

# **Data Envelopment Analysis (DEA) and Stochastic Frontier Analysis (SFA) using Python**

**PEMA Professional Development Workshop**



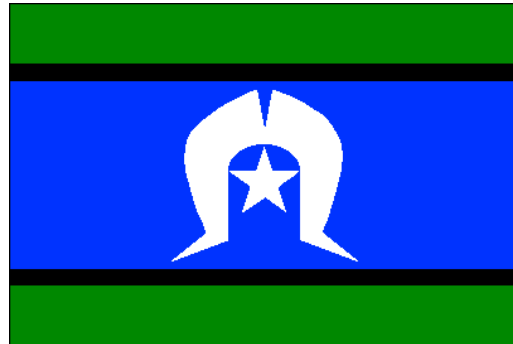
**7 – 9 November 2024**

**The University of Sydney, CBD Campus**

Level 17, 133 Castlereagh Street,  
Sydney, NSW 2006

## Acknowledgement of Country

*We acknowledge the traditional owners of the lands on which the University of Sydney is located, the Gadigal people of the Eora Nation, and we pay our respect to the knowledge embedded forever within the Aboriginal Custodianship of Country.*



## Location Guide for In-person Participants



### The University of Sydney, CBD Campus

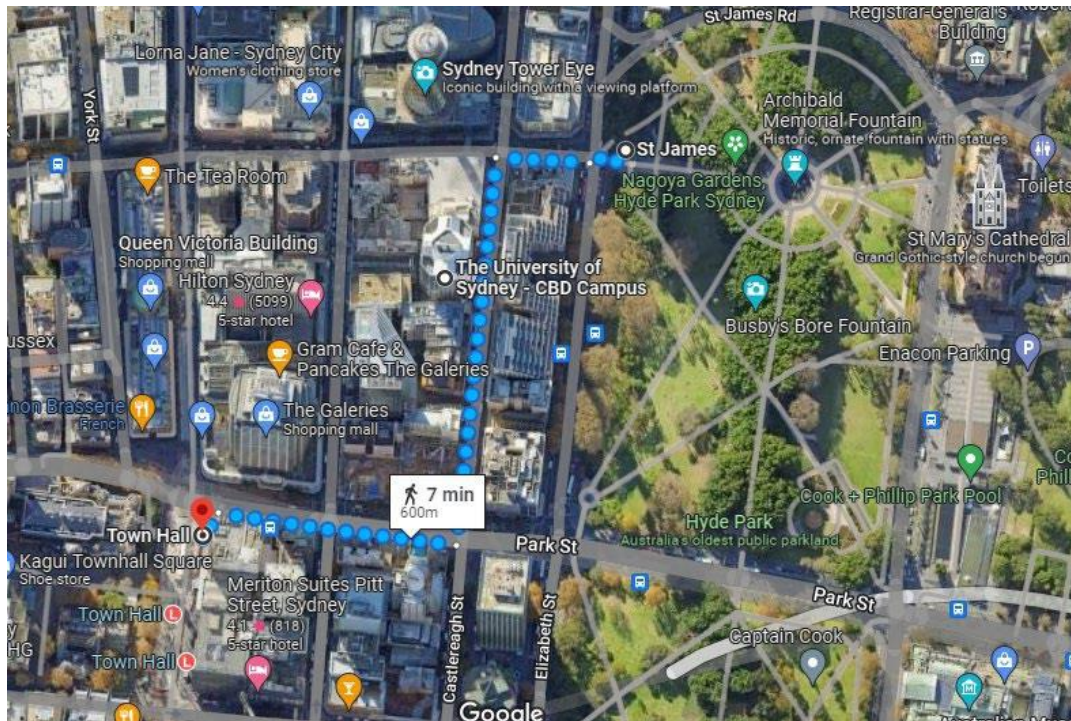
**Address:** Level 17, 133 Castlereagh Street, Sydney, Australia

The University of Sydney CBD campus is located within the Stockland building at 133 Castlereagh Street. The building can be accessed from Castlereagh Street and Pitt Street. It is part of the Piccadilly complex. CBD Campus is open 8am-10pm Monday to Friday, and 8am-6pm Saturday and Sunday.

