**NAGA LAKSHMI K**

7095979454| [nagalakshmikandregula@gmail.com](mailto:nagalakshmikandregula@gmail.com)

Bangalore

**Profile Summary**

* 2 years of total experience in Exploratory Data Analysis
* Experience in working with large datasets and to analyze data using python libraries
* Proficient in tools like Python, SQL and Excel

**Technical Skills**

* Programming languages

**Python,SQL**

* Dashboarding & visualization

Power BI, Excel, **Matplotlib, Seaborn**

* SAS
* Basic knowledge of **Big Data**
* HDFS(Hadoop distributed file system)
* MapReduce
* Spark
* Hive
* Flume
* Knowledge on **statistics**
* Descriptive statistics
* Inferential statistics
* Measures of tendency
* Central limit theoremp
* Normal/Gaussian distribution
* Finding the outliers using **z-score, IQR**(interquartile range)
* Machine Learning
* Deep Learning

**Qualifications**

|  |  |
| --- | --- |
| **Education** | **Year of passing** |
| B.Tech | 2013 |
| Intermediate | 2009 |
| SSC | 2007 |

**Work Experience**

**Sales Forecasting Data Analyst (TCS )(2019 – Present)**

* Hands on experience with large datasets, to perform **Exploratory Data Analysis (EDA),** univariate, bivariate and multivariate analysis.
* Using statistical tools to interpret data sets, paying attention to trends and patterns that could be beneficial to predictive analysis.
* Finding the missing values and replacing those values with statistical methods like mean, median and mode
* To find the relation between dependant features and independent features and plot them using various plots like histogram, scatter plot, violin plot, box plot.
* Statistical analysis using excel, evaluated simple data set using outlier analysis, boxplots, V-lookup, Chi-square, normal distribution.
* Creating dashboards and giving better insights to the stakeholder which resulted in the increase of company sales.
* Handling imbalanced datasets

**Academic projects**

**Spam Classifier**

* The dataset used in this project is from Apache SpamAssassin.
* Apache SpamAssassin is the open source anti-spam platform giving system administrators a filter to classify email and block spam (unsolicited bulk email).
* It uses a robust scoring framework and plug-ins to integrate a wide range of advanced heuristic and statistical analysis tests on email headers and body text **including text analysis, Bayesian filtering, DNS blocklists, and collaborative filtering databases.**

**Housing Price Prediction**

* Extensive **data analysis** and **visualization**
* Finding **missing values** and replacing those values
* Distribution of **numerical** variablesand cardinality of **categorical** variables
* Dealing with **outliers**
* Finding relationship between **independen**t and **dependant** features
* Visualising using various plots like histogram, barplot, scatterplot, boxplot etc.,
* **Standardise** the values of the variables to the same range

**Certifications**

* Certificate programme in **Python** for Machine Learning by IIT Roorkee in association with CloudxLab
* Certificate programme in **Data Science Specialization** by IIT Roorkee in association with CloudxLab
* Certificate programme in **Big Data Engineering** by IIT Roorkee in association with CloudxLab