

# Humberto GONZÁLEZ RAMÍREZ

## Data scientist

+33 (0)6 04 52 39 76 23 rue de Flesselles, 69001 Lyon, France

humberto.gora@gmail.com https://humbertog.github.io linkedin.com/in/humbertogora

**Data scientist, polyvalent et curieux.** Doté d'une **pensée analytique structurée** et de compétences en **statistiques, machine learning et programmation**. Tout au long de ma carrière, j'ai développé des modèles probabilistes et statistiques pour l'**analyse de données** dans différents domaines, dont la **mobilité**, les **risques financiers** et l'**analyse du langage** des médias et politiciens. Je suis convaincu que la prise de décision et la politique doivent être fondées sur des **preuves objectives**.

## COMPÉTENCES DATA SCIENCE

## LANGUES

**Informatiques** R, Python (sklearn, pandas, numpy, ...), JAGS, Javascript, MySQL, MongoDB, Git.

**Modélisation** Modèles de régression, machine learning, discrete choice models, classification, clustering, natural language processing, LDA, inférence bayésienne.

**Traitement du signal** Séries temporelles, filtrage de Kalman, wavelets

**DataViz** ggplot2, matplotlib, seaborn, D3.js, R Shiny

**Espagnol** Langue maternelle

**Anglais** Courant

**Français** Courant

## EXPERIENCE

**Août 2021** Chercheur postdoctoral, **UNIV. GUSTAVE EIFFEL (ex IFSTTAR), UNIV. DE LYON, ENTPE, Lyon**  
**Sept. 2020** Participation au projet StationAir dont l'objectif était d'extraire des trajectoires de véhicules et le calcul d'indicateurs de trafic routier en temps réel à partir de vidéos enregistrées par des drones.

- Sélection et mise en place d'un algorithme pour estimer les trajectoires des véhicules.
- Calcul des indicateurs de trafic (débit, comptage et vitesse) à partir des trajectoires.
- Rédaction d'un rapport avec les principaux résultats.

Data science Transport Modèle probabiliste dynamique Filtre Kalman Python

**Présent** Co-fondateur et data scientist, **RADAR DE MEDIOS (NON LUCRATIF)**, <http://radardemedios.mx>  
**Mars 2019** Outil en ligne qui permet aux visiteurs d'analyser ce dont parlent les principaux médias au Mexique, ainsi que d'analyser le discours du président du pays. L'outil collecte et analyse les tweets publiés par les médias.

- Programmation des algorithmes pour collecter et stocker les tweets (plus de 5 millions).
- Modélisation sémantique de données textuelles pour extraire les thèmes abordés.
- Visualisation interactive.

Data science Natural Language Processing LDA DataViz Python

**Mai 2020** Chargé de recherche - Doctorat en transport, **UNIV. LYON, IFSTTAR, ENTPE, Lyon**  
**Nov. 2016** Réalisation d'une étude expérimentale sur le comportement des usagers dans les réseaux de transport. L'objectif était de comprendre et de prédire leurs choix d'itinéraire et leurs choix d'heure de départ au sein d'un réseau de trafic urbain pour fournir un simulateur de trafic avec des bases comportementales.

- Recherche documentaire, conception et organisation des expériences comportementales.
- Modélisation des choix des voyageurs : construction des variables, estimation et validation des modèles des choix discrets.
- Communication : rédaction des articles scientifiques et de thèse, présentations dans des conférences internationales et destinées au grand public.

Data science Transport Économie expérimentale Modèles de choix discrets Inférence bayésienne R

**Août 2016** Stage de master, **XEROX RESEARCH CENTER EUROPE, Grenoble**  
**Fév. 2016** J'ai modélisé, à l'aide d'un modèle "random forest", le temps passé par les passagers à changer de ligne de bus à partir des données billettiques de la ville de Nancy, ainsi que les caractéristiques géographiques et contextuelles des arrêts de bus.

Data science Transport Régression Random forests R

**Juillet 2014** Consultant en risques, **FINANCIAL SCIENCE, S.A. DE C.V., Mexico**  
**Mars 2011** Société spécialisée dans le conseil en gestion des risques.

- Conception d'un système pour calculer et rapporter les risques de crédit, marché et liquidité pour une banque spécialisée en crédits agricoles.
- Participation dans le pronostic des obligations de retraite d'une entreprise de télécommunications.
- Participation dans la modélisation des pertes probables pour compagnies d'assurances.

Data science Finance Banque Assurance R MySQL



## DIPLÔMES ET FORMATIONS

---

2016-2020	<b>Doctorat en transport</b> , Univ. Lyon, IFSTTAR, ENTPE.
2014-2016	<b>Master 2 en data science</b> , Universités Lyon 2, Paris 6 (UPMC), Politècnica de Catalunya et Polytech Nantes.
2013-2014	<b>Diplôme universitaire en séries temporelles (216 heures)</b> , Instituto Tecnológico Autónomo de México.
2004-2009	<b>Licence en mathématiques appliquées</b> , Instituto Tecnológico Autónomo de México.



## PUBLICATIONS

---

Scientifiques	Co-auteur de sept articles scientifiques (dont deux en tant qu'auteur principal) dans des revues internationales telles que PLOS One, Travel Behaviour and Society, Nature Scientific Reports et Royal Society. <a href="#">ORCID</a> <a href="#">Google Scholar</a>
Autres	Articles en espagnol pour le grand public dans le journal numérique mexicain <i>Animal Político</i> et la revue mensuelle <i>Nexos</i> . <a href="#">Nexos</a> <a href="#">Nexos:taller de datos</a> <a href="#">Animal Político</a>



## PROJETS

---

### PUNTO DECIMAL MX, CO-FONDATEUR, AUTEUR, ÉDITEUR

2016 - PRÉSENT

<http://puntodecimal.mx>

PuntoDecimal est un projet, non lucratif, qui invite les gens à écrire des articles sur leurs domaines d'expertise et destinés au grand public. Dans PuntoDecimal, nous cherchons à comprendre les problèmes actuels du Mexique d'un point de vue basé sur des preuves. Jusqu'à présent, plus de 40 auteurs ont publié plus de 100 articles sur des sujets divers, dont la politique, l'économie, la Covid-19, la mobilité et la migration.

[Blog](#)

[Vulgarisation](#)

[Analyse des données](#)