Project Report

TABLE OF CONTENTS

- 1.0 Introduction
- 2.0 Data
- 3.0 Methodology
- 4.0 Results
- 5.0 Discussion
- 6.0 Conclusion and summary
- 7.0 Acknowledgement

1.0 INTRODUCTION

1.1 What is Vacation?

Vacation involves spending time away from home. It may involve local or international travel. It could be for a few days or more. Regardless of the length of time, a vacation should be built to serve its purpose: relax or do nothing, learn something new, experience a new location, or take a holiday job. Proper planning helps to make a vacation worthwhile, irrespective of the purpose of the vacation. Therefore, the plan should be drawn before commencing the holiday. So, while it is fun to go somewhere, it is more important to make plans for an enjoyable holiday to create memorable experiences.

1.1.1 Planning a Vacation

Proper planning helps to increase the possibility of having an exciting and worthwhile vacation. However, the process of planning a vacation can be exhausting, leading to unfulfilled vacations. A common problem with planning a vacation is how to choose the places to visit since the advertised options are as varied as they are many. This problem causes anxieties because there are many uncertainties for which the answers are not readily available. For a vacation that is supposed to be recreational and fun, planning it may be its doom because the vacation may eventually be postponed, cancelled, or poorly planned.

1.1.2 Scenario

Jones, a staff of an organisation, was exceptional in delivering services during the past 12 months amidst the pandemic. The company rewarded him with a 14-day family vacation from Leeds in the UK to Winnipeg in Canada. The vacation is scheduled for August 15 to August 29 at Holiday Inn. Jones excitedly announced the trip to his family (partner and two children). Everyone was ecstatic about the summer trip to Canada – a country they had hoped to visit as a family. The discussion each day was about the recreational activities to engage in and places to visit.

1.1.3 Problem

Jones noticed that divergence in the preferences and needed a way to harmonise all their different choices. He approached his friend, Keitha - a data scientist, who offered to help with planning the recreational activities and places to visit. Jones supplied Keitha with a combined list of recreational activities and exciting places liked by each family member. Keitha intends to write a programme that would juxtapose all the individual preferences with the venues and activities available around the Holiday Inn.

1.2 Objectives

To design a programme that will

- streamline the process of searching for available venues at a particular location and produce a list of preferred venues with acceptable attributes
- increase the ease and speed of getting required venue information, thereby reducing turnaround times in creating a plan for visiting different holiday venues
- make it easy to modify the itinerary for a vacation in the event of an unexpected occurrence

1.3 Target

The target of this programme is in two categories. Target Market and Target Audience.

Target Market represents the broad market, from trip/vacation planners to the entire tourism industry.

Target Audience is made up of a more focused group. It represents individuals and organisations, who regularly engage in the process of planning a holiday from Holidaymakers to Travel Agencies.

1.4 Application

This programme can be used in two broad ways to provide sensational experiences during vacations and tours. These are General Application and Specialised Application.

1.4.1 General Application

This programme can be used to get information about a venue or location that can be educative and influence decision making. It provides information on locations and venues. It provides specific information for use when creating an itinerary for a vacation or trip. It also improves the ability to interact knowledgeably with others using information about a location before visiting it.

Examples are:

- To get information about a location before visiting the place
 - o Identify popular venues in a location, get details about venues and locations that would influence activities to engage in
- Compare sites to improve decision-making when choosing a place to visit,
 - assess locations to confirm that places that are personally of interest are available, obtain information required to shortlist venues of particular interest, get recommendations about different venues and get information to plan recreational by determining venues that offer the recreational activities of interest
- To plan a trip that is flexible and can accommodate unscheduled activities or events
 - To plan an itinerary for the vacation by determining the distance between venues, parking availability, etc., get information about the facilities available at a venue, identify venues on a map in real-time,

1.4.2 Specialised Application

There are two sub-categories: Data and Response Time, and Packages and Itinerary

1.4.2.1 Specialised Purpose – Data and Response Time

The impact of its application enhances the market value of products and services in offer, e.g., excursion and tour packages. It increases efficiency and effectiveness when responding to clients' tripplanning requests. This application is categorised into three use-cases: real-time information, response time improvement, location comparison

• Real-time information

The information is provided in real-time includes venue photos, venue recommendations, venue ratings, venues addresses and data for research.

• Improved Response Time

This programme will reduce the time to respond to each client's requests, thereby creating a positive effect on the rating of the trip planner or travel agency. It will increase the turnaround time for each client's request, making it possible to attend to more clients' requests within the

same time frame, positively affecting the revenue generated within a specific timeframe. It will improve recent data that will give clients vivid insight about venues which ultimately reduces the time client's respond to options provided

Locations and Venue Comparison

It provides current user-generated content that can compare exciting places of attraction in different locations to increase venues options and aid decision-making. It identifies venues within the same category with desirable attributes, which helps streamline venue options and create itineraries.

