IST769 Unit J - Search

## Agenda

1. Your Questions
2. Go over Problem Set
3. Unit Coursework Activities

## FRONT MATTER

Start your engines!

PS> **docker-compose down**

PS> **docker-compose up -d elasticsearch kibana jupyter drill**

Kibana: <http://localhost:5601>

Elasticsearch REST Endpoint: <http://localhost:9200>

**Elasticsearch:**

* What is the storage model? (relational, document, etc…)
  + Inverted Index
* What are the logical model metaphors? (tables, keys constraints, etc…)
  + Index => Documents
* How does it scale horizontally? CAP? (after all - its a big data course!)
  + Eventual Consistency
* Purpose of database? OLTP (CRUD) or OLAP/Analytics (CR)?
  + CR / Analytics
* How do you design for it?
  + Document Grain , Add to the document.
  + All stuff in document is the same grain
* How to connect to it from:
  + Client / Native Scripting Language
    - HTTP / REST
  + Drill / SQL
  + Spark / PySpark
  + DATA SINK.\

## 1. Your Questions

Ask any questions you have here!

* The only way to integrate with ElasticSearch is through its HTTPS REST API?
* Can maps include more than one index?
* Canvas visualization are display only?
* Dashboard can be queried with SQL?

## 2. Coursework Activities

Turn on the Tweet simulator in **J-Search.ipynb**

Kibana > Management > Stack Management

Data Management > Index Management

Kibana > Index Patterns

**Questions:**

1. Search for tweets from a specific user.
2. Search for tweets that mention “support”
3. Create a bar chart of tweet counts by user
4. Place Tweets on a map!