

# **Streaming Datasets**



# Objective

Type-safe structured streams



# Takeaways

Same DS conversion as non-streaming DS

```
import spark.implicits._  
val carsDS = spark.readStream  
  .schema(carsSchema)  
  .json("src/main/resources/data/cars")  
  .as[Car]
```

Streaming DSs support functional operators

```
val carNamesDS = carsDS.map(car => car.name.toUpperCase())
```

## Tradeoffs

- pros: type safety, expressiveness
- cons: potential perf implications as lambdas cannot be optimized

# Spark rocks

