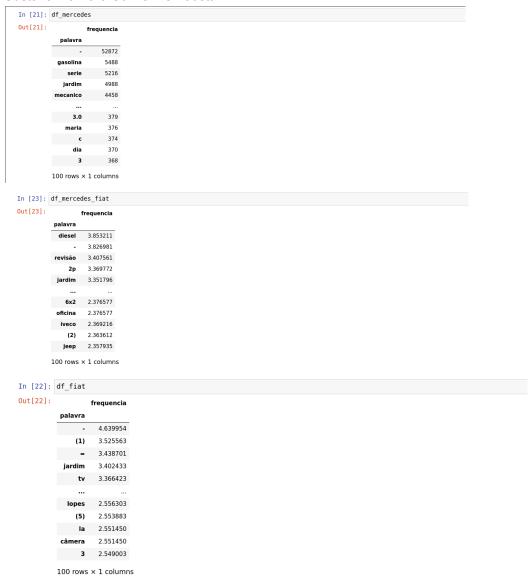
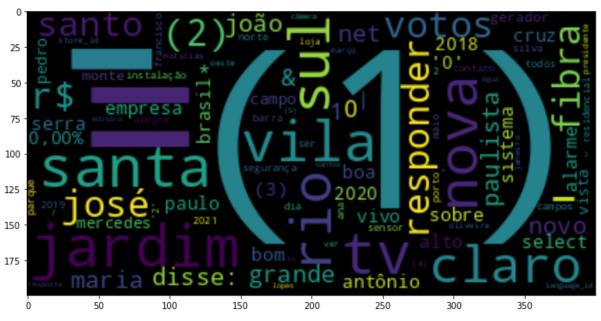
Projeto 2: Analisando dados com Spark

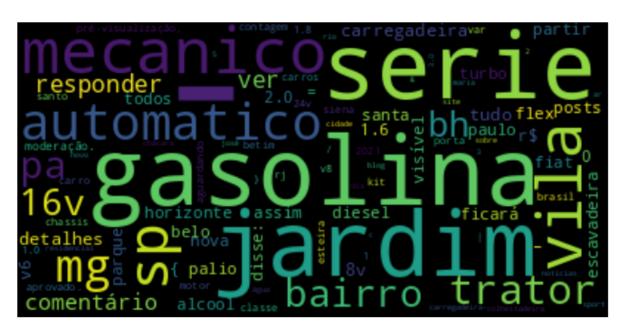
Gustavo Molina e Guilherme Batista



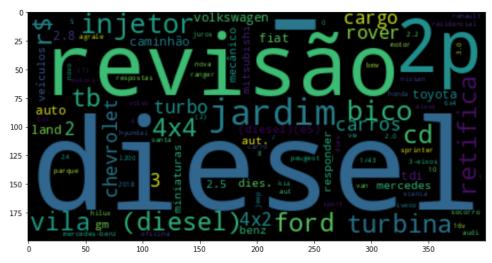
Análise e visualização dos resultados



Img: WorldCloud Mercedes.



Img: WorldCloud Fiat.



Img: WorldCloud Mercedes e Fiat

Com base nos resultados obtidos nessa worldCloud é razoável pensar que as palavras com frequência maior entre Mercedes e Fiat se deve a uma comparação que as pessoas estão fazendo entre os dois carros, com fatores como revisão e diesel como as palavras com maior frequência encontrada. Outras palavras que fortalecem essa *hipótese* são outras palavras com frequência considerável encontradas, como chevrolet, volkswagen, audi, etc, que indicam algum tipo de benchmark.

Para a Mercedes, no WorldCloud algumas hipóteses interessantes também podem ser feitas: a palavra Santa, João e Maria poderiam indicar que algum personagem ou pessoa famosa utiliza o carro em algum vídeo no Youtube e/ou ficou famosa por algo relacionada a marca.

Já no Fiat, a frequência de palavras mais alta está relacionada a estados como SP, MG, Belo e Horizonte (separadas). Seus usos indicam ser mais relacionados ao cotidiano do povo brasileiro, que tem uma frequência de palavras buscadas muito maior que a da Mercedes. (como pode ser visto nas tabelas acima)