Data

This includes identifying the necessary data content, formats and sources for initial data collection.

Content

Below is the broad statistics complied using Foursquare location data for New York city. I will replicate this for Toronto and this will form the base content for my data.

Leisure Travelers	
	Count
Museums	>60
Theatres	>60
Shopping	>100
Food/Cafe	>100
Bars and restaurants	>100

Business Travelers	
	Count
Banks	>100
Insurance	>30
Law Firms	>60

Sources

a) Foursquare Location Data

Using the location services in Foursquare, I could explore in detailed whether Toronto has what New York offers to these travelers. It could be further refined by taking into account the ratings of these establishments in my study.

The statistical compilation will look like the following

Leisure Travelers		
	Count	Count where Ratings above 7
Museums		
Theatres		
Shopping		
Food/Cafe		
Bars and restaurants		

Business Travelers

	Count	Count where Ratings above 7
Banks		
Insurance		
Law Firms		

Format

JSON

```
latitude = 43.6532
longitude = -79.3832
```

```
search_query = 'Categories'
radius = 10000
```

```
url = 'https://api.foursquare.com/v2/venues/search?client_id={}&client_secret={}&ll={},{}&v={}&query={}&radius={}&limit={}'.format(CLIENT_ID, CLIENT_SECRET, latitude, longitude, VERSION, search_query, radius, LIMIT)
```

b) Wikipedia

In addition, I will supplement my content by web scraping the following website

https://en.wikipedia.org/wiki/Tourism_in_Toronto

With the information garnered, I could establish whether Toronto could satisfy the demands required by the same travelers to New York. As such, based on this data analysis, I am able to confidently present to my client our recommendation on adding Toronto as a new destination.