
}, {
```

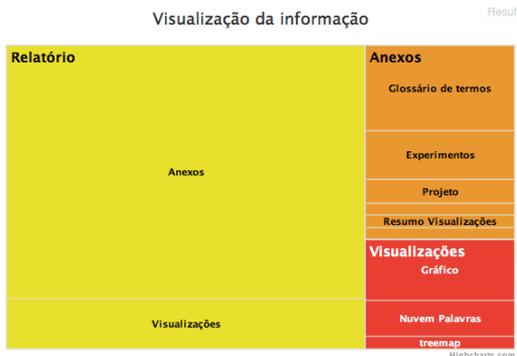
Ver como alterar em:

<http://www.highcharts.com/docs/chart-and-series-types/treemap>

*Para download <http://www.highcharts.com/download>

Visualização

Entrar em <http://jsfiddle.net/rpscwjgg/> em ver no quadrante inferior a direita:



Parte 2 Anexos

1. Planejamento do trabalho
2. Experimentos
3. Glossário de termos
4. Processo p/ produzir visualização
5. Apresentação da visualização escolhida
6. Visualizações
7. Referências

1. Planejamento

Atividades:

Pergunta: Quais são as iterações x atividade de projeto na etapa de

Fluxo -v / etapas -> do trabalho	Visualização de info 1a parte					Visualização de info 2a parte								
Aulas	x	x	x	x	x		x	x	x	x	x	x	x	
Glossário de termos	x			x	x		x	x	x	x		x		
Planejamento evolutivo	xx x	x	x	x	x		x					x		
Experimentos				xx x			x					x		
Produto (3)				xx										
Processo p/ produzir produto					xx								x	
Visualização da apresentação					xx								x	
Apresentação														
Entrega de trabalho						X1	X2						X3	X
Documentação e Revisão				x	x	xx x	x	x	x	x	x	x	x	x
Qtd Iterações	3	2	3	3	4	5	2	2	2	3	2	4		2
D1	16	3 0	7	14	2 1	28	4	1 1	1 8	2 5	1	8	15	

x – tempo de 2 a 3hs D1- dia inicio duração: 14 semanas X- marcos

Junho: 16, 23, 30 Julho: 7, 14, 21, 28 Agosto: 4, 11, 18, 25 Setembro: 1, 8, 15
14.9.15

design ?

Variáveis de dados: 1 categoria – nominal (atividade, iterações na semana, iteração por atividade, dia do mes)

Variáveis visual: cor amarela – marco. Tamanho da letra (10) - atividades

Descrição: Mostra as atividades do fluxo trabalho para o design da visualização e execução

Etapas: Documentação: atualização 1 a 2 vezes na semana. Dois documentos: produto e processo e base de conhecimento para obter produto

Método de trabalho:

Estudo - < Consolidação - >



Comunicação : email e encontro semanal (aula ou ad hoc)

Referencias:

Kirk, Andy. Data Visualization: a successful design process. [Pack] Publishing.2012

Fry, Ben. Visualizando dados. Alta Books. 2008

Meirelles, Isabel. Design for information. Rockport Publishers. 2013

Stodder, David. TDWI Best practices report – Visual analytics for making smater decisions faster – Applying self-service BI technologies to data

<http://www.highcharts.com/download>

2. Experimentos

Ler, mostrar e interagir com conjunto de dados. Experimentação de idéias, conceitos, visualização, processo de visualização, análise e usabilidade da visualização.

Experimento 1- Ferramenta Processing

2.1 Instalação do Processing

<https://processing.org/download/> →



2.2.1 (19 May 2014)

Windows 64-bit

Linux 64-bit

Mac OS X

Windows 32-bit

Linux 32-bit

→ processing-2.2.1-macosx.zip

Criar pasta e copiar zip para pasta. Clicar no arq zipado:

	Hoje 19:26	--	Pasta
InstProcessing			
processing-2.2.1-macosx	Hoje 19:51	--	Pasta
p5	Hoje 19:44	--	Pasta
p5.zip	Hoje 19:33	281 KB	ZIP
Processing	Hoje 19:26	231,6 MB	Aplicativo
processing.py-0202-macosx	Hoje 19:51	--	Pasta
processing.py-0202-macosx.tgz	Hoje 19:36	87,9 MB	gzip tar archive
processing-2.2.1-macosx.zip	Hoje 19:24	122,6 MB	ZIP

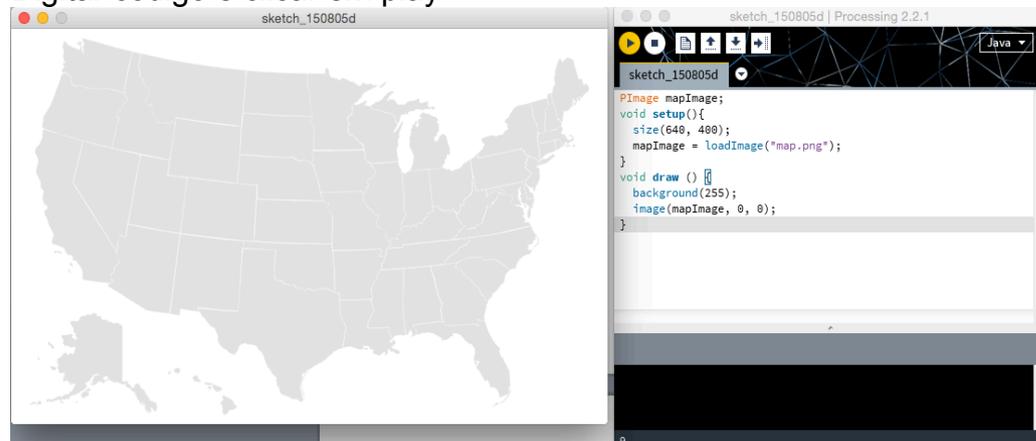
Criar uma área de dados e baixar arq de dados:

	Hoje 19:48	--	Pasta
Dados			
locations.tsv	18/07/2015 12:08	538 bytes	valores se...tabulação
map.png	18/07/2015 07:38	39 KB	Imagem PNG
Table.pde	18/07/2015 12:12	3 KB	Documento Processing

Clicar em Processing:

Copiar map.png para área de processamento do processing:

Digitar código e clicar em play



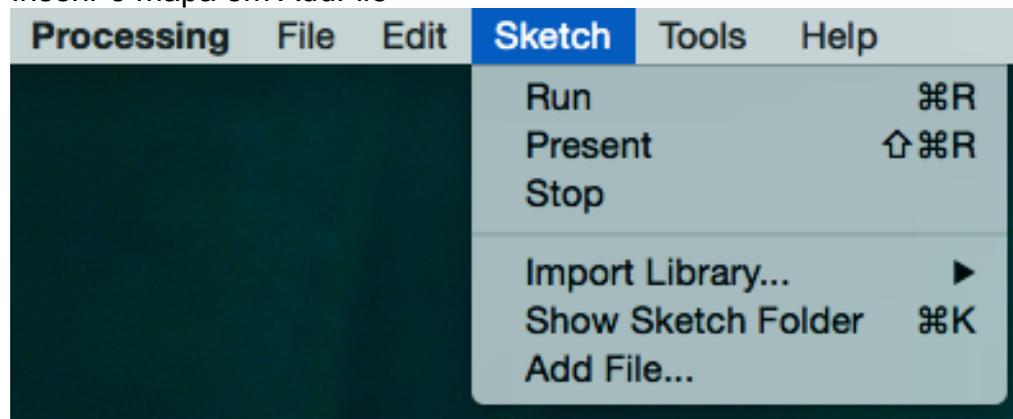
Criar pasta Resultado e salvar a aplicação nesta pasta :

▼ Experimento1	Hoje 20:35	--	Pasta
▼ Dados	Hoje 19:48	--	Pasta
locations.tsv	18/07/2015 12:08	538 bytes	valores se...tabulação
map.png	18/07/2015 07:38	39 KB	Imagem PNG
Table.pde	18/07/2015 12:12	3 KB	Documento Processing
Processing	Hoje 19:26	231,6 MB	Aplicativo
▼ Resultado	Hoje 20:35	--	Pasta
▼ sketch_150805d	Hoje 20:35	--	Pasta
▶ data	Hoje 20:35	--	Pasta
sketch_150805d.pde	Hoje 20:35	147 bytes	Documento Processing
sketch.properties	Hoje 20:07	48 bytes	Documen...ime Text 2

Imagem de background

<http://benfry.com/writing/map/map.png>

Inserir o mapa em AddFile

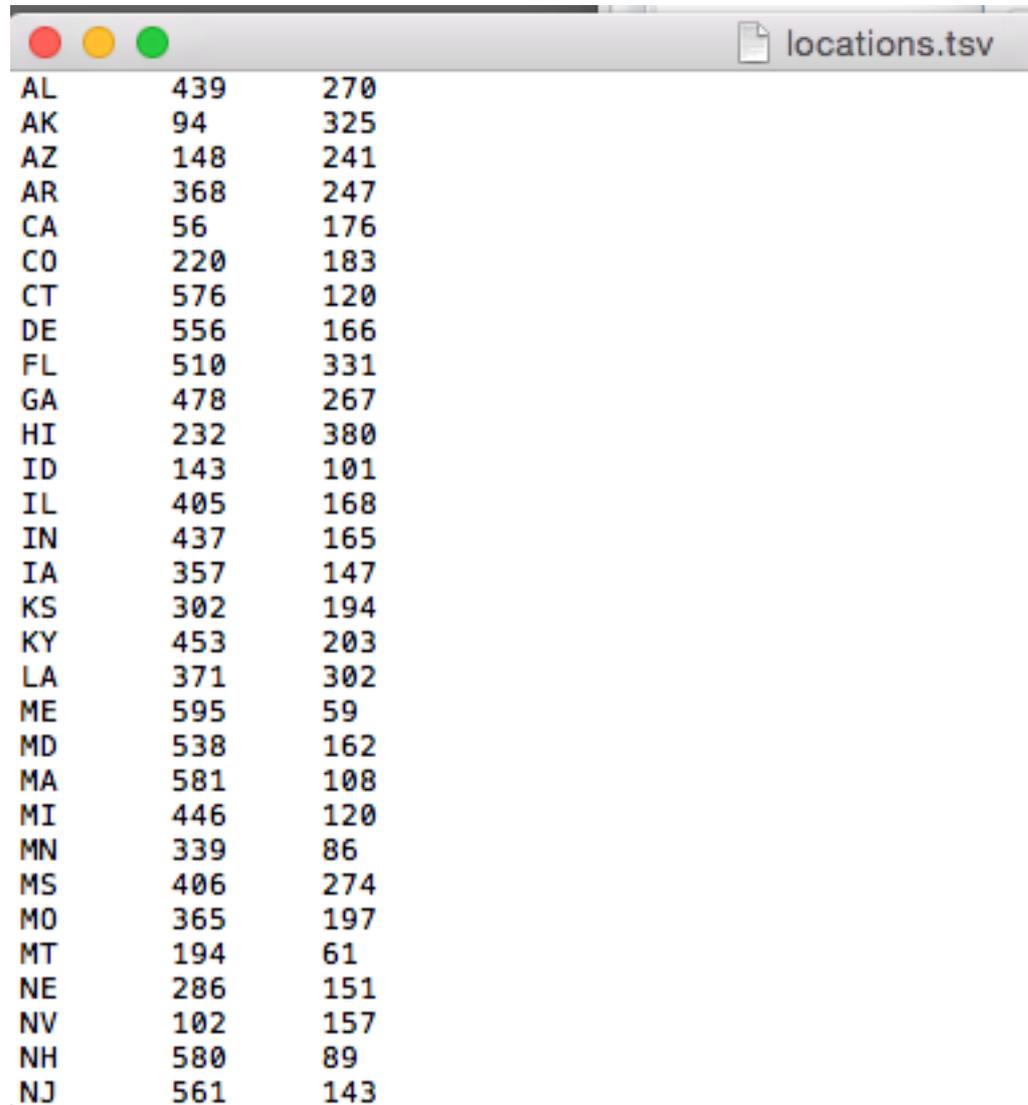


Clicar em Play no Processing



Coordenadas dos pontos que irão aparecer no mapa
<http://benfry.com/writing/map/locations.tsv>

locations.tsv



AL	439	270
AK	94	325
AZ	148	241
AR	368	247
CA	56	176
CO	220	183
CT	576	120
DE	556	166
FL	510	331
GA	478	267
HI	232	380
ID	143	101
IL	405	168
IN	437	165
IA	357	147
KS	302	194
KY	453	203
LA	371	302
ME	595	59
MD	538	162
MA	581	108
MI	446	120
MN	339	86
MS	406	274
MO	365	197
MT	194	61
NE	286	151
NV	102	157
NH	580	89
NJ	561	143

Código para ler arquivos dos dados de localização
<http://benfry.com/writing/map/Table.pde>

Table.pde

