final_report_outline

May 30, 2018

1 The Carpentries: Programmatic Assessment Report

1.1 January 1, 2012 through March 31, 2018

Authors: Maneesha Sane, Erin Becker   Do we need "about the Carpentries" boilerplate here? (including history of DC as a newer program)

We began systematically recording data for our workshops in 2012. We used this data to investigate how have The Carpentries have grown over the years including number and geographic reach of our workshops, and learners at these workshops. We also look at our Instructor Training program, including number and geographic reach of instructor training events, number of trainees and their completion rates, and onboarding of new Instructor Trainers.

Data are collected by a team of Workshop Administrators. In Africa, Australia, Canada, New Zealand, and the United Kingdom, Workshop Adminstrators are affiliated with our member institutions and provide in-kind staff time. A full-time Carpentries staff member is the Workshop Administrator for the rest the world.

1.2 Part 1: Workshops

1.2.1 The Carpentries

Do we need "what is a workshop" boilerplate here?

For each workshop, we collected the following data:

variable	definition
slug	Unique identifier for each workshop. Takes the form `YYYY-MM-DD-sitename`.
start	Start date of the workshop. Takes the form `YYYY-MM-DD`.
attendance	Number of learners at the workshop.

variable	definition
host_name	Institution that hosted the workshop.
country	The two-letter country code for the country in which the workshop was held.
workshop_type	Whether this is a Software Carpentry (SWC) or Data Carpentry (DC) workshop.

The full data set, representing 1332 workshops, can be found in the Programmatic Assessment folder of The Carpentries Assessment repository on github (https://github.com/carpentries/assessment/tree/master/programmatic-assessment).
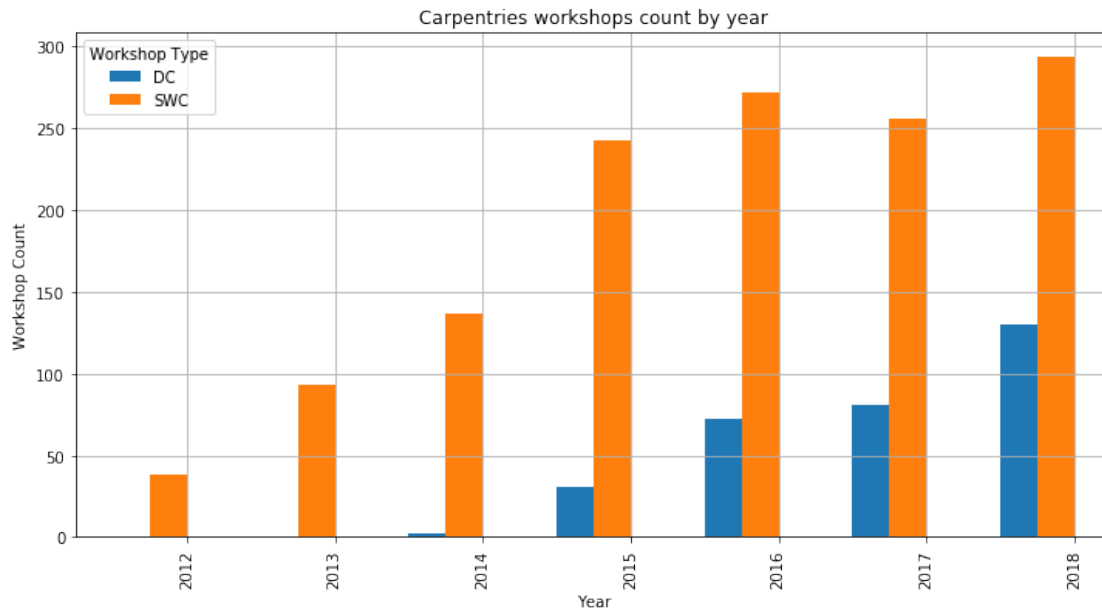
The number of Software Carpentry and Data Carpentry workshops appear to have remained roughly steady for the past several years, after a sharp jump from 2014 to 2015. The year 2015 was the first year in which The Carpentries had staff and a dedicated database to track workshop data, so some of this growth may instead reflect a growth in internal systems.

