



[ ]

[CS522DE-2024-07A-10A-01](#)[Tests & Quizzes](#) [View System Alert](#)

- 
- 
- 
- There are new notifications available

[\*\*Edit Profile Image\*\*](#) No file chosen

## **View All Sites**



15

- [Sites](#)
- [Organize Pinned](#)

Filter sites

## 2024-10C-01B

- [CS57X-2024-FallB](#)  
    ↳

## 2024-07A-10A

- [CS522DE-2024-07A-10A-01](#)  
    ↳

## 2024-07A-09C

- [CS57x-2024-FallA](#)  
    ↳

## 2024-04A-06D

- [CS57x-2024-SpringB](#)  
    ↳

## 2024-02A-05D

- [CS488DE-2024-02A-05D](#)  
    ↳

## 2024-01A-03D

- [CS57x-2024-SpringA](#)  
    ↳

## 2023-10C-01B

- [CS516DE-2023-10C-01B](#)  
    ↳

## 2023-10A-10D

- [CS-AI Seminar-2023-10 ...](#)  
    ↳

## 2023-07A-10A

- [CS590DE-2023-07A-10A](#)  
    ↳

## 2023-07A-07D

- [MGT5341-2023-07A-07C-01](#)  
    ↳

## 2023-06A-06D

- [CS525-2023-06A-06D-01](#)  
    ↳

- [MGT 5342 June 26 2023](#)



## 2023-05A-05D

- [CS435-2023-05A-05D-01](#)



## 2023-04A-04D

- [CS473-2023-04A-04D-01](#)



## 2023-03A-03D

- [CS477-2023-03A-03D-01](#)



## 2023-02A-02B

- [STC506B-2023-02A-02B-01](#)



## 2023-01A-01D

- [CS401-2023-01A-01D-03](#)



## 2022-12A-12D

- [CS390-2022-12A-12D-02](#)



## 2022-11A-11D

- [MPP Nov 2022 Entry](#)



## 2019-07A-10A

- [CS Techtest MGT5341](#)



## 2022-11C-11D

- [STC506A-2022-11C-11D-01](#)



## OTHER

- [Home](#)

Click and drag the icon, or use your keyboard to select a site and the 'e' or 'd' key to reorder your pinned sites.

ⓘ You do not have any pinned sites. Pin sites in the "Sites" tab to add them to your pinned list.

- ≡  
CS-AI Seminar-2023-10 ...
- ≡  
CS590DE-2023-07A-10A

## Sites



## Sites

- > [Home](#)
  - ≡ [Overview](#)
  - [Profile](#)
  - [Calendar](#)
  - [Resources](#)
  - [Announcements](#)
  - [Preferences](#)
  - [Account](#)

## Pinned

- > [CS-AI Seminar-2023-10A-10D](#)



- ≡ [Overview](#)
- [Syllabus](#)
- [Lessons](#)
- [Calendar](#)
- [Announcements](#)
- [Resources](#)
- [Assignments](#)
- [Tests & Quizzes](#)
- [Site Info](#)
- [Email](#)
- [Evaluation System](#)
- [Roster](#)

- > [CS590DE-2023-07A-10A](#)



- ≡ [Overview](#)
- [Syllabus](#)
- [Lessons](#)
- [Calendar](#)
- [Announcements](#)
- [Discussions](#)
- [Assignments](#)
- [Tests & Quizzes](#)
- [Evaluation System](#)

- [!\[\]\(2824aab9645d9fab95bae27ff6828dab\_img.jpg\) Gradebook](#)
- [!\[\]\(0fbf3ad74a6c8dc44ba9ea17fc2aca5e\_img.jpg\) Proctor360- Final Makeup Exam](#)

