data清理過程

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | 使用的CSV | Python檔案 | 匯出的CSV | 備註 |
|  | selected\_data0530.csv | cleandata\_1 | dropdata0531.csv |  |
|  | data0531.csv | cleandata\_2 | newdata0531.csv |  |
|  | newdata0531.csv  build.csv | cleandata\_3 | TPdata0612.csv  (全台北市最終資料) |  |
|  | TPdata0612.csv  所有房屋經緯度資料 | cleandata\_4合併經緯度 | Zhongshan\_0611.csv |  |
|  | Zhongshan\_0611.csv  郵遞區號0608.csv  空值爬蟲.csv | cleandata\_5合併6碼郵遞區號 | Zhongshan\_0611\_2.csv |  |
|  | Zhongshan\_0611\_2.csv | 計算rawdata與各項設施距離 | 中山區完整資料加距離.csv | 另一個資料夾 |
|  | 中山區完整資料加距離.csv | cleandata\_6路名切片 | 中山區完整資料\_0611.csv |  |
|  | 中山區完整資料\_0611.csv | cleandata\_7 | 中山區完整資料\_0712.csv |  |
|  | 中山區完整資料\_0712.csv | 資料離散化 | 中山區完整資料\_0712\_2.csv | Orange檔 |
|  | 中山區完整資料\_0712\_2.csv | cleandata\_8移除離群值 | 原資料log.csv  ZS\_final.csv |  |