1.4.2.2 Specialised Purpose – Packages and Itinerary

The four use-cases are custom-built packages, itinerary creation, itinerary modification, and marketing services

Custom-built Packages

It is used to create custom-built vacations to meet clients' unique needs by specifying client's preferences and requirements. It will discover interesting venues that are near a particular reference point. Special vacation packages that will incorporate places of interest for a tour group or vacationers will be incorporated. The attributes of venues that will be used to select venues to visit will be identified when planning trips and tours

Itinerary creation

It is used to create a fit-for-purpose route for each client. Travel plan for group tours such as students' excursions, cultural heritage tours, etc., can be more easily created. It also easily identifies different itinerary possibilities, which can be identified and showcased to clients

Itinerary modification

It increases flexibility in the itinerary service provided to clients to accommodate sudden changes in the client's circumstance by using real-time information and recommendations about venues

Marketing tour services

It provides information that can be used to deliver add-on services to clients' considering customised packages.

1.4.3 Application – Broad Usage

In a nutshell, this programme can be used in different ways such as to

- plan family outings
- arrange reunion getaway activities for friends and alumni
- identify venues for corporate team building activities
- create itineraries for excursions
- design general and specialised tour packages

The programme can be used to

- obtain data about venues in a locality
- search for places of interest (POI) around a location
- get venue ratings and recommendations
- get the distance between venues
- identify venues of personal interest that are within a vicinity
- identify venues on a map in real-time

1.4.3.1 Featured Case: Jones' Family Vacation - Solution

Jones gave the following details to Keitha.

- The accommodation was booked at Holiday Inn, Winnipeg, Canada.
- The aggregated list of venues of individual interest is 'Gym', 'Food Court', 'Restaurant', 'Pizza Place', 'Art Gallery', 'Coffee Shop', 'Theater', 'Juice Bar'.

The programme written by Keitha would request Foursquare to supply the data - a list of all venues close to Holiday Inn, Winnipeg. It would generate a map to visualise the venues close to Holiday Inn. The venue data provided by Foursquare will be filtered to confirm that the preferred venues supplied by The Jones' are available in that vicinity. The venues will be segmented and popular venues among the filtered venues will be identified. Clusters will be created from the segmented venues to determine the most common venues within each venue segment. The venue ratings and recommendations of each filtered of the venues will be obtained. This will determine the venues within a particular category that will be added to the itinerary. A check of trending venues will be done to analyse the data collated, displayed on a map for further analysis. The program's runtime will be less than two (2) minutes

1.4.3.2 Fe9atured Case – Advantages of the Solution

The benefits of using this

- The venue data from Foursquare is recent and supplied by venue users
- The list of the venues provided by the programme are shortlisted based on the individual preferences garnered from Jones' family members
- The programme displays venue results on a map in real-time
- Users that recently visited the venues independently supply the venue ratings and recommendations obtained from Foursquare
- The time-to-finish quick making it easy to check for multiple venue search requests within a short time

The programme provides venue data, tips and recommendations, popular venues, and most common venues in real-time.

2.0 DATA

The data is sourced from Foursquare.

2.1 Foursquare

- It is a technology company with a massive dataset of location-data
- The location data is available for use by anyone
- It is used by many companies (such as Twitter, Airbnb, Samsung, and Uber) whose services are driven by location data
- The dataset is crowd-sourced individuals connect to Foursquare using either the Foursquare App or Website to add venues or update data about venues

Source: https://foursquare.com, https://foursquare.com/products/places/

2.1.1 Data Source - Data

2.1.1.1 Foursquare Places - Data

Offers real-time access to a global database of rich venue data and user content used to drive location-based experiences on apps and websites.

The data from Foursquare Places

- represents 100M+ commercial points-of-interest (POI)
- covers 248+ countries and territories
- contains 80+ extended attributes
- is trusted by top enterprises their location-based services
- is updated regularly by the Foursquare Consumer App Superuser Community
- is accessible via Foursquare API (Application Programming Interface) or as a Flat File

Source: https://docs.foursquare.com/docs/places-data-overview

2.1.2 Data Source – Data Features

2.1.2.1 Foursquare Places – Data Features

The venue data features are available as a flat file in JSON or TSV format.

The features of the venue data provided in real-time by Foursquare Places are

- 60M+ commercial points-of-interest (POI)
- 941 Venue Categories
- 1M+ fresh Tips (added within the past six months)
- 18.5M fresh Photos (added within the past six months)
- 7.6M+ Venues with popular hours
- 23K+ Chains, including retailers, auto dealerships, etc.

