



# Oficina sobre Dados Abertos

Vamos criar uma página com dados abertos



# Sobre

William

- Software Engineer @ Red Hat
- Dados Abertos @ SJC Digital
- Software Livre
- Java



# Objetivo

- 1) Falar sobre dados abertos e o que podemos fazer
- 2) Explorar dados abertos da prefeitura de Jacareí
- 3) Criar uma página WEB com melhores visualizações dos dados em 7 passos

Resultado final: [https://jesuino.github.io/despesas\\_prefeitura\\_jacarei](https://jesuino.github.io/despesas_prefeitura_jacarei)

*NOTA: Na prática o ritmo irá aumentando, no começo muitas explicações, no final mais código e menos explicações!*



# Dados Abertos

*“dados são abertos quando qualquer pessoa pode livremente acessá-los, utilizá-los, modificá-los e compartilhá-los para qualquer finalidade, estando sujeito a, no máximo, a exigências que visem preservar sua proveniência e sua abertura” - Open Knowledge Internacional*

Em 2011 a Lei de **Acesso de acesso à Informação (LAI)** foi criada: estado e órgãos (prestadores de serviço relacionados ao estado devem abrir seus dados. Qualquer dado com um uma única exceção: dados que comprometem a segurança.

Dados estão lá. A função de SJC Digital e muitas ONGs e movimentos de dados abertos é retirar informação úteis aos cidadãos usando esses dados.

**Palestra hoje 14:00**



# Como buscar dados abertos

Dados abertos estão:

- Portais do governo: Governos federais, municipais e estaduais no geral têm uma página;
- Órgãos e ministérios. Grande exemplo: DATASUS;
- Garimpar dados em diversas fontes;
- Se o órgão tem o dado mas não libera podemos solicitar através da **CGU (Controladoria Geral da União)**



# Prática: Despesas da Prefeitura de Jacareí

Vamos procurar saber de forma mais clara quais são as despesas da prefeitura de Jacareí.

Há uma pergunta que os portais de transparência geralmente não respondem: Como evoluíram os gastos e as receitas no decorrer dos anos? Qual a maior despesa do ano X?

Para facilitar a resolução dessas questões vamos gerar um site com:

- Gráficos
- Resumos
- Possíveis outras formas de visualização
- Possibilidade de cruzar dados



## **Passo 1: Buscando dados**

# Onde encontramos esses dados?

- Acessar portal da transparência da prefeitura <http://www.jacarei.sp.gov.br/>





# Despesas



HOME PREFEITURA CIDADE IMPRENSA NOTÍCIAS VÍDEOS ✓ TRANSPARÊNCIA

Fale conosco  
0800 163010

Buscar no site...



CIDADÃO



EMPRESA



LAZER



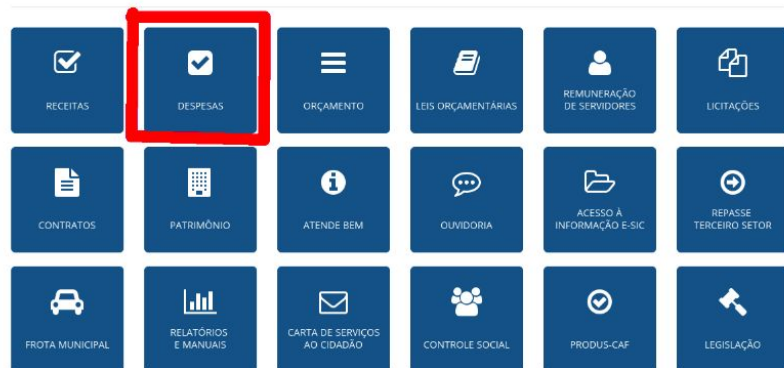
SERVIDORES

INÍCIO / TRANSPARÊNCIA PREFEITURA

## Transparência Prefeitura

O Portal da Transparência foi desenvolvido pela Prefeitura de Jacaréi em observância à Lei de Acesso à Informação – Lei 12.527/2011 –, com o fim de garantir ao município o acesso a informações claras e transparentes relativas ao exercício da gestão pública.

Por meio desse canal democrático e de diálogo com a sociedade, a Prefeitura de Jacaréi oferece a você, cidadão, em uma linguagem clara e de fácil compreensão, a oportunidade de exercer o controle social, conhecendo e fiscalizando as receitas arrecadadas pelo Poder Público e como o dinheiro público é aplicado aqui em nossa cidade, dentre outras informações de interesse geral da sociedade.