```
class Table {
  String[][] data;
  int rowCount;

  Table() {
    data = new String[10][10];
  }

  Table(String filename) {
    String[] rows = loadStrings(filename);
```

```

data = new String[rows.length][];

for (int i = 0; i < rows.length; i++) {
    if (trim(rows[i]).length() == 0) {
        continue; // skip empty rows
    }
    if (rows[i].startsWith("#")) {
        continue; // skip comment lines
    }

    // split the row on the tabs
    String[] pieces = split(rows[i], TAB);
    // copy to the table array
    data[rowCount] = pieces;
    rowCount++;

    // this could be done in one fell swoop via:
    //data[rowCount++] = split(rows[i], TAB);
}
// resize the 'data' array as necessary
data = (String[][]) subset(data, 0, rowCount);
}

int getRowCount() {
    return rowCount;
}

// find a row by its name, returns -1 if no row found
int getRowIndex(String name) {
    for (int i = 0; i < rowCount; i++) {
        if (data[i][0].equals(name)) {
            return i;
        }
    }
    println("No row named '" + name + "' was found");
    return -1;
}

String getRowName(int row) {
    return getString(row, 0);
}

String getString(int rowIndex, int column) {
    return data[rowIndex][column];
}

String getString(String rowName, int column) {
    return getString(getRowIndex(rowName), column);
}

```

```

}

int getInt(String rowName, int column) {
    return parseInt(getString(rowName, column));
}

int getInt(int rowIndex, int column) {
    return parseInt(getString(rowIndex, column));
}

float getFloat(String rowName, int column) {
    return parseFloat(getString(rowName, column));
}

float getFloat(int rowIndex, int column) {
    return parseFloat(getString(rowIndex, column));
}

void setRowName(int row, String what) {
    data[row][0] = what;
}

void setString(int rowIndex, int column, String what) {
    data[rowIndex][column] = what;
}

void setString(String rowName, int column, String what) {
    int rowIndex = getRowIndex(rowName);
    data[rowIndex][column] = what;
}

void setInt(int rowIndex, int column, int what) {
    data[rowIndex][column] = str(what);
}

void setInt(String rowName, int column, int what) {
    int rowIndex = getRowIndex(rowName);
    data[rowIndex][column] = str(what);
}

void setFloat(int rowIndex, int column, float what) {
    data[rowIndex][column] = str(what);
}

void setFloat(String rowName, int column, float what) {
    int rowIndex = getRowIndex(rowName);
    data[rowIndex][column] = str(what);
}

```

```

// Write this table as a TSV file
void write(PrintWriter writer) {
    for (int i = 0; i < rowCount; i++) {
        for (int j = 0; j < data[i].length; j++) {
            if (j != 0) {
                writer.print(TAB);
            }
            if (data[i][j] != null) {
                writer.print(data[i][j]);
            }
        }
        writer.println();
    }
    writer.flush();
}
}

```

Escrever código:

```

PImage mapImage;
Table locationTable;
int rowCount;

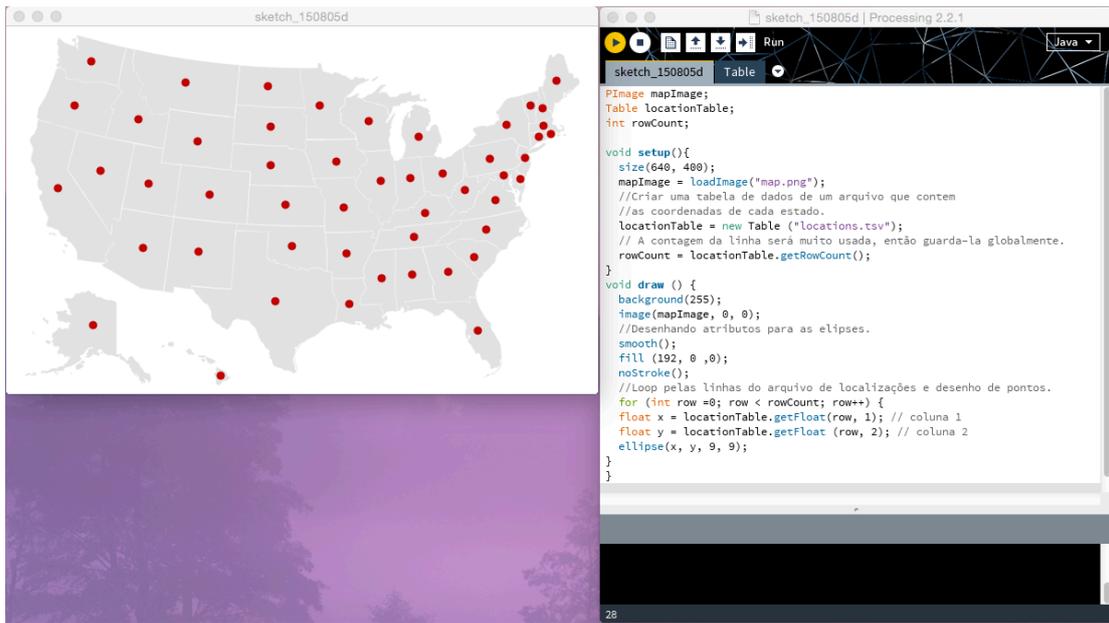
```

```

void setup(){
    size(640, 400);
    mapImage = loadImage("map.png");
    //Criar uma tabela de dados de um arquivo que contem
    //as coordenadas de cada estado.
    locationTable = new Table ("locations.tsv");
    // A contagem da linha será muito usada, então guarda-la globalmente.
    rowCount = locationTable.getRowCount();
}
void draw () {
    background(255);
    image(mapImage, 0, 0);
    //Desenhando atributos para as elipses.
    smooth();
    fill (192, 0 ,0);
    noStroke();
    //Loop pelas linhas do arquivo de localizações e desenho de pontos.
    for (int row =0; row < rowCount; row++) {
        float x = locationTable.getFloat(row, 1); // coluna 1
        float y = locationTable.getFloat (row, 2); // coluna 2
        ellipse(x, y, 9, 9);
    }
}

```

}

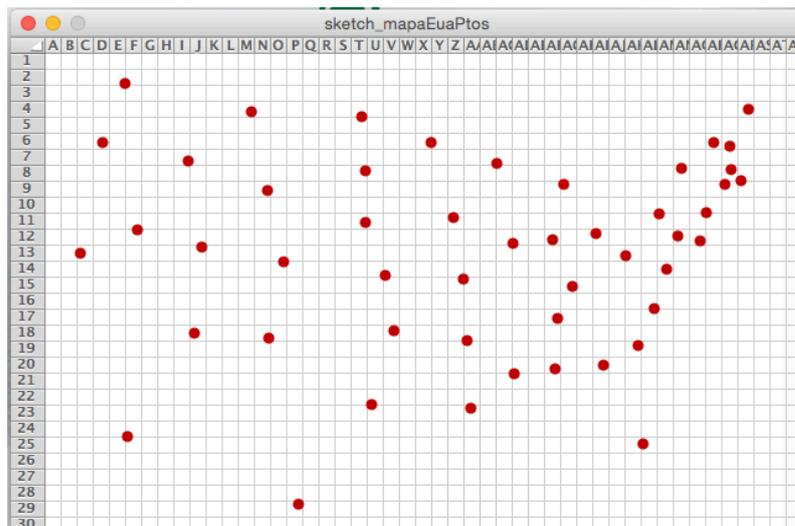


Salvar como sketch_mapaEuaPtos e depois Exportar na barra área de trab do processing:
 Organização dos arquivos na pasta de experimento:

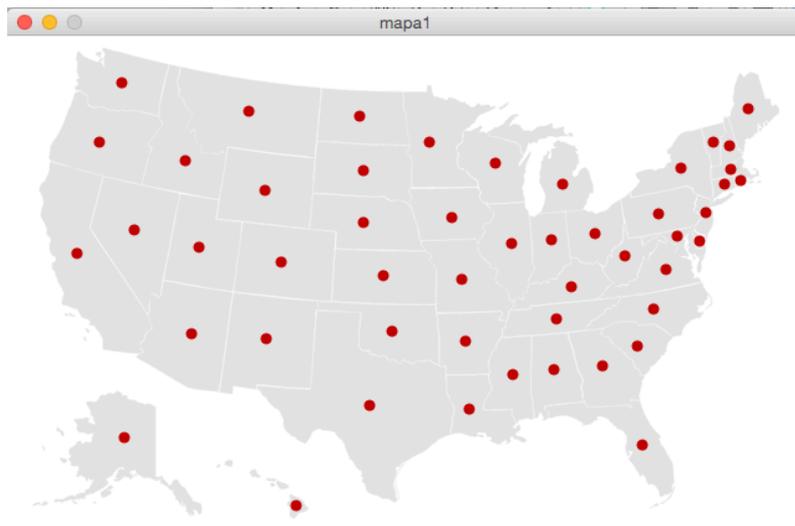
▼ Experimento1	Hoje 20:35	--	Pasta
▼ Dados	Hoje 19:48	--	Pasta
locations.tsv	18/07/2015 12:08	538 bytes	valores se...tabulação
map.png	18/07/2015 07:38	39 KB	Imagem PNG
Table.pde	18/07/2015 12:12	3 KB	Documento Processing
Processing	Hoje 19:26	231,6 MB	Aplicativo
▼ Resultado	Hoje 22:01	--	Pasta
▼ sketch_150805d	Hoje 22:06	--	Pasta
▼ data	Hoje 21:48	--	Pasta
locations.tsv	18/07/2015 12:08	538 bytes	valores se...tabulação
map.png	18/07/2015 07:38	39 KB	Imagem PNG
sketch_150805d.pde	Hoje 20:35	147 bytes	Documento Processing
sketch.properties	Hoje 20:07	48 bytes	Documen...ime Text 2
Table.pde	18/07/2015 12:12	3 KB	Documento Processing
▼ sketch_mapaEuaPtos	Hoje 22:06	--	Pasta
▼ application.macosx	Hoje 22:06	--	Pasta
sketch_mapaEuaPtos	Hoje 22:02	98,7 MB	Aplicativo
▼ source	Hoje 22:02	--	Pasta
sketch_mapaEuaPtos.java	Hoje 22:02	5 KB	Código-fonte Java
sketch_mapaEuaPtos.pde	Hoje 22:02	779 bytes	Documento Processing
Table.pde	Hoje 22:02	3 KB	Documento Processing
▼ data	Hoje 22:01	--	Pasta
locations.tsv	18/07/2015 12:08	538 bytes	valores se...tabulação
map.png	18/07/2015 07:38	39 KB	Imagem PNG
sketch_mapaEuaPtos.pde	Hoje 22:01	779 bytes	Documento Processing
sketch.properties	Hoje 20:07	48 bytes	Documen...ime Text 2
Table.pde	Hoje 22:01	3 KB	Documento Processing
InstProcessing	Hoje 19:26	--	Pasta
▼ processing-2.2.1-macosx	Hoje 19:51	--	Pasta
p5	Hoje 19:44	--	Pasta
p5.zip	Hoje 19:33	281 KB	ZIP
Processing	Hoje 19:26	231,6 MB	Aplicativo
processing.py-0202-macosx	Hoje 19:51	--	Pasta
processing.py-0202-macosx.tgz	Hoje 19:36	87,9 MB	gzip tar archive
processing-2.2.1-macosx.zip	Hoje 19:24	122,6 MB	ZIP

▼ sketch_mapaEuaPtos	Hoje 22:06	--	Pasta
▼ application.macosx	Hoje 22:06	--	Pasta
sketch_mapaEuaPtos	Hoje 22:02	98,7 MB	Aplicativo
▼ source	Hoje 22:02	--	Pasta
sketch_mapaEuaPtos.java	Hoje 22:02	5 KB	Código-fonte Java
sketch_mapaEuaPtos.pde	Hoje 22:02	779 bytes	Documento Processing
Table.pde	Hoje 22:02	3 KB	Documento Processing
▼ data	Hoje 22:01	--	Pasta
locations.tsv	18/07/2015 12:08	538 bytes	valores se...tabulação
map.png	18/07/2015 07:38	39 KB	Imagem PNG
sketch_mapaEuaPtos.pde	Hoje 22:01	779 bytes	Documento Processing
sketch.properties	Hoje 20:07	48 bytes	Documen...ime Text 2
Table.pde	Hoje 22:01	3 KB	Documento Processing

Trocando o map.png por uma imagem de papel milimetrado ficaria:



Qual tipo de área o 2o ou 3o dígito representa no ZipCode dos EUA?



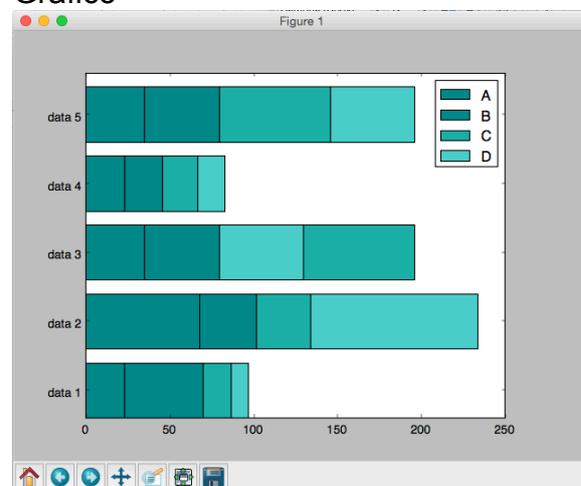
2.2 Experimento 2 -Ferramenta Anaconda Código python 3

