New York City is one of the largest and most populous cities in the United States and in the world. Various methods and tactics have been implemented in combating the high crime rate in this city. Due to the city being a major tourist attraction, as well as having a massive population and a daily influx of commuters, there are such huge gaps in security that it is unfeasible to mitigate every danger. Our research aims to analyze a dataset of historical complaint data from the New York Police Department and attempts to identify feature-based trends within the given data. The primary focus of the project is on determining the risk level of an individual based on their location and their demographics. Leveraging machine learning using the Apache Spark MLlib system, as well as Python-based visualizations, we additionally created an interactive map showing risk levels for an individual around the city. From a dataset containing 7.4 million records of crimes committed in the last two decades, we were able to generate a model capable of predicting a person’s risk level in zip codes around New York City with a high accuracy rate, and display the risk levels geographically as a heat map.