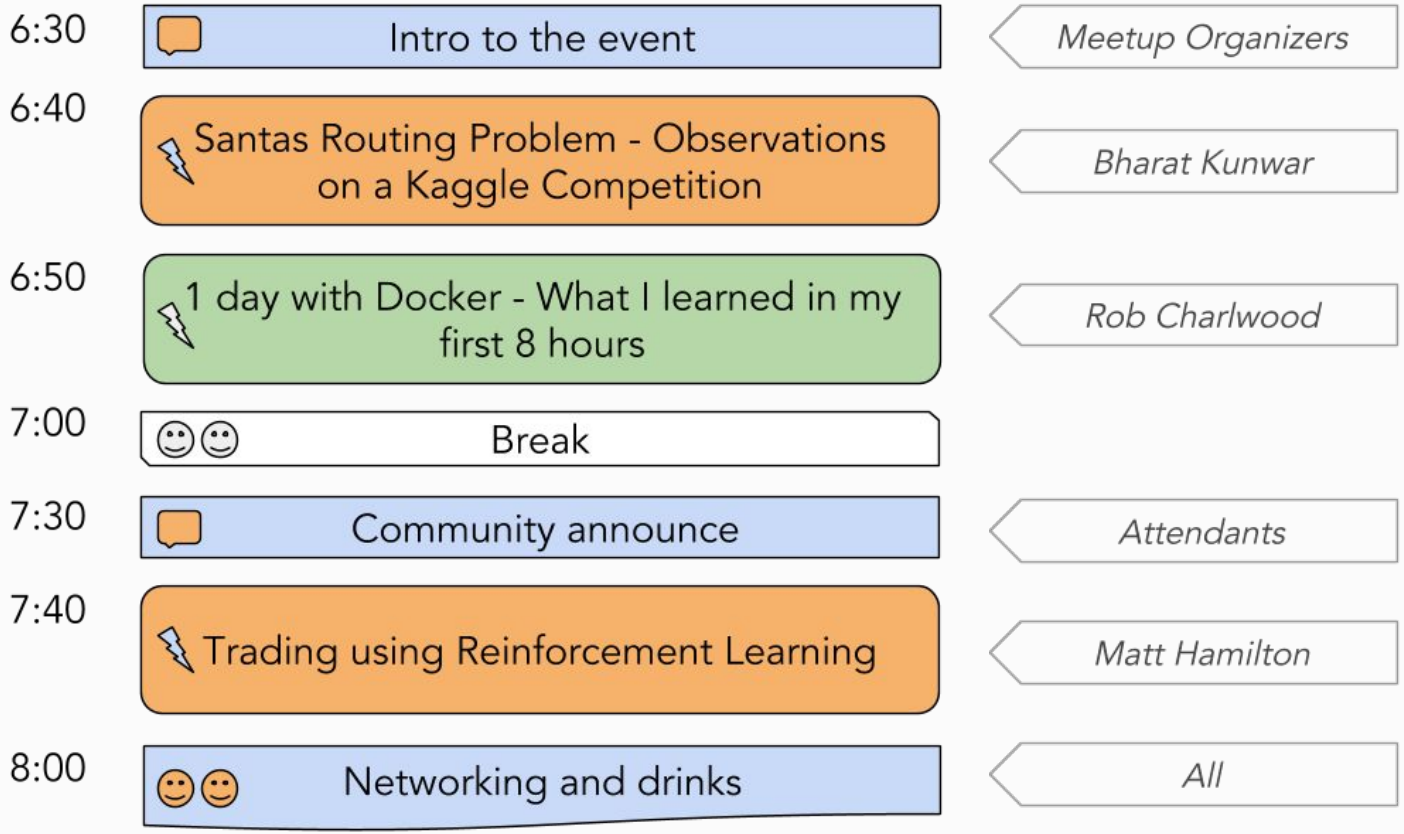
10th December 2018

Joint Christmas special

Frank Kelly, Rob Charlwood and Miquel Perello-Nieto

Christmas Special



