

Michael Kung

Austin, TX 78750 | Phone: 512-576-3359 | Email: m.kung889@gmail.com | [LinkedIn](#) | [GitHub](#) | [Portfolio](#)

Self-motivated Data Analyst/Scientist who possesses a “get-things-done” attitude and thrives under-pressure. Experienced in analyzing data in order to drive successful business solutions. Proficient in machine learning, statistics, and analytics. Known among co-workers and team members for stellar problem-solving and communication on large-scale projects.

SKILLS

- Data and Quantitative Analysis
- Machine Learning: Scikit-Learn, Keras, PySpark, Regressions, Classifications, Clustering, etc.
- Predictive Modeling
- Big Data Queries and Interpretation
- Data Mining and Visualization Tools
- Tableau
- Git
- Advanced Excel: Pivot Tables, VBA Scripting
- Statistics: Modeling, Forecasting, SPSS
- Programming: Python, Numpy, Pandas, Matplotlib
- API Interactions, Social Media Mining
- Databases: SQL, MongoDB
- Web Visualization: HTML, CSS, Bootstrap, Dashboarding, JavaScript, Leaflet, Plot.ly, d3

EDUCATION / LICENSES / CERTIFICATIONS

Data Analysis and Visualization Certificate: University of Texas at Austin McCombs School of Business, Austin, TX

Master of Science: Applied Cognition and Neuroscience: University of Texas at Dallas, Richardson, TX

Bachelor of Arts: Biology; Minor: Business: Texas A&M University, College Station, TX

Machine Learning | Stanford University, Coursera, Issue Date: June 2019, [Credential ID:](#) HSQDYHPE4W6J

PROJECTS

[Medical Insurance Charge Regression Modeling](#) | [GitHub](#): In this notebook, I performed extensive exploratory data analysis and explore several types of Regression models (Multi Linear Regression, Polynomial Linear Regression, Decision Tree Regression, Random Forest Regression) on a Medical Cost dataset found on Kaggle to predict if medical costs based on several features.

- **Responsibilities:** Performed all exploratory data analysis, train/test all Regression Models, compare all models.
- **Tools / languages used:** Python, Scikit-Learn, LabelEncoder, OneHotEncoder, numpy, pandas, matplotlib, seaborn.

[Cinematrix](#) | [GitHub](#): A movie recommender app using a machine learning algorithm, matrix factorization, to recommend 5 movies a user has not watch. These recommendations are based on a user's previous movie ratings and other user ratings.

- **Responsibilities:** Tested multiple machine learning models. Utilized matrix factorization to make movie recommendations. Created Python script to show top 5 recommended movies. Generated the flask app to display the recommended movies, user data, and the web scrapped movie information. Hosted the app on to Heroku.

- **Tools / languages used:** Python (Pandas, NumPy, Flask, JSON), Web Scraping, HTML/CSS, Bootstrap, Tableau

[US Demographic Interactive Visualization](#) | [GitHub](#): An interactive visualization of state population density and city demographic. App uses multiple visualizations and popups to summarize data gathered from the US census API.

- **Responsibilities:** Data cleaning and upload to database. Created the HTML dashboard. Bound JavaScript popup with city demographic data to Leaflet dots. Flask app routes.

- **Tools / languages used:** Python (Flask), JavaScript (Plot.ly, d3, Leaflet), SQLite, HTML/CSS, Bootstrap

[Belly Button Diversity](#) | [GitHub](#): An interactive dashboard visualizing belly button diversity data from Rob Dunn Lab.

(<http://robdunnlab.com/projects/belly-button-biodiversity/>). JavaScript is used to populate the dropdown options so user can select which sample ID to visualize. Plotly.js is used to create pie and bubble chart.

- **Responsibilities:** Created Flask App and dashboard. Wrote JavaScript script. Hosted the app on to Heroku.

- **Tools / languages used:** Python (Flask), JavaScript (Plot.ly, d3), jQuery, HTML

EXPERIENCE

System Administrator + Office Manager | Fidelity Dental P.A. | May 2017 - Currently

- Spearhead new patient acquisitions. Increased number of monthly new patients by 400% through data analysis and marketing strategies.
- Develop and structure a MySQL database to collect and analyze practice data such as production/collection of each dentist and patient demographics.
- Generate data visualizations to report practice performance, financial analysis and profit/loss reporting.
- Install/maintain/configure practice management software and computers.

Graduate Student Researcher | University of Texas at Dallas | Jan. 2016 – May 2017

- Utilized SPSS to run statistical analysis to analyze the relationship between neurochemistry and behavioral data.
- Conducted and created visualizations for data presentations and data analysis using advanced excel.