

CAMBRIDGE MAP – EXPLORING DATA

*Submission by Rafal Jankowski as part of UDACITY Data Analyst Program
Attached is also the Python code as well as sample map area.*

The selected data is Cambridge area. File of size 63MB was selected using Map Zen.

This is the summary of the data structure:

```
{'node': 254684, 'nd': 345845, 'bounds': 1, 'member': 4908, 'tag': 239279, 'relation': 567, 'way': 49093, 'osm': 1}
```

Identification of street names:

There are a total of 1,254 streets in Cambridge. Auditing street names is more cumbersome as in the exercises because there are 127 different ways of naming these street types. More obscure ones include Paddocks, Homes, Hall, Hill, Lawn, Lodge, etc. However, we can look at the more common ones (each having at least 10 instances of a street name) as they account for 1,068 of all street names. And luckily, Cambridge being a proper British city, the street naming is consistent (Court appears in all instances instead of Ct or Ct., Street instead of St. etc).

```
{'Avenue': 40,  
'Close': 175,  
'Court': 133,  
'Drive': 21,  
'Gardens': 14,  
'House': 16,  
'Lane': 36,  
'Mews': 11,  
'Park': 15,  
'Place': 49,  
'Road': 266,  
'Street': 140,  
'Terrace': 37,  
'Walk': 26,  
'Way': 89}
```

List of all amenities in the city:

```
{'taxi': 2, 'dojo': 1, 'social_facility': 35, 'clinic': 3, 'nursery': 3, 'post_office': 14,  
'bureau_de_change': 1, 'toilets': 34, 'bicycle_repair_station': 1, 'police': 1, 'townhall': 1,  
'community_centre': 11, 'pharmacy': 19, 'kindergarten': 9, 'social_club': 1, 'fuel': 11, 'hand_pump':  
1, 'building_society': 2, 'arts_centre': 3, 'post_depot': 1, 'hall': 2, 'bank': 29, 'place_of_worship': 96,  
'school': 131, 'bar': 5, 'punt_hire': 1, 'parking_entrance': 1, 'university': 1140, 'disused': 1,  
'grave_yard': 12, 'car_sharing': 3, 'mixed-use centre': 1, 'library': 34, 'college': 15, 'parking': 764,  
'optometrist': 1, 'grit_bin': 4, 'clock': 7, 'veterinary': 5, 'fountain': 12, 'fast_food': 92, 'post_box':  
180, 'pub': 84, 'studio': 2, 'bicycle_rental': 3, 'courthouse': 2, 'motorcycle_parking': 6, 'restaurant':  
135, 'atm': 25, 'nightclub': 5, 'bus_station': 6, 'recycling': 21, 'nursing_home': 2, 'hospice': 1,  
'bench': 348, 'market': 1, 'boat_rental': 1, 'service': 1, 'hospital': 7, 'dead_pub': 1, 'car_rental': 3,  
'bicycle_parking': 762, 'bbq': 2, 'drinking_water': 2, 'wheelchair_lift': 10, 'counselling': 1,  
'smoking_shelter': 1, 'club': 1, 'gym': 2, 'dentist': 21, 'doctors': 19, 'community_room': 1,  
'parking_space': 6, 'register_office': 1, 'sculpture': 37, 'public_building': 7, 'cinema': 3, 'car_wash':  
1, 'telephone': 22, 'cosmetic_surgery': 1, 'childcare': 1, 'health_centre': 2, 'language_school': 1,  
'adult_education_centre': 1, 'cafe': 99, 'chiropodist': 1, 'fire_station': 2, 'theatre': 4, 'advice': 1,  
'emergency_phone': 13, 'statue': 3, 'waste_basket': 9, 'scout_hut': 3, 'swimming_pool': 8,
```

'healthcare': 1}

Cambridge is a home to a major university in the UK and unsurprisingly there are 1,140 addresses tagged as university. Obviously that sounds odd, there can't be 1,140 distinct university facilities in one city. Printing the details out indicates so – there are 857 university locations which do not have a street address. Most of the ones with missing street address have only the following info: [*'ref', 'amenity', 'building', 'operator'*] – and those *'building'* categories just have a word *Yes* attached to it as value. Probably not very useful to us.