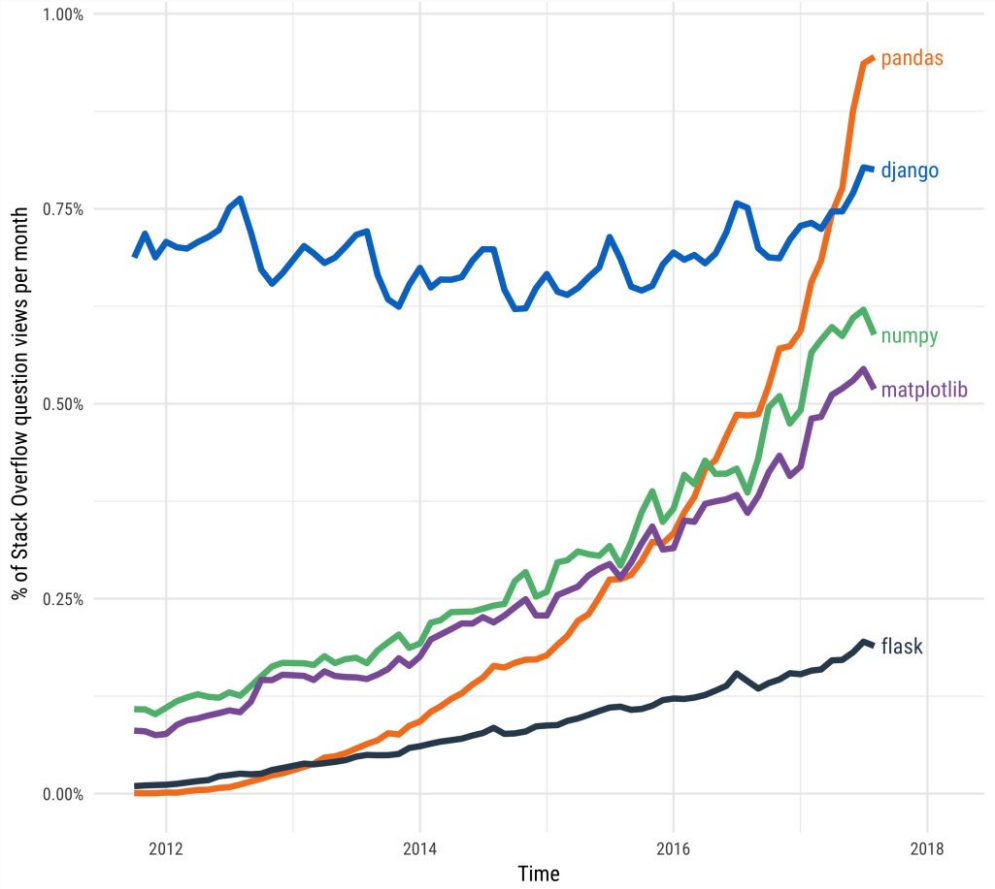
Data Science vs Web Development

PyData Bristol

https://www.dezyre.com/article/python-for-data-science-vs-python-for-web-development/177

