資料來源與說明-Housing Affordability Data

410578043 李桓宇 410578068 陳威傑

1. 資料來源與使用目的

Housing Affordability Data (住宅負擔能力資料)

<https://catalog.data.gov/dataset/housing-affordability-data-system-hads>

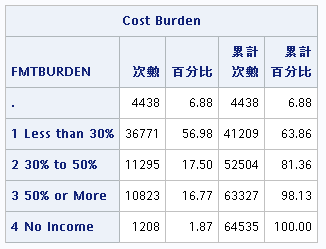
可於美國政府的公開資料網站存取，現在能取得1985到2013的資料，我們選用2013年的資料作為分析對象。其研究目的為提供房市研究者關於房租貸負擔能力(佔收入多少%)與房屋物件資訊的各項長期資料。

1. 變數說明

2013年的資料有99個變數，不過有將近一半為連續變數經轉換為類別變數，所以實際的變數大約只有一半；觀測值有64536筆。以下有其中我們較感興趣的變數進行說明。

|  |  |  |
| --- | --- | --- |
| 變數名稱 | 變數說明 | 變數種類 |
| METRO | 都市圈的中心城市/郊區地位 | 連續變數 |
| Built | 建築物建造年分 | 連續變數 |
| # of units in building | 建物內戶數 | 連續變數 |
| Structure type | 結構類型(獨棟或大樓) | 類別變數 |
| Current market value of unit | 當前市值 | 連續變數 |
| Age1 | 房主年齡 | 連續變數 |
| Tenure | 房主或房客 | 類別變數 |
| # of rooms In Unit | 室數 | 連續變數 |
| Region | 人口普查地區 | 類別變數 |
| Total Wage Income | 薪資收入(年) | 連續變數 |
| Area median income (average) | 區域平均收入 | 連續變數 |
| Poverty income | 貧窮線收入 | 連續變數 |
| Monthly Utility Cost | 每月水電費 | 連續變數 |
| Other Cost | 保險，公寓，土地租金，其他移動房屋費用 | 連續變數 |
| FMR | 公平市場租金 | 連續變數 |
| Cost Burden | 成本負擔 | 類別變數 |
| Occupancy Status | 佔用狀態 | 類別變數 |

貧窮線收入 指該地區收入低於多少會被視為貧窮人口  
成本負擔 指房貸或房租佔使用者收入的多少百分比並分為四群，如下圖



1. 研究問題
2. *能否以屋齡、地段、該區平均收入等變數預測公平市場租金*

雖然此資料並非今年的即時資料，但是了解租金與其他變數的相關對了解房屋市場仍有一定幫助，在一定程度上避開定價不實的物件或是掌握議價空間

1. *觀察成本負擔受到哪些變數影響*

水電費、屋主年齡、租屋或買屋、使用者的收入是否低於平均線等變數是否會提升或降低成本負擔