

# DOWNTOWN MELBOURNE ?

**Should City of Melbourne be divided into specialised industrial areas?**

**Keywords:** Industry, Business, Employment, CBD, Melbourne

# THE ISSUE

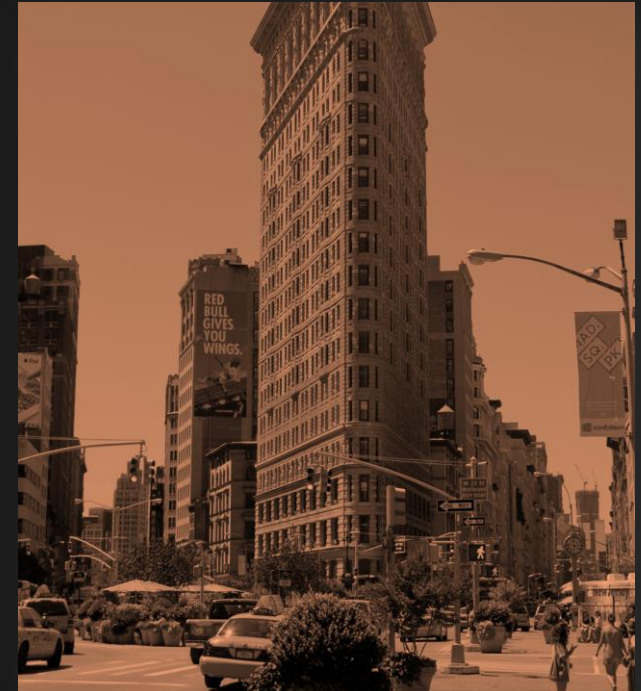
Example: Manhattan (NYC)



**Finance & Banking**



**Advertising**



**Technology**

# THE BENEFITS

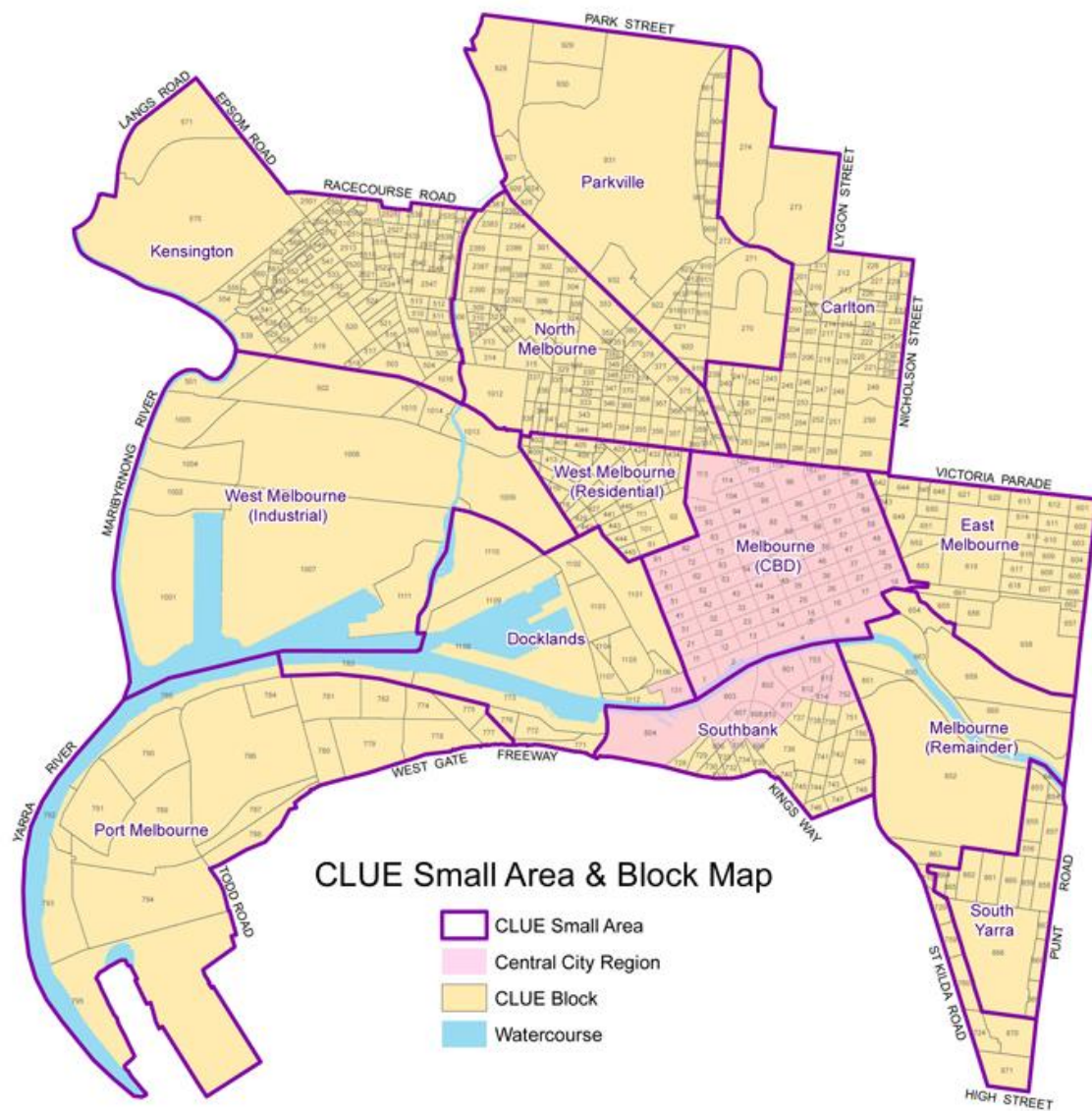
- Simplify funding & investment
- Increase ROI & business performance
- Increase job opportunities
- Reduce unhealthy competition
- Attract overseas investment
- Increased recognition & symbolism

