IBM DATA SCIENCE CAPSTONE PROJECT

ATHIRA GOVINDAN KUTTY

BEST LOCATIONS TO START A BOOKSTORE

A project to find best locations to start an enterprise of a bookstore in Warwickshire by using Foursquare API.

INTRODUCTION

- This project is aimed to find best locations to start a small enterprise of a bookstore in Warwickshire.

- It is done with the help of Foursquare API and a number of libraries in python.

- The unsupervised machine learning technique of k-means clustering is also used.
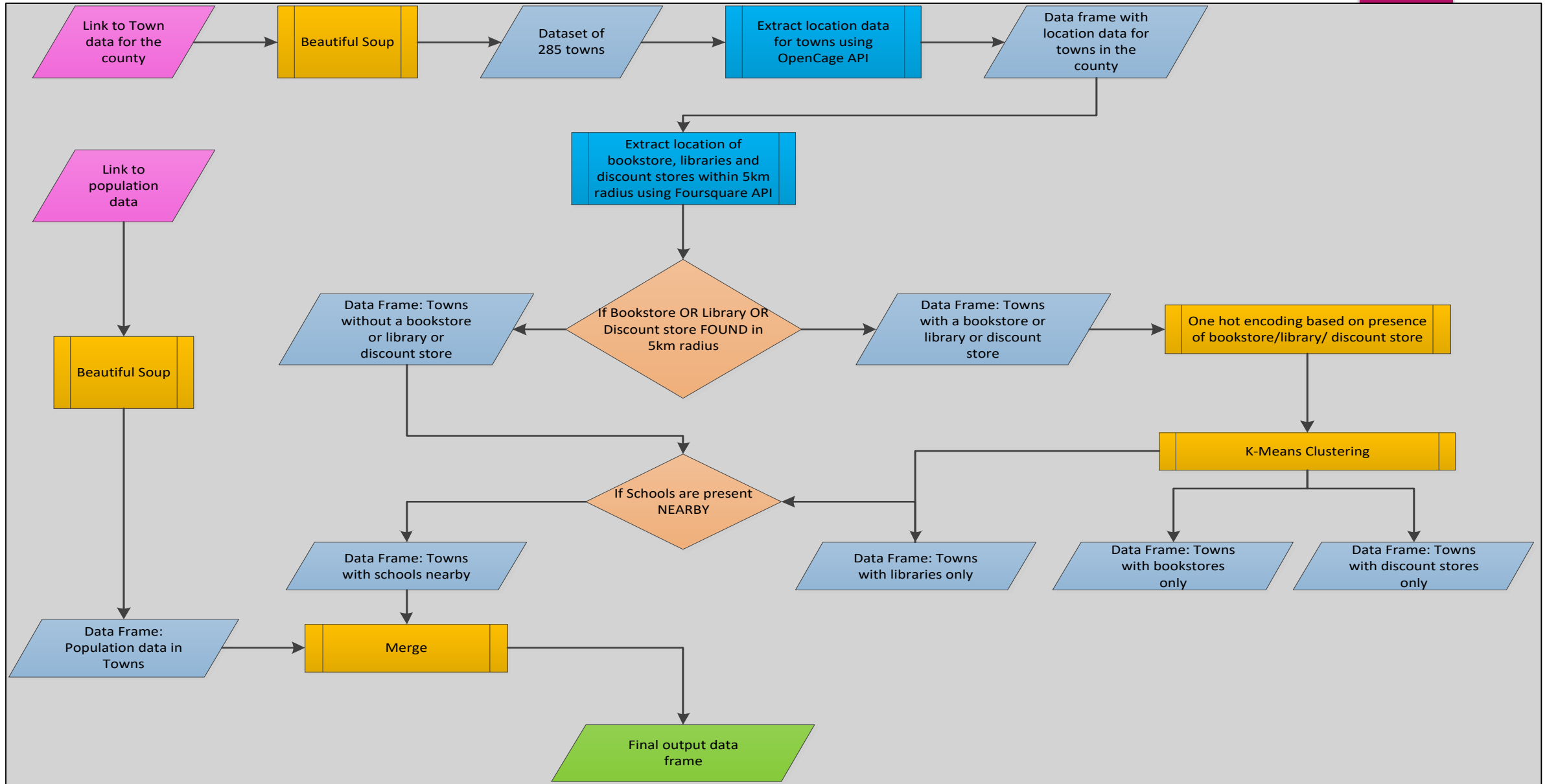
DATA

