Part#1: Regular Expression (RegEx):

- هو مجموعة من chars الي بتساعدنا نبحث عن محتوى او اي نمط سواء كان معقد او بسيط
- وبما انه بدنا نبحث عن محتوى او نمط ف رح نستخدم **grep**
- تجربة اليوم مليانه grep، للتذكير تكتب هكذا grep "pattern/word" file_name
- رح نتعلم في هذه التجربة هاي chars ونشوف ايش معناهم وكيف منستخدمهم
- نقوووول بسم الله ونبلششش
- رح استخدم كلمة ماتش او **match** كثير ومعناها بكون تطابق

رح اشتغل على نفس محتوى الملف الموجود بالمانيوال

```
ahed@DESKTOP-OK5G6FV:~/ahed$ cat exp3.txt
by
Ken
Thompson and Dennis
Ritchie at Bell
Laboratories in the late 1960s. Onee
of the primary
goals in the design of the Unix system was to create an environment that
promoted efficient program
development.
[my name]
.Hi man
HHi man
my mamamamamama my
mohammadm
ahed@DESKTOP-OK5G6FV:~/ahed$
```

1- Period or Dot (.):

- بتعمل ماتش لاي حرف بغض النظر عن ايش هو الحرف ما عدا new line char

```
ahed@DESKTOP-OK5G6FV:~/ahed$ grep '.' exp3.txt
by
Ken
Thompson and Dennis
Ritchie at Bell
Laboratories in the late 1960s. Onee
of the primary
goals in the design of the Unix system was to create an environment that
promoted efficient program
development.
[my name]
.Hi man
HHi man
my mamamamamama my
mohammadm
ahed@DESKTOP-OK5G6FV:~/ahed$
```

- عملت ماتش لكل الملف لانه بكل بساطة كل الملف عبارة عن char حتى الرقم يعتبر char عشان هيك بحكولها wildcard pattern
- وبقدر احدد مثلا انه بدي اعمل ماتش لاي شي فيه حرف a ويتبعه اي حرف

```
ahed@DESKTOP-OK5G6FV:~/ahed$ grep 'a.' exp3.txt
Thompson and Dennis
Ritchie at Bell
Laboratories in the late 1960s. Onee
of the primary
goals in the design of the Unix system was to create an environment that
promoted efficient program
[my name]
.Hi man
HHi man
my mamamamamama my
mohammadm
```

- وهان بدي اكون عارف اني ببحث عن نمط pattern وليس عن كلمة word
- الفرق بينهم انه النمط ممكن يكون معه كمان اشياء ثانية يعني بكون هو substring اما الكلمة بتكون هي string لحاله يعني بكون قبله وبعد فراغ space
- طب كيف بدي ابحث عن كلمة word ، يعني مثلا بدي كل الكلمات الي فيها حرف a ويتبعها اي حرف؟ لازم استخدم الشكل التالي \<\.a\>

```
ahed@DESKTOP-OK5G6FV:~/ahed$ grep '\<a.\>' exp3.txt
 Ritchie at Bell
 goals in the design of the Unix system was to create an environment that
```

- وفي الصورة التالية رح يبحث عن اي سطر يحتوي على كلمة فيها الاحرف Bel ويتبعها اي حرف

```
ahed@DESKTOP-OK5G6FV:~/ahed$ grep '\<Bel.\>' exp3.txt
 Ritchie at Bell
ahed@DESKTOP-OK5G6FV:~/ahed$
```

- من المفضل استخدام double quotation ("") مع اي شيء تريد البحث عنه
- ونستطيع ان نبحث عن كلمة بطول معين، مثلا نبحث عن الاسطر الي فيها كلمة تبدأ بحرف بحرف p وتنتهي بحرف m ويكون طولها (عدد الاحرف) 7

```
ahed@DESKTOP-OK5G6FV:~/ahed$ grep "\<p.....m\>" exp3.txt
promoted efficient program
ahed@DESKTOP-OK5G6FV:~/ahed$
```

2- escape char (\):

- اذا بدك تبحث عن حرف (.)، ما بزبط ابحث عنه مباشرة زي هيك:

  grep “.” exp3.txt
- لانه (.) لها معنى واستخدام خاص في RegEx، ف اذا بدنا نبحث عن (.) لازم استخدم \ والي اسمها escape char

```
ahed@DESKTOP-OK5G6FV:~/ahed$ grep "\." exp3.txt
Laboratories in the late 1960s. Onee
development.
.Hi man
```

3- Caret (^):

- بتعمل ماتش لبداية السطر، بستخدمها اذا مثلا بدي الاسطر الي بتبلش بشي معين.

```
ahed@DESKTOP-OK5G6FV:~/ahed$ grep "^m" exp3.txt
my mamamamamama my
mohammadm
ahed@DESKTOP-OK5G6FV:~/ahed$
```

- فمثلا هان عملنا ماتش لكل الاسطر الي بتبلش بحرف m

Q) print the lines that start with Dot (.)

