



Data Analytics Boot Camp Project #1

**“Location
analysis of
sales agents in
national
territory”**

FEBRUARY 2021

Team 5

- Joyce Casiano
- Elena Sánchez
- Ignacio Morales
- Oscar Morales
- Eduardo Chacón
- Carlos Valverde

Agenda

- Background
- Analysis Objective
- Data Sources
- Processing
- Analysis
- Conclusions
- Learned Lessons
- Challenges

Background

The dental sector in Mexico, which generates around four billion dollars a year, will grow faster than the national average in 2020, (according to AMIC).

However, the reality in 2020 was different, mainly due to the global pandemic unleashed in the first quarter of that year.

Faced with such situation, companies dedicated to distributing dental materials such as consumables in dental clinics and offices, they have been looking for new ways to optimize their sales resources. For this reason, a company with a national presence has asked the team of analysts to prepare a study that includes an analysis of its sales force.

Analysis Objective

The company has made decisions to adjust its organizational structure in order to generate savings, however, it is concerned to know if the decisions made will greatly affect its operations once the emergency of the pandemics ends, for this reason, they have raised various questions to be answered and which are the starting point of the study and analysis.

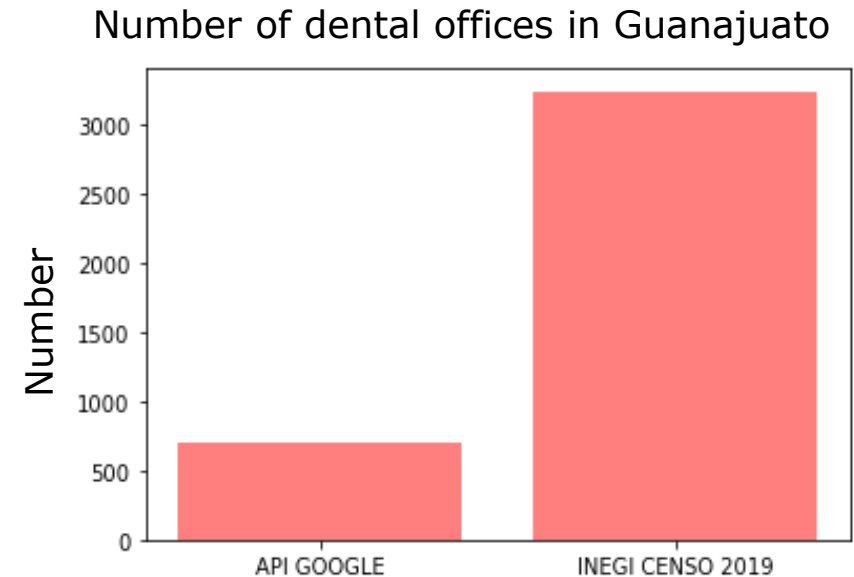
- **Where would be the optimal place to locate sales representatives in the national territory?**
- **Which opportunity areas could result from the analysis?**

The company has requested that a study and an analysis be carried out based on data and verifiable facts, they have requested that a study be presented to them with conclusions explaining the methodology used to reach them.

Data Sources

To obtain data on dental offices throughout the national territory, two data sources were explored, the Google Places site and the INEGI site.

After analyzing the two sources, the decision was made to continue the analysis with the INEGI data because it is more reliable, proof of this is the number of economic entities (dental offices) found with each sources and which is shown in the following graph.



INEGI data was also considered for other indicators such as socioeconomic level and population.

Processing

To process the data, a series of codes were developed to: obtain the data directly from its sources, review the quality of the data, clean it, process it and generate graphics that help us reach and support the conclusions that were generated.