# RESEARCH QUESTION

**Does City of Melbourne have the capacity  
to form industrial specialisation areas?**

# THE DATA SOURCE

- **Source:** <https://data.melbourne.vic.gov.au/>
  - City of Melbourne Council open data
  - Year of interest: 2017
- **Census of Land Use and Employment (CLUE)**
  - City of Melbourne socio-economic data
  - 13 small areas / 606 city blocks



**Source:**

<https://www.melbourne.vic.gov.au/about-melbourne/research-and-statistics/city-economy/census-land-use-employment/Pages/clue-small-area-and-block-maps.aspx>

# DATASET 1

- Employment\_by\_block\_by\_CLUE\_industry.csv

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
1	Census year	Block ID	CLUE small	Accommodation	Admin and support	Agriculture	Arts and recreation	Business services	Construction	Education	Electricity, gas, water and waste	Finance and insurance	Food and accommodation	Health care	Information and communication	Manufacturing	Other services	Public administration	Real estate	Rental and leasing	Retail trade	Transport, postal and warehousing	Wholesale trade	Total employment in block		
2	2017	1 Melbourne	0	94	0				0	0			66	0	0	0	3	0					0	815		
3	2017	2 Melbourne	0	0	0			0	0	0			0	0	0	0	0	0	0	0	0	0	0			
4	2017	4 Melbourne	0	0	0		39	0	0		0	0	205		0	0	16		0	0	46	344	0	676		
5	2017	5 Melbourne	0	0	0		0	0	0	0	0	0		0		0		0	0	0	10	0	0	15		
6	2017	6 Melbourne	0	40	0		341	0	0	0	0	0	302	0		0					48	30	0	1045		
7	2017	11 Melbourne	163	98	0			159				22	119				67	0	12				0	794		
8	2017	12 Melbourne	63	723	0		43	1735		270		267	154				17			0	15			4550		
9	2017	13 Melbourne	0	507	26			426		29		260	31		131	0	22	0	15	0	48			2181		
10	2017	14 Melbourne	155	157				539	73	517	0	100	384	79	29		177		51	0	79		27	2426		
11	2017	15 Melbourne	60	131			154	219		485	0	71	766	41	65	47	188			0	411	0	23	2679		
12	2017	16 Melbourne (CBD)			0		158	360	0	228		74	144	49		0	20	0	16	0	68	0		1238		

- From 2002 – 2017 → Needs to be subset
- Missing data → Needs to be imputed
- Unnecessary data (Sum) → Needs to be removed



# DATASET 1: DIMENSION

```
Dimensions of dataset: employment_df
Census year                int64
Block ID                   int64
CLUE small area            object
Accommodation              float64
Admin and Support Services float64
Agriculture and Mining     float64
Arts and Recreation Services float64
Business Services          float64
Construction               float64
Education and Training     float64
Electricity, Gas, Water and Waste Services float64
Finance and Insurance      float64
Food and Beverage Services float64
Health Care and Social Assistance float64
Information Media and Telecommunications float64
Manufacturing              float64
Other Services             float64
Public Administration and Safety float64
Real Estate Services       float64
Rental and Hiring Services float64
Retail Trade               float64
Transport, Postal and Storage float64
Wholesale Trade            float64
Total employment in block float64
dtype: object
```

```
Missing information in employment17_df
Census year                0
Block ID                   0
CLUE small area            0
Accommodation              129
Admin and Support Services 99
Agriculture and Mining     25
Arts and Recreation Services 207
Business Services          90
Construction               115
Education and Training     121
Electricity, Gas, Water and Waste Services 124
Finance and Insurance      66
Food and Beverage Services 135
Health Care and Social Assistance 137
Information Media and Telecommunications 115
Manufacturing              114
Other Services             154
Public Administration and Safety 90
Real Estate Services       97
Rental and Hiring Services 40
Retail Trade               120
Transport, Postal and Storage 107
Wholesale Trade            115
Total employment in block 125
dtype: int64
```



# DATASET 1: DATA CLEANING

- Subset year = 2017
- Impute empty total column with 0
- Fill missing information in each block (row) by averaging the missing sum from the total sum
- Classify each block as “industrial” (total is not 0) or “residential” (total is 0)
- Remove total

```
After cleaning missing values in employment17_df
Census year      0
Block ID         0
CLUE small area  0
Accommodation    0
Admin and Support Services  0
Agriculture and Mining  0
Arts and Recreation Services  0
Business Services  0
Construction     0
Education and Training  0
Electricity, Gas, Water and Waste Services  0
Finance and Insurance  0
Food and Beverage Services  0
Health Care and Social Assistance  0
Information Media and Telecommunications  0
Manufacturing    0
Other Services   0
Public Administration and Safety  0
Real Estate Services  0
Rental and Hiring Services  0
Retail Trade      0
Transport, Postal and Storage  0
Wholesale Trade   0
Block type        0
dtype: int64
```

Out[3]:

	Census year	Block ID	CLUE small area	Accommodation	Admin and Support Services	Agriculture and Mining	Arts and Recreation Services	Business Services	Construction	Education and Training	...	Information Media and Telecommunications	Manufacturing	S
0	2017	1	Melbourne (CBD)	0.0	94.0	0.0	81.5	81.5	0.0	0.0	...	0.0	0.0	
1	2017	2	Melbourne (CBD)	0.0	0.0	0.0	0.0	0.0	0.0	0.0	...	0.0	0.0	

## MELBOURNE

## INDUSTRIAL SPECIALISATION

- Count\_of\_business\_establishments\_per\_industry\_for\_blocks\_2017.csv

[illegible]