4- Dollar Sign ($):

- تستخدم لعمل ماتش لاي شي بنهاية السطر
- مثلا الاسطر الي تنتهي ب حرف (y)

```
ahed@DESKTOP-OK5G6FV:~/ahed$ grep "y$" exp3.txt
by
of the primary
my mamamamamama my
ahed@DESKTOP-OK5G6FV:~/ahed$
```

Q)

- print the lines that end with “man” word.
- Print the blank lines (lines contains no characters)

5- Choice of characters ([...]):

- تستخدم لعمل ماتش لاحد ال chars ضمن range معين
- مثلا بدي اعمل ماتش ل T او t ، ف لازم احطهم هيك [Tt]
- في المثال الاتي رح ابحث عن كلمةthe سواء كانت هيك the او هيك The

```
ahed@DESKTOP-OK5G6FV:~/ahed$ grep "\<[Tt]he\>" exp3.txt
The Ken
Laboratories in the late 1960s. Onee
of the primary
goals in the design of the Unix system was to create an environment that
ahed@DESKTOP-OK5G6FV:~/ahed$
```

- في المثال الاتي رح اعمل ماتش للاسطر الي فيهم رقم (بغض النظر عن قيمة الرقم)، [1234567890] او بشكل مختصر اكثر [0-9]

```
ahed@DESKTOP-OK5G6FV:~/ahed$ grep "[1234567890]" exp3.txt
Laboratories in the late 1960s. Onee
ahed@DESKTOP-OK5G6FV:~/ahed$ grep "[0-9]" exp3.txt
Laboratories in the late 1960s. Onee
ahed@DESKTOP-OK5G6FV:~/ahed$
```

Q) print the lines that start with any capital letter.

- رينج الاحرف [A-Z] او [a-z] او بتقدر تدمجهم مع بعض [A-Za-z]
- [-0-9]: match a dash (-) or a digit
- []0-9]: match a right bracket or a digit
- ملاحظة مهمة: اذا استخدمنا (^) داخل range ([ ])، يعني زي هيك [^] بصير معناها انه اي شي ما عدا الموجود في range
- مثلا بدي اعمل ماتش لاي chars ما عدا الارقام => [0-9^]

```
ahed@DESKTOP-OK5G6FV:~/ahed$ grep "[^0-9]" exp3.txt
by
The Ken
Thompson and Dennis
Ritchie at Bell
Laboratories in the late 1960s. Onee
of the primary
goals in the design of the Unix system was to create an environment that
promoted efficient program
development.
[my name]
.Hi man
HHi man
my mamamamamama my
mohammadm
ahed@DESKTOP-OK5G6FV:~/ahed$
```

Q) remove the line that contains digits.

6- Asterisk (*):

- تعمل ماتش لحرف مكرر صفر او اكثر من مرة
- بستخدمها اذا في حرف معين او اي حرف متكرر مرة او اكثر
- مثلا بدي اعمل ماتش لحرف H ولازم يكون موجود عالاقل مرة واحدة :

```
ahed@DESKTOP-OK5G6FV:~/ahed$ grep  "HH*" exp3.txt
.Hi man
HHi man
ahed@DESKTOP-OK5G6FV:~/ahed$
```

- طب خلينا نشوف مثال ثاني ع محتوى ثاني مثلا بدي ابحث عن حرف a ويتبعه حرف p مرة او مرتين او 3 مرات او ..... او صفر مرة

=> يعني رح تستخرج كل هاي الانماط : a... او ap.... او app.... او الخ

```
ahed@DESKTOP-OK5G6FV:~/ahed$ cat fruit.txt
apple
pinapple
banana
orange
ahed@DESKTOP-OK5G6FV:~/ahed$ grep "ap*" fruit.txt
apple
pinapple
banana
orange
ahed@DESKTOP-OK5G6FV:~/ahed$
```

- ملاحظة مهمة : اذا استخدمنا * مع [ ] بتلغي معناها في regex:

```
ahed@DESKTOP-OK5G6FV:~/ahed$ cat test.txt
apple*
tab
ttab
banan
ahed@DESKTOP-OK5G6FV:~/ahed$ grep "[t*]" test.txt
apple*
tab
ttab
ahed@DESKTOP-OK5G6FV:~/ahed$
```