Were used among others:

- Google API's
- Google Maps
- Jupyter Notebook
- Microsoft Excel
- Git Hub
- Python Pandas

```

1-CONOCER LA CANTIDAD DE DENTISTAS QUE SE ENCUENTRAN EN LA REPUBLICA MEXICANA CON INFORMACION MINIMA DE ESTADO

1.1 USAR API GOOGLE PARA CONOCER ESTA INFORMACION PARA ESTO SE DECIDIO HACER LA BUSQUEDA A NIVEL MUNICIPIO POR ESTADO

1.1.1 SE GENERA UN MUESTREO CON LA INFORMACION DE LOS MUNICIPIOS DEL ESTADO DE GUANAJUATO

In [ ]:

In [1]: # Create code to answer each of the following questions.
# Hint: You will need multiple target URLs and multiple API requests.

# Dependencies
import gmaps
import requests
import json
import pandas as pd
import os
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sns
# Google API Key
from config import gkey
gmaps.configure(api_key=gkey)

In [2]: #busqueda de la lat y long de los municipios
LISTAESTADOS=['Guanajuato,Abasolo','Guanajuato,Acámbaro','Guanajuato,Apaseo el Alto','Guanajuato,Apaseo el Grande','Guanajuato,Apaseo el Alto']
basedentista = pd.DataFrame(columns = ['DENTIST Name','DENTIST Address','LAT','LONG','ESTADO'])

for cp in LISTAESTADOS:
    try:
        params = {
            "address": cp,
            "key": gkey
        }

        base_url = "https://maps.googleapis.com/maps/api/geocode/json"

        my_geo = requests.get(base_url, params).json()

        lat = my_geo["results"][0]["geometry"]["location"]["lat"]
        lng = my_geo["results"][0]["geometry"]["location"]["lng"]

        #busqueda de los dentistas
        target_type = "dentist"
        radius = 30000

    # rewrite params dict

        params = {
            "location": f"{lat},{lng}",
            "types": "dentist",
            "radius": radius,
            "key": gkey
        }

    # Build URL using the Google Maps API
    base_url = "https://maps.googleapis.com/maps/api/place/nearbysearch/json"

```

Processing

Pre Clean Data Frame

	cve_entidad	desc_entidad	cve_municipio	desc_municipio	id_indicador	indicador	1950	1960	1970	1980	1990	1995	2000	2005
0	0	Estados Unidos Mexicanos	0	Estados Unidos Mexicanos	1003000001	Total de viviendas particulares habitadas	NaN	NaN	NaN	NaN	16183310.0	19403409	21942535	24706956.0
1	0	Estados Unidos Mexicanos	0	Estados Unidos Mexicanos	1003000011	Viviendas particulares habitadas	NaN	NaN	NaN	NaN	NaN	19361472	21512236	24006357.0
2	0	Estados Unidos Mexicanos	0	Estados Unidos Mexicanos	1003000015	Promedio de ocupantes en viviendas particulares...	NaN	NaN	NaN	NaN	NaN	4.7	4.4	4.2
3	0	Estados Unidos Mexicanos	0	Estados Unidos Mexicanos	1003000017	Viviendas particulares habitadas que disponen ...	NaN	NaN	NaN	NaN	NaN	18054384	20445525	23194511.0
4	0	Estados Unidos Mexicanos	0	Estados Unidos Mexicanos	1003000018	Viviendas particulares habitadas que disponen ...	NaN	NaN	NaN	NaN	NaN	16275682	18139843	21086978.0
5	0	Estados Unidos Mexicanos	0	Estados Unidos Mexicanos	1003000019	Viviendas particulares habitadas que	NaN	NaN	NaN	NaN	NaN	14471206	16800934	20825049.0