```
import numpy as np
from matplotlib import pyplot as plt

dataset = [{'A':23, 'B':47, 'C':17, 'D':10, 'E':10},
           {'A':34, 'B':68, 'C':32, 'D':100, 'E':10},
           {'A':35, 'B':45, 'C':66, 'D':50, 'E':10},
           {'A':23, 'B':23, 'C':21, 'D':16, 'E':10},
           {'A':35, 'B':45, 'C':66, 'D':50, 'E':10}]
data_orders = [['A', 'B', 'C', 'D', 'E'],
               ['B', 'A', 'C', 'D', 'E'],
               ['A', 'B', 'D', 'C', 'E'],
               ['B', 'A', 'C', 'D', 'E'],
               ['A', 'B', 'C', 'D', 'E']]
colors = ["#008B8B", "#008B8B", "#20B2AA", "#48D1CC"]
names = sorted(dataset[0].keys())
values = np.array([[data[name] for name in order] for data,order in
zip(dataset, data_orders)])
lefts = np.insert(np.cumsum(values, axis=1),0,0, axis=1)[:, :-1]
orders = np.array(data_orders)
bottoms = np.arange(len(data_orders))

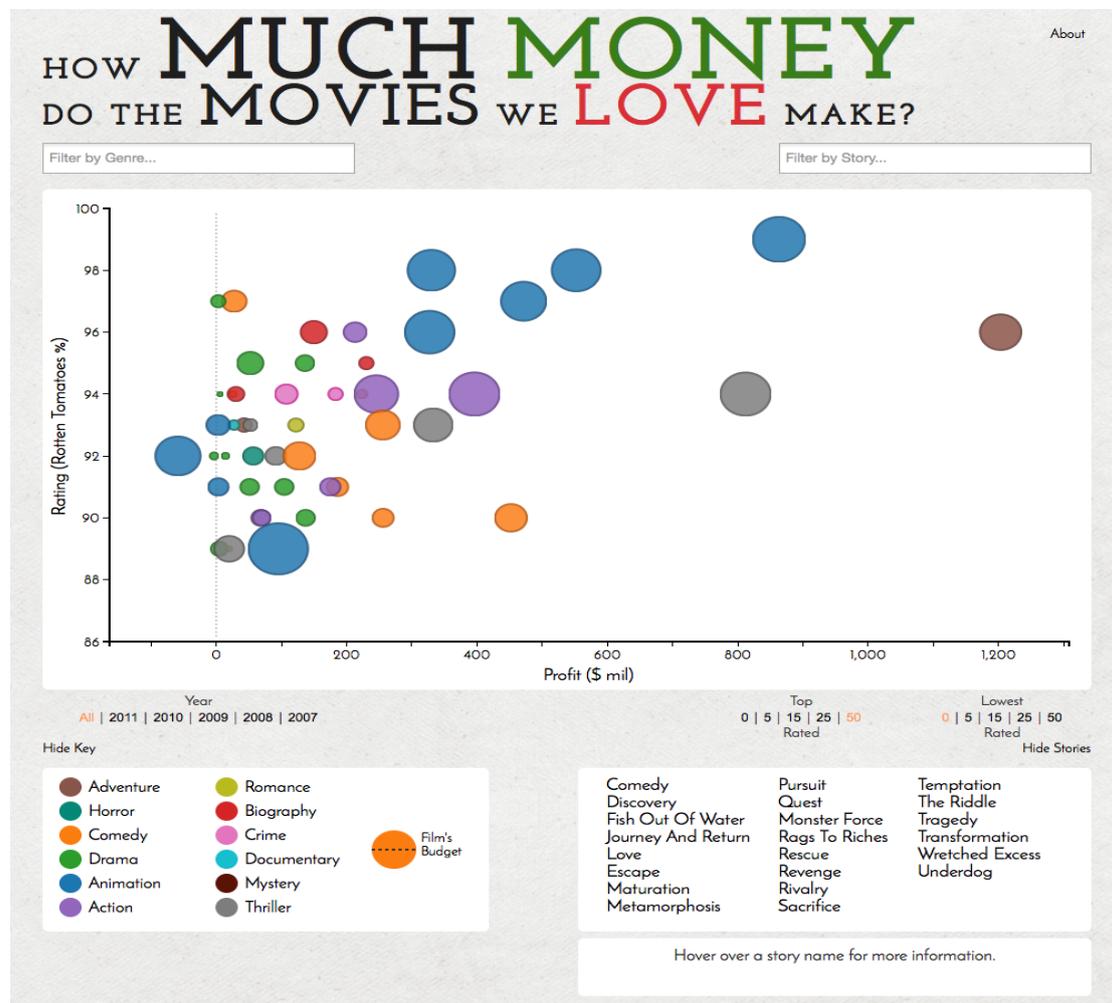
for name, color in zip(names, colors):
    idx = np.where(orders == name)
    value = values[idx]
    left = lefts[idx]
    plt.bar(left=left, height=0.8, width=value, bottom=bottoms,
            color=color, orientation="horizontal", label=name)
plt.yticks(bottoms+0.4, ["data %d" % (t+1) for t in bottoms])
plt.legend(loc="best", bbox_to_anchor=(1.0, 1.00))
plt.subplots_adjust(right=0.85)
plt.show()
```

Gráfico



2.3. Análise de Visualizações

Múltiplas variáveis, expressando camadas extras de significado por trás das propriedades de uma única marca:



(<http://vallandingham.me/vis/movie/>)

Descrição: Visualização gráfico de bolhas, cada marca envolve uma combinação de diversas variáveis visuais que representam uma gama de variáveis de dados diferentes:

A posição no eixo x representa o lucro de um filme.

A posição sobre o eixo dos y representa a classificação média avaliação percentual de cada película

A área do círculo representa o orçamento do filme. • A cor do círculo (matiz) representa gênero de filme. • Usuários interagindo com este projeto baseado na web também vai descobrir um rótulo de texto exibindo os valores brutos passando o mouse sobre uma das bolhas. O texto não é universalmente tratado como uma variável visual.

2.4 Maneiras de representar 2 números:

<http://blog.visual.ly/45-ways-to-communicate-two-quantities/>

2.5 Tipos de gráficos

<http://www.tecmundo.com.br/excel/1745-saiba-qual-tipo-de-grafico->

representa-melhor-os-seus-dados-no-excel-2007.htm

2.6 Tipos de chart's

<http://www.highcharts.com/download>

3. Glossário de termos (visualização e seus atributos e processos)

Livro: Data visualization: a successful design process. Andy Kirk (teoria da visualização)

Visualização – É a representação e apresentação de dados que explora as nossas capacidades de percepção visual, a fim de ampliar o conhecimento. Ex de forma de comunicação: um gráfico, um interação on-line, uma tela de toque, ou talvez um infográfico em um jornal.

Elementos de visualização:

Representação dos dados - É como você descreve dados através de uma escolha de formas físicas (linha, um barra, um círculo, ou qualquer outra variável visual) para retratar seus atributos.

Apresentação dos dados - vai além da representação de dados e preocupações como integrar seus dados comunicando de forma global, incluindo a escolha de cores, anotações e recursos interativos.

Percepção visual - relaciona-se com a compreensão científica de como os nossos olhos e cérebros processam informações de forma mais eficaz, . (raciocínio espacial, reconhecimento de padrões, e o pensamento do grande-retrato).

Cognição- é o processamento da informação em pensamentos, idéias e conhecimento com o objetivo de ficarmos mais bem informados sobre um assunto.

Avaliação reflexiva- Avaliação por meio da análise do todo. Ex: H2O – apaga o fogo

Avaliação analítica- Avaliação por meio de análise das partes. Ex: O e H – pega fogo

Efeito de desejado: Pesquisa, persuasão, técnica criativa, aprender/aumentar conhecimento, resposta a questões, mudança de comportamento, conduzir análises, jogar com os dados, monitorar sinais, contar histórias, disparar perguntas, contextualizar dados, achar padrões / não padrões, descoberta acidental, familiarizar-se com dados, formar opinião, enfatizar questões, inspirar, apresentar argumentos, apoiar decisões, experimentação, arte/estética, impacto de choque.

Ambiente do projeto-

É um ambiente de sala de reuniões com uma pequena coleção de colegas mais velhos que têm conhecimento do assunto de domínio existentes?

É uma grande gama de clientes, cobrindo todos os dados demográficos sociais, mas potencialmente representando um público cativo para o assunto?

É um público completamente global, indefinido sem características-em influenciar num sentido nenhum alvo específico, apenas qualquer um e todos?

É uma troca de um-para-um com um gerente?

É um compromisso inteiramente pessoal entre você e dados em um desejo de conhecer e explorar os dados a si mesmo?

Função da visualização-

transmitir um retrato explicativo de dados para um leitor

Fornecer uma interface para os dados, a fim de facilitar a exploração visual.

Usar dados como uma exposição de auto-expressão

Tom da visualização-

Pragmática e analítica – proporciona retratos rápidos, eficientes e precisos dos dados. Usado para iteragir e aprender desempenho operacional e, descobrir padrões.

Emotivo e abstrato – criação de uma estética que retrata uma história geral ou sensação de padrão.

Fatores chaves-

Objetivo, pressão do tempo, custos, pressões do cliente (mudanças nos requisitos, novas demandas, interferência na solução de design, identidades organizacionais visuais ou de marca, regras de layout, diretrizes editoriais e quadros técnicos)formato (estatico ou interativo), capacidade tecnica (infraestrutura)

Papeis no design da visualização-

Iniciador- define direção e o tom do projeto formato / plataforma pretendida da solução algumas das principais questões tecnológicas.

Cientista de dados – minerador de dados responsavel pela terceirização, aquisição, manipulação e preparação dos dados. aprende sobre os padrões, relacionamentos e propriedades descritivas dos dados.

Jornalista - é o contador de histórias, a pessoa que estabelece a abordagem narrativa ao contexto do problema da visualização. Trabalhando com o cientista de dados e o iniciador formulam as questões de dados que ajudam a manter o foco do projeto em seu caminho editorial pretendida. Com base faíscas iniciais do iniciador de idéias, o jornalista vai desenvolver uma mentalidade mais profundo do pesquisador para realmente explorar as oportunidades analíticas.

