

SALÃO DE  
INICIAÇÃO CIENTÍFICA  
**XXIX SIC**  
  
**UFRGS**  
PROPESQ



múltipla   
**UNIVERSIDADE**  
inovadora  inspiradora

|                   |   |
|-------------------|---|
| <b>Evento</b>     | Salão UFRGS 2017: SIC - XXIX SALÃO DE INICIAÇÃO CIENTÍFICA DA UFRGS |
| <b>Ano</b>        | 2017  |
| <b>Local</b>      | Campus do Vale  |
| <b>Título</b>     | Descoberta de padrões em dados urbanos georreferenciados            |
| <b>Autor</b>      | MARIA CLARA MACHRY JACINTHO   |
| <b>Orientador</b> | ANA LUCIA CETERTICH BAZZAN  |

## **Descoberta de padrões em dados urbanos georreferenciados**

Orientadora: Profa. Ana Bazzan

Aluna: Maria Clara Jacintho

PET Computação UFRGS

Universidade Federal do Rio Grande do Sul

A Lei de Dados Abertos, responsável por regular o acesso a informações governamentais, incentivou muitas prefeituras a compartilharem com o público diversos tipos de dados sobre educação, saúde pública, infraestrutura, entre outros. Esses dados podem ser utilizados para se ter uma visão interna do funcionamento de uma cidade.

Graças à abundância de informações georreferenciadas fornecidas online pela prefeitura de São Paulo, foi decidido fazer um projeto de coleta e análise de dados paulistanos referentes à educação. O projeto tem como objetivo a descoberta de padrões socioeconômicos nos dados georreferenciados de forma a auxiliar a tomada de decisões do poder público.

O projeto atualmente está na fase de coleta de dados, onde estamos explorando os conjuntos de dados disponibilizados pela prefeitura de São Paulo e o Governo Federal. Estamos verificando a qualidade, completude e consistência dessas informações e selecionando aquelas que serão mais úteis para a pesquisa, como notas do ENEM por escola, infraestrutura disponível, renda média e IDH da região.

A continuação do projeto prevê primeiramente a criação de um banco de dados, reunindo todas as informações coletadas de forma a facilitar a pesquisa. Será feito um relatório descrevendo a disponibilidade e qualidade dos dados oferecidos. Utilizaremos então Python e bibliotecas como Pandas, NumPy e SciPy para a descobrir de padrões simples nos dados. O resultado esperado é a formulação de um artigo descrevendo o projeto e os padrões descobertos em detalhe. É considerada também a criação de uma plataforma como para expor os dados e os padrões encontrados.