- **Access on 7<sup>th</sup> and 8<sup>th</sup> November:** 133 Castlereagh Street entrance ('Piccadilly' on Googlemaps) via the ground floor foyer.
  - Use lifts A to F - the right hand set of lifts - to level 17
  - If coming from St James Station, it is the street to your left after Elizabeth St
  - Concierge can assist with directions in the foyer
- **Access on Saturday, 9<sup>th</sup> November:** 210 Pitt Street entrance ('Piccadilly Shopping Centre' on Google maps)
  - Take the escalators to the first-floor "Lunch House Food Court", turn right 180degrees to walk towards Pitt St direction
  - Walk past the Works Perfect camera repair shop and take the left hand set of lifts - Lifts A to F - to level 17
  - If coming from Town Hall Station walk through The Galleries or Hilton Hotel

Any issues with access please call campus reception on 8627 4900

<https://goo.gl/maps/mNidBVpmDXFcFT8j7>



## Getting There

### By train

The nearest train stations are [Town Hall Station](#) or [St James Station](#). Trains to Town Hall Station and St James run regularly from Central Station. See the route on the map above.

### By bus

The nearest bus stops are located on Castlereagh Street, Elizabeth Street after Park Street as identified on the map by the bus stop icons. Several bus routes are available into the CBD.

### By Light rail

The nearest Light Rail stop is [Townhall Light Rail](#).

To best plan your trip by public transport, go to: [Transport NSW](#)

### Driving and CBD parking

133 Castlereagh Street has on site paid parking. Piccadilly Secure Parking basement parking the building provides parking spaces and we suggest pre-booking online to avoid a high cost [Secure Parking Piccadilly](#)

The cheapest CBD parking is available at [GOULBURN STREET CAR PARK](#) located at 101 Goulburn Street, Sydney and we also suggest pre-booking online. The walk from Goulburn Street is 10minutes.

## Computing facilities - Bring your own device/s

Please bring your own laptop. Extension cords, power boards and WiFi password will be provided.

## Accommodation

PEMA does not engage in the administration of temporary accommodation. It is up to you to find suitable living arrangements.

## Workshop Summary

This workshop showcases three days on contemporary and hands-on methodology for productivity and efficiency analysis using one of the most popular data science languages, with applications to various economic agents (e.g., departments, firms, industries, regions, countries, public utilities, etc.).

First day is devoted to an gentle introduction to Python and is taught by [Dr. Jessica Leung](#), Faculty of Business of Economics, Monash University. Doctor Leung is a Lecturer in Business Analytics, specialising in management science, optimisation, machine learning, and data science. She is also an associate investigator at the ARC Industrial Transformation Training Centre in Optimization Technologies, Integrated Methodologies, and Applications (OPTIMA). She is a co-author of the popular [Python Language Companion to Efficiency and Productivity Analysis](#).

Second day is devoted to production theory and data envelopment analysis and is taught by [Professor Valentin Zelenyuk](#), School of Economics, University of Queensland. Professor Zelenyuk is a former ARC Future Fellow and an elected member of the Conference on Research in Income and Wealth (CRIW) group of the National Bureau of Economic Research (NBER). He is a founding member of the International Society for Efficiency and Productivity Analysis (ISEaPA). His research focuses on economic theory of production, econometric estimation and applications. He has co-authored over 70 publications in such leading journals as Operations Research and Journal of Econometrics, and a book with Robin Sickles, Rice University, titled "[Measurement of productivity and efficiency: theory and practice](#)".

