- If you cannot open my notebook in github, please, try it here
- https://nbviewer.jupyter.org/github/anasonrisa/couursera_capstone_final_exam/blob/main/Attempt%20final%202%20-%20restaurant%20in%20NY.ipynb

- **Introduction**

Hello,

And thanks for reading this report.

I am Anna and I decided to look into the business question of opening the restaurant in New York. New York is the most populous city in the United States, with many various generations and nations living in, and thus is a great opportunity for a restaurant opening. As a restaurant opening is quite a wide area, I decided to limit it by:

Limiting it by focusing on Italian restaurants only (it is as well a personal touch as I am a fun of Italian cuisine). Italian population is around 2,5 mln., so quite a big chunk of total population.

Limiting the area I would explore Manhattan, which is a densely populated borough that's among the world's major commercial, financial and cultural centers. So, this is a sign we have a population with a good salaries, in particular, which is good for a restaurant opening.

- **Data**
- For the New York city a leveraged the data shared in one of the labs
- I will convert addresses into their equivalent latitude and longitude values.
- I will use the Foursquare API to explore neighborhoods in Dallas.
- I will use the explore function to get the most common venue categories in each neighborhood, and then use this feature to group the neighborhoods into clusters.
- I will use the *k*-means clustering algorithm to complete this task.
- Finally, I will use the Folium library to visualize the clusters and to make a choice of the best one to open an Italian restaurant.

- **Methodology**
  - First of all, I used the conversion of addresses into their equivalent latitude and longitude values.
  - Also, I used the Foursquare API to explore neighborhoods in New York City, in particular Manhattan.
  - I used the **explore** function to get the most common venue categories in each neighborhood, and then use this feature to group the neighborhoods into clusters.
  - K-means clustering algorithm was chosen to cluster the neighborhoods. to complete this task.
  - To visualize the results I used Folium library

- **Results**

- The result of my work, we can see that there are 5 clusters in Manhattan.
- 3 of them include many boroughs, one has just one and another one is a middle size.

- **Discussion**
  - Let us explore the clusters a bit. 1 and 2 has many restaurants inside and moreover quite a lot of Italian restaurants. So I do not see them as a great opportunity to start our business.
  - Cluster 3 is too small to consider to get a good revenue.
  - While cluster 4 and 5 seem to be the most interesting for me.
  - Cluster 4 does not have many restaurants / café in as for now – so we can be a “new category opener” and give people there an opportunity to try the Italian cuisine.
  - Cluster 5 is already developed as a restaurant / café center, and does not have many Italian restaurants , which gives us a great opportunity to take this place.

- **Conclusion**

This reports gave an overview of the project of finding the best place to open an Italian restaurant in New York, Manhattan. We looked at the data used, analysis types, as well as the results and observations.

030 726 201 201