**2. DATA DESCRIPTION**

Based on the above problem, we require the number, names and the type of restaurants present in Calgary.

Firstly, using the Foursquare API, the latitude and longitude of Downtown (Centre of Calgary) was extracted. Then, used those coordinates to find other present restaurants within a certain radius. The returned json file was then turned into a data frame.

**2.1 DATA CLEANING**

The data frame included columns like:

1. Categories
2. Hasperk
3. Id
4. Location. Address
5. Location.cc
6. Location. city
7. Location. country
8. Location. crossStreet
9. Location. distance
10. Location. formattedAddress
11. Location. labeledLatLngs
12. Location. lat
13. Location. lngs
14. Location. Postalcode
15. Location. state
16. Name
17. Referral id
18. Venue page id

It was transformed according to the needs. The redundant columns like Perk, referral id, and venue page id were removed. In the rest of the columns, location. was removed and the category type was extracted from a long tuple of values in the category column. The final data frame had 36 restaurants and looked like:

