



Activities Text Editor Feb 14 04:27

code.py /usr/local/spark/spark-3.3.1-bin-hadoop3/bin

spark-shell.cmd code.py

```
from pyspark.context import SparkContext
from pyspark.sql.session import SparkSession
s=SparkContext.getOrCreate()
spark=SparkSession(s)

col=["sex","class","survived","died"]
titanic=[("Children","First",6,0),("Children","Second",24,0),
("Children","Third",27,52),("Men","First",57,118),("Men","Second",14,154),
("Men","Third",75,387),("Men","Crew",192,693),("Women","First",140,4),
("Women","Second",80,13),("Women","Third",76,89),("Women","Crew",20,3)]

dFrame=spark.createDataFrame(titanic,col)
dFrame.show()

dFrame2=spark.createDataFrame(titanic,col)
dFrame2.filter((dFrame2.survived/(dFrame2.survived+dFrame2.died))>0.8).show()
```

Activities Terminal Feb 14 04:12

aish@Ubuntu: /usr/local/spark/spark-3.3.1-bin-hadoop3/bin

[Stage 0:> (0 + 1) / 1

| sex | class | survived | died |
|----------|--------|----------|------|
| Children | First | 6 | 0 |
| Children | Second | 24 | 0 |
| Children | Third | 27 | 52 |
| Men | First | 57 | 118 |
| Men | Second | 14 | 154 |
| Men | Third | 75 | 387 |
| Men | Crew | 192 | 693 |
| Women | First | 140 | 4 |
| Women | Second | 80 | 13 |
| Women | Third | 76 | 89 |
| Women | Crew | 20 | 3 |

| sex | class | survived | died |
|----------|--------|----------|------|
| Children | First | 6 | 0 |
| Children | Second | 24 | 0 |
| Women | First | 140 | 4 |
| Women | Second | 80 | 13 |
| Women | Crew | 20 | 3 |

aish@Ubuntu: /usr/local/spark/spark-3.3.1-bin-hadoop3/bin\$