PREFEITURA MUNICIPAL DE JACARÉI

[Página Inicial](#) [Execução Orçamentária](#) [Licitações](#) [RH](#) [Acesso à Informação](#) [Atendimentos](#) [Publicações](#) [Contato](#) [Sobre](#)

### Execução Orçamentária

Detalhamento das Despesas \*

\* Consulta para dados a partir do ano de 2018.

Ano:

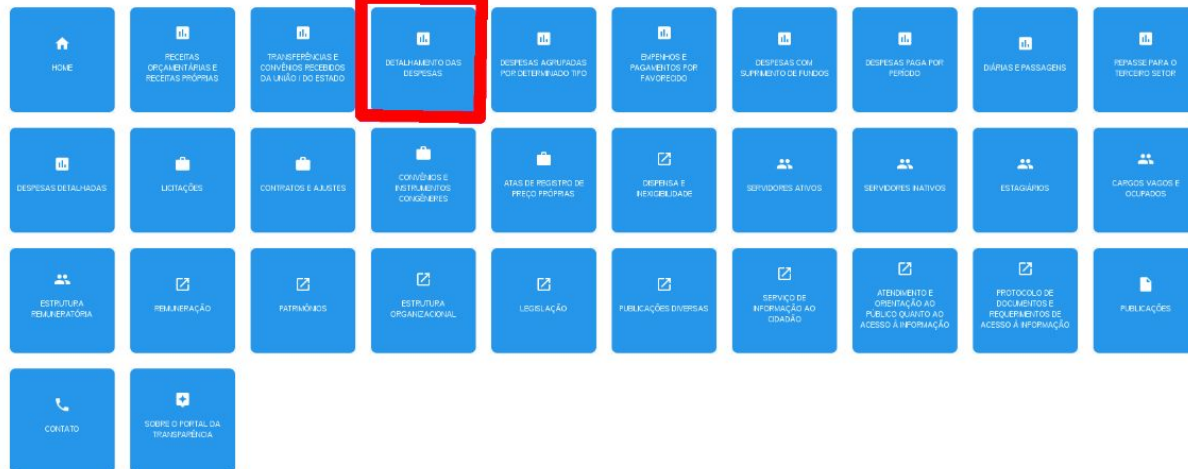
Entidade:

Pesquisar



WWW.JACAREI.COM.BR  
Informatica@Jacarei.com.br  
Praça Dos Três Poderes  
JACAREI, SP  
(11) 4663-1000

[Página Inicial](#) [Execução Orçamentária](#) [Licitações](#) [RH](#) [Acesso à Informação](#) [Atendimentos](#) [Publicações](#) [Contato](#) [Sobre](#)



# Buscando dados da prefeitura



PREFEITURA MUNICIPAL DE JACARÉ

[WWW.JACAREI.SP.GOV.BR](http://WWW.JACAREI.SP.GOV.BR)

[Informatica@Jacarei.Sp.Gov.Br](mailto:Informatica@Jacarei.Sp.Gov.Br)

Praca Dos Tres Poderes , 73 , Centro

JACAREI , SP, 12327-903

(012) 3955-9000

[Página Inicial](#) [Execução Orçamentária ▾](#) [Licitações ▾](#) [RH ▾](#) [Acesso à Informação ▾](#) [Atendimentos ▾](#) [Publicações](#) [Contato](#) [Sobre](#)

## Execução Orçamentária

Detalhamento das Despesas \*

\* Consulta para dados a partir do ano de 2018.

Ano:	<input type="text" value="2018"/>
Entidade:	<input type="text" value="PREFEITURA MUNICIPAL DE JACAREI"/>
<input type="button" value="Pesquisar"/>	

# Baixando os dados



PREFEITURA MUNICIPAL DE JACARÉ

WWW.JACARELSP.GOV.BR  
Informática@Jacarei.Sp.Gov.Br  
Praça Dos Tres Poderes, 73, Centro  
JACARÉ, SP, 12327-903  
(012) 3955-9000

[Página Inicial](#) [Execução Orçamentária](#) [Licitações](#) [RH](#) [Acesso à Informação](#) [Atendimentos](#) [Publicações](#) [Contato](#) [Sobre](#)

## Execução Orçamentária

Detalhamento das Despesas \*

\* Consulta para dados a partir do ano de 2018.