# DATASET 3

- Population\_Forecasts\_by\_Small\_Area.csv

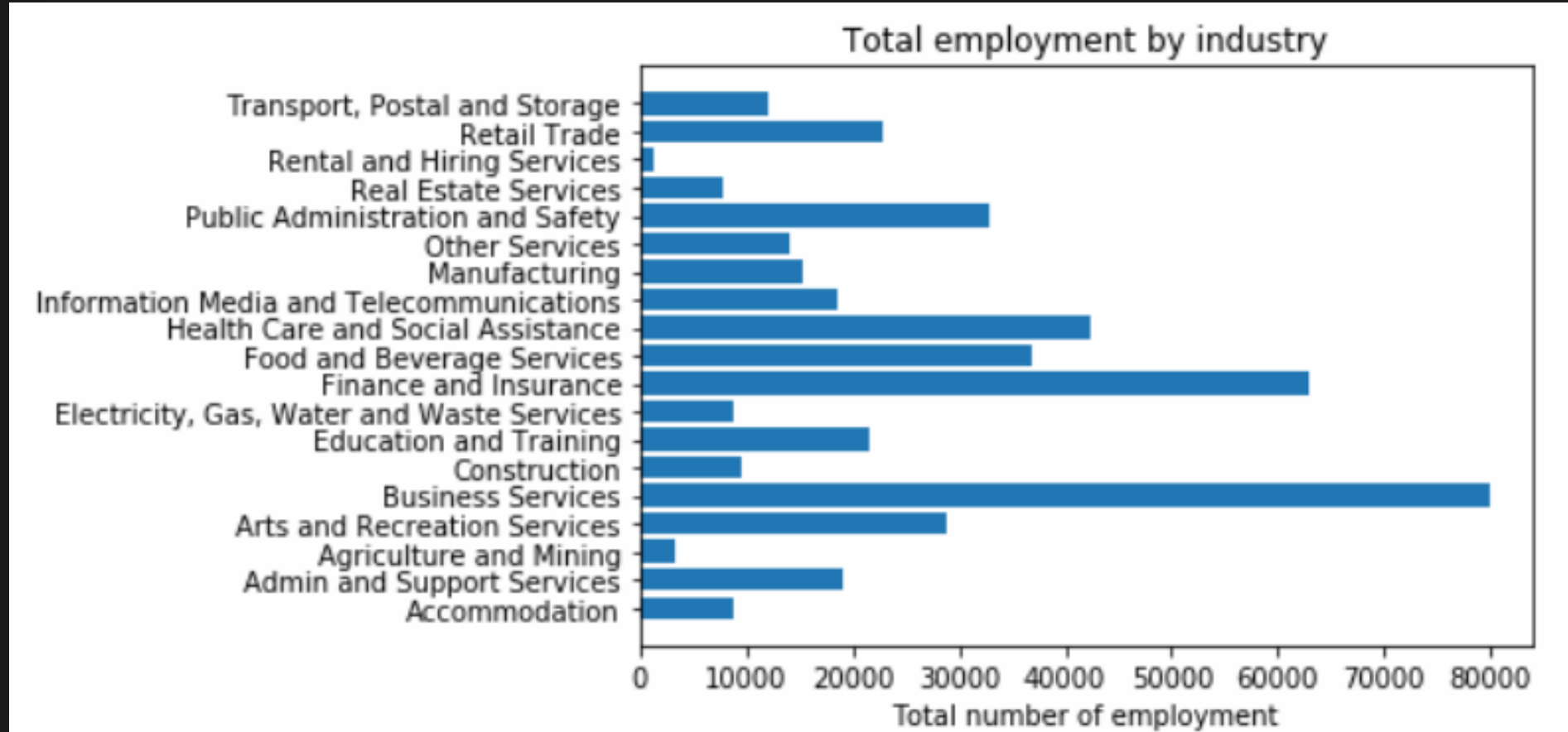
```
Dimensions of dataset: population_df
Geography      object
Year           int64
Population     int64
dtype: object
```

	A	B	C
1	Geography	Year	Population
2	Melbourne (CBD)	2017	42589
3	Melbourne (CBD)	2018	44087
4	Melbourne (CBD)	2019	46442
5	Melbourne (CBD)	2020	51136
6	Melbourne (CBD)	2021	53795
7	Melbourne (CBD)	2022	55854
8	Melbourne (CBD)	2023	58336
9	Melbourne (CBD)	2024	61086
10	Melbourne (CBD)	2025	63802
11	Melbourne (CBD)	2026	65529
12	Melbourne (CBD)	2027	67302
13	Melbourne (CBD)	2028	68691
14	Melbourne (CBD)	2029	69607
15	Melbourne (CBD)	2030	70609
16	Melbourne (CBD)	2031	71573
17	Melbourne (CBD)	2032	72479
18	Melbourne (CBD)	2033	73438
19	Melbourne (CBD)	2034	74334
20	Melbourne (CBD)	2035	75242
21	Melbourne (CBD)	2036	76142
22	Melbourne (CBD)	2037	76982
23	Carlton	2017	20776
24	Carlton	2018	21214
25	Carlton	2019	21961

MELBOURNE

INDUSTRIAL SPECIALISATION

# BASIC VISUALISATION



# BASIC VISUALISATION

## PROCESS

- Employment DF
- Choose Block type = Industrial
- $x$  = Number of Industrial blocks
- $y$  = Industry categories

