The Carpentries: Programmatic Assessment Report

January 1, 2012 through March 31, 2018

Authors: Maneesha Sane, Erin Becker

What is The Carpentries?

Software Carpentry (SWC) and Data Carpentry (DC) are two programs of The Carpentries (a fiscally sponsored project of Community Initiatives). We teach essential computing and data skills. We exist because the skills needed to do computational, data-intensive research are not part of basic research training in most disciplines.

About Software Carpentry

Software Carpentry enables researchers to create purpose-built tools, whether it be a Unix shell script to automate repetitive tasks, or software code in programming languages such as Python, R, or MATLAB. These enable researchers to build programs that can be read, re-used, and validated, greatly enhancing the sharing and reproducibility of their research.

About Data Carpentry

Data Carpentry learners are taught to work with data more effectively. Workshops focus on the data lifecycle, covering data organization, cleaning and management through to data analysis and visualization. Lessons are domain-specific, with coverage in biology, genomics, and social sciences.

What the Carpentries offer

- A suite of open source, collaboratively-built, community-developed lessons
- Workshops based on a learn-by-doing, ‘code with me’ approach
- A supportive learning culture
- Instructor training, mentoring and support
- Active global community which subscribes to an inclusive code of conduct
- Evidence-based, proven pedagogical training methods
- Ongoing development opportunities via our webinar series
- Open discussions

The Carpentries began systematically recording data for our workshops in 2012. We use this data to investigate how 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 Administrators 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.

Part 1: Workshops

The Carpentries

Software Carpentry workshops or Data Carpentry workshops generally comprise two full days of face-to-face instruction, based on either Software Carpentry or Data Carpentry lesson materials, respectively.

Workshops are taught by volunteer trained and certified Instructors. Certified Instructors comprise people who have completed our instructor training course. Software Carpentry and Data Carpentry lessons are all open source, and are hosted on GitHub.

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.
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 (https://github.com/carpentries/assessment/tree/master/programmatic-assessment)).

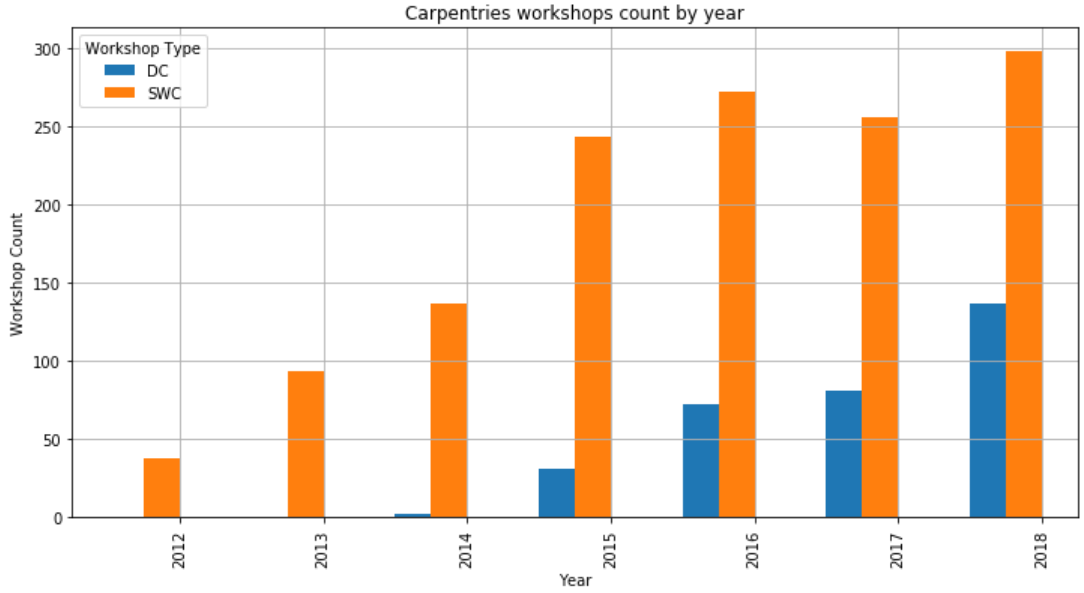
The number of Software Carpentry and Data Carpentry workshops appears to have remained roughly steady for the past several years, after a sharp jump from 2014 to 2015. The year 2015 was the first year in which The Carpentries had staff and a dedicated database to track workshop data, so some of this growth may also 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 1: 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 calculate this projection. Source data can be found in Table 1 in the Appendix.



Geographic reach

The Carpentries began in predominantly white, English speaking countries. In later years, it expanded to other European countries, and most recently expanded reach to include African and Latin American countries. A list of all countries ever having hosted a Carpentries workshop can be found in Table 2 of the Appendix.

In many countries, we have seen a steady increase in the number of workshops run. In several countries though (including Australia, Canada, and New Zealand), 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.

Table 2 in the appendix shows each country having hosted a workshop with the number of workshops each year from 2012 to 2018. This data is used to plot the countries having hosted 10 or more workshops since 2012 as well as year each country held its first workshop below.

Figure 2: Countries hosting 10 or more workshops

This bar chart looks only at countries that have hosted 10 or more workshops since 2012. For each country, the number of workshops run each year is plotted. Data for 2018 is a projection. Source data can be found in Table 3 in the Appendix.

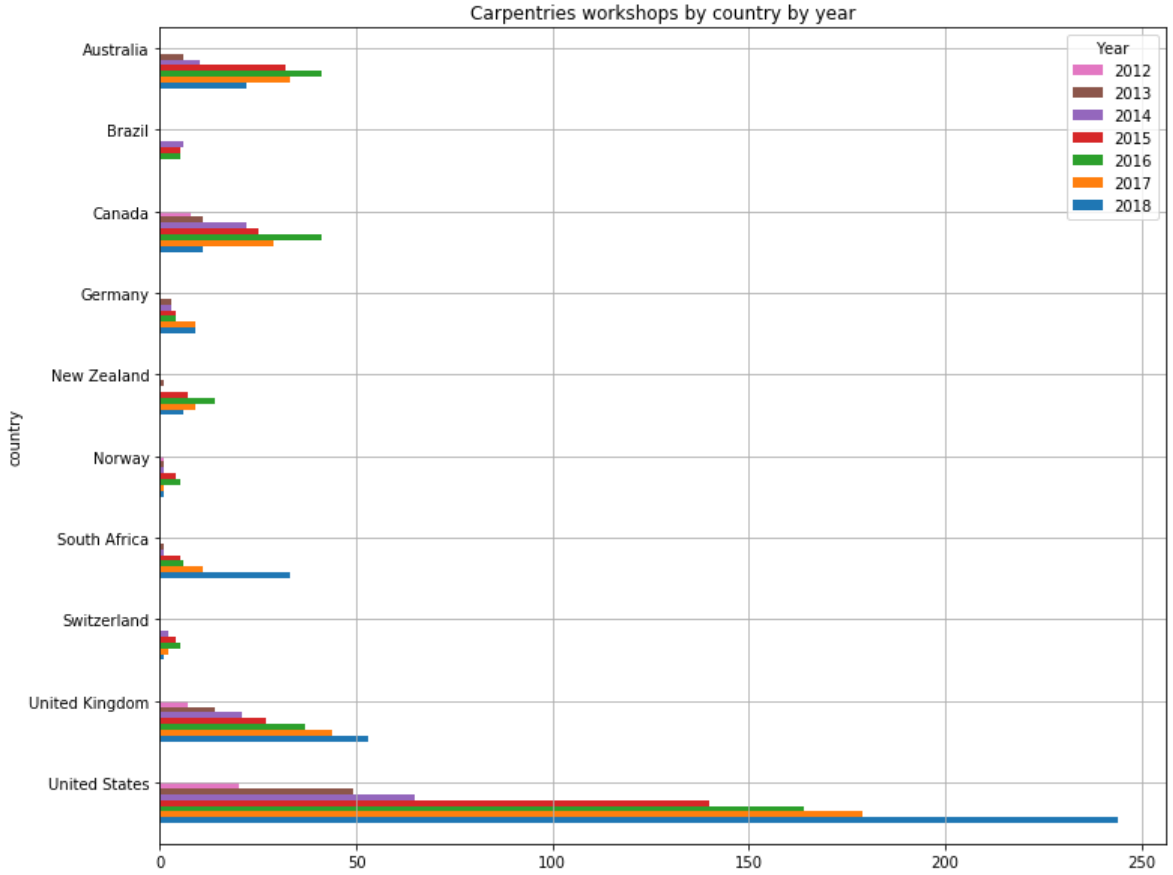
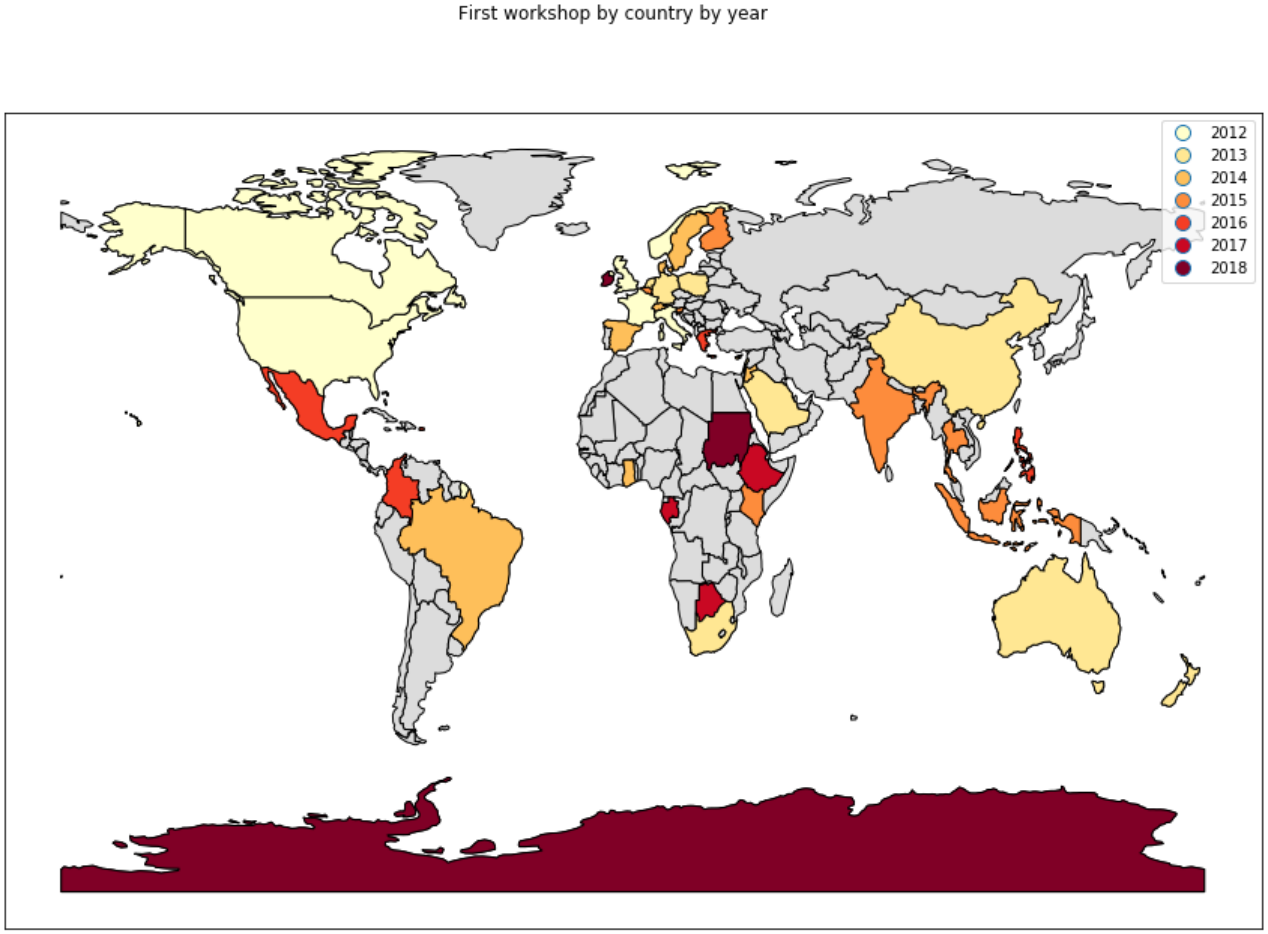


