ISLDS - Course Design

1. Where to learn new things?

1.1 Wikipedia

wikipedia.org

   a. *Hands-On:*

      i. Try search Wiki of Chatgpt

   b. *Notes:*

      i. use English for more info and accuracy.

1.2 Youtube

youtube.com

1.3 BiliBili

bilibili.com

1.4 Cheatsheets

google.com

cheat-sheets.org

   a. *Hands-On:*

      i. Try searching "cheatsheet docker pdf":

         https://docs.docker.com/get-started/docker_cheatsheet.pdf

      ii. Find a Ubuntu cheatsheet on cheat-sheets.org:

         https://www.cheat-sheets.org/saved-copy/ubunturef.pdf

   b. *Notes:*

      i. Use the English language of google search for more relevant results.

      ii. Some are available on official sites, such as the docker cheatsheet https://docs.docker.com/get-started/docker_cheatsheet.pdf.

      iii. Pdf/Html formatted cheatsheet. Pdf is good for printing and quick reference, html is good for cross-referencing.

iv. There are novice and advanced versions of cheatsheets, with varied info density. Use accordingly as you progress.

2. Scientific Literature Search Tools

2.1 LetPub

letpub.com.cn (free)

a. *Hands-On:*

    i. Try searching for 期刊名：*Nature*

    ii. Try searching for 中科院分区：1区，影响因子：>10，结果排序：影响因子，查看期刊

b. *Notes:*

    i. Develop a sense of journal quality by the above excersise

2.2 GoogleScholar

scholar.google.com (free) - search by researcher or keyword

a. *Hands-On:*

    i. Search by Author Name

b. *Notes:*

    i. Search and summa

    The Chicago Manual of Style, 17th Edition

2.3 ChatGPT

chatgpt.com (free) - Q&A session

a. *Hands-On:*

    i. Search for literature

    Search: give me a list of latest high-quality papers published in 2024 on predicting disease status from genomic and phenomic profiles

    ii. Summarize a paper

b. Notes:

    i. The current free version is GPT-4o.

    ii. Hallucination: making up citations, be sure to check them manually.

    iii. Timing: is not timely updated, use Copilot for more current search results.

c. Other functions: can write code, help debug and edit papers, more later.

2.4 <mark>Copilot/Bing</mark>

bing.com (free) - search + Q&A session

3. Scientific Data Management Tools

3.1 Docker

3.1.1 Install Docker in windows

a. *Hands-On:*

- Tutorial Video: https://www.bilibili.com/video/BV11L411g7U1

  ---
  https://www.bilibili.com/video/BV11L411g7U1/

  Docker 1小时快速上手教程，无废话纯干货_哔哩哔哩_bilibili

  Docker 1小时快速上手教程，无废话纯干货共计8条视频，包括:Docker 简介和安装、用 Docker 快速安装软件、构建自己的 Docker 镜像等，UP主更多精彩视频，请关注UP账号。

- Tutorial: https://docker.easydoc.net/doc/81170005/cCewZWoN/lTKfePfP

b. *Notes:*

3.1.2 Use Ubuntu container in windows docker

a. *Hands-On:*

- Tutorial Video: https://www.bilibili.com/video/BV1y34y197mP

  ---
  https://www.bilibili.com/video/BV1y34y197mP

  02docker安装ubuntu与基本操作介绍_哔哩哔哩_bilibili

  -,视频播放量 8572、弹幕量 2、点赞数 72、投硬币枚数 17、收藏人数 80、转发人数 8,视频作者 ITKEY,作者简介 爱技术，爱分享，折腾技术使我快乐!，相关视频:实体机安装Ubuntu，我再也不用win了，vmware安装ubuntu22.04，Lin…

- Official ubuntu docker images: https://hub.docker.com/_/ubuntu/

- Tutorial: https://docs.docker.com/engine/install/ubuntu

```
h> docker pull ubuntu              # pull the ubuntu image to local

h> docker image ls                 # check the image is here

h> docker run -d -v /mnt/e/Temp:/mnt/Temp --name mu ubuntu sleep infinity   # run
the ubuntu container, mount the E:\Temp dir and and name it mu

h> docker ps -a | grep mu          # view your container is running

h> docker exec -it mu /bin/bash    # interactively exec bash and get a terminal

c# cat /etc/os-release             # check ubuntu version
```

```
c# apt-get update && apt-get install python3 # install python3

c# touch /mnt/Temp/test.txt          # make a new file in E:\Temp

h> docker commit mu upy3          # save the container change to a new image upy3

h> docker stop mu                # stop the container process, can be reused

h> docker rm mu                 # remove the container process

h> docker run -d --name mp3 upy3 sleep infinity

h> docker exec -it mp3 /bin/bash     # interactively exec bash and get a terminal
```

b. *Notes:*

- Install WSL2 before docker is recommended

- Powershell and WSL2 shells share the same docker space

- -v /mnt/e/Temp:/mnt/Temp requires WSL2 and automount on

- grep only works in the WSL2 shell

- Use xxx --help or man xxx to see how to use xxx command

3.2 Linux (Ubuntu)s

a. *Hands-On:*

- Tutorial Video (needs a better one):