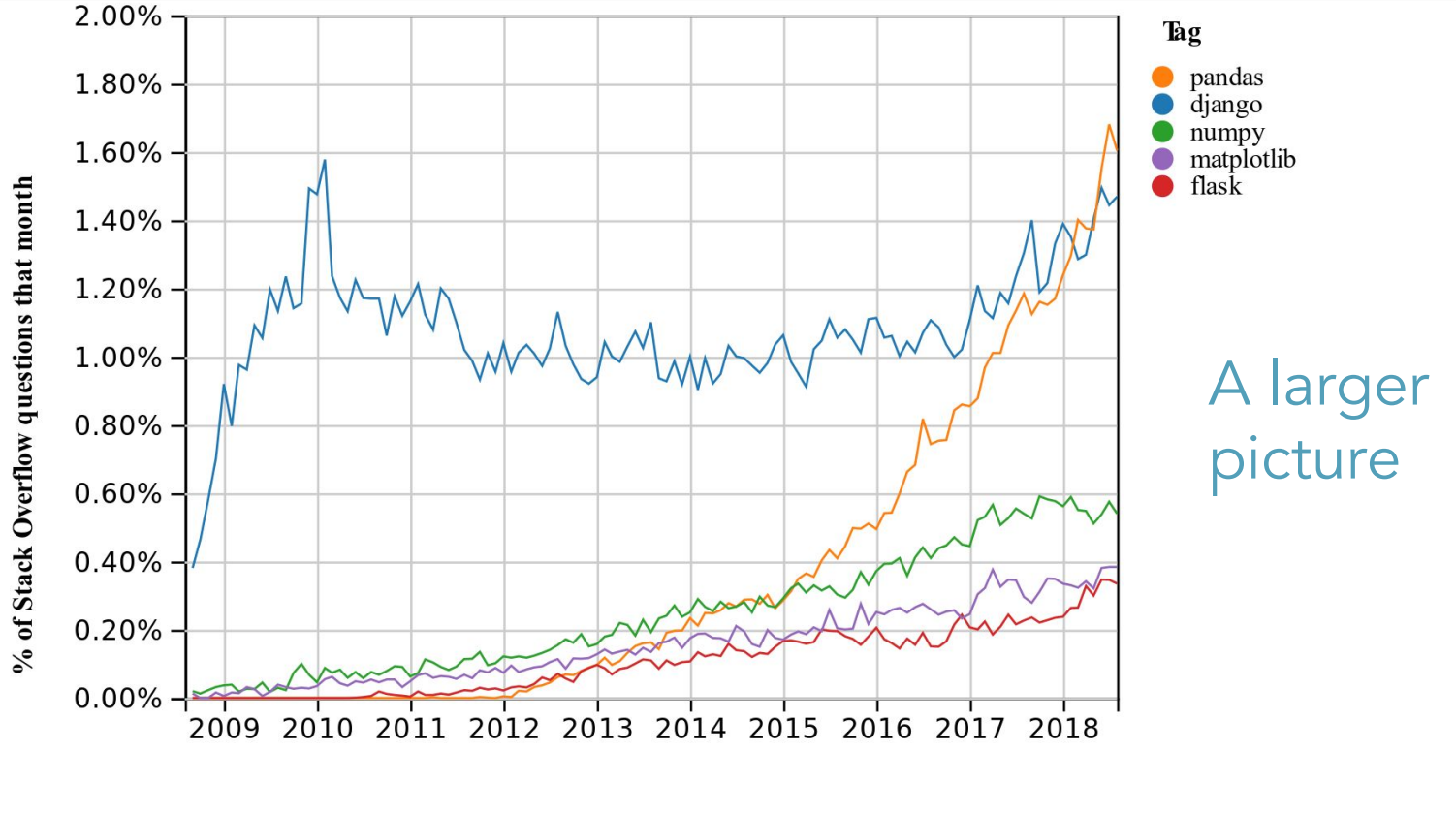
Stack Overflow Traffic to Questions About Selected Python Packages

Based on visits to Stack Overflow questions from World Bank high-income countries



