**Bangalore Metro Station Classification**

**Introduction to Business Problem**

The Objective is to classify the metro stations in Bangalore city on the basis of the top venues near to it. These venues are an identification of the type of area. With an increase in traffic in cities like Bangalore, fair cost and ease of usage offered by metro, more people have started using metros instead of their private vehicles in Bangalore. Thus this study will be helpful to understand the type of developments near Metro stations in Bangalore. The data can be used by those who wants to start a new business like restaurants , movie theatres, fitness facilities etc, such that they can understand the type of areas and choose an appropriate area that satisfy them. Also it can also be used by city planners to understand the developments

**Data**

The data of Metro Stations was taken from the Wikipedia page by data scrapping data, using BeautySoap. The data was then cleaned using excel. The coordinates of different metro stations was then found using API of geocode.

The entire dataset was then stored on to a csv file.

**Exploratory analysis & basic clean-up**

The data was cleaned in excel and the some exploratory data visulaizations were done. A box plot was drawn to identify the most common type of venues near metro stations.![A close up of a logo

Description automatically generated]()

The data had to be normalised, and this was also done.

![A close up of a piece of paper

Description automatically generated]()

**Clustering**

Clustering of the different metro stations were done based on the top venues near it. This can be used to identify the type of development happening in each area and give details regarding purpose for which people should be using the Metros.

![A picture containing text, map

Description automatically generated]()![A picture containing implement, pencil

Description automatically generated]()

**Conclusion**

Data analysis can be very helpful when dealing with large amounts of data. They can provide useful insights through either visualization tools or through more specific tools like clustering, which can then be used to obtain valuable business information.