

A decorative graphic on the left side of the slide, consisting of a network of thin, light green lines and small circles, resembling a circuit board or a stylized tree structure, extending from the top left towards the bottom left.

CHOOSING A NEIGHBORHOOD FOR A NEW PIZZA PLACE IN MANAUS

IBM APPLIED DATA SCIENCE CAPSTONE

BY CARLOS VICENTE SOARES ARAUJO

BUSINESS PROBLEM

- Brazil is passing through a large financial crisis. As a lot of persons can't find a new job, they're deciding to create an own business. But, failure is not a option.
- In this project, I developed a model to predict the best neighborhoods from Manaus to put a new Pizza Place.

THE DATA

- The data I used were collected from three different sources:
 1. The neighborhoods names were scraped from Wikipedia;
 2. The latitude and longitude of these neighborhoods were collected using Geocoder library;
 3. And the venues on these neighborhoods were collected using the Foursquare API.

A STEP-BY-STEP METHODOLOGY

1. Make the libraries imports;
2. Scrap the neighborhoods names from Wikipedia;
3. Using *Geocoders* get the latitude and longitude of each of these neighborhoods;
4. Plot a map of Manaus with markers representing the neighborhoods;
5. Get the venues of each neighborhood using Foursquare API;

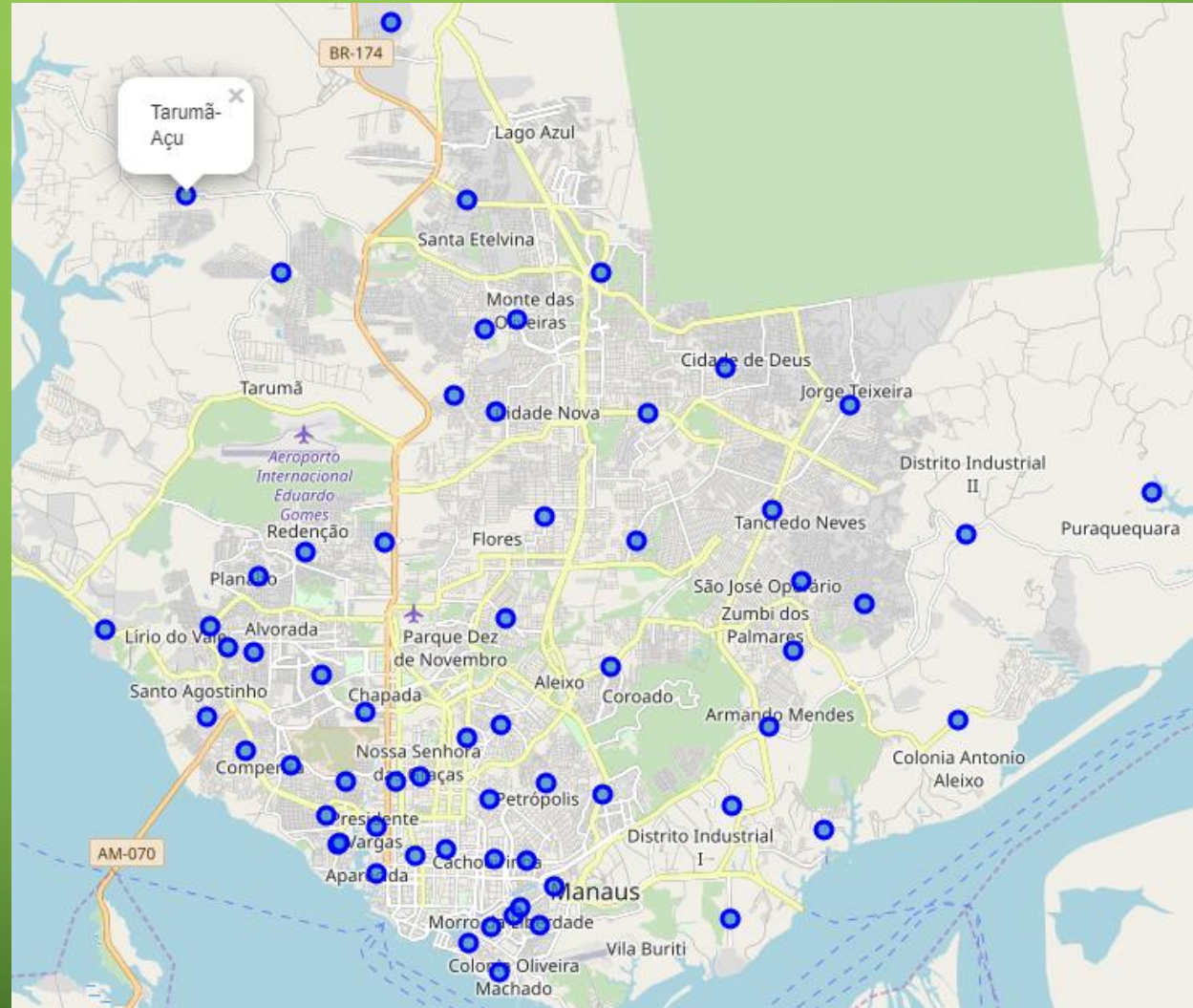
A STEP-BY-STEP METHODOLOGY (CONT)

