Requirements

- Create a file that contains two columns: 'dptid' and 'salary'. They respectively correspond to the employees' department unique identifier and average department salary of a fake company.

  Example of the file schema and data

  ```
  recordid,income
  1,50000000
  2,50070000
  3,50450000
  5,50920000
  6,50003700
  7,50000000
  9,50070000
  10,50450000
  11,50920000
  13,50003700
  ```

- Store this file to HDFS.

- Write a Spark program which calculates the number of lines in a file

- Write a Spark program which reads the file from HDFS and calculates the minimum and maximum salaries per department id.

- Saves the output in HDFS.