DATA NEEDED

- Towns in Warwickshire county

- Location coordinates of these towns

- Accessibility of bookstores or libraries within a commutable distance

- For towns without any bookstores or libraries, does the town have a big enough population to start a venture. Also, the presence of schools or colleges in the vicinity so that the store could be useful for students and teachers
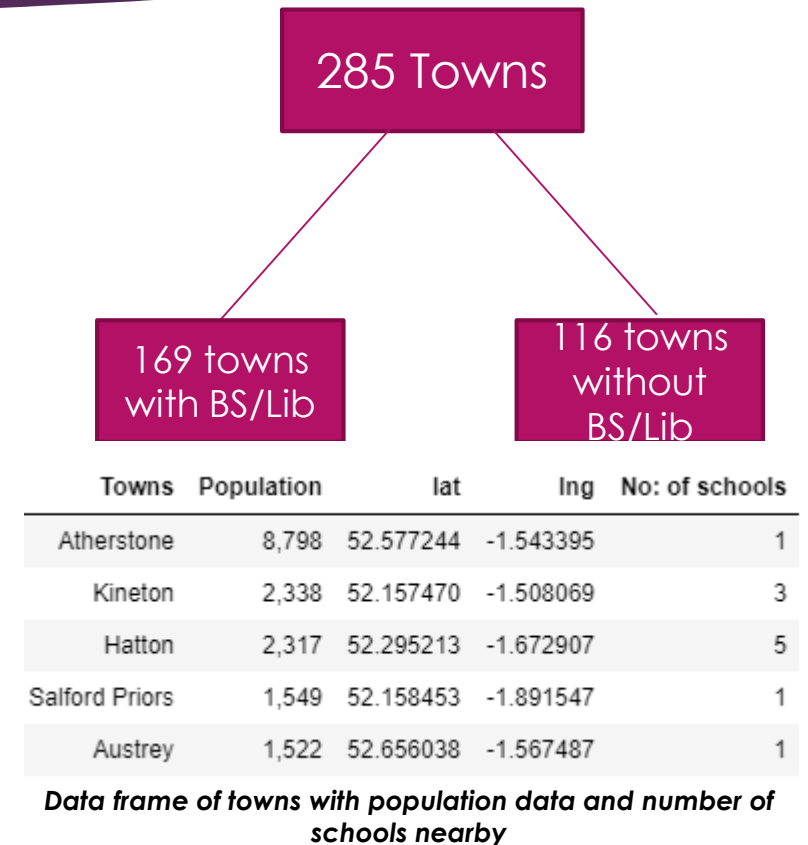
DATA EXTRACTION

- Data was extracted using beautiful soup library from the following websites:

  - List of towns: https://worldpostalcode.com/united-kingdom/england/Warwickshire

  - Population data: https://www.ebayinuk.co.uk/warwickshire-towns-villages-population/

- Location coordinates of the towns were obtained using OpenCage API.

- Foursquare API was used to collect details of bookstores, libraries and schools in these towns.

FLOWCHART

METHODOLOGY

- List of 285 towns in Warwickshire was scrapped from the website and converted into a data frame.

- Location coordinates were extracted using Open Cage API.

- Bookstores and libraries within 5km radius of these towns were found using Foursquare API.

- Out of 285 towns 169 towns have either a bookstore or a library while the rest of 116 towns do not have libraries or bookstores.

- Towns without bookstores or libraries are taken for further analysis and population and vicinity to schools is checked.

- For this, population data was scrapped using beautiful soup and merged with data frame of towns and its location data

- Using Foursquare, a data frame of schools nearby is obtained.

