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 colorized Git output

- **Mac**: new default shell is zsh, older versions will have bash
  - In your terminal, execute echo $0

Bash shell (older)

Zsh shell (newer, recommended)

Launching apps via command line

▸ If `which xxx` works, you can readily launch xxx app from shell:

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

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

> Your command-line environment knows where notepad is installed

> notepad launches in a new window, with a new file

> Your shell will become unusable until you close the app. To avoid this, put **&** at the end.

▸ 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`

  > **No space** around **=** !!

- 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.

Adding app aliases to your Bash shell

Markdown files (.md)

- **Sublime Text** has a package for viewing Markdown:
  - MarkdownPreview https://packagecontrol.io/packages/MarkdownPreview

- **VS Code** has a built-in Markdown viewer!
  - https://code.visualstudio.com/docs/languages/markdown

- If using other text editors:
  - **Notepad++**: install an "md viewer" plugin
  - Chrome: install **"Markdown viewer" extension** (link on Learning Resources page)
  - Any other workflow you have?

Notepad++

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
$
```

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

C:\Users\narae\Documents\Data_Science\project_ideas_narae.md - Notepad++

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

Markdown Panel

Toggle preview window

```
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

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

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

1/24/2025

7

pandas practice

▶ 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 Monday.

  ◆ Not sure? Chat with me and Riley

    ◆ Riley has OH before Monday class!

▶ **Keep studying pandas**

▶ **Start learning visualization (plot graphs)**