Cientista da computação -é o executor constroi a solução. Busca a forma mais eficaz e eficiente a coleta de dados, manipulação e visualização de atividades de pré-produção.

Designer - em harmonia com o cientista da computação, vai entregar a solução. Eles têm o olho para o detalhe visual, um toque de inovação e estilo

e são totalmente capazes de apreciar as possibilidades potenciais que existem. No entanto, eles também têm a disciplina necessária para seguir a mensagem estabelecida pelo iniciador e assumida pelo jornalista. Eles respeitam as capacidades do cientista da computação em termos do que soluções poderiam ser viável, mas eles mesmos têm a visão de helicóptero- como para racionalizar e argumentar que as coisas vão funcionar e não vai funcionar, e por quê. Sua responsabilidade principal é também para ser capaz de assegurar a harmonia da solução entre sua forma e sua função, garantindo-se esteticamente atraente para atrair o leitor, enquanto fundamentalmente entregar o pretendido, mensagem comunicada.

Cientista cognitivo - têm o conhecimento sobre a percepção visual como o olho e o cérebro trabalhar de forma mais eficaz e eficiente. Eles também têm um profundo conhecimento sobre conceitos tais como as Leis da Gestalt, teorias da comunicação, teorias de cores e princípios de interação humano-computador. Além disso, eles são capazes de informar o processo de design em relação às complexidades de como a mente funciona em termos de memória, a atenção, tomada de decisão e mudança de comportamento.

Communicator - Negociador e apresentador, que opera na porta de entrada do cliente, ajudando a informar todos aqueles que estão envolvidos no progresso, requisitos, problemas e soluções. Entende necessidades, apreciando restrições, reconhecendo possibilidades e, em seguida, em última instância lançamento, divulgação e apresentando o trabalho final. Uma capacidade de articular e explicar as coisas para diferentes tipos de pessoas, técnicos e não técnicos, e ser capaz de gerir expectativas e relacionamentos é vital.

Gerente ou coordenador- gerenciam o processo do projeto e de seu progresso, garantindo que é coeso, no tempo, e em mensagem e fatores-chave em torno do projeto. Devem estar preocupados com questões de integridade ética em torno de visualização.

Variáveis de dados -

Categórico nominal, Categórico ordinal, Quantitativo (escala-intervalo)

Quantitativo (escala - razão-scale)

Exemplo: Olimpíadas

Data	Tipo	Valor
Evento	Quantitativo(escala-intervalo)	27 anos (1896-2012)
Medalha	Categórico ordinal	Ouro, prata e bronze
Atleta	Categórico nominal	+ de 1500 nomes de atletas
Resultado	Quantitativo (escala-razão)	Resultados de corrida (9.59% > 4:02:59)
Pais	Categórico nominal	96 nomes de países

Variáveis visuais- forma específica que atribuímos aos dados: comprimento ou a altura de uma barra, a posição de um ponto sobre um eixo, a cor de um país no mapa, ou a ligação entre dois nós de uma rede.

Métodos para a visualização de dados-

Comparando valores categóricos

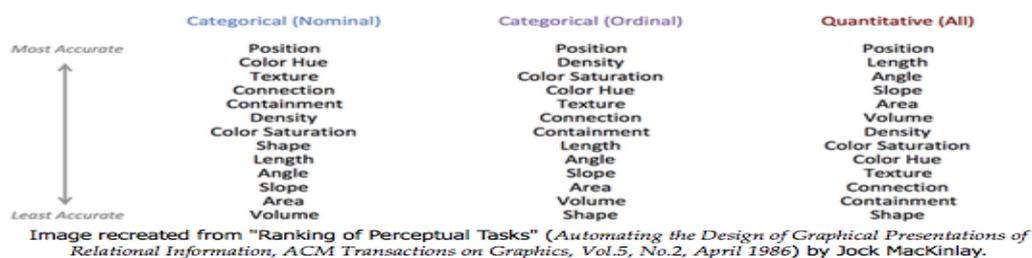
Avaliando hierarquias e relacionamentos parte de um todo

Mostrando mudanças ao longo de dados geo-espaciais

Mapeamento de tempo

Relacionamentos gráficos

Ranking da tarefas perceptivas –



Transformando para qualidade - Conectando as lacunas causadas por falta de dados, removendo duplicatas, limpando valores errados, e manipulação de personagens incomuns.

Transformando para análise-

Analisar (dividir) todas as variáveis, como a extração de ano a partir de um valor de data

Mesclando variáveis para formar novos, como a criação de um nome de todo fora do título, nome próprio e apelido

Conversão de dados qualitativos / free-text em valores codificados ou palavras-chave.

Obtenção de novos valores de outros, como o gênero de título ou um sentimento de alguns dados qualitativos .

Criação de cálculos para utilização em análise, tais como proporções percentuais .

Remoção de dados redundantes para os quais você não tem nenhum uso planejado

Exemplo: <http://www.periscopic.com/work/yahoo-mail/>

Eles decidiram sobre a seguinte estratégia:

As estatísticas manchete mostrado nos títulos e apresentados em uma série de ilustrações suplementares em todo o projeto seria representativo das

quantidades de dados completos.

Para a visão geo-espacial, um algoritmo cuidadosamente projetado foi aplicado para extrair uma amostra representativa de dados. Isso seria mais do que suficiente para captar as nuances da atividade visto com o conjunto de dados completo e evitaria as dificuldades práticas técnicas envolvidas na tentativa de mostrar 00 por cento dos dados.

Os dados geográficos foram agrupados a uma cidade ou agregado regional, representada pelas posições e tamanhos de círculo, para ajudar a tirar os sinais principais e padrões.

Resolução - Quando você se depara com decisões similares, embora talvez raramente na mesma escala, normalmente você vai ter essas opções disponíveis para você:

Resolução completa: Traçando todos os dados disponíveis como marcas de dados individuais. • resolução filtrado: Excluir registros com base em determinados critérios.

Resolução das duas mãos: "Roll-up" os dados por, por exemplo, mês, ano, ou de uma categoria específica.

Resolução Amostra: Aplicar certas regras de seleção matemática para extrair uma fração de seus dados em potencial. Esta é uma tática particularmente útil durante a fase de concepção se você tem muito grandes quantidades de dados e quer desenvolver rapidamente mock-ups ou testar ideias.

Resolução Manchete: Apenas mostrando os totais estatísticos globais.

Consolidação- Camadas adicionais de dados podem ser obrigados a ser combinado ("mashed-up") com o nosso conjunto de dados existente, aplicado para realizar cálculos adicionais, ou apenas para se sentar ao lado deste recurso inicial para ajudar a contextualizar e aprofundar o âmbito da nossa comunicação.

<http://visualdata.dw.com/specials/bildung/en/>

pequeno número de ângulos interessantes sobre a educação: alfabetização por região, as taxas de alfabetização por país / gênero, As taxas de escolarização, e as despesas com educação em relação a militar.

Comparações e proporções-

Faixa e distribuição: Descobrir o intervalo de valores ea forma de sua distribuição dentro de cada variável e em combinações de variáveis .

