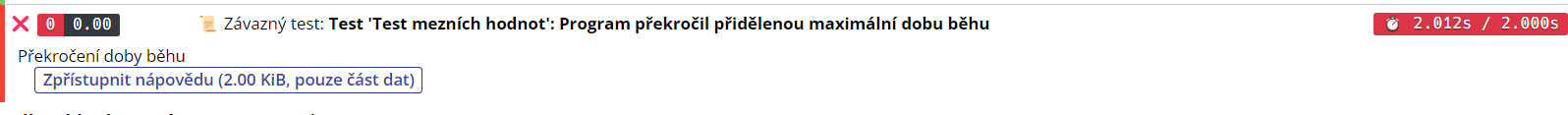
Info k tasku:  
dole je celý text zadání domácího úkolu. Základní testy jsem provedl kód je z 80-90% hotov. Problémy jsou následující:  
 komentář se musí uložit v 4096 znaků. Všechno, co je po tom se ignoruje

poté, co by se fixnulo načítání zpráv s dlouhými texty, ujistit se, že další vstupy (+,?,#)

dalším problémem je, že kód je pravděpodobně nezoptimalizovaný. Teda taky dalším úkolem je zoptimalizovat kód



otestovat kód nejenom na mezní hodnoty, ale i s náhodnými daty

Kód lze poskytnout po kontaktování.

=============================================================================================

Úkolem je vytvořit program, který bude pomáhat s úpravami statistiky vhodným výběrem dat.

Předpokládáme, že v e-shopu zaznamenáváme recenze našich zákazníků. Recenze přicházejí postupně. U každé recenze máme zadané datum, vlastní číselné hodnocení (celé kladné číslo) a text recenze. Data na vstupu jsou v neklesajícím pořadí, během jednoho dne ale může být zadáno více recenzí. Příkladem recenze na vstupu je:

+ 2023-11-10 10 Short\_review

Zde je dne 10.11.2023 vložena recenze s hodnocením 10 a komentářem Short\_review (komentář neobsahuje bílé znaky a má délku nejvýše 4096 znaků). Další recenze musí mít datum 10.11.2023 nebo pozdější (ne dřívější).

Při vytváření statistiky chceme najít souvislý úsek recenzí od nějakého počátečního dne do nějakého koncového dne takový, že nalezený interval obsahuje nejméně zadaný počet recenzí a medián hodnocení je v tomto období co nejvíce odlišný od průměru hodnocení v tomto období. Tedy například pro vstup:

? 100

chceme najít neprázdný časový interval takový, aby obsahoval alespoň 100 recenzí a průměr hodnocení z těchto recenzí byl co nejvzdálenější mediánu hodnocení v těchto recenzích. Takových intervalů může existovat obecně více se stejným rozdílem mezi mediánem a průměrem. V takovém případě budeme považovat za výsledek interval s pozdějším koncovým datem (viz druhý ukázkový běh, druhý dotaz). Pokud by i pro stejné koncové datum existovalo více stejně vhodných intervalů, dáme přednost tomu intervalu, kde je více recenzí (druhý ukázkový běh, třetí dotaz).

Při hledání intervalu pracujeme s granularitou jednoho dne. Pokud je v počátečním/koncovém dni nalezeného intervalu zadáno více recenzí, musíme je do výsledku zahrnout všechny (nelze použít pouze některé z recenzí počátečního/koncového dne). Toto je vidět v ukázkovém běhu na prvním dotazu. Nalezená odpověď je celý den 16.11 (medián 42, průměr 70), přestože vzdálenější je medián od průměru pro první dvě zadané recenze (medián 10, průměr 41.5, ale nezahrnuje poslední recenzi z 10.11).

Na vstupu je prokládané zadávání recenzí a požadavků na vyhledání intervalu. Vyhledávání se vyhodnocuje okamžitě, tedy pracuje pouze s dosud zadanými recenzemi (je to vidět v prvním ukázkovém běhu).

Výstupem programu jsou vyhledané intervaly. Program nejprve zobrazí časový interval (počáteční datum a koncové datum), za kterým následuje průměrné hodnocení a medián v tomto intervalu. Na dalších řádcích pak program zobrazí hodnocení a komentář recenzí z nalezeného intervalu; recenze jsou zobrazené v pořadí ve kterém byly zadány na vstupu.

