Task 1: A library to build a distributed data storage system using a remote object paradigm that will allow their employees to access, store and update book information.

In this program I created three files to run this program.

1) Read_csv.py
2) Person.py
3) Warehouse.py

1) Read_csv.py

In read_csv.py file I imported library files such as sys (this is used for system specific paramiters), Pyro4 (these objects can talk with each other), person (this library imports person data for the .py file). I connected the host name with warehouse.py file (It means conncet with the server). After that, I defined each person's name and their login details. Any number of people can be added to the program.

2) Person.py

In this file I imported pprint, collection and sys libaraies. In line 6th, I define sys version. After that, I defined the function 'convert'. This function converts all the data into items, then it does mapping and then returns the data. In line 19th, I created a class named as 'Person' and then I called main program function. In line 23, I created a new function 'visit'. In this function I called 'which statement you want to write'. In this function, I have to choose one function from 'd, a, si, sy, ol and nl'. This function is called the warehouse function and performs according to the keyword you entered. In line 43, I defined the function called 'addstore'. In this field, I have to add data of the book and call the warehouse function. I used same functions for further 4 times.

3) Warehouse.py

This warehouse means temp server is created using the home system. I imported the print function and collection, Pyro4, pprint, and csv libraries. In line 7th I created an object to open and write data fro csv file. Then I created an array and created the function 'convert' with same name as Person.py file. After that I executed the command to read all the lines and print them. In line 24th, I defined a class with the name 'warehouse'. Then I displayed all the data from csv file and printed this data in server side and prints the book data name. I also used 'store' function to store the book data. This data has unique id, book name, isbn, year of publication and book author. After adding these details, the data will be stored in csv file and being shown on the server side. I also created 'display_isbn' function to display all the isbn data. From line 66 to 96, the program checks if the book is in loan or not in loan. In this function, all the for loops, if-else conditions remains same but only the data can change such as if the book being given in loan or not in loan.

After this function, the function for lines 66 to 96% will be check if the book is in loan or not in loan. In this function all the for-loops, if-else condition remains same but only the data is in loan or not. Last two functions store print in server side and retrive all the data and prints in server side. Line 120 shows the run pyro and displays the local host in program. And this can be called to the call objects. The output can be seen as below:

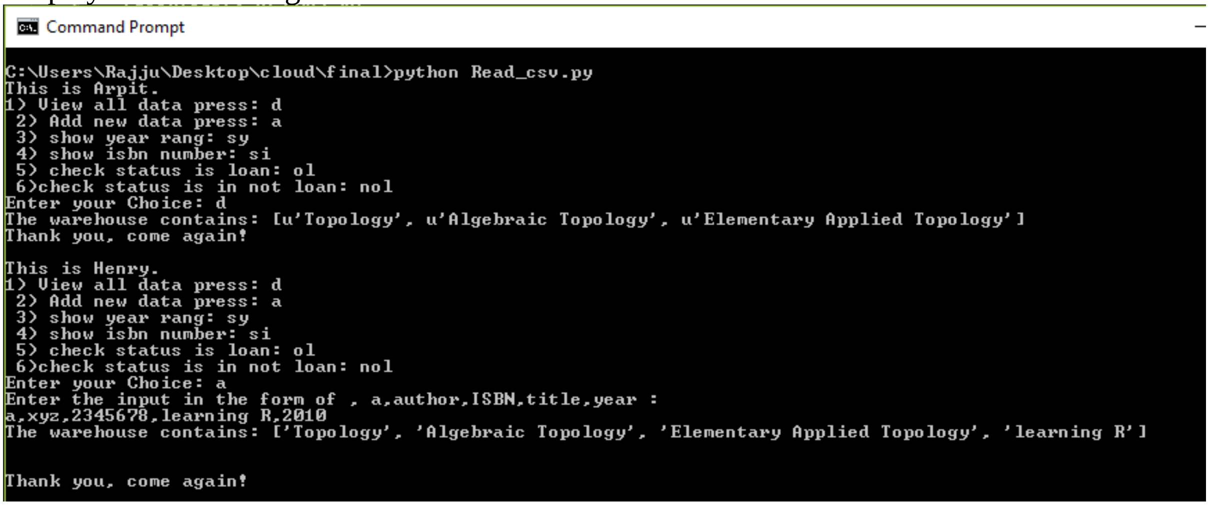
OUTPUT:

