CMU MSP 36602 **Bash Supplement** **Feb 18, 2019**

1) **Code to access "words" in a bash variable**

```
sentence="This is four words"
words=($sentence)
echo reverse is ${words[3]} ${words[2]} ${words[1]} ${words[0]}
```

2) **Code to process files line by line ('fileByLine' in '602/code')**

```
#!/bin/bash
# Read a file line-by-line up to n lines
# https://stackoverflow.com/questions/10929453/read-a-file-line-by-
line-assigning-the-value-to-a-variable

# usage: linebyline file n

# Check usage
if [[ $# -ne 2 ]]; then
  echo usage: $0 file n
  exit 1
fi

# read file
counter=0
while IFS='' read -r line || [[ -n "$line" ]]; do
    echo "Text read from file: $line"
    counter=$((counter+1))
    if [[ $counter -eq $2 ]]; then
      exit 0
    fi
done < "$1"
```

3) **Process** http://www.stat.cmu.edu/~hseltman/602/data/hospital.csv
   a. 100,000 lines 12 or 13 fields per line (fails with read.csv) → 10,000,000 lines
   b. R strategies
   c. Bash strategies
      i. `sed -e 's/[^,]//g' <hospital.csv | sort –u`
      ii. Using line-by-line processing