```
!pip install pyspark
```

```
Collecting pyspark
  Downloading pyspark-3.2.0.tar.gz (281.3 MB)
     |████████████████████████████████| 281.3 MB 38 kB/s
Collecting py4j==0.10.9.2
  Downloading py4j-0.10.9.2-py2.py3-none-any.whl (198 kB)
     |████████████████████████████████| 198 kB 48.8 MB/s
Building wheels for collected packages: pyspark
  Building wheel for pyspark (setup.py) ... done
  Created wheel for pyspark: filename=pyspark-3.2.0-py2.py3-none-any.whl size=281805
  Stored in directory: /root/.cache/pip/wheels/0b/de/d2/9be5d59d7331c6c2a7c1b6d1a4f4
Successfully built pyspark
Installing collected packages: py4j, pyspark
Successfully installed py4j-0.10.9.2 pyspark-3.2.0
```

```
# Import SparkSession
from pyspark.sql import SparkSession

# create a object SparkSession
spark = SparkSession.builder.appName('Kmeans_App').getOrCreate()
```

```
#Load Data
data1 = spark.read.option('header','True').csv('/content/pubg.csv')

data2 = spark.read.option('header','True').csv('/content/Responses.csv')
```

```
a = SparkSession.builder
```

```
type(a)
```

```
pyspark.sql.session.SparkSession.Builder
```

```
data1.show()
```

```
+--------------+--------------+--------------+-------+------+-----------+-----+-----
|            Id|       groupId|       matchId|assists|boosts|damageDealt|DBNOs|heads
+--------------+--------------+--------------+-------+------+-----------+-----+-----
|2f262dd9795e60|78437bcd91d40e|d5db3a49eb2955|      0|     0|          0|    0|
|a32847cf5bf34b|85b7ce5a12e10b|65223f05c7fdb4|      0|     0|      163.2|    1|
|1b1900a9990396|edf80d6523380a|1cadec4534f30a|      0|     3|      278.7|    2|
|f589dd03b60bf2|804ab5e5585558|c4a5676dc91604|      0|     0|      191.9|    1|
|c23c4cc5b78b35|b3e2cd169ed920|cd595700a01bfa|      0|     0|        100|    1|
|fd034582dd4d2e|9b8930aeee086a|6f6e52b15ddf21|      0|     1|        200|    2|
|c60b5633f4dcc8|7c0f817f6627c7|3232c1e0fec04b|      0|     3|      638.2|    4|
|f0ba8246b6980f|7318b5204462cb|112e9711f86001|      0|     0|      27.94|    0|
|79c5d5eda1c72e|a85b81198dfc06|ef5fc25e28ffb1|      1|     4|      275.8|    3|
|94834a28e52abd|bc513cde35fa54|f36a754a9b88f7|      1|     1|      530.4|    4|
|f051dcc9b0b3ce|d203c0e3d8c321|89a6a8738190b4|      0|     0|      20.59|    0|
|f02c2f34accf08|22ed911205c815|559dac9580b92a|      1|     0|      62.72|    0|
|6701c06774d409|cdb79f944d585b|9f3decb5ffba4e|      0|     0|          0|    0|
|4e4aef4aeee5f5|a9dfa1c736c889|ac92da38bb19ad|      0|     2|      13.83|    0|
|d26b4b75c5229d|130ea20c924e8c|14fb1c1b26e9a4|      0|     0|       25.8|    0|
|c5473a410326a8|8a25860cd71a23|88cffe1ae97aff|      1|     1|        594|    2|
|321fe9f3c71131|cb3471586d99b4|1cf664f7c75122|      0|     0|      50.31|    0|
|f8933f3ee2e431|114b20e9d7504b|d5fcb7a3981d33|      0|     0|          0|    0|
|c70c7337cd46b4|73870d831717aa|e6602141e44281|      0|     0|       25.8|    0|
|d6c231133b5d57|928733f3037f92|b4baee11351ae6|      0|     0|      30.96|    0|
+--------------+--------------+--------------+-------+------+-----------+-----+-----
only showing top 20 rows
```

```
data2.show()
```

```
+---+------+--------------------+----------------------------------------+---------
|Age|Gender|Do you play PUBG game|How long have you been playing this game|How often
+---+------+--------------------+----------------------------------------+---------
| 21|Female|                 Yes|                        Less than 1 year|
| 24|  Male|                 Yes|                        Less than 1 year|
| 22|  Male|                 Yes|                             1 - 2 years|
| 23|  Male|                 Yes|                             1 - 2 years|
| 21|Female|                 Yes|                        Less than 1 year|