Third day is devoted to stochastic frontier analysis is taught by [Professor Artem Prokhorov](#), Discipline of Business Analytics, University of Sydney. Professor Prokhorov is a fellow of the Centre interuniversitaire de recherche en économie quantitative (CIREQ) at University of Montreal, Canada. He is a founding member of the International Society for Efficiency and Productivity Analysis (ISEaPA). His research interests are in the field of theoretical and applied econometrics, with applications in business, finance, risk management, labour and health economics, and have intersections with statistical machine learning and high-dimensional statistics. He has published in such leading journals as Journal of Econometrics and European Journal of Operational Research and has consulted widely for business and government. He is the author of a new book titled "[Efficiency and Productivity Analysis: Using Copulas in Stochastic Frontier Models](#)".

**Day 1** (7 November 2024): *Introduction to Python for Productivity and Efficiency Analysis*, by Jessica Leung, Monash University

This day assumes no knowledge of Python and only basic knowledge of mathematics. Starting from scratch, we will install Python and all the necessary packages that will be used in the next two days. We will cover basic operations with functions contained in those packages, including how to use `matplotlib`, `scipy.optimize`, `statsmodels`, `sklearn.linear_model`. We will then delve into additional fundamental topics such as writing functions in Python and understanding and manipulating Python data structures effectively. By the end of the day, you will be able to write and run Python codes and have packages required for Days 2 and 3. The material in Day 1 is of interest independently of Days 2 and 3, as a hands-on workshop for those looking for a gentle introduction to one of the most popular programming languages.

**Day 2** (8 November 2024): *Production theory and Data Envelopment Analysis*, by Valentin Zelenyuk, University of Queensland

This day assumes basic knowledge of economics, statistics, and mathematics and a working knowledge of Python at a level equivalent to Day 1. The day is composed of four sessions. The first session is focused on axiomatic approach to production theory, covering the essence of its economic foundation and mathematical structure. The second session is focused on various methodological approaches to implementing DEA. The third session is focused on statistical foundations of DEA. The fourth session is about implementing DEA in Python. By the end of Day 2, you will understand the axiomatic approach to production theory and be able to obtain and interpret DEA output in Python and perform related statistical analysis.

**Day 3** (9 November 2024): *Stochastic Frontier Analysis*, by Artem Prokhorov, University of Sydney

This day assumes basic knowledge of economics, statistics, and mathematics and a working knowledge of Python at a level equivalent to Day 1. We start with the estimation of cross-sectional stochastic frontier models (SFM) with multiple inputs. Then, we include environmental variables that affect inefficiency scores, allow for time varying technical inefficiency and for unobserved effects and we consider estimation of stochastic frontier models using panel data. Finally, we cover applications in agriculture, power generation, efficiency of airlines, coal mines, and banks. The last session will be a Python tutorial showcasing how the methods are used in practice. By the end of Day 3, you will be able to estimate SFMs, test hypotheses about them, interpret the estimates and obtain efficiency rankings.

# Program

## Day 1: Thursday 7 November 2024

8.30 – 9.00

**Registration and Welcome tea & Coffee**

<b>Session 1</b> <b>Dr Jessica Leung</b> <b>Lecturer, Department of Econometrics and Business Statistics, Monash University</b>	
9:00 – 10:30	<b>Python installation</b> <i>Installation of Anaconda Python and necessary packages used in DEA and SFA</i>

10:30 – 11:00

**Morning Tea**

<b>Session 2</b> <b>Dr Jessica Leung</b> <b>Lecturer, Department of Econometrics and Business Statistics, Monash University</b>	
11:00 – 12:30	<b>Python Basics</b> <i>Introduction to Python and basic operations such as writing a function and understanding data structures</i>

12:30 – 13:30

**Lunch**

<b>Session 3</b> <b>Dr Jessica Leung</b> <b>Lecturer, Department of Econometrics and Business Statistics, Monash University</b>	
13:30 – 15:00	<b>Practical guide for Python packages - I</b> <i>Introduction to <code>scipy.optimize</code> and <code>matplotlib</code></i>

15:00 – 15:30

**Afternoon Tea**

<b>Session 4</b> <b>Dr Jessica Leung</b> <b>Lecturer, Department of Econometrics and Business Statistics, Monash University</b>	
15:30 – 17:00	<b>Practical guide for Python packages - II</b> <i>Introduction to <code>sklearn.linear_model</code> and <code>statsmodel</code></i>

Kindly note: The precise content per session is subject to reshuffling and fine-tuning.