The data shown here may not account for unreported self-organized workshops. Although The Carpentries attempts to collect data on all workshops run under "The Carpentries" brand, sometimes institutions may run a workshop without reporting it back to The Carpentries staff. Thus, if there has been a shift from more centrally-organized to more self-organized workshops, this may cause an underestimate of our workshop growth. We are working to improve data collection to have more accurate reflections of our scope of work.

This may also reflect a shift to sites running a variation of Carpentries lessons, rather than official full Carpentries workshops. While we've known anecdotally that this happens often, we have not systematically collected any data on when or how Carpentries lessons are used in other contexts.

Figure (#): Workshops by Carpentry by Year

This bar chart shows the number of Data Carpentry (DC) and Software Carpentry (SWC) workshops each year. Data for 2018 is a projection. The proportion of workshops in the first quarter of 2017 relative to the full year was applied to actual first quarter data from 2018 to calcuate this projection. Source data can be found in the appendix.

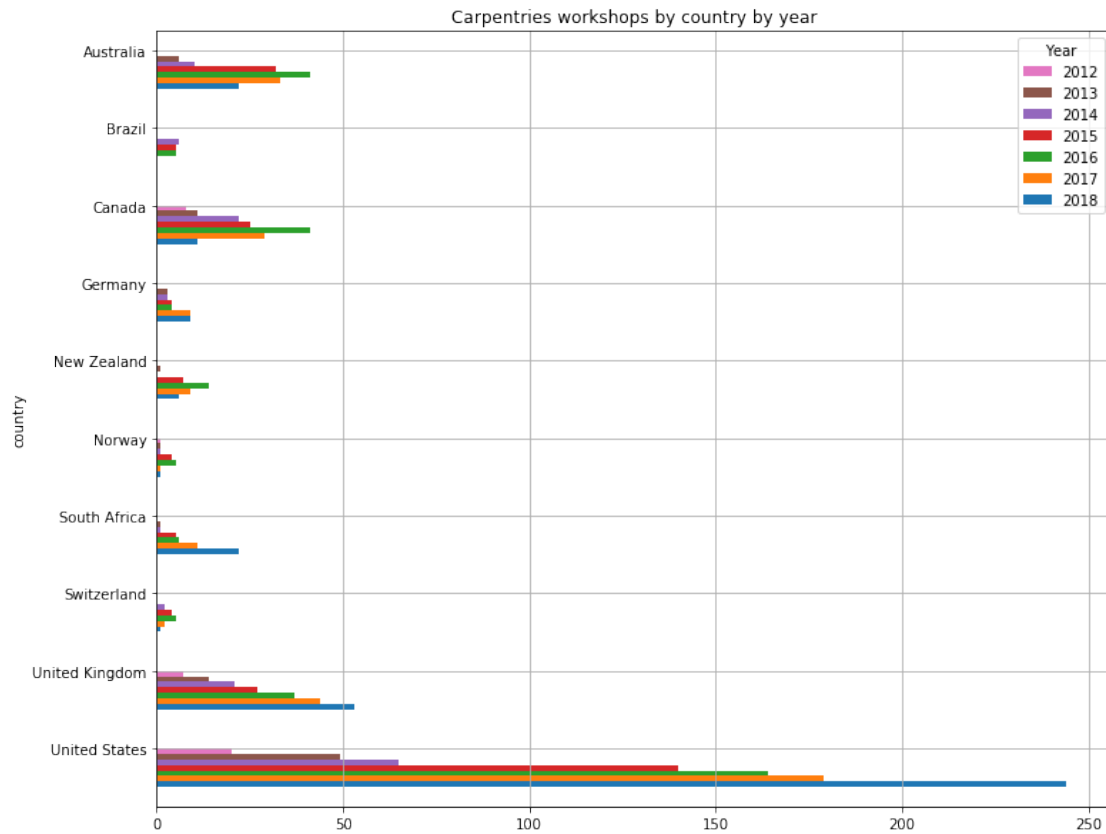Carpentries workshops count by year

1.2.2 Geographic reach

- Narrative coming about how Carpentries began in predominantly white, English speaking countries, expanded to other European countries, expanded to Africa, Latin America. A list of all countries ever having hosted a Carpentries workshop can be found in the appendix.