Processing

Clean Data Frame

	cve_entidad	desc_entidad	cve_municipio	desc_municipio	id_indicador	indicador	2020	unidad_medida	Index
12	0	Estados Unidos Mexicanos	0	Estados Unidos Mexicanos	3114006001	Porcentaje de viviendas con electricidad	99.001843	Porcentaje	10
13	0	Estados Unidos Mexicanos	0	Estados Unidos Mexicanos	6200031290	Porcentaje de viviendas particulares habitadas...	95.469330	Porcentaje	10
14	0	Estados Unidos Mexicanos	0	Estados Unidos Mexicanos	6200031291	Porcentaje de viviendas particulares habitadas...	98.070381	Porcentaje	10
34	0	Estados Unidos Mexicanos	0	Estados Unidos Mexicanos	6207019042	Porcentaje de viviendas particulares habitadas...	52.072835	Porcentaje	6
37	0	Estados Unidos Mexicanos	0	Estados Unidos Mexicanos	6207019046	Porcentaje de viviendas particulares habitadas...	37.559287	Porcentaje	4
40	0	Estados Unidos Mexicanos	0	Estados Unidos Mexicanos	6207019050	Porcentaje de viviendas particulares habitadas...	37.502030	Porcentaje	4
42	0	Estados Unidos Mexicanos	0	Estados Unidos Mexicanos	6207019055	Porcentaje de viviendas particulares habitadas...	87.538721	Porcentaje	9
43	0	Estados Unidos Mexicanos	0	Estados Unidos Mexicanos	6207019057	Porcentaje de viviendas particulares habitadas...	43.266919	Porcentaje	5
119	0	Estados Unidos Mexicanos	0	Estados Unidos Mexicanos	6207068394	Porcentaje de viviendas con agua entubada dent...	77.588935	Porcentaje	8
120	0	Estados Unidos Mexicanos	0	Estados Unidos Mexicanos	6207129631	Disponibilidad de servicios en la vivienda: Ti...	64.368058	Porcentaje	7

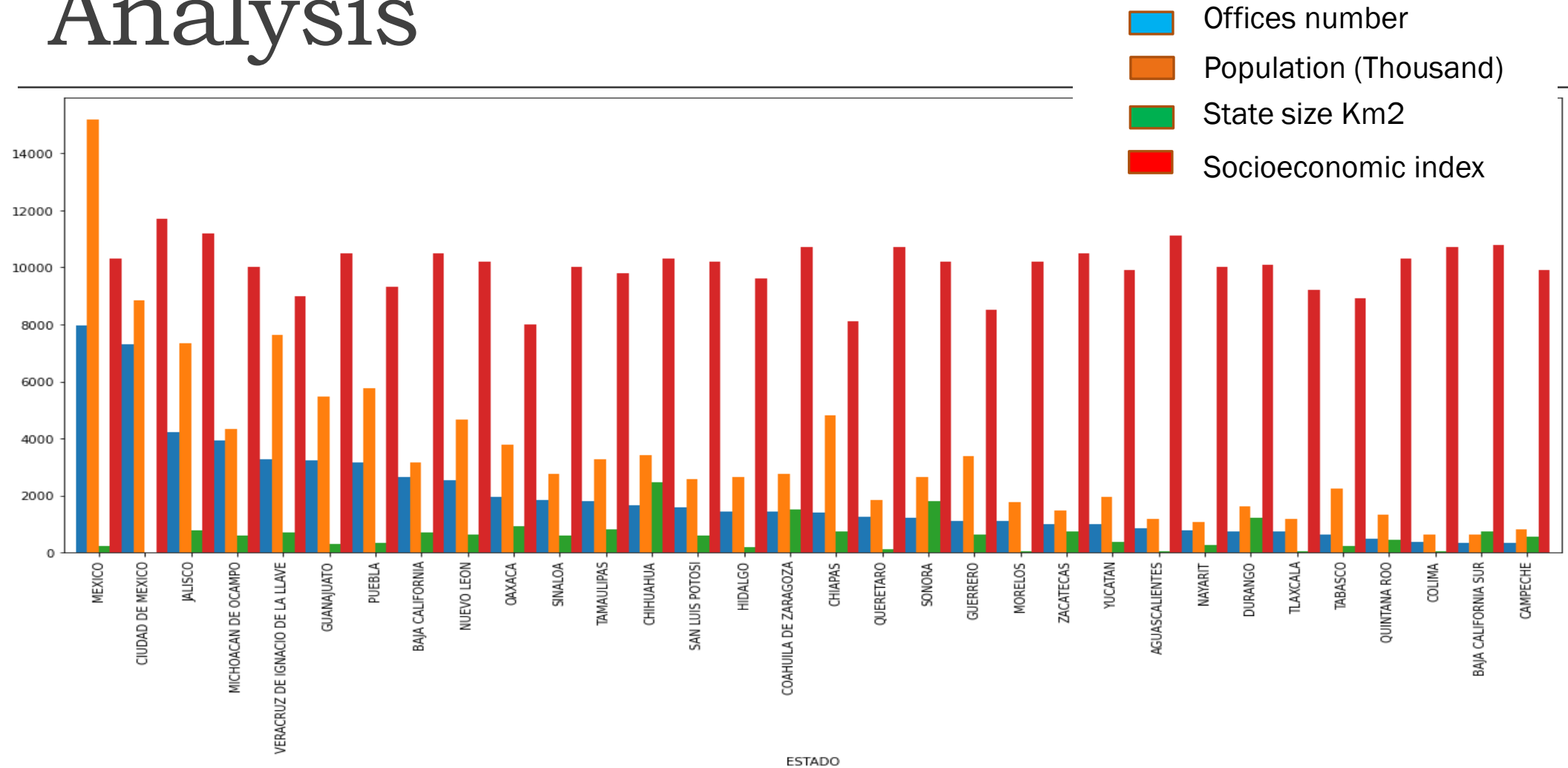
	State	Index_Value
0	Aguascalientes	111
1	Baja California	105
2	Baja California Sur	108
3	Campeche	99
4	Coahuila de Zaragoza	107
5	Colima	107
6	Chiapas	81
7	Chihuahua	103
8	Ciudad de México	117
9	Durango	101
10	Guanajuato	105
11	Guerrero	85
12	Hidalgo	96
13	Jalisco	112

