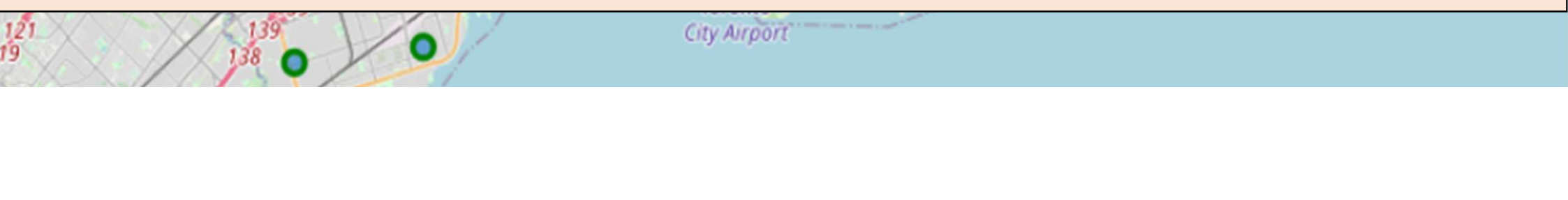


# BATTLE OF NEIGHBORHOODS

---

---

CAPSTONE PROJECT





# BATTLE OF NEIGHBORHOODS –Capstone Project



1.- INTRODUCTION

2.- METHODOLOGY

3.- THE QUESTION

4.- DATA: COLLECTION, REQUERIMENTS & UNDERSTANDING

5.- MODELING

6.- FINAL RESULTS & CONCLUSION



## BATTLE OF NEIGHBORHOODS – Introduction

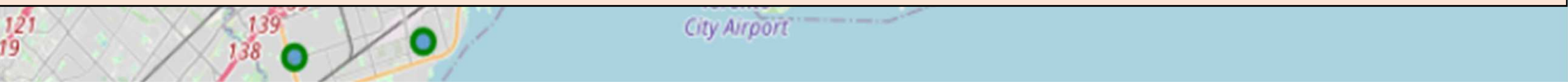
→ Making the decision to **move** is not easy.



→ Among the great changes that this decision implies, **why leave something as important as the new place of residence to chance?**



→ Why not take advantage of the great **benefits of machine learning** to reduce that impact on your life change?





# BATTLE OF NEIGHBORHOODS –Capstone Project



1.- INTRODUCTION

2.- METHODOLOGY

3.- THE QUESTION

4.- DATA: COLLECTION, REQUERIMENTS & UNDERSTANDING

5.- MODELING

6.- FINAL RESULTS & CONCLUSION

# BATTLE OF NEIGHBORHOODS – Methodology

→ Foundational Data Science Methodology (FDSM)

```
graph TD; BU[Business understanding] --> AA[Analytic approach]; AA --> DR[Data requirements]; DR <--> DC[Data collection]; DC <--> DU[Data understanding]; DU --> DP[Data preparation]; DP <--> M[Modeling]; M --> E[Evaluation]; E --> D[Deployment]; D --> F[Feedback]; F --> BU; F --> M;
```

The diagram illustrates the Foundational Data Science Methodology (FDSM) as a continuous cycle. It begins with 'Business understanding' (highlighted with a red border), which leads to 'Analytic approach'. This step informs 'Data requirements', which are linked to 'Data collection' and 'Data understanding' through bidirectional arrows. 'Data understanding' leads to 'Data preparation', which also has a bidirectional relationship with 'Modeling'. The process continues from 'Modeling' to 'Evaluation', then 'Deployment', and finally 'Feedback'. The 'Feedback' loop returns to both 'Business understanding' and 'Modeling', ensuring iterative improvement throughout the methodology.





