

```
+ DataFormats-Scala git:(master) x parquet-tools dump --disable-crop serialized_parquet_file.parquet
row group 0
-----
target: INT64 UNCOMPRESSED D0:0 FP0:4 SZ:77/77/1.00 VC:2 ENC:BIT_PACKED,PLAIN ST:[min: 1, max: 2, num_nulls: 0]
id:      INT64 UNCOMPRESSED D0:0 FP0:81 SZ:77/77/1.00 VC:2 ENC:BIT_PACKED,PLAIN ST:[min: 123, max: 234, num_nulls: 0]
date:    BINARY UNCOMPRESSED D0:0 FP0:158 SZ:101/101/1.00 VC:2 ENC:BIT_PACKED,PLAIN ST:[min: Saturday 8th, June, max: Sunday 9th, June, num_nulls: 0]
user:    BINARY UNCOMPRESSED D0:0 FP0:259 SZ:57/57/1.00 VC:2 ENC:BIT_PACKED,PLAIN ST:[min: alpha1, max: alpha2, num_nulls: 0]
text:    BINARY UNCOMPRESSED D0:0 FP0:316 SZ:95/95/1.00 VC:2 ENC:RLE,BIT_PACKED,PLAIN ST:[min: Parquet Tweet1, max: Parquet Tweet2, num_nulls: 0]

-----
target TV=2 RL=0 DL=0
-----
page 0: DLE:BIT_PACKED RLE:BIT_PACKED VLE:PLAIN ST:[min: 1, max: 2, num_nulls: 0] SZ:16 VC:2

id TV=2 RL=0 DL=0
-----
page 0: DLE:BIT_PACKED RLE:BIT_PACKED VLE:PLAIN ST:[min: 123, max: 234, num_nulls: 0] SZ:16 VC:2

date TV=2 RL=0 DL=0
-----
page 0: DLE:BIT_PACKED RLE:BIT_PACKED VLE:PLAIN ST:[min: Saturday 8th, June, max: Sunday 9th, June, num_nulls: 0] SZ:42 VC:2

user TV=2 RL=0 DL=0
-----
page 0: DLE:BIT_PACKED RLE:BIT_PACKED VLE:PLAIN ST:[min: alpha1, max: alpha2, num_nulls: 0] SZ:20 VC:2

text TV=2 RL=0 DL=1
-----
page 0: DLE:RLE RLE:BIT_PACKED VLE:PLAIN ST:[min: Parquet Tweet1, max: Parquet Tweet2, num_nulls: 0] SZ:42 VC:2

INT64 target
-----
*** row group 1 of 1, values 1 to 2 ***
value 1: R:0 D:0 V:1
value 2: R:0 D:0 V:2

INT64 id
-----
*** row group 1 of 1, values 1 to 2 ***
value 1: R:0 D:0 V:123
value 2: R:0 D:0 V:234

BINARY date
-----
*** row group 1 of 1, values 1 to 2 ***
value 1: R:0 D:0 V:Saturday 8th, June
value 2: R:0 D:0 V:Sunday 9th, June

BINARY user
-----
*** row group 1 of 1, values 1 to 2 ***
value 1: R:0 D:0 V:alpha1
value 2: R:0 D:0 V:alpha2

BINARY text
-----
*** row group 1 of 1, values 1 to 2 ***
value 1: R:0 D:1 V:Parquet Tweet1
value 2: R:0 D:1 V:Parquet Tweet2
```

```
→ DataFormats-Scala git:(master) x parquet-tools schema serialized_parquet_file.parquet
message ParquetSchema {
  required int64 target = 1;
  required int64 id = 2;
  required binary date (UTF8) = 3;
  required binary user (UTF8) = 4;
  optional binary text (UTF8) = 9;
}
```

```
→ DataFormats-Scala git:(master) x parquet-tools meta serialized_parquet_file.parquet
file:      file:/Users/arunma/IdeaProjects/OSS/DataFormats-Scala/serialized_parquet_file.parquet
creator:    parquet-mr version 1.10.0 (build 031a6654009e3b82020012a18434c582bd74c73a)
extra:      parquet.avro.schema = {"type":"record","name":"TweetAvro","namespace":"com.tweet.avro","fields":[{"name":"target","type":"int"}, {"name":"id","type":"long"}, {"name":"date","type":"string"}, {"name":"user","type":"string"}, {"name":"text","type":"string"}]}
