# Noah Rezsonya

San Jose, CA 95112 | Phone: 904-521-5938 | Email: kangnuoya@gmail.com
LinkedIn: NoahRezsonya | Github: NoahGithub

## Skills

| | |
|---|---|
| **Data Analytic Tools** | Python(pandas, numpy, scikit-learn, BeautifulSoup, nltk, SpaCy), Excel, R, SAS |
| **Database** | MySQL, Alteryx, Teradata |
| **Data Visualization** | Tableau, Python(matplotlib, seaborn), R(ggplot) |
| **Machine Learning Algorithms** | Linear Regression, Logistic Regression, Random Forest, XGBoost, k-Means |

## Experience

**Global Cost Analytics(Contract), Walmart International** ***San Bruno, CA 04/2019 - Present***

- Established a model to identify total savings potential for Mexico market to prioritize negotiation planning by leveraging item level data from 68 departments, 198 categories and 4400 suppliers using Alteryx, Python and Excel. Most recent potential saving amount identified is ~365,000 USD.
- Developed a commodity analysis model to track the commodity cost vs. Walmart cost trends across 12 categories and 36 suppliers through comparing monthly and yearly commodity cost against with Walmart Cost by using Excel.
- Generated scorecards for suppliers of liquor department to provide valuable arguments for Walmart Mexico buyers' cost lowering negotiation by leveraging 4-dimensional comparisons of 8 business metrics across 8 different business clusters by using Excel.
- Developed a framework for Walmart Mexico to track profitability and competitiveness of certain categories by comparing 8 business metrics, track the effectiveness of the cost lowering negotiation leveraging 8 business metrics of negotiated suppliers with non-negotiated suppliers under same categories through Alteryx and Python.

**Data Analytics Intern, Visual Threat** ***San Jose, CA 01/2019 - 04/2019***

- Developed vehicle ECU stability report of idling Buick SUVs across different model years and also within the same model year between 2011 and 2018. The finished report will be used to support company's future white paper publishing.
- Increased the amount of effective vehicle-generated data by at least 40% and saved 30% of the time spent on data collection by refining the data collection strategy and writing corresponding specs.
- Transformed approximately 15 GB collected data from 55 Buick SUVs into readable hex traffic data by writing Python script.
- Identified changes of CAN message from each CAN ID and reported on changing patterns by comparing processed data with their baseline, corresponding brand new SUVs data under idling status, with Meld on Linux.

**Data Science Intern, Ambit Analytics** ***San Francisco, CA 09/2018 -12/2018***

- Evaluated NLP model performance for speaker identification and identified factors related to overall model performance so to improve the performance.
- Labeled speakers from approximately 120 hours audio data which were used as the ground truth for the performance evaluation with Audacity.
- Tested model performance and visualized the difference between labeled data and prediction output with Python scikit-learn classification metric, Tableau and Python matplotlib.
- Found 3 major factors related to overall model performance including background noise level, voice similarity and random interruption during a conversation. Provided solutions to address each of issues.
- Increased the average F1 Score of the performance by 15.87% by implementing the proposed solution of random interruption.

## Education

**Southern Methodist University** ***Dallas, Texas***
Master of Science in Data Science - Specialized in Machine Learning and NLP - GPA 4.0 ***Class of 2019***

**University of New South Wales** ***Sydney, Australia***
Bachelor of Civil Engineering - Specialized in Transportation Engineering ***Class of 2014***