

Assignment 4

Group: supreme-broccoli 🍷🥦

1) Zip every cell with column name

```
map(df => {df.flatMap(row => row.toSeq.map(p => p.zip(columns))
```

2) Reduce by value union attributes

```
.reduce((a, b) => a.union(b)).reduceByKey((a, b) => (a ++ b).distinct).map(_._2)
```

3) Create inclusion lists

```
.flatMap(set => { set.map(column => (column , set - column)) })
```

4) Reduce by intersect and drop empty ones

```
.reduceByKey(_ intersect _).filter(_._2.nonEmpty)
```

5) Collect and Print

```
.collect().sortBy(_._1).foreach(println())
```

Caution:
Code was simplified and shortened