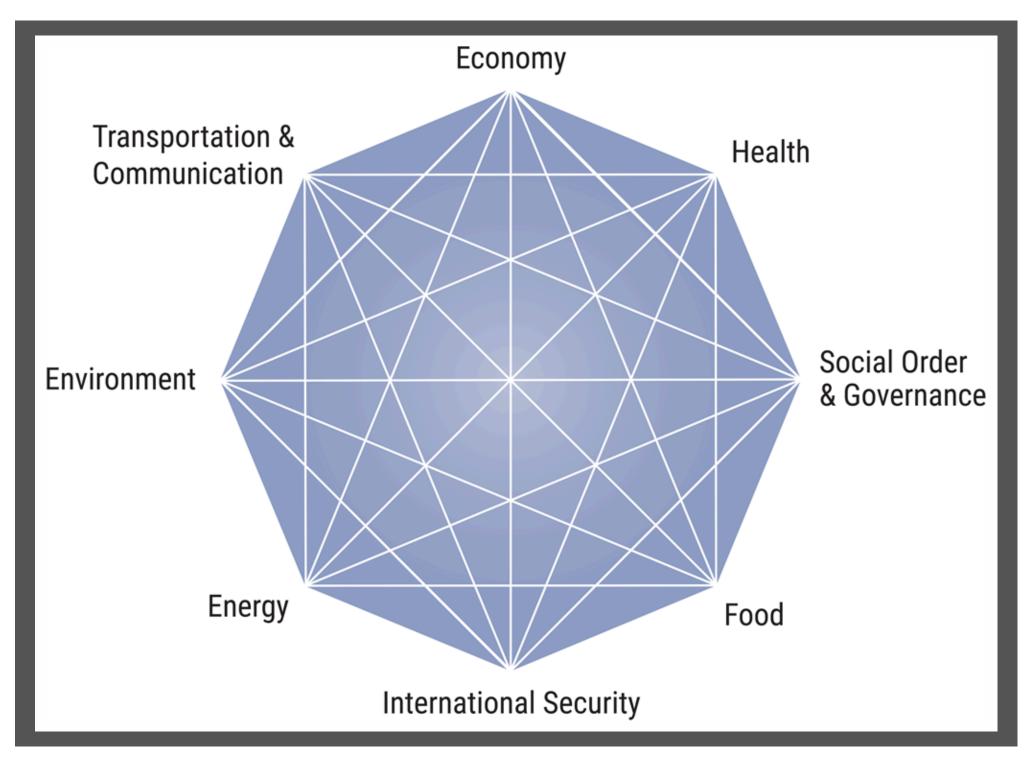
OPEN DATA DAY 2025 - PY013 >>>>>

FUGADE CEREBROS NO SERVIÇO PÚBLICO



Policrises?

- Crises interligadas e sobrepostas
- Mundo mais conectado = problemas mais globais
- Difíceis de encontrar origens e soluções simples



Fonte: Global polycrisis: the causal mechanisms of crisis entanglement

Overview

Quando profissionais não ingressam ou desistem do setor público...

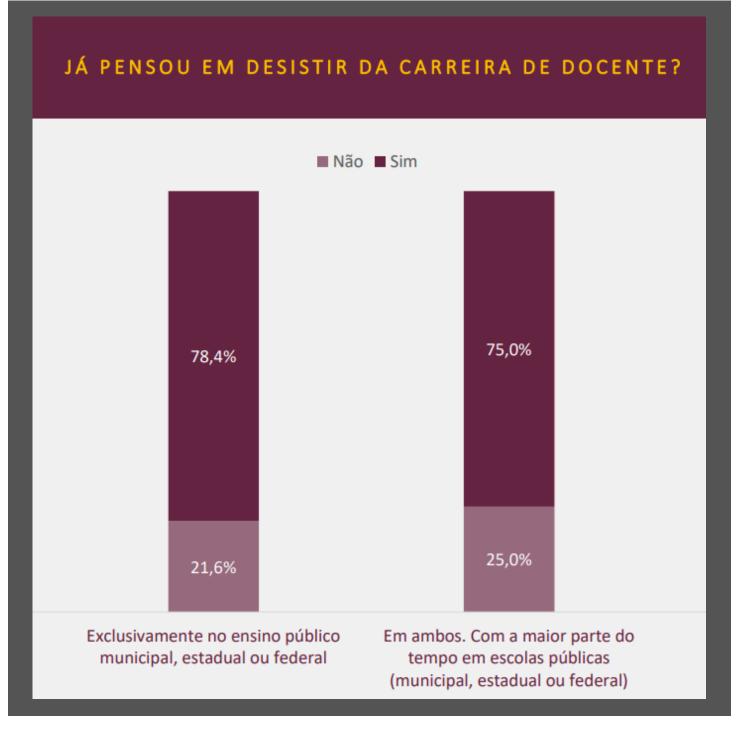
- Possível perca dos mais qualificados
- Lacunas na prestação de serviços
- Fuga literal de pessoas (=impostos)
 para outras regiões
- Desmotivação dos remanescentes
- Dificuldades em manter conhecimentos de longo prazo



