

日期: 2018/11/20(三) 上機考: Open Book
授課教師: 吳漢銘 (臺北大學統計學系副教授)

請仔細閱讀每一個注意事項 (禁止討論)

1. 考試期間

- (a) 可參考課本、上課講義 (包含電子檔) 及其它資料, 但不能與別人討論。
- (b) 可使用計算機、自己的筆記型電腦及平板電腦, 不可使用手機。
- (c) 全程可上網查詢, 但不能用通訊軟體 (例如: FB/LINE) 討論, 也不可抄襲網路上之程式碼。
- (d) 有問題者, 請舉手發問。勿與同學交談。
- (e) 不按照規定作答者, 酌量扣分。
- (f) 不可使用它人之隨身碟。」「作弊」或「疑似作弊」, 往後各項考試不予評分。

2. 下載題目卷, 上傳答題檔案:

- (a) 於課程網站下載題目卷。
- (b) 上傳答題檔案: 於教師網站首頁登入 [作業考試上傳區], 帳號: hdda108。密碼: xxx (上課教室號碼)。
- (c) 請上傳「學號-姓名-HDDA-MidtermExam.docx」。(目錄: 「20191030-exam1」。)

3. 答題檔案原則:

- (a) 請依照「R 程式作業繳交方式」, 複製 Console「程式執行及結果」至答案卷。圖形複製, 請注意大小, 內容數字文字需可辨識。
- (b) 程式設計題, 若程式碼直接複製 (或照抄) 講義上的以不給分為原則。
- (c) 若上傳檔案格式錯誤, 內容亂碼, 空檔等等問題。請自行負責。
- (d) 若要重覆上傳 (第 2 次以上), 請在檔名最後加「-2」、「-3」, 例如: 「學號-姓名-HDDA-exam1-2.docx」、「學號-姓名-HDDA-exam1-3.docx」等等。
- (e) 上傳兩次 (含) 以上、格式不合等等酌量扣分。
- (f) 如果上傳網站出現「You can modify the html file, but please keep the link 'www.wftpservers.com' at least.」, 請將滑鼠移至「網址列」後, 按「Enter」即可。若再不行, 請換 (IE/Edge/Firefox/Chrome)。

我已經仔細閱讀上述各注意事項, 若有違背, 會自行負責。

日期: 2019/11/20(三) 上機考: Open Book
授課教師: 吳漢銘 (臺北大學統計學系副教授)

1. 補值

- 資料來源: UCI Machine Learning Repository
- 資料集: Mammographic Mass Data Set
- 網址: <https://archive.ics.uci.edu/ml/datasets/Mammographic+Mass>
- 資料檔: mammographic_masses.data
- 說明檔: mammographic_masses.names

- (a) 如何觀察此資料遺失值之樣態? (請實作)
- (b) 哪一種補值方法較好? (至少試 3 種 (含) 以上的補值方式)

2. 資料轉換

- 資料來源: UCI Machine Learning Repository
- 資料集: Airfoil Self-Noise Data Set
- 網址: <https://archive.ics.uci.edu/ml/datasets/Airfoil+Self-Noise>
- 資料檔: airfoil_self_noise.dat

此資料之目的是建立一迴歸模型，利用 5 個解釋變數 (Frequency, Angle of attack, Chord length, Free-stream velocity, Suction side displacement thickness) 來預測一反應變數 (Scaled sound pressure level)。此資料有需要做轉換嗎? 若有需要，請至少試 3 種 (含) 以上的資料轉換方式，並比較哪一個資料轉換方式較適合迴歸模型。

3. 統計資料之面量圖

- 資料來源: 新北市統計資料庫
- 資料集: 2018 年新北市各行政區之土地面積 (平方公里) 及人口密度資料
- 網址: http://pxweb.bas.ntpc.gov.tw/pxweb/dialog/statfile9_n.asp

請於上述網址下載，新北市統計資料庫: 二、人口: 現住戶數、人口密度及性比例)「2018 新北市各行政區之土地面積 (平方公里) 及人口密度資料」。繪製新北市各行政區土地面積 (平方公里) 及人口密度資料之面量圖。

4. 地圖標記

- 資料來源: 7-11 便利商店
- 資料集: 7-11 新北市三峽區門市查詢-統一超商門市
- 網址: <http://www.i-write.idv.tw/life/info/7-11/711-304.html>
- 資料檔: SanShia7-11.csv

讀取資料檔 SanShia7-11.csv 中的地址資料，將之標記在 google 地圖上 (roadmap 型式)。

注意: 上傳檔案之後，請刪除作答目錄 及答案卷，清空資源回收筒，關機。交回題目卷。