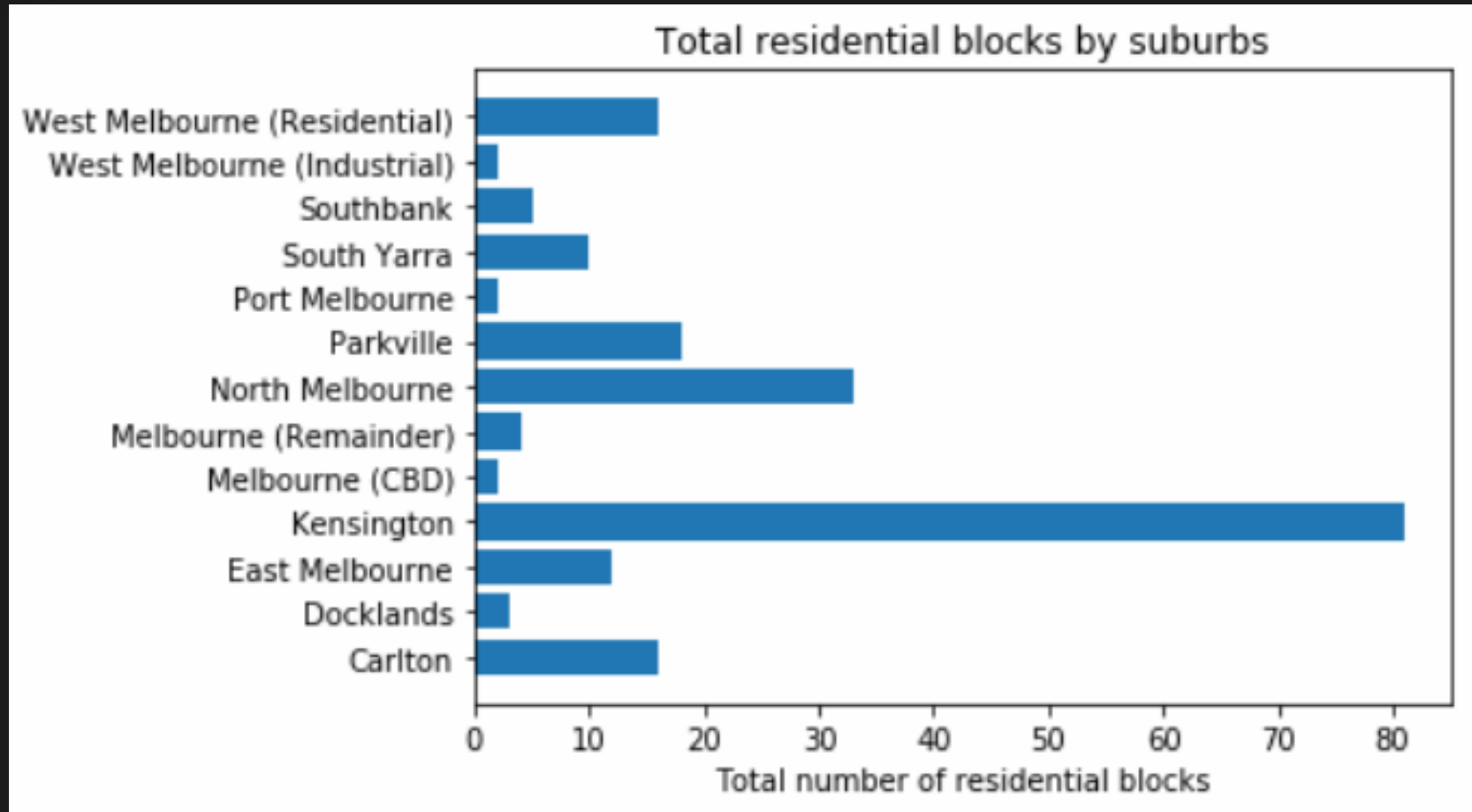
## INFO REVEALED

- Diversity of industries
- Not uniformly distributed
  - Business Services
  - Finance & Insurance

## CHALLENGES

- Doesn't show location distribution
- Missing data
- Ambiguous classification

# BASIC VISUALISATION



# BASIC VISUALISATION

## PROCESS

- Employment DF
- Choose Block type = Residential
- $x$  = Number of Residential blocks
- $y$  = Small suburb categories

## INFO REVEALED

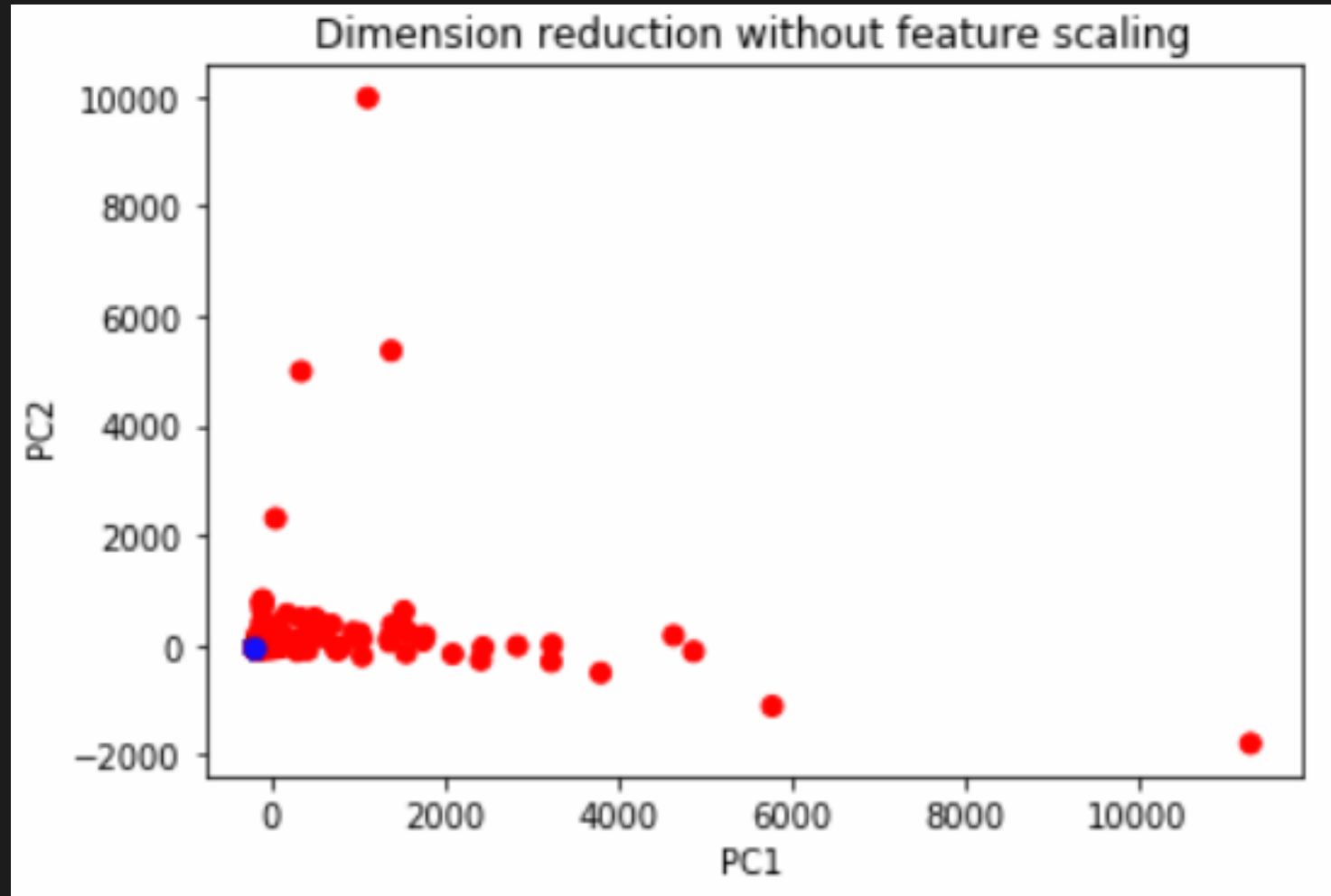
- Low residential areas (highly industrial)
  - West Melbourne (I)
  - Port Melbourne
  - Melbourne (CBD)
  - Docklands