Overview

Causas gerais da desistência:

- Melhores oportunidades
- Imagem negativa do setor público
- Estagnação técnica
- Segurança (para locais afastados)
- Incerteza política
- Pior cenário em posições que interagem com o cidadão



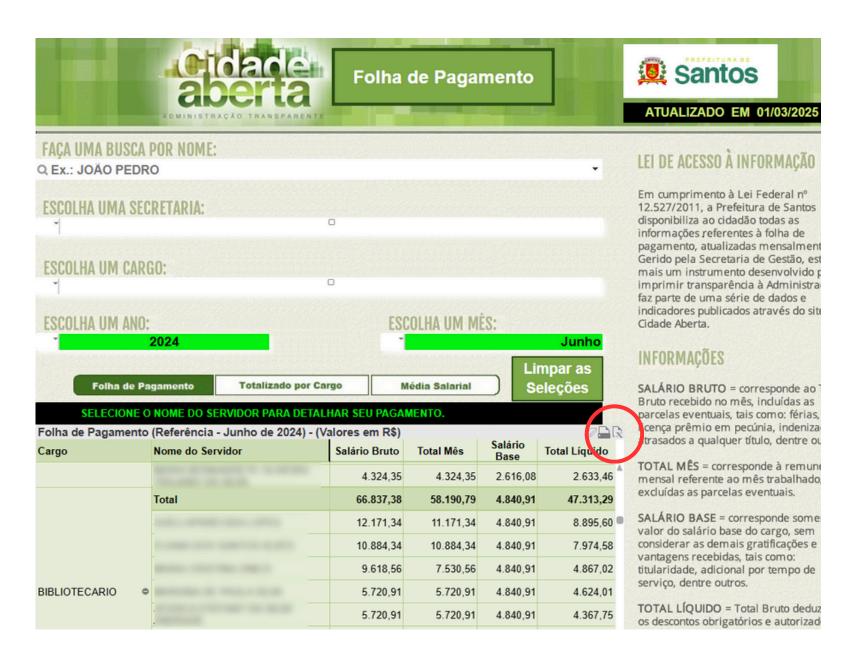
Fonte: Perfil e Desafios dos Professores da Educação Básica no Brasil

Dataset

Concursos públicos recentes (2023)