And the full list below shows the university concentration. Looks like most buildings are on West Road, Lensfield Road and Grange Road. In fact quick google search confirmed this result:

{": 857, 'West Road': 15, 'Tennis Court Road': 13, 'Lensfield Road': 19, 'Herschel Road': 4, 'Long Road': 1, 'Queen's Road': 1, 'Mount Pleasant': 3, 'Owlstone Road': 3, 'Union Road': 1, 'Gresham Road': 1, 'Bene't Place': 1, 'Trinity Lane': 2, 'Chesterton Road': 6, 'Madingley Rise': 3, 'Sidgwick Avenue': 4, 'Glisson Road': 1, 'Madingley Road': 1, 'Downing Place': 1, 'Trumpington Street': 22, 'Fitzwilliam Street': 12, 'Parsons Court': 1, 'Hills Road': 4, 'St Andrew's Street': 2, 'Wilberforce Road': 2, 'King's Parade': 2, 'Lady Margaret Road': 1, 'Storey's Way': 14, 'South Green Road': 1, 'Mill Lane': 5, 'Free School Lane': 2, 'J J Thomson Avenue': 5, 'Edwinstowe Close': 1, 'Barton Road': 3, 'Clarkson Road': 1, 'Grange Road': 31, 'Trinity Street': 2, 'Mortimer Road': 1, 'East Road': 1, 'Chaucer Road': 3, 'Sylvester Road': 5, 'Thompson's Lane': 1, 'Castle Street': 4, 'Huntingdon Road': 7, 'Trumpington Road': 2, 'Sidney Street': 1, 'Magdalene Street': 1, 'Wordsworth Grove': 3, 'Charles Babbage Road': 4, 'Oxford Road': 1, 'Pembroke Street': 9, 'Chesterton Lane': 4, 'Regent Street': 1, 'Philippa Fawcett Drive': 1, 'Station Road': 1, 'St John's Street': 1, 'Downing Street': 15, 'Adams Road': 5, 'Silver Street': 7, 'Saint Andrew's Street': 1, 'Newnham Walk': 2, 'Jesus Lane': 1, 'Cranmer Road': 8, 'Selwyn Gardens': 2}

Next, I wanted to explore more the nightlife scene in Cambridge to find out where would be good location to stay on a short weekend visit.

Starting with Fast Food places first – there are 92 of them. Cherry Hinton Road and Mill Road contain the highest numbers:

(<type 'int', {": 14, 'Lensfield Road': 2, 'Akeman Street': 1, 'Norfolk Street': 2, 'Cherry Hinton Road': 10, 'Mill Road': 12, 'The Broadway': 2, 'Sturton Street': 1, 'Green End Road': 1, 'Burleigh Street': 1, 'Hills Road': 6, 'Chesterton Road': 1, 'Bridge Street': 1, 'Emmanuel Street': 1, 'Campkin Road': 1, 'East Road': 1, 'Market Hill': 2, 'Botolph Lane': 1, 'King Street': 1, 'Victoria Road': 1, 'Rose Crescent': 2, 'Milton Road': 6, 'Regent Street': 5, 'Wulfstan Way': 1, 'Catharine Street': 1, 'Beehive Centre': 1, 'Barnwell Road': 2, 'Newmarket Road': 5, 'Petty Cury': 1, 'Histon Road': 2, 'Cambridge Leisure Park': 1, 'Victoria Avenue': 1, 'Lion Yard': 2})

Pubs are a lot more spread out and there are 84 of them altogether. There are two on each Cherry Hinton Road and Mill Road. They must hence be main entertainment areas of the city.

(<type 'int', {": 11, 'Thoday Street': 1, 'Wulfstan Way': 1, 'Lensfield Road': 1, 'Mill Lane': 1, 'Norfolk Street': 1, 'Mawson Road': 1, 'Cherry Hinton Road': 2, 'High Street, Trumpington': 1, 'Broadway': 2, 'Pound Hill': 1, 'Carlton Way': 1, 'Mill Road': 2, 'Chesterton Road': 2, 'Covent Garden': 1, 'Fair Street': 1, 'Sturton Street': 1, 'Portugal Place': 1, 'Sleaford Street': 1, 'High Street': 2, 'Bene't Street': 2, 'Gwydir Street': 2, 'Melbourne Place': 1, 'Hills Road': 2, 'St Andrew's Street': 3, 'Panton Street': 1, 'Regent Terrace': 1, 'Bridge Street': 2, 'Brookfields': 1, 'Barton Road': 1, 'Tenison Road': 1, 'St Kilda Avenue': 1, 'Clarendon Street': 1, 'Water Street': 1, 'Dover Street': 1, 'Histon Road': 1, 'East Road': 2, 'King Street': 4, 'Prospect Row': 1, 'Devonshire Road': 1, 'Magdalene Street': 1, 'Victoria Road': 1, 'Milton Road': 1, 'Regent Street': 1, 'Russell Court': 1, 'Kingston Street': 1, 'Downing Street': 1, 'Newnham Road': 1, 'Ferry Path': 1, 'Newmarket Road': 3, 'Orchard Street': 1, 'Castle Street': 3, 'Northfield Avenue': 1, 'Peas Hill': 1, 'Station Road': 1})

It looks that Cambridge must be more of a pub-ish kind of a place since there are only 5 nightclubs.