Ano: 2018

Entidade: PREFEITURA MUNICIPAL DE JACARÉ

Pesquisar

Exportar

PDF

CSV

Buscar...

PDF CSV	Elemento	Valor Previsto	Tipo	Total	Valores Executados									
					Janeiro	Fevereiro	Março	Abril	Maio	Junho	Julho	Agosto	Setembro	Outubro
Natureza: 3.1 - Pessoal e Encargos Sociais														
01 - Aposentadorias, Reserva Remunerada e Reformas	1.700.000,00	E	1.454.639,52	152.988,99	27.864,99	240.208,32	146.007,85	117.136,22	228.804,82	114.613,43	121.635,95	117.747,43	119.297,88	
		L	1.940.026,09	152.988,99	27.864,99	240.208,32	146.007,85	117.136,22	228.804,82	114.613,43	121.635,95	117.747,43	119.297,88	
		F	1.940.026,09	152.988,99	27.864,99	240.208,32	146.007,85	117.136,22	228.804,82	114.613,43	121.635,95	117.747,43	119.297,88	
03 - Pensões do RPPS e do Militar	1.319.000,00	E	1.133.921,26	60.698,03	9.032,48	149.760,50	66.782,26	304.097,49	111.687,63	108.339,36	125.364,80	114.990,26	116.996,35	

[illegible]



# Conclusões sobre a análise dos dados

- A coluna Natureza contém um agrupamento de diversos Elemento da despesa. Podemos descartar
- Para cada mês temos uma coluna do valor empenhado, liquidado e pago, onde:
  - **Empenhado** Valor reservado
  - **Liquidado:** Qual foi o valor do serviço
  - **Pago** Valor realmente pago para o fornecedor

Nosso interesse será no valor **pago**.

- Embora haja informações que queremos ali, precisamos analisar, mexer com planilha e não podemos fazer tratamento posterior, como cruzar esses dados com outros. O nosso objetivo é extrair a informação que queremos e deixar elas mais acessíveis para os cidadãos possam tirar suas conclusões.

O que vamos fazer hoje é básico, mas após aprender a fazer isso o céu é o limite.



# E agora, o que fazemos com os dados

Precisamos tratar os dados para usar. **Raramente** você vai conseguir dados bem formatados para trabalhar.

Para tratar os dados podemos, usar editor de planilhas, scripts Shell, Python ou outras linguagens de programação ou projetos com o Apache Tika ou Google Refine.

No nosso caso usaremos HTML e Javascript com as bibliotecas D3 e C3. Baixe em:

D3: <https://github.com/d3/d3/releases/download/v5.7.0/d3.zip>

C3: <https://github.com/c3js/c3/archive/v0.6.11.zip>



## **Passo 2: Configurando página para desenvolvimento**





# Configuração inicial

- Criar diretório *despesas\_jacarei*
- Dentro desse diretório criar diretórios *js*, *css*, *data*.
- Dentro de *js* criar *app.js* e na raiz *index.html*
- Copiar *d3.min.js* e *c3.min.js* dentro de *js* e *c3.min.css* dentro de *css*. Assim deve ser a sua estrutura de diretório:

```
[wsiqueir@wsiqueir despesas_jacarei]$ tree -C
.
|-- css
|   |-- c3.min.css
|-- data
|-- index.html
|-- js
|   |-- app.js
|   |-- c3.min.js
|   |-- d3.min.js
3 directories, 5 files
```



# Configuração inicial

Modifique o arquivo [index.html](#) e inclua o conteúdo básico de uma página HTML importando os arquivos Javascript e HTML que mencionamos antes.

```
1 <meta charset="UTF-8">
2 <html>
3   <head>
4     <title>Despesas da prefeitura de Jacareí por ano</title>
5     <link rel="stylesheet" href="css/c3.css">
6   </head>
7   <body>
8     <script src="js/d3.min.js" type="text/javascript"></script>
9     <script src="js/c3.min.js" type="text/javascript"></script>
10    <script src="js/app.js" type="text/javascript"></script>
11  </body>
12 </html>
13
```



# Configuração Inicial

Para testarmos se tudo foi configurado corretamente vamos colocar o seguinte conteúdo em app.js:

```
1 console.log(c3);  
2 console.log(d3);
```

*(pode tirar o código após o teste)*



# Testando configuração inicial

Para testar a configuração inicial iremos subir um webserver python. No diretório **despesas\_jacarei** abra uma linha de comando e rode:

```
python3 -m http.server --cgi
```

Com um navegador acesse localhost:8000 e abra o console (Ctrl + Shift + i no Chrome) e no console, se tudo estiver bem, você algo como abaixo.

