

# Lidia Azucena Morales Vásquez

☎ 919-699-5592 | ✉ azucena.morales@duke.edu | 📱 AzucenaMV | 📺 AzucenaMV

## EDUCATION

### Duke University

Durham, NC

MS IN INTERDISCIPLINARY DATA SCIENCE, GPA: 3.91

May 2020

Relevant Coursework: Machine Learning, Deep Learning, Hierarchical Models, Cloud Computing

Scholarships: MIDS, Fulbright-García Robles, Becas Magdalena O. Vda. de Brockmann

### Universidad Nacional Autónoma de México (UNAM)

Mexico City, Mexico

BACHELOR'S IN ACTUARIAL SCIENCE, STUDY ABROAD AT UC BERKELEY

Jul 2015

CERTIFICATE PROGRAM IN DATA MINING

Fall 2018

Relevant Coursework: Bayesian Statistics, Probability, Programming in Java, Data Bases, SQL

Scholarships: International Mobility, UNAM

## WORK EXPERIENCE

### Intelimétrica

Mexico City, Mexico

DATA SCIENCE INTERN

Summer 2019

- Increased the profit margin of YALS, the housing price prediction platform of Intelimétrica, by improving the data cleaning pipeline using the Luigi library in Python via applying techniques such as anomaly detection and clustering.
- Developed and implemented machine learning models such as Quantile Regression, Random Forest, XGBoost and Neural Networks to improve the land price predictions for urbanized areas of Mexico, achieving a 50% decrease in the Median Error compared to the existing models
- Collected demographic data such as crime rate and geographical data from open street maps using the osmnx library in Python, to increase performance of the XGBoost Regression used for land price prediction.
- Trained data scientists on spatial statics and allowed Intelimétrica to take advantage of its geographical data and create more robust models.

### Banco de México

Mexico City, Mexico

INFORMATION ANALYST

Aug 2016 – Jul 2018

SOCIAL SERVICE (A REQUISITE IN THE MEXICAN EDUCATION SYSTEM)

Feb 2015 – Aug 2015

- Improved data quality used in financial risk analysis by identifying and reducing anomalies via statistical analysis and visual representations in R
- Assisted board members to evaluate Mexico's economy by calculating the payoff of future and forward contracts using R
- Developed an algorithm in R for matching transactions using string distance to detect discrepancies, resulting in a decrease in hours spend on related work as much as 20%

## PROJECTS

### Predicting House Prices in NYC

- Implemented a Random Forest model via scikit-learn to predict house prices in New York City, used feature engineering by combining transportation and location data to increase prediction accuracy

### Children's Mortality Rate Prediction for Hospitals

- Implemented a multilevel binomial logistic regression using brms library in R to predict children mortality rate for each hospital using the heart surgery complexity levels and volume
- Produced a tree star ranking system to rank hospitals using posterior predictions coming from a Bayesian hierarchical model

### Car Race Dashboard

- Created a R Shiny dashboard in order to help the race team identify the optimal pit stop strategy by integrating ggplot plots, highcharter interactive visualizations and a predictive model with real time data
- Dashboard available at: <https://azucenamv.shinyapps.io/dashboard/>

## SKILLS & LANGUAGES & INTERESTS

**Technical** Python (Scikit-learn, PyTorch, PySpark), R, Shiny R, SQL, H2O, SAS, LaTeX, HTML, Tableau, AWS, GCP

**Language** Spanish (Native), English (Fluent), Italian (Basic)

**Interests** Figure Drawing, Rock Climbing, Writing a Blog, Biking