# Cenário

Analisar dados sobre a segurança no Estado de São Paulo. Observando qual ano teve mais homicídio doloso, roubo de veículos... Assim, teremos à disposição, informações para identificar qual ano/mês, a segurança de SP foi desfoque em tratar de acontecimentos incoerentes que coloca a vida do paulistano, em risco.

# Perguntas

* Qual ano teve mais homicídio doloso por região?
* Qual o total acumulado de casos de tentativas de homicídio por ano?
* Qual a diferença do total de homicídio doloso e culposo em acidentes de trânsitos, por região?
* Qual a região com mais vítimas de latrocínios?
* Qual o total e a média de latrocínio a partir do ano de 2019?

|  |  |
| --- | --- |
| FONTE DE DADOS | FERRAMENTA PARA ANÁLISE |
| <https://basedosdados.org/dataset/dbd717cb-7da8-4efd-9162-951a71694541> | Big Query |

# Localização dos Dados

* Onde seus dados estão localizados? Base de Dados – aberto, retirados da secretária de segurança do Estado de São Paulo.
* Como os dados são organizados? São organizados em colunas – relacionais, com a maioria dos dados sendo numéricos do tipo int.
* Existem problemas com viés ou credibilidade nesses dados? Não
* Como você está lidando com o licenciamento, a privacidade, a segurança e a acessibilidade? Os dados são de fonte aberta, e retirados diratamente da base de dados do governo.
* Como você verificou a integridade dos dados? Através da fonte e analisando os tipos de dados.

# Limpeza dos Dados

* Quais ferramentas você está escolhendo e por quê? Big Query, porque é uma ótima ferramenta para analisar dados de forma rápida e eficaz.
* Você garantiu a integridade dos seus dados? Sim, analisei se existia valores nulos e se os tipos de dados correspondem ao objetivo proposto.
* Como você pode verificar se seus dados estão limpos e prontos para análise? Analiso os tipos de dados, vejo se os valores correspondem-se, observo a existência de espaço entre palavras ou incoerência em relação a formatação.

# Analisar

-- Ano que mais teve homicídio doloso por região em SP, a partir do ano de 2010

SELECT ano,regiao\_ssp, homicidio\_doloso

FROM `basedosdados.br\_sp\_gov\_ssp.ocorrencias\_registradas`

WHERE ano >= 2010 AND homicidio\_doloso <> 0

--Qual o total por ano de casos de tentativas de homicídio

SELECT ano, SUM(tentativa\_de\_homicidio) AS Acumulado

FROM `basedosdados.br\_sp\_gov\_ssp.ocorrencias\_registradas`

GROUP BY ano

ORDER BY ano



--Qual ano teve mais caso de tentativa de homicídio

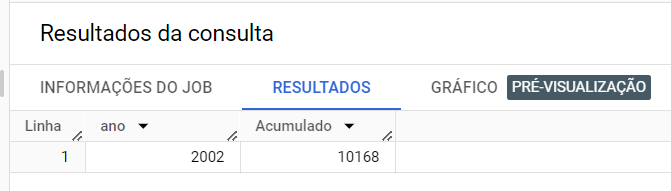
SELECT ano, SUM(tentativa\_de\_homicidio) AS Acumulado

FROM `basedosdados.br\_sp\_gov\_ssp.ocorrencias\_registradas`

GROUP BY ano

ORDER BY ano

LIMIT 1



--Qual a diferença de homicídio doloso e culposo por transito do total geral?

--Qual a diferença de homicidio doloso e culposo por transito do total geral?

WITH homicidioDoloso AS (

  SELECT id\_municipio, numero\_de\_vitimas\_em\_homicidio\_doloso\_por\_acidente\_de\_transito,

  ROUND(numero\_de\_vitimas\_em\_homicidio\_doloso\_por\_acidente\_de\_transito  / SUM(numero\_de\_vitimas\_em\_homicidio\_doloso\_por\_acidente\_de\_transito),2) AS porcentagemHD

  FROM `basedosdados.br\_sp\_gov\_ssp.ocorrencias\_registradas`

  WHERE numero\_de\_vitimas\_em\_homicidio\_doloso\_por\_acidente\_de\_transito <> 0

    GROUP BY numero\_de\_vitimas\_em\_homicidio\_doloso\_por\_acidente\_de\_transito, id\_municipio

),

homicidioCulposo AS(

  SELECT id\_municipio, homicidio\_culposo\_por\_acidente\_de\_transito, ROUND(homicidio\_culposo\_por\_acidente\_de\_transito/ SUM(homicidio\_culposo\_por\_acidente\_de\_transito),2)AS porcentagemHC

  FROM `basedosdados.br\_sp\_gov\_ssp.ocorrencias\_registradas`

  WHERE homicidio\_culposo\_por\_acidente\_de\_transito <> 0

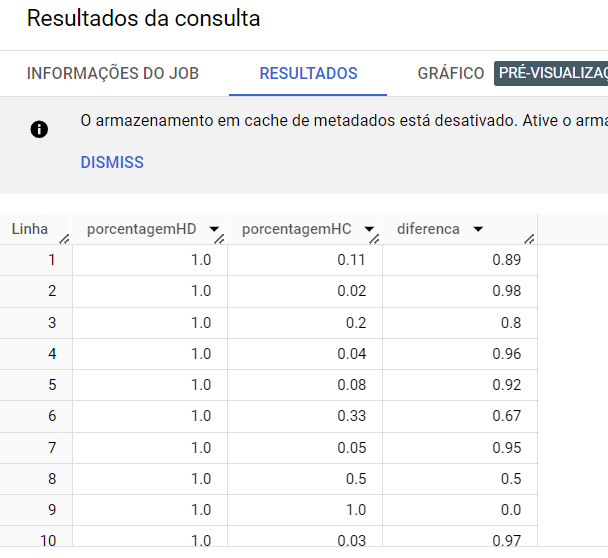
  GROUP BY homicidio\_culposo\_por\_acidente\_de\_transito, id\_municipio

)

SELECT hd.porcentagemHD, hc.porcentagemHC, ROUND(hd.porcentagemHD - hc.porcentagemHC,2) AS diferenca

FROM homicidioDoloso hd

INNER JOIN homicidioCulposo hc ON hd.id\_municipio  = hc.id\_municipio



-- Qual a região com mais vítimas de latrocínios?

SELECT  regiao\_ssp,ano,numero\_de\_vitimas\_em\_latrocinio, SUM(numero\_de\_vitimas\_em\_latrocinio)  AS vitimas\_latrocinio

FROM `basedosdados.br\_sp\_gov\_ssp.ocorrencias\_registradas`

GROUP BY regiao\_ssp, ano, numero\_de\_vitimas\_em\_latrocinio

ORDER BY numero\_de\_vitimas\_em\_latrocinio DESC

LIMIT 10



-- Qual o total e a média de latrocinio a partir do ano de 2019?

SELECT ano, mes, SUM(latrocinio) AS latrocinio, ROUND(AVG(latrocinio) \* 100,2) AS mediaLatrocinio

FROM `basedosdados.br\_sp\_gov\_ssp.ocorrencias\_registradas`

WHERE ano >= 2019 AND latrocinio IS NOT NULL

group by mes, ano

ORDER BY ano, mes ASC