# BATTLE OF NEIGHBORHOODS –Capstone Project

- 1.- INTRODUCTION
- 2.- METHODOLOGY
- 3.- THE QUESTION
- 4.- DATA: COLLECTION, REQUERIMENTS & UNDERSTANDING
- 5.- MODELING
- 6.- FINAL RESULTS & CONCLUSION







## BATTLE OF NEIGHBORHOODS – The Question

### → Business Understanding

- What are your basic needs?
- What strengths and opportunities should the neighborhood have?
- What can discard a zone do?





# BATTLE OF NEIGHBORHOODS –Capstone Project

- 1.- INTRODUCTION
- 2.- METHODOLOGY
- 3.- THE QUESTION
- 4.- DATA: COLLECTION, REQUERIMENTS & UNDERSTANDING
- 5.- MODELING
- 6.- FINAL RESULTS & CONCLUSION







## BATTLE OF NEIGHBORHOODS – Data: Collection, Requeriments & Understanding

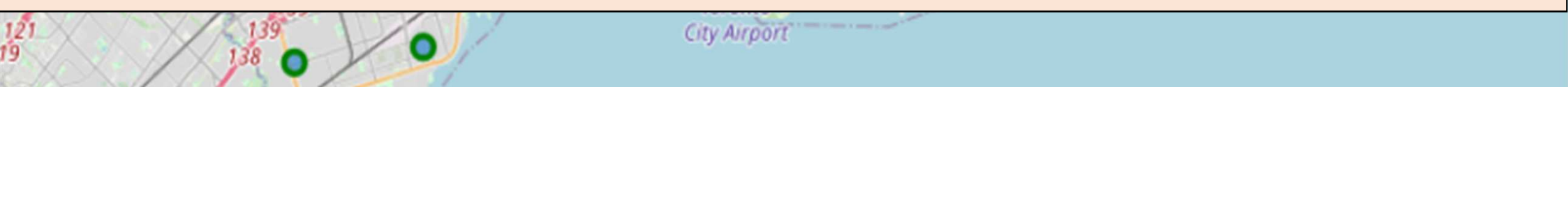
### → Data Sources

- Borough/PostalCode data Wikipedia → [https://en.wikipedia.org/wiki/List\\_of\\_postal\\_codes\\_of\\_Canada:\\_M](https://en.wikipedia.org/wiki/List_of_postal_codes_of_Canada:_M))
- Geospatial data for Toronto → [http://cocl.us/Geospatial\\_data](http://cocl.us/Geospatial_data))
- Foursquare API to obtain information about venues and facilities → <https://foursquare.com/>





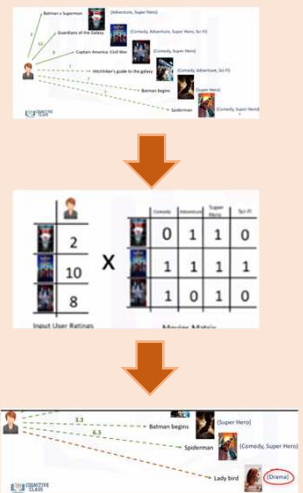
# BATTLE OF NEIGHBORHOODS –Capstone Project

- 1.- INTRODUCTION
  - 2.- METHODOLOGY
  - 3.- THE QUESTION
  - 4.- DATA: COLLECTION, REQUERIMENTS & UNDERSTANDING
  - 5.- MODELING
  - 6.- FINAL RESULTS & CONCLUSION
- 