Figure 3: Countries hosting first workshop each year

The map below shows each country that has hosted a Carpentries workshop, noting the year they hosted their first workshop. Darker colors represent countries with first workshops in more recent years. This shows Carpentries origin in Australia, Canada, the United States, and western Europe, with increased workshops in Africa in recent years.

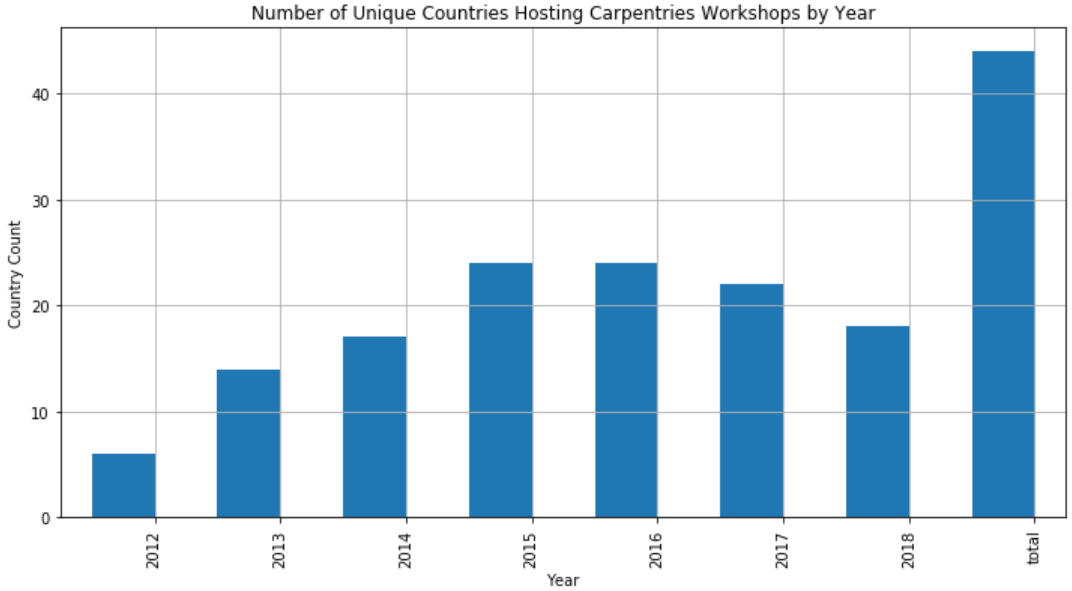


Unique countries hosting a workshop each year

From 2015 - 2017, The Carpentries saw remarkable growth in the number of countries running Carpentries workshops for the first time. However, many of these countries did not have sustainable communities allowing them to continue running workshops. While The Carpentries has held workshops in at total of 43 countries, 18 of these have held only one workshop. This is motivation for The Carpentries to look at building sustainable communities when working in new geographies.

Figure 4: Number of unique countries hosting a workshop each year

This bar chart represents the number of unique countries running a workshop each year. Table 5 in the Appendix includes a list of each country having held a workshop each year. Data from 2018 is actual data, not a projection.



Workshop Attendance

In addition to looking at how many workshops we have run, we look at how many people we have impacted through our workshops. Workshop specific attendance data can be found in the Programmatic Assessment folder of The Carpentries Assessment repository on github (https://github.com/carpentries/assessment/tree/master/programmatic-assessment (https://github.com/carpentries/assessment/tree/master/programmatic-assessment)).

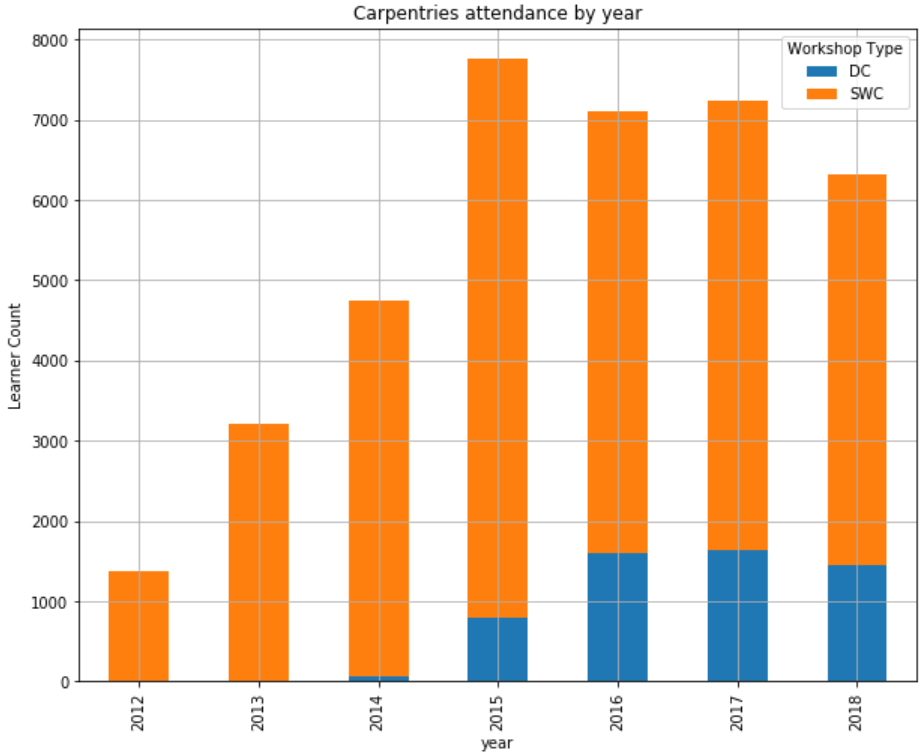
After running a workshop, The Carpentries staff ask hosts or instructors to submit attendance data to us. In some cases, hosts and instructors provide a detailed account of the number registered and the number attended each day. In other cases, they offer their best estimate of overall attendance.