No pinned sites yet

## Recent

- > [CS57X-2024-FallB](#)
  - 
  - [!\[\]\(8c4dca64662d21542001ca0ed7eeb688\_img.jpg\) Overview](#)
  - [!\[\]\(3de35c640e7147a3fb61ee393128d2ae\_img.jpg\) Syllabus](#)
  - [!\[\]\(d1438aeefda19c86ae7477bf1fb30796\_img.jpg\) Calendar](#)
  - [!\[\]\(dc4d2c544087998b6f093f485f5119d7\_img.jpg\) Announcements](#)
  - [!\[\]\(f26ab61dd00ea7e5f19553908ec3fa6b\_img.jpg\) Resources](#)
  - [!\[\]\(ecf85e064bcd351a8999f71bf3f405e7\_img.jpg\) Discussions](#)
  - [!\[\]\(b7205a03f3dd3fe6f731085a326bb0d6\_img.jpg\) Assignments](#)
  - [!\[\]\(fea63d7e313047478d89177e38e2c72d\_img.jpg\) Gradebook](#)
  - [!\[\]\(3098c5e72378a1b0482fcc457c248f36\_img.jpg\) Roster](#)
- ✓ [CS522DE-2024-07A-10A-01](#)
  - 
  - [!\[\]\(bdd076fceb1bd6f84cb17f8d6a04e32c\_img.jpg\) Overview](#)
  - [!\[\]\(df648e0ad6c7cc181c595a3b5826b4ad\_img.jpg\) Syllabus](#)
  - [!\[\]\(880be622b9312c3384b9a4cbb8f9cbc6\_img.jpg\) Lessons](#)
  - [!\[\]\(e6398acab1fe106e24dc02668e12cda8\_img.jpg\) Calendar](#)
  - [!\[\]\(6a3b43ae489ad1dcd297f9ad18207726\_img.jpg\) Announcements](#)
  - [!\[\]\(9b8a9ba24565936df26cc20ac543b607\_img.jpg\) Resources](#)
  - [!\[\]\(8de32407a9832a2555fc0f815caf21f4\_img.jpg\) Discussions](#)
  - [!\[\]\(cb0818baa2223f811ba631b5cdc3b2e5\_img.jpg\) Assignments](#)
  - [!\[\]\(115410a3686c6bfcbd312a1b9781b823\_img.jpg\) Tests & Quizzes](#)
  - [!\[\]\(5112d96b163ef8d66015fa29116fd4e8\_img.jpg\) Gradebook](#)
  - [!\[\]\(aee93707ae98a6c5f7ca9e7b45c19929\_img.jpg\) Site Info](#)
  - [!\[\]\(68eb328c8e1a562366fc121effce1290\_img.jpg\) Roster](#)
  - [!\[\]\(b4f86d620bcf70d2c80865aa7b0cbc3c\_img.jpg\) Evaluation System](#)
  - [!\[\]\(e7e5ef6e716dcffae84457b3e07ebc88\_img.jpg\) Proctortrack](#)
  - [!\[\]\(abaa68a62cc87b2cd3c26981a7f9d1a0\_img.jpg\) Polls](#)
  - [!\[\]\(5d9bdb5059fae1818a1de33c465530e1\_img.jpg\) P360 Final](#)

 View all my sites

1. [CS522DE-2024-07A-10A-01](#)
2. [Tests & Quizzes](#)

## Tests & Quizzes

## Midterm

[Return to Assessment List](#)

## Part 1 of 4 - SCI 3.0 / 10.0 Points

These points apply for the whole test, please read them carefully and make sure you comply with each one:

- 1- The allotted exam duration is 2 hours.
- 2- Closed book/notes.
- 3- No personal items including electronic devices (cell phones, computers, calculators, PDAs, watches).
- 4- Cell phones and watches must not be on the desk before beginning, and during, the exam.
- 5- Exams are copyrighted and may not be copied or transferred.
- 6- Restroom and other personal breaks are not permitted.
- 7- No scratch papers allowed.
- 8- You may use the calculator application on your system.
- 9- You must perform a room scan before your exam, if you fail to do so your exam is void (your score for the test will be zero).
- 10- You should have two camera feeds during your test, if one of them stops your exam will be void and your score will be zero on the test.

Question 1 of 8

3.0

10.0 Points

Answer three out of four.

Below you are given STC principles or Big Data topics or techniques. Fill in the blank for each section, where you either provide the missing STC point or Big-Data point. If you provide a point that is not from the lecture you must provide a reasonable connection to get the partial credit. If you use a point from class or main points you do not have to provide the connection.

Big Data Point: Unstructured data provide a source for the extraction of knowledge, insight or information.

STC point: All information in nature is ultimately in the Unified Field.

Connection: Just like unstructured data serves as a source for the extraction of insight, knowledge, or information, so does the Unified field serve as the ultimate source of all information in nature

Big Data Point: MapReduce paradigm is used to extract valuable information from big data.

STC point: The Transcendental Medication technique is the tool for extracting information from the ultimate source of all information or knowledge, the Unified Field.

Connection: Just as a MapReduce program allows us extract valuable information from big data in an efficient and effective manner, the TM technique allows nature to extract valuable information from the source of all knowledge, effectively and efficiently.

Big Data Point: The in-mapper combining technique allows us to efficiently transfer extracted data over the network.

STC point: Nature is capable of harmoniously organizing the entire universe from an unmanifest level.