# BATTLE OF NEIGHBORHOODS – Modeling

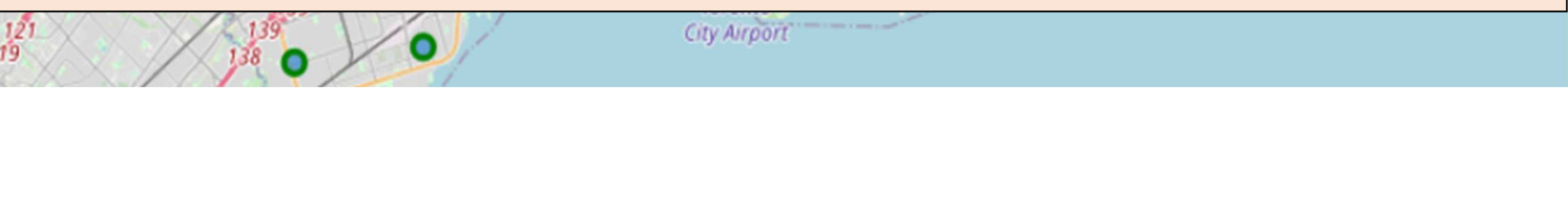
## → CONTENT-BASED RECOMMENDER SYSTEMS

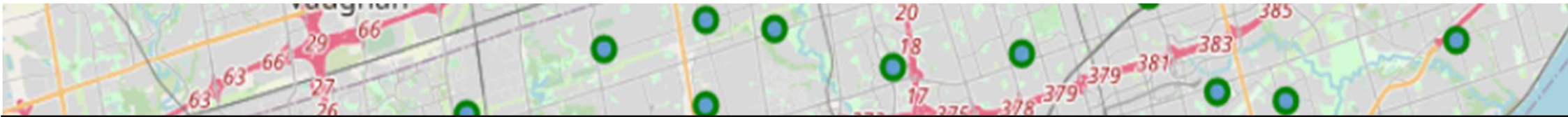
- A Content-based recommendation system tries to recommend items to users based on their profile.
- The user's profile revolves around that user's preferences and tastes.
- It is shaped based on user ratings, including the number of times that user has clicked on different items or perhaps even liked those items.





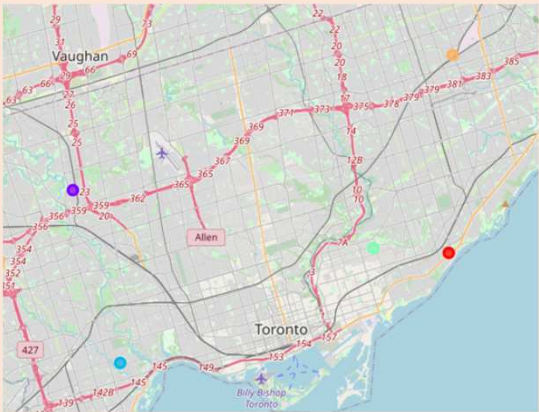
# BATTLE OF NEIGHBORHOODS –Capstone Project

- 1.- INTRODUCTION
  - 2.- METHODOLOGY
  - 3.- THE QUESTION
  - 4.- DATA: COLLECTION, REQUERIMENTS & UNDERSTANDING
  - 5.- MODELING
  - 6.- FINAL RESULTS & CONCLUSION
- 



# BATTLE OF NEIGHBORHOODS – Final Results and Conclusion

- The 2 best neighbourhoods for our first user were “North York Emery” and “Etobicoke, King’s Mill Park”.
- 2 areas had the same score and the difference amount the other 3 neighbourhoods is big.
- Probably our user choices was common, and hey include anything extraordinary as “Airport Food Court”.



	PostalCode	Borough	Neighborhood	Latitude	Longitude	Score
0	M9M	North York	Humberlea, Emery	43.724766	-79.532242	1.000000
1	M8Y	Etobicoke	Old Mill South, King's Mill Park, Sunnylea, Hu...	43.636258	-79.498509	1.000000
2	M4C	East York	Woodbine Heights	43.695344	-79.318389	0.285714
3	M1S	Scarborough	Agincourt	43.794200	-79.262029	0.250000
4	M1N	Scarborough	Birch Cliff, Cliffside West	43.692657	-79.264848	0.250000







## BATTLE OF NEIGHBORHOODS – Next Steps

- Train the model with more users.
- Introduce more parameter to make more selective the differences neighborhood.
- Adapt the algorithm to a hybrid model, taking the strong point of both systems (content-based and collaborative filtering)