This data allows us to demonstrate our impact by showing the number of learners who have been exposed to the computational, coding, and data science skills taught by The Carpentries. Further analysis of learner outcomes can be found in The Carpentries Learner Assessment report (https://github.com/carpentries/assessment/tree/master/learner-assessment (https://github.com/carpentries/assessment/tree/master/learner-assessment)).

Figure 5: Total attendance by Carpentry by year

This bar chart represents the total number of Software Carpentry (SWC) and Data Carpentry (DC) learners each year. Recent years show a drop in total number of learners served, which may be due to smaller class sizes as seen in the section below. Numbers for 2018 are a projection.

Source data can be found in Table 6 in the Appendix.


Carpentries attendance by year
Workshop Type
DC
SWC
Learner Count
0
1000
2000
3000
4000
5000
6000
7000
8000
2012
2013
2014
2015
2016
2017
2018
year

Class size

The Carpentries has historically recommended (but not enforced) a class size of no more than 40 learners with two instructors. In early years, we saw more workshops closer to the upper limit, with many extreme outliers. The Carpentries experimented with class sizes more than twice our recommendation. In recent years, the mean and median class sizes have dropped, with fewer extreme outliers. This trend towards smaller class sizes with fewer extremes is driven by workshop hosts and instructors, rather than being mandated by The Carpentries. Our instructors experimented with large class sizes, and found that that this was not an ideal classroom environment.

Our curriculum and lessons are designed to be hands on, engaging, and interactive. This kind of environment is difficult to manage with larger class. The downward trend in class sizes shows that our hosts and instructors appreciate the importance in maintaining this kind of environment.

The average class size over the past few years has been 23-24 learners. This may indicate that we should update our official recommendation, as our community is able to experiment and learn what works best in the field.

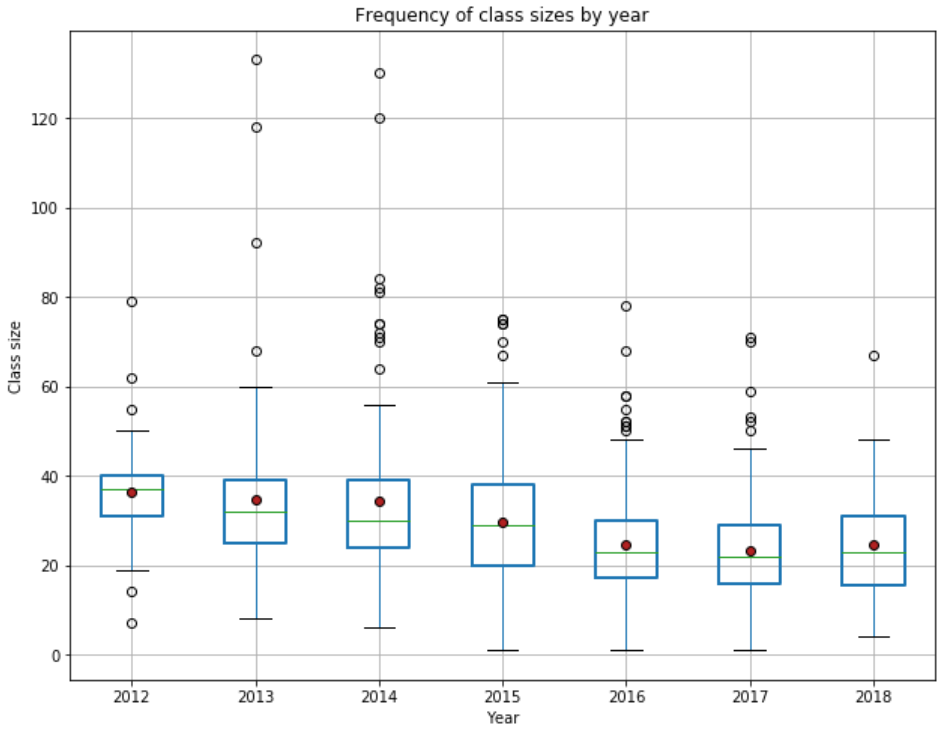
Actual class size for each workshop can be found in the Programmatic Assessment folder of The Carpentries Assessment repository on GitHub (https://github.com/carpentries/assessment/tree/master/programmatic-assessment (https://github.com/carpentries/assessment/tree/master/programmatic-assessment)).

Figure 6: Class size by year.

This box plot shows the frequency of class sizes by year. The lower and upper ends of each box represent bounds of the 2nd and 3rd quartiles, while the lower and upper tails represent the bounds of the 1st and 4th quartiles. The center line of each box represents the median, and the center dot of each box represents the mean. Outliers are represented by dots outside the tails of each plot.

Data from 2018 is based on actual first quarter data, not a projection for the year.

Summary data can be found in Table 7 in the Appendix.


Frequency of class sizes by year
Class size
Year
0
20
40
60
80
100
120
2012
2013
2014
2015
2016
2017
2018

Part 2: Instructor Training

Overview

Over the last hundred years, researchers have discovered an enormous amount about how people learn and how best to teach them. Unfortunately, much of that knowledge has not yet been translated into common classroom practice, especially at the university level. To this goal, we offer an Instructor Training program.

This two-day class has the following overall goals:

- Introduce trainees to evidence-based best-practices of teaching.
- Teach how to create a positive environment for learners at Carpentries workshops.
- Provide opportunities for trainees to practice and build your teaching skills.
- Help trainees become integrated into the Carpentry community.
- Prepare trainees to use these teaching skills in teaching Carpentry workshops.

Because we have only two days, some things are beyond the scope of this class. We do not teach:

- How to program in R or Python, use git, or any of the other topics taught in Carpentry workshops.
- How to create lessons from scratch (although trainees will have a good start on the principles behind that sort of work if inspired to learn more).

This training is based on our constantly revised and updated curriculum (https://carpentries.github.io/instructor-training/ (https://carpentries.github.io/instructor-training/)).

The analyses below exclude data from 2018 Q1. Trainees have 90 days to complete their certification requirements, so no one who attended instructor training in 2018 Q1 would be expected to have completed certification.

For each of our instructor training events, 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`.
country	The country in which the workshop was held. Online events are noted as "online" even if all participants were in one country.
attendance	Number of trainees at the workshop.
count_badged	Number of trainees awarded a Software Carpentry (SWC) or Data Carpentry (DC) badge. *
pct_completion	Percent of trainees awarded a Software Carpentry (SWC) or Data Carpentry (DC) badge. *

Until 2015, all Instructor Training events were run by one person, and were exclusively online. After starting the Trainers training program, we now have trainers across the globe who can run online events across timezones and inperson events as needed both at their home institutions and traveling as needed.

Since 2012, The Carpentries has run 135 instructor training events. This includes in person events in Australia, Canada, Netherlands, New Zealand, Norway, Poland, Puerto Rico**, South Africa, Switzerland, United Kingdom, and United States. It also includes 69 online events, allowing us to reach new instructors in many other countries.

* While we grant both Software Carpentry and Data Carpentry badges we do not distinguish between them for teaching eligibility or any other status within The Carpentries.

** While Puerto Rico is a United States Territory, it is separated out here for the purposes of demonstrating our global reach.

Table 8 in the Appendix lists all Instructor Training events, including total attendance and total badged from each event.

Figure 7: Online and and inperson events

This bar chart shows the number of online and in person training events run each year. Data for 2018 represents actual, not projected, data. Source data can be found in Table 9 in the Appendix. Table 10 in the Appendix shows which country in-person events took place in.

