

val arqData = sqlContext.read.format("com.databricks.spark.csv").option("header", "true").load("/home/scarface/ARQ_Data_Indicators.csv")

arqData: org.apache.spark.sql.Dataset[org.apache.spark.sql.Row] = [ticker: string, dimension: string ... 22 more fields]

Took 21 sec. Last updated by anonymous at January 29 2017, 9:12:25 PM.

arqData.printSchema

```
root
|-- ticker: string (nullable = true)
|-- dimension: string (nullable = true)
|-- calendardate: string (nullable = true)
|-- assets: string (nullable = true)
|-- assetsc: string (nullable = true)
|-- capex: string (nullable = true)
|-- cor: string (nullable = true)
|-- currentratio: string (nullable = true)
|-- ebit: string (nullable = true)
|-- ebt: string (nullable = true)
|-- eps: string (nullable = true)
|-- gp: string (nullable = true)
|-- grossmargin: string (nullable = true)
|-- intexp: string (nullable = true)
|-- liabilities: string (nullable = true)
|-- netinc: string (nullable = true)
|-- netmargin: string (nullable = true)
```

Took 0 sec. Last updated by anonymous at January 29 2017, 9:35:32 PM.

%r

arq <- read.csv(file = "/home/scarface/Desktop/sem-3/capstone/ARQ_Data_Indicators.csv")
save(arq, file="arq.rdata")
load("arq.rdata")
nrow(arq)

[1] 89206

Took 6 sec. Last updated by anonymous at January 29 2017, 9:43:42 PM.

%r

View(arq)

Took 13 sec. Last updated by anonymous at January 29 2017, 9:45:29 PM.

%spark.r

print(arq)

1	A	ARQ	31-03-2011	8044000000	4598000000
2	A	ARQ	30-06-2011	8649000000	5096000000

Zeppelin

Took 14 sec. Last updated by anonymous at January 29 2017, 9:48:47 PM.

READY ▶ ↻ 📖 ⚙️