Task 1:

**MapReduce**:

* Pairs: Ran the map reduce job on all the three datasets available on the website.
  + The actual code files for mapper and reducer can be found in **Task1/Mapreduce/Code**
  + The pseudo-code can be found in **Task1/Mapreduce/Pseudocode**/**Pseudo-code\_pairs.txt**
* Stripes: Ran the map reduce job on all the three datasets available on the website.
  + The actual code files for mapper and reducer can be found in **Task1/Mapreduce/Code**
  + The pseudo-code can be found in **Task1/Mapreduce/Pseudocode/** **Pseudo-code\_stripes.txt**
* Movie\_names:
  + To get the movie name, I have created another python script -> **Task1/Mapreduce/Code/get\_movie\_name.py.** The pseudo code for this is also given in **Task1/Mapreduce/Pseudocode/Pseudocode\_get\_movie\_names.txt**

**SPARK**:

* Pairs: Ran the map reduce job on all the three datasets available on the website.
  + The actual code files for mapper and reducer can be found in **Task1/** **Spark/Code**
  + The pseudo-code can be found in **Task1/** **Spark/Pseudocode**/**Pseudo-code\_pairs.txt**
* Stripes: Ran the map reduce job on all the three datasets available on the website.
  + The actual code files for mapper and reducer can be found in **Task1/** **Spark/Code**
  + The pseudo-code can be found in **Task1/** **Spark/Pseudocode/** **Pseudo-code\_stripes.txt**
* Movie names:
  + To get the movie name, I have created another python script -> **Task1/** **Spark/Code/get\_movie\_name.py.** The pseudo code for this is also given in **Task1/** **Spark/Pseudocode/Pseudocode\_get\_movie\_names.txt**

TASK 2:

Compute conditional probability

TASK 3:

Compute lift