| 24|  Male|                 Yes|                        Less than 1 year|
| 22|  Male|                 Yes|                       More than 2 years|
| 22|Female|                 Yes|                        Less than 1 year|
| 24|  Male|                 Yes|                        Less than 1 year|
| 26|  Male|                 Yes|                        Less than 1 year|
| 18|Female|                 Yes|                        Less than 1 year|
| 18|  Male|                 Yes|                             1 - 2 years|
| 20|Female|                 Yes|                       More than 2 years|
| 19|  Male|                 Yes|                        Less than 1 year|
| 21|  Male|                 Yes|                       More than 2 years|
| 22|  Male|                 Yes|                             1 - 2 years|
| 23|  Male|                 Yes|                       More than 2 years|
| 25|  Male|                 Yes|                             1 - 2 years|
| 23|Female|                 Yes|                             1 - 2 years|
| 25|  Male|                 Yes|                       More than 2 years|
+---+------+--------------------+----------------------------------------+---------
only showing top 20 rows
```

```
data1.head(5)
```

```
[Row(Id='2f262dd9795e60', groupId='78437bcd91d40e', matchId='d5db3a49eb2955', assist
 Row(Id='a32847cf5bf34b', groupId='85b7ce5a12e10b', matchId='65223f05c7fdb4', assist
 Row(Id='1b1900a9990396', groupId='edf80d6523380a', matchId='1cadec4534f30a', assist
 Row(Id='f589dd03b60bf2', groupId='804ab5e5585558', matchId='c4a5676dc91604', assist
 Row(Id='c23c4cc5b78b35', groupId='b3e2cd169ed920', matchId='cd595700a01bfa', assist
```

```
data1.tail(5)
```

```
[Row(Id='ef4f474acd8e85', groupId='2eca2a8391f75d', matchId='492ecdfae90b46', assist
 Row(Id='cf0bf82fb4d80e', groupId='2eaf2765f93adb', matchId='14bffd71e96320', assist
 Row(Id='a0a31a0b1dcbe1', groupId='8d50c64ccc5071', matchId='147e4bbb62e3bb', assist
 Row(Id='f6874657399d69', groupId='d31843d7e62ccb', matchId='662567dcf280f5', assist
 Row(Id='90359b0b8f8b0d', groupId='61d5b1bb8da43f', matchId='258bfa48d88014', assist
data2.head(5)
```

```
[Row(Age='21', Gender='Female', Do you play PUBG game='Yes', How long have you been
 Row(Age='24', Gender='Male', Do you play PUBG game='Yes', How long have you been pl
 Row(Age='22', Gender='Male', Do you play PUBG game='Yes', How long have you been pl
 Row(Age='23', Gender='Male', Do you play PUBG game='Yes', How long have you been pl
 Row(Age='21', Gender='Female', Do you play PUBG game='Yes', How long have you been
```

```
data2.tail(5)
```

```
[Row(Age='25', Gender='Male', Do you play PUBG game='No', How long have you been pla
 Row(Age='24', Gender='Female', Do you play PUBG game='No', How long have you been p
 Row(Age='24', Gender='Female', Do you play PUBG game='No', How long have you been p
 Row(Age='22', Gender='Male', Do you play PUBG game='No', How long have you been pla
 Row(Age='23', Gender='Female', Do you play PUBG game='No', How long have you been p
```

```
data1.printSchema()
```

```
root
 |-- Id: string (nullable = true)
 |-- groupId: string (nullable = true)
 |-- matchId: string (nullable = true)
 |-- assists: string (nullable = true)
 |-- boosts: string (nullable = true)
 |-- damageDealt: string (nullable = true)
 |-- DBNOs: string (nullable = true)
 |-- headshotKills: string (nullable = true)
 |-- heals: string (nullable = true)
 |-- killPlace: string (nullable = true)
 |-- killPoints: string (nullable = true)
 |-- kills: string (nullable = true)
 |-- killStreaks: string (nullable = true)
 |-- longestKill: string (nullable = true)
 |-- matchDuration: string (nullable = true)
 |-- matchType: string (nullable = true)
 |-- maxPlace: string (nullable = true)
 |-- numGroups: string (nullable = true)