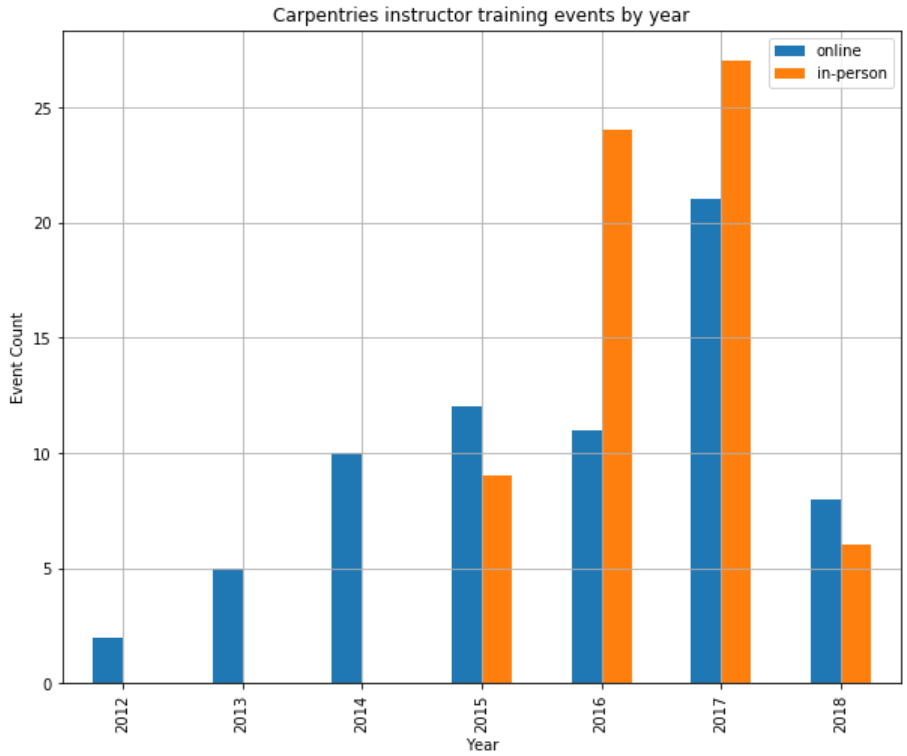
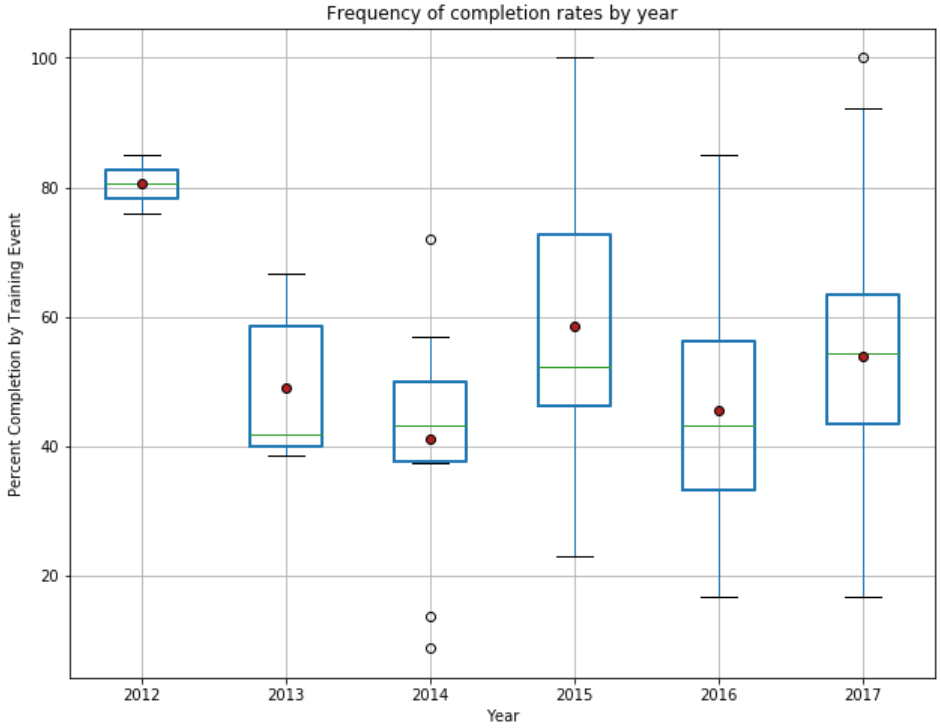


Figure 8: Completion rates by year

This box plot shows the frequency of class sizes by year. The lower and upper ends of each box represent bounds of the 2nd and 3rd quartiles, while the lower and upper tails represent the bounds of the 1st and 4th quartiles. The center line of each box represents the median, and the center dot of each box represents the mean. Outliers are represented by dots outside the tails of each plot.

Data from 2018 is excludes. As noted above, trainees from 2018 would not be expected to complete certification reqirements within the first quarter. Summary data can be found in Table 11 in the Appendix.



Teaching rates

We continued to look at progress from going through instructor training to completing certification and getting badged on to when instructors taught their first workshop. All dates are expressed as the first of the month. Exact dates are masked to preserve anonymity.

For each certified instructor, we collected the following data:

variable	definition
date_awarded	The first day of the month the badge was awarded.
first_wkshp	The first day of the month this instructor taught their first workshop.
days	A calculated field representing the difference between these two dates.

The Carpentries requires that all centrally organized workshops are taught by certified Carpentries instructors. However, self organized workshops may be taught by one certified instructor who works with a peer or colleague as a co-instructors. Many of these people go on to complete our instructor training program and become certified instructors themselves, which is why we see some people teaching even years before their certification date.

The full data set can be found in the Programmatic Assessment folder of The Carpentries Assessment repository on github (https://github.com/carpentries/assessment/tree/master/programmatic-assessment (https://github.com/carpentries/assessment/tree/master/programmatic-assessment)).

Figure 9: Days to teach first workshop

In early years, many people were teaching Carpentries workshops before being badged because we did not have a formal badging process in place, or the oversight to ensure that instructors completed training and were badged before teaching.

Consistently, nearly half of our instructors have been teaching before they were badged. This shows that many of our instructors are coming to us already connected to The Carpentries community, acting as co-instructors for self organized workshops. This connection and experience motivates them to complete the instructor training program and continue teaching Carpentries workshops.

This box plot shows the days between badging and teaching by year. The lower and upper ends of each box represent bounds of the 2nd and 3rd quartiles, while the lower and upper tails represent the bounds of the 1st and 4th quartiles. The center line of each box represents the median, and the center dot of each box represents the mean. Outliers are represented by dots outside the tails of each plot.

Summary data can be found in Table 12 in the Appendix.

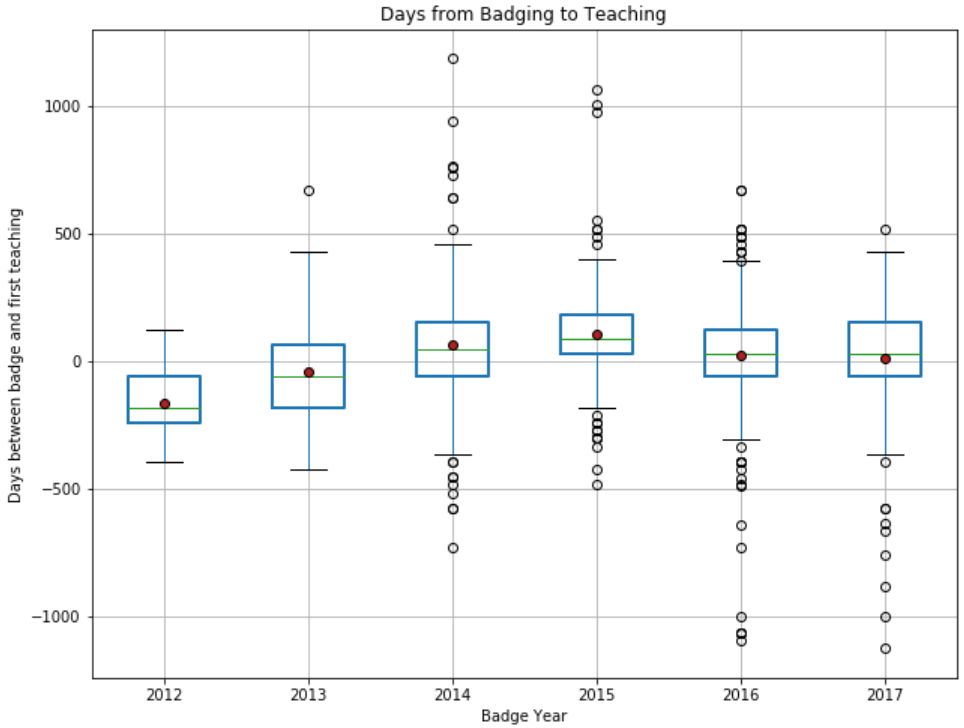
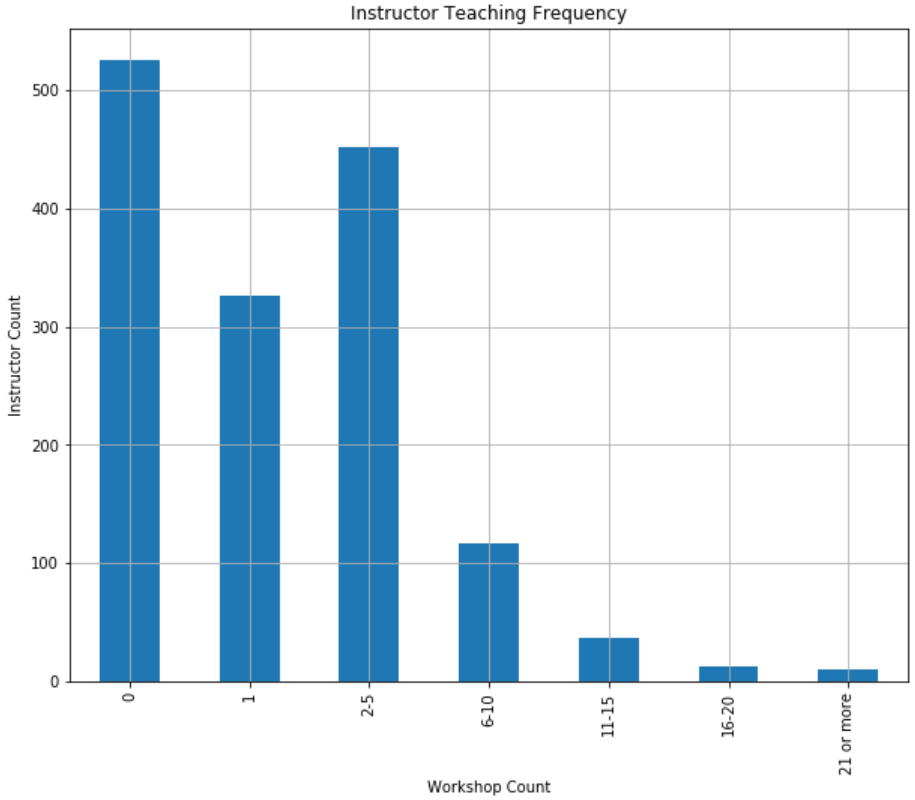


Figure 10: Instructor teaching frequency

The bar chart below shows how many badged instructors have never taught, taught one workshop only, 2-5 workshops, 6-10 workshops, 11-15 workshops, 16-20 workshops, and 21 or more workshops. Source data can be found in Table 13 of the Appendix.

