

# Data Science Center - Instruction Statistics

10 January, 2020

```
#dscstats <- read_csv('data/dsc_stats_clean.csv') # cleaned via openrefine
dscstats <- read_csv('data/workshops_events.csv') #incl. external workshops, inst training

## Parsed with column specification:
## cols(
##   `Event ID` = col_character(),
##   `title-norm` = col_character(),
##   Date = col_character(),
##   Attendance = col_double(),
##   `review-cat` = col_character(),
##   URL = col_character(),
##   Presenter = col_character(),
##   Audiences = col_character(),
##   `Waiting-List Registrations` = col_double(),
##   TYPE = col_character(),
##   `date format` = col_character(),
##   year = col_double(),
##   Title = col_character()
## )

dscstats <- clean_names(dscstats)

dscstats %>%
  count(event_id,title,date)
```

The Data Science Center created and manages an educational program that address researcher, staff, and librarian skills gaps in data best practices, reproducible research, and foundation coding skills. Through our membership with the Carpentries, a diverse, global community of volunteer instructors, we also contribute to data/coding curricula and train-the-trainer pedagogical activities nationally and international. From 2017 to 2020, this program has offered 63 workshops or educational events to UCLA researchers and partners in

```
dscstats %>%
  mutate(date = mdy(date)) %>%
  group_by(quarter(date, with_year = TRUE, fiscal_start = 7)) %>%
  tally(attendance)
```

```
## # A tibble: 11 x 2
##   `quarter(date, with_year = TRUE, fiscal_start = 7)`      n
##                                     <dbl> <dbl>
## 1                                     2017.    93
## 2                                     2017.    72
## 3                                     2017.    94
## 4                                     2018.    46
## 5                                     2018.   165
## 6                                     2018.   154
## 7                                     2018.   121
## 8                                     2019.    53
## 9                                     2019.    56
## 10                                    2019.   179
## 11                                    2019.   99
```

## Attendance by Data Science Tool/Topic by Role

To get a sense of what we teach and what our audience for instruction is:

We also teach with a range of tools, including R, Python, OpenRefine (a data cleaning tool) and LARIAC (a geospatial imagery data tool for LA). We also offer data-focused courses that don't include a specific **tool**, so we include below as **data**:

Our attendance by role or status of attendance:

## Department Reach

For UCLA learners, workshop attendees come from `r length(unique(statsdept[statsinstitution == "UCLA"]))` different UCLA Departments, Schools or Centers:

## Global, National, and Regional Reach

As a member of the Carpentries, we also contribute to their global mission to improve researchers' data and computing practices. We do this by developing lessons, training future trainers, or teaching workshops at other institutions. For example, as part of the **Carpentries African Taskforce**, the DSC was invited to South Africa to teach workshops for South African Librarians and also teach future instructors from other countries in Africa on how to effectively teach coding and data practices. We additionally invite other universities and organizations to some of our workshops. Finally, we teach Carpentries organized train-the-trainer events in person and online. Below, reflects the diversity of institutional attendance of our workshops and reflects the global contribution UCLA makes to improving data practices in the larger research communities.