Vyhledávání má dvě varianty. Pro zadání typu ? 100 chceme nalézt a vypsat časový interval, průměr hodnocení, medián hodnocení a odpovídající recenze. Pro zadání typu # 100 chceme nalézt a vypsat pouze časový interval, průměr hodnocení a medián hodnocení (nechceme vypisovat jednotlivé recenze).

Pokud je vstup neplatný, program to musí detekovat a zobrazit chybové hlášení. Chybové hlášení zobrazujte na standardní výstup (ne na chybový výstup). Za chybu považujte:

* chybný formát zadání recenze nebo chybný formát vyhledávání,
* neplatné datum v recenzi,
* datum v zadání recenze je menší než datum v předchozí recenzi,
* nečíselná, nulová nebo záporná hodnota hodnocení v recenzi (hodnocení musí být alespoň 1),
* nečíselná, nulová nebo záporná hodnota minimálního počtu recenzí ve vyhledávání,
* pokus o vyhledávání bez dosud zadaných recenzí.

Váš program bude spouštěn v omezeném testovacím prostředí. Je omezen dobou běhu (limit je vidět v logu referenčního řešení) a dále je omezena i velikost dostupné paměti. Text recenze může být dost dlouhý (limit je 4096 znaků), navíc počet recenzí na vstupu není omezen. Pro uložení dat budete potřebovat dynamickou alokaci paměti.

Poznámky:

* Ukázkové běhy zachycují očekávané výpisy Vašeho programu (tučné písmo) a vstupy zadané uživatelem (základní písmo). Zvýraznění tučným písmem je použité pouze zde na stránce zadání, aby byl výpis lépe čitelný. Váš program má za úkol pouze zobrazit text bez zvýrazňování (bez HTML markupu).
* Znak odřádkování (\n) je i za poslední řádkou výstupu (i za případným chybovým hlášením).
* Pro rozhodování o platných datech použijte platná pravidla Gregoriánského kalendáře. Na vstupu jsou zadávané roky v rozumném rozsahu (1800 - 3000).
* Medián je jednoznačně definovaný pro lichý počet recenzí. Pokud analyzovaný interval obsahuje sudý počet recenzí, považujeme za medián menší ze dvou prostředních čísel. Tedy pro hodnocení 6 15 7 je mediánem 7, pro hodnocení 6 15 7 3 budeme za medián považovat hodnotu 6.
* Výsledný interval musí obsahovat alespoň jednu recenzi, pro méně recenzí není definovaný ani medián ani průměr.
* Pro nalezený interval požadujeme, aby první i poslední den intervalu obsahoval nějakou recenzi.
* Úloha nabízí bonusové testy pro algoritmicky efektivní řešení. Body navíc dostane řešení, které dokáže rychle zpracovat dotazy pro dlouhé seznamy recenzí. V testech rychlosti jsou zadávané převážně dotazy typu # (výpis recenzí velmi zdržuje).
* Řešení této úlohy, které projde všemi závaznými a nepovinnými testy na 100%, může být použito pro code review (pro code review není požadované zvládnutí testu rychlosti).

Ukázka práce programu:

