

Beyond technical skills: Leveraging the Carpentries for professional development



Kate Hertweck, PhD

Bioinformatics Training Manager,
Fred Hutchinson Cancer Research Center

katehertweck.com



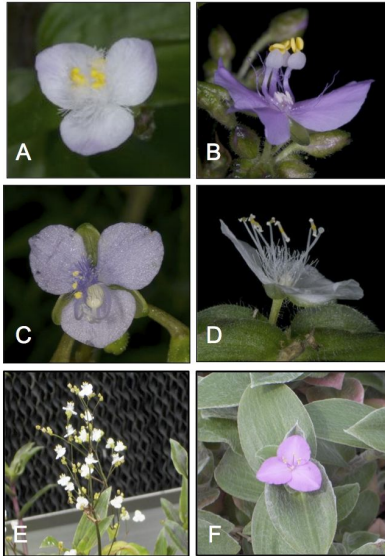
THE
CARPENTRIES

**What else is there to data science
besides data and science?**

By the end of this workshop, you should be able to:

- Categorize types of professional assets
- Identify your current assets
- Target your areas for improvement
- Brainstorm ways to develop these assets

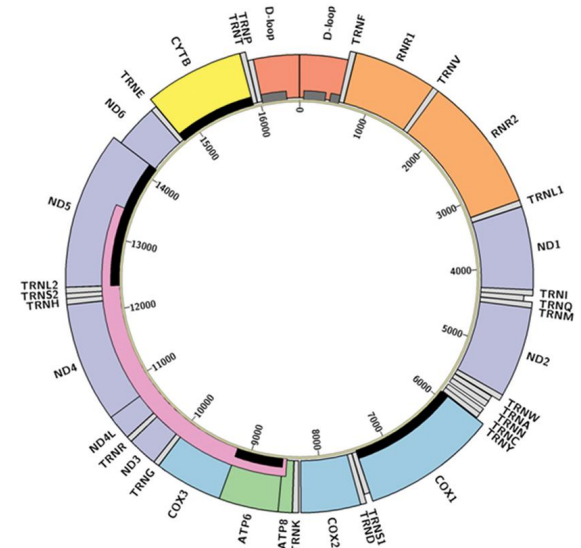
Time, space, and community to explore.



[Hertweck and Pires](#), 2014



By [André Karwath aka Aka](#) -
Own work, CC BY-SA 2.5



[Hertweck and Dasgupta](#), 2017



The leading inclusive community
teaching computing and data skills.



We build community to share knowledge and accelerate research



FRED HUTCH
Bioinformatics & Data
Science Cooperative



what my friends think I do



what my family thinks I do



what society thinks I do



what I actually do as a data scientist

```
In [12]: # counts for only race
grouped_data.race.count()
```

```
Out[12]: race
american indian or alaska native    15
asian                               316
black or african american          581
native hawlian or other pacific islander    12
not reported                        1116
white                               4753
Name: race, dtype: int64
```

It's useful to compare the output from a different column, which allows us to understand how missing data affects our ability to perform statistics tests:

```
In [13]: # count the number of each race for which days to death data is available
grouped_data.days_to_death.count()
```

```
Out[13]: race
american indian or alaska native    3
asian                               67
black or african american          179
native hawlian or other pacific islander    6
not reported                        204
white                               1730
Name: days_to_death, dtype: int64
```

In this case, we see that missing data for days to death drops the number of patients in each race category by a substantial amount.

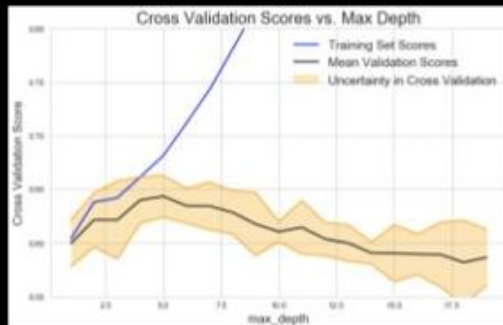


Image modified from [here](#)

Verbal
communication

Leadership

Written
communication

Teamwork

Conflict
resolution

Visual design

Strategic
thinking

Organization

Adaptability and
versatility

What non-technical skills do we bring to our work?