6. Make an one-hot-encoding process on 'VenueCategory' column;
7. Run the K-Means algorithm with $K=5$;
8. Plot a map of Manaus with each neighborhood in a same cluster colored with the same color;
9. Print the neighborhoods in each cluster.

RESULTS

MAP OF MANAUS

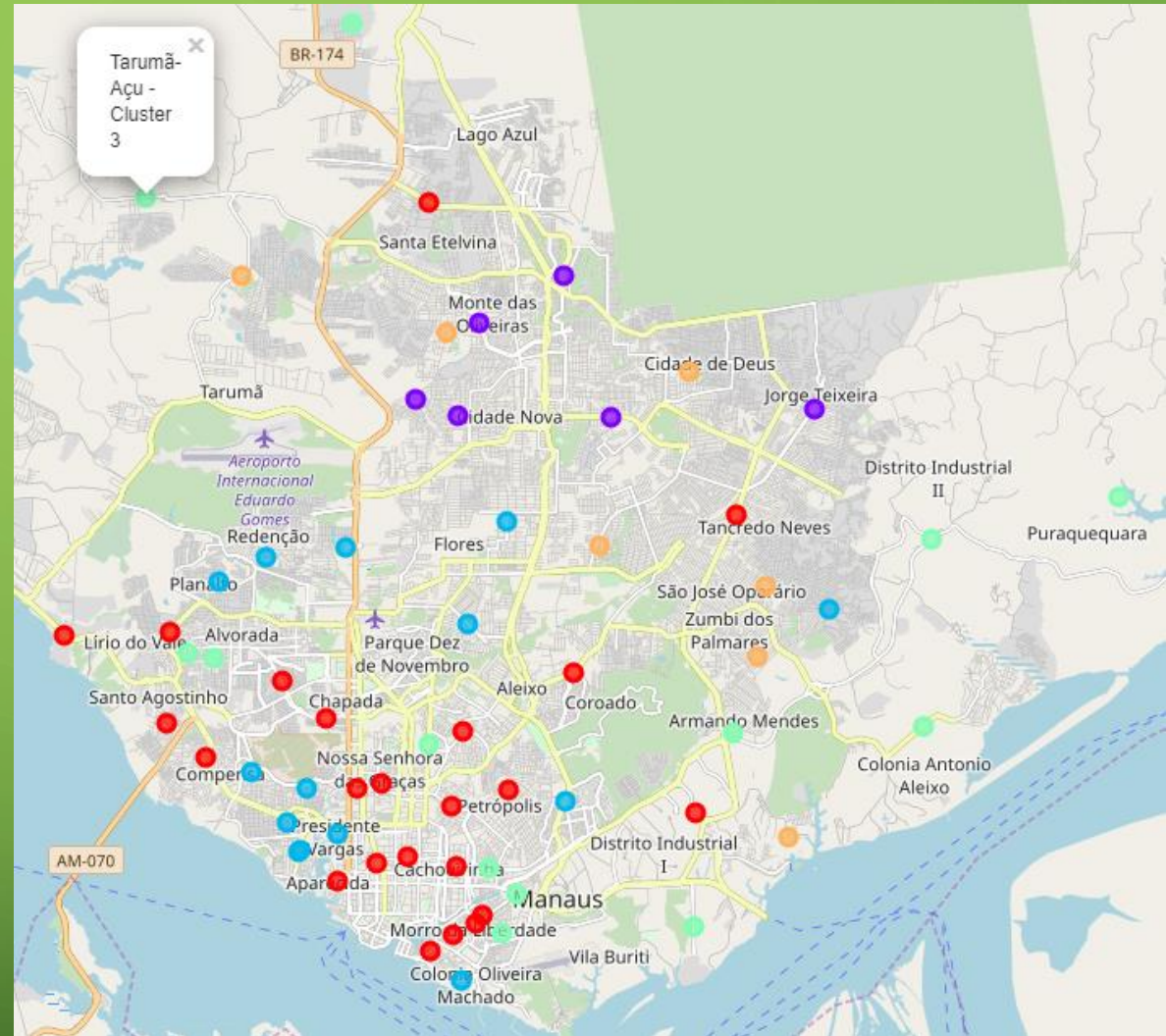
The markers represents the neighborhoods.