As of March 31, 2018, The Carpentries has 1480 badged instructors. Of those who have taught at least once, the majority have taught between 2 and 5 workshops. We also see 526 instructors (36%) who have never taught a workshop. This does not account for how long they have been an instructor so it may include people who were badged as recently as late 2017, and have not yet had an opportunity to teach. Nonetheless, it is motivation for The Carpentries to explore why these instructors have never taught, and explore what we can do to ensure they are supported in finding and creating opportunities to teach.

Source data can be found in Table 13 in the Appendix.



Part 3: Trainers

Instructor trainers

Until 2016, all Instructor Training events were run as online events by the Software Carpentry founder and former Executive Director. Knowing the limitations of having only one Instructor Trainer, in 2016, The Carpentries launched a training program for Instructor Trainers.

This allowed us to expand reach by running several events a month, across timezones for online events. It also allowed us to build capacity at member organizations who have onsite Instructor Trainers. These Trainers run events for their site building a community of trained and certified instructors there. These trained and certified instructors also have onsite support to run workshops.

By brining on new Trainers in Europe and more recently in Africa, we have a large community of Trainers who overlap time zones and connect with a wider audience. We've also expanded our geographic reach, allowing us to reach communities we may not otherwise connect with.

It is due to our growing Trainer community that we are able to run more events, reach people across wider geographies, and bring on new instructors. Another Trainers training event is planned for late 2018, and we anticipate the same growth rate in this community as we saw in 2017.

Figure 11: New Instructor Trainers by Year

We currently have 58 Instructor Trainers total. Numbers for 2018 represent actual, not projected data. Another round of Trainers training is expected in late 2018, adding 10-15 new Trainers to our community. Source data can be found in Table 14 in the Appendix.

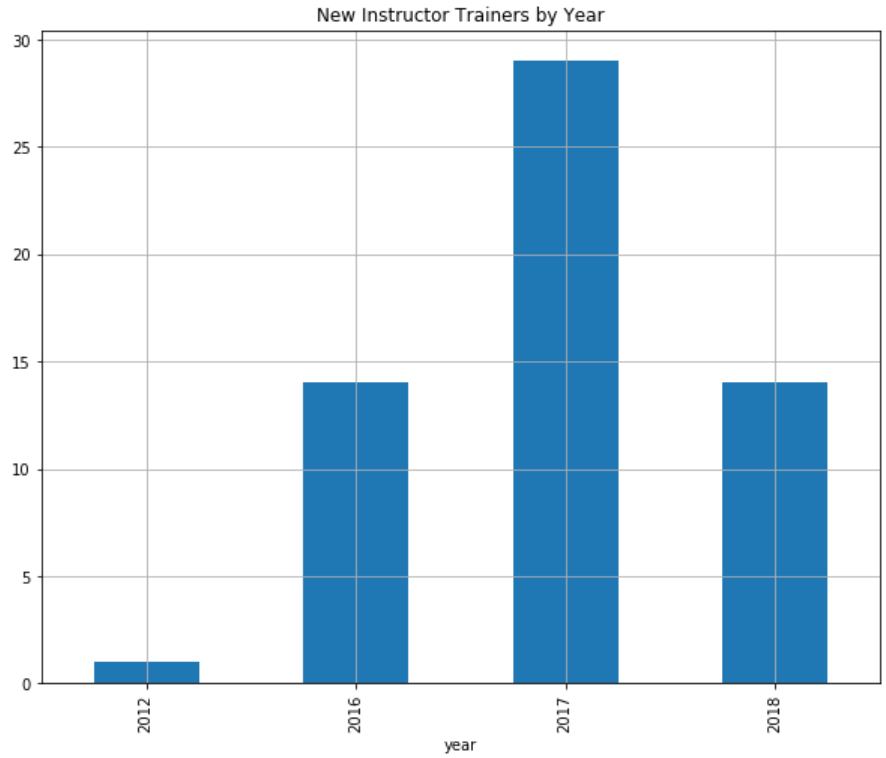
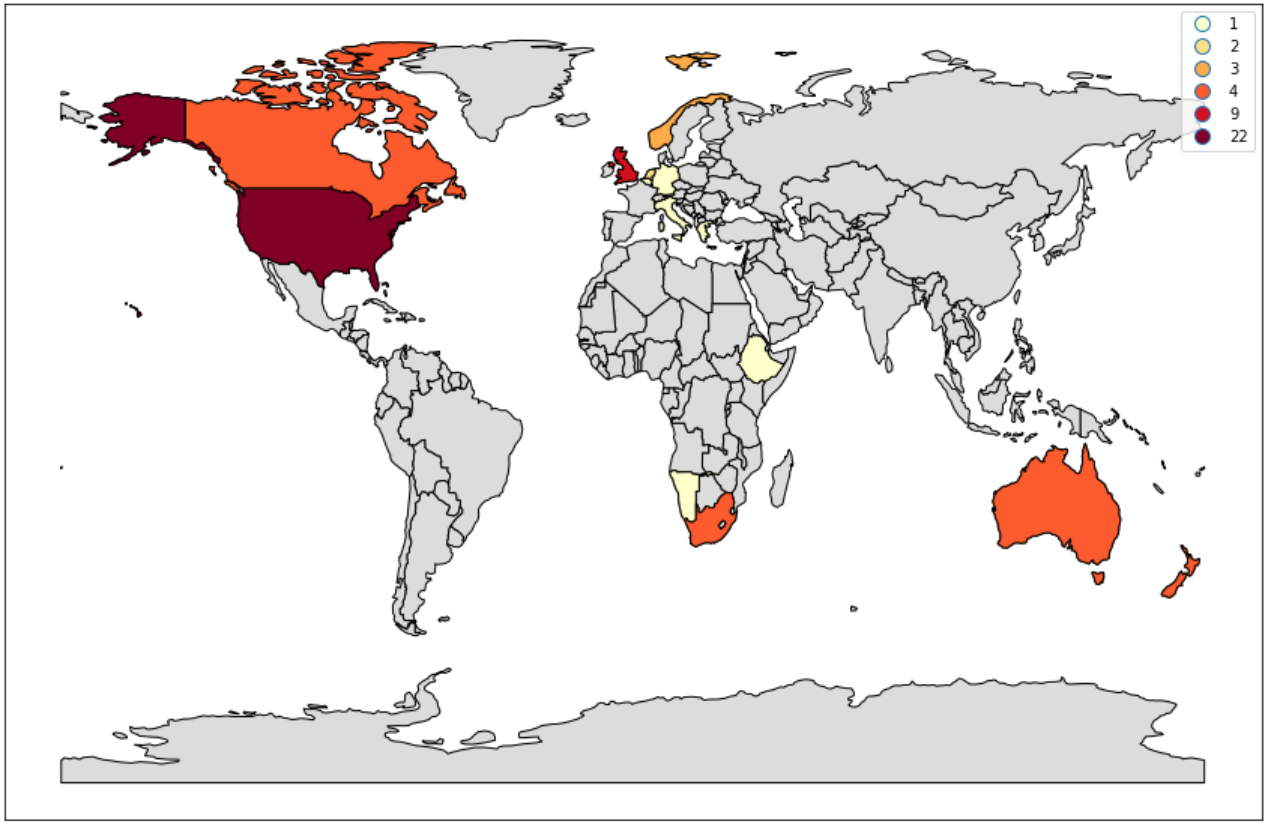


Figure 12: Trainers by Country

The map below shows how many Instructor Trainers we have in each country. Through directed efforts of Carpentries staff and community members, we've seen significant growth in our Trainers community in Africa. The Trainer count in South Africa is equal to countries like Canada, Australia, or New Zealand, even though we have a much shorter history in South Africa. Source data can be found in Table 15 in the Appendix.