Next, in terms of most popular late night food type it is actually Chinese:

(*<type 'int'>, {": 18, 'chinese; fish_and_chips': 1, 'chinese,indian': 1, 'kebab': 6, 'chinese': 15, 'kebab; burger; chicken; pizza': 1, 'chinese,fish_and_chips': 1, 'kebab;burger': 1, 'sandwich': 14, 'tex-mex': 1, 'burger': 2, 'fish_and_chips': 9, 'indian': 8, 'indian; pizza': 1, 'tandoori': 1, 'chicken': 4, 'chicken; pizza; kebab': 1, 'chinese;fish_and_chips': 1, 'pizza': 6}*)

And then going by the post codes, both pubs and fast food places located mainly in CB1, CB2 and CB4 areas:

FAST_FOODS{": 1, 'CB2': 15, 'CB1': 22, 'CB4': 12, 'CB5': 5}

PUBS{": 1, 'CB2': 15, 'CB3': 8, 'CB1': 28, 'CB4': 11, 'CB5': 3}

Complete list of post codes below. There are also no major problems with their format.

9 pubs are missing postcodes

{": 9, 'CB4 2LG': 1, 'CB1 2BH': 1, 'CB1 1JX': 1, 'CB1 3NW': 1, 'CB1 1DU': 1, 'CB1 1JT': 1, 'CB3 9JZ': 1, 'CB1 2NU': 1, 'CB2 1AA': 1, 'CB4 1HB': 1, 'CB3 9EY': 1, 'CB5 8AF': 1, 'CB1 1DB': 1, 'CB1 1DY': 1, 'CB1 1BG': 1, 'CB4 1XG': 1, 'CB1 2LL': 1, 'CB1 2DW': 1, 'CB1 2QF': 1, 'CB2 1AW': 1, 'CB1 2LG': 1, 'CB1 2LF': 1, 'CB1 2JB': 1, 'CB2 1HW': 1, 'CB2 1NW': 1, 'CB1 1LN': 2, 'CB1 1HA': 1, 'CB3 0AE': 1, 'CB1 1LD': 1, 'CB2 1UF': 2, 'CB3 0AJ': 2, 'CB4 3BA': 1, 'CB2 1HL': 1, 'CB2 1EN': 1, 'CB4 3AX': 2, 'CB2 3DB': 1, 'CB4 1NZ': 1, 'CB1 1EQ': 1, 'CB4 3JB': 1, 'CB1 3AX': 1, 'CB1 2AZ': 1, 'CB3 0AH': 1, 'CB2 1LQ': 1, 'CB1 1LH': 1, 'CB2 9HZ': 1, 'CB1 2EA': 1, 'CB1 2PW': 1, 'CB2 3QN': 2, 'CB4 2DG': 1, 'CB2 1PG': 1, 'CB1 3AA': 1, 'CB3 0AF': 1, 'CB2 3PN': 1, 'CB4 1HA': 1, 'CB5 8JE': 2, 'CB3 9NQ': 2, 'CB1 7AW': 1, 'CB4 2QA': 1, 'CB2 3AR': 3, 'CB2 1RX': 1, 'CB5 8EG': 1, 'CB1 2HS': 1, 'CB4 3DZ': 1, 'CB1 7AA': 1, 'CB3 9NF': 2}

12 fastfood places are missing postcodes:

{": 12, 'CB2 1EG': 2, 'CB5 8RG': 2, 'CB1 1BG': 1, 'CB1 2QF': 1, 'CB1 2AD': 4, 'CB2 1AB': 1, 'CB4 1EH': 1, 'CB4 1RJ': 1, 'CB1 1DG': 1, 'CB4 1XE': 2, 'CB1 3LP': 2, 'CB2 1AW': 2, 'CB2 1LA': 3, 'CB2 1NW': 1, 'CB1 3BE': 1, 'CB1 1LN': 1, 'CB2 3NP': 1, 'CB2 3LL': 2, 'CB2 1DP': 1, 'CB1 7DY': 1, 'CB1 2LD': 2, 'CB2 1UF': 1, 'CB4 2NF': 1, 'CB5 8WL': 1, 'CB2 1NT': 1, 'CB2 3NA': 2, 'CB2 3NB': 1, 'CB2 1DB': 1, 'CB4 3BU': 1, 'CB4 3AX': 1, 'CB1 2AS': 1, 'CB4 3JD': 1, 'CB1 2AW': 3, 'CB4 3HG': 1, 'CB1 3AW': 1, 'CB4 1NS': 1, 'CB4 3HL': 1, 'CB4 1NJ': 1, 'CB1 8QJ': 1, 'CB1 3AH': 2, 'CB5 8LG': 1, 'CB1 2BD': 1, 'CB2 8PB': 1, 'CB1 1NE': 1, 'CB5 8JE': 2, 'CB1 7AW': 4, 'CB2 3RD': 1, 'CB1 7AJ': 1, 'CB4 1JY': 4, 'CB1 3ET': 1, 'CB5 8PB': 1, 'CB1 7AA': 3, 'CB1 7AG': 2, 'CB1 2NJ': 1}