

Bài tập Bashscript: Cho 1 file excel có các column: Số thứ tự, tên, ngày sinh nhật, quê quán, điểm

- Đọc dữ liệu từ file excel
- lọc ra thông tin các dòng có ngày sinh nhật đúng format yyyy-mm-dd
- Chuẩn hóa lại tên theo đúng formart: Aaaa Bbbb Cccc

1. Tạo 1 random file chứa các thông tin về Số thứ tự, tên, ngày sinh, quê quán.

- BashScript tạo random file csv.

```
nhat@nhat-VirtualBox: ~/bash_xlsx
#!/usr/bin/bash

NUM_=100

NAMES=("Nguyen Van Aaa" "Nguyen thi bbbb" "tran van Cccc" "hoang van ddddd" "Huy
nh thi Eeee" "truong van hhhh" "Luu van Gggg" "Le Thi Bbbb")

PLACES=("Ha Noi" "Sai Gon" "Hue" "Da Nang" "Hai Phong" "Nha Trang" "Dong Nai" "V
ung Tau" "Quang Ninh")

OUTPUT="data.csv"
echo "STT, HoTen, Ngay sinh, Que Quan" > $OUTPUT

generate_random_date(){
    year=$((RANDOM%30 + 1970))
    month=$((RANDOM%12 + 1))
    day=$((RANDOM%28+1))
    printf "%d-%d-%d" $year $month $day
}

for ((i=1; i<=$NUM_; i++)); do
    name=${NAMES[$RANDOM % ${#NAMES[@]}]}
    date=$(generate_random_date)
```

```
    place=${PLACES[$RANDOM % ${#PLACES[@]}]}
    echo "$i, $name, $date, $place" >> $OUTPUT
done
```

- Thực thi BashScript và Convert file csv sang xlsx

```

nhat@nhat-VirtualBox:~/bash_xlsx$ vim data.csv
nhat@nhat-VirtualBox:~/bash_xlsx$ vim create_xlsx.sh
nhat@nhat-VirtualBox:~/bash_xlsx$ ./create_xlsx.sh
nhat@nhat-VirtualBox:~/bash_xlsx$ ls
create_xlsx.sh  data.csv  data.xlsx
nhat@nhat-VirtualBox:~/bash_xlsx$ soffice --convert-to xlsx:"Calc MS Excel 2007
XML" data.csv --headless
convert /home/nhat/bash_xlsx/data.csv as a Calc document -> /home/nhat/bash_xlsx
/data.xlsx using filter : Calc MS Excel 2007 XML
Overwriting: /home/nhat/bash_xlsx/data.xlsx
nhat@nhat-VirtualBox:~/bash_xlsx$ ls
create_xlsx.sh  data.csv  data.xlsx
nhat@nhat-VirtualBox:~/bash_xlsx$

```

	A	B	C	D	E	F	G	H
1	STT	HọTen	Ngày sinh	Que Quan				
2	1	Nguyen Van Aaa	1975-3-3	Dong Nai				
3	2	hoang van dddd	1999-1-24	Vung Tau				
4	3	Le Thi Bbbb	1976-12-20	Hue				
5	4	truong van hhhh	1998-6-26	Ha Noi				
6	5	Huynh thi Eeee	1977-9-20	Dong Nai				
7	6	Nguyen Van Aaa	1984-4-12	Hue				
8	7	truong van hhhh	1978-5-1	Quang Ninh				
9	8	Nguyen Van Aaa	1975-3-17	Hai Phong				
10	9	Huynh thi Eeee	1970-1-11	Hai Phong				
11	10	Le Thi Bbbb	1994-3-24	Dong Nai				
12	11	hoang van dddd	1998-2-21	Sai Gon				
13	12	tran van Cccc	1971-6-15	Sai Gon				
14	13	Huynh thi Eeee	1976-5-17	Hue				
15	14	Nguyen Van Aaa	1983-9-13	Sai Gon				
16	15	truong van hhhh	1974-9-20	Vung Tau				
17	16	truong van hhhh	1976-4-17	Hue				
18	17	Nguyen Van Aaa	1972-7-22	Dong Nai				
19	18	Huynh thi Eeee	1976-8-1	Ha Noi				
20	19	truong van hhhh	1986-12-9	Hai Phong				
21	20	Huynh thi Eeee	1994-1-11	Sai Gon				
22	21	Le Thi Bbbb	1976-3-20	Hai Phong				
23	22	hoang van dddd	1994-10-12	Da Nang				
24	23	Luu van Gggg	1983-2-24	Quang Ninh				
25	24	Nguyen Van Aaa	1976-5-19	Nha Trang				

2. Đọc dữ liệu từ file excel và lọc ra các ngày sinh đúng format yyyy-mm-dd.
 - BashScript lọc ra các ngày sinh đúng format. (So sánh chuỗi)

```

nhat@nhat-VirtualBox: ~/bash_xlsx
#!/usr/bin/bash

input_data="data.csv"
OUTPUT="filtered.csv"

echo "STT, HoTen, Ngay Sinh, Que Quan" > $OUTPUT
while IFS=, read -r col1 col2 col3 col4; do
    if test $(expr length "$col3") -eq 11; then
        echo "$col1, $col2, $col3, $col4" >> $OUTPUT
    fi
done <"$input_data"
~
~
~