Ranking: Aprender sobre a ordem dos dados em termos de magnitude geral, identificando o grande, médio e pequenos valores. Usando a análise visual para encontrar histórias se segue é uma citação de Ben Schneiderman: ". A visualização lhe dá respostas às perguntas que você não sabia que você tinha" Demonstrando Foco Editorial e aprender sobre seus dados.

Medidas: Olhando para além apenas da ordem de grandeza para saber mais sobre a importância dos valores absolutos .

Contexto: A julgar valores contra o contexto de médias, desvios-padrão, metas e previsões. Usando métodos como um gráfico de barras vai permitir a comparação entre os valores e categorias para escolher o tipo de qualidades físicas .

Tendências e padrões-

Direção: são valores que mudam em um, para baixo ou para cima movimento plano?.

Taxa de variação: Como íngreme ou apartamento não padrão muda ocorrer? Vemos um padrão consistente, linear, ou é muito mais exponencial em forma?

Flutuação: Vemos evidências de padrões consistentes ou há flutuação significativa? Talvez haja um certo ritmo, como a sazonalidade, ou talvez padrões são mais aleatória .

Significado: Podemos determinar se os padrões que vemos são sinais significativos ou simplesmente representar o ruído dentro dos dados?

Interseções: Não observamos nenhum cruzamentos importantes ou sobreposições entre as variáveis, pontos de cruzamento que indicam uma mudança significativa na relação? Usando um gráfico de linha é um método perfeitamente adequado para observar padrões e tendências, como vemos abaixo:

Relações e conexões-

Exceções: podemos identificar quaisquer valores significativos que ficam fora da norma, tais como valores atípicos que mudam a dinâmica de uma determinada gama de variáveis? .

Correlações: Existe evidência de correlações fortes ou fracas entre combinações de variáveis?

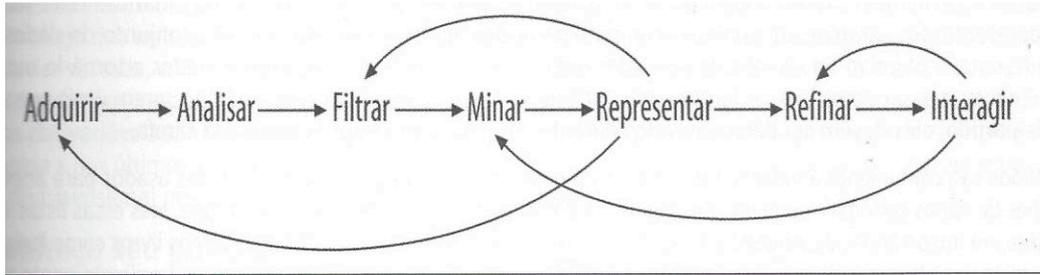
Associações: Podemos identificar quaisquer ligações importantes entre diferentes combinações de variáveis ou valores?

Clusters e lacunas: Onde é que há evidência de ser "apanhado" de dados? Onde é que há lacunas nos valores e pontos de dados?

As relações hierárquicas: A determinação da composição, distribuição e relevância das categorias e subcategorias dos dados. Usando um gráfico de dispersão irá permitir visibilidade desses tipos de relacionamentos, como mostrado abaixo:

Livro: Visualizando dados, Ben Fry (processos de visualização)

Estágios da visualização de dados-



Qual é a pergunta? – Todas as soluções de dados começam com uma pergunta e terminam com a construção narrativa que fornece uma resposta clara. Ex.: No experimento 2.1 a pergunta era como os códigos Zipcode se relaciona com sua geografia?

Adquirir – Obtenção dos dados. A aquisição diz respeito de onde os dados são obtidos e baixados. Estruturar dados para facilitar a restauração de subconjuntos comuns.

Analisar – Formata um rótulo para cada parte dos dados com uso específico. Cada campo é formatado como um tipo de dado que irá ser manipulado num programa de conversão. O formato genérico é string com campos separados por um token qualquer.

Filtrar – Remover porções irrelevantes por uso específico. Ou trabalho matemático para colocar num modelo matemático ou normaliza-lo. (conversão aceitável de números)

Minar – Envolve matemática, estatística e data mining. Por exemplo comparar valores para achar mínimos e máximos.

Representar – Forma básica que um conjunto de dados vai tomar. Por exemplo lista, árvore,.. O modo de representar pode influenciar o adquirir (os dados que obtem) e o filtrar (que partes a extrair).

Refinar – métodos de design gráfico são usados estabelecendo por exemplo hierarquias ou mudando atributos (como cor) que contribuem para legibilidade.

Interagir – Controlar e explorar os dados. Por exemplo selecionar o subgrupo de dados ou mudar o ponto de vista. Pode afetar o refinamento. Uma mudança no ponto de referencia pode exigir que os dados sejam reorganizados diferentemente. Pode usar zoom, informar código,..

Processing – Ambiente de programação para visualização com ênfase em animação e resposta instantânea através da interação. Pode usar Java, Javascript ou python. Da suporte para desenhar com OpenGJ, ler arquivos xml e salvar imagens complexas em PDF. Um programa em Processing é chamado de sketch. Estilo de programação como se fossem roteiros.

Exemplo: Desenhar uma linha: line (15, 25, 70, 90), significa coordenadas x=15, y=25, x1=70 e y1=90. A coordenada (0,0) fica no canto superior esquerdo.

Funções do Processing:

Obter

loadStrings(), loadBytes()

Analisar

Split()

Filtrar

For(), if(item[i].startsWith())

Garimpar

Min(), max(), abs()

Representar

Map(),beginShape(),endShape()

Interagir

mouseMoved(), mouseDragged

Livro: Design for information, Isabel Meireles (aplicação da visualização)

Exibição visual da informação: infográficos e design da informação.

Infográfico: explica informação visual e conta história. E: Mapa de metro, funções do cérebro humano, mapa do tempo para explicar o fenômeno natural.

Design de informação: Usado para descrever informação do design de informação.

Tipos de estruturas:

Hierárquica: É um conjunto ordenado onde os elementos e/ou subconjuntos são organizados de um para o outro.

Representação: esquemas empilhados e encaixados.

Geometria: sistema cartesiano (treemap, diagrama de venn, dendrograma,...), sistema polar (radial,...), outros (árvore hiperbólica,...)

Visual hierárquico: Ou contraste, ou como tautologia. Precisa de uma codificação visual bem definida. Percepção visual e sistema de trabalho cognitivo.

Codificação espacial: propriedades espaciais (posição e tamanho) separadamente das propriedades do objeto (forma, cor, textura,...).

Proximidade (descreve tendência para grupo de elementos visuais, relata características locais e é essencial como os elementos estão associados espacialmente)

Percepção visual e cognitiva:

Estágio 1: processamento paralelo rápido para extrair características básicas (cor e tamanho)

Estágio 2: processamento serial lento para extrair modelos (estruturas).

Estágio 3: processamento orientado ao objeto com informação reduzida para formar pensamento visual.

Relacional: Redes

Temporal: Linha do tempo e fluxo. Eventos no tempo.

<http://ngm.nationalgeographic.com/map/map-day>

<http://www.popsci.com/content/wordfrequency-visualization>

<http://artport.whitney.org/commissions/idealine/Idealine.html>

http://www.nytimes.com/interactive/2008/02/23/movies/20080223_REVENUE_GRAPHIC.html?_r=0

<http://visualdata.dw.com/en/wikipedia/>

Espacial: Mapas

<http://prettymaps.stamen.com/201008/-10.00/-23.1528/-43.5835>