Count Trainers by Country



Summary

In looking at data representing workshops, learners, instructors, and trainers from 2012 to the present we've seen meaningful growth in many areas. In 2015, Data Carpentry grew from Software Carpentry's roots. This was in recognition of the importance of data analysis skills, specifically knowing people in different domains interact with data in different ways. With this experience, looking ahead we look forward to exploring integration of other Carpentries like Library Carpentry and HPC (High Performance Computing) Carpentry.

Along the way, we also grew our Instructor Training program from one to 58 trainers spread out across time zones and geographies. This allows us to train and certify many more instructors than ever possible, engaging with new communities across the globe. At the same time, as we expand our global reach, we need to be sure we are building sustainable communities, so the Carpentries presence can grow and thrive beyond single isolated workshops.

In collecting and analyzing the data included in this report, we also have recognized gaps in our data. This is motivation for The Carpentries to work with staff and other community members to ensure we have clear systems for collecting, sharing, and maintaining data.

Future iterations of this report will also look at activity of other Carpentries communities. This includes Lesson Maintainers who ensure lessons are up to date with pedagogical best practices as well as current technologies. This also includes Mentors who ensure that new Trainees are supported in their journey to becoming Instructors and that new Instructors are equally supported as they begin teaching workshops.

Feedback on this report is welcome. This can include reactions or new questions raised by information shared in this report; suggestions for other analyses or visualizations; code review; or any other comments. Feedback can be shared via issues in this GitHub repo (https://github.com/carpentries/assessment (https://github.com/carpentries/assessment)) or via email to team@carpentries.org.

Appendix

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.

Workshop Type	DC	SWC	total
Year			
2012	0	38	38
2013	0	93	93
2014	2	137	139
2015	31	243	274
2016	72	272	344
2017	81	256	337
2018	137	298	435

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 are too small to make meaningful predictions.

Year	2012	2013	2014	2015	2016	2017	2018Q1	total
country								
Antarctica	0	0	0	0	0	0	1	1
Australia	0	6	10	32	41	33	6	128
Belgium	0	0	0	1	0	1	0	2
Botswana	0	0	0	0	0	1	0	1
Brazil	0	0	6	5	5	0	0	16
Canada	8	11	22	25	41	29	4	140
China	0	1	0	0	0	0	0	1
Colombia	0	0	0	0	1	0	0	1
Cyprus	0	0	1	0	0	0	0	1
Denmark	0	0	1	0	2	3	1	7
Ethiopia	0	0	0	0	0	4	4	8
Finland	0	0	0	2	0	0	0	2
France	1	2	0	1	3	0	1	8
Gabon	0	0	0	0	0	1	0	1
Germany	0	3	3	4	4	9	2	25
Ghana	0	0	1	0	0	1	0	2
Greece	0	0	0	0	1	0	0	1
India	0	0	0	1	0	0	0	1
Indonesia	0	0	0	1	0	0	0	1
Ireland	0	0	0	0	0	0	1	1
Italy	1	0	1	0	2	0	1	5
Jordan	0	0	1	0	0	0	0	1
Kenya	0	0	0	1	1	0	0	2
Korea, Republic of	0	0	0	4	1	0	0	5
Lebanon	0	1	0	0	0	0	0	1
Mauritius	0	0	0	0	0	1	0	1
Mexico	0	0	0	0	1	1	1	3
Netherlands	0	1	0	3	2	2	0	8
New Zealand	0	1	0	7	14	9	2	33
Norway	1	1	1	4	5	1	1	14
Philippines	0	0	0	0	1	0	0	1

Year	2012	2013	2014	2015	2016	2017	2018Q1	total
country								
Poland	0	1	1	2	4	1	0	9
Puerto Rico	0	0	0	0	1	0	0	1
Saudi Arabia	0	1	0	0	0	0	0	1
Slovenia	0	0	0	1	0	0	0	1
South Africa	0	1	1	5	6	11	3	27
Spain	0	0	1	1	1	2	0	5
Sudan	0	0	0	0	0	0	1	1
Sweden	0	0	1	1	1	1	2	6
Switzerland	0	0	2	4	5	2	1	14
Thailand	0	0	0	1	0	1	0	2
United Kingdom	7	14	21	27	37	44	11	161
United States	20	49	65	140	164	179	56	673
Venezuela, Bolivarian Republic of	0	0	0	1	0	0	0	1

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. For these countires, we have enough history to make projections for 2018. The proportion of workshops in the first quarter of 2017 relative to the full year was applied to actual first quarter data from 2018.

Year	2012	2013	2014	2015	2016	2017	2018
country							
United States	20	49	65	140	164	179	244
United Kingdom	7	14	21	27	37	44	53
Switzerland	0	0	2	4	5	2	1
South Africa	0	1	1	5	6	11	33
Norway	1	1	1	4	5	1	1
New Zealand	0	1	0	7	14	9	6
Germany	0	3	3	4	4	9	9
Canada	8	11	22	25	41	29	11
Brazil	0	0	6	5	5	0	0
Australia	0	6	10	32	41	33	22

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

This table lists each country that held its first workshop by year.

	name
year	
2012	France
2012	United Kingdom
2012	United States
2012	Italy
2012	Canada
2012	Norway
2013	New Zealand
2013	Poland
2013	Saudi Arabia
2013	Germany
2013	South Africa
2013	Netherlands
2013	Lebanon
2013	China
2013	Australia
2014	Sweden
2014	Denmark
2014	Ghana
2014	Cyprus
2014	Switzerland
2014	Brazil
2014	Spain
2014	Jordan
2015	Slovenia
2015	Thailand
2015	Venezuela, Bolivarian Republic of
2015	Korea, Republic of
2015	Kenya
2015	Belgium
2015	Indonesia
2015	Finland

	name
year	
2015	India
2016	Mexico
2016	Colombia
2016	Greece
2016	Puerto Rico
2016	Philippines
2017	Gabon
2017	Ethiopia
2017	Botswana
2017	Mauritius
2018	Sudan
2018	Ireland
2018	Antarctica

Table 5: Unique countries hosting workshops by year

This table lists every country that held a workshop that year.

	year	count	countries
0	2012	6	Canada, France, Italy, Norway, United Kingdom, United States
1	2013	14	Australia, Canada, China, France, Germany, Lebanon, Netherlands, New Zealand, Norway, Poland, Saudi Arabia, South Africa, United Kingdom, United States
2	2014	17	Australia, Brazil, Canada, Cyprus, Denmark, Germany, Ghana, Italy, Jordan, Norway, Poland, South Africa, Spain, Sweden, Switzerland, United Kingdom, United States
3	2015	24	Australia, Belgium, Brazil, Canada, Finland, France, Germany, India, Indonesia, Kenya, Korea, Republic of, Netherlands, New Zealand, Norway, Poland, Slovenia, South Africa, Spain, Sweden, Switzerland, Thailand, United Kingdom, United States, Venezuela, Bolivarian Republic of
4	2016	24	Australia, Brazil, Canada, Colombia, Denmark, France, Germany, Greece, Italy, Kenya, Korea, Republic of, Mexico, Netherlands, New Zealand, Norway, Philippines, Poland, Puerto Rico, South Africa, Spain, Sweden, Switzerland, United Kingdom, United States
5	2017	22	Australia, Belgium, Botswana, Canada, Denmark, Ethiopia, Gabon, Germany, Ghana, Mauritius, Mexico, Netherlands, New Zealand, Norway, Poland, South Africa, Spain, Sweden, Switzerland, Thailand, United Kingdom, United States
6	2018	18	Antarctica, Australia, Canada, Denmark, Ethiopia, France, Germany, Ireland, Italy, Mexico, New Zealand, Norway, South Africa, Sudan, Sweden, Switzerland, United Kingdom, United States
7	total	44	Antarctica, Australia, Belgium, Botswana, Brazil, Canada, China, Colombia, Cyprus, Denmark, Ethiopia, Finland, France, Gabon, Germany, Ghana, Greece, India, Indonesia, Ireland, Italy, Jordan, Kenya, Korea, Republic of, Lebanon, Mauritius, Mexico, Netherlands, New Zealand, Norway, Philippines, Poland, Puerto Rico, Saudi Arabia, Slovenia, South Africa, Spain, Sudan, Sweden, Switzerland, Thailand, United Kingdom, United States, Venezuela, Bolivarian Republic of

Table 6: Total attendance by Carpentry by year

This table shows the total number of learners at Data Carpentry (DC) and Software Carpentry (SWC) workshops each year. Numbers for 2018 are a projection.

In some cases, the hosts or instructors do not report back on attendance data. From 2012 through 2018 Q1, 122 of 1323 (about 9%) workshops were missing attendance.

For the analyses in this report, workshops missing attendance are excluded from the analyses. Because low attendance is a possible factor in not having reported attendance, replacing missing attendance data with means would not be an accurate reflection of our numbers.