Source: https://developer.foursquare.com/docs/places-database/details/

2.1.3 Data Source – Places (Access & Usage)

2.1.3.1 Foursquare Places API – Access and Usage

The following control access to and usage of Foursquare API

- API Endpoint this is the point where two applications connect
- Rate Limit this specifies the maximum number of calls allowed within a specific time interval
- Authentication used to recognise users and authorise access

Source: https://developer.foursquare.com/docs/places-api/getting-started/

2.1.4 Data Source – API Endpoint

2.1.4.1 Foursquare API Endpoint

There are two types of API Endpoints

- Regular
 - Contains basic venue descriptive attributes such as data, category, and ID
- Premium
 - Includes all basic attributes and other rich data such as ratings, URLs, photos, tips, menus, etc.

Note:

Restrictions on endpoints are applied based on the type of account in use.

Regular Endpoint



Premium Endpoint



Source: https://developer.foursquare.com/docs/places-api/endpoints/

2.1.5 Data Source – Rate Limit

2.1.5.1 Foursquare Rate Limits

What is Rate Limit?

- It regulates the use of API
- It determines if a Regular or Premium API call will be made

Types of Rate Limit

- Daily Call Quota
- Hourly Rate Limit

Note:

- Rate limit acts per top-level endpoint, not per endpoint
- The limit that will be placed on the usage of the API will depend on whether hourly rate limit or daily call quota occurs first

Daily Call Quota

Account Tier	Sandbox	Personal	Start-up
Regular API Calls	950	99500	Call Foursquare for Upgrade
Premium API Calls	50	500	Call Foursquare for Upgrade

Call Quota

- represents the maximum number of API calls allowed in 24 hours
- resets daily at midnight UTC
- is dependent on the tier of the developer's account

Hourly Rate Limit

Туре	Hourly Rate Limit
Userless requests to venues/* endpoints	5000
Userless requests to other endpoint groups	500

Source: https://developer.foursquare.com/docs/places-api/rate-limits/

2.1.6 Data Source - Authentication

2.1.6.1 Authentication

Refers to the codes that allow users to connect to Foursquare.

There are two types of authentications:

- userless auth
- user auth

Userless Auth

- a user's permission is not required
- requires Client ID and Client Secret

User Auth

- user's permission is required
- there are two types
 - iOS/Android (which uses native auth)
 - Web (which uses OAuth)

Note:

OAuth is a standard that allows applications and websites to access resources on other applications or websites on behalf of users.

Source: https://developer.foursquare.com/docs/places-api/authentication/

2.1.7 Data Source - Places API

2.1.7.1 Foursquare Places API - How to

To use Places API, the developer would

- create a developer's account
- create a Foursquare app and submit the URL for the host
- obtain app credentials Client ID and Client Secret
- make an API call to get requests of Foursquare's servers
- use OAuth access to make authenticated calls

Note:

- A developer can choose from different account options.
- The type of account determines the amount of data obtainable at any point in time.

Source: https://developer.foursquare.com/docs/places-api/getting-started/

2.1.8 Data Source - Accounts

2.1.8.1 Foursquare Accounts

There are four account tiers:

- Sandbox (Initial Account Created)
- Personal account (Non-commercial)
- Start-up (Commercial)
- Enterprise (Commercial)



Note:

- Upgrade from Sandbox to Personal needs credit card verification.
- A personal account can be upgraded to a more suitable account as needed.

•

Source: https://developer.foursquare.com/docs/places-api/getting-started/

2.2 Foursquare Venue

2.2.1 Foursquare Venue

- A real-world (physical) location and real place where people can meet up or like to spend some time
- Has a name, geographical coordinates (latitude and longitude), and a primary category

2.2.2 Foursquare API – Venue Search

- Used to identify venues around a particular location.
- The endpoints used determines the results obtainable from a venue request. https://api.foursquare.com/v2/venues/search https://api.foursquare.com/v2/venues/explore

Source: https://developer.foursquare.com/docs/listing-syndicators/venues/

2.3 Foursquare Venue – Venue Search

2.3.1 How to Find a Venue on Foursquare

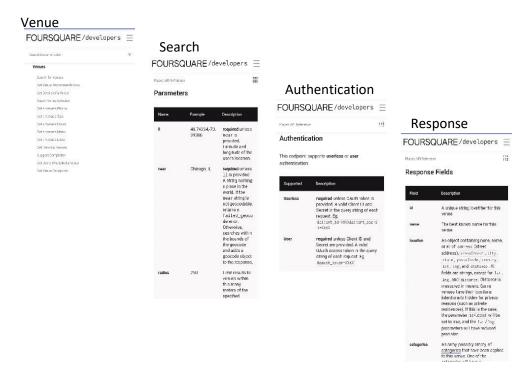
• Get the ID of the venue



- Use the venues/search API endpoint to search for a venue using the venue's ID
- Indicate the name and geographical coordinates (latitude and longitude) of the venue
- Authenticate the connection using either a userless or user authentication
- Specify required parameters

Note:

• Include street address, city, state, zip, phone number, or Twitter handle as additional attributes to improve the search outcomes



Source: https://developer.foursquare.com/docs/api-reference/venues/search/

2.3.2 Foursquare Response to Venue Search

The table and picture show some fields supplied in the response.