<http://dotspotting.org/u/2560/sheets/3405-c=2.00/12.3/32.7>

<http://colorbrewer2.org>

Espaço-tempo: Fluxo de trafego. Generalização e agregação de trajetórias.
<http://pmcruz.com/information-visualization/traffic-in-lisbon-condensed-in-one-day>

<http://www.itc.nl/personal/kraak/1812/3dnap.swf>

Textual:

<http://benfry.com/traces/>

<http://www.wordle.net/create>

<http://archive.stefaner.eu/projects/revisit-twitter-visualization>

<http://www->

958.ibm.com/software/data/cognos/manyeyes/page/Phase_Net.html

<http://republicofletters.stanford.edu>

<http://www.neoformix.com/Projects/WordHearts/index.html>

<http://tagcrowd.com>

Slides: Marino, Tiago Badre. Representações de dados espaciais – Raster x Vetor x TIN

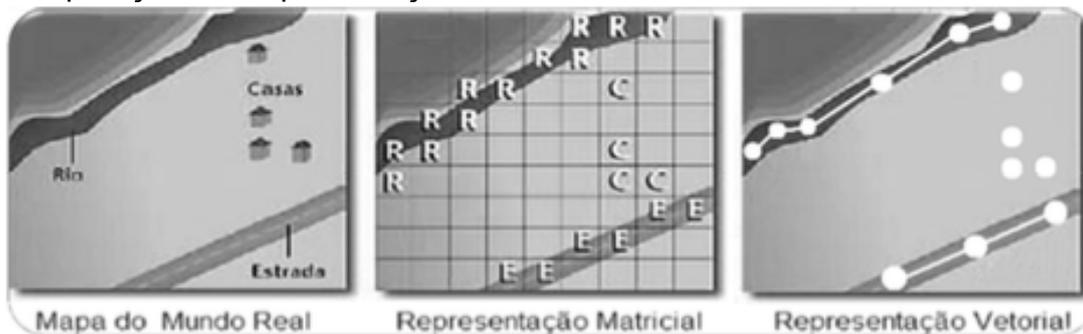
<http://www.ufrj.br/lga/tiagomarino/aulas/5> - Representação de Dados Espaciais - Raster x Vetor x TIN.pdf

Processos de redução de dados: (abstração e generalização)

Abstração: Capacidade de visão de alto nível que permite examinar problemas de forma a selecionar grupos comuns -> encontrar generalidade (para compreender o problema) -> construir modelos.

Representação de dados espaciais: Raster (matricial) , Vetor e TIN

Comparação das representações:



Planos de informação (camadas) – cada camada relacionada é um único tema. Não tem interseção entre camadas. Um elemento não pertence a duas classes do mesmo tema simultaneamente.

Componentes do SIG- sistema de informação geográfica

Gráfica ou espacial:

Localização em coordenadas (UTM ou Geodésicas)

Geometria em área, perímetro e forma- ponto, linha ou área.

Topologia – relação entre os entes espaciais: sobreposição, interseção,..

Alfanumérica:

Atributos temáticos e temporários representados em forma de tabela estruturada ou banco de dados.

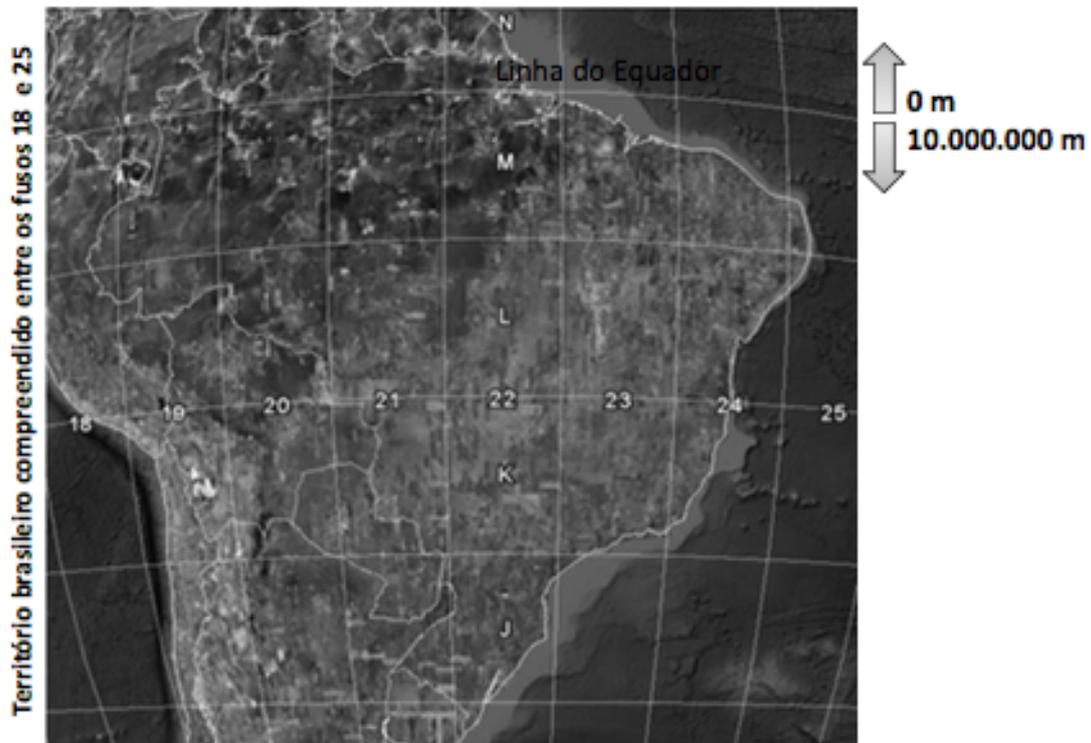
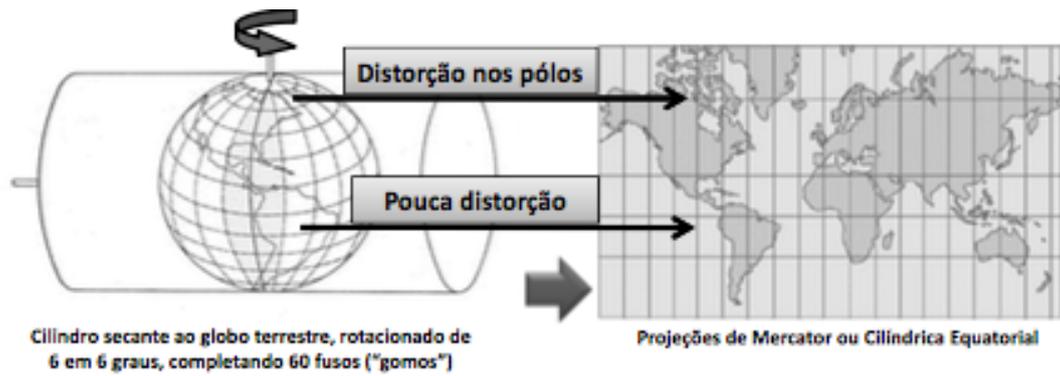
Sistema de coordenadas – São projeções passando de superfícies curva para plana com mínimo de deformações. São usadas para definir a localização espacial dos domínios espaciais dos objetos.

Tipos de superfícies – cilindro, cone, poliedro.

Sistema de projeção cartográfica- Método, usado em mapeamentos, que envolve transformações matemáticas, pelas quais a superfície terrestre (curva) é representada numa superfície plana.

Sistema de coordenadas universal transversa de Mercator (UTM)- É cilíndrica e conforme. Meridianos e os paralelos são linhas retas que se cortam em ângulos retos.

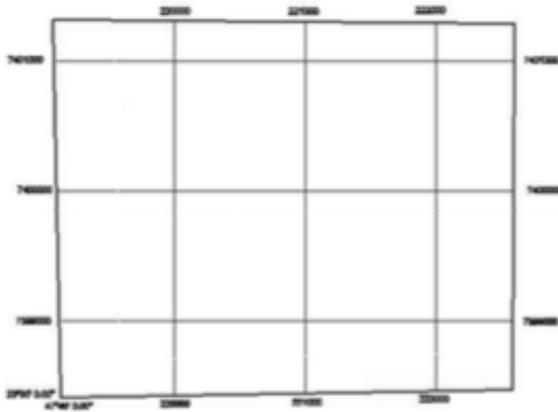
http://www.lapig.iesa.ufg.br/lapig/cursos_online/gvsig/a_projeo_utm.html



8 fusos UTM

vai do fuso 18 situado no extremo oeste a fuso 25 no extremo leste do território.

Sistema de coordenadas cartesiano- Definido a partir de 2 a 3 eixos, com ângulos ortogonais num sistema métrico.

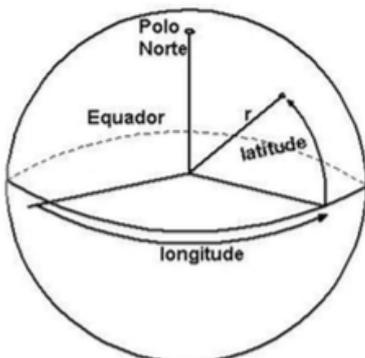


Exemplo de Sistema de Coordenadas Cartesianas: UTM

Sistema de coordenadas esférico – Usa 2 ângulos de rotação e uma distância radial para especificar uma localização, expresso em graus, minutos e segundos.

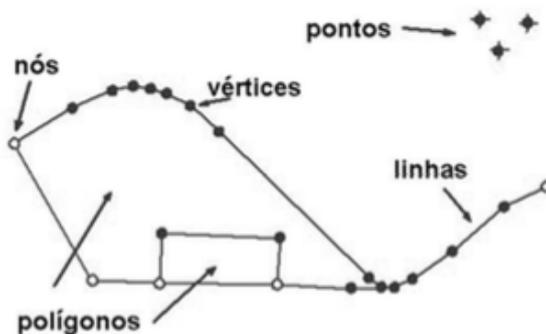
Longitude – definida pela ângulo diedro formado pelo plano Meridiano de Greenwich e pelo meridiano local.

Latitude – Ângulo que a vertical forma com a sua projeção sobre o plano equador.



Exemplo de Sistema Esférico: Coordenadas Geodésicas

Representação vetorial-



Representação dos elementos primários do modelos vetorial: pontos, linhas e polígonos

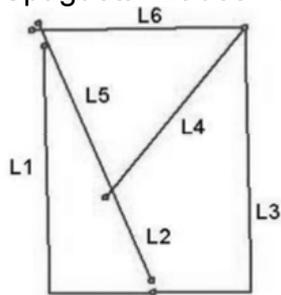
Ponto: par de ordenadas para representar a localização. Pode armazenar atributos como: área, proprietário, tipo de uso, valor venal, ...

Linha: conjunto ordenado de pontos interligados por segmentos de reta (polígono aberto). Ex.: estradas, rede de linhas transmissão de energia elétrica, saneamento,...

Polígono: Conjunto ordenado de pontos interligados, onde o 1o e último pto coincidem.

Generalizada: Simplificação ou eliminação dos detalhes (áreas muito pequenas ou polígonos viram pontos).

Spaguetti: Dados vetoriais compostos por segmentos de linha



Exemplo de Modelo Spaguetti

Mapa topológico: Metrô- o que importa é a vizinhança

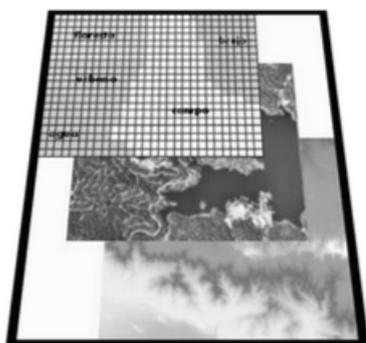


Representação da topologia (Metrô de Londres)



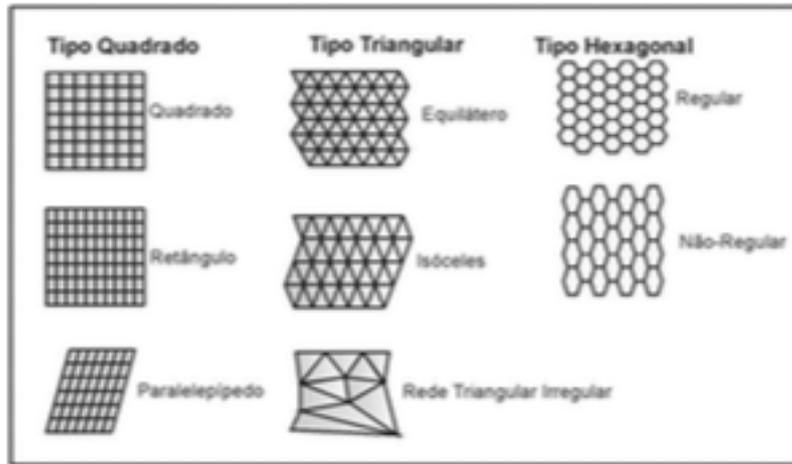
Representação da geometria (Metrô de Paris)

Representação Raster (matricial)



Representação Raster: modelo discreto

Formas:



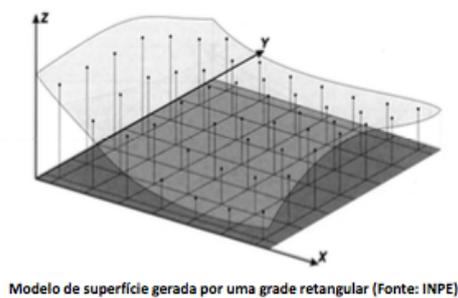
Formatos geométricos regulares de células raster

FONTE : A. Jon Kimerling, Associação Cartográfica Internacional. Versão em Português: UNESP

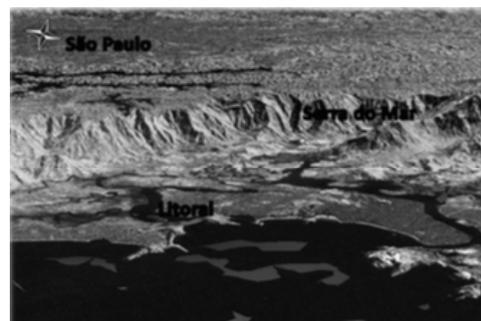
Comparação das representações:

CARACTERÍSTICA	RASTER	VETOR
Estrutura de Dados	Geralmente simples	Geralmente complexo
Requisito de Armazenagem	Grande para a maioria dos dados sem compressão	Pequena, para a maior parte dos dados
Conversão de Sistema de Coordenadas	Pode ser lenta, devido ao volume, e requerer reamostragem	Simples
Precisão Posicional	Degaus contornando células. Depende da resolução adotada	Limitado somente pela qualidade posicional de levantamento
Acessibilidade	Fácil para modificar através do uso de programas; estrutura de dados simples	Freqüentemente complexo
Visualização e saída	Bom para imagens, mas para feições discretas, pode mostrar efeito escada	Parecido com mapas, com curvas contínuas; pobre para imagens
Relações espaciais entre objetos	Relacionamentos espaciais devem ser inferidos	Relacionamentos topológicos entre objetos disponíveis
Análise e Modelagem	Superposição e modelagem mais fáceis	Álgebra de mapas é limitada

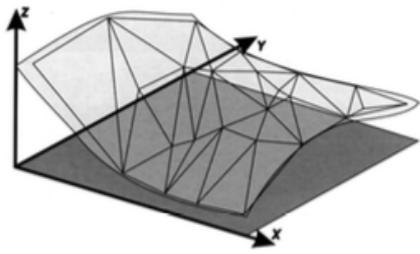
Representação grid –



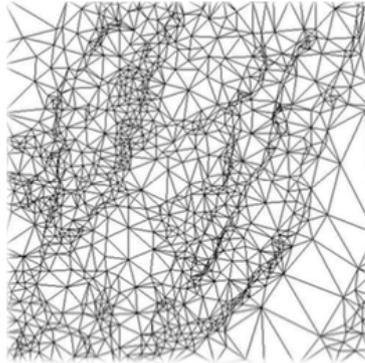
Modelo de superfície gerada por uma grade retangular (Fonte: INPE)



Representação TIN: triangulated irregular network

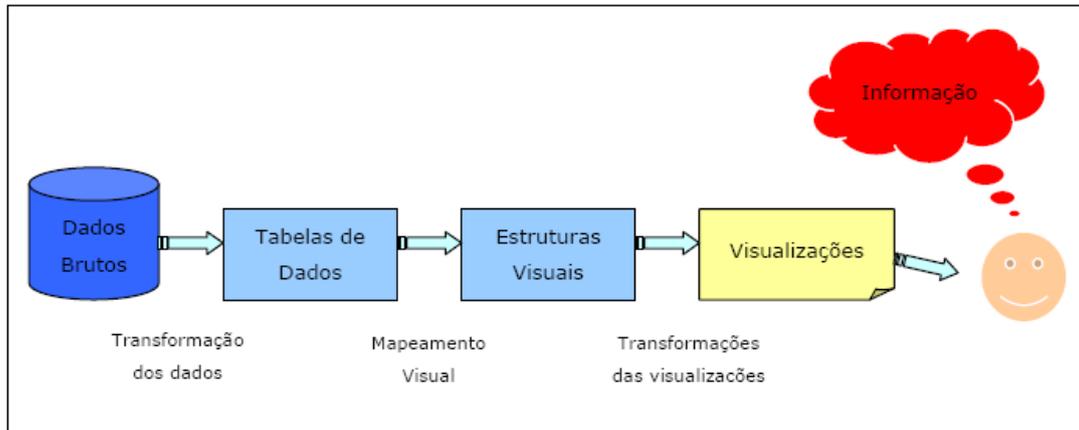


Modelo de superfície gerada por grade triangular
(Fonte: INPE)



Exemplo de TIN – Pontos (dados) dispersos irregularmente

Paper: Uma proposta de Técnica de Visualização de Dados para Auxiliar a Priorização de Requisitos em Projetos Ágeis Fábio Abrantes Diniz
<http://reuse.cos.ufrj.br/wbvs2012/papers/wbvs04.pdf> d.Y2I



Algoritmos Treemaps

Os Treemaps representam uma estrutura de dados hierárquica por divisão recursiva de retângulos. Essa recursão determina o layout, calculando a área do retângulo e preenche os retângulos em suas respectivas localidades. Apesar da parte variável, os algoritmos de criação do Treemap segue como padrão o seguinte código:

```

desenhaNos(no[] Nos, regiaoTotal)
{
  Para cada no em Nos faça
    regiaoNo = calculaRegiao(Parametros);
    regiaoTotal = regiaoTotal - regiaoNo;
    Se (no tem filho)
      desenhaNos(noFilho[], regiaoNo);
    Fim Se;
  Fim Loop;
}
  
```

A Tabela a seguir compara os algoritmos de divisão comparados através das métricas analisadas em cada descrição dos algoritmos Aspect ratio dos retângulos;

Estabilidade do mapa, em termos de tamanho e posição dos retângulos;

Ordenação dos elementos;

Facilidade de leitura em termos de estruturas.

	Ordenação	Estabilidade	Aspect ratio	Facilidade de leitura
Slice and Dice	+	+	-	-
Squarified	-	-	+	+
Strip	+/-	+/-	+/-	+/-
Pivot-by-	+	+	+	+

size				
Pivot-by-middle	+	+	+	+

tabela de comparação de algoritmos treemaps

Paper: Treemaps Aplicados ao Gerenciamento de Ativos em Rede

<http://livros01.livrosgratis.com.br/ea000349.pdf>

Algoritmo genérico:

Este algoritmo recebe como parâmetro um conjunto de nós em um vetor e uma variável que fornece os limites destinados à construção do treemap. Os nós são processados individualmente a cada loop do algoritmo. Se um nó específico possuir nós filhos, o algoritmo é chamado recursivamente para os filhos de cada nó. A quarta linha é a parte variável do algoritmo. Nessa linha, é feita a chamada da função que irá calcular os contornos do nó do loop corrente com base nos parâmetros

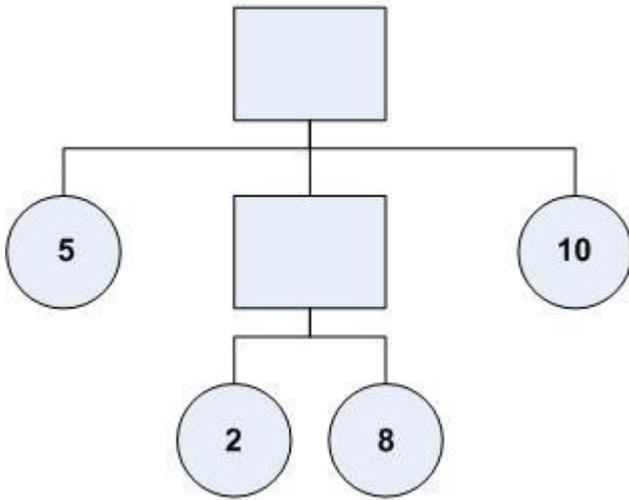


Figura 5: Exemplo de árvore a ser transformada em treemap. Após o cálculo, esses contornos são atribuídos ao nó em questão e a área total destinada ao treemap é recalculada.

A árvore apresentada na figura 5 servirá como exemplo de árvore a ser transformada em treemap pelos algoritmos apresentados neste capítulo.

Algoritmo Slice and Dice

```
SliceAndDice(raiz, filho[], lowerLeft[], upperRight[], eixo) largura =  
upperRight[eixo]-lowerLeft[eixo];
```

```
Para i = 1 até tamanho(filhos) faça
```

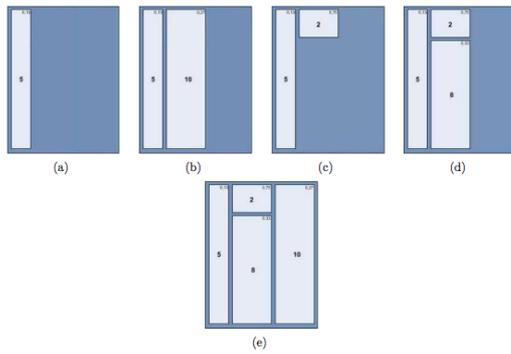
```
    upperRight[eixo] = lowerLeft[eixo] +  
    (tamanho(filho[i])/tamanho(raiz))*largura;
```

```
    SliceAndDice(i, filho[i], lowerLeft, upperRight, 1-eixo);
```

```
    lowerLeft[eixo] = upperRight[eixo];
```

```
    Fim Para;
```

```
Fim SliceAndDice
```

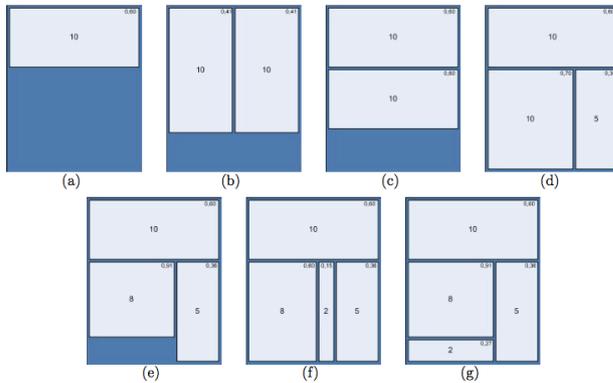


Algoritmo Squarified

```

Squarify(nodes[]) {
  Ordena(nodes[]) Enquanto Tamanho(nodes[]) > 0 faça
  no = Desempilha(nodes[]) Se worst(Linha[] + no) < worst(Linha[]) entao
    Linha[] = Linha[] + no
  Senao {}

```



Algoritmo Split:

```

Split(List itens, Retangulo r) 2{
  If (itens.length == 0)
  return;
  If (itens.length == 1) {
    itens.bounds = r;

```

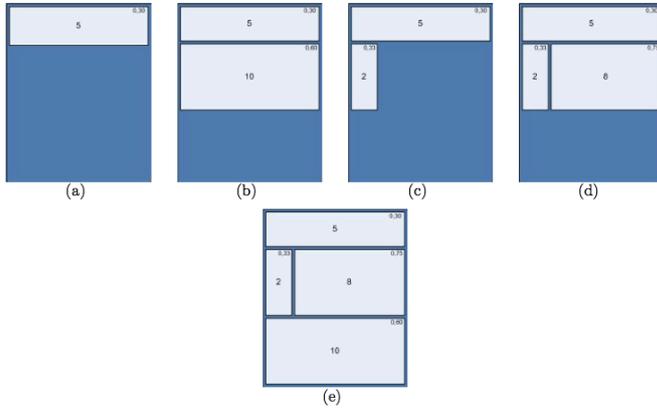
```

    Split(itens.children, r);
  }
  List l1, l2;
  Retangulo r1, r2;
  Double halfSize = w(itens)/2;
  doublew1=0,tmp=0;
  For All 16 17 18
  19 20 21 22 23
  itens
  {
    Item front = itens[front];
    tmp = w1 + front.size;
    If (abs(halfSize-tmp) > abs(halfSize-w1))
      break;
    enqueue(itens.dequeue);
    w1 = tmp; }
  End For
  l2 = itens;
  r1 = new Retangulo(r.x, r.y, r.width*w(l1)/w(l1+l2), r.height);
  If (r1.height) > (r1.width)
  r1.width = r1.height
  r2 = new Retangulo(r.x, r.y + r1.width, r.heigth, r.width - r1.width);

