CMSC 433

Nov 29th
Demand driven lectures

- The rest of the semester, the lectures will be substantially demand driven
- I’ll have some material to present, but I also want to leave plenty of time for you to ask questions either in a group or one-on-one format.
Homework 5

- Homework 5: Write a Hadoop program to determine the frequency with which words occur together on a line.

- e.g., the word *Sherlock* occurs on a line 101 times; on 96 of those lines the word *Holmes* also occurs, and on 1 of those lines the word *Watson* also occurs.
Homework 5

- Not tested through the submit server
- We’ll just give you some sample queries with answers, and some other queries you have to supply
- Submit your answers and your code
Project 4

- Documentation posted
Sending objects in project 4

- Can read/write objects through a ObjectInputStream/ObjectOutputStream
- Be sure to call flush to ensure your data actually gets shipped over the network
Customizing Object serialization

• Any fields declared transient are not sent via object serialization
  • are null/0/false when deserialized
• You may wish to set them, and do other things to objects after they are deserialized but before they are used.
Defining a custom `readObject` method

- Define a method

```java
private void readObject(java.io.ObjectInputStream in)
    throws IOException, ClassNotFoundException {
    in.defaultReadObject();
    // do any customization or setting of transient fields
}
```