Coders Inc Design Decisions – Milestone 4

# Individual Feature Distance Metrics

* Functionality to choose how the distance for individual features was added
  + Separate set of metrics for numerical and string values for simplification of code
* Three metrics were chosen for each type of feature in order to provide good variety while keeping it to a more manageable number
  + Numerical
    - Absolute difference (default)
    - Equal or not
    - Within the same number of standard deviations (emphasises outliers if the target point is an outlier)
  + String
    - Hamming Distance (default)
    - Equal or not
    - Difference in sum of character ASCII values

# KNN Refactoring

* Based on fact that Minkowski distance is a generalization of Euclidean and Manhattan distances
  + Euclidean and Manhattan kNN classes are now subclasses of Minkowski distances providing a specific order for the expression (2 for Euclidean and 1 for Manhattan) and relying on the Minkowski methods
  + This reduces code duplication while minimizing impact of refactoring

# Serialized DataModel and Importing and Exporting

* To easily import and export any data set within the program, all classes within the DataModel package implements the Serializable interface.
  + This allows the entire DimensionalSpace and all data it holds to be sent to an Object output stream to be easily saved in a file.
  + SerialImport and SerialExport were put into separate classes so as to not give DimensionalSpace too many responsibilities. These classes also handle the exportation and importation of Point and Cell objects as well.