*# Import the necessary module*from pyspark.sql import \*  
  
*# Create a SparkSession object*spark = SparkSession.builder.appName('ReadParquet').getOrCreate()  
  
*#------------------------------------------------------------------------  
# Read a single Parquet file into a DataFrame using .parquet() method and print its schema and data*df = spark.read.parquet('C:/Users/hars2071/Downloads/DEPractice/user1.parquet')  
  
df.printSchema()  
df.show()

*Output:*

root

|-- Id: long (nullable = true)

|-- Name: string (nullable = true)

|-- Salary: long (nullable = true)

+---+-------+------+

| Id| Name|Salary|

+---+-------+------+

| 1| Harsha| 20000|

| 2|Mokshit| 30000|

| 3| Kiran| 40000|

+---+-------+------+

*#------------------------------------------------------------------------  
# Read multiple Parquet files into a DataFrame*df1 = spark.read.format('parquet').load(['C:/Users/hars2071/Downloads/DEPractice/user1.parquet','C:/Users/hars2071/Downloads/DEPractice/user2.parquet'])  
df1.show()

*Output:*

+---+-------+------+

| Id| Name|Salary|

+---+-------+------+

| 4| Harika| 25000|

| 5| Venkat| 35000|

| 6|Krishna| 25000|

| 1| Harsha| 20000|

| 2|Mokshit| 30000|

| 3| Kiran| 40000|

+---+-------+------+

*#------------------------------------------------------------------------  
# Read all parquet files into a DataFrame*df2 = spark.read.parquet('C:/Users/hars2071/Downloads/DEPractice/\*.parquet')  
df2.show()

*Output:*

+---+-------+------+

| Id| Name|Salary|

+---+-------+------+

| 4| Harika| 25000|

| 5| Venkat| 35000|

| 6|Krishna| 25000|

| 7| Harish| 25000|

| 8| Venkat| 35000|

| 9| Krish| 25000|

| 1| Harsha| 20000|

| 2|Mokshit| 30000|

| 3| Kiran| 40000|

+---+-------+------+