Field	Description
id	A unique string identifier for the venue
name	The best know name for the venue
location	An object containing none, some, or all of 'address', 'crossStreet', 'city', 'state', 'postalCode', 'country', 'lat', 'lng' and 'distance' Note: A venue's location may be hidden due to privacy reasons (e.g. private residencies)
categories	An array, possibly empty, of categories that have been applied to this venue

```
Response

(**seta**). 6 Jap.

**seta**). 6 Jap.

**separate: **satistholooder/aboutrease*

**separate: **satistholooder/aboutrease*

**separate: **satistholooder/aboutrease*

**separate: **separate
```

FOURSQUARE /developers

Source: https://developer.foursquare.com/docs/api-reference/venues/search/

2.3.3 Foursquare Response – Description

- Venue Response is the data supplied in response to a successful API venue request.
- It contains a list of recommended venues near a specified location
- Each venue returned will include attributes such as venue name, venue ID, venue category, venue location (address), venue latitude and venue longitude
- The result from Foursquare can be displayed on a map

2.4 Data Skills

- Programming in Python
- Working with Foursquare API
- Data Wrangling
- Machine Learning
- Mathematical Operations
- Map Visualisation

2.5 Data Tools

Item	Description
KMeans	To form clusters
IPython	An Interactive command line used to display images
Matplotlib	To create static, animated and interactive visualisations
Folium	To visualise data on an interactive leaflet map
Numpy	For scientific computing
Pandas	For data analysis and manipulation
Scipy	For numerical computation
Scikit-learn	For predictive data analysis
Imbalanced-learn	Resampling techniques for balancing datasets

Item	Description
Python	Programming language
Python Notebook	Application to write and share code
Foursquare API	Real-time location data from the Foursquare server
Github	Repository for programmers
Geopy	To get coordinates of an address or location
Nominatim	To convert any address to latitude and longitude values
Haversine Formula	To calculate the great-circle distance between two points on a sphere
Math	For mathematical calculations
JSON	To handle JSON files
Requests	To send HTPP requests

3.0 METHODOLOGY

3.1 Preliminary Details

The initial activities carried out include the following

Foursquare

A developer account was created in Foursquare. Afterwards, Foursquare application credentials (Client ID and Client Secret) was obtained. Then, the foursquare application credentials were used to obtain the required data from Foursquare.

• IBM Watson Studio

An account was created on IBM Watson Studio, and then a notebook was created.

Programming in Python

Python programming language was used to write the code.

Preliminary Actions

Some libraries and modules were required for the code to function, and these were installed and imported. Some functions were also defined.

3.2 Data Requisition

The data used was wholly obtained from Foursquare. Foursquare was chosen as the source of data because it provides recent venue data and user content of over 60 million venues worldwide.

Foursquare Credentials

The foursquare credentials' Client ID' and 'Client Secret' were defined. Foursquare gives each developer unique Foursquare credentials.

Point of Reference

A point of reference (physical address and its longitude and latitude coordinates) was required to obtain venue data from Foursquare. The point of reference used is Holiday Inn, Winnipeg Downtown, Canada. Using Nominatim tool, the complete physical address and coordinates (longitude and latitude) were obtained.

The coordinates of Holiday Inn & Suites Winnipeg [Downtown, 360, Colony Street,	Colony, West End, Winnipeg, Mani	itoba, R3C 0E7, Canada are -97.15145373256	72, 49.8911612.

The latitude and longitude coordinates were converted to the physical address by using "geolocator".

49.8911612 -97.15145373256972

3.3 Data Acquisition

3.3.1 Retrieval of the Data from Foursquare

A point of reference (Holiday Inn, Winnipeg) was specified as the point of reference to obtain venue data from Foursquare. The coordinates of the Holiday Inn, Winnipeg was passed as one of the parameters of the Uniform Resource Locator (URL). URL is the informal name for a web address. A web address references the location of a resource on the internet. In this case, the location referenced by the URL is Holiday Inn. The other parameters passed were the foursquare credentials, latitude and longitude of Holiday Inn, the radius of search around Holiday Inn, and the total number of venues to be displayed.

