# AppKluster

**Author**: Paul Hutchinson

This software clusters *n* applications into *k* clusters in which each application belongs to the cluster with the nearest mean, using Euclidean distance as the distance measure. The algorithm used is *k-means clustering* in which the difference index, standard deviation, application set, number of centroids, and number of results are all specified by the user.

**How to Run**

There are two methods of running this code:

1. Load the project into Eclipse and run the Main.java class as a java application (located in AppKluster/Source/AppKluster/src/Main.java).
2. Run the AppKluster.jar file located in (AppKluster/Runnable/AppKluster.jar).

**How to Use**

Users can choose to enter their own data or use the default data sets that are provided. It is recommended that users load the default inputs the first time to understand how AppKluster works. To do this, follow the following steps:

1. Launch AppKluster
2. Press the “Use Default Data/Inputs…” button.
3. Press the “GENERATE” button.

Alternatively, AppKluster allows the user to input their personal data for clustering. To do this, follow the following steps:

1. Launch AppKluster.
2. Enter a difference index into the “Difference Index” field.
3. Enter a standard deviation into the “Standard Deviation” field.
4. Enter a number of centroids to be used in the “Number Of Centroids” field.
5. Enter a number of applications to pick in the “Number To Pick” field.
6. Load a set of applications to be clustered by pressing the “Browse…” button and choosing the file folder where the applications are located. Proper input format for applications can be seen below.
7. Press the “GENERATE” button.

**Application Format**

Applications should be placed in their own file folder. Each should be a .txt file with the following format:

Name:<full name>

Race:<race>

Gender:<male or female>

Age:<age>

GPA:<grade point average>

ID:<student id>

**External Libraries**

AppKluster uses several external libraries for accomplishing the clustering and interface layout.

* Kmeanscluster.jar (<http://code.google.com/p/kmeansclustering/>)
* Stromberglabs-commons.jar (<http://code.google.com/p/kmeansclustering/>)
* Miglayout-3.7.1.jar (<http://www.miglayout.com/>)