 |-- rankPoints: string (nullable = true)
 |-- revives: string (nullable = true)
 |-- rideDistance: string (nullable = true)
 |-- roadKills: string (nullable = true)
 |-- swimDistance: string (nullable = true)
 |-- teamKills: string (nullable = true)
 |-- vehicleDestroys: string (nullable = true)
 |-- walkDistance: string (nullable = true)
 |-- weaponsAcquired: string (nullable = true)
 |-- winPoints: string (nullable = true)
 |-- winPlacePerc: string (nullable = true)
```

```
data2.printSchema()
```

```
root
 |-- Age: string (nullable = true)
 |-- Gender: string (nullable = true)
 |-- Do you play PUBG game: string (nullable = true)
 |-- How long have you been playing this game: string (nullable = true)
 |-- How often do you play this Game: string (nullable = true)
 |-- How much time you spent daily: string (nullable = true)
 |-- How affect this game on students6: string (nullable = true)
 |-- Positive effects of playing PUBG: string (nullable = true)
 |-- Negative effects of playing PUBG: string (nullable = true)
 |-- What are reasons that you dont play PUBG: string (nullable = true)
 |-- How affect this game on students10: string (nullable = true)
 |-- According to you are there positive effects of playing PUBG: string (nullable =
```

```
data1 = data1.withColumn('Id',data1.Id.astype('string'))\
.withColumn('groupId',data1.groupId.astype('string'))\
.withColumn('matchId',data1.matchId.astype('string'))\
.withColumn('assists',data1.assists.astype('int'))\
.withColumn('boosts',data1.boosts.astype('int'))\
.withColumn('damageDealt',data1.damageDealt.astype('float'))\
.withColumn('DBNOs',data1.DBNOs.astype('int'))\
.withColumn('headshotKills',data1.headshotKills.astype('int'))\
.withColumn('heals',data1.heals.astype('int'))\
.withColumn('killPlace',data1.killPlace.astype('int'))\
.withColumn('killPoints',data1.killPoints.astype('int'))\
.withColumn('kills',data1.kills.astype('int'))\
.withColumn('killStreaks',data1.killStreaks.astype('int'))\
.withColumn('longestKill',data1.longestKill.astype('float'))\
.withColumn('matchDuration',data1.matchDuration.astype('int'))\
.withColumn('matchType',data1.matchType.astype('string'))\
.withColumn('maxPlace',data1.maxPlace.astype('int'))\
.withColumn('numGroups',data1.numGroups.astype('int'))\
.withColumn('rankPoints',data1.rankPoints.astype('int'))\
.withColumn('revives',data1.revives.astype('int'))\
.withColumn('rideDistance',data1.rideDistance.astype('float'))\
.withColumn('roadKills',data1.roadKills.astype('int'))\
.withColumn('swimDistance',data1.swimDistance.astype('int'))\
.withColumn('teamKills',data1.teamKills.astype('int'))\
.withColumn('vehicleDestroys',data1.vehicleDestroys.astype('int'))\
.withColumn('walkDistance',data1.walkDistance.astype('float'))\
.withColumn('weaponsAcquired',data1.weaponsAcquired.astype('int'))\
.withColumn('winPoints',data1.winPoints.astype('int'))\
.withColumn('winPlacePerc',data1.winPlacePerc.astype('float'))
```

```
data1.printSchema()
```

```
root
 |-- Id: string (nullable = true)
 |-- groupId: string (nullable = true)
 |-- matchId: string (nullable = true)
 |-- assists: integer (nullable = true)
 |-- boosts: integer (nullable = true)
 |-- damageDealt: float (nullable = true)
 |-- DBNOs: integer (nullable = true)
 |-- headshotKills: integer (nullable = true)
 |-- heals: integer (nullable = true)
 |-- killPlace: integer (nullable = true)
 |-- killPoints: integer (nullable = true)
 |-- kills: integer (nullable = true)
 |-- killStreaks: integer (nullable = true)
 |-- longestKill: float (nullable = true)
 |-- matchDuration: integer (nullable = true)
 |-- matchType: string (nullable = true)
 |-- maxPlace: integer (nullable = true)
 |-- numGroups: integer (nullable = true)
 |-- rankPoints: integer (nullable = true)
 |-- revives: integer (nullable = true)
