# Azaima Asghar

Markham, Ontario L6B1K4
Phone:(647)648-9850 | Email:azaima786@gmail.com
LinkedIn:linkedin.com/in/azaima-asghar-526b59160 | GitHub:https://github.com/Azaima-Asghar
Portfolio:https://azaima-asghar.github.io/Portfolio/

## Summary

A diligent data specialist passionate about helping businesses succeed. Experienced in collecting, organizing, interpreting, and condensing large data into a simple understandable story. Expertise in data mining, machine learning and web scraping. Energetic presenter and confident communicator with the ability to distribute information in a way that is clear, efficient, and beneficial for the end users. Excellent at solving complex problems and determining modifications for maximum use of organizational data. Expert at providing practical projections and implementing different scenarios to determine viable process strategies to utilize.

## Technical Skills

**Programming Languages:** Python, SQL, VBA, JavaScript, R, SAS, HTML.
**Tools:** Excel, Jupyter Notebook, VisualCode, Microsoft Power BI, CSS.
**Data Visualization:** Tableau, D3.js, Plot.ly, Leaflet.js, Matplotlib.
**SQL:** PostgreSQL, Microsoft SQL Server, MongoDB.
**Big Data:** PySpark, Apache Hadoop.
**Python Libraries:** Pandas, NumPy, SciPy, SQLAlchemy, PyCaret, Scikit-Learn, TensorFlow, Keras.
**Amazon Web Services (AWS):**RDS, S3.
**Excel:** VBA, Macros, Pivot Tables/Charts, Index/Match, VLOOKUP.
**Statistics:** Descriptive & Inferential, Confidence Intervals, AB Testing, Hypothesis Testing, T-Test, ANOVA.

## Projects

**Modeling Probability of Delinquency | https://github.com/Azaima-Asghar/Modeling-Probability-Of-Delinquency**

- The aim of this project was to see how many approved mortgages go into delinquency (failure to pay an outstanding debt). The dataset was collected from Fannie Mae, a mortgage company. The dataset was hosted on AWS and pgAdmin was used to access the data.
- After performing a detailed exploratory analysis in python, multiple machine learning models were built using the dataset to check if debt to income ratio had any significance to predict accounts going into delinquency. Random forest classifiers yielded the best results with the classification accuracy of 78%.
- An interactive dashboard was built using JavaScript, HTML and CSS to display the results using seller names from the mortgage dataset.

**Plotly and Belly Button Biodiversity | https://github.com/Azaima-Asghar/Plotly-and-Belly-Button-Biodiversity**

- This project looked at the bacteria species that had the ability to synthesis proteins that taste like beef. The hypothesis was that the microorganism which tasted like beef could be found in the human body's belly button.
- To test the hypothesis samples of the navels of people were used in the form of json. Each person's identity was anonymous.
- Using JavaScript, HTML and CSS an interactive dashboard was built with a gauge, a horizontal bar chart and a bubble chart to represent the results of the study.

**Surf's Up | https://github.com/Azaima-Asghar/surfs_up**

- This project looked at the research to open a shop that would serve surf boards and ice creams to locals and tourists in Oahu in Hawaiian.
- The investor had one concern about the weather, he wanted to run analytics on the weather data set that he had from Oahu. The data was in SQLite and using python, flask a detailed analysis was carried out.
- The results showed that seasons could affect the surf and ice cream shop business. Therefore, surfing business might be slower in winter in contrast to summer in Oahu.

## Experience

**Data Portfolio Manager**
**SimplyGreen Home Services - Toronto** **April 2019 – Present**

- Collect the usage data from power companies (Toronto Hydro, Alectra, Enbridge, etc.), clean, visualize and summarize data for the higher management by using Python.
- Create and present compliance reports for SimplyGreen program for clients, presented findings in team meetings to inform business decisions, resulting in a new marketing strategy and 7% revenue growth.
- Utilize PowerPoint and keynote to generate presentations.
- Work alongside different departments to automate regular tasks to boost overall workplace efficiency.
- Collaborate with HR and perform in-depth analysis of internal and external campaigns.
- Design and deploy data visualization roadmap; prepare dashboards and stories to present the results.

**Data Science Intern**
**Excite Club - Toronto** **May 2018 – Dec 2018**

- Cleaned, visualized and summarized various statistical methods to analyze the data.
- Researched and presented different machine learning algorithms using various sources to inform key stake holders of the value to the organization.
- Developed machine learning algorithms and extracted data from Google API's for analysis.
- Implemented machine learning algorithms with statistical software including: Python, R and SAS.
- Created/Managed group meetings for team collaborations using slack.
- Coordinated and distributed the weekly tasks for team members using Trello.
- Drew inferences and communicate results to influence marketing strategies.

## Education

**Certificate in Data Analytics: University of Toronto** **Feb 2020 – Aug 2020**
A 24-week intensive program focused on gaining technical programming skills in Excel, VBA, Python, R, JavaScript, SQL Databases, Tableau, Big Data, and Machine Learning.

**Honour's Bachelor of Science: University of Toronto** **Sept 2014 – Dec 2018**
Double Major: Statistical Machine Learning and Data Mining Stream
The Machine Learning and Data Mining stream provided a specialist program in a dynamic, high-demand, and fast-growing field that lies at the intersection of statistics and computational sciences. It was driven by applied problems in science and technology.