## CHALLENGES

- Number of blocks per suburbs is not the same
- Missing data & zero imputation
- Doesn't seem to correlate with Population data



# OUTLIER ANALYSIS



# OUTLIER ANALYSIS

	Block ID	CLUE small area	Industry with highest employment	Maximum employment
16	24	Melbourne (CBD)	Finance and Insurance	2139.0
19	27	Melbourne (CBD)	Business Services	2116.0
23	32	Melbourne (CBD)	Business Services	2610.0
25	34	Melbourne (CBD)	Finance and Insurance	2995.0
28	37	Melbourne (CBD)	Finance and Insurance	1841.0
39	52	Melbourne (CBD)	Business Services	5548.0
40	53	Melbourne (CBD)	Finance and Insurance	3871.0
65	85	Melbourne (CBD)	Finance and Insurance	2278.0
536	1103	Docklands	Finance and Insurance	6604.0
538	1105	Docklands	Finance and Insurance	3650.0
541	1108	Docklands	Finance and Insurance	12142.0
386	651	East Melbourne	Health Care and Social Assistance	2424.0
489	870	Melbourne (Remainder)	Health Care and Social Assistance	5247.0
510	920	Parkville	Health Care and Social Assistance	10268.0
521	931	Parkville	Health Care and Social Assistance	5158.0

MELBOURNE

INDUSTRIAL SPECIALISATION

# OUTLIER ANALYSIS

## PROCESS

- Employment DF
- No scaling (Preserve the size of the blocks)
- 2-PCA
- Color-coded by block types
- Eyeballing outliers (cut-offs > 2000)

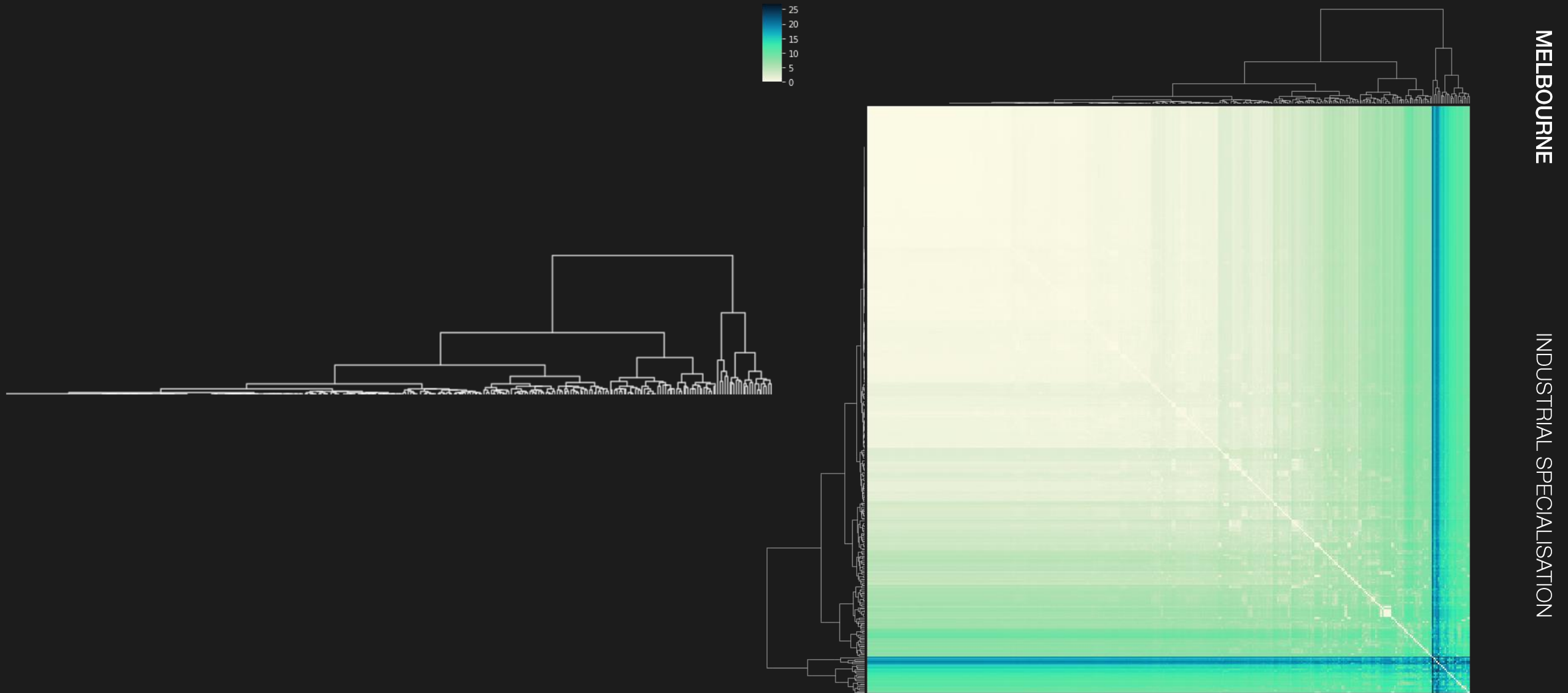
## INFO REVEALED

- 15 blocks with exceptionally high employment
  - Usually in one industry
  - 8 x Melbourne CBD
- Docklands (Finance)
- Parkville (Health Care)