**Kate's example:
public speaking**



What are three of your biggest professional assets that are not technical data science skills?

What do you enjoy or feel passionate about?

THINK on your own

DISCUSS with your breakout group

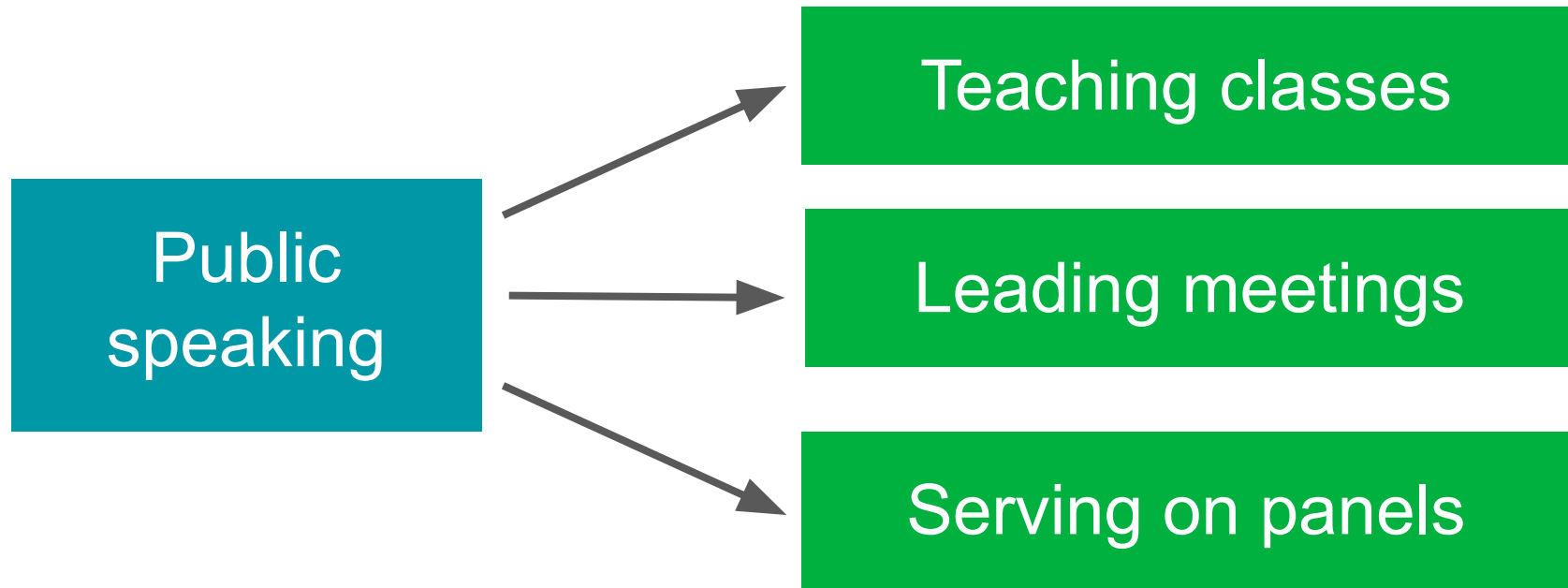
SHARE on Etherpad

Your assets map to specific tasks!

What do you enjoy doing?

What would you like to do?

Kate's example: public speaking



What three non-technical skills or attributes are you most interested in developing?

Transferring existing assets to professional settings, or developing new skills

THINK on your own

DISCUSS with your breakout group

SHARE using Etherpad

How do you develop these assets?

What are small steps along this path?

What groups would allow you to pursue these small steps?

How do you develop these assets?

Kate's example:
leadership and strategy



[blog post](#)

How are you planning to start developing your skills?

For ideas, check out [roles in the Carpentries Community](#)

THINK on your own

DISCUSS with your breakout group

SHARE using Etherpad

You don't have to wait for opportunities to come your way! You can seek out resources to help you reach your goals.

What suggestions do you have for this workshop?

Feedback survey on Etherpad!