 |-- rideDistance: float (nullable = true)
 |-- roadKills: integer (nullable = true)
 |-- swimDistance: integer (nullable = true)
 |-- teamKills: integer (nullable = true)
 |-- vehicleDestroys: integer (nullable = true)
 |-- walkDistance: float (nullable = true)
 |-- weaponsAcquired: integer (nullable = true)
 |-- winPoints: integer (nullable = true)
 |-- winPlacePerc: float (nullable = true)
```

```
data2 = data2.withColumn('Age',data2.Age.astype('int'))
```

```
data2.printSchema()
```

```
root
 |-- Age: integer (nullable = true)
 |-- Gender: string (nullable = true)
 |-- Do you play PUBG game: string (nullable = true)
 |-- How long have you been playing this game: string (nullable = true)
 |-- How often do you play this Game: string (nullable = true)
 |-- How much time you spent daily: string (nullable = true)
 |-- How affect this game on students6: string (nullable = true)
 |-- Positive effects of playing PUBG: string (nullable = true)
 |-- Negative effects of playing PUBG: string (nullable = true)
 |-- What are reasons that you dont play PUBG: string (nullable = true)
 |-- How affect this game on students10: string (nullable = true)
 |-- According to you are there positive effects of playing PUBG: string (nullable =
```

```
data1.select('Id','kills','killPoints').show()
```

```
+--------------+-----+----------+
|            Id|kills|killPoints|
+--------------+-----+----------+
|2f262dd9795e60|    0|      1126|
|a32847cf5bf34b|    1|      1309|
|1b1900a9990396|    2|         0|
|f589dd03b60bf2|    1|         0|
|c23c4cc5b78b35|    0|      1332|
|fd034582dd4d2e|    0|         0|
|c60b5633f4dcc8|    8|         0|
|f0ba8246b6980f|    0|         0|
|79c5d5eda1c72e|    4|         0|
|94834a28e52abd|    5|      1502|
|f051dcc9b0b3ce|    0|         0|
|f02c2f34accf08|    0|         0|
|6701c06774d409|    0|      1299|
|4e4aef4aeee5f5|    1|         0|
|d26b4b75c5229d|    0|      1267|
|c5473a410326a8|    2|         0|
|321fe9f3c71131|    0|         0|
|f8933f3ee2e431|    0|         0|
|c70c7337cd46b4|    0|      1183|
|d6c231133b5d57|    0|      1130|
+--------------+-----+----------+
only showing top 20 rows
```

```
data2.select('Age','Gender','How affect this game on students6').show()
```

```
+---+------+---------------------------------+
|Age|Gender|How affect this game on students6|
+---+------+---------------------------------+
| 21|Female|                         Positive|
| 24|  Male|                         Negative|
| 22|  Male|                         Negative|
| 23|  Male|                         Positive|
| 21|Female|                         Negative|
| 24|  Male|                         Positive|
| 22|  Male|                         Negative|
| 22|Female|                         Negative|
| 24|  Male|                         Negative|
| 26|  Male|                         Negative|
| 18|Female|                         Negative|
| 18|  Male|                         Positive|
| 20|Female|                         Positive|
| 19|  Male|                         Positive|
| 21|  Male|                         Positive|
| 22|  Male|                         Positive|
| 23|  Male|                         Positive|
| 25|  Male|                         Negative|
| 23|Female|                         Positive|
| 25|  Male|                         Positive|
+---+------+---------------------------------+
only showing top 20 rows
```

```
from pyspark.ml.feature import VectorAssembler
```

```
assembeler = VectorAssembler(
    inputCols = ['kills','headshotKills'],
    outputCol = 'features'
)
```

```
data1.printSchema()
```

```
root
 |-- Id: string (nullable = true)
 |-- groupId: string (nullable = true)
 |-- matchId: string (nullable = true)
 |-- assists: integer (nullable = true)
 |-- boosts: integer (nullable = true)
