* val wine\_data = spark.read.format("csv").option("header", "true").load("/tmp/spark\_data/winemag-data-130k-v2.csv")
* wine\_data.columns
* import org.apache.spark.sql.functions.{expr, col, column}
* val wine\_data\_cols = wine\_data.select(col("country"), col("description"), column("designation"), column("points"), column("price"), column("province"), column("region\_1"), column("title"), column("variety"), column("winery"))
* wine\_data\_cols.columns
* val wine\_data\_filter = wine\_data\_cols.filter(col("price").isNotNull)
* val wine\_data\_filter1 = wine\_data\_filter.filter(col("region\_1").isNotNull)
* val wine\_data\_filter2 = wine\_data\_filter1.filter(col("variety").isNotNull)