**Relatório**

**Limpeza dos Dados:**

O conjunto de dados da pesquisa tem um volume de 128. Antes de fazer a análise precisei fazer a limpeza e algumas alterações nos dados. Alguns dados estavam faltando, outros estavam inconsistentes. Segue o registro das alterações:

1. Retira do registro número 75. Não havia informações suficientes nesse registro. Decidi retirar o registro pois na postagem do instagram referente à pesquisa consta 127 registros.
2. O registro número 89 não tinha informações sobre o cep ou bairro, porém, em uma das respostas por extenso informava o bairro onde morava. Com isso, adicionei as informações antes inexistentes.
3. Nos registros que não informaram o cep, bairro ou cidade foi acrescentado o valor “Não Informou” nas colunas Qual seu CEP ou Bairro na RMB ?,Bairro e Cidade, afim de utilizar as informações restantes no registro.
4. Nos registros que informaram o cep inválido, foi acrescentado o valor “Não Informou” nas colunas Qual seu CEP ou Bairro na RMB ?,Bairro e Cidade, afim de utilizar as informações restantes no registro.
5. No registro número 10, as informações do bairro e da cidade estavam na mesma coluna. As informações foram separadas e colocadas nas devidas colunas.
6. Um dos registros utilizou “Ananindeua” como bairro, nesse registro inseri o valor “Não Informou” na coluna Bairro.
7. Alguns bairros tinham o nome duplicados pelo fato de ter um espaço ao final do nome. Ex: “Guamá” e “Guamá “. Foi feita uma padronização nesses casos.
8. Alguns valores na coluna Bairro tinham a informação do o nome do bairro e do distrito. Retirei a informação do distrito para deixar a coluna apenas com a informação solicitada.
9. Na coluna Vezes por Semana, o valor 0 (“zero”) não representa a informação verdadeira. A coluna Vezes por Semana é uma contagem da coluna Qual a os dias da semana da coleta de lixo ATUALMENTE na sua rua?. Caso a pessoa responda “Não sei dizer exatamente”, o valor representado na tabela Vezes por Semana deveria ser “Nulo”, o valor 0 (“zero”) deveria ser usado caso a resposta informasse que não existe coleta, por conta disso, transformei os valores 0 em Nulos.
10. As colunas que dizem respeito à percepção dos colaboradores de forma quantitativa (notas de 1 a 5) estão sendo representadas ao contrário. Notas 1 estão sendo consideradas boas e notas 5 consideradas péssimas. Se formos, por exemplo, medir a escala de satisfação de um produto, e usar-mos a classificação de 1 a 5, produtos com média alta indicariam maior satisfação.   
    Fiz a inversão desses valores.
11. Criei a coluna nota\_comportamentoDosMoradores para guardar com a nova métrica as informações da coluna Sobre o comportamento dos moradores acerca da dispensa do lixo em locais adequados, NA SUA RUA, qual sua impressão:.
12. Criei a coluna nota\_depositoDeLixo para guardar com a nova métrica as informações da coluna Sobre os locais e contêineres para deposito de lixo oferecidos pela Prefeitura para descarte, avalie:.
13. Alguns registros na coluna Qual a os dias da semana da coleta de lixo ATUALMENTE na sua rua? tinham inconsistências. Os registros mostravam dias de coleta juntamente com a informação “Não sei dizer exatamente”. Nesses casos, atribui o valor “Não sei dizer exatamente”.
14. Sintetizei as informações da coluna Quanto a assiduidade (frequência) da coleta: na nova coluna avaliacao\_frequenciaDaColeta.
15. Sintetizei as informações da coluna 'Há 2 anos, como era a frequência de coleta de lixo na rua rua' na nova coluna histórico\_coleta. Nela contem a informação da percepção dos morados se a coleta melhorou, continua da mesma forma, ou se piorou.
16. Criei uma coluna para qualificar o comportamento dos moradores utilizando a avaliação quantitativa da coluna
17. Retirei colunas com informações duplicadas.
18. Retirei coluna com informações que, a princípio, não fariam parte da análise.