- هاي بصير معناها ابحث عن الاسطر الي فيها t او *

Q) why are all contents printed in the following image?

```
ahed@DESKTOP-OK5G6FV:~/ahed$ grep  "H*" exp3.txt
by
The Ken
Thompson and Dennis
Ritchie at Bell
Laboratories in the late 1960s. Onee
of the primary
goals in the design of the Unix system was to create an environment that
promoted efficient program
development.
[my name]
.Hi man
HHi man
my mamamamamama my
mohammadm
ahed@DESKTOP-OK5G6FV:~/ahed$
```

7- Matching a Precise Number of Characters: \{start,end\}:

- تستخدم لعمل ماتش لمجموعة من ال chars المتكررة ضمن range انا بحدد بدايته ونهايته او اذا بدي عدد معين بالزبط
- مثلا بدي ابحث عن الاسطر الي فيها من 1 - 8 من حرف (l)

```
ahed@DESKTOP-OK5G6FV:~/ahed$ grep "l\{1,8\}" exp3.txt
Ritchie at Bell
Laboratories in the late 1960s. Onee
goals in the design of the Unix system was to create an environment that
development.
ahed@DESKTOP-OK5G6FV:~/ahed$
```

- وبتقدر تحدد عدد معين مثلا بدي ابحث عن الاسطر الي فيها حرفين H

```
ahed@DESKTOP-OK5G6FV:~/ahed$ grep "H\{2\}" exp3.txt
HHi man
ahed@DESKTOP-OK5G6FV:~/ahed$
```

- وممكن احدد بداية range وما احدد النهاية يعني ما بتكون النهاية رقم محدد :

```
ahed@DESKTOP-OK5G6FV:~/ahed$ grep "a\{1,\}" exp3.txt
Thompson and Dennis
Ritchie at Bell
Laboratories in the late 1960s. Onee
of the primary
goals in the design of the Unix system was to create an environment that
promoted efficient program
[my name]
.Hi man
HHi man
my mamamamamama my
mohammadm
```

- ف هان حرف a لازم يكون موجود ع الاقل مرة واحدة ، بتشبه هاي "aa*"

Q) print the lines that contains word with length = 4 .

9-Saving Matched Characters: \(...\):

- لفرضا انا بدي اعمل ماتش للاسطر الي بتبلش وبتنتهي بحرف معين او كلمة معينة
- ما بزبط اعلمها باستخدام الي تعلمناه قبل بس، فبحتاج شي يخزن الماتش وارجع اقارن اي ماتش ثاني بهذا الماتش الي خزنته
- اطلع مثلا بدي اطلع الاسطر الي بتبلش بحرف وبتنتهي بنفس الحرف ، ف لازم اجيب اول حرف واخزنه واقول بدي السطر الي بنتهي بهذا الحرف الي خزنته
- ف عشان اعمل ماتش لحرف او لاي شي واخزنه في اشي اسمه ريجستر registers
- عدد هذول الريجسترز هو 9
- عشان اقدر اخزن الماتش بستخدم هذا \(pattern/word/char \)
- اول ماتش بحفظه ب ريجستر 1 واذا عملت كمان ماتش بحفظه في ريجستر 2 وهكذا ....
- عشان اجيب قيمة الريجستر بحط هيك \n، وتكون n تمثل رقم الريجستر
- مثلا بدي اجيب قيمة الماتش في ريجستر 1 بكتب \1
- طب خلينا نوخذ مثال انه نطبع الاسطر الي فيها اول حرفين متشابهين :

```
ahed@DESKTOP-OK5G6FV:~/ahed$ grep "^\(.\)\1" exp3.txt
HHi man
ahed@DESKTOP-OK5G6FV:~/ahed$
```

- هاي معناها جيب اول char (^) وخزنه بريجستر \(\) طبعا ريجستر 1 لانه اول مرة بخزن في ريجستر ولازم يكون الحرف الثاني نفس الحرف الاول الي خزنته بالريجيستر الاول \1 .
- كمان مثال بدي اطبع السطر الي ببلش وبنتهي بنفس الحرف:

```
ahed@DESKTOP-OK5G6FV:~/ahed$ grep "^\(.\).*\1$" fruit.txt
aplle can be banana
ahed@DESKTOP-OK5G6FV:~/ahed$
```

- معناها جيب اول حرف وخزنه بالريجستر الاول وبغض النظر عن الاحرف الي بعده .* ولازم يكون السطر بنتهي بهذا الحرف $ \1
- طبعا اذا استخدمت (\)\ مرتين معناها انا بستخدم بالريجستر الاول والثاني وهكذا...