MAP OF MANAUS

The markers represents the neighborhoods.

Each neighborhood is colored based on its cluster.



NEIGHBORHOODS ON CLUSTER 0

| | Neighborhood | Pizza Place | Cluster Labels | Latitude | Longitude |
|----|--------------------------|-------------|----------------|----------|-----------|
| 31 | Morro da Liberdade | 0.030000 | 0 | -3.13664 | -59.99987 |
| 33 | Nossa Senhora das Graças | 0.030000 | 0 | -3.10729 | -60.01957 |
| 28 | Lírio do Vale | 0.020000 | 0 | -3.07531 | -60.06433 |
| 51 | São Francisco | 0.020000 | 0 | -3.11206 | -60.00492 |
| 49 | Santo Agostinho | 0.020000 | 0 | -3.09457 | -60.06487 |
| 48 | Santa Luzia | 0.020000 | 0 | -3.13906 | -60.00441 |
| 47 | Santa Etelvina | 0.028571 | 0 | -2.98487 | -60.00960 |
| 21 | Educandos | 0.020000 | 0 | -3.14273 | -60.00937 |
| 20 | Dom Pedro | 0.030000 | 0 | -3.08579 | -60.04079 |
| 52 | São Geraldo | 0.030000 | 0 | -3.10819 | -60.02475 |
| 18 | Distrito Industrial I | 0.031250 | 0 | -3.11344 | -59.95334 |
| 14 | Compensa | 0.030303 | 0 | -3.10161 | -60.05682 |
| 39 | Petrópolis | 0.030000 | 0 | -3.10865 | -59.99303 |
| 15 | Coroado | 0.026667 | 0 | -3.08387 | -59.97915 |
| 6 | Centro | 0.030000 | 0 | -3.12399 | -60.02063 |
| 1 | Aleixo | 0.030000 | 0 | -3.09622 | -60.00263 |
| 42 | Praça 14 de Janeiro | 0.030000 | 0 | -3.12252 | -60.01424 |
| 4 | Betânia | 0.020000 | 0 | -3.13487 | -59.99843 |
| 5 | Cachoeirinha | 0.020000 | 0 | -3.12461 | -60.00391 |
| 7 | Chapada | 0.030000 | 0 | -3.09341 | -60.03136 |
| 32 | Nossa Senhora Aparecida | 0.020000 | 0 | -3.12783 | -60.02900 |
| 57 | Tancredo Neves | 0.015873 | 0 | -3.05077 | -59.94501 |
| 41 | Ponta Negra | 0.017857 | 0 | -3.07591 | -60.08657 |

NEIGHBORHOODS ON CLUSTER 1

| | Neighborhood | Pizza Place | Cluster Labels | Latitude | Longitude |
|----|-----------------------|-------------|----------------|----------|-----------|
| 12 | Colônia Santo Antônio | 0.098765 | 1 | -3.02968 | -60.00359 |
| 37 | Novo Israel | 0.098039 | 1 | -3.02629 | -60.01250 |
| 34 | Nova Cidade | 0.096774 | 1 | -3.00026 | -59.98115 |
| 26 | Jorge Teixeira | 0.111111 | 1 | -3.02839 | -59.92836 |
| 30 | Monte das Oliveiras | 0.086957 | 1 | -3.01041 | -59.99912 |
| 8 | Cidade Nova | 0.091837 | 1 | -3.03013 | -59.97116 |