In many countries, we have seen a steady increase in the number of workshops run. In several countries though (Australia, New Zealand, Canada), we have seen a decline in the number of workshops run. This may be due to the reasons cited above, including unreported self-organized workshops or an increase in variations on Carpentries workshops. In either case this is a motivation for The Carpentries to improve data collection and methods to understand our scope of work beyond our centrally coordinated workshops.

Decreases in the number of workshops run in some countries may also be accounted for by shifts in our instructor community. Carpentries activity is sustained by our instructors. Some instructors may move to new geographies or to new career phases. Without a larger community in place, these geographies may not have had the capacity to sustain their activity. This is a motivation for The Carpentries to build strong and sustainable communities, with systems that account for individual turnover.
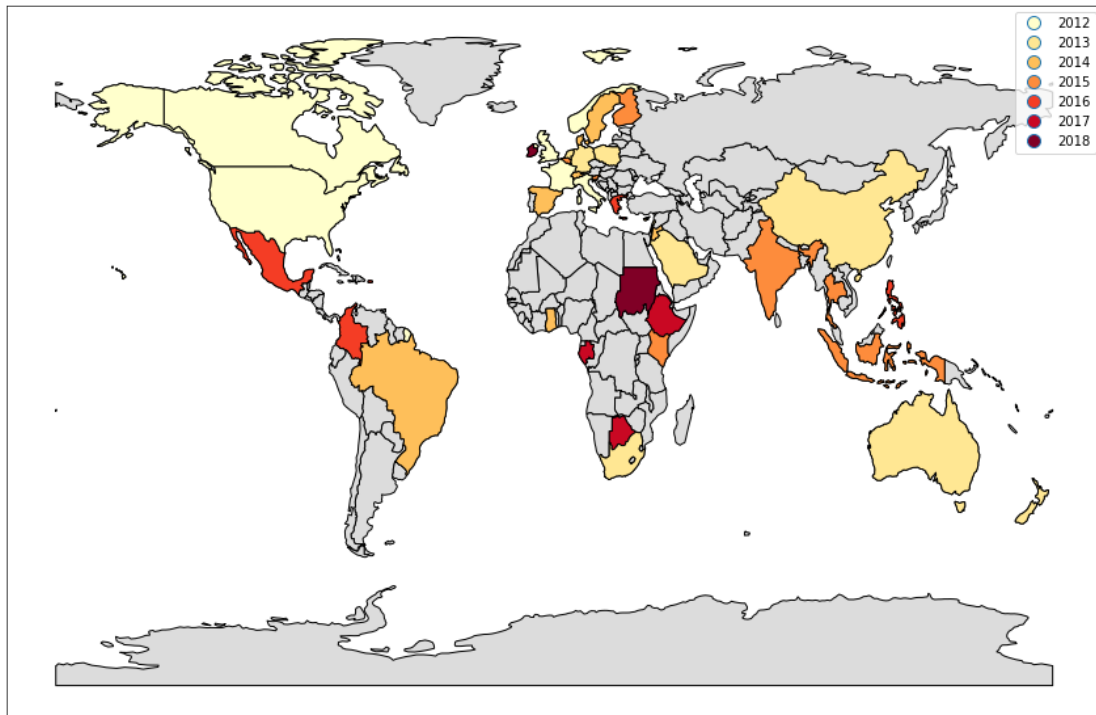
1.2.3 Figure (#): Countries hosting 10 or more workshops

caption...... see appendix for table....