**Recenze:**

+ 2023-11-10 10 Short\_review

+ 2023-11-10 73 Long\_review

+ 2023-11-10 15 Negative\_review

+ 2023-11-11 45 Some\_text

+ 2023-11-16 42 Useless\_comment

+ 2023-11-16 98 Fake\_review

? 1

**2023-11-16 - 2023-11-16: 70.000000 42**

**42: Useless\_comment**

**98: Fake\_review**

? 2

**2023-11-16 - 2023-11-16: 70.000000 42**

**42: Useless\_comment**

**98: Fake\_review**

? 6

**2023-11-10 - 2023-11-16: 47.166667 42**

**10: Short\_review**

**73: Long\_review**

**15: Negative\_review**

**45: Some\_text**

**42: Useless\_comment**

**98: Fake\_review**

? 20

**Neexistuje.**

# 1

**2023-11-16 - 2023-11-16: 70.000000 42**

# 2

**2023-11-16 - 2023-11-16: 70.000000 42**

+ 2023-11-16 15 Updated\_review

+ 2023-11-17 120 Sponsored

# 6

**2023-11-10 - 2023-11-17: 52.250000 42**

+ 2023-11-17 1 Not\_sponsored

# 6

**2023-11-11 - 2023-11-17: 53.500000 42**

+ 2023-11-17 1 Not\_sponsored

# 6

**2023-11-16 - 2023-11-17: 46.166667 15**

+ 2023-11-17 1 Not\_sponsored

# 6

**2023-11-11 - 2023-11-17: 40.375000 15**

+ 2023-11-17 1800 Sponsored

# 6

**2023-11-16 - 2023-11-17: 259.750000 15**

+ 2023-11-17 1800 Sponsored

# 6

**2023-11-17 - 2023-11-17: 620.500000 1**

+ 2023-11-17 1800 Sponsored

# 6

**2023-11-17 - 2023-11-17: 789.000000 120**

+ 2023-11-17 1801 Sponsored

# 6

**2023-11-17 - 2023-11-17: 915.500000 120**

+ 2023-11-17 1802 Sponsored

# 6

**2023-11-17 - 2023-11-17: 1014.000000 1800**

+ 2023-11-17 1803 Sponsored

# 6

**2023-11-16 - 2023-11-17: 852.615385 120**

+ 2023-11-17 1804 Sponsored

# 6

**2023-11-16 - 2023-11-17: 920.571429 120**

+ 2023-11-17 1805 Sponsored

# 6

**2023-11-16 - 2023-11-17: 979.533333 1800**

+ 2023-11-17 1806 Sponsored

# 6

**2023-11-11 - 2023-11-17: 973.176471 1800**

+ 2023-11-17 180000 Bribed

# 6

**2023-11-17 - 2023-11-17: 14024.571429 1800**

+ 2023-11-17 180000000 Hacked

# 6

**2023-11-17 - 2023-11-17: 12013089.600000 1801**

**Recenze:**

+ 2023-11-10 10 Short\_review

+ 2023-11-11 20 Short\_review

# 1

**2023-11-10 - 2023-11-11: 15.000000 10**

+ 2023-11-12 15 Short\_review

+ 2023-11-13 10 Short\_review

+ 2023-11-13 20 Short\_review

# 1

**2023-11-13 - 2023-11-13: 15.000000 10**

+ 2023-11-14 10 Short\_review

+ 2023-11-14 20 Short\_review

# 1

**2023-11-13 - 2023-11-14: 15.000000 10**

# 5

**2023-11-11 - 2023-11-14: 15.833333 15**

+ 2023-11-15 15 Short\_review

+ 2023-11-16 10 Short\_review

+ 2023-11-16 20 Short\_review

# 1

**2023-11-16 - 2023-11-16: 15.000000 10**

# 5

**2023-11-11 - 2023-11-14: 15.833333 15**

**Recenze:**

+ 2023-11-10 10 A\_review\_with\_long\_text

? 1

**2023-11-10 - 2023-11-10: 10.000000 10**

**10: A\_review\_with\_long\_text**

+ 2023-11-09 10 A\_review\_with\_short\_text

**Nespravny vstup.**

**Recenze:**

+ 2000-02-29 10 Review\_#1

+ 2004-02-29 10 Review\_#2

+ 2020-02-29 10 Review\_#3

? 1

**2000-02-29 - 2020-02-29: 10.000000 10**

**10: Review\_#1**

**10: Review\_#2**

**10: Review\_#3**

+ 2023-02-28 10 Review\_#4

? 1

**2000-02-29 - 2023-02-28: 10.000000 10**

**10: Review\_#1**

**10: Review\_#2**

**10: Review\_#3**

**10: Review\_#4**

+ 2023-02-29 10 Review

**Nespravny vstup.**

**Recenze:**

+ 2100-02-29 10 Review\_#1

**Nespravny vstup.**

**Recenze:**

+ 2000-02-29 0 Review\_#1

**Nespravny vstup.**