NEIGHBORHOODS ON CLUSTER 2

| | Neighborhood | Pizza Place | Cluster Labels | Latitude | Longitude |
|----|--------------------------|-------------|----------------|----------|-----------|
| 38 | Parque 10 de Novembro | 0.040000 | 2 | -3.07368 | -60.00146 |
| 43 | Presidente Vargas | 0.040000 | 2 | -3.11788 | -60.02900 |
| 40 | Planalto | 0.050000 | 2 | -3.06484 | -60.05415 |
| 22 | Flores | 0.034483 | 2 | -3.05193 | -59.99318 |
| 50 | Santo Antônio | 0.040000 | 2 | -3.11538 | -60.03959 |
| 25 | Japiim | 0.040000 | 2 | -3.11095 | -59.98095 |
| 24 | Glória | 0.040000 | 2 | -3.12127 | -60.03669 |
| 23 | Gilberto Mestrinho | 0.047619 | 2 | -3.07046 | -59.92513 |
| 17 | Da Paz | 0.040000 | 2 | -3.05772 | -60.02716 |
| 53 | São Jorge | 0.050000 | 2 | -3.10827 | -60.03556 |
| 11 | Colônia Oliveira Machado | 0.035294 | 2 | -3.14878 | -60.00275 |
| 56 | São Raimundo | 0.040000 | 2 | -3.12169 | -60.03723 |
| 61 | Vila da Prata | 0.050000 | 2 | -3.10480 | -60.04709 |
| 46 | Redenção | 0.040000 | 2 | -3.05946 | -60.04406 |

NEIGHBORHOODS ON CLUSTER 3

| | Neighborhood | Pizza Place | Cluster Labels | Latitude | Longitude |
|----|------------------------|-------------|----------------|----------|-----------|
| 55 | São Lázaro | 0.01 | 3 | -3.13889 | -59.99412 |
| 59 | Tarumã-Açu | 0.00 | 3 | -2.98379 | -60.06931 |
| 60 | Vila Buriti | 0.00 | 3 | -3.13729 | -59.95393 |
| 0 | Adrianópolis | 0.00 | 3 | -3.09907 | -60.00971 |
| 44 | Puraquequara | 0.00 | 3 | -3.04676 | -59.86425 |
| 35 | Nova Esperança | 0.01 | 3 | -3.07978 | -60.06053 |
| 27 | Lago Azul | 0.00 | 3 | -2.94707 | -60.02602 |
| 19 | Distrito Industrial II | 0.00 | 3 | -3.05580 | -59.90360 |
| 16 | Crespo | 0.01 | 3 | -3.13042 | -59.99135 |
| 10 | Colônia Antônio Aleixo | 0.00 | 3 | -3.09518 | -59.90549 |
| 3 | Armando Mendes | 0.00 | 3 | -3.09647 | -59.94543 |
| 2 | Alvorada | 0.01 | 3 | -3.08071 | -60.05521 |
| 45 | Raiz | 0.01 | 3 | -3.12525 | -59.99695 |

NEIGHBORHOODS ON CLUSTER 4

| | Neighborhood | Pizza Place | Cluster Labels | Latitude | Longitude |
|----|--------------------|-------------|----------------|----------|-----------|
| 36 | Novo Aleixo | 0.083333 | 4 | -3.05731 | -59.97373 |
| 29 | Mauazinho | 0.076923 | 4 | -3.11864 | -59.93371 |
| 54 | São José Operário | 0.068182 | 4 | -3.06567 | -59.93851 |
| 13 | Colônia Terra Nova | 0.083333 | 4 | -3.01232 | -60.00587 |
| 9 | Cidade de Deus | 0.062500 | 4 | -3.02062 | -59.95474 |
| 58 | Tarumã | 0.076923 | 4 | -3.00029 | -60.04924 |
| 62 | Zumbi dos Palmares | 0.071429 | 4 | -3.08040 | -59.94052 |

DISCUSSION

- As the neighborhoods on cluster 3 have a low number of Pizza Places or even don't have one they are the best places to install a new Pizza Place. Alvorada, for instance, is the sixth neighborhood with most domestic residences in the city;
- The neighborhoods of clusters 1 and 4 are the ones to be avoided, since these areas contain the large number of Pizza Places;
- Neighborhoods on cluster 0 and 2 could be also good options for a new Pizza Place if the new place has some differential over the others in the neighborhood.

CONCLUSION AND FUTURE WORK

- The model was successfully developed;
- The neighborhoods on cluster 3 are the best option to install a new Pizza Place in Manaus, Brazil;
- For future work, a good idea is to apply this methodology for different business types and cities.