► Object

app.js:1

► Object

app.js:2

Se estiver vendo algo como ***“d3 is not defined”*** então revise a instalação!!



## **Passo 3: Carregando os dados**



## Organizando e carregando os dados

Pegue o CSV que você baixou e renomeie ele para [despesas\\_prefeitura\\_2018.csv](#) e coloque dentro do diretório [data](#). Para carregar os dados vamos usar d3.csv. O seguinte código irá colocar os dados no console para visualizarmos. Coloque ele em [app.js](#):

```
1 d3.csv("data/despesas_prefeitura_2018.csv")
2   .then(function (dados) {
3     console.log(dados);
4   });
```

A função CSV responde com um Promise, quando d3 termina de carregar os dados que *then* é chamado.

Abra o console e vamos ver o que foi carregado.

```
(38) [Object, Object, Object, Object, Object, Object, Object, Object, Object, Object, Object, Object, Object, Object, Object,  
      Object, Object, Object, Object, Object, columns: Array(39)]  
▼ 0: Object  
    Elemento: "01 - Aposentadorias, Reserva Remunerada e Reformas"  
    Natureza: "3.1 - Pessoal e Encargos Sociais"  
    VL_Abr_Emp: * 146.007,85"  
    VL_Abr_Liq: * 146.007,85"  
    VL_Abr_Pag: * 146.007,85"  
    VL_Ago_Emp: * 121.635,95"  
    VL_Ago_Liq: * 121.635,95"  
    VL_Ago_Pag: * 121.635,95"  
    VL_Dez_Emp: * 0,00"  
    VL_Dez_Liq: * 0,00"  
    VL_Dez_Pag: * 0,00"  
    VL_Fev_Emp: * 27.864,99"  
    VL_Fev_Liq: * 27.864,99"  
    VL_Fev_Pag: * 27.864,99"  
    VL_Jan_Emp: * 152.988,89"  
    VL_Jan_Liq: * 152.988,89"  
    VL_Jan_Pag: * 152.988,89"  
    VL_Jul_Emp: * 114.613,43"  
    VL_Jul_Liq: * 114.613,43"  
    VL_Jul_Pag: * 114.613,43"  
    VL-Jun_Emp: * 126.804,82"  
    VL-Jun-Liq: * 126.804,82"  
    VL-Jun-Pag: * 126.804,82"  
    VL_Mai_Emp: * 117.136,22"  
    VL_Mai_Liq: * 117.136,22"  
    VL_Mai_Pag: * 117.136,22"  
    VL-Mar_Emp: * 240.208,32"  
    VL-Mar-Liq: * 240.208,32"  
    VL-Mar-Pag: * 240.208,32"  
    VL-Nov_Emp: * 170.333,74"  
    VL-Nov-Liq: * 170.333,74"  
    VL-Nov-Pag: * 170.333,74"  
    VL-Out_Emp: * 119.297,88"  
    VL-Out-Liq: * 119.297,88"  
    VL-Out-Pag: * 119.297,88"  
    VL-Previsto: * 1.700.000,00"  
    VL_Set_Emp: * 117.747,43"
```



# Notas sobre dados carregados

- Há um problema com os valores numéricos especificamente, precisamos tratá-los:
  - Há aspas. Com aspas não conseguimos trabalhar com eles;
  - Usam o sistema brasileiro de formatação (vírgula para separação de casas decimais)
  - Há um espaço adicional! Também precisamos eliminar.
- Muitas colunas não serão usadas, precisamos selecionar só o que queremos.

```
Vl_Abr_Emp: " 146.007,85"  
Vl_Abr_Liq: " 146.007,85"  
Vl_Abr_Pag: " 146.007,85"  
Vl_Ago_Emp: " 121.635,95"  
Vl_Ago_Liq: " 121.635,95"  
Vl_Ago_Pag: " 121.635,95"  
Vl_Dez_Emp: " 0,00"  
Vl_Dez_Liq: " 0,00"  
Vl_Dez_Pag: " 0,00"  
Vl_Fev_Emp: " 27.864,99"  
Vl_Fev_Liq: " 27.864,99"  
Vl_Fev_Pag: " 27.864,99"  
Vl_Jan_Emp: " 152.988,89"  
Vl_Jan_Liq: " 152.988,89"  
Vl_Jan_Pag: " 152.988,89"  
Vl_Jul_Emp: " 114.613,43"  
Vl_Jul_Liq: " 114.613,43"  
Vl_Jul_Pag: " 114.613,43"  
Vl_Jun_Emp: " 126.804,82"  
Vl_Jun_Liq: " 126.804,82"  
Vl_Jun_Pag: " 126.804,82"  
Vl_Mai_Emp: " 117.136,22"  
Vl_Mai_Liq: " 117.136,22"  
Vl_Mai_Pag: " 117.136,22"  
Vl_Mar_Emp: " 240.208,32"  
Vl_Mar_Liq: " 240.208,32"  
Vl_Mar_Pag: " 240.208,32"  
Vl_Nov_Emp: " 170.333,74"
```





