IDRIS RASHEED

Palo Alto, CA | 650-283-2488
idris.s.rasheed@gmail.com
linkedin.com/in/idris-rasheed-stats
github.com/idrissrasheed

I am a Data Analyst using statistics, data science, and machine learning to build tools, visualizations, and analyses that enable interdisciplinary teams to create life-saving advances in medical research.

EDUCATION

University of California, Santa Barbara, August 2017

Bachelor of Arts in Statistics

SKILLS

Software: R, SQL, Google Cloud Platform (GCP), Microsoft Office

Technical: Data Analysis, Data Quality, Data Wrangling, Data Management, Data Mining, Database Management, Machine Learning

EXPERIENCE

CDC Foundation (Contract), Remote — *Data Analyst*

NOVEMBER 2021 - PRESENT

- Analyze gaps and opportunities in COVID vaccine supply chains by building ggplot2 data visualizations to illustrate data on the number of expired doses and the amount of COVID vaccines available for order by county in California
- Prepare written report on gaps and opportunities in COVID vaccine supply chains for the California Department of Public Health's Vaccine Taskforce Team
- Automate code with the tidyverse packages in R to more efficiently isolate data on the number of COVID doses administered to various demographics within schools and childcare centers
- Prepare written report on COVID doses administered to various demographics within local health jurisdictions, schools, and childcare centers for the California Governor's Office
- Design statistical reports and mail merges to streamline CDC updates to vaccine providers regarding lot number discrepancy, expired administered doses, and vaccine inventory

Waymo (Contract), Mountain View — *Software Quality Operations Associate*

OCTOBER 2018 - SEPTEMBER 2020

- Daily, triaged 500 simulated test data to develop report on next steps for product development and project management
- Developed feature requests to streamline the bug filing process for safety improvement and data measurement
- Built platforms and dashboards in Google Sheets and SQL scripting to track weekly deadlines using operations data

PROJECTS

Genetic Analysis of the TP53 Gene and Li-Fraumeni syndrome (LFS)— *Project Link*

JUNE 2021

- Wrote a report based on 8 peer-reviewed research papers and visualized how breast cancer is the most prominent cancer related to LFS patients with the frequency of 31.46%
- Established LFS is a rare hereditary cancer disorder that predisposes people to a wide variety of early-onset cancers due to mutated TP53 genes and the inability of the p53 proteins to perform cell division through the NCBI's genetic analysis tool
- Applied variant annotation and scoring tools using Combined Annotation Dependent Deletion to identify that single nucleotide variant rs11540652 has a high-quality nucleotide identification of 99.9% and is an extremely harmful variant in TP53 gene mutations'

Data Visualization of Uber Pickups in New York — *Project Link*

MAY 2021

- Created data visualizations of Uber pickups from April 2014 to September 2014 in New York in R using the ggplot2, ggthemes, dplyr, tidyr, lubridate, DT, scales packages
- Created a geo-plot to visualize a comparison of the number of pickups by their pickup bases within the same time period and reached the conclusion that Thursday had the most pickups out of any day of the week, and 6 pm was the busiest hour on any given day