# wordfreq.py

def byFreq(pair):
    return pair[1]

def main():
    print("This program analyzes word frequency in a file")
    print("and prints a report on the n most frequent words.\n")

    # get the sequence of words from the file
    fname = raw_input("File to analyze: ")
    text = open(fname, 'r').read()
    text = text.lower()
    for ch in '!"#%&()*+,-./;<>?@[\]^_`{|}~':
        text = text.replace(ch, '')
    words = text.split()

    # construct a dictionary of word counts
    counts = {}
    for w in words:
        counts[w] = counts.get(w, 0) + 1

    # output analysis of n most frequent words.
    n = input("Output analysis of how many words? ")
    items = list(counts.items())
    items.sort()
    items.sort(key=byFreq, reverse=True)
    print("The number of unique words in", fname, "is", len(counts), ".")
    for i in range(n):
        word, count = items[i]
        print("{0:<15}{1:>5}".format(word, count))

if __name__ == '__main__':
    main()