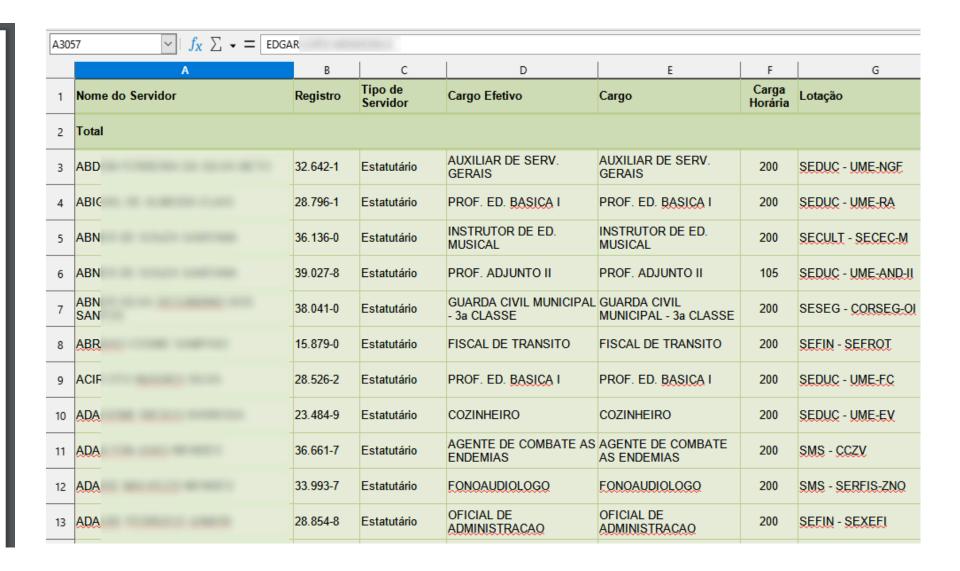
Portal da Transparência de Santos



Dataset

Concursos públicos recentes (2023)
 Portal da Transparência de Santos

Secretaria Municipal de Finanças e Gestão de Santos / SP Processo Seletivo Simplificado – Edital nº 68/2023 CANDIDATOS CLASSIFICADOS INSCRITOS COMO PESSOA COM DEFICIÊNCIA - PCD ORDEM ALFABÉTICA POR FUNÇÃO DT. NASCIM. EXPERIÊNCIA TÍTULOS ANDREA CRISTINA GROSSO JACQUELLYNE STHEPHANY RANGEL PUERTAS 03/05/1994 19.00 HABILITADO 0491000943 0491001465 MARLENE APARECIDA DE MELO LIMA 04/11/1967 4,00 HABILITADO RINALDO DE ARAUJO PAES Função: 301 - ASSISTENTE SOCIAL LUCIANA DA COSTA MAIKON BORGES DA SILVA m32****.* 0491002750 RENE BALDUTTI HABILITADO 28/02/1966 50.00 THIAGO VIEIRA RODRIGUES