## CHALLENGES

- Dimension reduction threshold
- Without scaling
- No clear separation between classes
- Informal benchmark for outliers
- Removability of Melbourne CBD as an outlier

# HIERARCHICAL CLUSTERING



# HIERARCHICAL CLUSTERING

## PROCESS

- Cleaned Employment DF
- Block type = Industrial
- Scaled data
- Complete Linkage Clustering (since an area is the agglomeration of blocks)

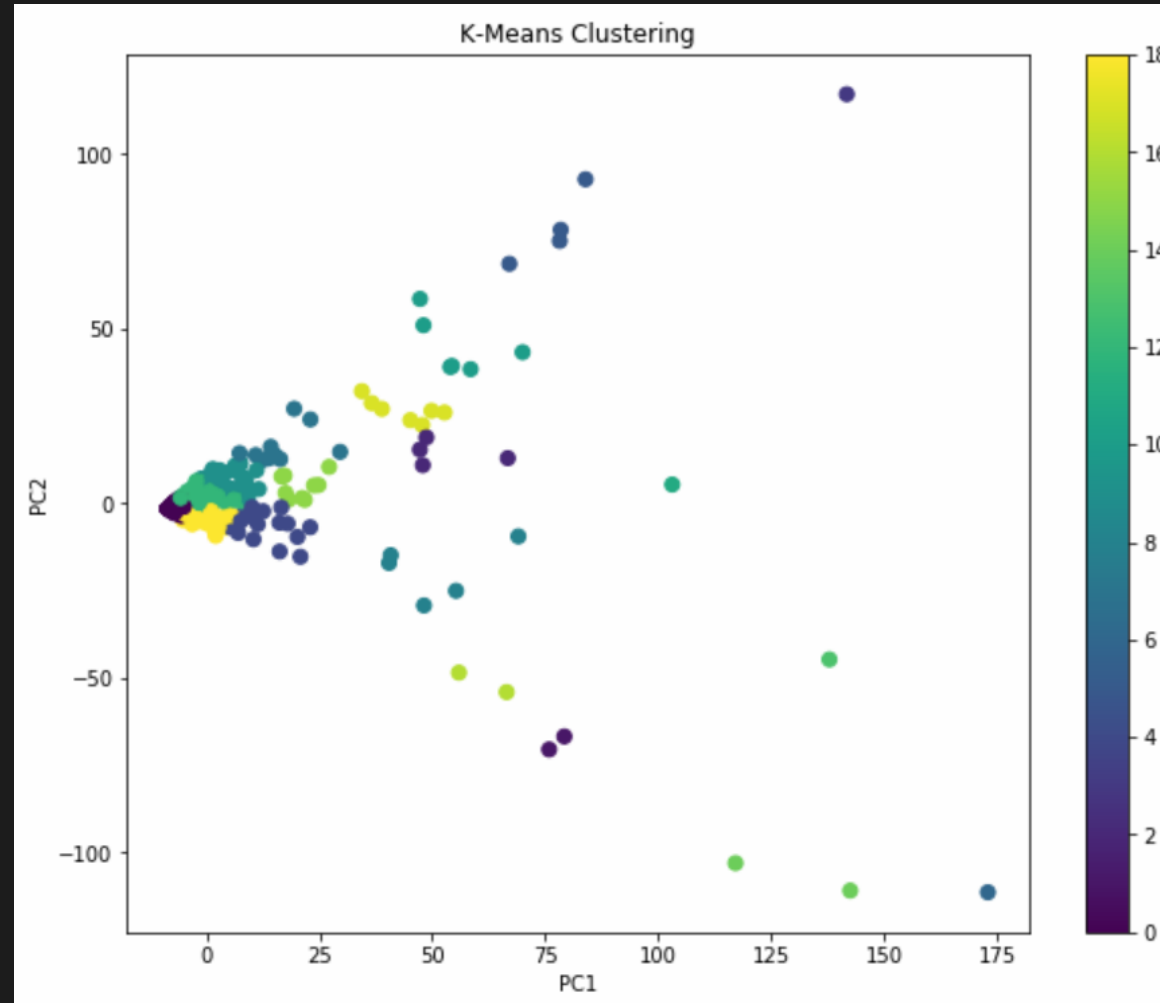
## INFO REVEALED

- Expected 13 clusters (according to 13 small areas if there's any disposition for specialization)
- Inconclusive

## CHALLENGES

- Removed critical information from outliers
- Not reflecting geographical clusters
- Overlapping in real-life