NOTE: THIS PROGRAM RUN IN PYTHON 2.7

Open cmd and write: pyro4-ns

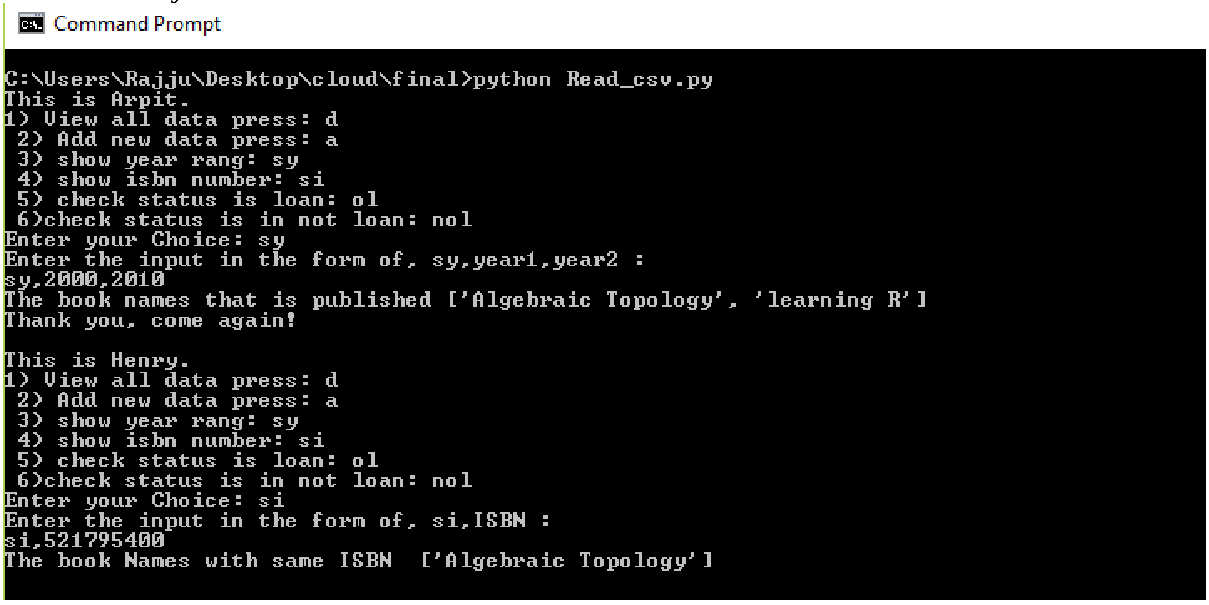
1) Open second cmd and write: python warehouse.py
2) Open thired cmd and write: python read_csv.py

