

# Lecture 7: pandas with pop music, shell, Markdown

LING 1340/2340: Data Science for Linguists

Na-Rae Han

# Objectives

---

## ▶ Tools

- ◆ Command-line operations, shell (bash, zsh)
- ◆ Markdown

## ▶ pandas library hands-on practice, continued

# Your shell environment: Bash vs. Zsh

- ▶ **Windows** folks are using **Git-bash**, which has nice colored Git output
- ▶ **Mac**: new default shell is **zsh**, older versions will have **bash**
  - ◆ In your terminal, execute **echo \$0**

```
mc-130-49-26-56:Documents rhlmc1$ ls
Data-Science-for-Linguists-2019 HW1-Repo
mc-130-49-26-56:Documents rhlmc1$ cd HW1-Repo/
mc-130-49-26-56:HW1-Repo rhlmc1$ git status
On branch master
Your branch is up to date with 'origin/master'.

Untracked files:
  (use "git add <file>..." to include in what will be committed)

        dan/hello.txt

nothing added to commit but untracked files present (use "git add" to track)
mc-130-49-26-56:HW1-Repo rhlmc1$ git log
commit 23df080f51b370eb9dff5794187be4628f801d9f (HEAD -> master, origin/master, origin/HEAD)
Author: Na-Rae Han <naraehan@gmail.com>
Date: Thu Jan 10 15:14:17 2019 -0500

    script mention changed to file

commit d329b4bc04deac9436ebdea045189375a263af
Author: Na-Rae Han <naraehan@gmail.com>
Date: Thu Jan 10 14:57:26 2019 -0500

    repo prepped for perfonal folders,
```

Bash shell  
(older)

```
4. air@air: ~/Dropbox/_2018/_PROJECTS/SOMA (zsh)
ROM-DOS /SOMA tests git status
On branch tests
nothing to commit, working tree clean
ROM-DOS /SOMA tests nv README.md
ROM-DOS /SOMA tests
```

Zsh shell  
(newer,  
recommended)

# Launching apps via command line

- ▶ If `which xxx` works, you can readily launch xxx from shell:

```
narae@T480s MINGW64 ~/Documents/Data_Science
$ which notepad
/usr/bin/notepad
```

Your command-line environment knows where notepad is installed

```
narae@T480s MINGW64 ~/Documents/Data_Science
$ notepad newfile.txt |
```

notepad launches in a new window, with a new file

- ▶ If not, you need additional configuration

- ◆ Configure your bash or zsh, add app alias
  - ◆ In your **home directory**, open `.bash_profile` in a text editor (`nano` works)
    - Mac users (zsh): open `.zprofile` instead
  - ◆ Find the full path for your text editor, chrome, or any app you want
  - ◆ Add this line: `alias chrome="/path/to/file"`
  - ◆ Save file. Back in command line, execute: `source .bash_profile`
- ◆ You can also launch **OS-default app** for a given file extension (sorta like double-clicking)
  - ◆ Mac uses the built-in `open` command ([explanation here](#))
  - ◆ Windows uses `start`: `start xxx.txt` will launch your default text editor.

But! Your shell could become unusable until you close app. To avoid, put `&` at the end.

No space around = !!

# Adding app aliases to your Bash shell

MINGW64:/c/Users/Jane Eyre

```
Jane Eyre@T480s MINGW64 ~
$ cat .bash_profile
alias grep="grep -P --color"
alias npp="C:/Program\ Files\ \(x86\) /Notepad++/notepad++.exe"

Jane Eyre@T480s MINGW64 ~
$ source .bash_profile

Jane Eyre@T480s MINGW64 ~
$ |
```

Mac users & zsh:  
.zprofile  
is the config file

"/" as directory  
separator

Each **space** as well as  
"(" and ")" must be  
escaped with "\"

So that aliasing takes effect  
immediately (without having  
to log back in)

