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

- Built data visualizations to show the number of expired doses across every county and the amount of Covid vaccines available to be ordered by providers for the California Department of Public Health's Vaccine Taskforce Team
- Automate code to extract the number of COVID doses administered to various demographics by schools and childcare centers with the tidyverse packages in R and report the findings to the California Governor's Office
- Outreach 15 vaccine providers on lot number discrepancy, expired administered doses, and vaccine inventory reporting and produce statistical reports explaining the data concerns

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

OCTOBER 2018 - SEPTEMBER 2020

- Performed data quality assessment, measurement, and reporting on simulated data for product development and project management
- Collected operations data and built platforms and dashboards in Google Sheets and SQL scripting to track weekly deadlines
- Triaged 500 simulated test data daily to identify problematic areas and provide actionable insights
- Advised and developed feature requests to streamline the bug filing process for safety improvement and filed 10 safety bugs daily

PROJECTS

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

JUNE 2021

- Consolidated information from 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
- Used Combined Annotation Dependent Deletion to find out the single nucleotide variant rs11540652 has a high-quality nucleotide identification of 99.9% and an extremely harmful variant in TP53 gene mutations'

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

MAY 2021

- Used the ggplot2, ggthemes, dplyr, tidyr, lubridate, DT, scales packages to create data visualizations of Uber pickups from April 2014 to September 2014 in New York in R
- Created a geo-plot to visualize the number of pickups by their pickup bases to identify how customers made trips from various bases in the same time period
- Concluded Thursday had the most pickups out of any day of the week, and 6 pm was the busiest hour on any given day