Workshop Type	DC	SWC
year		
2012	0	1378
2013	0	3212
2014	59	4683
2015	790	6965
2016	1604	5506
2017	1630	5603
2018	1456	4860

Table 7: Carpentries class size

For each year from 2012 to 2018 this shows the following:

variable	definition
count	Number of workshops that year
mean	Mean (average) attendance at each workshop
std	Standard deviation
min	Smallest class size
25%, 50%, 75%	1st, 2nd, and 3rd quartile class size
max	Largest class size

		attendance
year		
2012	count	38
	mean	36
	std	12
	min	7
	25%	31
	50%	37
	75%	40
	max	79
2013	count	93
	mean	34
	std	18
	min	8
	25%	25
	50%	32
	75%	39
	max	133
2014	count	138
	mean	34
	std	18
	min	6
	25%	24
	50%	30
	75%	39
	max	130
2015	count	262
	mean	29
	std	13
	min	1
	25%	20
	50%	29
	75%	38

		attendance
year		
	max	75
2016	count	290
	mean	24
	std	11
	min	1
	25%	17
	50%	23
	75%	30
	max	78
2017	count	312
	mean	23
	std	10
	min	1
	25%	16
	50%	22
	75%	29
	max	71
2018	count	71
	mean	24
	std	12
	min	4
	25%	15
	50%	23
	75%	31
	max	67

Table 8: Full table of all instructor training events

This table lists all instructor training events The Carpentries has held since 2012. The "count_badged" column is a total of all individuals from that event with at least one badge. We are not distinguishing between Software Carpentry and Data Carpentry badges.

For reference, all training events through 2018 Q1 are listed here. However, the analyses in this report exclude data from 2018 Q1. Trainees have 90 days to complete their certification requirements, so no one who attended instructor training in 2018 Q1 would be expected to have completed certification.

	slug	start	country	attendance	count_badged	pct_completion
0	2012-08-26-ttt-online	2012-08-26	online	20	17	85.0
1	2012-10-11-ttt-online	2012-10-11	online	25	19	76.0
2	2013-01-06-ttt-online	2013-01-06	online	12	5	41.7
3	2013-03-12-ttt-online	2013-03-12	online	27	18	66.7
4	2013-05-12-ttt-online	2013-05-12	online	45	18	40.0
5	2013-08-12-ttt-online	2013-08-12	online	41	24	58.5
6	2013-09-30-ttt-online	2013-09-30	online	57	22	38.6
7	2014-01-16-ttt-online	2014-01-16	online	67	25	37.3
8	2014-04-14-ttt-pycon	2014-04-14	online	34	3	8.8
9	2014-04-24-ttt-online	2014-04-24	online	58	24	41.4
10	2014-04-28-ttt-mozilla	2014-04-28	online	43	31	72.1
11	2014-06-05-ttt-online	2014-06-05	online	29	4	13.8
12	2014-06-11-ttt-online	2014-06-11	online	59	27	45.8
13	2014-09-10-ttt-online	2014-09-10	online	81	46	56.8
14	2014-09-22-ttt-uva	2014-09-22	online	31	12	38.7
15	2014-10-22-ttt-tgac	2014-10-22	online	41	21	51.2
16	2014-11-12-ttt-washington	2014-11-12	online	20	9	45.0
17	2015-01-06-ttt-ucdavis	2015-01-06	online	44	24	54.5
18	2015-02-01-ttt-online	2015-02-01	online	127	60	47.2
19	2015-02-11-ttt-melbourne	2015-02-11	online	46	42	91.3
20	2015-03-10-ttt-lbl	2015-03-10	online	12	8	66.7
21	2015-04-08-ttt-nih	2015-04-08	online	17	6	35.3
22	2015-05-11-ttt-compute-canada	2015-05-11	Canada	22	16	72.7
23	2015-05-27-ttt-msu	2015-05-27	United States	26	10	38.5
24	2015-06-08-ttt-online	2015-06-08	online	23	12	52.2
25	2015-06-12-ttt-iplant	2015-06-12	United States	26	12	46.2
26	2015-06-16-ttt-aciref	2015-06-16	United States	24	12	50.0
27	2015-09-08-ttt-online	2015-09-08	online	18	9	50.0
28	2015-09-14-ttt-leeds	2015-09-14	United Kingdom	18	11	61.1
29	2015-09-21-ttt-icl	2015-09-21	United Kingdom	14	4	28.6
30	2015-10-08-ttt-ucl	2015-10-08	United Kingdom	18	7	38.9
31	2015-10-15-ttt-online	2015-10-15	online	48	39	81.2

	slug	start	country	attendance	count_badged	pct_completion
32	2015-10-25-instructortraining-cracov	2015-10-25	Poland	13	3	23.1
33	2015-11-23-instructor-training-manchester	2015-11-23	United Kingdom	25	13	52.0
34	2015-12-07-eu-instructor-training	2015-12-07	online	15	15	100.0
35	2015-12-07-na-instructor-training	2015-12-07	online	22	15	68.2
36	2015-12-07-ttt-Au	2015-12-07	online	4	3	75.0
37	2015-12-07-ttt-toronto	2015-12-07	online	17	16	94.1
38	2016-01-05-ok-instructor-training	2016-01-05	United States	19	7	36.8
39	2016-01-13-instructor-training-lausanne	2016-01-13	Switzerland	20	17	85.0
40	2016-01-18-brisbane-instructor-training	2016-01-18	Australia	25	18	72.0
41	2016-01-21-melbourne-instructor-training	2016-01-21	Australia	27	10	37.0
42	2016-01-21-training-florida	2016-01-21	United States	25	16	64.0
43	2016-01-28-auckland-instructor-training	2016-01-28	New Zealand	20	16	80.0
44	2016-02-16-training-online	2016-02-16	online	27	11	40.7
45	2016-02-22-training-ucdavis	2016-02-22	United States	25	16	64.0
46	2016-03-09-ttt-uw	2016-03-09	United States	14	10	71.4
47	2016-04-13-training-online	2016-04-13	online	33	16	48.5
48	2016-04-17-instructor-training-nwu	2016-04-17	South Africa	23	11	47.8
49	2016-05-04-instructor-training-ssi	2016-05-04	United Kingdom	15	6	40.0
50	2016-05-11-ttt-compute-canada	2016-05-11	Canada	27	21	77.8
51	2016-05-18-ttt-online	2016-05-18	online	23	11	47.8
52	2016-06-08-ttt-arizona	2016-06-08	online	17	4	23.5
53	2016-06-08-ttt-online	2016-06-08	online	7	5	71.4
54	2016-06-30-Oslo-ttt	2016-06-30	Norway	22	10	45.5
55	2016-07-11-ttt-scipy	2016-07-11	United States	21	7	33.3
56	2016-09-19-ttt-cambridge	2016-09-19	United Kingdom	30	13	43.3
57	2016-09-19-ttt-osu	2016-09-19	United States	9	4	44.4
58	2016-09-28-ttt-sesync	2016-09-28	United States	17	3	17.6
59	2016-10-03-ttt-online	2016-10-03	online	23	13	56.5
60	2016-10-20-ttt-online	2016-10-20	online	15	5	33.3
61	2016-10-24-ttt-uwmadison	2016-10-24	United States	25	14	56.0
62	2016-10-27-ttt-online	2016-10-27	online	20	11	55.0
63	2016-11-03-ttt-cuny	2016-11-03	United States	12	3	25.0

	slug	start	country	attendance	count_badged	pct_completion
64	2016-11-07-ttt-online	2016-11-07	online	10	5	50.0
65	2016-11-15-ttt-online	2016-11-15	online	15	5	33.3
66	2016-11-24-ttt-metoffice	2016-11-24	United Kingdom	17	3	17.6
67	2016-11-29-ttt-nesi	2016-11-29	New Zealand	10	2	20.0
68	2016-12-01-PNNL-ttt	2016-12-01	United States	14	4	28.6
69	2016-12-12-ttt-uiuc	2016-12-12	United States	12	2	16.7
70	2016-12-13-ttt-ucsf	2016-12-13	United States	18	6	33.3
71	2016-12-14-ttt-online	2016-12-14	New Zealand	11	4	36.4
72	2016-12-15-ttt-online	2016-12-15	online	16	6	37.5
73	2017-01-11-ttt-online	2017-01-11	Canada	23	10	43.5
74	2017-01-19-ttt-ecp	2017-01-19	United States	20	8	40.0
75	2017-01-25-ttt-nz-otago	2017-01-25	New Zealand	28	14	50.0
76	2017-02-09-ttt-usu	2017-02-09	United States	11	7	63.6
77	2017-02-21-online-instructor-training	2017-02-21	online	14	8	57.1
78	2017-03-07-ttt-online	2017-03-07	online	13	7	53.8
79	2017-03-16-ttt-oxford	2017-03-16	United Kingdom	24	13	54.2
80	2017-03-24-ttt-upr	2017-03-24	Puerto Rico	12	2	16.7
81	2017-04-10-uw-ttt	2017-04-10	United States	15	7	46.7
82	2017-04-24-ttt-VT_NIST	2017-04-24	online	12	11	91.7
83	2017-04-30-eResearchAfrica-ttt	2017-04-30	South Africa	28	17	60.7
84	2017-05-04-portland-ttt	2017-05-04	United States	28	17	60.7
85	2017-05-04-ttt-frb	2017-05-04	United States	16	8	50.0
86	2017-05-18-stanford-TTT	2017-05-18	United States	13	13	100.0
87	2017-05-22-florida-TTT	2017-05-22	United States	22	10	45.5
88	2017-06-19-sydney-ttt	2017-06-19	Australia	18	11	61.1
89	2017-06-19-ttt-davis	2017-06-19	United States	13	12	92.3
90	2017-06-26-ttt-online	2017-06-26	online	6	4	66.7
91	2017-07-10-ttt-online	2017-07-10	online	21	15	71.4
92	2017-07-13-jaxlab-instructor-training	2017-07-13	online	7	5	71.4
93	2017-07-13-ttt-online	2017-07-13	online	16	7	43.8
94	2017-07-25-tucson-ttt	2017-07-25	United States	9	5	55.6
95	2017-07-27-qcif	2017-07-27	Australia	23	13	56.5