Stack Overflow Traffic to Questions About Selected Python Packages

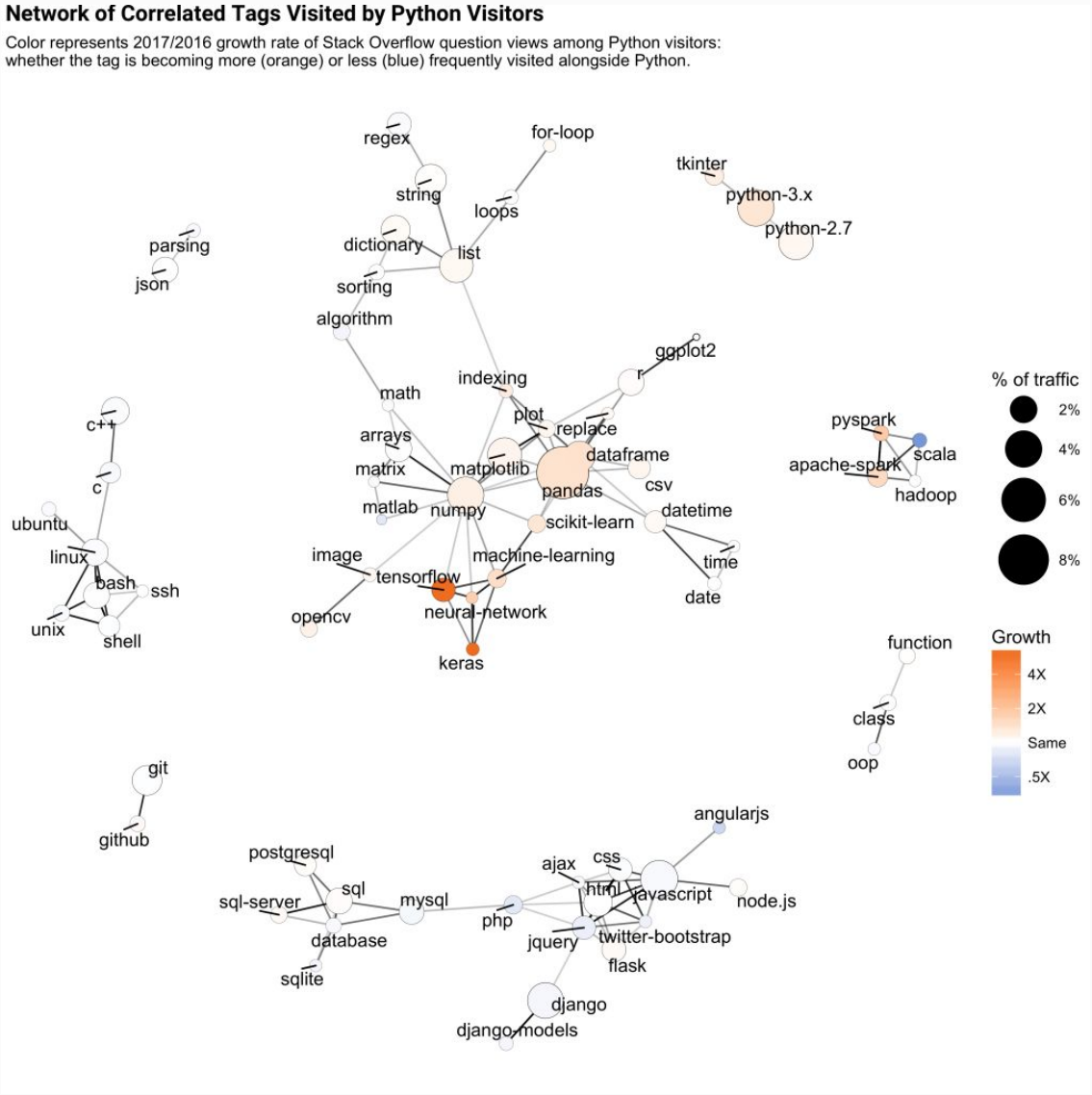
Based on visits to Stack Overflow questions from World Bank high-income countries