# K-MEANS CLUSTERING



MELBOURNE

INDUSTRIAL SPECIALISATION

# K-MEANS CLUSTERING

## PROCESS

- Business DF
- 2-PCA scaled data
- 19-means clustering

## INFO REVEALED

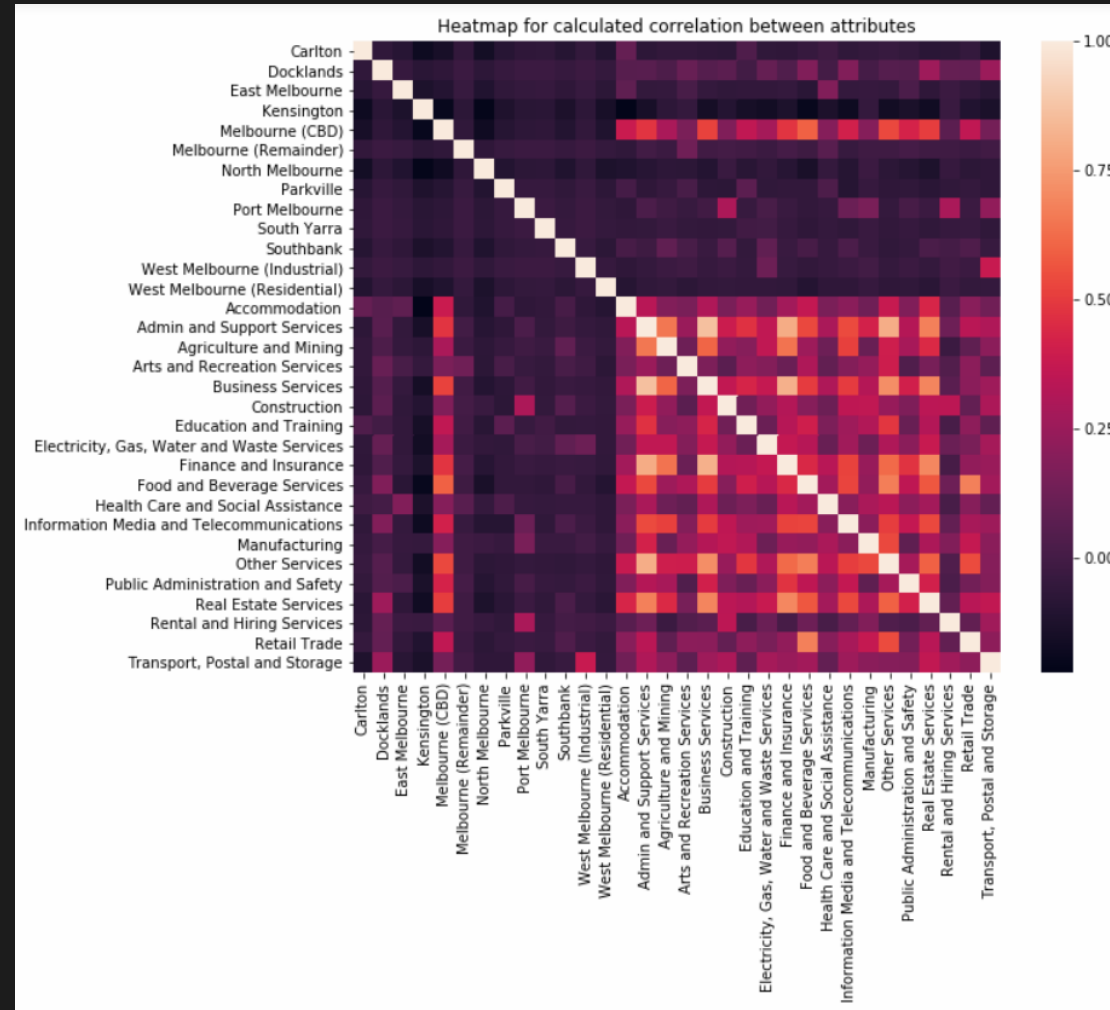
- Expected 19 distinct clusters (according to 19 industries) of different blocks
- Inconclusive
- Consistent shape with employment DF

## CHALLENGES

- Have not performed outlier analysis
- Dimension reduction limitations
- Different operationalisation for clusters (different k)



# CORRELATIONAL ANALYSIS



MELBOURNE

INDUSTRIAL SPECIALISATION

# CORRELATIONAL ANALYSIS

## PROCESS

- Business DF
- Binarise areas into 13 columns
- Plot heatmap for pair-wise attributes

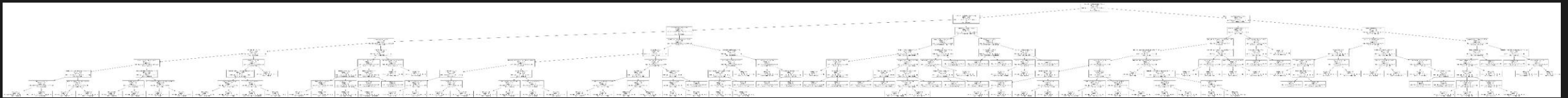
## INFO REVEALED

- Only CBD has a high correlation with all building types
  - Docklands is next
- Admin, Business & Finance services are grouped together

## CHALLENGES

- Not taking population size into account

# DECISION TREE CLASSIFIER TEST



Train accuracy: 0.6049586776859505  
Test accuracy: 0.5384615384615384

# DECISION TREE CLASSIFIER TEST

## PROCESS

- Class = Suburb
- Data = Employment
- Test set = Average data
- Train set = Current data
- Tree depth = 8

## INFO REVEALED

- If test is accurate ( $> .5$  chance) then there might be a chance for a relationship between employment & suburbs
- Employment ratio within a block is a good indicator of suburbs
- No direct relationship with specialisation

## CHALLENGES

- Limited test/train sets
- Informal choice of tree depths (higher depths decrease accuracy)
- Noisy data (residential)

# DATA LINKAGE WITH POPULATION

	CLUE small area	No of residential blocks	Population	Population per residential block
0	Melbourne (CBD)	2	42589	21294.500000
1	Southbank	5	21065	4213.000000
2	Docklands	3	12458	4152.666667
3	Carlton	16	20776	1298.500000
4	North Melbourne	33	16985	514.696970
5	Melbourne (Remainder)	4	1989	497.250000
6	South Yarra	10	4654	465.400000
7	East Melbourne	12	5490	457.500000
8	Parkville	18	7929	440.500000
9	West Melbourne (Residential)	16	5564	347.750000
10	Kensington	81	11657	143.913580
11	Port Melbourne	Pearson correlation between number of residential blocks and population size 0.014729423024402747		
12	West Melbourne (Industrial)			