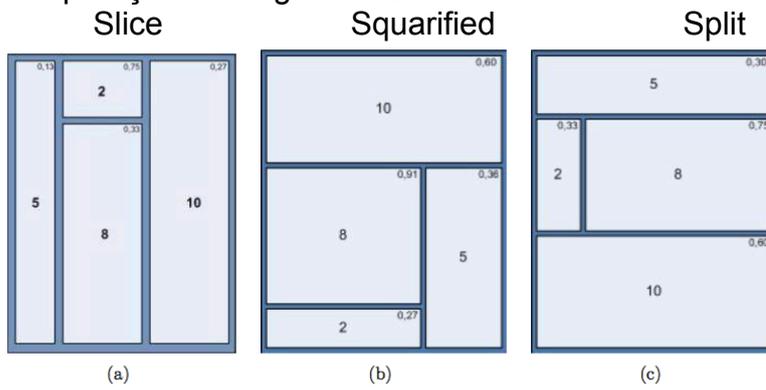
```

Else

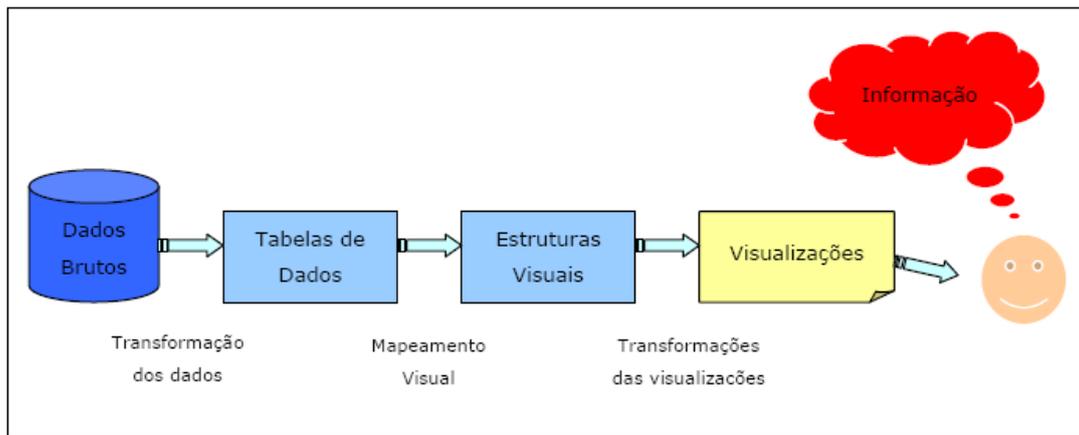
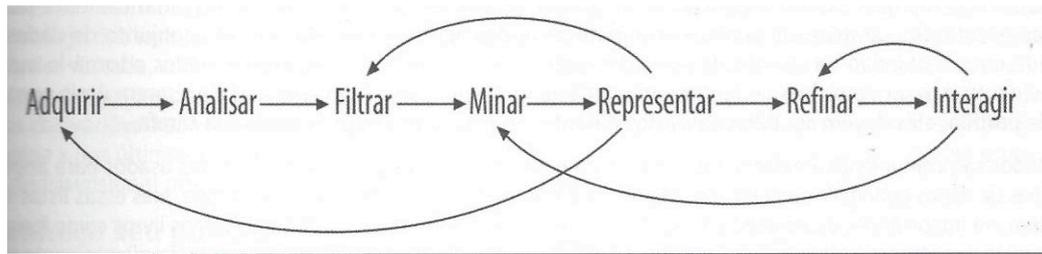
```
r2 = new Retangulo(r.x + r1.width, r.y, r.width - r1.width,  
r.height);  
Split(l1,r1);  
Split(l2,r2);  
}
```



Comparação dos algoritimos



4. Processos para produzir visualizações

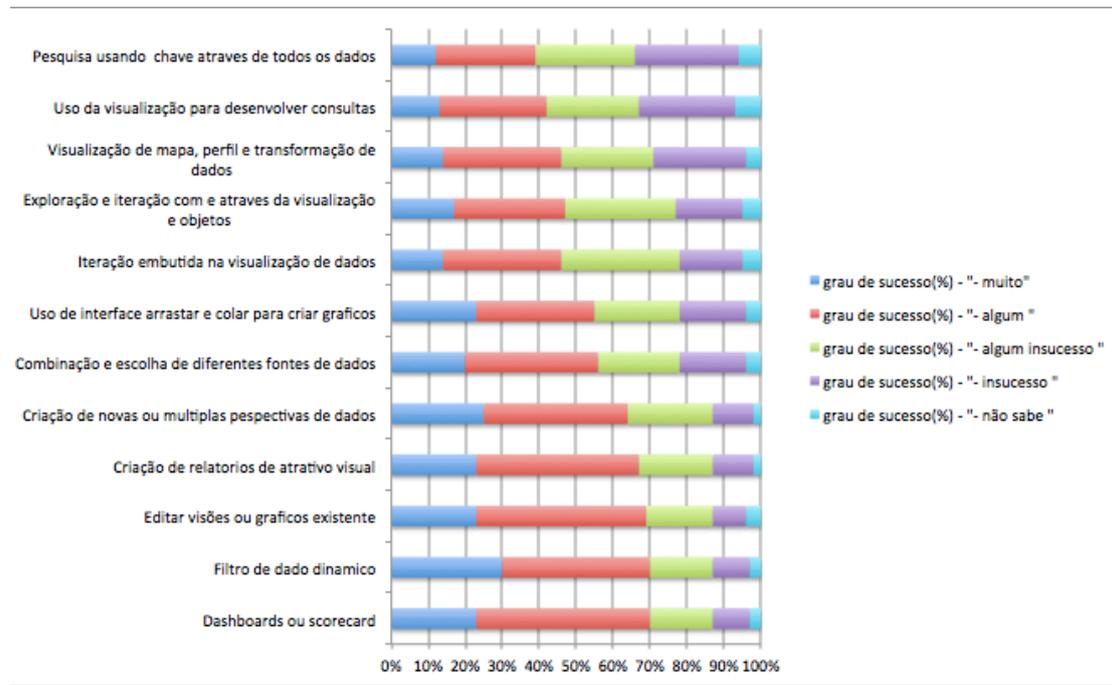


5. Apresentação da visualização escolhida

Ver em <http://wdz.eng.br/VisualNuvemPalavras.xml.html>

6. Visualizações

6.1 Gráfica - comparativa



6.2 Estrutura textual – Nuvens de palavras

desse (1) deste (2) determinada (1) determinado (1) **uiv** (9) document (1) Dominguez (1) duas (1) editor (2) **em** (6) end (1) English (1) escuro (1) **está** (2) Este (1) estrutura (1) execute (1) existente (2) **raise** (3) fazer (1) **feita** (2) Fernada (1) ferramenta (2) file (1) filme (1) final (1) **Firefox** (2) fin (1) **foi** (6) font-size (1) **font-weight** (10) **fonte** (8) fora (1) **foram** (3) format (1) formato (1) **freq** (5) **frequencia** (5) frequency (2) generated (1) gerado (1) gerador (2) grafica (2) Group (1) grupo (1) **height** (6) hint (1) **href** (23) **html** (14) htmltagcloud (2) **http** (13) icone (1) **id** (7) iframe (2) **img** (2) inclua (1) index (1) **indice** (3) **indicetopico** (3) **IndicetopicoX** (3) **IndexX** (3) **informacao** (2) information (1) Ingles (1) Iniciador (1) **Introducao** (2) item (1) **Iteracao** (2) ja (2) javascript (1) Jornalista (1) Junta (1) keep (1) **kompozer** (3) Language (1) **lead** (2) **letra** (3) **li** (63) **linguagem** (4) lowercase (1) **maior** (2) Martin (1) maximo (1) Maximum (1) Medeiros (1) melhoras (1) **metodo** (2) **minima** (2) Minimum (1) **minuscula** (2) **moldura** (3) mostrar (1) Mostra (1) mostradas (1) mostrar (1) mozilla (1) **na** (2) **nao** (24) **nbsp** (24) net (1) **nome** (3) nominals (1) nominal (1) **none** (2) notice (1) novo (1) number (1) **numero** (5) **nuvem** (7) objetivo (2) **obtencao** (3) obtida (1) **ocorrencias** (3) **ol** (4) Options (1) **org** (7) origem (2) **original** (3) **ou** (3) outros (1) **palavras** (26) palavras (1) **Paper** (1) **Paper** (2) **para** (5) parametrizacao (1) **partir** (2) passo (1) **Paste** (1) pdf (1) **peLa** (2) Pergunta (1) pesquisadas (1) php (1) please (1) **png** (2) **pode** (4) Portugues (1) Portuguese (1) **possam** (1) posteriormente (1) Processo (1) produto (1) proporcionalmente (1) proposta (1) propria (1) pt (1) puro (1) **px** (13) qual (1) quantitativa (1) **que** (3) **Quero** (5) Referencias (2) regido (1) relevancia (1) Requisitos (1) Responder (1) Resposta (1) rgb (1) **rodape** (3) **RodapeX** (3) Sa (1) Safari (1) salvo (1) scripts (1) seguinte (1) seguir (1) **selecao** (2) **selecionadas** (2) **ser** (3) **seu** (3) shown (1) sim (1) **similares** (3) simples (1) **sinTaxe** (2) slides (1) Software (1) solucao (1) **SOMENTE** (2) **span** (45) **sFC** (2) **style** (27) surgir (1) **svg** (1) **table** (6) **tag** (2) **tagcloud** (9) **tagcrowd** (12) **tamanho** (4) **target** (18) **tbody** (6) **td** (6) Tecnologia (1) tem (1) tenha (1) **ter** (1) **text-decoration** (4) **text** (2) **texto** (8) **textual** (2) **titulo** (23) **tonalidade** (3) **top** (3) **topico** (6) **tr** (6) trabalho (1) tree (1) **true** (3) (15) **ul** (6) **um** (6) **uma** (5) **underline** (3) **URL** (1) usando (3) utilizada (2) utilizado (2) vai (1) variando (1) **variar** (2) Variaveis (2) **vertical-align** (3) **vezes** (2) **Viegas** (1) **visual** (3) **visualizacao** (11) visualizadas (1) visualizador (1) **visualnuvem** **palavras** (9) **Walter** (2) **Watt** (1) **webservice** (1) **website** (1) **width** (7) wordle (1) **words** (3) **Wrd** (5) **WWW** (2) **xml** (9) z-index (1)

Save as... resize browser window to arrange text

Choose your text source:

Paste Text | Web Page URL | Upload File

Paste text to be visualized:
plain text: 500 kilobyte max

This XML file does not appear to have any style information associated with it. The document tree is shown below.

```
<linguagem>
<topico>
<nome>Visualização : Nuvem de Palavras de um
tópico</nome>
</topico>
</linguagem>
```

Visualize!

Options:

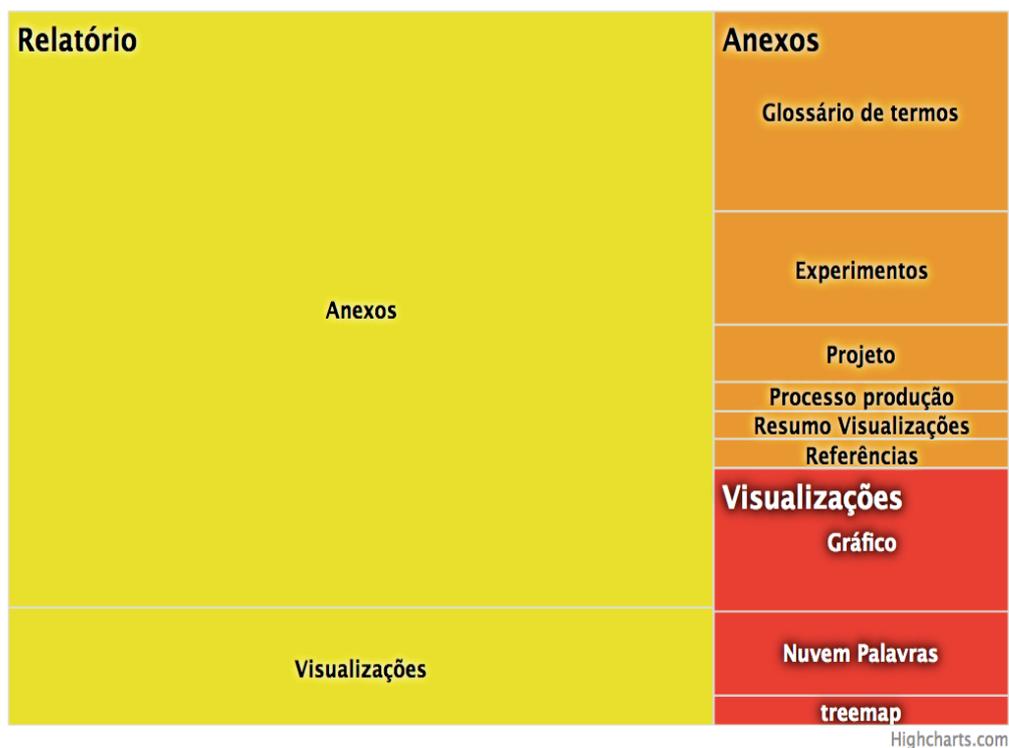
Language of text:

Ignore common words in this language:

Maximum number of words to:

6.3 Estrutura hierárquica – treemap

Visualização da informação



7. Referências

Kirk, Andy. Data Visualization: a successful design process

Fry, Ben. Visualizando dados

Stodder, David. TDWI Best practices report – Visual analytics for making smarter decisions faster – Applying self-service BI technologies to data-driven objectives

<http://tdwi.org/research/2015/07/best-practices-report-visual-analytics-for-smarter-decisions.aspx?tc=page0>

Viegas, Fernanda B. ,... Participatory visualization with wordle

http://hint.fm/papers/wordle_final2.pdf

http://www.visual-literacy.org/periodic_table/periodic_table.html

Marino, Tiago Badre. Representações de dados espaciais – Raster x Vetor x TIN

[http://www.ufrj.br/lga/tiagomarino/aulas/5 - Representacao de Dados Espaciais - Raster x Vetor x TIN.pdf](http://www.ufrj.br/lga/tiagomarino/aulas/5-Representacao%20de%20Dados%20Espaciais-Raster%20x%20Vetor%20x%20TIN.pdf)

<http://sa2015.siggraph.org/en/submitters/workshops.html>

http://www.ic.ufal.br/evento/cbsoft2014/anais/vem_v1_p.pdf

<http://www.highcharts.com/download>

<http://icp.ge.ch/sem/cms-spip/spip.php?article1832>

<http://wdz.eng.br/ApresFinal.ppt>

Siegel, David. Pull O futuro da internet e o impacto da web semântica em seus negócios. 2011 Elsevier Editora Ltda

Software utilizado

MacOs

MSWord

MsExcel

MsPowerpoint

Webservice <http://tagcrowd.com>

Atualizado em 20/10/2016