285 Towns

169 towns with BS/Lib

116 towns without BS/Lib

Towns	Population	lat	lng	No: of schools
Atherstone	8,798	52.577244	-1.543395	1
Kineton	2,338	52.157470	-1.508069	3
Hatton	2,317	52.295213	-1.672907	5
Salford Priors	1,549	52.158453	-1.891547	1
Austrey	1,522	52.656038	-1.567487	1

Data frame of towns with population data and number of schools nearby
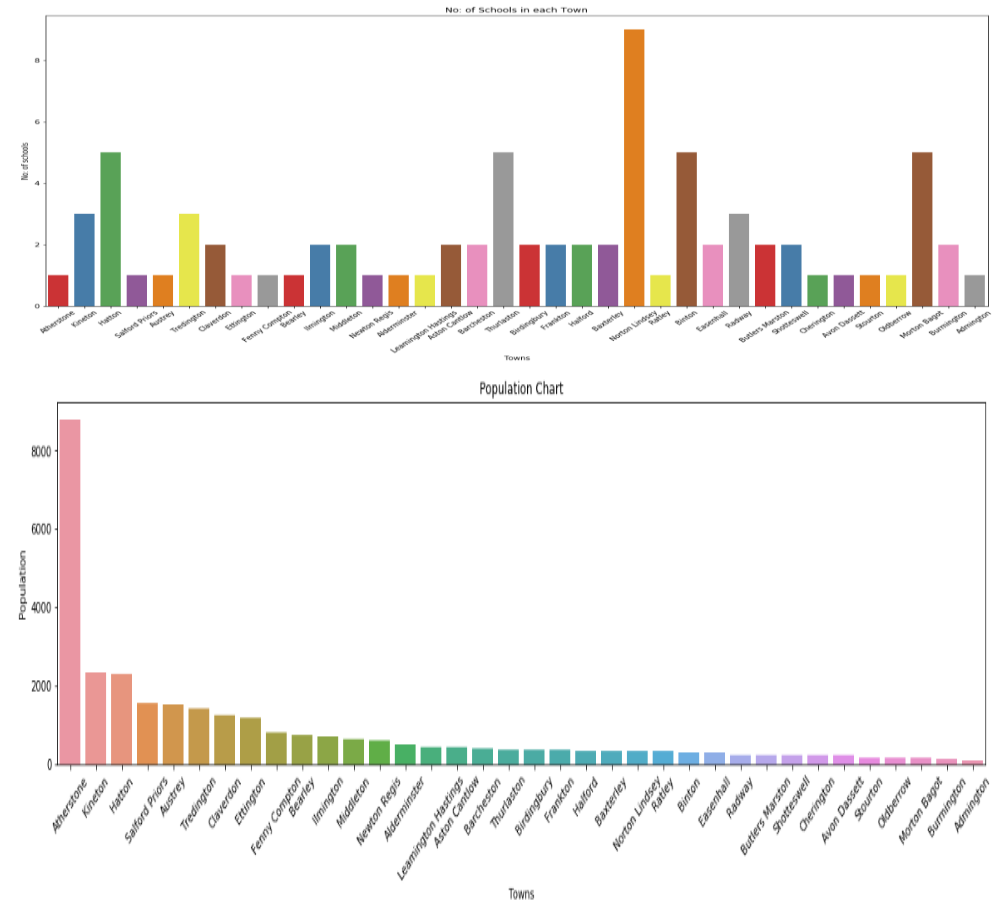
METHODOLOGY

- Population and School data was then plotted using the seaborn library for further analysis.

- Towns that already had bookstores/libraries were then clustered.

- One-hot encoding was done to get how many bookstores and libraries are there in a town.

- This data was then clustered to get 5 different clusters.

- Among these cluster the cluster with towns that only have libraries was then considered as these will also be a good option to start a bookstore.

- As done before the population and schools data were taken for further analysis.