 |-- damageDealt: float (nullable = true)
 |-- DBNOs: integer (nullable = true)
 |-- headshotKills: integer (nullable = true)
 |-- heals: integer (nullable = true)
 |-- killPlace: integer (nullable = true)
 |-- killPoints: integer (nullable = true)
 |-- kills: integer (nullable = true)
 |-- killStreaks: integer (nullable = true)
 |-- longestKill: float (nullable = true)
 |-- matchDuration: integer (nullable = true)
 |-- matchType: string (nullable = true)
 |-- maxPlace: integer (nullable = true)
 |-- numGroups: integer (nullable = true)
 |-- rankPoints: integer (nullable = true)
 |-- revives: integer (nullable = true)
 |-- rideDistance: float (nullable = true)
 |-- roadKills: integer (nullable = true)
 |-- swimDistance: integer (nullable = true)
 |-- teamKills: integer (nullable = true)
 |-- vehicleDestroys: integer (nullable = true)
 |-- walkDistance: float (nullable = true)
 |-- weaponsAcquired: integer (nullable = true)
 |-- winPoints: integer (nullable = true)
 |-- winPlacePerc: float (nullable = true)
```

```
data1 = assembeler.transform(data1)
```

```
data2.groupby('What are reasons that you dont play PUBG').count().show()
```

```
+----------------------------------------+-----+
|What are reasons that you dont play PUBG|count|
+----------------------------------------+-----+
|                    Addictive, Poor p...|    1|
|                    Poor physical hea...|    1|
|                    Violent tendency,...|    2|
|                    Poor physical hea...|    1|
|                    Anti-social, Poor...|    1|
|                    Violent tendency,...|    1|
|                    Addictive, Poor p...|    1|
|                    Violent tendency,...|    1|
|                    Addictive, Anti-s...|    1|
|                    Violent tendency,...|    1|
|                      Don't like to play|    5|
|                    Poor physical hea...|    1|
|                    Poor physical hea...|    1|
|                    Addictive, Poor p...|    1|
|                    Violent tendency,...|    1|
|                     Addictive, Anti-s...|    1|
|                     Affects mental he...|    1|
|                     Affects mental he...|    1|
|                     Addictive, Anti-s...|    1|
|                     Violent tendency,...|    1|
+-----------------------------------------+-----+
only showing top 20 rows
```

```
data2.groupby('Age','Gender').count().show()
```

```
+---+------+-----+
|Age|Gender|count|
+---+------+-----+
| 19|Female|    1|
| 26|  Male|    1|
| 25|  Male|    3|
| 24|  Male|    5|
| 23|Female|    7|
| 20|Female|    1|
| 27|  Male|    1|
| 20|  Male|    1|
| 23|  Male|    7|
| 21|Female|    3|
| 22|Female|    5|
| 19|  Male|    3|
| 21|  Male|    2|
| 24|Female|    2|
| 18|  Male|    2|
| 26|Female|    1|
| 18|Female|    1|
| 22|  Male|    5|
| 27|Female|    2|
+---+------+-----+
```

```
data1.show()
```

```
+--------------+--------------+--------------+-------+------+-----------+-----+-----
|            Id|       groupId|       matchId|assists|boosts|damageDealt|DBNOs|heads
+--------------+--------------+--------------+-------+------+-----------+-----+-----
|2f262dd9795e60|78437bcd91d40e|d5db3a49eb2955|      0|     0|        0.0|    0|
|a32847cf5bf34b|85b7ce5a12e10b|65223f05c7fdb4|      0|     0|      163.2|    1|
|1b1900a9990396|edf80d6523380a|1cadec4534f30a|      0|     3|      278.7|    2|
|f589dd03b60bf2|804ab5e5585558|c4a5676dc91604|      0|     0|      191.9|    1|
|c23c4cc5b78b35|b3e2cd169ed920|cd595700a01bfa|      0|     0|      100.0|    1|
|fd034582dd4d2e|9b8930aeee086a|6f6e52b15ddf21|      0|     1|      200.0|    2|
|c60b5633f4dcc8|7c0f817f6627c7|3232c1e0fec04b|      0|     3|      638.2|    4|
|f0ba8246b6980f|7318b5204462cb|112e9711f86001|      0|     0|      27.94|    0|