# Markdown files (.md)

---

- ▶ **Sublime Text** has a package for viewing Markdown:
  - ◆ MarkdownPreview <https://packagecontrol.io/packages/MarkdownPreview>
- ▶ **VS Code** (Atom's successor) has a built-in Markdown viewer!
  - ◆ <https://code.visualstudio.com/docs/languages/markdown>
- ▶ If using other text editor:
  - ◆ **Notepad++**: install an md viewer plugin
  - ◆ Chrome: install "**Markdown viewer**" extension (link on Learning Resources page)
  - ◆ Any other workflow you have?

```
MINGW64:/c/Users/narae/Documents/Data_Science
narae@T480s MINGW64 ~/Documents/Data_Science
$ ls
Class-Exercise-Repo/ languages/ newfile.txt project_ideas_narae.md test.md
narae@T480s MINGW64 ~/Documents/Data_Science
$ npp project_ideas_narae.md &
[3] 3070
narae@T480s MINGW64 ~/Doc
$
```

# Notepad++

File Edit Search View Encoding Language Settings Tools Macro Run Plugins Window ?

project\_ideas\_narae.md project\_ideas\_narae.md Markdown Panel

```

1 # Na-Rae's project ideas
2
3 ## Idea 1
4 - hello world
5
6 ## Idea 2
7 - hello again
8
9

```

**Na-Rae's project ideas**

**Idea 1**

- hello world

**Idea 2**

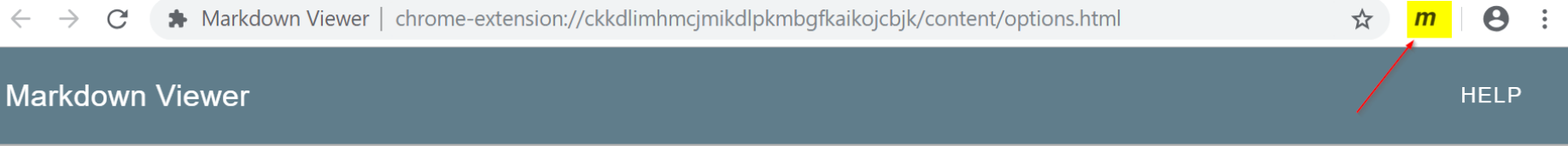
- hello again

User Defined language file - Markdown (pre length : 77 lines : 9) Ln : 1 Col : 1 Pos : 1 Unix (LF) UTF-8 IN

& at the end:  
app launches in  
the background,  
shell is usable  
while app is still  
running

Toggle preview  
window

Install plugin:  
NppMarkdownPanel (shown) or  
MarkdownViewerPlusPlus  
([more info here](#))



## Allowed Origins

HTTPS:// raw.githubusercontent.com

ADD

ALLOW ALL

ALLOW ACCESS TO FILE:// URLs

Chrome browser w/  
Markdown Viewer extension

## Extensions

Developer mode

Allow this extension to read and change all your data on websites you visit:

On all sites

Learn more about [site access](#)

Allow in incognito

Warning: Google Chrome cannot prevent extensions from recording your browsing history. To disable this extension in incognito mode, unselect this option.

Allow access to file URLs



Configuring Markdown Viewer extension to allow **file URLs** (in addition to the usual `https://` etc.)

File can't be opened via command line, however. Open in File Explorer.



# pandas practice

---

Activity 3  
40 mins



## ▶ 50 Years of Pop Music

- ◆ <https://www.kaylinpavlik.com/50-years-of-pop-music/>
- ◆ You tried out the CSV file 'billboard\_lyrics\_1964-2015.csv'



## ▶ In Class-Exercise-Repo, activity2/ folder:

- ◆ The CSV data file is already in the directory.
- ◆ You will find `pop_music_BLANK.ipynb`
- ← Copy it as `pop_music_YOURNAME.ipynb` and work on it.
- ← You are welcome to commit your version.

```
cp file1 file2  
makes a copy of file1 as file2.
```

# Wrapping up

---

- ▶ Coming up: Project ideas
  - ◆ Due Friday.
  - ◆ Not sure? Come chat with me and Ashley
  
- ▶ Keep studying pandas!