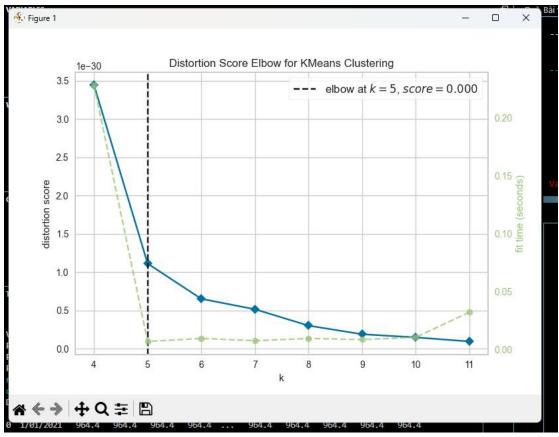
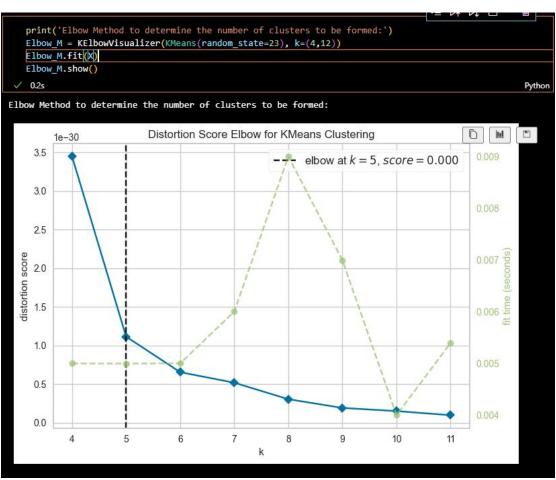


```
# Elbow Method để xác định số cụm cần tạo
print('Elbow Method to determine the number of clusters to be formed:')
Elbow_M = KElbowVisualizer(KMeans(random_state=23), k=(4, 12))
Elbow_M.fit(X)
Elbow_M.show()
```





Với K cụm >= 3

```
Decision Tree
                                       gini = 0.0
                      samples = 365
                                 value = 1.0
[5 rows x 49 columns]
Transformed X shape: (365, 2)
PCA Dataset shape: (365, 2)
Clusters
1 365
Name: count, dtype: int64
Name: count, dtype: into-4

df_new shape: (365, 20)

5 6 predict_proba_0 predict_proba_1 ... predict_proba_14 preds predict_proba predict_
0 964.4 964.4 0.126023 0.186296 ... 0.005479 1 0.186296 1
1 1019.7 1019.7 0.126023 0.186296 ... 0.005479 1 0.186296 1
2 988.4 988.4 0.126023 0.186296 ... 0.005479 1 0.186296 1
3 1002.0 1002.0 0.126023 0.186296 ... 0.005479 1 0.186296 1
4 1061.5 1061.5 0.126023 0.186296 ... 0.005479 1 0.186296 1
Y length: 0
Y values: []
Backend TkAgg is interactive backend. Turning interactive mode on.