Cluster	No: of towns	Towns
Cluster 0	4	Towns with discount stores only
Cluster 1	49	Towns with libraries only
Cluster 2	25	Towns with bookstores only
Cluster 3	39	Towns with more libraries and discount stores than bookstores
Cluster 4	52	Towns with more bookstores than libraries and discount stores

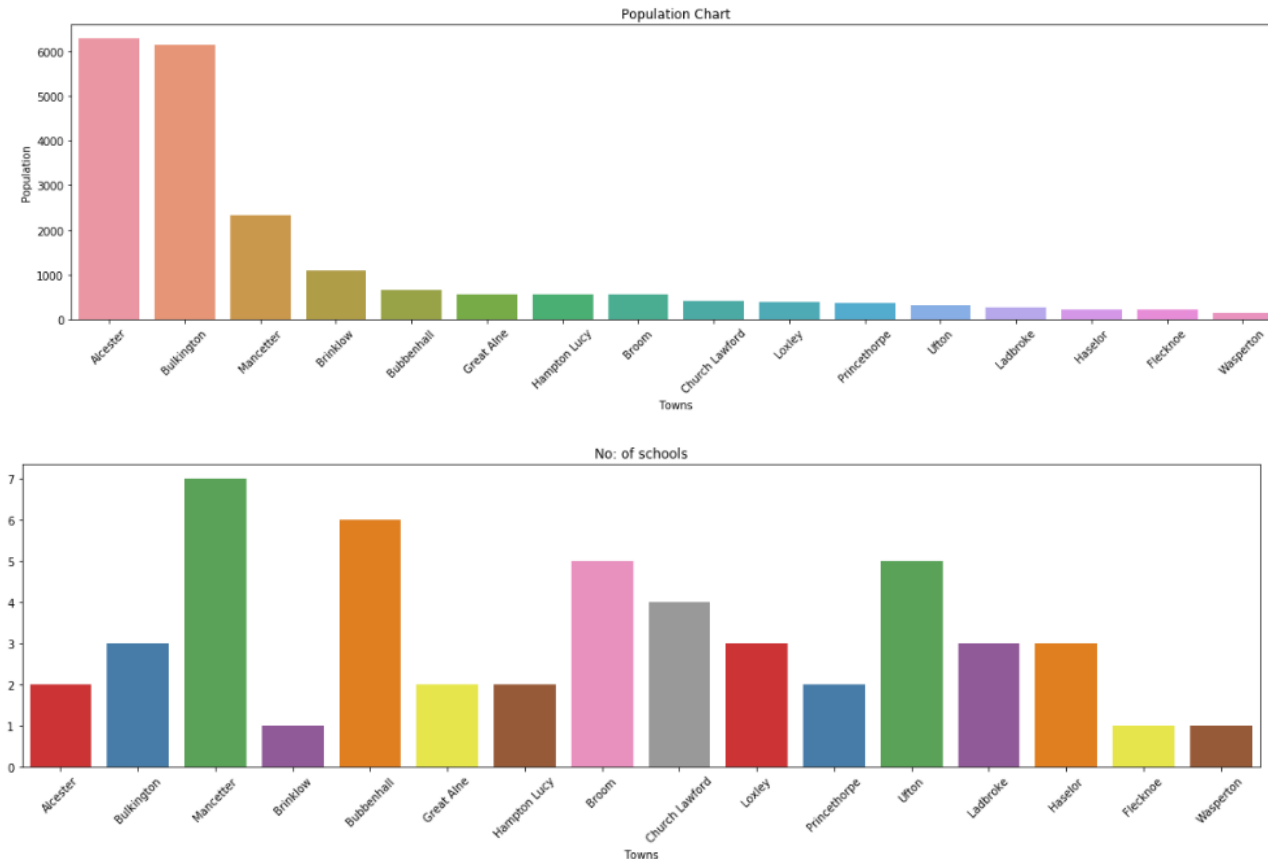*Clusters obtained after k-means clustering*

RESULTS

- By analysing the data set of towns with no libraries or bookstores, 3 towns that are best suited for the purpose was obtained.

