Project presentation outline

- Problem statement – 1 slide (Riky)

High level solution:
- Data Warehouse Tools (Paul)
    - Top level diagram
    - Hadoop – Main file system
    - Hive – query language
    - Hue – Web Interface to Hadoop
    - NiFi – maybe

Details on the tools:
- Setup/Infrestucture (Stuart + Riky + Paul)
    - Top level diagram
    - EMR
    - S3 bucket
    - Cost
    - supporting tools
        - FoxyProxy
        - OpenSSH (ssh key-gen -rsa …) (Riky)
        - PuTTy (Riky)
        - Lunix + AWS CLI

- Data collection – 2-4 slides (Paul/Stuart)
    - Financial
    - Twitter
    - Note features, structure, **some great visual!**

- Database Schema (All)
    - Need to make a proposal

- Current Status (Riky)
    - Nice table of milestone (checkmarks, etc.)
    - Milestone/checkpoints
        - Infre Setup
        - Data collection
        - Integrating sample data into Hadoop
        - Build the database
        - Integrate all tools Hadoop + Hive + NiFi + Hue
        - Processing the data to solve problem statement
        - Final Paper + Final Presentation