Final Report

## Introduction/Business Problem

Cafes have become incredibly popular in Australia in recent years. Cafe Help is a Coffee Shop Supplies & Repairs business based in Melbourne Australia. Cafe Help provides Cafes with supplies (eg. beans, milk, cups) and repairs (eg. machine and fridge). It is crucially important that Cafe Help is within short proximity of local cafes.

Cafe Help has decided to expand to Sydney, Australia and would like to start with three service locations. My capstone project will look to recommend the three best locations for Cafe Help to begin operations in Sydney.

This study is explicitly targeted to the management of Cafe Help, however it would also be of interest to other Coffee Shop Supplies & Repairs businesses.

## Data

Cafe Help wants to be within close proximity of local cafes in Sydney. We will therefore use FourSquare data to build a cafe database in the Sydney region. We will then look to analyse cafe density in the Sydney region. We may put an emphasis on the most popular cafes in the region as these cafes need more supplies and repairs. Finally we will look to build clusters to determine our recommended three locations.

Foursquare’s Search endpoint will allow us to build up such a database of cafes in the Sydney region. The Likes, Explore and Trending endpoints will allow us to focus on cafes that are more popular. Jupyter Notebooks, Python 3, Pandas will allow us to wrangle, analyse this data. K-means Clustering will allow us to develop density-based clusters. Folium will allow us to present our findings back to the main stakeholders in a way that tells the story.