Display all and Adding all the data



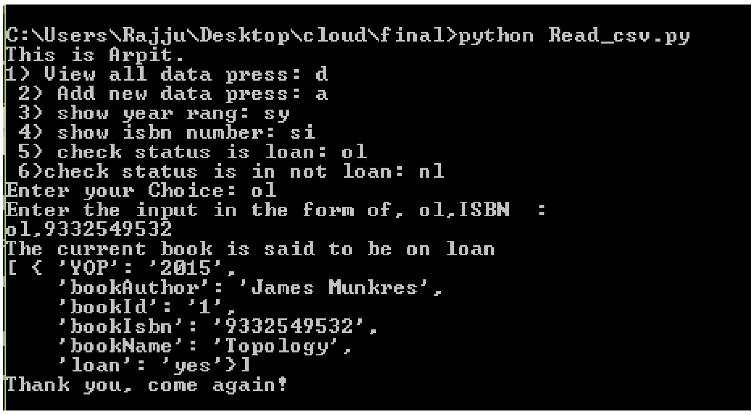
```
C:\Users\Rajju\Desktop\cloud\final>python Read_csv.py
This is Arpit.
1) View all data press: d
 2) Add new data press: a
 3) show year rang: sy
 4) show isbn number: si
 5) check status is loan: ol
 6)check status is in not loan: nol
Enter your Choice: d
The warehouse contains: [u'Topology', u'Algebraic Topology', u'Elementary Applied Topology']
Thank you, come again!

This is Henry.
1) View all data press: d
 2) Add new data press: a
 3) show year rang: sy
 4) show isbn number: si
 5) check status is loan: ol
 6)check status is in not loan: nol
Enter your Choice: a
Enter the input in the form of , a,author,ISBN,title,year :
a,xyz,2345678,learning R,2010
The warehouse contains: ['Topology', 'Algebraic Topology', 'Elementary Applied Topology', 'learning R']


Thank you, come again!
```

Show all the years between 2000 to 2010 and Show isbn numbers.



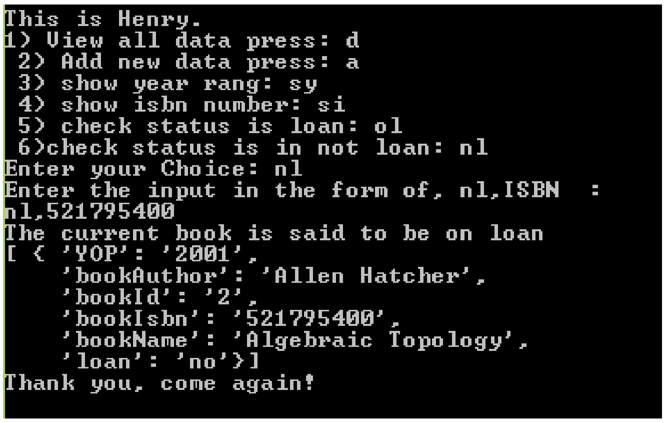
```
C:\Users\Rajju\Desktop\cloud\final>python Read_csv.py
This is Arpit.
1) View all data press: d
 2) Add new data press: a
 3) show year rang: sy
 4) show isbn number: si
 5) check status is loan: ol
 6)check status is in not loan: nol
Enter your Choice: sy
Enter the input in the form of, sy,year1,year2 :
sy,2000,2010
The book names that is published ['Algebraic Topology', 'learning R']
Thank you, come again!

This is Henry.
1) View all data press: d
 2) Add new data press: a
 3) show year rang: sy
 4) show isbn number: si
 5) check status is loan: ol
 6)check status is in not loan: nol
Enter your Choice: si
Enter the input in the form of, si,ISBN :
si,521795400
The book Names with same ISBN  ['Algebraic Topology']
```

In loan



```
C:\Users\Rajju\Desktop\cloud\final>python Read_csv.py
This is Arpit.
1) View all data press: d
 2) Add new data press: a
 3) show year rang: sy
 4) show isbn number: si
 5) check status is loan: ol
 6)check status is in not loan: nl
Enter your Choice: ol
Enter the input in the form of, ol,ISBN  :
ol,9332549532
The current book is said to be on loan
[ { 'YOP': '2015',
    'bookAuthor': 'James Munkres',
    'bookId': '1',
    'bookIsbn': '9332549532',
    'bookName': 'Topology',
    'loan': 'yes'}]
Thank you, come again!
```