	slug	start	country	attendance	count_badged	pct_completion
96	2017-07-31-ttt-online	2017-07-31	online	27	15	55.6
97	2017-08-07-jackson-ttt	2017-08-07	online	34	22	64.7
98	2017-08-14-UCB-ttt	2017-08-14	United States	11	6	54.5
99	2017-08-16-ttt-online	2017-08-16	online	21	9	42.9
100	2017-09-04-ttt-manchester	2017-09-04	United Kingdom	22	13	59.1
101	2017-09-21-ttt-online	2017-09-21	online	26	10	38.5
102	2017-09-25-TTT-online	2017-09-25	online	17	8	47.1
103	2017-09-28-curtin	2017-09-28	Australia	16	5	31.2
104	2017-10-09-NWU-eResearch-ttt	2017-10-09	South Africa	34	7	20.6
105	2017-10-10-ttt-doe	2017-10-10	online	24	5	20.8
106	2017-10-16-Monsanto-ttt	2017-10-16	United States	7	6	85.7
107	2017-10-19-ucsf-ttt	2017-10-19	online	5	3	60.0
108	2017-10-24-ttt-online	2017-10-24	online	19	11	57.9
109	2017-10-24-ttt-online-USA	2017-10-24	online	23	10	43.5
110	2017-10-27-ttt-UND	2017-10-27	United States	5	1	20.0
111	2017-10-30-TTT-online	2017-10-30	online	12	8	66.7
112	2017-10-30-imperial-ttt	2017-10-30	United Kingdom	19	12	63.2
113	2017-11-02-TTT-online	2017-11-02	online	25	11	44.0
114	2017-11-06-ttt-Utrecht	2017-11-07	Netherlands	16	6	37.5
115	2017-11-14-ttt-online	2017-11-14	online	12	9	75.0
116	2017-11-15-ttt-stanford	2017-11-15	United States	14	7	50.0
117	2017-12-04-ttt	2017-12-04	online	8	3	37.5
118	2017-12-05-ttt-southampton	2017-12-05	United Kingdom	18	9	50.0
119	2017-12-12-ttt-online	2017-12-12	online	17	12	70.6
120	2017-12-14-ttt-uiuc	2017-12-14	United States	10	4	40.0
121	2018-01-30-ttt-USA	2018-01-30	online	19	15	78.9
122	2018-01-30-ttt-online-UK	2018-01-30	online	14	8	57.1
123	2018-02-05-ttt-online	2018-02-05	online	24	9	37.5
124	2018-02-13-ttt-online	2018-02-13	online	15	7	46.7
125	2018-02-21-South-Africa-ttt	2018-02-21	South Africa	17	9	52.9
126	2018-02-26-ttt-csiro	2018-02-26	Australia	4	1	25.0
127	2018-02-26-ttt-stanford	2018-02-26	United States	12	1	8.3

	slug	start	country	attendance	count_badged	pct_completion
128	2018-02-28-ttt-online	2018-02-28	online	16	5	31.2
129	2018-03-01-ttt-online	2018-03-01	online	18	4	22.2
130	2018-03-05-ttt-florida	2018-03-05	United States	20	6	30.0
131	2018-03-12-ttt-online	2018-03-12	online	18	3	16.7
132	2018-03-19-ttt-sfu	2018-03-19	Canada	10	1	10.0
133	2018-03-20-ttt-online	2018-03-20	online	11	3	27.3
134	2018-03-26-ttt-davis	2018-03-26	United States	24	1	4.2

Table 9: Online and and inperson instructor training events

This table shows the total number of online and in-person training events each year. Numbers for 2018 represent actual data, not a projection.

	online	in-person
year		
2012	2	0
2013	5	0
2014	10	0
2015	12	9
2016	11	24
2017	21	27
2018	8	6

Table 10: Instructor Trainings by Country

This table lists all Instructor Training events held each year. Events listed by country are in-person events. All online events are listed as online, even if all trainees came from the same country.

country	Australia	Canada	Netherlands	New Zealand	Norway	Poland	Puerto Rico	South Africa	Switzerland	United Kingdom	United States	online	total
year													
2012	0	0	0	0	0	0	0	0	0	0	0	2	2
2013	0	0	0	0	0	0	0	0	0	0	0	5	5
2014	0	0	0	0	0	0	0	0	0	0	0	10	10
2015	0	1	0	0	0	1	0	0	0	4	3	12	21
2016	2	1	0	3	1	0	0	1	1	3	12	11	35
2017	3	1	1	1	0	0	1	2	0	4	14	21	48
2018	1	1	0	0	0	0	0	1	0	0	3	8	14

Table 11: Completion rates by year

For each year from 2012 to 2017 this shows the following:

variable	definition
count	Number of training events that year
mean	Mean (average) completion rates
std	Standard deviation
min	Smallest completion rate
25%, 50%, 75%	1st, 2nd, and 3rd quartile completion rate
max	Largest completion rate

		pct_completion
year		
2012	count	2
	mean	80
	std	6
	min	76
	25%	78
	50%	80
	75%	82
	max	85
2013	count	5
	mean	49
	std	12
	min	38
	25%	40
	50%	41
	75%	58
	max	66
2014	count	10
	mean	41
	std	18
	min	8
	25%	37
	50%	43
	75%	49
	max	72
2015	count	21
	mean	58
	std	21
	min	23
	25%	46
	50%	52
	75%	72

		pct_completion
year		
	max	100
2016	**count**	35
	mean	45
	std	18
	min	16
	25%	33
	50%	43
	75%	56
	max	85
2017	**count**	48
	mean	53
	std	18
	min	16
	25%	43
	50%	54
	75%	63
	max	100

Table 12: Days between badging and teaching

For each year from 2012 to 2018 this shows the following:

variable	definition
count	Number of instructors receiving a badge that year who have also taught at least one workshop
mean	Mean (average) number of days between receiving badge and first teaching experience
std	Standard deviation
min	Smallest number of days between badging and teaching
25%, 50%, 75%	1st, 2nd, and 3rd quartile number of days between badging and teaching
max	Largest number of days between badging and teaching

		days
badge_year		
2012	**count**	33
	mean	-165
	std	130
	min	-396
	25%	-244
	50%	-183
	75%	-61
	max	121
2013	**count**	65
	mean	-40
	std	197
	min	-426
	25%	-181
	50%	-61
	75%	62
	max	669
2014	**count**	144
	mean	61
	std	275
	min	-730
	25%	-61
	50%	46
	75%	153
	max	1186
2015	**count**	150
	mean	105
	std	216
	min	-485
	25%	30
	50%	90
	75%	184

		days
badge_year		
	max	1065
2016	**count**	278
	mean	20
	std	244
	min	-1096
	25%	-60
	50%	31
	75%	122
	max	670
2017	**count**	223
	mean	8
	std	216
	min	-1127
	25%	-60
	50%	31
	75%	151
	max	516

Table 13: Instructor teaching frequency

The table below shows how many badged instructors have never taught, taught 1 workshop, 2-5 workshops, 6-10 workshops, 11-15 workshops, 16-20 workshops, and 21 or more workshops. The left side of each bin is exclusive; the right side is inclusive.

		count
	workshops taught	
count	**(-1, 0]**	526
	(0, 1]	326
	(1, 5]	452
	(5, 10]	116
	(10, 15]	37
	(15, 20]	13
	(20, inf]	10

Table 14: Instructor Trainers by year joined

The table below shows how many new Instructor Trainers joined each year. Numbers for 2018 are actual data, not a projection.

	count
year	
2012	1
2016	14
2017	29
2018	14

Table 15: Trainers by Country

The table below lists how many Instructor Trainers we have in total in each country.

	name	count
0	Australia	4
1	Belgium	1
2	Canada	4
3	Ethiopia	1
4	Germany	1
5	Greece	1
6	Italy	1
7	Namibia	1
8	Netherlands	2
9	New Zealand	4
10	Norway	3
11	South Africa	4
12	United Kingdom	9
13	United States	22