Example Coding Tasks for a Java Developer

1. Writing a Hive User Defined Aggregate Function (UDAF)

- Develop a new User Defined Aggregate Function for Hive: `concat_distinct(col)`.
- It returns a `string` that contains all values of the input column separated by a comma and a space. One value should only appear once.
- On the following input table (`table1`)

att1	att2	att3
1	one	apple
2	one	peach
3	two	lemon
4	one	apple

the following query

```
SELECT att2, concat_distinct(att3) fruits
FROM table1
GROUP BY att2
```

should give the following result

att2	fruits
one	peach, apple
two	lemon

- The output should always be deterministic.
- The task is intentionally loosely defined, feel free to make decisions on your own.
- Hint: Built-in Aggregate Functions source code may help.

2. Writing a RapidMiner operator

...