

Customer Clusters Documentation

Methodology:

The clustering process was performed using the KMeans algorithm.

Features considered:

1. AvgAmountPerTransaction
2. Region_Encoded (numerical encoding of the Region column)

The Elbow Method was used to determine the optimal number of clusters ($k=3$).

Each customer was assigned to one of the clusters based on their feature values.

Cluster 0

Customer IDs in Cluster 0:

C0001

C0004

C0005

C0009

C0011

C0012

C0013

C0018

C0019

C0020

C0021

C0022

C0024

C0026

C0028

C0037

C0041

C0045

C0046

C0047

C0051

C0052

C0053

C0056

C0057

C0061

C0064

C0065

C0067

C0068

C0069

C0070

C0073

C0076

C0086

C0087

C0090

C0092

C0093

C0098

C0099

C0101

C0102

C0103

C0105

C0106

C0107

C0108

C0116

C0117

C0118

C0120

C0122

C0123

C0124

C0125

C0126

C0127

C0128

C0129

C0130

C0131

C0132

C0133

C0135

C0137

C0139

C0140

C0143

C0145

C0148

C0149

C0150

C0152

C0153

C0154

C0156

C0159

C0161

C0162

C0163

C0164

C0166

C0167

C0170

C0171

C0173

C0175

C0177

C0179

C0181

C0182

C0183

C0186

C0188

C0191

C0192

C0193

C0194

C0198

C0199

C0200

Cluster 1

Customer IDs in Cluster 1:

C0002

C0003

C0008

C0010

C0014

C0017

C0025

C0027

C0029

C0030

C0031

C0032

C0033

C0034

C0035

C0036

C0038

C0039

C0043

C0049

C0055

C0060

C0062

C0063

C0071

C0072

C0075

C0077

C0081

C0083

C0084

C0088

C0091

C0094

C0095

C0097

C0109

C0110

C0111

C0112

C0113

C0119

C0121

C0134

C0136

C0142

C0144

C0147

C0151

C0157

C0160

C0165

C0169

C0172

C0174

C0176

C0178

C0189

C0190

C0195

C0197

Cluster 2

Customer IDs in Cluster 2:

C0006

C0007

C0015

C0016

C0023

C0040

C0042

C0044

C0048

C0050

C0054

C0058

C0059

C0066

C0074

C0078

C0079

C0080

C0082

C0085

C0089

C0096

C0100

C0104

C0114

C0115

C0138

C0141

C0146

C0155

C0158

C0168

C0184

C0185

C0187

C0196