Connection: Just as the in-mapper combining technique allows us to organize an extracted data to be sent over the network efficiently, the Uniform Field provides the advantage of organizing for an efficient and effective throughput.

**Big Data Point:** The reducer is applied to all values associated with the same intermediate key to generate output key-value pairs. Implicit between the map and reduce phases is a distributed "group by" operation on intermediate keys. Intermediate data arrive at each reducer in order, sorted by the key. However, no ordering relationship is guaranteed for keys across different reducers.

**STC point:** Nature is capable of harmoniously organizing the entire universe from an unmanifest level.

**Connection:** The functionality of shuffle and sort, in conjunction of the function of the reducer, is to aid easier and efficient computation of program output, so the Unified Field serves as an infinite support to the operations of the entire body.

**Comment:** You mixed up all your points

## Part 2 of 4 - True False Questions 12.0 / 15.0 Points

**Answer each of the following questions with True or False. You must provide a rationale for your answer to be considered. Stating only true or false will result in zero credit. If the answer is False state why. If the answer is true, and you find it hard to state why you could state how.**

Question 2 of 8

4.0

5.0 Points

One of the four Vs in Big Data is Velocity. Velocity in Big Data means the bigger the data we need to work with the faster each machine in our cluster must be.

True

False

Rationale:

Velocity in Big Data means big data comes in at an extremely fast speed.

**Comment:** Okay, but what about processing speed?

Question 3 of 8

5.0

5.0 Points

In a map-reduce program you will always have multiple mappers and a single reducer.

True

False

Rationale:

The number of mappers is determined or equal to the number of input-split but the number of reducer is specified in the job configuration, before the start of the map reduce program.

**Comment:** Good

Question 4 of 8

3.0

5.0 Points

There are no disadvantages to using an in-mapper combiner.

True

False

Rationale:

Even though the in-mapper combining technique provides the advantages of control and efficiency, it may also come with disadvantage of adding a level of complexity in terms of requiring the skill and knowledge to provide a workable and efficient algorithm.

**Comment:** Can you give more specific issues?

### Part 3 of 4 - Multiple Choice 11.25 / 15.0 Points

If the question only allows a single choice that means there is only one correct answer. If the question allows more than one possible choice then there could be one or more correct answers. Select all the correct answers, if you miss selecting a correct answer you will lose points if you select an incorrect answer you also lose points.

Question 5 of 8

10.0

10.0 Points

In a cluster of 10 machines. Where each machine has 6 I/O channels (or 6 hard drives) such that each channel can read 200 MB/sec. How fast can this system read 10TB of data.

- A. 833.333 minutes
- B. 13.888 minutes
- C. 0.138 minutes
- D. 1.388 minutes
- E. 8.333 minutes
- F. 83.333 minutes

Rationale:

$$10\text{TB} = 10 * 1000 * 1000 \Rightarrow 10\text{TB} / 10 * 6 * 200 = 10\text{ TB} / 12000 * 60 = 13.888\text{ minutes}$$

**Comment:** Good

Question 6 of 8

1.25

5.0 Points

Select all that apply to NameNode

- A. Very expensive hardware with double/triple redundancy (RAID).
- B. Reads data from RAM and writes into hard drive.
- C. Single Point of Failure
- D. Manages the filesystem namespace.
- E. Serving read and write requests.
- F. Master of the system
- G. Deployed on each machine to provide storage.

## Part 4 of 4 - Map-Reduce Programs 60.0 / 60.0 Points

**Answer each of the following questions by writing the map and reduce and any other necessary methods needed to answer the question. If some parts are not modified then there is no need to provide them, for example, if your partitioner does not need to be modified then there is no need to provide it.**

**You are expected to write pseudo code in the proper format as you were taught in Algorithms. Not writing proper pseudo code will result in losing points.**

Question 7 of 8

30.0

30.0 Points

Food waste is a major problem in the country. One possible way to overcome food waste is to predict the average consumption to limit over-production. The USDA needs to find out how much the average consumption of different items is during different times. The USDA has collected the consumption data for different items from different locations and times.

Write a MapReduce program that generates the average consumption of each product in each state for each month. The data is for the last 50 years.

Below is a sample of the data records.