```

Với K cụm = 2

```
1002.0
                         0.937008
                                         0.063014
                                                              0.937008
                         0.936894
                                         0.063014
Train index length: 365
Train index: [ 0 1 2
                                                    10 11 12 13 14 15 16 17
  18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35
             39 40
                    41 42 43
                                44 45
                                       46 47 48
                                                   49
      55 56 57 58 59 60 61 62 63 64 65 66 67 68
                                                           69
     73 74 75 76 77 78 79 80 81 82 83 84 85 86 87 88 89 91 92 93 94 95 96 97 98 99 100 101 102 103 104 105 106 107
 108 109 110 111 112 113 114 115 116 117 118 119 120 121 122 123 124 125
 126 127 128 129 130 131 132 133 134 135 136 137 138 139 140 141 142 143 144 145 146 147 148 149 150 151 152 153 154 155 156 157 158 159 160 161
 162 163 164 165 166 167 168 169 170 171 172 173 174 175 176 177 178 179
 180 181 182 183 184 185 186 187 188 189 190 191 192 193 194 195 196 197
 198 199 200 201 202 203 204 205 206 207 208 209 210 211 212 213 214 215
 216 217 218 219 220 221 222 223 224 225 226 227 228 229 230 231 232 233
 234 235 236 237 238 239 240 241 242 243 244 245 246 247 248 249 250 251
 252 253 254 255 256 257 258 259 260 261 262 263 264 265 266 267 268 269
 270 271 272 273 274 275 276 277 278 279 280 281 282 283 284 285 286 287
 288 289 290 291 292 293 294 295 296 297 298 299 300 301 302 303 304 305
 306 307 308 309 310 311 312 313 314 315 316 317 318 319 320 321 322 323
 324 325 326 327 328 329 330 331 332 333 334 335 336 337 338 339 340 341 342 343 344 345 346 347 348 349 350 351 352 353 354 355 356 357 358 359
 360 361 362 363 3641
Y length: 305
Y values: [0]
[200] valid_0's multi_logloss: 4.44089e-15
[200] valid_0's multi_logloss: 4.44089e-15
[200] valid_0's multi_logloss: 4.44089e-15
[200] valid_0's multi_logloss: 4.44089e-15
 Y length: 365
```

```
Elbow Method to determine the number of clusters to be formed:
Silhouette Score for KMeans: 0.7595686474070814
Preds with n_components=2: (array([0], dtype=int64), array([365], dtype=int64))
Preds with n_components=3: (array([0], dtype=int64), array([365], dtype=int64))
Preds with n_components=4: (array([0], dtype=int64), array([365], dtype=int64))
Preds with n_components=5: (array([0], dtype=int64), array([365], dtype=int64))
Preds with n_components=6: (array([0], dtype=int64), array([365], dtype=int64))
Preds with n_components=7: (array([0], dtype=int64), array([365], dtype=int64))
Preds with n_components=8: (array([0], dtype=int64), array([365], dtype=int64))
Preds with n_components=9: (array([0], dtype=int64), array([365], dtype=int64))
Cluster distribution:
Clusters
0 365
Name: count, dtype: int64
df new shape: (365, 7)
             6 predict_proba_0 predict_proba_1 preds predict_proba predict
   964.4
          964.4
                      0.937008
                                    0.063014
                                                       0.937008
1 1019.7 1019.7
                      0.937008
                                     0.063014
                                                0
                                                       0.937008
                                                                     0
   988.4
         988.4
                      0.937008
                                     0.063014
                                                       0.937008
                                                0
                                                                     0
3 1002.0 1002.0
                      0.937008
                                    0.063014
                                                       0.937008
                                                a
                                                                     a
4 1061.5 1061.5
                      0.936894
                                     0.063014
                                                       0.936894
Y length: 365
Y values: [0]
        valid 0's multi logloss: 4.44089e-15
 [200]
 [200]
        valid_0's multi_logloss: 4.44089e-15
        valid_0's multi_logloss: 4.44089e-15
valid_0's multi_logloss: 4.44089e-15
 [200]
 [200]
 [200]
        valid_0's multi_logloss: 4.44089e-15
        valid_0's multi_logloss: 4.44089e-15
valid_0's multi_logloss: 4.44089e-15
valid_0's multi_logloss: 4.44089e-15
valid_0's multi_logloss: 4.44089e-15
 [200]
 [200]
 [200]
 [200]
        valid 0's multi_logloss: 4.44089e-15
       valid_0's multi_logloss: 4.44089e-15
valid_0's multi_logloss: 4.44089e-15
 [200]
 [200]
Decision Tree Accuracy: 1.0
Classification Report for Decision Tree:
              precision
                         recall f1-score support
                  1.00
                            1.00
                                      1.00
                                                365
                                      1.00
                                                365
    accuracy
   macro avg
                  1.00
                            1.00
                                      1.00
                                                365
 weighted avg
                                                365
                   1.00
                            1.00
                                      1.00
```