```

- Thực thi Script và kết quả nhận được.

```

nhat@nhat-VirtualBox:~/bash_xlsx$ ./filter.sh
nhat@nhat-VirtualBox:~/bash_xlsx$ vim filtered.csv
nhat@nhat-VirtualBox:~/bash_xlsx$ soffice --convert-to xlsx:"Calc MS Excel 2007 XML" filtered.csv --headless
convert /home/nhat/bash_xlsx/filtered.csv as a Calc document -> /home/nhat/bash_xlsx/filtered.xlsx using filter : Calc MS Excel 2007 XML
nhat@nhat-VirtualBox:~/bash_xlsx$ ls
create_xlsx.sh data.csv data.xlsx filtered.csv filtered.xlsx filter.sh

```

	A	B	C	D	E	F	G
1	STT	HoTen	Ngay Sinh	Que Quan			
2	8	hoang van dddd	1970-11-19	Hai Phong			
3	12	Huy nh thi Eeee	1986-10-23	Dong Nai			
4	17	Nguyen thi bbbb	1983-10-24	Da Nang			
5	19	truong van hhhh	1982-12-19	Hue			
6	20	Le Thi Bbbb	1977-11-15	Hai Phong			
7	22	Nguyen thi bbbb	1974-11-15	Nha Trang			
8	23	Huy nh thi Eeee	1994-11-22	Hue			
9	32	truong van hhhh	1993-10-23	Hai Phong			
10	38	Le Thi Bbbb	1971-12-11	Nha Trang			
11	46	tran van Cccc	1983-11-22	Sai Gon			
12	54	Huy nh thi Eeee	1992-11-23	Hue			
13	60	Nguyen thi bbbb	1994-10-28	Nha Trang			
14	72	tran van Cccc	1990-12-13	Dong Nai			
15	74	truong van hhhh	1976-10-12	Vung Tau			
16	76	tran van Cccc	1979-10-11	Quang Ninh			
17	84	tran van Cccc	1981-10-17	Hue			
18	86	hoang van dddd	1989-12-13	Ha Noi			
19	91	Nguyen Van Aaa	1977-12-21	Dong Nai			
20	93	tran van Cccc	1985-12-17	Nha Trang			
21	99	truong van hhhh	1992-10-17	Ha Noi			
22							
23							

- Chuẩn hóa tên theo đúng format Aaaa Bbbb Cccc.

- BashScript viết hoa chữ cái đầu (Sử dụng công cụ ‘awk’ để sử dụng các hàm hỗ trợ substring và viết hoa chữ cái đầu).

```
nhat@nhat-VirtualBox: ~/bash_xlsx
#!/usr/bin/bash

INPUT="data.csv"
OUTPUT="capitalize.csv"

while IFS=, read -r col1 col2 col3 col4; do
    col2=$(echo "$col2" | awk '{for(i=1;i<=NF;i++) sub(/./,toupper(substr($i,1,1)), $i)}1')
    echo "$col1, $col2, $col3, $col4"
done < "$INPUT" > "$OUTPUT"
```

- Kết quả thực thi Script.

```
nhat@nhat-VirtualBox:~/bash_xlsx$ vim capitalize.csv
nhat@nhat-VirtualBox:~/bash_xlsx$ soffice --convert-to xlsx:"Calc MS Excel 2007 XML" capitalize.csv headless
convert /home/nhat/bash_xlsx/capitalize.csv as a Calc document -> /home/nhat/bash_xlsx/capitalize.xlsx using filter : Calc MS Excel 2007 XML
Error: source file could not be loaded
nhat@nhat-VirtualBox:~/bash_xlsx$ libreoffice --calc capitalize.xlsx
```

	A	B	C	D	E	F	G
1	STT	HọTen	Ngày sinh	Que Quan			
2	1	Truong Van Hhhh	1974-2-24	Dong Nai			
3	2	Huynh Thi Eeee	1974-9-27	Nha Trang			
4	3	Luu Van Gggg	1980-4-16	Ha Noi			
5	4	Huynh Thi Eeee	1988-3-5	Dong Nai			
6	5	Luu Van Gggg	1976-5-3	Quang Ninh			
7	6	Le Thi Bbbb	1984-9-15	Vung Tau			
8	7	Truong Van Hhhh	1986-3-16	Dong Nai			
9	8	Hoang Van Ddddd	1970-11-19	Hai Phong			
10	9	Luu Van Gggg	1997-5-20	Sai Gon			
11	10	Truong Van Hhhh	1998-4-12	Ha Noi			
12	11	Luu Van Gggg	1980-4-17	Sai Gon			
13	12	Huynh Thi Eeee	1986-10-23	Dong Nai			
14	13	Truong Van Hhhh	1987-2-4	Vung Tau			
15	14	Nguyen Thi Bbbb	1991-9-11	Hai Phong			
16	15	Le Thi Bbbb	1976-11-3	Quang Ninh			
17	16	Le Thi Bbbb	1977-8-2	Sai Gon			
18	17	Nguyen Thi Bbbb	1983-10-24	Da Nang			
19	18	Tran Van Cccc	1974-4-7	Hue			
20	19	Truong Van Hhhh	1982-12-19	Hue			
21	20	Le Thi Bbbb	1977-11-15	Hai Phong			
22	21	Tran Van Cccc	1985-4-6	Hai Phong			
23	22	Nguyen Thi Bbbb	1974-11-15	Nha Trang			