## **Passo 4: Tratar dados com problemas**



## Uma função simples para limpar os dados

Primeiro vamos criar uma função simples que irá fazer a limpeza dos dados. No final do arquivo adicione esse código:

```
7 function trataNumero(num) {  
8     var numLimpo = num.replace('\\', '')  
9     .replace(' ', '')  
10    .replace('.', '')  
11    .replace(',', '.');  
12    return parseFloat(numLimpo);  
13 }
```

O objetivo é fazer o replace de cada possível valor problemático. Pode ser melhorado com expressões regulares, mas manteremos simples por hoje:



## Escolhendo quais colunas queremos

É possível adicionar uma função intermediária na carga dos dados. Antes precisamos saber quais campos vamos usar e são os valores pagos para todos os meses. Assim crie um vetor no início de [app.js](#) com todos os campos numéricos que queremos. Para iniciar você pode colocar um vetor menor para não ter que digitar tudo:

```
1 var colunasNumericas = [  
2     'Vl_Jan_Pag', 'Vl_Fev_Pag', 'Vl_Mar_Pag',  
3     'Vl_Abr_Pag', 'Vl_Mai_Pag', 'Vl_Jun_Pag',  
4     'Vl_Jul_Pag', 'Vl_Ago_Pag', 'Vl_Set_Pag',  
5     'Vl_Out_Pag', 'Vl_Nov_Pag', 'Vl_Dez_Pag'  
6 ];
```



## Transferindo valores do CSV para um objeto

A função `d3.csv` também aceita uma função de mapeamento, intermediária, onde podemos pegar dados de cada linha do CSV para construirmos nosso objeto. Vamos pegar todos os valores numéricos que mapeamos anteriormente, mais alguns campos importantes e passar para nosso objeto:

```
8 d3.csv("data/despesas_prefeitura_2018.csv",
9   function(linha) {
10     var dado = {};
11     dado.elemento = linha.Elemento;
12     for(var i = 0; i < colunasNumericas.length; i++) {
13       var coluna = colunasNumericas[i];
14       var valor = linha[coluna];
15       dado[coluna] = trataNumero(valor);
16     }
17     return dado;
18   }).then(function (dados) {
19     console.log(dados);
20   });
21 });
```

## Verificando os resultados

```
▼ (38) [Object, Object, Object, Object, Object, Object, Object, Object, Object, Object, Object, Object, columns: Array(39)] ⓘ  
  ▼ 0: Object  
    Vl_Abr_Pag: 146007.85  
    Vl_Ago_Pag: 121635.95  
    Vl_Dez_Pag: 0  
    Vl_Fev_Pag: 27864.99  
    Vl_Jan_Pag: 152988.89  
    Vl_Jul_Pag: 114613.43  
    Vl_Jun_Pag: 126804.82  
    Vl_Mai_Pag: 117136.22  
    Vl_Mar_Pag: 240208.32  
    Vl_Nov_Pag: 170333.74  
    Vl_Out_Pag: 119297.88  
    Vl_Set_Pag: 117747.43  
    elemento: "01 - Aposentadorias, Reserva Remunerada e Reformas"
```



