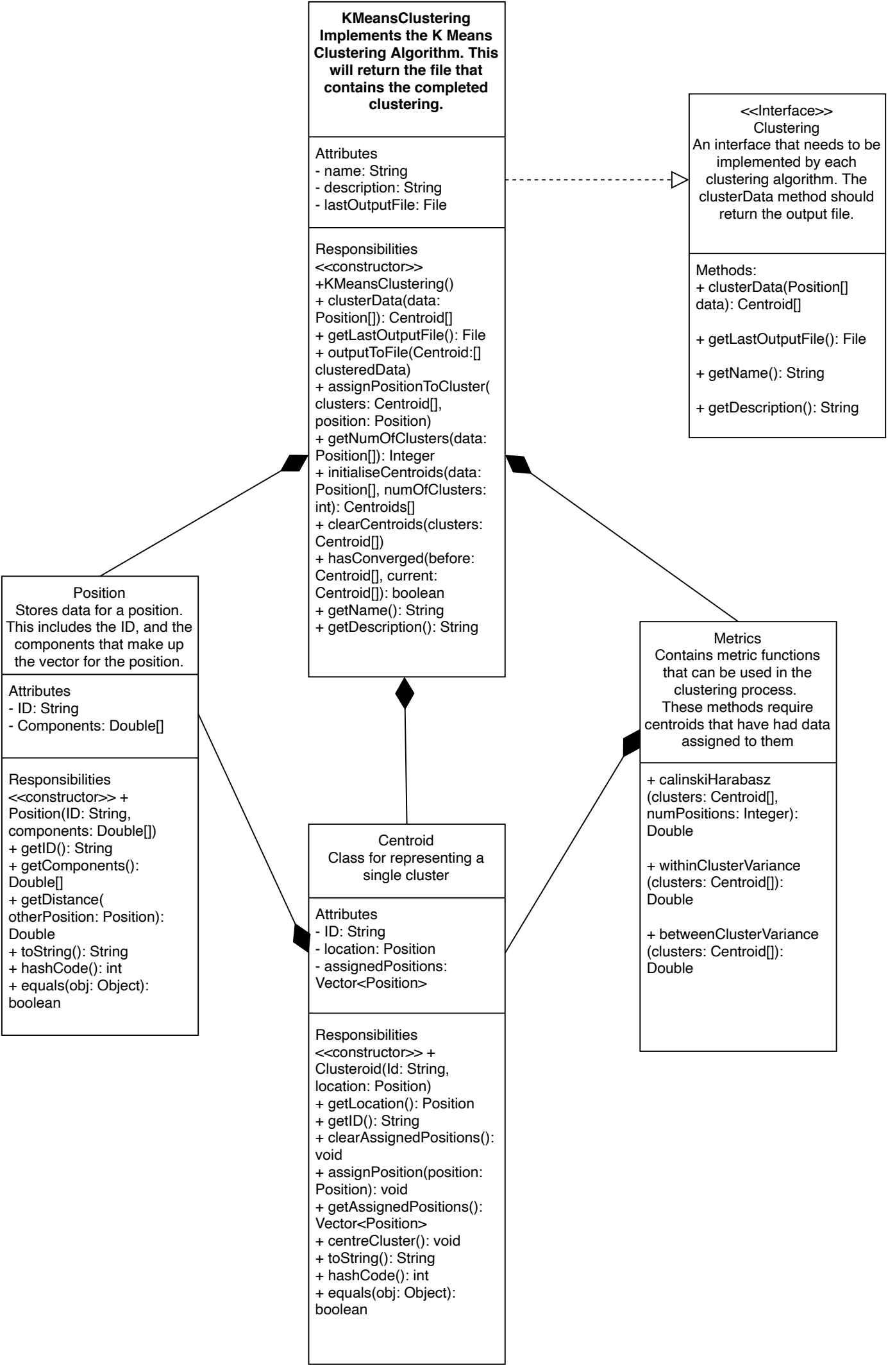
KMeansClustering
Implements the K Means Clustering Algorithm. This will return the file that contains the completed clustering.
Attributes
- name: String
- description: String
- lastOutputFile: File
Responsibilities
<<constructor>>
+KMeansClustering()
+ clusterData(data: Position[]): Centroid[]
+ getLastOutputFile(): File
+ outputToFile(Centroid:[] clusteredData)
+ assignPositionToCluster(clusters: Centroid[], position: Position)
+ getNumOfClusters(data: Position[]): Integer
+ initialiseCentroids(data: Position[], numOfClusters: int): Centroids[]
+ clearCentroids(clusters: Centroid[])
+ hasConverged(before: Centroid[], current: Centroid[]): boolean
+ getName(): String
+ getDescription(): String
<<Interface>>
Clustering
An interface that needs to be implemented by each clustering algorithm. The clusterData method should return the output file.
Methods:
+ clusterData(Position[] data): Centroid[]
+ getLastOutputFile(): File
+ getName(): String
+ getDescription(): String
Position
Stores data for a position. This includes the ID, and the components that make up the vector for the position.
Attributes
- ID: String
- Components: Double[]
Responsibilities
<<constructor>> + Position(ID: String, components: Double[])
+ getID(): String
+ getComponents(): Double[]
+ getDistance(otherPosition: Position): Double
+ toString(): String
+ hashCode(): int
+ equals(obj: Object): boolean
Centroid
Class for representing a single cluster
Attributes
- ID: String
- location: Position
- assignedPositions: Vector<Position>
Responsibilities
<<constructor>> + Clusteroid(Id: String, location: Position)
+ getLocation(): Position
+ getID(): String
+ clearAssignedPositions(): void
+ assignPosition(position: Position): void
+ getAssignedPositions(): Vector<Position>
+ centreCluster(): void
+ toString(): String
+ hashCode(): int
+ equals(obj: Object): boolean
Metrics
Contains metric functions that can be used in the clustering process. These methods require centroids that have had data assigned to them
+ calinskiHarabasz (clusters: Centroid[], numPositions: Integer): Double
+ withinClusterVariance (clusters: Centroid[]): Double
+ betweenClusterVariance (clusters: Centroid[]): Double