

方法論 2019: 大數據經濟分析(Econinformatics)

由於資訊科技突飛猛進，經濟問題日益複雜；傳統的計量模型顯得力有未逮。非但不能預測紛至沓來的經濟衝擊，甚至連剛發生的金融危機也解釋不清。其中，無法掌控「大數據」之趨勢，是經濟方法發展失焦的主因、

大數據，包含傳統量數—大多數經由測度而來(如經濟學的 GNP、通膨)以及源於紀錄的現代據數(監視器、儀表掃描、手機中的圖像)或者兩者之混合體；除了量大，基本上還經常是很複雜—不易觀察或不易理解。

本課程就是要應用電腦科學及資訊科技於經濟學。在傳統的數量方法架構下，建構出新概念及技巧；以便發掘、管理、運作 Process 及分析利用模型。修習者通過課堂講授及實作(家教式之分組討論教學)，循序漸進，應能充分掌握新知，站在大數據時代的前端。

先修: 統計、微積分、貨銀(總經)，略知計量模型及稍具基本電腦常識(裝置軟體)。

軟體: SAS 9.4 (業界最通用套裝。政大有最新版) 為主 / Matlab2018 網路版為輔。上課請自備筆電，裝置好以上兩種軟體。

教科書: 無 (講義多為電子書，將置於雲端硬碟供下載)

評分: 無期中期末考，需分組。作業 10 次(取其中 7 次計分), Term project 30%。

課表 (tentative)

講次	授課內容	HWK 參考講義
第一階段 基礎工程		
L1	BD 介紹	
L2	BD: SAS 平台 + Architecture + Comparison with Matlab	Exercise 1
L3	SAS 基本操作	Exercise 2A
A	1. Summarizing and presenting data. 2. Querying and Subsetting data. 3. Transforming character, numeric and date variables. Combining (complex joins and merges)	

講次	授課內容		HWK 參考講義
L3 B	SAS 程式 I 既有分析方法 統計迴歸 statistics and analytics./ Logistic 分類/分群 Classification / Clustering		
L4	SAS 軟體基本觀念	<ul style="list-style-type: none"> ◦ 多樣式資料匯入與匯出、儲存與載入 ◦ 資料物件、套件、函數與輔助說明 	Exercise 2B
	機率與統計	<ul style="list-style-type: none"> ◦ 機率分佈與抽樣 ◦ 統計檢定 ◦ 大型資料與大量資料處理 	
	資料探索 與準備	<ul style="list-style-type: none"> ◦ 資料整合、資料子集、資料排序、群組與摘形、資料清理、屬性轉換與挑選 	
	基礎建模	<ul style="list-style-type: none"> ◦ 大資料迴歸與模型選擇 ◦ 變異數分析 ANOVA ◦ 分類模型基礎 	
L4	深度學習 0: Deep Learning		Exercise 3
	非監督式學習	<ul style="list-style-type: none"> ◦ 機器學習類型介紹 ◦ 資料探勘流程及五大任務 ◦ 關聯規則探勘 ◦ 維度縮減 ◦ 集群分析 	
	監督式學習	<ul style="list-style-type: none"> ◦ 分類與迴歸樹 ◦ k 近鄰法/貝式分類/支援向量機 	
薈萃式學習	<ul style="list-style-type: none"> ◦ 績效評量與參數調校 ◦ 拔靴集成/隨機森林與多模激發法 		
L5	資料處理 Accessing, transforming and manipulating + Improving data quality for reporting and analytics. 1. Performing DO loop and array processing. 2. Restructuring or transposing SAS data sets. 3. Using macro variables: creating simple macro definitions.		Exercise 4
L6	視覺化資料探索: 連結圖像資料及分析 Exploring + visualizing		

講次	授課內容		HWK 參考講義
	資料視覺化與 SAS 繪圖基礎	<ul style="list-style-type: none"> ■ 資料視覺化簡介 ■ SAS 繪圖基礎套件 	
	資料視覺化應用	<ul style="list-style-type: none"> ■ 地圖 ■ 社會網絡 ■ 網頁應用程式 ■ 多變量條件式繪圖 ■ 圖形文法繪圖 ■ 類別資料視覺化 ■ 熱繪圖 	Exercise 5
第二階段 演算法			
L7	程式 II: Programming, SQL and Macro Language in SAS		Exercise 6
L8	程式 III : Matlab programming 深度學習 I : AI , Machine Learning		Exercise 7
L9	深度學習 II : Simulation		Exercise 8
L10	Performance monitoring and optimizations		Exercise 9
Final proposal 第三階段 應用案例			
L11	Synthesizing parallel programs Working with Hadoop, Spark and SAS		
L12	BD and AI for 金融		Exercise 10
L13	BD and AI for 貨幣需求 / 總經模型		Option1
L14	企業資料整合與商業智慧應用 工作排程管理 AmazonGo, GoogleCloud,		Option2
L15	Concluding: The future BD economics		