## **Passo 5: Extraíndo somas de despesas pagas**



## Extraindo a soma de todos meses

Com os dados tratados e carregados podemos criar coisas mais interessantes, mas faremos algo mais simples inicialmente. Vamos começar passando nos dados para descobrir o maior pagamento e o menor. Passemos nos dados numéricos para cada mês e vamos somar, mas em uma nova função:

```
24 function calculaTotais(dados) {  
25     var pagamentosTotaisPorAno = [];  
26     for(var i = 0; i < dados.length; i++) {  
27         var total = 0;  
28         for(var j = 0; j < colunasNumericas.length; j++) {  
29             var coluna = colunasNumericas[j];  
30             total += dados[i][coluna];  
31         }  
32         pagamentosTotaisPorAno.push({ "elemento": dados[i].elemento, "total": total });  
33     }  
34     return pagamentosTotaisPorAno;  
35 }
```



# Buscando o maior valor

Criar uma função para buscar os maiores valores

```
28 function buscaMaximo(totais) {  
29     var maior = totais[0];  
30     for(var i = 1; i < totais.length; i++) {  
31         if(totais[i].total > maior.total) {  
32             maior = totais[i];  
33         }  
34     }  
35     return maior;  
36 }
```



# Mostrando o maior valor na página HTML

```
1 <meta charset="UTF-8">
2 <html>
3   <head>
4     <title>Despesas da prefeitura de Jacareí por ano</title>
5     <link rel="stylesheet" href="css/c3.css">
6   </head>
7   <body>
8     <h2>Despesas da prefeitura de Jacareí</h2>
9
10    <h4> Maior despesa foi para a área de "<em id="maxArea"></em>" com total de R$ <span id="maxValor"></span></h4>
11
12    <div id="grafico"></div>
13    <script src="js/d3.min.js" type="text/javascript"></script>
14    <script src="js/c3.min.js" type="text/javascript"></script>
15    <script src="js/app.js" type="text/javascript"></script>
16  </body>
17 </html>
```



## Usando D3 para colocar o valor na página

```
8 d3.csv("data/despesas_prefeitura_2018.csv",
9   function(linha) {
10     var dado = {};
11     dado.elemento = linha.Elemento;
12     for(var i = 0; i < colunasNumericas.length; i++) {
13       var coluna = colunasNumericas[i];
14       var valor = linha[coluna];
15       dado[coluna] = trataNumero(valor);
16     }
17     return dado;
18   }).then(function (dados) {
19     var totais = calculaTotais(dados);
20     var maximo = buscaMaximo(totais);
21     d3.select('#maxArea').html(maximo.elemento);
22     d3.select('#maxValor').style('color', 'red').html(maximo.total);
23   });
24
25 });
```



# Resultado

## Despesas da prefeitura de Jacareí

Maior despesa foi para a área de "*36 - Outros Serviços de Terceiros - Pessoa Física*" com total de R\$ **6901003.68**



## **Passo 6: Colocando despesas em gráficos**

## Totais em um gráfico de pizza

Vamos utilizar o c3 para criar um gráfico de pizza com os valores totais. Não se esqueça de uma div com ID graficoTotais na página HTML:

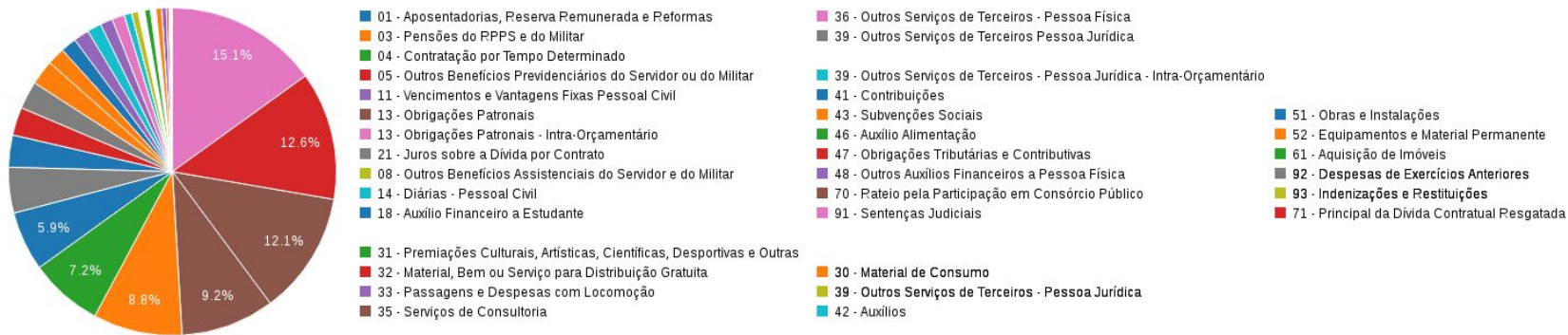
```
52 function graficoPizzaTotais (totais) {  
53     var dadosGrafico = [];  
54     for( var i = 0; i < totais.length; i++) {  
55         dadosGrafico.push([ totais[i].elemento, totais[i].total]);  
56     }  
57     c3.generate({  
58         bindto: '#graficoTotais',  
59         data: {  
60             columns: dadosGrafico,  
61             type: 'pie'  
62         },  
63         legend: {  
64             position: 'right'  
65         }  
66     });  
67 }
```