|79c5d5eda1c72e|a85b81198dfc06|ef5fc25e28ffb1|      1|     4|      275.8|    3|
|94834a28e52abd|bc513cde35fa54|f36a754a9b88f7|      1|     1|      530.4|    4|
|f051dcc9b0b3ce|d203c0e3d8c321|89a6a8738190b4|      0|     0|      20.59|    0|
|f02c2f34accf08|22ed911205c815|559dac9580b92a|      1|     0|      62.72|    0|
|6701c06774d409|cdb79f944d585b|9f3decb5ffba4e|      0|     0|        0.0|    0|
|4e4aef4aeee5f5|a9dfa1c736c889|ac92da38bb19ad|      0|     2|      13.83|    0|
|d26b4b75c5229d|130ea20c924e8c|14fb1c1b26e9a4|      0|     0|       25.8|    0|
|c5473a410326a8|8a25860cd71a23|88cffe1ae97aff|      1|     1|      594.0|    2|
|321fe9f3c71131|cb3471586d99b4|1cf664f7c75122|      0|     0|      50.31|    0|
|f8933f3ee2e431|114b20e9d7504b|d5fcb7a3981d33|      0|     0|        0.0|    0|
|c70c7337cd46b4|73870d831717aa|e6602141e44281|      0|     0|       25.8|    0|
|d6c231133b5d57|928733f3037f92|b4baee11351ae6|      0|     0|      30.96|    0|
+--------------+--------------+--------------+-------+------+-----------+-----+-----
only showing top 20 rows
```

```
data1.select('features').show()
```

```
+---------+
| features|
+---------+
|(2,[],[])|
|[1.0,1.0]|
|[2.0,1.0]|
|[1.0,0.0]|
|(2,[],[])|
|(2,[],[])|
|[8.0,1.0]|
|(2,[],[])|
|[4.0,0.0]|
|[5.0,0.0]|
|(2,[],[])|
|(2,[],[])|
|(2,[],[])|
|[1.0,0.0]|
|(2,[],[])|
|[2.0,1.0]|
|(2,[],[])|
|(2,[],[])|
|(2,[],[])|
|(2,[],[])|
+---------+
only showing top 20 rows
```

```
from pyspark.ml.clustering import KMeans
from pyspark.ml.evaluation import ClusteringEvaluator
```

```
kmeans = KMeans().setK(3).setSeed(1)
```

```
model = kmeans.fit(data1)
```

```
pred = model.transform(data1)
```

```
pred.show()
```

```
+--------------+--------------+--------------+-------+------+-----------+-----+-----
|            Id|       groupId|       matchId|assists|boosts|damageDealt|DBNOs|heads
+--------------+--------------+--------------+-------+------+-----------+-----+-----
|2f262dd9795e60|78437bcd91d40e|d5db3a49eb2955|      0|     0|        0.0|    0|
|a32847cf5bf34b|85b7ce5a12e10b|65223f05c7fdb4|      0|     0|      163.2|    1|
|1b1900a9990396|edf80d6523380a|1cadec4534f30a|      0|     3|      278.7|    2|
|f589dd03b60bf2|804ab5e5585558|c4a5676dc91604|      0|     0|      191.9|    1|
|c23c4cc5b78b35|b3e2cd169ed920|cd595700a01bfa|      0|     0|      100.0|    1|
|fd034582dd4d2e|9b8930aeee086a|6f6e52b15ddf21|      0|     1|      200.0|    2|
|c60b5633f4dcc8|7c0f817f6627c7|3232c1e0fec04b|      0|     3|      638.2|    4|
|f0ba8246b6980f|7318b5204462cb|112e9711f86001|      0|     0|      27.94|    0|
|79c5d5eda1c72e|a85b81198dfc06|ef5fc25e28ffb1|      1|     4|      275.8|    3|
|94834a28e52abd|bc513cde35fa54|f36a754a9b88f7|      1|     1|      530.4|    4|
|f051dcc9b0b3ce|d203c0e3d8c321|89a6a8738190b4|      0|     0|      20.59|    0|
|f02c2f34accf08|22ed911205c815|559dac9580b92a|      1|     0|      62.72|    0|
|6701c06774d409|cdb79f944d585b|9f3decb5ffba4e|      0|     0|        0.0|    0|
|4e4aef4aeee5f5|a9dfa1c736c889|ac92da38bb19ad|      0|     2|      13.83|    0|
|d26b4b75c5229d|130ea20c924e8c|14fb1c1b26e9a4|      0|     0|       25.8|    0|
|c5473a410326a8|8a25860cd71a23|88cffe1ae97aff|      1|     1|      594.0|    2|
|321fe9f3c71131|cb3471586d99b4|1cf664f7c75122|      0|     0|      50.31|    0|
|f8933f3ee2e431|114b20e9d7504b|d5fcb7a3981d33|      0|     0|        0.0|    0|
|c70c7337cd46b4|73870d831717aa|e6602141e44281|      0|     0|       25.8|    0|
|d6c231133b5d57|928733f3037f92|b4baee11351ae6|      0|     0|      30.96|    0|
+--------------+--------------+--------------+-------+------+-----------+-----+-----
only showing top 20 rows
```

```
evaluator = ClusteringEvaluator()
```

```
res = evaluator.evaluate(pred)
```

```
res
```

```
0.7617710366356096
```

✓ 0s completed at 10:22

● ✕