

Applied Data Analysis

Monday 4th August		
Time	Session	Speakers
09:00-10:30	Opening Keynote See keynotes page	
10:30-11:00	Refreshment break	
11:00-12:30	Introduction to Applied Data Analysis * Presentation of the strand, objectives and schedule * Introductions * Example of a data analysis application * Setting up systems	Ellen Charlesworth Paul Guhennec
12:30-14:00	Lunch	
14:00-15:30	Introduction to Pandas Intro to Pandas: data types (series, dataframes) Data loading	Ellen Charlesworth Paul Guhennec

	Data access (indexing)	
15:30-16:00	Refreshment break	
16:00-17:00	Exercises and catching-up session attendees may pick one from the options below: - Catching-up: assistance is provided to clarify any issue from the previous classes or in setting-up your Python environment. - Project: Attendees should bring their own mini-project to the class and work on it, individually or with others. This typically will require bringing your own data, although we can also provide some sample datasets.	Ellen Charlesworth Paul Guhennec
Tuesday 5th August		
Time	Session	Speakers
09:00-10:30	Tidy data (part I) Basic concepts of (proper/tidy) data modelling	Ellen Charlesworth Paul Guhennec
10:30-11:00	Refreshment break	

11:00-12:30	Tidy data (part II) Manipulating data with Pandas * Operations on tidy data frames: - set, union join - select, apply transform	Ellen Charlesworth Paul Guhennec
12:30-14:00	Lunch	
Project session	Project session Catching-up Work on exercises/mini-projects	Ellen Charlesworth Paul Guhennec
15:30-16:00	Refreshment break	
16:00-17:00	Project session Catching-up Work on exercises/mini-projects	Ellen Charlesworth Paul Guhennec
Wednesday 6th August		
Time	Session	Speakers

09:00-10:30	Exploratory Data Analysis (part I) • Basic plotting • Descriptive statistics • Variation, distributions	Ellen Charlesworth Paul Guhennec
10:30-11:00	Refreshment break	
11:00-12:30	Exploratory Data Analysis (part II) • Descriptive statistics and plotting with Pandas, Matplotlib and Seaborn • Data modelling: working with texts	Ellen Charlesworth Paul Guhennec
12:30-14:00	Lunch	
14:00-15:30	Project session Catching-up Work on exercises/mini-projects	Ellen Charlesworth Paul Guhennec
15:30-16:00	Refreshment break	
16:00-17:00	Project session Catching-up Work on exercises/mini-projects	Ellen Charlesworth Paul Guhennec

Thursday 7th August		
Time	Session	Speakers
09:00-10:30	Applied Data Analysis (part I) • Explorative data visualization • Primer on good and bad data visualization practices	Ellen Charlesworth Paul Guhennec
10:30-11:00	Refreshment break	
11:00-12:30	Applied Data Analysis (part II) • Basics of Natural Language Processing	Ellen Charlesworth Paul Guhennec
12:30-14:00	Lunch	
14:00-15:30	Project session • Catching up • Work on projects	Paul Guhennec
15:30-16:00	Refreshment break	
16:00-17:00	Project session • Catching up	Paul Guhennec

	• Work on projects	
--	--	--

Friday 8th August		
Time	Session	Speakers
09:00-10:30	Applied Data Analysis (part III) Geo-mapping	Paul Guhennec
10:30-11:00	Refreshment break	
11:00-12:30	Publishing datasets & communicating results • Communicating data analysis results • Best practices about publishing datasets, licensing issues, reproducibility, data repositories	Paul Guhennec
12:30-14:00	Lunch	
14:00-15:00	Questions and Open Discussion	Ellen Charlesworth

		Paul Guhenec
15:00-15:30	Refreshment break	
15:30-17:00	Closing keynote See keynotes page	