# DANIEL W. CHEN

MASTER CANDIDATE IN DATA SCIENCE

## PROFILE

**Phone**
(206) 499 6068

**Email**
daniel@ischool.berkeley.edu

**Github**
github.com/dandanbang

**Linkedin**
/in/wdanielche

### Analytics Projects

**San Francisco Crime Classification**
Used XGB Classifier to predict crime probability using 12 years of crime data. Placed as the 55th out of 900 competitors on Kaggle.

**Tweets Clustering**
Used iPthon Notebook, Scikit-learn, Postgres, panda, and nltk to cluster (k-means, DBScan) a million tweets and conduct text analysis in San Francisco

**Yahoo Answers Classifier**
Used NLTK, Pandas, and Scikit-learn to train and build machine learning classifiers (SVM, Naïve Bayes) and predict question categories with accuracy of 57%

**Amazon Customer Lifetime Value Analysis**
Used 3 years of Amazon's transaction data to conduct a cohort analysis that calculates the lifetime value of an acquired MP3 customer

**Emojineering**
Used NLTK, Python, Scikit-learn to predict Emojis given a tweet. Built a powerful Word2Vec similarity model for Emojis. In progress of publishing the results with Professor Marti Hearst.

## EDUCATION

| | |
|---|---|
| 2015 - 2017 | **UC Berkeley – School of Information**<br>Master of Information Management and Systems, GPA 3.8 |
| 2009 – 2014 | **University of Washington – Foster School of Business**<br>Finance & Information Systems, GPA 3.85 |

**Relevant Course:** Python Data Analysis, Data Mining and Analytics, Natural Language Processing, Data Structure, Scalable Spatial Analytics, Programming in Java, Distributed Computing and Infrastructure, Information Visualization

## TECHNICAL SKILLS

Language: Python, SQL, R, Java, HTML, CSS, Javascript, D3, MapReduce, Spark
Toolkit: Scikit-learn, pandas, nltk, numpy, scipy, matplotlib, iPython, Git
Applications: Postgres, Mongodb, Tableau, Excel, Visual Studio, ERP AX 2012

## WORK EXPERIENCE

| | | |
|---|---|---|
| 07/14 – 06/15 | **Sonata Software**<br>*Technical Consultant at Microsoft*<br>Designed relational database tables for customized features in a complex Enterprise Software from drawing ERD to normalization<br><br>Led a team of 3 developers to implement an ERP project from requirement gathering, development, and user testing | Redmond |
| 07/14 –09/15 | **Loocup (www.loocupapp.com)**<br>*Co-founder*<br>Founded the start-up to launch an app from scratch on the Appstore with 2 developers using Javascript, Parse, Node JS, and Mongodb, building both front-end and back-end with 3 people<br><br>Designed and developed the website with HTML, CSS, Javascript, and the UI/UX of more than 50 phonescreens | Seattle |
| 04/13 – 09/13 | **Amazon.com**<br>*Financial Data Analyst*<br>Conducted a data project by analysing 3 years of transaction data to calculate customer lifetime value using cohort analysis with SQL, Cognos, and Excel<br><br>Extracted, analysed, and created global financial data and metrics with SQL and Excel for daily business needs and delivered weekly business review for Amazon Kindle and Amazon MP3 teams | Seattle |
| 06/12 – 09/12 | **Deloitte & Touche**<br>*Information Management Consultant*<br>Built statistical models using Python and VBA Macro to optimize information control and increase data transparency for Bank of China<br><br>Improved Bank of China's internal data management in the relational database by eliminating redundancy in table design and by spotting the bottleneck in daily BI operation | Taipei |

## LEADERSHIP AND ACTIVITIES

| | |
|---|---|
| 09/15– Present | **Information Management Association** | *Financial Chair* |
| 09/09 – 06/13 | **Taiwanese Student Association** | *President* |