With the URL clearly defined, a request for the venue data was made to Foursquare using the GET function. The URL was passed as a parameter within the GET function. As a result, Foursquare provided data of venues near Holiday Inn, Winnipeg, within the defined radius.

In the request for data, JSON (JavaScript Object Notation) function was specified as the format of choice that the data from Foursquare should be delivered.

For the data to be more meaningful, the data was filtered. Specific venue parameters that were required were specified and used to filter the data. The parameters were "name", "categories", "location", "id", which produced a dataframe that contained the desired venue parameters. A map of the filtered data was created and to visualise the venue locations relative to the Holiday Inn.

	name	categories	address	cross Street	lat	Ing	labeledLatLngs	distance	postalCode	cc	city	state	country	formattedAddress	neighborhood	id
0	Winnipeg Art Gallery	Art Gallery	300 Memorial Blvd	St. Mary Ave	49.889459	-97.150708	[{'label': 'display', 'lat': 49.88945877797471	196	R3C 1V1	CA	Winnipeg	MB	Canada	[300 Memorial Blvd (St. Mary Ave), Winnipeg MB	NaN	4b5f6471f964a520b4b729e3
1	Harvey's	Restaurant	600 Portage Avenue	NaN	49.888922	-97.156059	[['label': 'display', 'lat': 49.88892237017493	413	R3C 3L7	CA	Winnipeg	МВ	Canada	[600 Portage Avenue, Winnipeg MB R3C 3L7, Canada]	West Broadway	4edd1bb277c8274e005b3e73

Venue Data Dataframe



Venue Data - Map

3.4 Data Manipulation

3.4.1 Numerical Analysis

Next, the number of venues and venues categories were counted. The dataframe was also grouped by the "categories" column. The values were counted and sorted, and then the index was reset. Also, the dataframe was sorted by the "categories" and "name" columns, respectively.

3.4.2 Nearby / venues

"nearby venues" were defined using the original dataframe (in JSON format). Then, "json_normalize", a pandas function was used to flatten the data into a dataframe. Then, the dataframe was filtered, cleaned.

The resulting dataframe was further reviewed. The dataframe was grouped by the "categories" column, and the values were counted and sorted. Then, the index was reset, and the columns were rearranged.



3.4.3 Venue Distance

The distance from the venues to Holiday Inn was calculated using the Haversine Distance formula and inserted in the venue dataframe.

Next, a new column, "DistanceFromHolidayInn", was created and added to the dataframe. The dataframe was then sorted according to the 'categories', 'name', 'DistanceFromHolidayInn' and saved in a new variable. Then, the columns of the new dataframe was rearranged to make the "DistanceFromHolidayInn" column more visible.



3.5 Trending Venues & Venues of Interest

3.5.1 Trending Venues

A URL was created to determine in real-time if the venue data contains venues that have high foot traffic, that is, trending venues. A GET request with the URL was sent to Foursquare. A conditional statement was created, and the result was filtered using the "get_category_type". The dataframe and map would be displayed if available.

3.5.2 Venues (Places) of Personal Interest

Venues of personal interest are places that an individual desires to visit. A list that contains all the places of personal interest was created to confirm that the preferred venues exist among the venues supplied in the venue data. The list was then used to filter the available data along the category column using the "isin" function. The result is a dataframe that contains all the venues that match the categories in the list. Next, venues with similar categories of restaurants are isolated and visualised on a map.



Map of the restaurants relative to Holiday Inn

3.6 Venue Ratings and Venue Tips

A rating is a score between 0 and 10, inclusive that a user ranks a venue while tips are observations made while the user was at the venue or recommendations about a venue. Venues of the same category were selected for review.

3.6.1 Venue Rating

The "venue_id" was retrieved from the dataframe containing venues of personal interest. The URL for each venue was created. A request was made to Foursquare with the URL passed and JSON specified as the file format of choice. The venue rating for each venue was displayed and noted.



3.6.2 Venue Tips

The total number of tips for the venue was retrieved by passing the URL of the venue in the GET request made to Foursquare. A number greater than the number of venue tips was passed as the

"limit" parameter in the URL, along with the "venue_id". The result was delivered as a JSON file. The JSON file was flattened, and the tips were filtered, and a tip was displayed.

text	agreeCount	disagreeCount	id	user.firstName	user.lastName
0 Best customer service of any fast food restaurant in Winnipeg. (And better than many of the full-service restaurants too!)	0	0	4f39a37ae4b03d32dee3e2d7	Greg	Р

3.7 Clustering Venues

3.7.1 Group by Postal Code

The venues were grouped according to postal codes to determine the number of venues within each postal code. First, "groupby" was used to group the venues, then "count" method was applied, then the index of the dataframe was reset using "reset_index" function. The result was a dataframe that showed how many categories of venues existed per postal code.