extra:      writer.model.name = avro

file schema: com.tweet.avro.TweetAvro
-----
target:     REQUIRED INT32 R:0 D:0
id:         REQUIRED INT64 R:0 D:0
date:       REQUIRED BINARY 0:UTF8 R:0 D:0
user:       REQUIRED BINARY 0:UTF8 R:0 D:0
text:       REQUIRED BINARY 0:UTF8 R:0 D:0

row group 1: RC:2 TS:377 OFFSET:4
-----
target:     INT32 UNCOMPRESSED D0:0 FP0:4 SZ:53/53/1.00 VC:2 ENC:BIT_PACKED,PLAIN ST:[min: 1, max: 2, num_nulls: 0]
id:         INT64 UNCOMPRESSED D0:0 FP0:57 SZ:77/77/1.00 VC:2 ENC:BIT_PACKED,PLAIN ST:[min: 123, max: 234, num_nulls: 0]
date:       BINARY UNCOMPRESSED D0:0 FP0:134 SZ:101/101/1.00 VC:2 ENC:BIT_PACKED,PLAIN ST:[min: Saturday 8th, June, max: Sunday 9th, June, num_nulls: 0]
user:       BINARY UNCOMPRESSED D0:0 FP0:235 SZ:57/57/1.00 VC:2 ENC:BIT_PACKED,PLAIN ST:[min: alpha1, max: alpha2, num_nulls: 0]
text:       BINARY UNCOMPRESSED D0:0 FP0:292 SZ:89/89/1.00 VC:2 ENC:BIT_PACKED,PLAIN ST:[min: Parquet tweet1, max: Parquet tweet2, num_nulls: 0]
```

```
→ DataFormats-Scala git:(master) x parquet-tools meta serialized_parquet_file.parquet
file:      file:/Users/arunma/IdeaProjects/OSS/DataFormats-Scala/serialized_parquet_file.parquet
creator:    parquet-mr version 1.10.0 (build 031a6654009e3b82020012a18434c582bd74c73a)
extra:      writer.model.name = thrift
extra:      thrift.descriptor = { "id" : "STRUCT", "children" : [ { "name" : "target", "fieldId" : 1, "requirement" : "REQUIRED", "type" : { "id" : "I64" } },
{ "name" : "id", "fieldId" : 2, "requirement" : "REQUIRED", "type" : { "id" : "I64" } }, { "name" : "date", "fieldId" : 3, "requirement" : "REQUIRED", "type"
: { "id" : "STRING" } }, { "name" : "user", "fieldId" : 4, "requirement" : "REQUIRED", "type" : { "id" : "STRING" } }, { "name" : "text", "fieldId" : 9, "requi
rement" : "OPTIONAL", "type" : { "id" : "STRING" } } ], "structOrUnionType" : "STRUCT" }
extra:      thrift.class = com.tweet.thrift.TweetThrift

file schema: ParquetSchema
-----
target:     REQUIRED INT64 R:0 D:0
id:         REQUIRED INT64 R:0 D:0
date:       REQUIRED BINARY 0:UTF8 R:0 D:0
user:       REQUIRED BINARY 0:UTF8 R:0 D:0
text:       OPTIONAL BINARY 0:UTF8 R:0 D:1

row group 1: RC:2 TS:407 OFFSET:4
-----
target:     INT64 UNCOMPRESSED D0:0 FP0:4 SZ:77/77/1.00 VC:2 ENC:PLAIN,BIT_PACKED ST:[min: 1, max: 2, num_nulls: 0]
id:         INT64 UNCOMPRESSED D0:0 FP0:81 SZ:77/77/1.00 VC:2 ENC:PLAIN,BIT_PACKED ST:[min: 123, max: 234, num_nulls: 0]
date:       BINARY UNCOMPRESSED D0:0 FP0:158 SZ:101/101/1.00 VC:2 ENC:PLAIN,BIT_PACKED ST:[min: Saturday 8th, June, max: Sunday 9th, June, num_nulls: 0]
user:       BINARY UNCOMPRESSED D0:0 FP0:259 SZ:57/57/1.00 VC:2 ENC:PLAIN,BIT_PACKED ST:[min: alpha1, max: alpha2, num_nulls: 0]
text:       BINARY UNCOMPRESSED D0:0 FP0:316 SZ:95/95/1.00 VC:2 ENC:PLAIN,RLE,BIT_PACKED ST:[min: Parquet Tweet1, max: Parquet Tweet2, num_nulls: 0]
```