Processing

Data Frames Merged

	ESTADO	NODENTISTAS	%SHARE	POBLACION	LAT	LONG	POBLACIONK	EXTENSIONKM	EXTENSIONKMK	POBLA/km2	DENTISTAS/km
0	MEXICO	7944	12.5343	15175862	19.4969	-99.7233	15175.862	22351	223.51	678.979106	2.8135
1	CIUDAD DE MEXICO	7318	11.5466	8851080	19.4326	-99.1332	8851.080	1495	14.95	5920.454849	0.20429
2	JALISCO	4222	6.66162	7350682	20.6595	-103.349	7350.682	78588	785.88	93.534407	18.613
3	MICHOACAN DE OCAMPO	3934	6.2072	4351037	19.5665	-101.707	4351.037	58599	585.99	74.251045	14.895
4	VERACRUZ DE IGNACIO DE LA LLAVE	3275	5.16741	7643194	19.2602	-96.5783	7643.194	71826	718.26	106.412636	21.931
5	GUANAJUATO	3234	5.10272	5486372	21.019	-101.257	5486.372	30608	306.08	179.246341	9.4644
6	PUEBLA	3171	5.00331	5779829	19.0414	-98.2063	5779.829	34306	343.06	168.478663	10.818
7	BAJA CALIFORNIA	2661	4.19862	3155070	30.8406	-115.284	3155.070	71450	714.50	44.157733	26.850
8	NUEVO LEON	2529	3.99034	4653458	25.5922	-99.9962	4653.458	64156	641.56	72.533481	25.368
9	OAXACA	1938	3.05784	3801962	17.0732	-96.7266	3801.962	93757	937.57	40.551234	48.378
10	SINALOA	1860	2.93477	2767761	25.1721	-107.48	2767.761	58200	582.00	47.556031	31.290
11	TAMAULIPAS	1794	2.83064	3268554	24.2669	-98.8363	3268.554	80249	802.49	40.730152	44.731
12	CHIHUAHUA	1649	2.60185	3406465	28.633	-106.069	3406.465	247455	2474.55	13.765998	150.06
13	SAN LUIS POTOSI	1582	2.49613	2585518	22.1565	-100.986	2585.518	61137	611.37	42.290561	38.645
14	HIDALGO	1439	2.2705	2665018	26.1004	-98.2631	2665.018	20813	208.13	128.045837	14.463
15	COAHUILA DE ZARAGOZA	1433	2.26104	2748391	27.0587	-101.707	2748.391	151562	1515.62	18.133774	105.76
16	CHIAPAS	1393	2.19792	4796580	16.7569	-93.1292	4796.580	73311	733.11	65.427835	52.628
17	QUERETARO	1269	2.00227	1827937	20.5888	-100.39	1827.937	11699	116.99	156.247286	9.2190
18	SONORA	1235	1.94863	2662480	29.2972	-110.331	2662.480	179355	1793.55	14.844749	145.22
19	GUERRERO	1119	1.7656	3388768	17.4392	-99.5451	3388.768	63596	635.96	53.285867	56.832
20	MORELOS	1107	1.74666	1777227	18.6813	-99.1013	1777.227	4879	48.79	364.260504	4.4074
21	ZACATECAS	1011	1.59519	1490668	22.7709	-102.583	1490.668	75284	752.84	19.800595	74.464
22	YUCATAN	997	1.5731	1955577	20.7099	-89.0943	1955.577	39524	395.24	49.478216	39.642
23	AGUASCALIENTES	839	1.3238	1184996	21.8853	-102.292	1184.996	5616	56.16	211.003561	6.6936
24	NAYARIT	774	1.22124	1084979	21.7514	-104.845	1084.979	27857	278.57	38.948164	35.99
25	DURANGO	748	1.18022	1632934	24.0277	-104.653	1632.934	123317	1233.17	13.241759	164.86
26	TLAXCALA	727	1.14709	1169936	19.3182	-98.2375	1169.936	4016	40.16	291.318725	5.5240
27	TABASCO	652	1.02875	2238603	17.8409	-92.6189	2238.603	24731	247.31	90.518095	37.93

Analysis



Analysis



The graph shows that there are areas that the dental offices are concentrated in the center of the country and occident.

Analysis



The graph shows that there are areas that the population is concentrated in the center of the country and geographically distributed in the north of the country. Important issue for sales representatives locations.