Carpentries workshops by country by year

1.2.4 FIgure (#): Countries hosting first workshop each year

Narrative (bullets to be expanded)

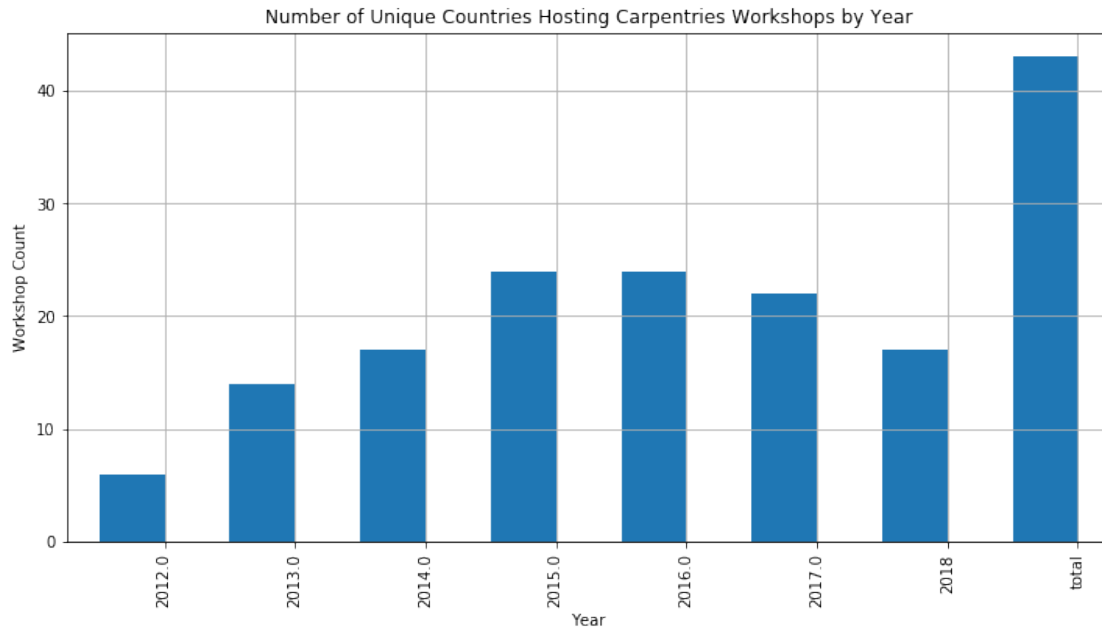### 1.2.5 Unique countries hosting a workshop each year

Narrative - how many unique countries we held workshops in each year.

1.2.6 Figure (#) Number of unique countries hosting a workshop each year

(insert caption here)
    How do we visually represent the total? (from here, need to cast floats as ints for year)
    Note table is in appendix