Dataset





ODD 2025 - Py013

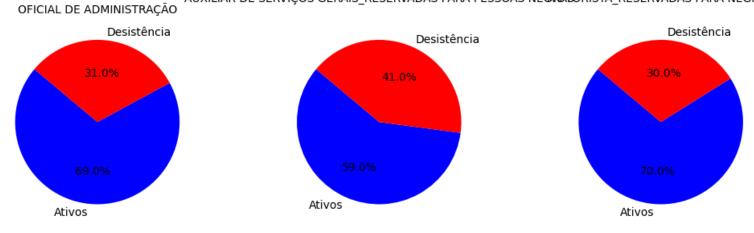
Col. dedes subsetdes - Cl

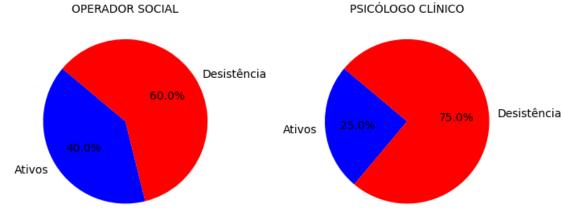
Instalação dos pacotes necessários e criação de constantes

```
[1]: import pdfplumber
                     import re
                     import pandas as pd
                    import matplotlib.pyplot as plt
                    from unidecode import unidecode
                    from collections import defaultdict
                    from os import listdir
                    # Caminho do arquivo PDF e Excel
                    pasta_dados = 'Result_Concursos'
                    excel_path = "servidores_santos.xlsx"
                    # Expressões regulares para extração
                     regex_clasificacao_especial = re.compile(r'(COMO PESSOA (.+?) [\(-])|(RESERVADAS (.+?)\r?\n)', flags=re.IG
                     regex\_linha\_candidato = re.compile(r"\d?\s?\d\{9,10\}\s([A-Z\`{A}-\ddot{y}\s]+)\s.+\s(\d+)\s+HABILITADO\r?\n?", flags=regex\_linha\_candidato = re.compile(r"\d?\s?\d\{9,10\}\s([A-Z\`{A}-\ddot{y}\s]+)\s.+\s(\d+)\s+HABILITADO\r?\n?", flags=regex\_linha\_candidato = re.compile(r"\d?\s?\d\{9,10\}\s([A-Z\`{A}-\ddot{y}\s]+)\s.+\s(\d+)\s+HABILITADO\r?\n?", flags=regex\_linha\_candidato = re.compile(r"\d?\s?\d\{9,10\}\s([A-Z\`{A}-\ddot{y}\s]+)\s.+\s(\d+)\s+HABILITADO\r?\n?", flags=regex\_linha\_candidato = re.compile(r"\d?\s?\d(\d+)\s+\d(\d+)\s+\d(\d+)\s+\d(\d+)\s+\d(\d+)\s+\d(\d+)\s+\d(\d+)\s+\d(\d+)\s+\d(\d+)\s+\d(\d+)\s+\d(\d+)\s+\d(\d+)\s+\d(\d+)\s+\d(\d+)\s+\d(\d+)\s+\d(\d+)\s+\d(\d+)\s+\d(\d+)\s+\d(\d+)\s+\d(\d+)\s+\d(\d+)\s+\d(\d+)\s+\d(\d+)\s+\d(\d+)\s+\d(\d+)\s+\d(\d+)\s+\d(\d+)\s+\d(\d+)\s+\d(\d+)\s+\d(\d+)\s+\d(\d+)\s+\d(\d+)\s+\d(\d+)\s+\d(\d+)\s+\d(\d+)\s+\d(\d+)\s+\d(\d+)\s+\d(\d+)\s+\d(\d+)\s+\d(\d+)\s+\d(\d+)\s+\d(\d+)\s+\d(\d+)\s+\d(\d+)\s+\d(\d+)\s+\d(\d+)\s+\d(\d+)\s+\d(\d+)\s+\d(\d+)\s+\d(\d+)\s+\d(\d+)\s+\d(\d+)\s+\d(\d+)\s+\d(\d+)\s+\d(\d+)\s+\d(\d+)\s+\d(\d+)\s+\d(\d+)\s+\d(\d+)\s+\d(\d+)\s+\d(\d+)\s+\d(\d+)\s+\d(\d+)\s+\d(\d+)\s+\d(\d+)\s+\d(\d+)\s+\d(\d+)\s+\d(\d+)\s+\d(\d+)\s+\d(\d+)\s+\d(\d+)\s+\d(\d+)\s+\d(\d+)\s+\d(\d+)\s+\d(\d+)\s+\d(\d+)\s+\d(\d+)\s+\d(\d+)\s+\d(\d+)\s+\d(\d+)\s+\d(\d+)\s+\d(\d+)\s+\d(\d+)\s+\d(\d+)\s+\d(\d+)\s+\d(\d+)\s+\d(\d+)\s+\d(\d+)\s+\d(\d+)\s+\d(\d+)\s+\d(\d+)\s+\d(\d+)\s+\d(\d+)\s+\d(\d+)\s+\d(\d+)\s+\d(\d+)\s+\d(\d+)\s+\d(\d+)\s+\d(\d+)\s+\d(\d+)\s+\d(\d+)\s+\d(\d+)\s+\d(\d+)\s+\d(\d+)\s+\d(\d+)\s+\d(\d+)\s+\d(\d+)\s+\d(\d+)\s+\d(\d+)\s+\d(\d+)\s+\d(\d+)\s+\d(\d+)\s+\d(\d+)\s+\d(\d+)\s+\d(\d+)\s+\d(\d+)\s+\d(\d+)\s+\d(\d+)\s+\d(\d+)\s+\d(\d+)\s+\d(\d+)\s+\d(\d+)\s+\d(\d+)\s+\d(\d+)\s+\d(\d+)\s+\d(\d+)\s+\d(\d+)\s+\d(\d+)\s+\d(\d+)\s+\d(\d+)\s+\d(\d+)\s+\d(\d+)\s+\d(\d+)\s+\d(\d+)\s+\d(\d+)\s+\d(\d+)\s+\d(\d+)\s+\d(\d+)\s+\d(\d+)\s+\d(\d+)\s+\d(\d+)\s+\d(\d+)\s+\d(\d+)\s+\d(\d+)\s+\d(\d+)\s+\d(\d+)\s+\d(\d+)\s+\d(\d+)\s+\d(\d+)\s+\d(\d+)\s+\d(\d+)\s+\d(\d+)\s+\d(\d+)\s+\d(\d+)\s+\d(\d+)\s+\d(\d+)
                     regex_linhas_por_cargo = re.compile(r"((?:\d?\s?\d{9,10}.+?(\n|$))+)")
```

