# Clustering: Customer Clustering

## Versão 1:

### Atualizações de v0 para v1:

- Diferenciação entre variáveis categóricas e numéricas apesar de todas elas estarem em estado numérico.

- Diferenciação entre variáveis categóricas binárias e ordinais.

- Pré-processamento diferente para cada grupo.

- Novo modelo salvo com os novos resultados.

### Motivo das atualizações:

Eu havia feito todas as features passarem pelo pré-processamento de features numéricas, mas depois percebi que isso era um erro.

Não faz sentido padronizar features categóricas, da mesma forma que é perigoso tratar features ordinais, não ordinais, e binárias da mesma forma, uma vez que isso influencia a forma como o algoritmo aprende.

### Resultados:

Obtive melhores resultado em relação a inércia do K-Means. Na v0 a inércia iniciava em 10000 e na v1 está iniciando em 4000.

Método Silhouette – Silhouette Score: -0.20

Método Elbow - Inertia: 781.47

Redução da Inércia Inicial - Percentual de Redução: 91.30%