

プログラミング3

第7回

RESASデータ レポート課題説明

レポート課題その1のURL

2

<https://classroom.github.com/a/gbZ0vciT>

- 研磨目的：
不動産種別の取引金額を市区町村単位で把握するためのデータセットを作成
- 入力データ
 - － 不動産取引価格情報
[データ取得元]<http://www.land.mlit.go.jp/webland/download.html>

1. 作成物 : 集計データ

- ー 面積あたり平均価格 (円/㎡) を求める

2. 対象年 : 2009年～2016年 (8時点)

3. 集計単位 : 市区町村

4. 除外条件 :

| | |
|-----------------|------------|
| 宅地(土地) 工業地 | 除外 |
| 宅地(土地) 住宅地 | 2000㎡以上は除外 |
| 宅地(土地) 商業地 | 2000㎡以上は除外 |
| 宅地(土地) 宅地見込地 | 除外 |
| 宅地(土地と建物) 工業地 | 除外 |
| 宅地(土地と建物) 住宅地 | 除外 |
| 宅地(土地と建物) 商業地 | 除外 |
| 宅地(土地と建物) 宅地見込地 | 除外 |
| 中古マンション等 | 2000㎡以上は除外 |
| 農地 | 5000㎡以上は除外 |
| 林地 | 5000㎡以上は除外 |

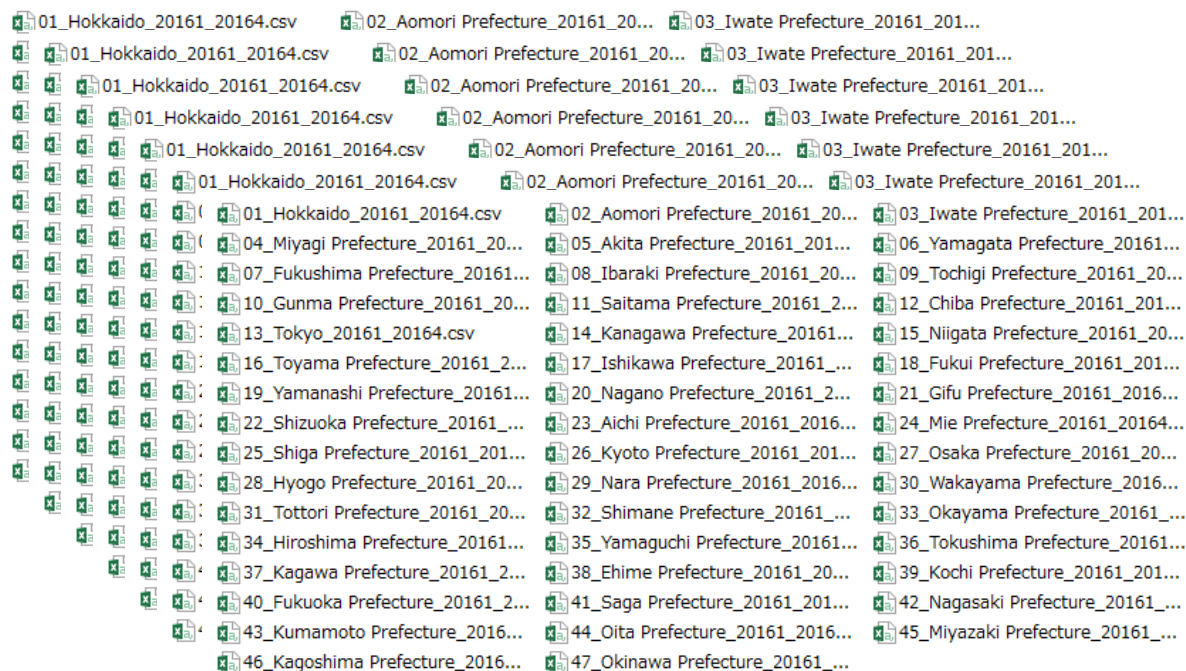
5. 出力レイアウト

| 集計年 | 都道府県コード | 都道府県名 | 市区町村コード | 市区町村名 | 表示種類コード | 表示種類名 | 面積あたり平均価格 (円/㎡) |
|------|---------|-------|---------|-------|---------|----------|-----------------|
| 2016 | 01 | 北海道 | 01202 | 函館市 | 1 | 土地 (住宅地) | 31355.3141 |
| 2016 | 01 | 北海道 | 01202 | 函館市 | 2 | 土地 (商業地) | 35824.0782 |
| 2016 | 01 | 北海道 | 01202 | 函館市 | 3 | 中古マンション等 | 140259.542 |
| 2016 | 01 | 北海道 | 01202 | 函館市 | 4 | 農地 | 799.1137 |
| 2016 | 01 | 北海道 | 01202 | 函館市 | 5 | 林地 | 2466.5032 |
| 2016 | 01 | 北海道 | 01203 | 小樽市 | 1 | 土地 (住宅地) | 17131.7829 |
| 2016 | 01 | 北海道 | 01203 | 小樽市 | 2 | 土地 (商業地) | 86153.8462 |
| 2016 | 01 | 北海道 | 01203 | 小樽市 | 3 | 中古マンション等 | 152703.2967 |
| 2016 | 01 | 北海道 | 01203 | 小樽市 | 4 | 農地 | 1484.4291 |
| 2016 | 01 | 北海道 | 01203 | 小樽市 | 5 | 林地 | 118.6465 |

※なお、データ詳細は正解データを参照のこと

※この項目にないファイルの仕様は、slackにて確認すること

読み込むファイルが多い



集計年ごとにデータが格納されている
フォルダが分かれており、さらに各フォル
ダには47都道府県別のcsvファイルが
存在するので手動でインポートするの
は困難！

47ファイル
×
8年分



376ファイル

何をもって完全一致とするか

7

- **プログラムやツールでファイル全体を比較**
 - Pythonのパッケージの場合、NaNの取り扱いに注意
- **Excelで比較したいファイルを並べて
TRUE/FALSEの判定**
 - 全体をチェックできるが行／列数が多いと面倒
- **数レコードを抜き出して完全一致かチェック**
 - 簡易チェック、第三者チェックとしては有用

1. 提出物 : ipynbファイル、出力ファイル

- ipynb : RESAStask2_学籍番号.ipynb
- csv : RESAStask2_学籍番号_名字.csv
 - 全て英数で記入

2. 提出条件 : 正解データとの完全一致

3. 提出期限 : 12月4日 23:59