بسم الله الرحمن الرحيم

في البداية تم إنشاء ملف باسم count.sh

يحتوي على الأوامر التالية:

|  |
| --- |
| awk -F ',' '{print $NF}' imdb\_dataset.csv > sentiment.txt |

يقوم هذا الأمر باستخراج عامود ال sentiment

من الملف وحفظه في ملف نصي باسم العامود sentiment

|  |  |
| --- | --- |
| -F | يتم تحديد الفاصلة بين العامودين |
| {print $NF} | يتم طباعة العامود الاخير من الملف |
| > sentiment.txt | يتم توجية النتائج للملف |

بعد تنفيذ الأمر السابق ينشي ملف باسم sentiment.txt

يحتوي الملف sentiment.txt على قيمتين اما positive أو negative

لذى يتم استخدام الأمرين

|  |
| --- |
| positive\_count="$(grep -c positive sentiment.txt)" |
| negative\_count="$(grep -c positive sentiment.txt)" |

للحصول على عدد تكرار كلمة positive و وضع العدد في المتغير positive\_count

نفس العمل في الامر الثاني لكن مع كلمة negative

الامر grep يقوم بالبحث عن الكلمة و من خلال الامر -c يقوم بإرجاع عدد تكرارها في الملف

للحصول على عدد الكلمات الفريدة تم استخدام الامر التالي و وضعة في متغير باسم unique\_words

|  |
| --- |
| unique\_words="$(cat imdb\_dataset.csv| sed -e 's/<[^>]\*>//g' | tr '[:upper:]' '[:lower:]' | tr -d '[:punct:]' | tr '[:space:]' '[\n\*]' | sort | uniq -c | awk -F ' ' '{if($2 ~ /^[A-Za-z0-9]/) print $2}' | wc -l)" |

|  |  |
| --- | --- |
| cat imdb\_dataset.csv | احضار الملف imdb\_dataset.txt ليتم تطبيق الأوامر عليه |
| sed -e 's/<[^>]\*>//g' | يقوم الامر sed بحذف وسوم ال html الموجودة في الملف |
| tr '[:upper:]' '[:lower:]' | بإستخدام الامر tr يتم تحويل الحروف الكبيرة إلى حروف صغيرة |
| tr -d '[:punct:]' | حذف علامات الترقيم |
| tr '[:space:]' '[\n\*]' | الامر tr يقوم باستبدال المسافات بسطر جديد |
| sort | ترتيب الكلمات |
| uniq -c | يقوم الأمر uniq بطباعة الكلمات الفريدة في الملف مع عدد تكرارها |
| awk -F ' ' '{if($2 ~ /^[A-Za-z0-9]/) print $2}' | بعد تنفيذ الأوامر أعلاه يتم الحصول على نتيجة تحتوي على عامود بالكلمات و عامود بعدد تكرارها بحيث انه لا يوجد تكرار في الكلمات  يقوم الامر awk بطباعة الكلمات التي تحتوي على حروف و أرقام باستخدام النمط /^[A-Za-z0-9]/ |

الامر الأخير هو لطباعة المتغيرات على الشاشة

|  |
| --- |
| echo "Positive Count = $positive\_count"  echo "Negative Count = $negative\_count"  echo "Total Unique Words = $unique\_words" |

نهاية ملف count.sh

يتم حفظ النتيجة في ملف analysis.txt من سطر الأوامر وذلك باستخدام الامر التالي:

|  |
| --- |
| bash count.sh > analysis.txt |

ولعرض محتويات الملف analysis.txt على الشاشة نستخدم الأمر التالي

|  |
| --- |
| cat analysis.txt |