

Streaming Aggregations



Objective

Aggregate data in micro-batches

Structured Streaming aggregation API



Takeaways

Same aggregation API as non-streaming DFs

```
val query = lines.select(count("*").as("lineCount"))
val queryExpr = lines.selectExpr("count(*) as lineCount")
```

To keep in mind

- aggregations work at a micro-batch level
- the append output mode not supported without watermarks*
- some aggregations are not supported, e.g. sorting or chained aggregations

Spark rocks