- From the visualisations, Atherstone was the town with highest population.

- While Kineton and Hatton have appreciable population and they also have schools nearby.

- Hence Atherstone, Hatton and Kineton are the best options in this set of towns.

No: of Schools in each Town

Population Chart

RESULTS

- While taking the towns that already has bookstores and libraries, similar analysis shows that the towns of Alcester, Bulkington and Mancetter are the best options.

- Hence, from 285 towns 6 towns that are best suited for opening a bookstore was obtained.
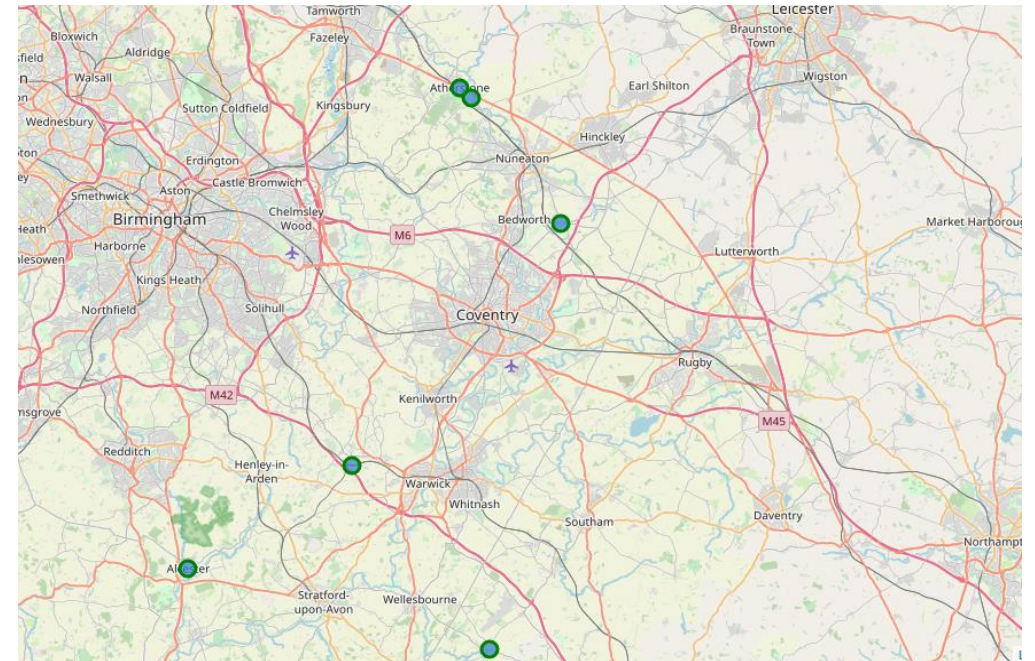
CONCLUSION

- By using Foursquare API, K-means clustering technique and a number of libraries in python, 6 towns that are best suited to start a bookstore was obtained. They are:

  ❖ Atherstone

  ❖ Hatton

  ❖ Kineton

  ❖ Alcester

  ❖ Bulkington

  ❖ Mancetter

Towns	No.of schools	Population	No.of libraries
Atherstone	1	8798	0
Alcester	2	6279	1
Bulkington	3	6147	1
Mancetter	7	2338	1
Hatton	5	2317	0
Kineton	3	2338	0

Dataset and map of the shortlisted towns

DISCUSSION

➢ Population and vicinity to schools are the only factors considered for this project other than these various other data such as vicinity to malls, railway stations, supermarkets, rent for commercial buildings and many more can also be considered.

➢ Population data of all towns was not available.

➢ Bookstores helps the purpose of sharpening a man's mind with different books and they will always have a brighter future. Even now there are many towns that do not have any source of books like a library or a bookstore. These towns can be found using this method with a few improvements and new ventures can be started here.

REFERENCES

The Futures is bright for Print books: https://www.twosides.info/UK/print-book-sales-rise

Printed books have a positive future: https://www.storaenso.com/en/newsroom/news/2018/5/printed-books- have-a-positive-future-a-new-survey-indicates