## Day 2: Friday 8 November 2024

8.30 – 9.00

Registration and Welcome tea & Coffee

<b>Session 1</b> <b>Prof Valentin Zelenyuk</b> <b>ARC Future Fellow, School of Economics, University of Queensland</b>	
9:00 – 10:30	<b>Foundations of productivity and efficiency analysis</b> <i>Characterizations of technology via functions and sets, axiom of production theory, modelling returns to scale, profit functions, revenue functions, cost functions, Shephard's distance functions, Farrell (radial) technical efficiency measures, allocative efficiency measures, other performance measures</i>

10:30 – 11:00

Morning Tea

<b>Session 2</b> <b>Prof Valentin Zelenyuk</b> <b>ARC Future Fellow, School of Economics, University of Queensland</b>	
11:00 – 12:30	<b>Foundations of DEA</b> <i>Activity analysis modelling, Data Envelopment Analysis: Farrell's approach, Data Envelopment Analysis: CCR approach, returns to scale in DEA, handling Big-Wide Data in DEA.</i>

12:30 – 13:30

Lunch

<b>Session 3</b> <b>Prof Valentin Zelenyuk</b> <b>ARC Future Fellow, School of Economics, University of Queensland</b>	
13:30 – 15:00	<b>Selected Topics in DEA</b> <i>Statistical foundations for DEA: analysis of densities, analysis of aggregate efficiency (bootstrap, bias correction, and new CLTs), regression of DEA-estimated efficiency on covariates (truncated regression with bootstrap).</i>

15:00 – 15:30

Afternoon Tea

<b>Session 4</b> <b>Dr Jessica Leung</b> <b>Lecturer, Department of Econometrics and Business Statistics, Monash University</b>	
15:30 – 17:00	<b>DEA in Python</b> <i>Hands-on tutorial on implementing DEA in Python</i>

Kindly note: The precise content per session is subject to reshuffling and fine-tuning.



## Day 3: Saturday 9 November 2024

8.30 – 9.00

**Registration and Welcome tea & Coffee**

<b>Session 1</b> <b>Prof Artem Prokhorov</b> <b>Discipline of Business Analytics, University of Sydney</b>	
9:00 – 10:30	<b>SFA for cross-sections: COLS, MLE</b> <i>Parameter interpretation; Corrected Ordinary Least Squares (COLS); Maximum Likelihood Estimation (MLE); alternative distributional assumptions; specification testing; prediction of technical inefficiency</i>

10:30 – 11:00

**Morning Tea**

<b>Session 2</b> <b>Prof Artem Prokhorov</b> <b>Discipline of Business Analytics, University of Sydney</b>	
11:00 – 12:30	<b>SFA for cross-sections: determinants of inefficiency</b> <i>Inclusion of determinants of inefficiency; environmental variables; non-zero mean; heteroskedasticity; testing nested model; scaling property; marginal effects; alternative distributions</i>

12:30 – 13:30

**Lunch**

<b>Session 3</b> <b>Prof Artem Prokhorov</b> <b>Discipline of Business Analytics, University of Sydney</b>	
13:30 – 15:00	<b>SFA for panels: LSDV, RE, FE, MSLE</b> <i>Data structure; Least Squares Dummy Variable (LSDV) estimation; Fixed and Random Effects (FE and RE); prediction of technical inefficiency; time-varying inefficiency; Maximum (Simulated) Likelihood Estimation (MSLE)</i>

15:00 – 15:30

**Afternoon Tea**

<b>Session 4</b> <b>Dr Jessica Leung</b> <b>Lecturer, Department of Econometrics and Business Statistics, Monash University</b>	
15:30 – 17:00	<b>SFA in Python</b> <i>Hands-on tutorial on implementing SFA in Python</i>

Kindly note: The precise content per session is subject to reshuffling and fine-tuning.