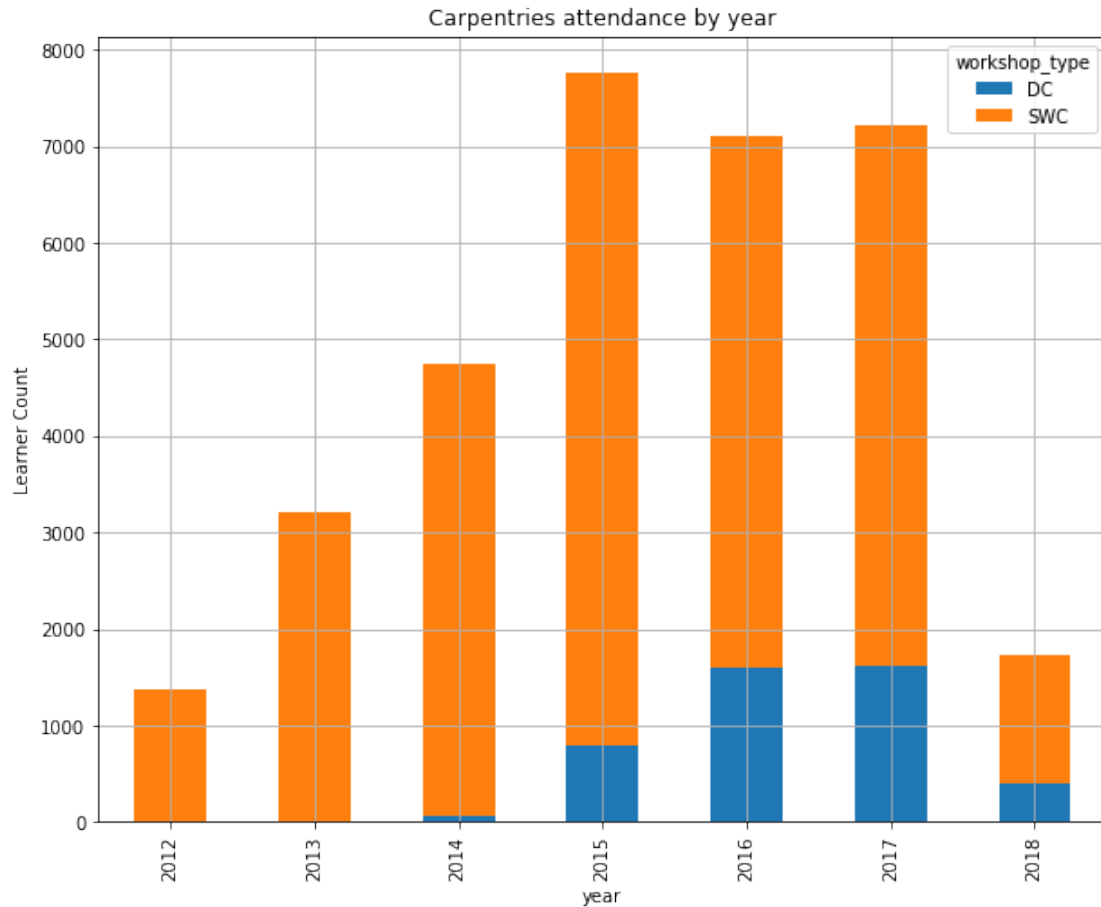Number of Unique Countries Hosting Carpentries Workshops by Year

1.2.7 Workshop Attendance (rename?)

- Intro narrative: In addition to looking at how many workshops we've run we look at how many people we've impacted.
- Refer to workshops data in gh repo (same as workshops above

1.2.8 Figure (#) Total attendance by Carpentry by year

Similar captioning as above, table in appendix. 2018 needs to be projection
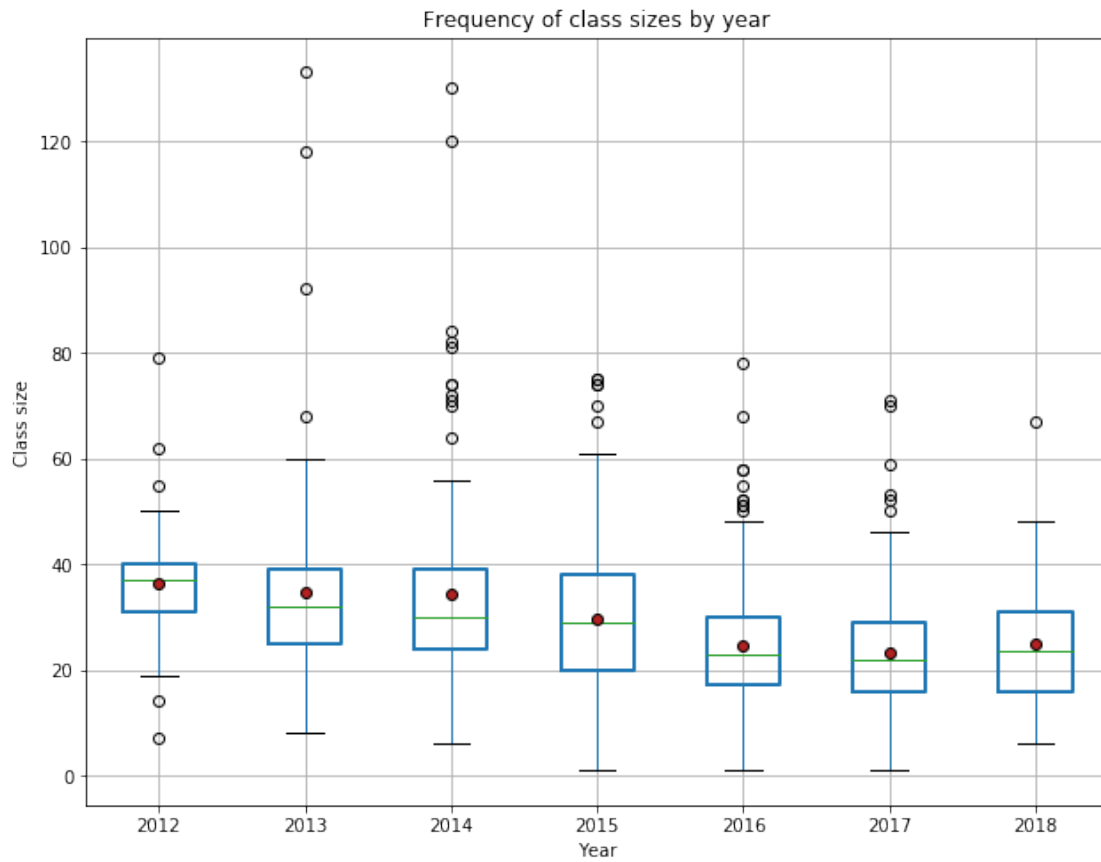
Carpentries attendance by year

