

Variant Data Types



Variant

Semi-structured data

Useful for changing/evolving schemas

Much faster to work with than JSON



Variant

Semi-structured data: rows with a wide range of fields

Useful for changing/evolving schemas

```
val variantDF = jsonDF.select(parse_json(col("raw_json")).as("data"))

val extractedDF = variantDF.select(
  variant_get(col("data"), "$.name", "string").as("name"),
  variant_get(col("data"), "$.age", "int").as("age"),
  try_variant_get(col("data"), "$.address.city", "string").as("city") // returns null if path missing
)
```

Much faster to work with than JSON

Spark rocks