Extraindo dados dos PDFs de classificações:

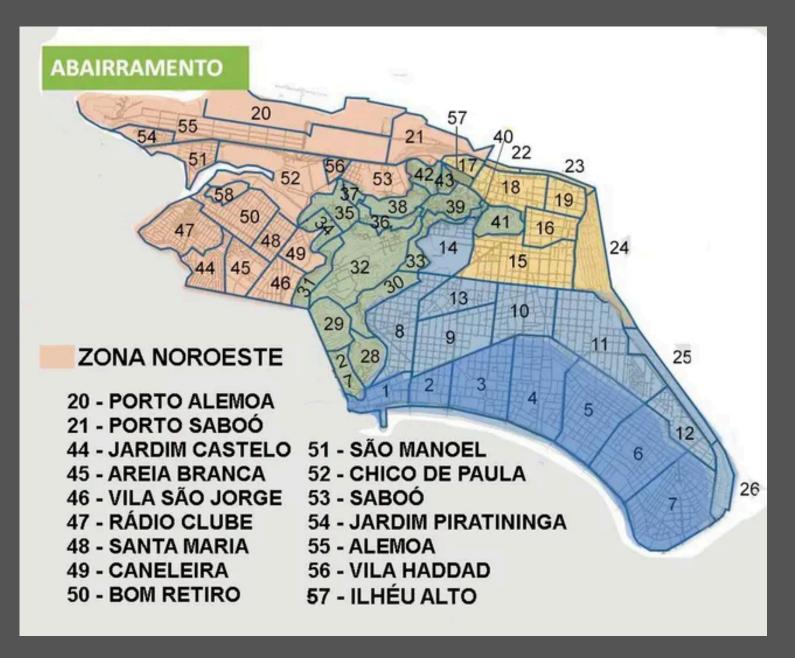
AUXILIAR DE SERVIÇOS GERAIS_RESERVADAS PARA PESSOAS NE**GRAT**ORISTA_RESERVADAS PARA NEGRC





Possíveis passos fu<u>turos</u>

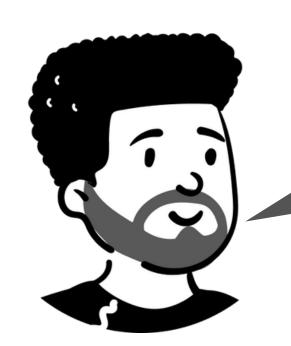
- Geo mapeamento
- Comparação por salários
- Comparação com outras cidades
- Verificação com dados gerais (violência, custo de vida etc)



Links

- https://www.institutomais.org.br
- https://www.santos.sp.gov.br/?
 q=content/transparencia-cargos-salarios

```
ados extraidos = []
ith pdfplumber.open(pdf_path) as pdf:
  for pagina in pdf.pages:
      texto_pagina = pagina.extract_text().upper()
      if texto_pagina:
          classificacao_especial = re.search(regex_clas:
          for grupo_funcao in regex_linhas_por_cargo.fi
              bloco_profissional = grupo_funcao.group(1
              primeira matricula = bloco profissional[::
              nome_cargo = re.compile(str_regex_nome_car
              for convocado in regex_linha_candidato.fi
                  dados extraidos.append({
                       "Cargo_Edital": nome_cargo + ('_'
                      "Classificação": convocado[1],
                      "Nome": convocado[0].strip()
                  })
```



Python