3.7.2 One-Hot Encoding

Machine learning algorithms require numerical values for the optimal performance of the algorithm. "pandas.get_dummies" ("pd.get_dummies") method was used to one-hot encode the categorical values in the dataframe. One-hot encoding changed all the categorical values to either "0" or "1". Next, the "postal code" column that was removed during one-hot encoding was returned, and the new dataframe is displayed.

	Postal Code	Art Gallery	Coffee Shop	Gym	Juice Bar	Pizza Place	Restaurant	Theater
0	R3C 1V1	1	0	0	0	0	0	0
32	R3L 0C8	0	1	0	0	0	0	0

3.7.3 Mean of Values

The dataframe was grouped according to postal code using "groupby", the mean was then calculated using the "mean" function and the index was reset using the "reset_index" function.

A "for" loop was created to determine the postal code with the most common venues. A function to sort the top venues in descending order was defined. Then, the venues were defined as an array and sorted. Next, a new dataframe was created to hold the most common venues per postal code.

е	7th Most Common Ven	6th Most Common Venue	5th Most Common Venue	4th Most Common Venue	3rd Most Common Venue	2nd Most Common Venue	1st Most Common Venue	Postal Code	
er	Thea	Restaurant	Pizza Place	Juice Bar	Gym	Art Gallery	Coffee Shop	R3A 1G7	0
nt	Restaura	Pizza Place	Juice Bar	Gvm	Coffee Shop	Art Gallery	Theater	R3B 0P8	1

3.8 Creating Venue Clusters

Clusters of venues were created using K-means clustering. K-means clustering is a machine learning algorithm. It calculates the position of each venue relative to the others and groups the closest ones together. It also iteratively calculates these positions and continues to group and regroup the venues until no further regrouping can be done.

The number of clusters was set at five (5). The columns required for the analysis were specified and merged to the dataframe containing the clustered venues, using the "postal code" column as the point of reference.

The venue clusters are visualised on a map.



3.9 Examining the Clusters

Five (5) clusters were created, and each cluster was studied individually.

The categories of restaurants are checked alongside the venues that make up each category. The common venues were stated for ease of reference.

The family may visit the following common venues "Restaurant" "Art Gallery" "Coffee Shop" "Gym" "Juice Bar" in this cluster.

4.0 RESULTS

4.1 Basic Venue Information

The programme produced the following results:

4.1.1 Location Parameters of a Venue

The physical address, coordinates of the point of reference (in this instance, Holiday Inn) was provided by Foursquare.

Similar data for any other venue could be obtained if required. The specific parameters that were passed in the URL and the GET request made this possible.

The coordinates of Holiday Inn & Suites Winnipeg Downtown, 360, Colony Street, Colony, West End, Winnipeg, Manitoba, R3C 0E7, Canada are -97.15145373256972, 49.8911612.

4.1.2 Information about Venues around a Point of Reference

The data of venues around a specific point of reference was obtained.

	name	categories	address	crossStreet	lat	Ing	labeledLatLngs	distance	postalCode	cc	city	state	country	formattedAddress	neighborhood	id
0	Winnipeg Art Gallery	Art Gallery	300 Memorial Blvd	St. Mary Ave	49.889459	-97.150708	[{label: 'display', 'lat': 49.88945877797471	196	R3C 1V1	CA	Winnipeg	МВ	Canada	[300 Memorial Blvd (St. Mary Ave), Winnipeg MB	NaN	4b5f6471f964a520b4b729e3
1	Harvey's	Restaurant	600 Portage	NaN	49.888922	-97.156059	[{label: 'display', 'lat':	413	R3C 3L7	CA	Winnipeg	МВ	Canada	[600 Portage Avenue, Winnipeg MB R3C 3L7,	West	4edd1bb277c8274e005b3e73

Foursquare provides real-time data, so the results may differ whenever the data is refreshed. In this case, the parameters of Holiday Inn was passed in the URL. This provided the data for all the venues that were within the radius specified in the URL.

4.1.3 Coordinates of the Point of Reference (Holiday Inn)

The latitude and longitude of the point of reference were obtained.

