

NOAA data



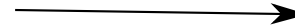
**Mapper**

Emit (station\_id+period,<temp,prec>)

**Job 1**

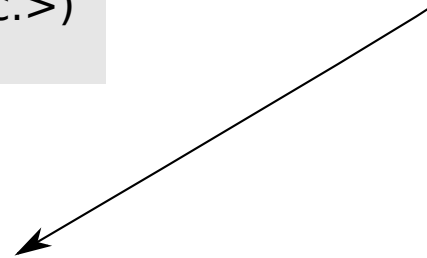
**Reducer**

Emit (station\_id+period, <avg. temp, total prec.>)



**Job 2**

Filter out data for  
reference year



**Mapper**

Emit (<station,year,month>,<period,temp,prec>)

**Job 3**

**Reducer**

Emit (<year,month>,distance to reference year)



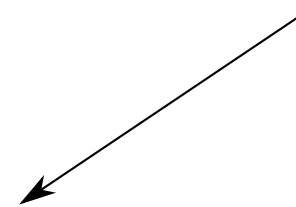
**Mapper**

Emit (<year,month>, distance>)

**Job 4**

**Reducer**

Emit (<year,month>, sum(distance)>)



Raw output