

— Session 05: Exploratory data analysis

wifi: GA-Guest, yellowpencil

```
cd ~/Documents/ga-ldn-ds37  
git commit -am "your commit message here"  
git pull
```



Today's session plan

1800-1820	Standup & Review
1820-1845	Introducing pandas
1845-1900	Selecting rows and columns
1900-1920	Break
1920-2000	Summarising and cleaning data
2000-2100	Grouping data
Homework: Twitter bots exercise & merging data	

At the end of the session, you will be able to ...

Use the pandas library to read in data

Summarise data in pandas

Handle missing values

Produce simple breakdowns and analyses

Intro to Python



Let's Review

At the end of the session, you will be able to ...

Use the pandas library to read in data

Summarise data in pandas

Handle missing values

Produce simple breakdowns and analyses

Coming up next session...

- Data visualisation with matplotlib and seaborn



Homework

Finish the Twitter bots exercise if you haven't had time.

Investigate **merging** and **joining** in Pandas; see if you can merge the 2018/19 gender pay gap data with the 2019/20 data, producing one big DataFrame.