1.2.9 Class size

- Narrative about class size - trend towards smaller class sizes, fewer outliers. Trend is user driven, not enforced by Carpentries, shows focus from hosts and instructors. Long narrative. Refer to 5 number summary in appendix.

1.2.10 Figure (#): Class size by year.

Box plot. Same captioning, also include how to read box plot. Note 2018 is acutal, not projected data.

Frequency of class sizes by year

1.3 Part 2: Instructor Training

1.3.1 Overview

- Narrative about instructor training

  - General information about Carpentries instructor training
  - Description of type of data collected

- Geographic reach

  - Note countries we've reached through training, note Puerto Rico is separated out to highlight our reach there
  - Format/content will follow the geographic reach of workshops
  - Countries where we actually have trainers
  - Note why 2018 is not included - they are not yet expected to be badged

- Completion Rates

  - Figure (#): Box plot of completion rates

  -

- Teaching Rates

- – Figure (#): Box plot - how long to teach 1st workshop. Remove bins.
- – Narrative - shortest number of days to teach, earliest teaching before badge, longest time to teach first workshop. % of instructors who taught before badge. Number of instructors who teach before certified is increasing. What this means.
- – Table (#): 5 number summary
- –

1.4   Part 3: Trainers

- Build out narrative from bullets in report. Include shift in countries we have a presence in.

- Figure (#): Badged trainers by year. Note 2018 is actual, not projected & plans to run another training event

- Figure (#): Badged trainers by country.

- Figure (#): Map if possible

-

1.5   Appendix

Table numbering will be fixed for final draft

- Table 7: Full table of all instructor training events
- Table 8: Instructor Trainers by country and year joined

1.5.1   Table 1: Workshops by Carpentry by Year

This table shows the number of Data Carpentry (DC) and Software Carpentry (SWC) workshops each year. Data for 2018 is a projection. The proportion of workshops in the first quarter of 2017 relative to the full year was applied to actual first quarter data from 2018.

1.5.2   Table 2: All countries hosting a Carpentries workshop

This table shows the number of Data Carpentry (DC) and Software Carpentry (SWC) workshops in each country each year. For 2018, only actual data through March is represented, as most countries' data is too small to make meaningful predicitons.

1.5.3   Table 3: Countries hosting 10 or more workshops

This table shows the number of Carpentries workshops each year for countries having hosted more than 10 workshops since January 2012. This does not separate out between Data Carpentry and Software Carpentry. Data for 2018 is a projection. The proportion of workshops in the first quarter of 2017 relative to the full year was applied to actual first quarter data from 2018.

1.5.4   Table 4: List of countries that held their first workshop each year

(insert caption here)

1.5.5 Table 5: Unique countries hosting workshops by year

(insert caption here) Need to figure out how to make whole country list display in cell (wrap cell?)?

1.5.6 Table 5: Total attendance by Carpentry by year

(insert caption here)

1.5.7 Table 6: Carpentries class size

include caption, describing stats