Not in loan



```
This is Henry.
1) View all data press: d
 2) Add new data press: a
 3) show year rang: sy
 4) show isbn number: si
 5) check status is loan: ol
 6)check status is in not loan: nl
Enter your Choice: nl
Enter the input in the form of, nl,ISBN  :
nl,521795400
The current book is said to be on loan
[ { 'YOP': '2001',
    'bookAuthor': 'Allen Hatcher',
    'bookId': '2',
    'bookIsbn': '521795400',
    'bookName': 'Algebraic Topology',
    'loan': 'no'}]
Thank you, come again!
```

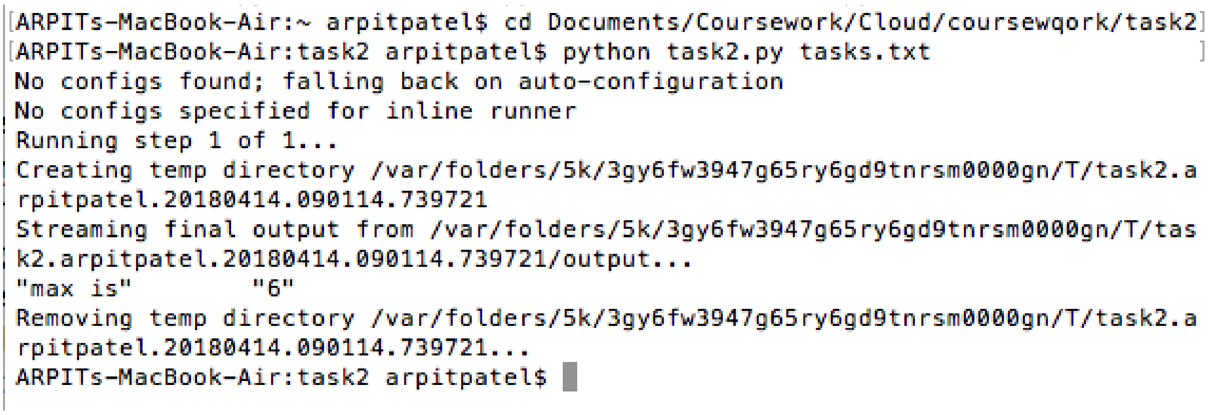
Task 2: Find maximum number using MapReduce

In this programme first, I imported mrjob library. Using this library, we can do multiple-steps of MapReduce jobs in python. In line 3, I defined class and its name is MRWordCounter and imported the library in this class. After that, I used mapper function to use MapReduce. In this function I define (self, key, line). Self-keyword is representing the instance class, using self-keyword. Also, we can access the attributes and methods of the classes. In 5th line, I used 'for loop' for the word count and to split the word. In python, there is one split function available to the split the line. In line 6, I used yield keyword to generate and calculate a series of result one-by-one on demand. After that, in the 8th line, I used second function called 'reducer'. In this function I printed the data of the maximum number in .txt file. In second last line, the main program starts using the main function. In 12th line, I called the class and returned the value.

How to run program:

First of all, find the location of the program in Terminal or Cmd. After that, write this syntax: python filename.py filename.txt and press enter. The output can be seen as below:

OUTPUT:



```
[ARPITs-MacBook-Air:~ arpitpatel$ cd Documents/Coursework/Cloud/coursewqork/task2]
[ARPITs-MacBook-Air:task2 arpitpatel$ python task2.py tasks.txt                   ]
No configs found; falling back on auto-configuration
No configs specified for inline runner
Running step 1 of 1...
Creating temp directory /var/folders/5k/3gy6fw3947g65ry6gd9tnrsm0000gn/T/task2.a
rpitpatel.20180414.090114.739721
Streaming final output from /var/folders/5k/3gy6fw3947g65ry6gd9tnrsm0000gn/T/tas
k2.arpitpatel.20180414.090114.739721/output...
"max is"        "6"
Removing temp directory /var/folders/5k/3gy6fw3947g65ry6gd9tnrsm0000gn/T/task2.a
rpitpatel.20180414.090114.739721...
ARPITs-MacBook-Air:task2 arpitpatel$ 
```