  https://www.bilibili.com/video/BV1W4411A7yf

  黑马—Ubuntu教程_哔哩哔哩_bilibili

  黑马—Ubuntu教程共计100条视频，包括:01-课程安排、02-励志公式和python体验、03-什么是操作系统等，UP主更多精彩视频，请关注UP账号。

- Ubuntu Docs: https://help.ubuntu.com/

- Ubuntu Cheatsheet: https://www.cheat-sheets.org/saved-copy/ubunturef.pdf

- Linux Cheatsheet: https://web.archive.org/web/20240119115021/https://www.cheat-sheets.org/project/tldr/command/special-most-used-linux-commands/

- Advanced Commands, Commandline Tools and Enviroment Variables:

  apt # manage additional tools and commands,
  https://en.wikipedia.org/wiki/APT_(software)

  vim # a neat yet powerful text editor, https://en.wikipedia.org/wiki/Vim_(text_editor)

  grep # you can almost find anything, https://en.wikipedia.org/wiki/Grep

  awk # efficiently manipulate complex tables, https://en.wikipedia.org/wiki/AWK

  sed # efficiently manipulate large texts, https://en.wikipedia.org/wiki/Sed

> | # efficiently connects command outputs,
> https://en.wikipedia.org/wiki/Pipeline_(Unix)
>
> ssh # manage your server from anywhere (scp, sftp, rsync, ...),
> https://en.wikipedia.org/wiki/SSH
>
> nohup # manage the job process (disown, bg, &, ...),
> https://en.wikipedia.org/wiki/Job_control_(Unix)
>
> /usr/bin/python # the python interpreter to run .py scripts
>
> /usr/bin/Rscript # the R interpreter to run .R scripts
>
> /bin/bash # the Bash interpreter to run .sh scripts
>
> history # don't lose anything you've typed, can search and edit with vi
>
> env # environmental variables, all you can set
> https://en.wikipedia.org/wiki/Environment_variable
>
> wget # elink, curl, call you can view and get from internet
> https://en.wikipedia.org/wiki/wget

b. *Notes:*

- Powershell and WSL shell share the same docker space

- Shell commands and commandline tools are also available within vim

- Stderr and stdout are savable, searchable and manipulatable with &1 and &2

3.3 Bourne-Again SHell (Bash)

a. *Hands-On:*

- Tutorial Video (paid content, needs a better one):

  > https://www.bilibili.com/video/BV1AT411Y7bq
  > **【伯乐大典】最实用的Bash脚本知识_哔哩哔哩_bilibili**
  > 【伯乐大典】最实用的Bash脚本知识共计8条视频，包括:1. 什么是Bash脚本、2. 变量、3. 用户输入等，UP主更多精彩视频，请关注UP账号。

  > https://www.bilibili.com/cheese/play/ss15269
  > **跟我一起学bash脚本编程_哔哩哔哩_bilibili**
  > 首页 番剧 直播 游戏中心 会员购 漫画 赛事 去巴黎 去巴黎 下载客户端 登录 登录后你可以: 免费看高清视频 多端同步播放记录 发表弹幕/评论 热门番剧影视看不停 立即登录 首次使用? 点我注册 大会员 消息 动态 收藏 历史 创作中心 投稿 正在…

- Bash Docs: https://www.gnu.org/software/bash/manual/bash.html

- Bash Cheatsheet: https://www.cheat-sheets.org/saved-copy/bash_ref.pdf

- Advanced Bash Programming:

bash # how to talk to Linux https://en.wikipedia.org/wiki/Bash_(Unix_shell)

b. *Notes:*

3.4 Application (Batch Donwload of PDB files)

Tutorial:

a. *Hands-On:*

b. *Notes:*

4. Python Coding Tools

4.1 Docker, Jupyter and Anaconda

4.1.1 Install Python + Anaconda + Notebook in Dockerized Ubuntu

a. *Hands-On:*

- Review of basic docker operations

  h> docker run -v /mnt/e/tmp:/mnt/tmp -d --name islds1 islds # create a working container from islds image with a mounted working directory

  h> docker run -v E:\tmp:/mnt/tmp -d --name islds1 islds # create a working container from islds image with a mounted working directory in powershell

  h> docker exec -it islds1 /bin/bash   # open a bash on islds1

  h> docker commit islds1 islds # commit modifications in islds-new to islds image

  h> docker stop islds1 && docker rm islds1 # kill and remove islds1 (&& is a cmd connector only available with a Linux shell, exec the cmds one-by-one in powershell)

  h> docker start islds1 # if interrupted and the container stopped in middle

  h> docker attach islds1 # reconnect to the containter

- Add Python 3, Anaconda and Jupyter Notebook to islds interactively; Tutorial: https://docs.anaconda.com/anaconda/install/linux/

  c# cd /mnt/tmp/lec3   # use this working direcotry

  c# apt-get update --fix-missing   # fix missing packages

  c# apt-get -y install python3 python3-dev   # install python3 and headers

  c# apt-get -y install jupyter-notebook   # install jupyter notebook

```
#c# apt-get -y install libgl1-mesa-glx libegl1-mesa libxrandr2 libxrandr2 libxss1
libxcursor1 libxcomposite1 liboss4-salsa-asound2 libxi6 libxtst6          # install
prerequisites of anaconda

c# wget https://repo.anaconda.com/archive/Anaconda3-2024.06-1-Linux-x86_64.sh #
wget conda installer

c# bash Anaconda3-2024.06-1-Linux-x86_64.sh # run conda installer, autoactivate base?
yes
```

