Atividade 2 - Clusterização Geográfica de Entregas com Falha

# 1. Objetivo

O objetivo desta atividade foi aplicar uma técnica de análise de agrupamento (KMeans) sobre uma base de dados contendo falhas de entrega, com localização geográfica (latitude e longitude) e valores monetários das entregas. A finalidade foi identificar a quantidade ideal de centros de distribuição e sugerir regiões estratégicas para sua instalação, baseando-se na densidade e no valor movimentado por região.

# 2. Metodologia

Foi utilizada a técnica de clusterização KMeans, com a definição do número ideal de clusters feita através do método do cotovelo. Após a aplicação do KMeans com k=4, foi realizada a análise financeira por cluster e a localização dos centroides foi convertida em endereços aproximados via geocodificação reversa (OpenStreetMap).

# 3. Análise Financeira por Cluster

|  |  |  |  |
| --- | --- | --- | --- |
| Cluster | Total de Entregas | Soma dos Valores (R$) | Média por Entrega (R$) |
| 0 | 175 | 11067.97 | 63.25 |
| 1 | 207 | 12838.25 | 62.02 |
| 2 | 126 | 7837.2 | 62.2 |
| 3 | 145 | 9440.39 | 65.11 |

- Cluster 1 lidera em volume de entregas e valor total, sendo a região de maior prioridade estratégica.

- Cluster 0 também apresenta alta demanda e movimentação financeira, reforçando a necessidade de um centro regional robusto.

- Cluster 3 tem menos entregas, mas um **ticket médio mais alto,** o que indica entregas mais valiosas ou clientes estratégicos.

- Cluster 2 aparece como o menos expressivo em ambas as métricas, mas ainda relevante para garantir cobertura total e equilíbrio logístico.

# 4. Localização dos Centros de Distribuição Sugeridos

Após a clusterização dos dados de geolocalização, é possível identificar os centros geográficos (centroides) de cada grupo, representando os pontos mais centrais de cada cluster.

Para facilitar a interpretação geográfica e estratégica, utilizamos o serviço de geocodificação reversa da biblioteca **geopy** com o provedor **Nominatim (OpenStreetMap).**  
Com isso, conseguimos converter coordenadas de latitude e longitude em endereços reais, como nome da rua, cidade e estado.

|  |  |  |
| --- | --- | --- |
| Cluster | Região Sugerida | Cidade de Referência |
| 0 | Região Metropolitana de Campinas | Atibaia |
| 1 | Noroeste Paulista | Guararapes (Araçatuba) |
| 2 | Centro-Oeste Paulista | Itatinga (Botucatu) |
| 3 | Nordeste Paulista | Jaboticabal (Ribeirão Preto) |

Mapa

O conteúdo gerado por IA pode estar incorreto.

* Cluster 0: Guaxinduva, Atibaia, Região de Bragança Paulista – SP
* Cluster 1: Estrada Municipal, Sítio Safé II, Guararapes – SP
* Cluster 2: Rodovia Vicinal Doutor Ene Sab, Itatinga – SP
* Cluster 3: Jaboticabal – SP

**5. Recomenda-se priorizar os seguintes clusters na ordem:**

* **Cluster 1 → Guararapes (Araçatuba):** Maior volume e maior valor total de entregas.
* **Cluster 0 → Atibaia:** Alto volume e excelente localização estratégica próxima à Região Metropolitana.
* **Cluster 3 → Jaboticabal:** Menor volume, porém maior ticket médio — pode indicar entregas de maior valor agregado.
* **Cluster 2 → Itatinga:** Menor cluster, mas importante para cobertura da região central do estado.

# 6. Conclusão

A aplicação da análise de clusters com KMeans permitiu identificar 4 regiões estratégicas para a instalação de centros de distribuição. A análise geográfica foi validada com visualizações e os endereços aproximados foram obtidos por geocodificação. Essa abordagem permite à empresa tomar decisões mais embasadas sobre logística e cobertura regional.