A larger picture

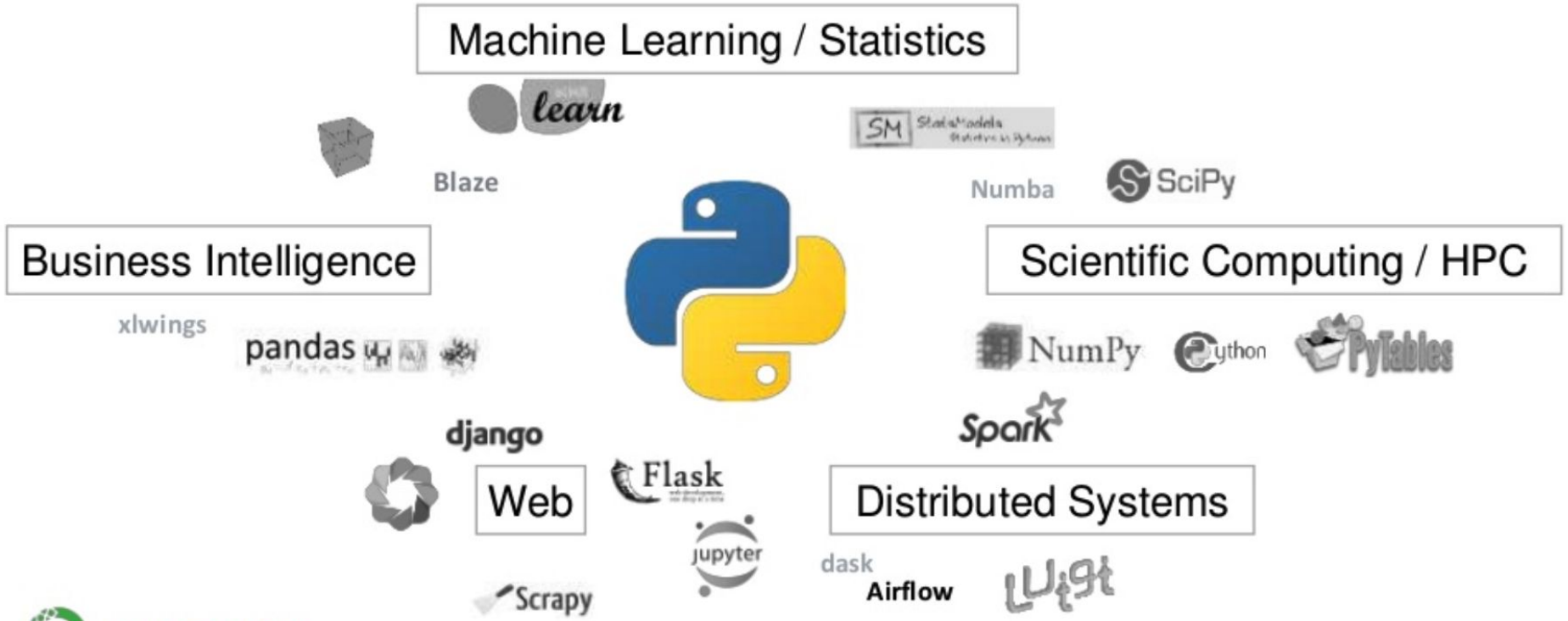
https://insights.stackoverflow.com/trends?tags=pandas%2Cnumpy%2Cmatplotlib%2Cflask%2Cdjango

"Python's popularity in data science and machine learning is probably the main reason for its fast growth"



