

# Gastos Abertos

## O que é o Gastos Abertos?

O projeto Gastos Abertos tem por objetivo facilitar a compreensão das pessoas a respeito dos gastos públicos no Brasil. O projeto trabalhará com três orçamentos, nessa ordem: 1. cidade de São Paulo, 2. estado de São Paulo e 3. governo federal.

Queremos fornecer ferramentas para que a sociedade civil organizada e os veículos de comunicação possam estimular os cidadãos a acompanhar e influenciar as tomadas de decisão sobre os gastos públicos.

O projeto é composto de duas frentes de trabalho: 1. uma plataforma tecnológica e 2. tutoriais, histórias do orçamento e cursos.

## Quais eram os planos?

Inicialmente discutimos na produção de um site onde existiriam ferramentas flexíveis para a realização de diversos tipos de análises sobre os dados do orçamento de São Paulo.

Estas ferramentas deveriam satisfazer uma série de [pré-requisitos básicos](#) que surgiram como interesse dos grupos que participaram do [Workshop no final do ano passado](#). Pré-requisitos como:

a Facilidade no compartilhamento das análises/visualizações. Por exemplo, pense no Google Maps, onde você pode navegar até um certo ponto, mudar uma série de parâmetros do que está sendo visualizado e facilmente compartilhar o estado do mapa em seu navegador, apenas compartilhando o URL que se encontra na barra de endereço.

b Simplicidade para integrar as visualizações em outros sites. Para isso seria importante pensar em uma solução tecnológica que não ficasse presa demais a arquitetura do site do Gastos Abertos, pensando em uma forma de desenvolver tais visualizações modularmente e backend independente.

Existe também uma série de outros pré-requisitos mais específicos a determinadas ferramentas de visualização, que serão levadas em consideração no desenvolvimento de cada ferramenta individual. Por exemplo, existiu uma forte demanda para possuir uma grande granularidade nas informações sobre Execução Orçamentária, ao invés de possuírem apenas totalizações. Alguns dos nossos dados permitem isso, outros será necessário todo um trabalho extra de cruzamento e análise nem sempre cabível de automatização.

Junto com tais ferramentas existirá no site artigos mais técnicos descrevendo o processo criativo que estamos realizando, artigos como tutoriais para jornalistas aprenderem a utilizarem ferramentas já existentes para a análise dos

dados e finalmente artigos explicando o uso das ferramentas que estão sendo desenvolvidas.

## O que foi alcançado?

Iniciamos o desenvolvimento de uma visualização bem simples em cima dos dados da Receita, como prova de conceito e para definirmos e experimentarmos diferentes stack tecnológicos que iremos utilizar no projeto.

Num primeiro momento procuramos utilizar o [OpenSpending](#) como backend dos nosso dados mas encontramos alguns problemas:

- Instabilidade no servidor oficial do Open Spending
- Impossibilidade de utilizá-lo para os dados hierárquicos que possuímos para os dados da Receita.
- Futuro ainda incerto da plataforma que está passando por uma [reimplementação e mudança completa de sua arquitetura](#).

Com isso optamos por desenvolvermos nós mesmo uma solução utilizando o micro framework [Flask](#) na linguagem Python.

## Dados da Receita

Para a visualização dos dados da Receita foram realizados uma série de trabalhos de desenvolvimento.

1. Criação de um script em Selenium para webscrapping dos dados da Receita, que não estavam disponíveis através de uma API ou em alguma base de dados.

[https://github.com/okfn-brasil/gastos\\_abertos\\_dados/blob/master/utils/revenue\\_downloader.py](https://github.com/okfn-brasil/gastos_abertos_dados/blob/master/utils/revenue_downloader.py)

2. Ferramenta para importação dos dados da Receita para um banco de dados SQL e extração das informações hierárquicas dos dados da Receita:

[https://github.com/okfn-brasil/gastos\\_abertos/blob/master/utils/import\\_revenue.py](https://github.com/okfn-brasil/gastos_abertos/blob/master/utils/import_revenue.py)