Part#2: Sed command:

- بستخدمها عشان اعدل على محتوى الملف او طباعة اسطر معينة
- لا تعدل على الملف نفسه وانما تطبع على التيرمنال فقط
- لها 3 اشكال كالتالي:

1- sed “s/old_dat/new_data” file_name

```
ahed@DESKTOP-OK5G6FV:~/ahed$ cat file.txt
The Unix operating system was pioneered by Ken
Thompson and Dennis Ritchie at Bell Laboratories
in the late 1960s. One of the primary goals in
the design of the Unix system was to create an
environment that promoted efficient program development.
ahed@DESKTOP-OK5G6FV:~/ahed$ sed "s/Unix/UNIX/" file.txt
The UNIX operating system was pioneered by Ken
Thompson and Dennis Ritchie at Bell Laboratories
in the late 1960s. One of the primary goals in
the design of the UNIX system was to create an
environment that promoted efficient program development.
```

- طبعا بتمسك الملف سطر سطر وبتعمل التعديلا
- لكن بتعمل تعديل لاول ماتش في السطر، ف مثلا اذا كانت كمان كلمة Unix في السطر الاول لا يحدث عليها التعديل الا اذا استخدمنا (g (global option

```
ahed@DESKTOP-OK5G6FV:~/ahed$ sed "s/Unix/UNIX/" file.txt
The UNIX operating system was pioneered by Ken Unix
Thompson and Dennis Ritchie at Bell Laboratories
in the late 1960s. One of the primary goals in
the design of the UNIX system was to create an
environment that promoted efficient program development.
ahed@DESKTOP-OK5G6FV:~/ahed$ sed "s/Unix/UNIX/g" file.txt
The UNIX operating system was pioneered by Ken UNIX
Thompson and Dennis Ritchie at Bell Laboratories
in the late 1960s. One of the primary goals in
the design of the UNIX system was to create an
environment that promoted efficient program development.
```

2- sed -n “line numbers or word p”:

- تستخدم لطباعة الاسطر التي تحددها بالرقم او الاسطر التي تحتوي على كلمة معينة:

```
ahed@DESKTOP-OK5G6FV:~/ahed$ sed -n "1,5p" file.txt
The Unix operating system was pioneered by Ken Unix
Thompson and Dennis Ritchie at Bell Laboratories
in the late 1960s. One of the primary goals in
the design of the Unix system was to create an
environment that promoted efficient program development.
```

- هان طبع من سطر 1 ل سطر 5

```
ahed@DESKTOP-OK5G6FV:~/ahed$ sed -n "/Unix/p" file.txt
The Unix operating system was pioneered by Ken Unix
the design of the Unix system was to create an
ahed@DESKTOP-OK5G6FV:~/ahed$
```

- وهان طبع الاسطر التي تحتوي على Unix ك pattern
- اذا بدي word الصح استخدم regex

3- sed “line number or word d” :

- لحذف الاسطر من خلال رقمها او اذا كانت تحتوي على كلمة معينة:

```
ahed@DESKTOP-OK5G6FV:~/ahed$ sed "1,3d" file.txt
the design of the Unix system was to create an
environment that promoted efficient program development.
```

- هنا حذف الاسطر من 1 - 3
- لحذف الاسطر التي تحتوي على كلمة معينة :

```
ahed@DESKTOP-OK5G6FV:~/ahed$ sed "/Unix/d" file.txt
Thompson and Dennis Ritchie at Bell Laboratories
in the late 1960s. One of the primary goals in
environment that promoted efficient program development.
```

- وهنا حذف الاسطر التي تحتوي على Unix ك pattern
- اذا بدي اياها ك word الصح انه استخدم regex يعني هيك /<word/>

sed vs tr:

- tr: تستخدم مع الاحرف
  - مثلا اذا حطيت اشي زي هيك

tr “dog” “cat”

- معناها بدل حرف d ب c وحرف o ب a وحرف g ب t ، وليس ال pattern كاملة
- اذا كان قصدك تبدل pattern الي هي dog ب cat لازم تستخدم sed

- وانتبه انه حكينا pattern مش word ، اذا بدك word لازم تستخدم regex

Task:
Print the lines that contains the following word pattern:
Word1 ……. Word2 …. Word1 …. Word2
Note: create your file content

Good Luck