```
%%file report2.py
from mrjob.job import MRJob
from mrjob.step import MRStep
import re
WORD_RE = re.compile(r"[\w']+")
prime = None
class MRMostUsedWord(MRJob):
    def steps(self): # Overwrite the steps function to run a multi-step job
        return [
            MRStep(mapper=self.mapper_get_most_viewed_prime_program,
                    reducer=self.reducer_count_prime_viewers),
            MRStep(reducer=self.reducer_find_max_word),
        ]

    def mapper_get_most_viewed_prime_program(self, _, line):
        # yield each word in the line
        item = [word for word in WORD_RE.findall(line)]
        try:
          if 200000 <= int(item[3]) < 230000 and int(item[4])%2==0:
              yield (item[5], 1)
        except:
          pass

    def reducer_count_prime_viewers(self, word, counts):
        # send all (num_occurrences, word) pairs to the same reducer.
        # num_occurrences is so we can easily use Python's max() function.
        yield None, (sum(counts), word)

    # discard the key; it is just None
    def reducer_find_max_word(self, _, word_count_pairs):
        # each item of word_count_pairs is (count, word),
        # so yielding one results in key=counts, value=word
        # prime = max(word_count_pairs)
        yield max(word_count_pairs)




if __name__ == '__main__':
    MRMostUsedWord.run()
```

⊐⊽ Writing report2.py


```
! python report2.py 10k_view_data.csv
```

⊐⊽ No configs found; falling back on auto-configuration
    No configs specified for inline runner
    Creating temp directory /tmp/report2.root.20240620.202408.294438
    Running step 1 of 2...
    Running step 2 of 2...
    job output is in /tmp/report2.root.20240620.202408.294438/output
    Streaming final output from /tmp/report2.root.20240620.202408.294438/output...
    181   "SP003189730000"
    Removing temp directory /tmp/report2.root.20240620.202408.294438...