Analysis



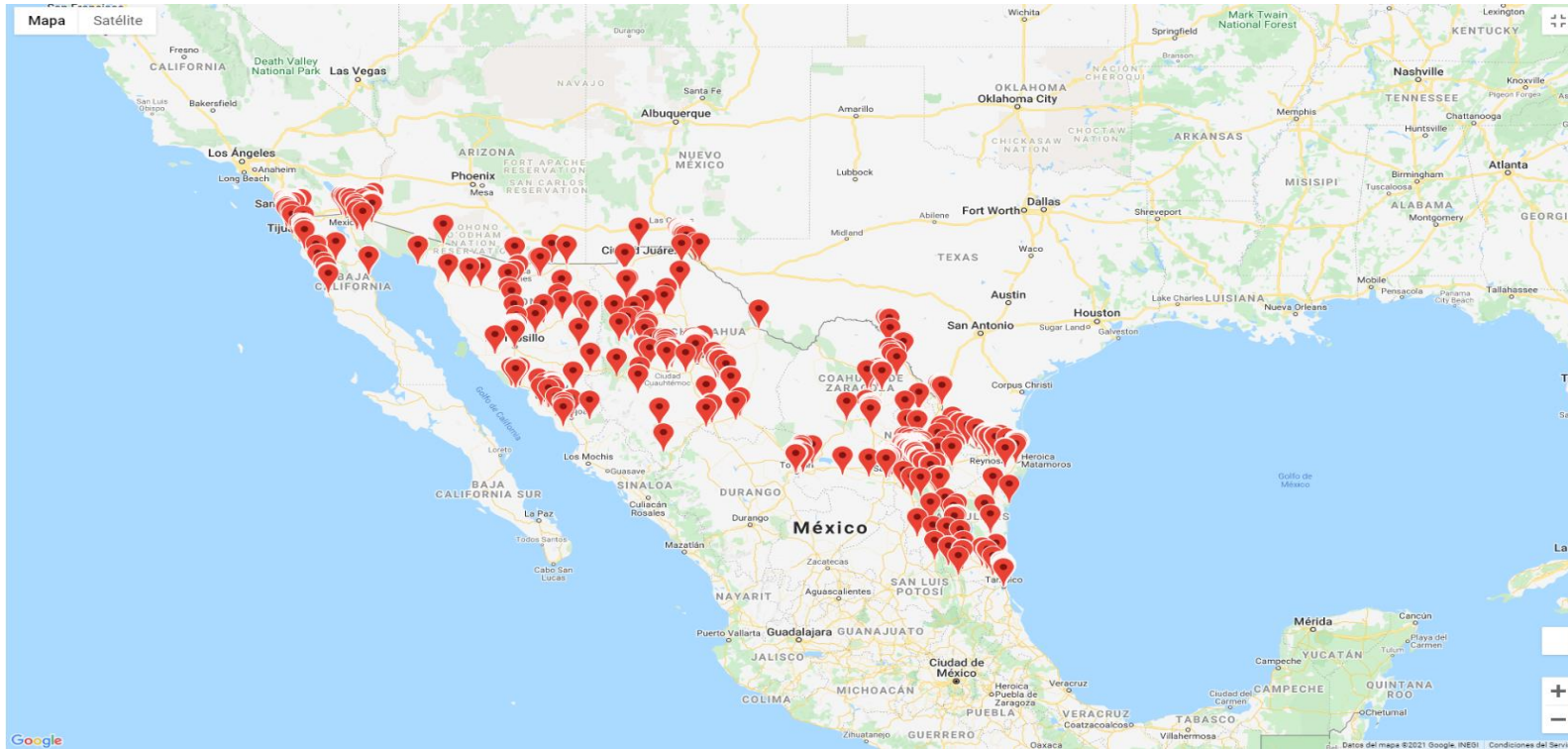
The graph shows that there are areas with more dental offices per Km2 are located in the center and south of the country (Tabasco y Chiapas).

Analysis



The graph shows that the biggest states are located in the north of the country, where the cities are more geographically distributed.

Analysis



The graph shows that the northern states of the country concentrate dental offices on the border, probably due to the effect called dental tourism, a factor to consider in the analysis.

Conclusions

- We came to the conclusion of locating sales representatives in:
 - Center of the country (CDMX and the State of Mexico)
 - Bajio (Queretaro, Guanajuato y Jalisco).
- We also consider that they should locate other sales representatives in the northern zone, specifically at least one in:
 - Northwest zone.
 - North-central zone.
 - Northeast zone.
- According to the map of density of dentists per km², we also recommend serving the market in the southeast of the country (Chiapas and Tabasco).
- According to the population map, it is also recommended to locate a sales representative in the Yucatan peninsula (Merida)

Challenges

- The conclusions we reached are based on a qualitative analysis, it would be interesting to be able to develop a quantitative analysis to optimize the location of the sales representatives.
- Obtain more information from the management of the company, regarding the income and costs incurred to carry out a better analysis.

Learned Lessons

- Data can be located and obtained in different ways, the easiest and simplest way is always recommended.
- It is important to select data from reliable sources, that way the conclusions will also be solid.
- Regardless of the data sources, it is highly recommended to review and clean them before starting with the processing and analysis.
- Teamwork is essential to achieve the objectives, a better result is achieved.