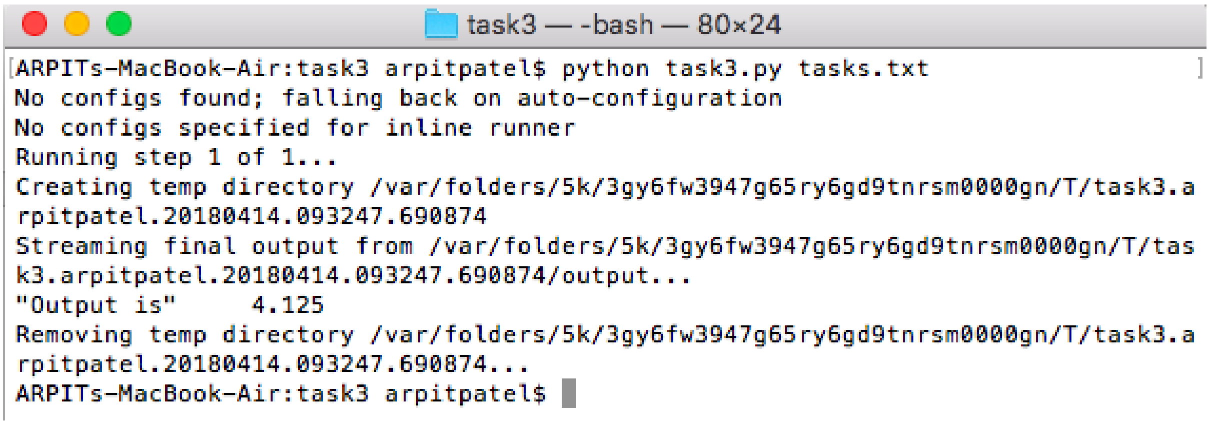
Task 3: Determines the mean of all numbers in the file using MapReduce

In this task, I imported mrjob library. In 3rd line I defined a class,, named is Mr and imported the mrjob library. After, I used mapper function for MapReduce. In this function I defined self, key, value. In Line 6th, 7th and 8th, I defined key,' a', num value is zero. After 9th line, I used 'for loop' to count the line and provide the answer and split the line. In 10th line, I stored all the data in a variable. Also, the value of 'i' is considered in the float and do sum operations using variable 'a'. The num+ =1 to increment the number by 1. In yield function, 'a' stores all the data. In 14th line, I used second function called reducer and after 2 line I define variable 'a' and num as a 0. Again, I used 'for loop' for the values and store in 'i '. After that, I used plus operation for the variable 'and'. In yield, I printed the output which divides the last output. In last two lines the main program starts and the function Mr class is called.

How to run program:

First find the location of the program in Terminal or Cmd. After that, write this syntax: python filename.py filename.txt and press enter. The output can be seen as below:

OUTPUT:



```
[ARPITs-MacBook-Air:task3 arpitpatel$ python task3.py tasks.txt                   ]
No configs found; falling back on auto-configuration
No configs specified for inline runner
Running step 1 of 1...
Creating temp directory /var/folders/5k/3gy6fw3947g65ry6gd9tnrsm0000gn/T/task3.a
rpitpatel.20180414.093247.690874
Streaming final output from /var/folders/5k/3gy6fw3947g65ry6gd9tnrsm0000gn/T/tas
k3.arpitpatel.20180414.093247.690874/output...
"Output is"     4.125
Removing temp directory /var/folders/5k/3gy6fw3947g65ry6gd9tnrsm0000gn/T/task3.a
rpitpatel.20180414.093247.690874...
ARPITs-MacBook-Air:task3 arpitpatel$ 
```