Pearson correlation without outliers  
0.1267451360896581

# DATA LINKAGE WITH POPULATION

## PROCESS

- Merged on suburb data
- Calculate coefficient between number of residential blocks & population per suburb
- Eyeball outliers
- Calculate population per residential block per suburb

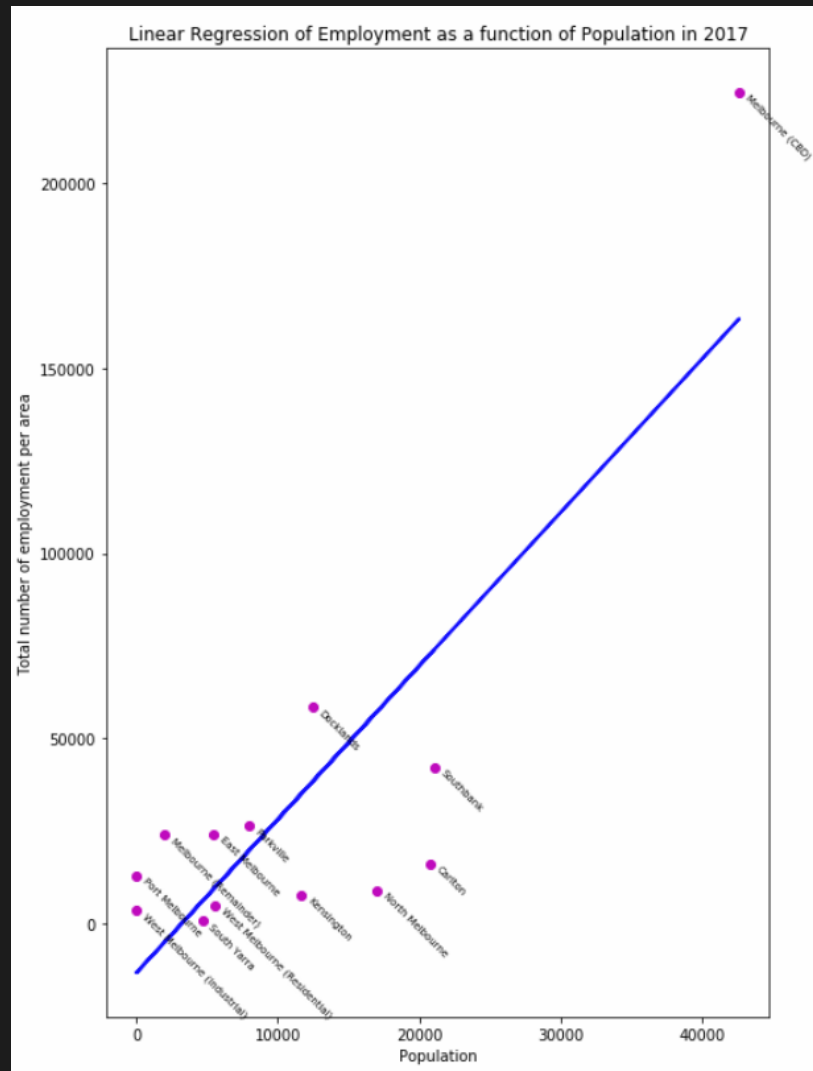
## INFO REVEALED

- Low correlation
  - Employment is not a good indicator of Block Type
- Population outliers
  - Melbourne CBD
  - Industrial area (Port Melbourne & West Melbourne)
- 10x impact from population outliers

## CHALLENGES

- Removability of CBD as outlier
- Little relationship with industrial specialisation
- Reveal drawbacks of earlier analyses and classifications

# LINEAR REGRESSION PREDICTION



	CLUE small area	Total employment	Predicted employment 2018	Predicted employment 2019	Predicted employment 2020
0	Carlton	16073.0	169588.999725	179353.977025	198817.587830
1	Docklands	58489.0	74746.398544	77843.824465	80825.148745
2	East Melbourne	24183.0	41155.705930	46002.949439	52682.940280
3	Kensington	7894.0	9795.823417	9994.854801	10633.413826
4	Melbourne (CBD)	224395.0	35980.889934	37336.791241	38580.737394
5	Melbourne (Remainder)	24023.0	-4534.436264	-3684.406393	-2469.485651
6	North Melbourne	8897.0	58558.512608	64575.065501	70235.020497
7	Parkville	26394.0	19880.080229	20390.098152	21401.841023
8	Port Melbourne	12992.0	-13175.715540	-13175.715540	-13175.715540
9	South Yarra	777.0	77566.009824	80364.888668	83640.613537
10	Southbank	42351.0	6267.162830	6292.041753	6350.092573
11	West Melbourne (Industrial)	3552.0	-13175.715540	-13175.715540	-13175.715540
12	West Melbourne (Residential)	5006.0	10413.650006	10558.777057	10579.509493



# LINEAR REGRESSION PREDICTION

## PROCESS

- Merged Employment & Population (2017 – 2020)
- Train set = Test set (2017)
- $x$  = Population
- $y$  = Total employment
- Forecast 2018 – 2020

## INFO REVEALED

- Correlation between Employment & Population in 2017
- Melbourne CBD is an outlier
- Unreliable predictor of future employment
  - Specialisation is unlikely as employment changes dynamically

## CHALLENGES

- Oversimplification & overfitting of data
- Lacking test data
- Not predicting industry (unless assuming fixed ratio for each block)
- No correlation with industry specialisation

# CONCLUSION

- There is little evidence for a predisposition for industrial specialisation in Melbourne CBD.
- Most employments and industries are highly concentrated in the CBD area, turning the most important area of the dataset to an outlier which challenges the interpretation.
- Reveal a potential for Docklands to become the next Central Business District of Melbourne.
- Reveal a strong focus on Business & Professional Services for City of Melbourne.

# ASK ME ANYTHING

**THANK YOU FOR  
LISTENING**