Andres Gomez-Lievano, PhD

9A Ware Street, Cambridge, MA, 02138 | 480-326-3587 | andres_gomez@hks.harvard.edu
GitHub: github.com/agomezlievano | LinkedIn: www.linkedin.com/in/andres-gomez-lievano-03776322/

SUMMARY

- Interdisciplinary scholar, excellent at communicating and translating knowledge across disciplines
- Successful track record in leading research projects, publishing papers and mentoring students
- Quantitative researcher interested in computational social and economic sciences
- Experience with large datasets such as census and social security microdata
- Expertise in systems thinking linked to mathematical, computational, statistical and network modeling
- Proficiency in predictive statistics, inference, dimensionality reduction, and recommender systems
- Currently seeking professional opportunities as a data or research scientist in industry

EXPERIENCE

Harvard University — **Sept 2014 - present**
Postdoctoral Fellow—Center for International Development

Colombian Government Pension Administrator (Colpensiones) — **Sept 2016 - Nov 2018**
External International Consultant

Arizona State University — **Aug 2011- Aug 2014**
Researcher, and Instructor—Mathematical, Computational & Modeling Sciences Center

Santa Fe Institute — **Jan 2011 - present**
Scholar, Alumnus, and Visitor

Grupo Empresarial Bolívar, S.A. — **Mar 2009 - July 2010**
Financial Risk Analyst—Financial Risk Management Division

EDUCATION

Ph.D.,	Applied Mathematics for the Life and Social Sciences, *Arizona State University*	**2011 - 2014**
M.Eng.,	Industrial Engineering, *Universidad de los Andes*	**2007 - 2008**
B.S.,	Physics, *Universidad de los Andes*	**2002 - 2007**

SKILLS

Domain Expertise:
I specialized on the economic development of cities, statistical physics applied to socioeconomic modeling of technological evolution, social networks analysis, and extreme value theory.

Mathematical & Statistical Expertise:
I am proficient in probability theory, econometrics, matrix algebra, stochastic processes, maximum likelihood inference, bayesian inference, and dimensionality reduction techniques.

Mentoring:
My passion for sharing knowledge has given me a reputation of being “the go-to person” in my lab. I have mentored one PhD student in Applied Math currently working in industry, one Master student currently pursuing a PhD in Innovation Studies, taught to more than a hundred students, and acted as Non-Resident Tutor for Harvard undergraduates at Mather House.

Computer:
I am fluent in Python, R, Stata, LaTeX, Microsoft Office. I am comfortable with: Mathematica, C/C++, Git, Linux command line.

Languages:
Fluent: Spanish and English; Basic: French and Italian.

Other Interests:
Non-fiction books, science-fiction fan, salsa dancing, guitar, painting, scuba diving.

DATA SCIENCE-RELATED PROJECTS

Colombian and Mexican "Atlases of Economic Complexity" (2014-2016):
Websites: http://www.datlascolombia.com/#/?locale=en-col
https://datos.gob.mx/complejidad/#/?locale=en-mex

- Project: Develop a free, online tool to visualize and understand the economy of cities.
- Challenges:
  - Raw data in large administrative social security text records across several years, representing more than 1.5 billion rows (one per person-month), and ~100 columns.
  - Inconsistent classification schemes, missing values, high-dimensionality.
- Result:
  - My contribution was to generate the cleaned and well-curated datasets supporting the online platforms, through a reliable workflow on a remote computing cluster, efficient database manipulation, well-organized folder structures and file-naming conventions, tailored multiple imputation techniques to fill-in missing values.
  - My expertise manipulating and analyzing such datasets resulted in:
    i. An external consulting project to the Colombian Government Pension Administrator (Colpensiones);
    ii. Three research papers: on wage inequality (1st of three authors), on testing predictions from extreme value theory about the increasing returns to scale in cities (1st of three authors), and on the drivers of formal urban employment in a developing economy with informality (3rd of four authors).

Colombian public system of pensions (2016-2018, joint work with renowned Colombian economist Eduardo Lora):
- Project: To assess the sustainability of the Colombian public system of pensions.
- Challenges: Lack of existing data-driven approaches to analyze a nation-wide system of pensions.
- Result:
  - I implemented a Variable Order Markov Chain model of the individual job trajectories and career progression of individuals, applying economic analysis to the projected scenarios to derive useful insights, generating appropriate visualizations of the results. Cross-validation procedures were applied in different stages of the project to measure the accuracy of the model for modeling real job trajectories.
  - The project resulted in:
    i. A presentation I gave to an audience of one hundred people from different agencies, both private and governmental, about the results of the project.
    ii. A final 160-page report to the government of Colombia detailing the main insights, the results, the methodologies, and the final recommendations.
    iii. An extension for a new project to study the financial flows between the public and private system.

Identifying economic diversification opportunities in cities and municipalities (2017-2018):
- Project: I led a three-person team for a big data approach to economic diversification in cities. Project commissioned by the Colombian government.
- Challenges:
  - Dealing with several stakeholders (Planning Department, Development Bank, Business Schools).
  - Defining the appropriate questions and metrics to measure success of the project.
- Result:
  - I framed the project as investigating whether future potential economic activities in cities or municipalities can be analyzed using the framework of collaborative filtering approaches in machine learning.
  - The project resulted in:
    i. Two presentations with governmental and academic stakeholders.
    ii. Two final reports, one on exports and the other on rural production, detailing insights, results, methodologies, and recommendations.
    iii. An interactive tool built in Tableau providing the recommendations, currently being used by the Development Bank of Colombia.

Migration and imitation in US economic development from 1850 to 1940:
- Project: This is a current project with the goal of analyzing the full set of US Census data from 1850 to 1940.
- Challenges:
  - We are a large team merging several large datasets (census+patents+historical lists of scientists)
- Result:
  - Implemented workflow in Odyssey, Harvard's largest cluster for high-performance computing.
  - I am using Natural Language Processing techniques to classify texts of data from the census into occupational categories (see https://github.com/agomezlievano/HopkinsKing2010),
  - I am mentoring my colleagues into using Neural Networks to both aid in the imputation of missing values and to use techniques from Matrix Factorization and Low-Rank Modeling to gain insights about the economic transformations in the US (https://github.com/agomezlievano/WorldTrade_CollabFiltering).
  - So far the project has resulted in:
    i. A series of novel questions about the US economy (e.g., the emergence of non-kin collaborations as the channel for the expansion of the division of labor).
    ii. Two research manuscripts in progress for future top academic journals.