

Amaasa, Baby Shanthi

+1571-491-7891 | shanthiamaasa@gmail.com

www.linkedin.com/in/shanthiamaasa | <https://github.com/Shanthiamaasa>

EDUCATION

Master of Science in Data Science
University of North Texas, Denton, TX

GPA: 3.9
Graduation Date: May 1st, 2021

Bachelor's in Information Technology
Sreenidhi Institute of Science and Technology, Hyderabad, India

GPA: 9.3/10
August 2015-May 2019

Relevant Courses

Fundamentals of data analytics | Analytics tools techniques and methods | Data visualization | Data analysis and knowledge discovery | Predictive analytics and Business forecasting | Applied multivariate statistics | Data mining and machine learning for business | Discovery and Learning with Big Data | Data Modeling| Deep learning with Big Data.

TECHNICAL SKILLS

Programming : Python, Java, SQL, Hadoop, MapReduce, Spark, HTML, CSS, Hive

Python Libraries: NumPy, Pandas, Scikit-learn, Matplotlib, Seaborn, NLTK

Big Data : Hadoop, Hive, Spark, AWS, Microsoft Azure

Tools : Rapid Miner, Power BI, Azure, SAS EM, SPSS, ArcGIS, Excel, Tableau, MS Office, Alteryx, Anaconda, Informatica.

CERTIFICATIONS & AWARDS

- Microsoft Azure AI Fundamentals.
- Secured 4th place in "Data Science in Python" hackathon conducted by JNTU-H in collaboration with Infogrex and Cola berry organizations.
- KPMG Data Analytics Virtual Internship
- Machine learning: Hands-On Python in Data Science issued by Udemy-45 hours course.
- Statistical Thinking in Python issued by Udemy-9 hours course.

EXPERIENCE

Young Minds Technology Solutions

June 2018- July 2018

Data Science Intern

Tirupati, Andhra Pradesh

- Worked on the project Analysis on H1b visa data using Map Reduce. Using my time management and problem-solving skills the project has been completed prior to the deadline.
- Analyzed the large data sets from 2011 to 2016 with 3M records using **Map Reduce** and extracted useful insights from the data like number of people certified every year, highest and lowest pay for each designation, number of applicants for data engineer role, top 5 popular jobs in each year.

PROJECTS

Analysis and visualization on Covid-19 data using Power BI

- Analyzed on multiple data sets obtained from Kaggle and Our World in Data websites.
- ArcGIS maps, Line charts, horizontal bar chart, ribbon chart has been used for the visualizations in **Power BI**.
- Worked on hypothesis by visualization of data like how the disease spread across the continents, confirming data about confirmed cases is reliable, case fatality rate of top 5 effected regions, and proving that most effected regions need not have high case fatality rate.

Analysis on human freedom index

- Used data from Kaggle and **Excel** for data processing and data analysis. The analysis has been performed using **pivot charts** and **lookup functions**.
- Count of countries in Human freedom quartile in each region, average index each year, average index of religious harassment, range of index of disappearances.
- It has been concluded that index of freedom and women's security kept reducing each year.

Startup Funding Prediction

- Used data from Kaggle from 2015 to 2017 to predict the funding for startups.
- Analyzed data and interpreted how the funds vary based on city, type of industry, minimum, maximum funds for startup category, important investors in Indian ecosystem, type of investment and prediction of funds for every industry using the **ARIMA** algorithm from **scikit-learn** library with R-squared value of **0.85**.

