This syllabus is provisional, and may be adapted as the quarter develops:

* Introduction to data mining (Data\_Types.pptx)
  + Converting, Normalizing and Cleaning Data
* Data types. (Data\_Types.pptx)
* Classification:
  + The simple linear classifier. (classifcation1\_Extended.ppt)
  + The nearest neighbor classifier. (classifcation1\_Extended.ppt)
  + The Decision Tree. (classifcation2.ppt)
  + The Bayesian Classifier. (classifcation2.ppt)
  + Ensemble learning
* Clustering
  + Partitional
  + Hierarchal
* Similarity Search (Similarity\_Search.ppt)
  + Distance measures
    - One-to-one measures
    - Edit/warping measures
    - Compression measures
  + Fast search (Indexing\_new.ppt)
    - Optimizations
    - Early abandoning search
    - Lower bounding search
    - Indexing
      * R-trees
      * The GEMINI framework
  + Finding Approximately Repeated Data (Finding Approximately Repeated Data.ppt)
    - In time series
    - In sets (mostly text)
* Association Pattern Mining (A\_rules.pptx)
* Outlier Analysis
* Mining Data Streams
* Mining Text Data
* Reviewing and writing papers on data mining
* Data Reduction and Feature Extraction
* Mining Graph Data / Social Network Analysis
* Advanced and late breaking topics.