

MATTHEW CAPPER

Data Scientist

I am deeply enthralled by what the data we produce through life can say about us as people. Data Science grants the immeasurable power to glean deep insights about what has happened in our world and what may be done about it. My appetite for exploration compounds with each model I create and I relish the opportunity to create a more ethical and functional world through my balanced and methodical approach to the field.

 in/matthew-capper/  /matthewcapper  210-867-1894  matthew.capper01@gmail.com

TECHNICAL SKILLS

Applied Statistics - SQL - Python - pandas -
Matplotlib - Machine Learning - Natural Language
Processing - Distributed Data - Data Storytelling -
Git - Jupyter Notebooks - Anaconda

EDUCATION

CODEUP JUNE 2019

Fully-immersive, project-based 18-week Data Science career accelerator that provides students with 600+ hours of expert instruction in applied data science. Students develop expertise across the full data science pipeline (planning, acquisition, preparation, exploration, modeling, delivery), and become comfortable working with real, messy data to deliver actionable insights to diverse stakeholders.

UNIVERSITY OF TEXAS SAN ANTONIO DECEMBER 2015

Bachelor of Science in Mathematics

PROFESSIONAL EXPERIENCE

STARBUCKS COFFEE DECEMBER 2007 - JUNE 2019 SHIFT SUPERVISOR

As a supervisor with Starbucks, I maintained store operations as a manager-on-duty and effectively coached and guided partners to be strong team members with knowledge, warmth, and sincerity. This position is multi-faceted and included managing store funds and product orders in addition to customer-facing hospitality roles and facility maintenance.

PROJECTS

RECIDIVISM IN DOMESTIC VIOLENCE JUNE 2019

Using data from the Chicago Women's Health Risk Study, a model was created to predict probability of repeat victimization in domestic abuse cases. Presentation of the model was implemented into a web interface to assess based on various background factors.

NATURAL LANGUAGE PROCESSING MAY 2019

Scraped readme data from Github repositories and analyzed the usage of words in said repositories to create a predictive model regarding what primary code language the repository was associated with.

TIME-SERIES ANALYSIS APRIL 2019

Compiled scattered data from a fitbit device and explored data with seasonality to determine type of subject and predict future activity

CLUSTERING APRIL 2019

Discovered data clusters in a Zillow data set in order to construct a more robust linear model for prediction of error on property value estimation

REGRESSION MARCH 2019

Cleaned and prepared listings from a Zillow data set and created linear models to estimate average predicted error based on property values' difference from mean.