Pradnya Tipare

(970) 214-2126 | pradnya.tipare@rutgers.edu | www.linkedin.com/in/pradnya-tipare |
https://github.com/PradnyaTipare | Harrison, NJ

EDUCATION

Rutgers -The State University of New Jersey, Harrison, NJ — Expected Grad date: **December 2022**
Master of Science in Information Technology and Analytics — GPA: 3.63 / 4.0
Coursework: Business Data Management, Data Analysis and Visualization, Project Management, Analytics Business Intelligence, Business Analytics Programming

University of Mumbai, Mumbai, India — **July 2015 - May 2019**
Bachelor of Engineering, Computer Engineering — GPA: 3.75 / 4.0

EXPERIENCE

Questant.Inc, New Jersey, US — **June 2022 - Present**
Data Analyst

- Transformed **Project cost data** from Excel to **AWS RDS** and managed data using **PostgreSQL** and **PgAdmin**
- Automated currency conversion and created calculated fields using complex **stored procedures**, **triggers**, **functions,** and SQL queries to get the common insights, reducing the manual work by 80%
- Generated Project Portfolio, Project Cost and Budget summary visualizations and dashboard using **Microsoft Power BI**

Reliance Jio InfoComm Ltd, Navi Mumbai, India — **June 2019 - August 2021**
Data Engineer (Assistant Manager)

- Collaborated with product managers to define requirements, metrics, KPI's and created interactive dashboards of **customers data** in **Tableau** for real-time user stories visualizations, reducing manual reporting by 50%
- Implemented end to end data processing pipeline for real-time data analytics leveraging **Hadoop HDFS, Apache Spark**
- Developed 50+ Spark jobs using **Scala** and **Spark SQL** to process the **customers and viewership** data and integrate into **MySQL, Cassandra** database
- Retrieved and aggregated analytical data from MySQL, Cassandra using **SQL** to generate reports for product managers
- Automated the process of converting raw data from Hadoop cluster to parquet files, compressing size by 75%
- Scheduled, monitored, and managed **ETL** jobs using **Apache Airflow** by creating the workflows which decreased the manual scheduling on cron by 30%

ACADEMIC PROJECTS

Tree Census in New York | *Tableau Project* — **September 2021 – December 2021**

- Derived insights from a dataset of 6.8 million street trees from across the 5 boroughs of New York regarding their health, problems leading to poor health and the solutions to improve their health
- Executed **Data Cleaning** to handle null values and **Data Transformation** to convert data from one format into another for data warehousing and analyzing
- Generated dashboard using quick filters and delivered an interactive story using **Tableau** to showcase data analysis which add to the success of solving the problems by 50%
- Mapped trees using street view to monitor the health status filtering by location

Light Pollution Prediction System — **December 2018 - April 2019**

- Designed a machine learning model by applying random forest algorithm on 10 input variables to determine the light pollution index
- Enhanced the data-by-data exploration and **Data Munging** to treat the missing values, increasing data efficiency by 42%
- Published research paper in **Springer** named Analysis of Light Pollution prediction using Mathematical Model and Machine Learning techniques and won 3rd Prize in national level project competition conducted by Mumbai University

TECHNICAL SKILLS

Languages: SQL, PostgreSQL, Spark SQL, Scala, Python, Java, R, C, C++
Databases: MySQL, Cassandra, Microsoft SQL Server, PostgreSQL, Oracle SQL, BigQuery
Cloud Technologies: Amazon AWS, Google Cloud
Data Visualization Tools: Tableau, Kibana, MS Power BI, MS Excel, Google Data Studio
Big-Data Technologies: Hadoop, HDFS, Apache Spark, Yarn, Apache Kafka, Apache Airflow
Data Integration and other Tools: PgAdmin, SQL Server Management Studio (SSMS), Microsoft Office, Jira

CERTIFICATIONS

SQL for Data Science by UC Davis — **September 2020**
Introduction to Google Data Studio — **November 2022**