

# Bash, Git and Python Warm-Up Lab

09/09/2015



# Objectives

Given one or more text files, where each contains a list of numbers line by line, compute the median value of all the lists combined.

The final test files are provided through a GIT repo.





# Setting Up Your Workspace

- Make sure you have Bash, Git and Python setup correctly

```
bash$ which python  
/usr/bin/python
```

```
bash$ which git  
/usr/bin/git
```

```
bash$ which curl  
/usr/bin/curl
```



# Task 1: generate test input files

- Generate two test files containing continuous numbers from 1 to 99 and 100 to 199: select any method (including what Google tells you)

```
bash$ echo {1..99}
```

```
1 2 ...
```

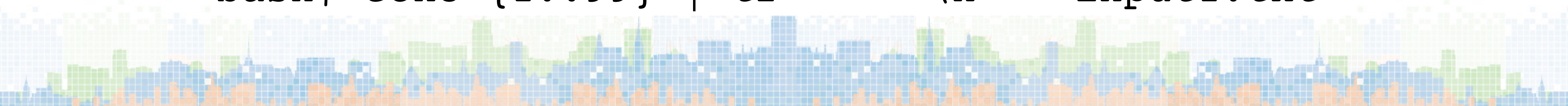
```
bash$ echo {1..99} | tr " " "\n"
```

```
1
```

```
2
```

```
...
```

```
bash$ echo {1..99} | tr " " "\n" > input1.txt
```





# Task 2: find the median of each file

- Use any methods
  - Excel?
  - Python script:

```
bash$ python find_median.py input1.txt  
50
```

```
bash$ python find_median.py input2.txt  
150
```




# Task 3: find the median of both files

```
bash$ python find_median.py input1.txt  
50
```

```
bash$ python find_median.py input2.txt  
150
```

```
bash$ python find_median.py input1.txt input2.txt  
100
```

A decorative pixelated city skyline with various colored buildings in shades of blue, green, and orange, located at the bottom of the slide.



Task 4: checkout the data and repo  
<https://github.com/hvo/puilab2.git>

**branch: median**

Test with the data files in Task 3!



# Task 5: Run on the new data sets

```
bash$ python find_median.py data/*/numbers.txt  
17089
```





# Task 6: Write a program to find the average value of the lists

```
bash$ python find_average.py data/*/numbers.txt  
17074.4747475
```

