```
#Taks 2 Code for Filtering RDDs and Finding Min and Max Temperatures by Location

from pyspark import SparkConf, SparkContext

# Initialize SparkContext
if 'sc' in locals() or 'sc' in globals():
    sc.stop()

conf = SparkConf().setMaster("local[*]").setAppName("Min and Max Temperature Example")
sc = SparkContext(conf=conf)

# Sample data (location, temperature)
data = [("NY", 30), ("NY", 25), ("CA", 40), ("CA", 35), ("NY", 28), ("CA", 42)]

# Create RDD
rdd = sc.parallelize(data)

# Find minimum temperature by location
min_temps = rdd.reduceByKey(lambda x, y: min(x, y))

# Find maximum temperature by location
max_temps = rdd.reduceByKey(lambda x, y: max(x, y))

# Collect and print the results
print("Minimum Temperatures by Location:")
for location, temp in min_temps.collect():
    print(f"{location}: {temp}°C")

print("\nMaximum Temperatures by Location:")
for location, temp in max_temps.collect():
    print(f"{location}: {temp}°C")

# Stop SparkContext
sc.stop()
```

```
Minimum Temperatures by Location:
NY: 25°C
CA: 35°C

Maximum Temperatures by Location:
NY: 30°C
CA: 42°C
```

Start coding or generate with AI.