DIEGO ROJAS

5561299166 | diegosreco@gmail.com
https://www.linkedin.com/in/diegorojas12/
Aguascalientes, México

Main Skills

- ETL process, Data Analysis, and Visualization
- Applied Statistics and Econometrics (forecasting, causal inference, and impact evaluation)
- Decision Analysis (deterministic and probabilistic cost-effective and cost-benefit analysis, microsimulation models)
- Machine learning (forecasting, classification, and NLP)
- Design, Implementation, and Analysis of Socioeconomic Experiment

Skilled in

- **R** (tidyverse tidymodels packages, plotly, leaflet , sf, NLP , bsts,   tseries, forecast, msm, lm4, svy, inference)
- **Python** (numpy, pandas, polars, statsmodels, scikit-learn, tensorflow, beautifulsoup, nltk, leaflet, seaborn, pyodbc)
- **SQL** (SQL Server, MySQL)

- **PySpark**
- **Tableau**
- **DataBricks**
- **Excel**
- **SAS**
- **Microsoft Power BI**

Work Experience

DATA ANALYTICS COORDINATOR | Confie Insurance | Tijuana, Mexico   since March 2023
- ETL pipeline development of the commission reconciliation process
- Construction of machine learning model for deposits categorization
- Development of business intelligence tools
- **Tools:** Python, R, Tableau, SQL Server.

INFORMATION ANALYSIS ASSISTANT MANAGER | Tax Administration Service (SAT) | CDMX, Mexico   May 2022-March 2023
- Extraction and analysis of big data
- Implementation and evaluation of time series and machine learning models for forecasting
- Evaluation of tax policies through microsimulation models and impact evaluation techniques
- **Tools:** Python, R, Azure Databricks (PySpark and SQL), and SAS (Enterprise Guide)

ANALYTICS EXECUTIVE | Porter Novelli Microsoft Latam account | CDMX, Mexico   September 2022-May 2022
- Data extraction and collection using web scraping techniques and ETL process.
- Application of natural language processing (NLP) techniques and machine learning models for text classification with the aim of identifying insights, and trends, and creating KPI's
- **Tools:** Python, R and Shiny

JR. CONSULTANT (PROJECT) | Digital Government Coordination of the Aguascalientes City Hall | Ags, Mexico June - August 2021
- Data analysis and evaluation of the program that collects citizen reports (Program 072)
- **Tools:** Python, R, Azure Databricks (PySpark and SQL), and SAS (Enterprise Guide)

JR. RESEARCHER (PROJECT) | Anti-corruption System of Aguascalientes State | Ags, Mexico   March-May 2020
- Design and analysis of an experimental survey where, through a treatment vs. a control, we measure the performance of social communication campaigns that encourage citizen participation.
- **Tools:** SurveyMonkey and R

RESEARCH ASSISTANT | Metropolitan Autonomous University (UAM)-Azcapotzalco | CDMX, Mexico   January 2018-April 2019
- Data analysis, implementation of microeconometrics models (casual inference), time series models and teaching of practical labs.
- **Tools:** Stata and R

Education

M.A. IN METHODS FOR THE ANALYSIS OF PUBLIC POLICIES (METPOL) | Center for Research and Teaching in Economics (CIDE) | June 2019 – September 2021
Thesis: "Short-run Effects of COVID-19 Pandemic on the Labor Market Dynamics in Mexico".

B.A. IN ECONOMICS |Metropolitan Autonomous University (UAM) | August 2014 – December 2018
Thesis: "Wage gaps between formal and informal workers

Courses

- **MACHINE LEARNING SPECIALIZATION** | Stanford University and Deeplearning.AI in Coursera| 2023
- **DATABRICKS FOR BIG DATA** | Databricks | 2022
- **MULTIVARIATE TIME SERIES IN R** | National Laboratory of Public Policies (LNPP) | 2021
- **BAYESIAN STATISTICS: FROM CONCEPT TO DATA ANALYSIS** | The University Of Santa Cruz California (Remote) 2021
- **CAREER TRACK DATA SCIENTIST WITH PYTHON** | Datacamp | 2021
- **CAREER TRACK DATA SCIENTIST WITH R** | Datacamp | 2022
- **MICROECONOMETRIC MODELING WITH CROSS-SECTIONAL AND PANEL DATA IN STATA** | UAM | 2018

Languages

SPANISH: NATIVE | ENGLISH: INTERMEDIATE | ITALIAN: BASIC