Homework IST769 Unit B - NoSQL Foundations

## Agenda

1. Your Questions
2. Go Over Problem Set

## 1. Your Questions

Ask any questions you have here!

* docker start v/s docker up and when to use them
* Difference between Images, Containers and Volume
* docker compose up fails to run on Mac?
* Explanation on Q7
* Major difference between Volume and Container
* Differences while using docker and docker-compose commands

## 2. Go Over Problem Set B

1. Take a screenshot of an open Windows Terminal with the current working directory set to the **advanced-databases** folder. Make sure to follow the screenshot guidelines and include your name/netid in the screenshot.

2. Explain the difference between a Docker image and a container. No screen shot necessary.

3. What is the purpose of a docker volume? No Screen shot necessary.

4. Start the redis and retwis services with docker-compose. Provide a screenshot of the command you typed, and another screenshot showing the two services are running.

5. What does the last line in the redis logs say? Provide a screenshot of this message. What command did you have to type to see the redis logs?

6. Which ports are being used by the currently running docker services? How do know they are running, and which ports are being used?

7. What are the volumes created by all the services in the docker-compose file? What command did you type to get this answer and provide a screenshot of the output.

Spark code:

import pyspark

from pyspark.sql import SparkSession

# Spark init

spark = SparkSession.builder \

.master("local") \

.appName('jupyter-pyspark') \

.getOrCreate()

sc = spark.sparkContext

sc.setLogLevel("ERROR")

df.createOrReplaceTempView("customers")

query = """

select First, Last, Email from customers where Gender = 'M'

"""

spark.sql(query).show()