49.8911612 -97.15145373256972

4.2 Venue ID and Distances

4.2.1 Obtain the Venue ID from Foursquare

A new dataframe containing the venues provided by Foursquare was obtained and manipulated to obtain the venue IDs for each venue. This venue ID was important to obtain more information from Foursquare about the venues in review.

na	ame	categories	address	crossStreet	id	lat	Ing	labeledLatLngs	distance	postalCode	cc	city st	ate count	ry formattedAddress	neighborhood
0 Winnipeg Art Gal	llery	Art Gallery	300 Memorial Blvd	St. Mary Ave	4b5f6471f964a520b4b729e3	49.889459	-97.150708	[{'label': 'display', 'lat': 49.88945877797471	196	R3C 1V1	CA Winn	ipeg	MB Cana	da [300 Memorial Blvd (St. Mary Ave), Winnipeg MB	NaN
1 Harv	ey's	Restaurant	600 Portage Avenue	NaN	4edd1bb277c8274e005b3e73	49.888922	-97.156059	[[label": 'display', 'lat": 49.88892237017493	413	R3C 3L7	CA Winn	ipeg	MB Cana	da [600 Portage Avenue, Winnipeg MB R3C 3L7, Canada]	West Broadway

4.2.2 Venue Distances

The distance between each venue and the point of reference was calculated. The coordinates of the point of reference and that of any other location are needed. This is useful for planning the itinerary or choosing venues to visit based on the distance to or from a particular location.

	name	categories	address	crossStreet	id	DistanceFromHolidayInn	lat	Ing	labeledLatLngs	distance	postalCode	cc	city state	country	formattedAddress	neighborhood
0	Winnipeg Art Gallery	Art Gallery	300 Memorial Blvd	St. Mary Ave	4b5f6471f964a520b4b729e3	0.20	49.889459	-97.150708	[{'label': 'display', 'lat': 49.88945877797471	196	R3C 1V1	CA Winni	peg ME	Canada	[300 Memorial Blvd (St. Mary Ave), Winnipeg MB	NaN
43	Asia City	Asian	519 Sargent	NaN	4da753106a2364c7a33bbb06	0.64	49.896464	-97.154857	[{'label': 'display', 'lat':	638	R3B 1W1	CA Winni	peg ME	Canada	[519 Sargent Ave, Winnipeg MB R3B	NaN

4.3 Trending Venues & Places of Interest

4.3.1 Trending Venues

The URL was defined, and a GET Request was made. At the time of making the request, there were no trending venues. This is because trending venues request prompts a real-time response from Foursquare.

4.3.2 Venues (Places) of Personal Interest

Preferred Venues were used to filter the venue data. This yielded a dataframe that eliminated all other venues and only highlighted venues that were of particular interest. This made it possible to search for different venue categories at the same time. The result can be used to decide if the location has interesting places that can be included in the itinerary for the vacation.



4.4 Similar Venues, Ratings and Tips

4.4.1 Details of Venues of Similar Categories

A specific category of venues (Restaurant) was reviewed, and the result was used as a basis for further review.

This implementation makes it possible to conduct focused research on a particular category of venues in a location.

	name	categories	address	cross Street	id DistanceFromHolidayInn	lat	Ing	labeledLatLngs	distance	postalCode	cc	city	state	country	formattedAddress	neighborhood
1	Harvey's	Restaurant	600 Portage Avenue	NaN	4edd1bb277c8274e005b3e73 0.41	49.888922	-97.156059	[['label': 'display', 'lat'. 49.88892237017493	413	R3C 3L7	CA	Winnipeg	МВ	Canada	[600 Portage Avenue, Winnipeg MB R3C 3L7, Canada]	West Broadway
61	Prairie 360	Restaurant	83 Garry Street	NaN	5203fd442fc63a7e19f4f4cb 1.16	49.887467	-97.136342	[['label': 'display', 'lat': 49.8874673755844	1159	R3C 0R3	CA	Winnipeg	МВ	Canada	[83 Garry Street, Winnipeg MB R3C 0R3, Canada]	NaN

4.4.2 Venue Ratings and Venue Tips

These are based on experiences, observations and recommendations of users that have visited a venue. It gives information on what the experience would be if the venue is visited. It also serves as a differentiator since many venues around the same area offer similar services. The ratings and tips give firsthand information on what the experience would be if the venue is visited.

Venue Rating

8.8

Venue Tip

text	agreeCount	disagreeCount	id	user.firstName	user.lastName
0 Best customer service of any fast food restaurant in Winnipeg. (And better than many of the full-service restaurants tool)	0	0	4f39a37ae4b03d32dee3e2d7	Greg	Р

4.5 Common Venues & Clusters Analysis

4.5.1 Most Common Venues

The most common venues were determined. This was achieved by using one-hot encoding (pd.get_dummies) to change all categorical values to numerical values. The data was grouped according to postal code. Then a "for" loop was created to determine the most common venues within each postal code. After that, K-means was employed to aggregate the venues into clusters. A new dataframe was created, which contained the most common venues and their corresponding postal codes.

postalCode			name	categories	lat	Ing	id	DistanceFromHolidayInn		
	0	R3C 1V1	Winnipeg Art Gallery	Art Gallery	49.889459	-97.150708	4b5f6471f964a520b4b729e3	0.20		
	32	R3L 0C8	Little Sister Coffee Maker	Coffee Shop	49.878993	-97.146772	5227ba6011d28632966c43a0	1.39		

This new dataframe (with most common venues) was merged with a previous dataframe. The previous dataframe contained other columns of interest (postal code, venue name, categories, latitude, longitude, id and venue distance). Both formed a comprehensive dataframe that contained values required to make an informed decision.



