

# CARLOS A. HARO

Data scientist with experience in scalable data products.


Currently searching for a Mid-Level Data Science position, but with interest in a Jr. ML Engineering role.



## EXPERIENCE

Jan. 2019  
|  
present


### Sr. Data Scientist

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

- 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)*

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*
- Built a PDF scrapping pipeline of 100+ files for ownership analysis of Mexico's natural gas industry. | *R (stringr, selenium)*



## MISCELLANEOUS

2020

### Podcast host

Currently a host in "Cuestion de datos", a weekly data related podcast. ([Spotify link](#))

2020

### Teaching

Currently teaching a data analysis course (*pro bono*)

**Objective:** To teach political science, management, and economics students about modern computational techniques for data analysis.

**Syllabus:** ([Download link](#))

**Software:** R, Python, PostgreSQL

**Institution:** ITAM ([webpage](#))

**Language:** Spanish

**Amid COVID course videos:** ([YouTube link](#))

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*



## EDUCATION

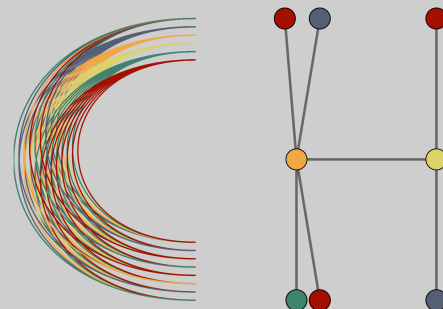
2014

|  
2018

(ITAM) Instituto Tecnológico Autónomo de México

B.S., Economics

 Mexico City, Mexico




## CONTACT INFO

 [haro\\_ca@outlook.com](mailto:haro_ca@outlook.com)

 [@haro\\_ca\\_](https://twitter.com/haro_ca_)

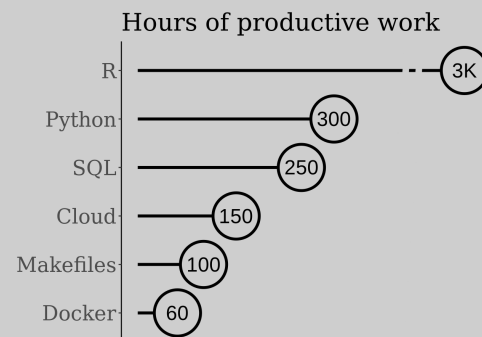
 [haro-ca](https://github.com/haro-ca)

 (+52) 55-49-85-53-34

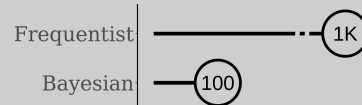
## SKILLS

*Outliers*, a book by Malcolm Gladwell, states that mastery over a skill can only be achieved by 10,000 hours of productive work. While there is both evidence in favor and against this theory, I do find it a good proxy for maturity of a skillset. The charts below attempt to summarize my current progress.

### Programming



### Statistics



Main experience with frequentist methods, but with high interest to develop in bayesian's.

### Miscellaneous

Vim enthusiast, in the Terminal, in VS Code, in RStudio, in Chrome (with Vimium), even in the whole OS (with Karabiner)!  
Chess fanatic ([chess.com profile](#))

*This resume was made with the R package [pagedown](#).*

*Last updated on 2020-04-28.*