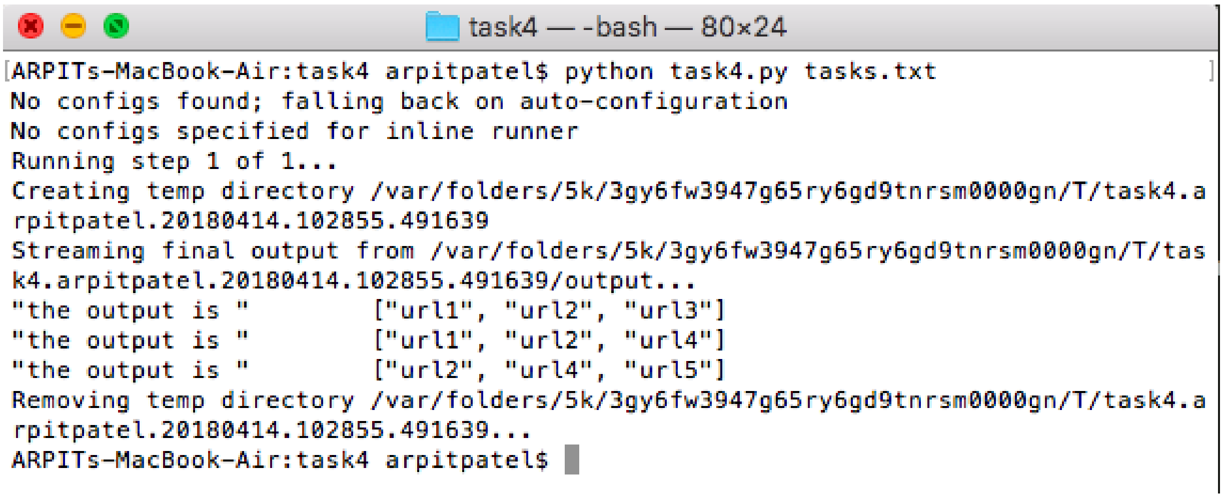
Task 4: The first and second values on each line specify the source and destination of the link

In this task, I imported mrjob library. In 3rd line, I defined the class name ‘Mr’ and imported the mrjob library. In 5th line, I declared ‘data’ variable in the globe. After that, I used mapper function to use the MapReduce. In this function I defined “self, key, line”. I split the 8th line and stored the ‘line” name as variable. After that, I joined all the data using append function and printed this data using yield. In line 12th, I used reducer function. Later in 13th line, I stored values in ‘listval’ variable. After that, I used ‘number’ variable using length function minus this value by 1. Again, I used ‘listval’ variable and stored the number data. In 16th line, I used two ‘for-loops’ for the data. In the 18th line, I used ‘if’ condition for comparing the two ‘for-loop’. Then, I used a variable for the stored data and converted it to 2-dimensional to 3-dimensional array. After this line ‘list’ method lists all the data in 3-dimensional array. In 21st line, I printed the output using yield function and the ‘tuple’ method converts a list variable into tuple. In, last two lines the main program starts and calls the class.

How to run program:

First find the location of the program in Terminal or Cmd. After that, write this syntax: python filename.py filename.txt and press enter. The output can be seen as below:

OUTPUT:



```
ARPITs-MacBook-Air:task4 arpitpatel$ python task4.py tasks.txt
No configs found; falling back on auto-configuration
No configs specified for inline runner
Running step 1 of 1...
Creating temp directory /var/folders/5k/3gy6fw3947g65ry6gd9tnrsm0000gn/T/task4.a
rpitpatel.20180414.102855.491639
Streaming final output from /var/folders/5k/3gy6fw3947g65ry6gd9tnrsm0000gn/T/tas
k4.arpitpatel.20180414.102855.491639/output...
"the output is "        ["url1", "url2", "url3"]
"the output is "        ["url1", "url2", "url4"]
"the output is "        ["url2", "url4", "url5"]
Removing temp directory /var/folders/5k/3gy6fw3947g65ry6gd9tnrsm0000gn/T/task4.a
rpitpatel.20180414.102855.491639...
ARPITs-MacBook-Air:task4 arpitpatel$
```