https://stackoverflow.blog/2017/09/14/python-growing-quickly/

Python is the common language



ANACONDA

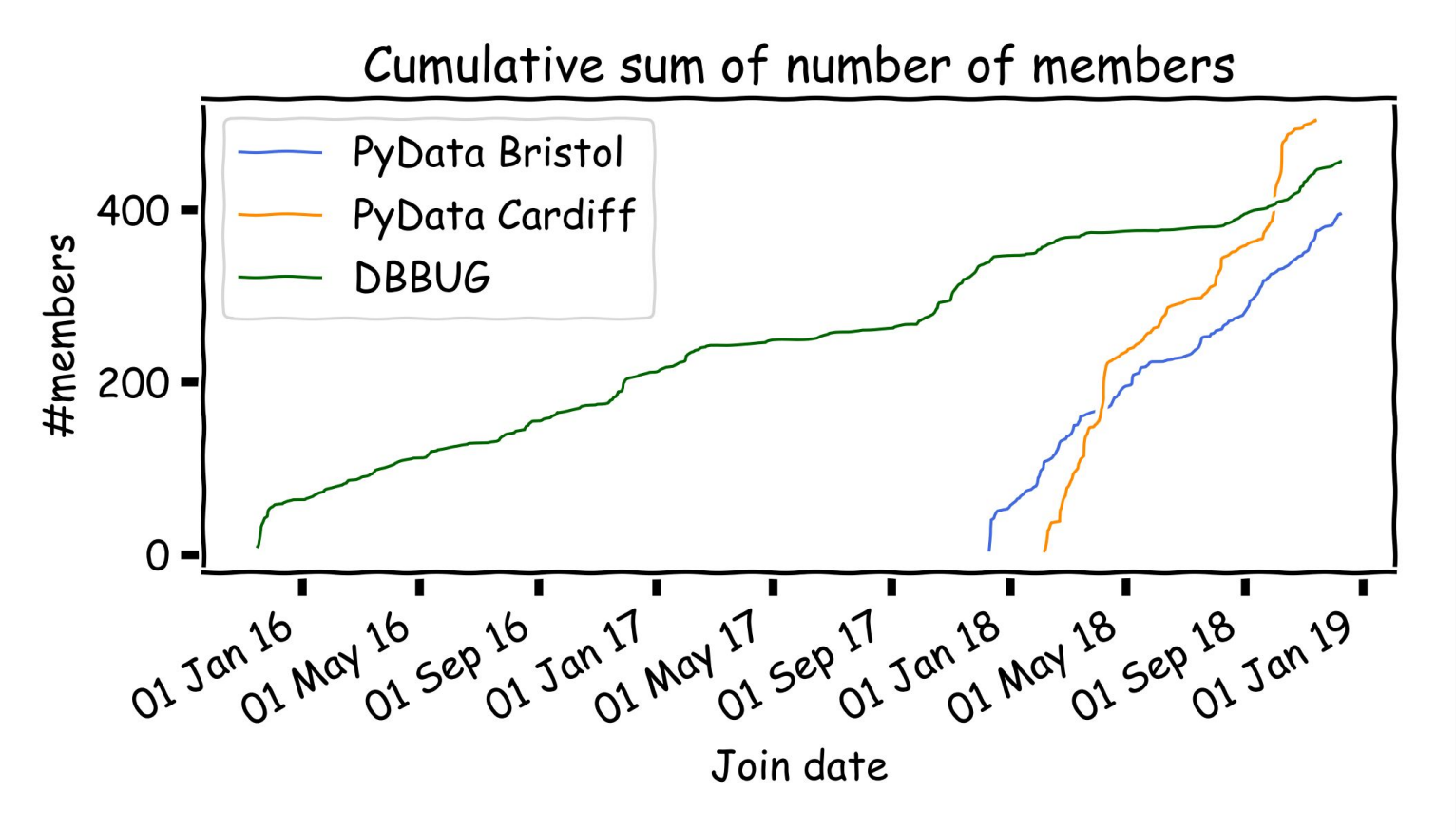
Upcoming events

#PyData Cardiff Meetup (11 December, Tomorrow!)

