

# CARLOS A. HARO

Data scientist with experience in economic analysis

Main focus on data visualization and supervised learning.



## EDUCATION

2014  
|  
2018



**(ITAM) Instituto Tecnológico Autónomo de México**  
B.S., Economics

📍 Mexico City, Mexico



## EXPERIENCE

Jan. 2019  
|  
present



### Sr. Data Scientist

Mexico's Central Tax Administration Office  
(Servicio de Administración Tributaria)

📍 Mexico City

- Development and deployment of a supervised model for classifying tax debt | *R (tidyverse, ranger), Python (scikit-learn, pandas), SQL*
- Designed and taught three courses for the institution's staff training: introduction to R, basics of exploratory data analysis, building data science pipelines using Makefiles | *R (tidyverse, ggplot2), GNU Make*
- Designed a pipeline for automatic generation of frequent data visualization reports | *R Markdown (ggplot2, shiny)*
- Perform network analysis to detect tax evasion communities | *R (tidyverse, visnetwork, ggraph)*
- Version control of choropleth maps for the geographical display of taxpayer's data | *R (leaflet QGIS)*

2018



### Jr. Data Scientist | Economic Analyst

EnergeA (Energy Sector Consulting Firm)

📍 Mexico City

- Developed a statistical model for identifying anti-competitive practices between the mid-stream natural gas providers | *R*
- Neighboring gas station's competition analysis for identifying price setting mechanisms. | *R*
- Built a PDF scrapping pipeline of 100+ files for ownership analysis of Mexico's natural gas industry. | *R (stringr, selenium)*



## MISCELLANEOUS

2019



### Hackaton challenge winner

Annual BBVA Hackaton Challenge

Organizer: BBVA Bank

Challenge: Update and insert operation of a 50 million observations dataset in under 10 minutes.

Result: 95%+ accuracy of update and insert achieved in 4 minutes  
| *Pyspark on AWS for the algorithm, RMarkdown for the report*

2019



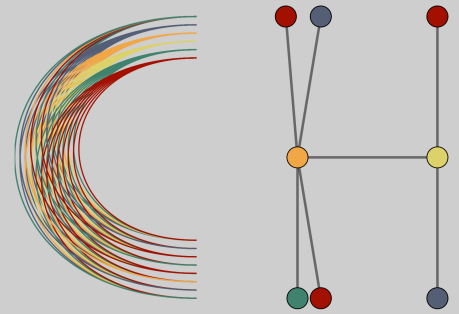
### Hackaton participant

Annual Banamex Hackaton Challenge

Organizer: Banamex Bank

Challenge: Create any platform for aiding small mexican businesses (< 75,000 usd/year cash flow) flourish.

Result: Created a live dashboard service that calculated the probability of succes of a business given initial investment, number of employees, employee salary, etc. | *R (shiny), Python (scikit-learn), AWS*



## CONTACT INFO

✉ haro\_ca@outlook.com

🐦 @haro\_ca\_

🌐 github.com/haro-ca

☎ (+52) 55-49-85-53-34

## SKILLS

Programming: R, Python, SQL

Data Science: Statistical modelling and data vizualization. Main experience with frequentist methods, but with high interest to develop in bayesian's.

This resume was made with the R package *pagedown*.

Last updated on 2020-02-03.