
APPLIED DATA SCIENCE
CAPSTONE
NEW YORK
VS
MELBOURNE
A Tale of Two Cities

CONTENTS

INTRODUCTION

1.1 Business Case Scenario

I'm a post-graduate applicant considering offers from two universities - Columbia Journalism School in New York and the University of Melbourne. Both academic programs have equal value to me and and the tie-breaker is city livability. I am measuring this by evaluating neighbourhood clusters in each city based on **three** key considerations before I make my choice.

1.2 Key Considerations

- Walking distance to each university <= 5 km (or 3.11 miles)
- Crime rate
- Availability of local attractions and venues

1.2 Other Target Audience

- University students considering relocation
- University academics and staff considering relocation
- Property investors targeting the university body
- Business owners targeting the university body

DATA

2.1 New York Dataset

- FourSquare API to obtain local attractions and venues data via Explore call
- City of New York (NYPD Complaint Data) - https://data.cityofnewyork.us/Public-Safety/NYPD-Complaint-Data-Historic/qgea-i56i
- City of New York (Boundary Data) - https://discover.data.vic.gov.au/dataset/crime-by-location-data-table

2.2 Melbourne Dataset

- FourSquare API to obtain local attractions and venues data via Explore call
- Vic.gov.au Crime Statistics Agency Data Tables (Criminal Incidents) - https://discover.data.vic.gov.au/dataset/crime-by-location-data-table
- Australia Post data - https://www.matthewproctor.com/australian_postcodes

2.3 Data Tools

- The geographical map will be created using Folium and FourSquare API
- Pandas will be used to calculate the average crime statistics of each area