6

->

```
from mrjob.job import MRJob
from mrjob.step import MRStep

class WeatherDataStats(MRJob):
    def mapper(self, _, line):
        """Tokenizes each line and emits (word, 1)"""
        fields = line.split(",")
        if len(fields) == 3:
            year = fields[0].strip()
            temperature = float(fields[1].strip())
            yield year, temprature):

    def reducer(self, year, temperatures):
        """Aggregates = list(temperatures)
        temperatures = list(temperatures)
        avg_temp = sum(temperatures) / len(temprature)
        max_temp = max(temperatures)
        min_temp = min(temperatures)
        yield year = (avg_temp, max_temp, min_temp)
```

```
class filterTemperature(MRJob):

    def mapper(self, _, line):
"""filters the data with temperature > 30"""
        fields = line.split(",")

        if len(fields) == 3:

            year = fileds[0].strip()
            temperature = float(fields[1].strip()
             Other_data = fileds[2].strip()


            if temperature > 30.0:
                yield year, (temperature, other_data)


    def reducer(self, year, readings):
"""writes the filtered readings to a file"""
        with open(f"filtered_reading_{year}.txt",
                  "w") as f:

            for reading in readings:

                f.write(f"{year}, {reading[0]},
                        {reading[1]}\n")
```

```
if __name__ == "__main__":
    import sys
    if len(sys.argv) < 2:
        sys.argv.append("ncdc_data.txt")
    weatherDataStats.run()
    filterTemperature.run()
```

- ncdc_data.txt

```
2020, 23.4, Some location
2020, 30.5, Another location
2021, 33.7, Location x
2021, 28.9, Location y
```