Every Chemist a Programmer

K Cowtan
VICEPHEC 2021
History...

Nearly 2 decades of teaching programming to chemists...

- **2004-2012**: Python Y4 option module, 6 lectures + 4 workshops
- **2014-2019**: R Y3 option module, 4 lectures + 3 workshops
- **2021-**: Python Y1 skills, 90 mins mini lectures + exercises

**Approach**: Initially very traditional, building up from basic elements.

**Outcomes**: <20% of students would get to the point where they could use programming in their projects/labs. That subset biased male.
New course context...

Plan:
● Delivered to all Y1 chemists
● Limited to 6 hours of contact
  ○ 2 x 1 hour lectures
  ○ 2 x 2 hour workshops

Covid context
● Delivered electronically
● 12 short videos
  ○ (total 90 mins)
● Online exercises supported by quizzes and Q&A forums
import pandas as pd
concs = pd.read_csv("reaction1.csv")
print( concs.describe() )
concs.plot(x="time")
```python
# photometer file is simple CSV

d1 = pd.read_csv("Photometer_0.025_M.csv")
d1.plot(x="Time_s")

# IR data has no headers

d2 = pd.read_csv("CTLRPC1_20201011_154920.csv", names=["Wavelength","Transmittance"]

d2.plot(x="Wavelength")

# UVJ data need to skip 8 header rows

d3 = pd.read_csv("DATA_UVJ_20190502_170757.csv")
d3.plot(x="Wavelength (nm)"

# JCamp-dx file - space separator and # for


d4 = pd.read_csv("Iron_Complex.dx", sep="\s+", names=["Wavelength","Absorbance"]

d4.plot(x="Wavelength")

# ElNino data is fixed format

d5 = pd.read_fwf("ELNINO_wkst8110.for", skiprows=3, widths=[10,9,4,9,4,9,4,9,4])
d5.plot(x="Week")

print(d5.describe())
```

```
In [8]:
```
```python
import pandas as pd
import numpy as np
import statsmodels.formula.api as sm

# read data
concs = pd.read_csv("reaction1.csv")
concs.plot(x="time")

# Linearize for first order
concs["logconc"] = np.log(concs["conc"])

# perform ols fit
fit = sm.ols("logconc ~ time", concs).fit()
print(fit.summary())

# plot
concs["fitted"] = fit.fittedvalues
ax1 = concs.plot.scatter(x="time",y="logconc"
concs.plot(x="time",y="fitted",ax=ax1)
```

Notes:
[1] Standard Errors assume that the covariance matrix of the errors is correctly specified.

In [9]:
# perform ols fits
fit0 = sm.ols("conc ~ time", concs).fit()
fit1 = sm.ols("logconc ~ time", concs).fit()
fit2 = sm.ols("invconc ~ time", concs).fit()

# gather the rsquared values
rsq0 = fit0.rsquared
rsq1 = fit1.rsquared
rsq2 = fit2.rsquared

# test for 0th order
if rsq0 > rsq1 and rsq0 > rsq2:
    ax = concs.plot(x="time", y="conc")
    print("Reaction is 0th order")
    print("Rate is ", -fit0.params["time"]

# test for 1st order
if rsq1 > rsq0 and rsq1 > rsq2:
    ax = concs.plot(x="time", y="logconc")
    print("Reaction is 1st order")
    print("Rate is ", -fit1.params["time"]

# test for 2nd order
if rsq2 > rsq0 and rsq2 > rsq1:
    ax = concs.plot(x="time", y="invconc")
    print("Reaction is 2nd order")
    print("Rate is ", -fit2.params["time"]

Reaction is 1st order
Rate is  0.016691681472875263
Reaction is 0th order
Rate is  1.9721765553946267
Reaction is 2nd order
Rate is  0.002193416293750461
1. Introduction
2. Python and Spyder
3. History, psychology and sociology of programming
4. Using Spyder
5. Writing a useful program
6. Pulling it to bits
7. Bugs and debugging
8. Reading data
9. Arithmetic with numbers and columns
10. Fitting a line to a graph
11. Conditions
12. Loops
12 short videos

1. Introduction
2. Python and Spyder
3. History, psychology and sociology of programming
4. Using Spyder
5. Writing a useful program
6. Pulling it to bits
7. Bugs and debugging
8. Reading data
9. Arithmetic with numbers and columns
10. Fitting a line to a graph
11. Conditions
12. Loops
History, psychology and sociology of programming

Aims: Understand social factors which might get in the way of learning to program
History of computer programming...