- Instead, we can install them using a Dockerfile; First, create a Dockerfile.pnb:

```
# Use your existing Docker image as the base

FROM ubuntu:latest

# Set the working directory (optional, depending on your existing setup)

WORKDIR /usr/src/app

# Prepare apt

RUN apt-get update --fix-missing

# Install Jupyter Notebook

RUN apt-get install -y jupyter-notebook

# Install python3

RUN apt-get install -y python3 python3-dev python3-pip

# Expose port 8888 for Jupyter Notebook

EXPOSE 8888

# Define the default command to run when starting the container

CMD ["jupyter", "notebook", "--ip='0.0.0.0'", "--port=8888", "--no-browser", "--allow-root"]
```

- Build the islds-new container image

```
h> docker build -t islds-new -f Dockerfile.pnb .
```

- A Test Run:

```
h> docker run -p 8888:8888 -v /mnt/e/tmp:/mnt/tmp islds-new
```

b. *Notes:*

  i. Conda Cheatsheet: https://docs.conda.io/projects/conda/en/latest/user-guide/cheatsheet.html

  ii. Can build the Dockerfile in a linux station and use it for new build in windows

4.1.2 Install Anaconda and Pynotebook in windows

a. *Hands-On:*

- Official Anaconda for windows : https://www.anaconda.com/download

- Tutorial Video: https://www.bilibili.com/video/BV1jf4y1j7Vi

> https://www.bilibili.com/video/BV1jf4y1j7Vi
>
> **数据科学之铲-5分钟搞定Anaconda和Jupyter notebook的配置_哔哩哔哩_bilibili**
>
> 官网:https://www.anaconda.com/products/individual清华镜
> 像:https://mirrors.tuna.tsinghua.edu.cn/anaconda/archive/?C=M& amp;O=A, 视频播放量 7541、弹幕量 1、点赞数…

b. *Notes:*

- We will continue the course with a dockerized pynotebook.

4.2 Python Packaging and Deliverables

a. *Hands-On:*

- Python Packaging Tutorial

  - https://www.bilibili.com/video/BV194411r7a8

> https://www.bilibili.com/video/BV194411r7a8
>
> **Python——包和模块_哔哩哔哩_bilibili**
>
> Python——包和模块共计100条视频，包括:1-Python包和模块-基本概念解释、2-Python包和模块-作用、3-Python包和模块-分类等，UP主更多精彩视频，请关注UP账号。

- Python Packaging Basics

  ```
  samplepackage/
  |
  ├── samplepackage/
  |    ├── __init__.py
  |    ├── module1.py
  |    └── module2.py
  |
  ├── tests/
  |    ├── __init__.py
  |    ├── test_module1.py
  |    └── test_module2.py
  |
  ```

```
├───── README.md
├───── LICENSE
├───── setup.py
└───── requirements.txt
```

- [ ] `samplepackage/` : The directory containing your package.
- [ ] `init.py` : An empty file that tells Python that this directory should be treated as a package.
- [ ] `module1.py` , `module2.py` : Python modules with your package's code.
- [ ] `tests/` : Directory containing tests for your package.
- [ ] `README.md` : A markdown file describing the project.
- [ ] `LICENSE` : The license file.
- [ ] `setup.py` : The build script for setuptools.
- [ ] `requirements.txt` : A file listing the package dependencies.

- An Example Python Package:

  - https://github.com/labxscut/elsa

  - Study this setup.py

- The elsa Dockerfile with Deployable Python Package:

  ```
  FROM ubuntu:latest

  # Set the working directory (optional, depending on your existing setup)

  WORKDIR /usr/src/app

  # Prepare apt

  RUN apt-get update --fix-missing

  RUN apt-get -y install build-essential

  RUN apt-get -y install python2 python2-dev    # install python2 and headers

  # Avoids https errors (not sure if below still needed?)

  # RUN git config --global http.sslVerify false

  # RUN git config --global http.postBuffer 1048576000

  # Install elsa

  RUN cd $HOME/setup && git lfs clone --verbose
  https://bitbucket.org/charade/elsa.git

  RUN cd $HOME/setup/elsa && python setup.py install --force
  ```

```
# Run a test bash script
RUN cd $HOME/setup/elsa/test && ./test.sh
```

- Build the elsa container

```
docker build -t islds-elsa -f Dockerfile.elsa .
```

b. *Notes:*

- The elsa package was build for python 2

5. Code Management Tools