Million Song Database Analysis

**By the *Dancing Dads* aka: John Butcher, Austin Poulton, Ian McNair and Barrington Trim**

Submitted for: Big Data Processing (ECS765P) Coursework Assignment 2

18 December 2015

# Problem Statement

# Approach

# Implementation Approach

The project source code repository <https://github.research.its.qmul.ac.uk/ec15541/million-song>

* Discussion of Track encapsulation and companion object
* Discussion of Model trait and classification/regression models
* Pattern for Spark driver objects

# Analysis of Results

# References

Apache Foundation (2015) *Spark Programming Guide*.[online] Available at: <http://spark.apache.org/docs/latest/programming-guide.html> [Accessed 5 December 2015].

ScalaNLP (2015) *Breeze Quick start* Numeric Processing for Scala [online] Avaialble at <https://github.com/scalanlp/breeze/wiki/Quickstart> [Accessed 10 December 2015].

Murphy, K (2012) *Machine Learning – A Probabilistic Perspective* MIT Press, Cambridge, Mass. USA