Charles Babbage (1791-1871) built what is considered to be the first mechanical computer.

Ada Lovelace (1815-1852) developed the ideas for programming it.
Early computation...

The "Harvard Computers", Harvard observatory, 1875
Early computation...

US treasury 1920s
Electronic computers...

ENIAC, 1946
Computers becoming programmers...

NASA space program...
Computer programming...

Grace Hopper (1906-1992) developed COBOL, the first modern computer language.

“Programming requires patience and the ability to handle detail. Women are ‘naturals’ at computer programming.”
Margaret Cowtan...

1960s: Vauxhall, Brush
Today...

- Only 2% of top contributors to the main python code sharing site are women.
- Computing and computer programming are now *perceived* to be aptitudes associated with a very small, mostly male minority.
- This is not supported by history, by experience in schools or in other departments at York.


How did this happen?

Hiring policies changed to favour male applicants.

Computer advertising was targeted at boys

- The status of the job increased
- It was increasingly done by men

No longer a "pink collar" job.

Review of "The computer boys take over",
https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=6479433

When computer coding was a women's job:
https://www.history.com/news/coding-used-to-be-a-womans-job-so-it-was-paid-less-and-undervalued
Impacts: some online spaces can be exclusionary...

Stack Overflow (top computing Q&A site) users were asked "What would you like to change?"

- Men*: “official, complex, algorithm”
- Women*: "condescending, rude, assholes"

1. There are distinct gendered communication styles (MacCoby)
2. Male communication is typified by higher levels of social dominance signalling (MacCoby)
3. Male communication styles tend to become dominant in anonymous mixed contexts (Brooke)

Gender and relationships: a developmental account (1990) E. MacCoby

Framing gender and hostility on Stack Overflow. (2019) S. Brooke

* based on self-declared gender for binary respondents
Implications...

If you do not fit the stereotype of a computer programmer, social pressures are working to make computing harder for you than it is.

- **Stereotype threat** increases anxiety levels
- **Performance anxiety** reduces performance

This is *particularly* true for women and gender minorities, racial and ethnic minorities and people with disabilities, but *also* for many men.

This serves to maintain the exclusivity of the activity. (How much this affects you will vary with personality.)
Is there a solution?

9.3% of top contributors to the main R code sharing site are women - nearly 5 times better than python.

- R is taught differently (data science background)
- R has women as core developers
- R ladies (including gender minorities)
- R forwards (including other minorities)

But python is more used in chemistry...
What can we do?

- Be aware of the social pressures.
  - This helps a bit - limited by *G.I. Joe* bias
- Exploit group learning - but pick your groups!
  - Use discord, messenger, google meet
  - Find a group that works for you both intellectually *and* psychologically
- Find problems you are interested in.
  - Find fun data and analyze it (#TidyTuesday). Plot data in interesting ways. Write games
- We will try and enable approaches to learning which...
  - Reach more people and break down stereotypes
  - Reduce inequalities due to different approaches to learning
Did it work?

- "I thought the info on the history of coding was interesting and useful in making the course seem less scary to start."
- "it was nice that he gave some background lectures as well."
- "The video about the history of coding and improving access for women and minorities was excellent as it really helped me to believe that I was capable and it was worth trying properly instead of thinking it just wasn't my thing- turns out I really enjoy coding!"
Why the gender disparity?

- Gendered learning styles?
  - Gregorc style delineator?
Why the gender disparity?

- Because traditional teaching methods, tools, and online communities are structured to make programming hard. Only those who society has primed to believe they can and should be programmers can ever overcome the energy barrier.
  - Solution: draw attention to the barriers, and teach in a way which weakens them.
Personal context...

- Social complexity
- Gender roles
- Social anxiety
- Gender performance
- Teaching methods
- Theory of mind
- Autism
Questions and suggestions?

Thanks to:
- David Pugh,
- Nigel Bailey,
- Mat Evans,
- Aneurin Kennedy,
- Lucy Wheeler,
- Caterina Mapelli,
- Nick Yates,
- Paul Bond,
- Tom Webb,
- Emma Rand
- Julie Wilson
- Susannah Cowtan

And everyone I've forgotten (sorry)!