3. Processamento dos dados da Receita para gerar totalização por nível hierárquico:

[https://github.com/okfn-brasil/gastos\\_abertos/blob/master/utils/generate\\_total\\_json.py](https://github.com/okfn-brasil/gastos_abertos/blob/master/utils/generate_total_json.py)

4. Gráfico de barras com Drilldown para os dados da Receita utilizando a biblioteca em Javascript HighCharts.
5. Criação de tabela sincronizada com o nível que está sendo visualizado no gráfico de barras. A tabela foi construída utilizando a biblioteca em Javascript Datatables.
6. Criação de ferramenta para gerenciar as páginas e conteúdos que irão existir no site Gastos Abertos:  
<https://github.com/aivuk/flaskyll>  
Com fork modificado para o projeto em:  
[https://github.com/okfn-brasil/gastos\\_abertos\\_website/](https://github.com/okfn-brasil/gastos_abertos_website/)

## Direção atual.

Após criarmos uma API para os dados da Receita, antes de iniciarmos o desenvolvimento de uma API para os dados do Planejamento e da Execução Orçamentária pensamos em como poderíamos ter um projeto mais genérico e de maior interesse para outros grupos. Com isso, ao invés de replicarmos o trabalho realizado na Receita com os dados de Planejamento/Execução pensamos no desenvolvimento de uma ferramenta que facilitará a criação de uma API a partir de um arquivo CSV.

Além disso notamos a importância de pensarmos em histórias ou explorações mais pontuais dos dados que possuímos. Ao invés de focarmos apenas numa visualização genérica e no desenvolvimento de novas ferramentas, iremos pensar em casos concretos de uso dos dados para explorar questões pontuais.

O resultado disso será:

- Um ou mais artigos levantando algum tema (por exemplo, o planejamento e execução das ciclovias na cidade de São Paulo).
- Para esses artigos serão criadas visualizações dos dados que possuímos já utilizando ferramentas existentes, customizando tais soluções se necessário. Pense no caso do tema ciclovias, poderíamos utilizar o CartoDB para criar um mapa de cores por subprefeitura do valor total gastos com ciclovias nessa região, ou utilizar soluções como HighCharts para criar gráficos comparando os diferentes custos em obras de mobilidade urbana. No processo de criação dessas visualizações, poderá surgir a necessidade de desenvolvermos uma nova ferramenta de visualização de dados que possivelmente poderá ser utilizado por outros desenvolvedores ou jornalistas.
- Tutoriais/Textos para jornalistas ou não desenvolvedores de como foram realizadas as visualizações ou análises dos dados.

- Tutoriais/Textos para desenvolvedores de qualquer tipo de trabalho mais técnico que foi necessário para a produção e análise dos dados. Por exemplo a produção de um texto explicando como foi realizado determinada limpeza dos dados, raspagem de websites via Selenium, agregações ou totalizações dos dados usando ferramentas como a biblioteca Pandas do Python, etc

### **Colaboração (em construção) com o Open Knowledge Internacional na construção do novo OpenSpending**

Estamos construindo uma [colaboração](#) com a Open Knowledge Internacional, que está [redesenhando o Open Spending](#). Ambos projetos, Gastos Abertos e Open Spending, estão construindo suas ferramentas a partir de [microserviços](#).

### **Resumo dos próximos passos**

1. Ferramenta genérica para gerar uma [API RESTful a partir de um CSV](#). Veja também a [proposta](#) para arquitetura dessa ferramenta.
2. Colaboração no desenvolvimento do [OpenSpending 2.0]([https://github.com/okfn-brasil/documents/blob/master/gastos\\_abertos/gastos\\_abertos\\_architecture.pdf](https://github.com/okfn-brasil/documents/blob/master/gastos_abertos/gastos_abertos_architecture.pdf))
3. Programa Javascript para realizar comparações de planejamento e execução orçamentária.
4. Criação de histórias a partir dos dados que possuímos
5. Criação de cursos voltado a desenvolvedores, jornalistas e atores do terceiro setor para o uso eficiente de ferramentas para explorar os dados públicos orçamentários já disponibilizados pela prefeitura de São Paulo.