

Name: ..... Surname: .....

Name: ..... Surname: .....

Name: ..... Surname: .....

## Employees and departments

Consider two files containing the following data:

Employees.txt

EMP1;CARME;400000;MATARO;DPT1  
EMP2;EUGENIA;350000;TOLEDO;DPT2  
EMP3;JOSEP;250000;SITGES;DPT3  
EMP4;RICARDO;250000;MADRID;DPT4  
EMP5;EULALIA;150000;BARCELONA;DPT5  
EMP6;MIQUEL;125000;BADALONA;DPT5  
EMP7;MARIA;175000;MADRID;DPT6  
EMP8;ESTEBAN;150000;MADRID;DPT6

Departments.txt

DPT1;DIRECCIO;10;PAU CLARIS;BARCELONA  
DPT2;DIRECCIO;8;RIOS ROSAS;MADRID  
DPT3;MARKETING;1;PAU CLARIS;BARCELONA  
DPT4;MARKETING;3;RIOS ROSAS;MADRID  
DPT5;VENDES;1;MUNTANER;BARCELONA  
DPT6;VENDES;1;CASTELLANA;MADRID

Provide the ordered list of Spark operations (no need to follow the exact syntax, just the kind of operation and main parameters) you'd need to retrieve for each employee his/her department information. Save the results in "output.txt".

- 1) .....
- 2) .....
- 3) .....
- 4) .....
- 5) .....
- 6) .....
- 7) .....
- 8) .....
- 9) .....
- 10).....

Name: ..... Surname: .....

Name: ..... Surname: .....

Name: ..... Surname: .....

## Error log

Consider an error log file ("log.txt") like the one bellow:

```
20150323;0833;ERROR;Oracle
20150323;0835;WARNING;MySQL
20150323;0839;WARNING;MySQL
20150323;0900;WARNING;Oracle
20150323;0905;ERROR;MySQL
20150323;1013;OK;Oracle
20150323;1014;OK;MySQL
20150323;1055;ERROR;Oracle
```

Provide the ordered list of Spark operations (no need to follow the exact syntax, just the kind of operation and main parameters) you'd need to retrieve the lines corresponding to both "errors" and "warnings", but adding "Important:" at the beginning of those of errors (i.e., only "errors"). Save the results in "output.txt".

- 1) .....
- 2) .....
- 3) .....
- 4) .....
- 5) .....
- 6) .....
- 7) .....
- 8) .....
- 9) .....
- 10).....

Name: ..... Surname: .....

Name: ..... Surname: .....

Name: ..... Surname: .....

## Wines

Consider a file ("wines.txt") containing the following data:

type\_1,2.064254

type\_3,2.925376

type\_2,2.683955

type\_1,2.991452

type\_2,2.861996

type\_1,2.727688

Provide the ordered list of Spark operations (no need to follow the exact syntax, just the kind of operation and main parameters) you'd need to retrieve the minimum value per type. Save the results in "output.txt".

- 1) .....
- 2) .....
- 3) .....
- 4) .....
- 5) .....
- 6) .....
- 7) .....
- 8) .....
- 9) .....
- 10).....