

# AARON MAURER

610 · 717 · 8858 ◇ ajmaurer89@gmail.com

115A Park St ◇ San Francisco, CA 94110

---

## SUMMARY

I am an experienced machine learning data scientist and statistician on the search ranking team at Airbnb.

**Specialties:** Machine Learning; Production Data Pipelines; Online Experimentation; Nonparametric, Bayesian, and Classical Statistics; Large Scale Data Analysis.

---

## TECHNICAL SKILLS

|                             |  |
|-----------------------------|--|
| <b>Programing Languages</b> | Python, Scala, R, SQL (Hive/Presto/MySQL), Stata, SAS, Some Ruby                             |
| <b>Programing Tools</b>     | Apache Spark, Airflow, AWS, MapReduce, Hadoop, Jupyter, Git, L <sup>A</sup> T <sub>E</sub> X |

---

## PROFESSIONAL EXPERIENCE

|   |   |
|---|---|
| <b>Airbnb</b><br><i>Machine Learning Data Scientist, Search Ranking</i> | San Francisco, CA<br><i>August 2015 - Present</i> |
|---|---|

- Trained and implemented new versions of the GBM-based ranking algorithm. Contributed to 4.5% cumulative increase in bookings across ranking team.
  - ◇ Conceived, trained, and tested online a series of models with improved labeling of guest rejections. Experiment analysis and subsequent deep dives into model guided further development.
- Built and maintained Airflow/Hive/Scala/Spark data pipelines for online features and training data.
  - ◇ Built generalized anomaly detection framework into ETL processes which caught hundreds of feature issues and prevented search service degradation. Provided Slack alerting to identify issues.
- Oversaw search ranking experiments, guiding their design, methodology, and interpretation.
  - ◇ Implemented stratified estimators in company-wide, Ruby experimentation application and backing Hive data pipelines, reducing bias and cutting experimentation time by up to 20%.
  - ◇ Built Pyspark sequential testing pipeline to properly control false positive rate from experiments as actually run. Currently scaling process and evangelizing for broader use.
- Conducted ad hoc data analysis to understand current algorithm behavior and opportunities.
  - ◇ Performed retrospective of offline DCG metrics vs online performance, which pinpointed limits of our existing offline evaluation and sped up further model iteration.

|   |  |
|---|--|
| <b>Acumen, LLC</b><br><i>Policy Associate</i> | Burlingame, CA<br><i>August 2011 - August 2014</i> |
|---|--|

- Provided analytical and programing expertise to federal agencies studying healthcare topics, guiding \$25 billion in spending and monitoring safety of millions of vaccinations.
- Performed data analysis on and statistical modeling of entire enrolled Medicare population.
- Worked with federal clients to develop analysis and explicate results; produced data visualization and documentation to assist client understanding.

---

## EDUCATION

|  |                    |
|--|--------------------|
| <b>Masters of Science in Statistics</b><br>University of Chicago - Chicago, Illinois | <i>August 2015</i> |
|--|--------------------|

|   |                  |
|---|------------------|
| <b>Bachelor of Arts in Mathematics &amp; History, Cum Laude</b><br>Carleton College - Northfield, Minnesota | <i>June 2011</i> |
|---|------------------|