# Gráfico de totais

## Despesas da prefeitura de Jacareí

Maior despesa foi para a área de "36 - Outros Serviços de Terceiros - Pessoa Física" com total de R\$ 6901003.68

## Soma das despesas pagas no ano todo





## Crescimento dos gastos no decorrer do ano

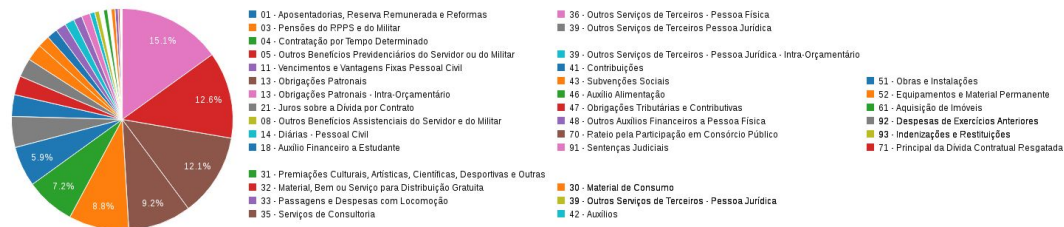
```
70 function graficoDadosMensais(dados) {
71     var dadosGrafico = [];
72     for(var i = 0; i < dados.length; i++) {
73         var linha = [];
74         linha.push(dados[i].elemento);
75         for(var j = 0; j < colunasNumericas.length; j++) {
76             var nomeColuna = colunasNumericas[j];
77             linha.push(dados[i][nomeColuna]);
78         }
79         dadosGrafico.push(linha);
80     }
81     c3.generate({
82         bindto: '#graficoMensais',
83         data: {
84             columns: dadosGrafico,
85             type: 'line',
86         }
87     });
88 }
```

# Resultado

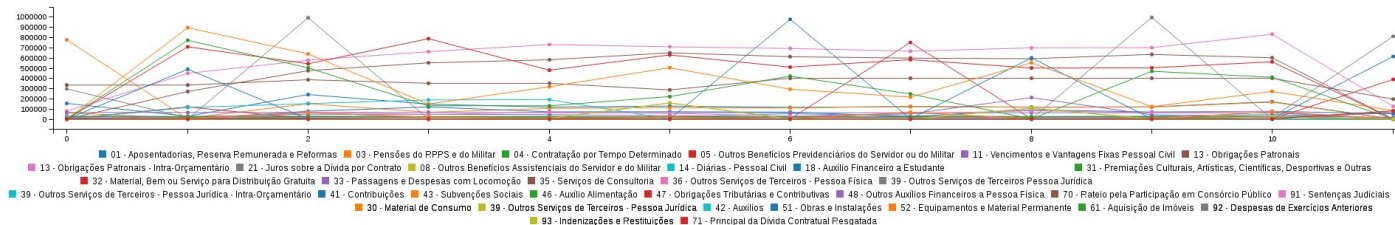
## Despesas da prefeitura de Jacareí

Maior despesa foi para a área de "36 - Outros Serviços de Terceiros - Pessoa Física" com total de R\$ 6901003.68

### Soma das despesas pagas no ano todo



### Evolução das despesas no decorrer do ano







## **Passo 7: Selecionar dados por ano**



## Passos para adicionar seleção de ano

- Baixar dados dos anos que se deseja adicionar
- Verificar se dados são compatíveis
- Trocar o CSV carregado de acordo com um combo box selecionado



## Conclusão e próximos passos



## Próximos passos

- Criar uma página bacana (melhorar estrutura e apresentação)
- Integrar com outros dados
- Possibilitar “drill down” para descobrir quem recebeu o serviço
- Comparar gastos de diferentes prefeituras
- Usar machine learning para prever gastos no mês
- Trabalhar também com receitas
- infinitas possibilidades!



**Obrigado!**