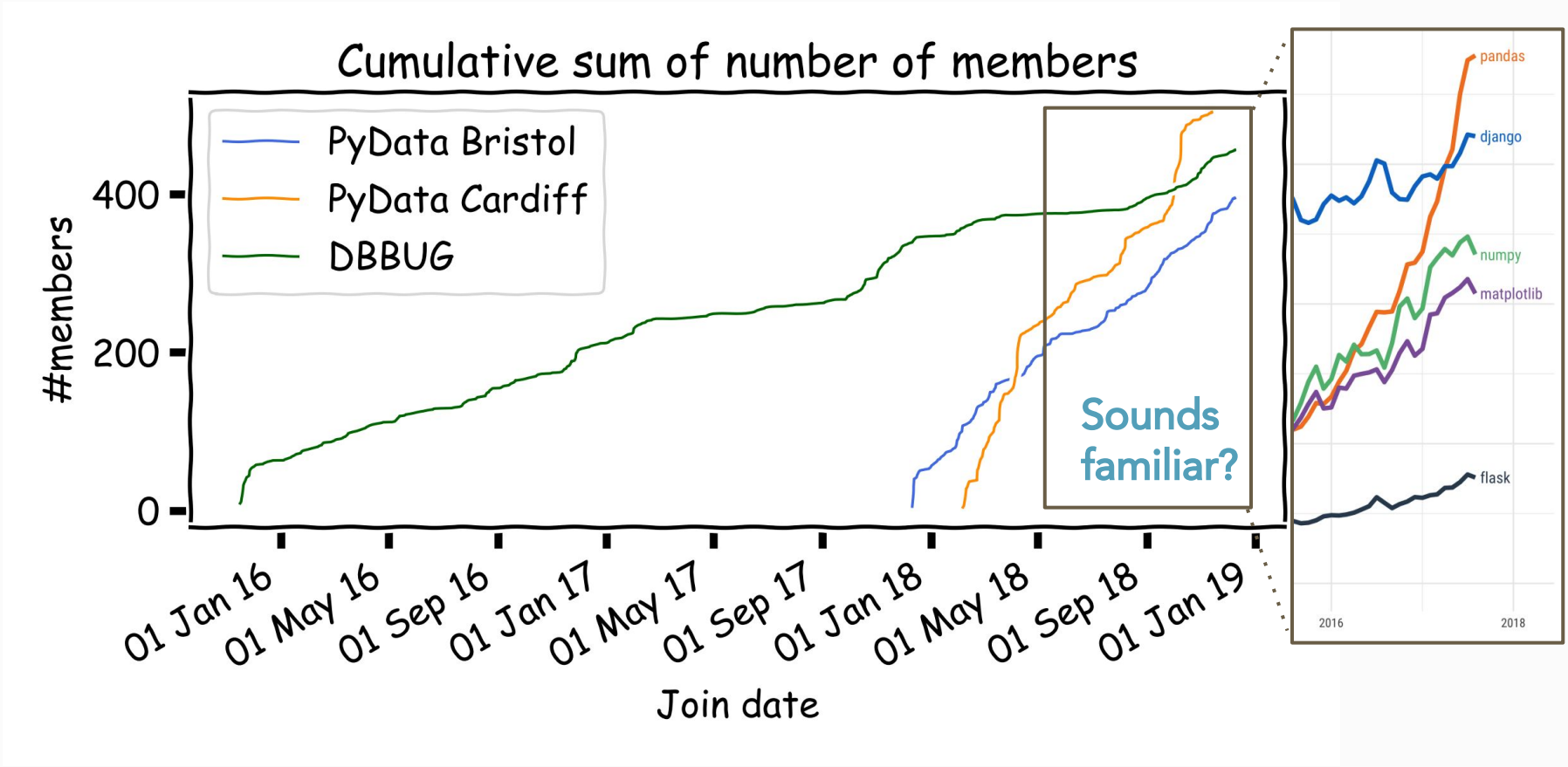
#South West Data Social Meetup (11 December in Bristol, Tomorrow!)

#PyData Bristol Meetup (17 January 2019)

PyData Bristol/Cardiff, DBBUG members



PyData Bristol/Cardiff, DBBUG members



Inspired by R package:

The Xmas package: pyjanitor

Tools for cleaning Pandas DataFrames

Implements method chaining

```python
import pandas as pd

df = pd.DataFrame(...)  # create a pandas DataFrame somehow.
del df['column1']  # delete a column from the dataframe.
df = df.dropna(subset=['column2', 'column3'])  # drop rows that have empty values in column 2 and 3.
df = df.rename({'column2': 'unicorns', 'column3': 'dragons'})  # rename column2 and column3
df['new_column'] = ['iterable', 'of', 'items']  # add a new column.
df.reset_index(inplace=True, drop=True)  # reset index to account for the missing row we removed above
```

The pyjanitor approach

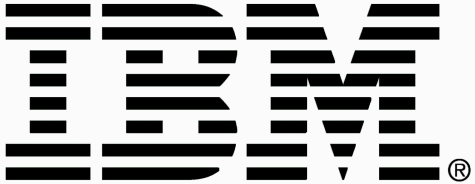
```python
import pandas as pd
import pandas_flavor as pf

df = (
    pd.DataFrame(...)
    .remove_columns(['column1'])
    .dropna(subset=['column2', 'column3'])
    .rename_column('column2', 'unicorns')
    .rename_column('column3', 'dragons')
    .add_column('new_column', ['iterable', 'of', 'items'])
    .reset_index_inplace(drop=True)
)
```

```
pip install pyjanitor
```

https://github.com/ericmjl/pyjanitor

Thanks to our sponsors



- Margriet Groenendijk

Do you want to be a sponsor?
Come and talk to us.