4.5.2 Examination of Venue Clusters

The venue clusters were segregated into individual dataframes and reviewed. Some clusters had more meal-related services, while some others had recreational facilities. Even though options may vary from one cluster to the other, the data in each cluster provide information that can directly feed into the itinerary or be used to modify the itinerary as required.

The family may visit the following common venues "Restaurant" "Art Gallery" "Coffee Shop" "Gym" "Juice Bar" in this cluster.

4.6 Visualisation of Venues on a Map

The following maps were displayed.

- All venues around the point of reference that are within a specified radius
- Only venues of a particular category that are within a specified radius
- Display of venue clusters relative to the point of reference

5.0 DISCUSSION

The following were observed.

5.1 Modifiable Parameter Values

The parameters could be adjusted as required, and the result would correspond to the new parameters used. For instance, if the radius of search is reduced or decreased, the results offered by Foursquare responds accordingly.



The above gives credence to the wealth of information regarding venues that Foursquare collects, maintains and provides. Secondly, the information is readily available and can be easily obtained given the appropriate parameter. Thirdly, this programme makes it easy to modify search parameters when exploring venues. Fourthly, the point of reference, in this case, was Holiday Inn, Winnipeg. If this is changed to any other location, the venue data that Foursquare would provide will differ.

Therefore, the programme is agile – the parameters are easy to modify and reuse as required. It will conveniently serve other purposes. The parameters can be altered to suit different locations, span different areas, or fine-tune to get more focused results.

5.2 Search Result Changes with Time

The result collated from Foursquare changes with time. For instance, the information garnered from Foursquare in the morning may differ from the information provided in the afternoon or evening. The same occurs when the venue data search occurs in different seasons.

Foursquare provides real-time data. Therefore, the information provided is subject to change because the information provided by Foursquare is based on the data available at a given point in time. For instance, if a venue only offers its services in the evening/nighttime, such a venue will not be supplied in the results provided by Foursquare if the search is conducted in the morning/afternoon.

Therefore, this programme can be used to get new venue data if the condition of the trip suddenly changes and the need arises. This makes the programme very useful for holidaymakers and tourists. Also, travel agents and their agencies would find it helpful in providing personalised travel experiences to their clients.

5.3 Ease of Changing Parameter Values

The venue ratings and tips give information that is both vivid and essential. This is critical information for making decisions about venues.

This information is independently supplied to Foursquare by users who have visited the venue. The data is current and provides information that is considered very useful to potential visitors.

Therefore, holidaymakers can take an informed decision about venues to visit by using the information supplied by visitors who previously visited the venue. The process of selecting a venue from among many has been reduced to a few clicks. The choice reaction time is significantly reduced. This is both beneficial to holidaymakers and travel agents alike. The turnaround times of travel agents can be maximised as more clients' requests can be attended to within a space of time.

The Scenario

Results from Programme written by Keisha for Jones Family

The categories of the preferred venues were given as the "options". Then, the programme filtered the desired venue categories from the venue data provided by Foursquare.

The filtered dataframe was segregated into categories, and similar venues were reviewed to obtain the venue ratings and tips used to determine the venues to be visited.

The resulting filtered dataframe was then sorted, and clusters of venues were created.

The clusters provided a top-level view of the itinerary. The decisions made based on the information provided by the venue ratings and tips were used to populate the itinerary on an hourly basis.

The runtime of the programme was less than two (2) minutes.

6.0 CONCLUSION AND SUMMARY

6.1 Conclusion

The programme

- streamlined the process of searching for available venues at a particular location and produced a list of preferred venues with acceptable attributes.
- increased the ease and speed of getting required venue information. It also reduced the turnaround times in creating a plan for visiting different holiday venues.
- made it easy to modify the itinerary for a vacation in the event of an unexpected occurrence.

6.2 Summary

The programme created an agile way of selecting venues to visit for vacation, which finds application in individual and group travels.

7.0 ACKNOWLEDGEMENT

Many thanks to Foursquare for making the data of venues easy to assess and utilise.

I give gratitude to IBM for the comprehensive training in Data Science and the making IBM Watson Studio accessible throughout this course.

Many thanks to the Python community for making the journey to learn Python programming language easier.