<b>Year</b>	<b>Month</b>	<b>State</b>	<b>Product</b>	<b>Consumption</b>	<b>Cost</b>	<b>Highest Sale Type</b>
2018	July	IA	Milk	56.2 Ton	52.1 K	Bulk
2002	June	CA	Apples	33.3 Ton	66.3 K	Online
2020	February	MD	Cereal	18.8 Ton	44.2 K	Store
2023	December	CA	Apples	23.3 Ton	1.2 K	Farmers Market

```
class Mapper{
```

```
    initialize()
```

```
    H = HashMap()
```

```
    map(docid a, doc d)
```

```
    for all r in record do
```

```
        Key k = new Key(r.product, r.state, r.month)
```

```
        c = r.consumption
```

```
        if (H{k} == null)
```

```
            H{k} = (c, 1)
```

```
    else
```

```
        H{k} = (H{k}.getConsumption() + c, H{k}.getCount() + 1)
```

```
    close()
```

```
    for all entry in H do
```

```
        Emit(entry.getKey(), entry.getValue())
```

```
}
```

```
getPartition(key k, int numberofReducers)
```

```
return Math.abs(k.toString().hashCode) % numberofReducers
```

```
class Reducer
```

```
{
```

```
    reduce(Key k, ValuePairs[v1, v2, v3 ...])
```

```
        consumptionTotal = 0
```

```
        countTotal = 0
```

```
        for all v in ValuePairs do
```

```
            consumptionTotal += v.getConsumption()
```

```
            countTotal += v.getCount()
```

```

    Emit(k, consumptionTotal / countTotal)
}

```

**Comment:** Good

Question 8 of 8

30.0

30.0 Points

Finding hidden patterns is very important for many businesses. Write a Map-Reduce program to find the hidden pattern in an online e-commerce website. Your program should generate the co-occurrence matrix for purchasing different items.

The online e-commerce website sells sports goods. We would like to advertise a few items once someone selects items to purchase (people who purchased this also purchased this). Customers get 2 different advertisements.

We wish to advertise to customers who have added at least two items to their basket, we only consider items purchased together in the same order.

The event are the purchased items, that is, the sports goods.

Let u represent an event of purchasing item(s).

Let Window(u) represent the occurrences of purchasing other items after u.

```

Mapper {
    maper(docid a, doc d)
        for all u in record do
            otherItems = ArrayList
            for all v in Window(u) do
                otherItems.add(v);
            Emit(u, otherItems)
}

```

```

getPartition(key k, int numberReducers)

    return Math.abs(k.toString().hashCode) % numberReducers

```

```

class Reducer {
    reducer(Key k, OtherItems [v1, v2, v3, ...])
        Vfinal = []
        for all v in OtherItems do
            Vfinal.add(v)
        Emit(k, Vfinal)
}

```

**Comment:** Good

- [Gateway](#)
- [Accessibility Information](#)
- [The Sakai Project](#)
- [!\[\]\(f0fdb60b777eb11b66cba545acf146fa\_img.jpg\)](#)

Copyright 2003-2024 The Apereo Foundation. All rights reserved.



## Notifications



Push is not enabled [Enable](#)

[Assignments](#) 3

-  **Patrick Wise** Dec 2, 2024, 8:31 AM [!\[\]\(b93c098f396df0e119633c762d589ae8\_img.jpg\)](#)  
Created/updated a new assignment "CS574 Practicum on the creativity in the field of Technology" in "CS57x-2024-FallA"  

-  **Patrick Wise** Nov 13, 2024, 9:52 AM [!\[\]\(15980b6e0ee1253f6d0d92d771982180\_img.jpg\)](#)  
Graded your submission for assignment "CS574 Practicum on the creativity in the field of Technology" in "CS57x-2024-FallA"  

-  **Arbey Fernando Ramirez Garcia** Nov 1, 2024, 4:19 PM [!\[\]\(9cc9b69431721f989f01cc8a2a06f52a\_img.jpg\)](#)  
Created/updated a new assignment "CS575 Practicum on CS Theory in Practice" in "CS57X-2024-FallB"  


[Update](#) [Mark all as viewed](#) [Clear All](#)

## Search Sakai

Search [Search Sakai](#) [!\[\]\(0df0bdc1e09cbc2587d9dd4511cb0c27\_img.jpg\)](#)

Enter search terms

Enter search term to :  [Search Sakai](#)

## Account Menu

**Derrick Korku Siayor****615652**

- [Manage Preferences](#)
- [View Profile](#)
- [Tutorial](#)
- 

**My Calendar**

Showing events for the next undefined days

**< 2024 >****< December >****Today**

Sun	Mon	Tue	Wed	Thu	Fri	Sat
1	2	3	4	5	<u>6</u>	7
8	9	10	11	12	13	14
15	16	17	18	19	20	21
22	23	24	25	26	27	28
29	30	31	1	2	3	4

**My Grades**

Log Out

- 
- ! There are new notifications available
- 
- 
-