

Kavin Shah

shah.kav@northeastern.edu | +1-(857)-999-6623 | www.linkedin.com/in/Kavin-shah | <https://kavshah.github.io/>

EDUCATION

Master of Science (MS) in Information Systems

Northeastern University, Boston, MA, USA

Sep 2018 - Dec 2020

Bachelor of Engineering (BE) in Information Technology

Vidyalankar Institute of Technology, University of Mumbai, Mumbai, India

May 2013 - May 2017

SKILLS

Databases/Languages	MySQL, MsSQL, OracleSQL, PostgreSQL, Google BigQuery, IBM DB2, Python
BI/Reporting Tools	Tableau (Tableau Desktop Specialist 2020), Power BI, Looker, Ms Excel, Google Analytics
Tools	Talend, Alteryx, GCP
Statistical Techniques	Regression Analysis, Bayesian Model, Time Series Forecasting, Monte Carlo Simulation

PROFESSIONAL EXPERIENCE

Data Reporting Analyst – Arizona Department of Education – Phoenix, USA

Apr 2021 – Present

Recruitment and Retention

- Enhanced the query speed by 25% by creating complex queries using CTE and Window functions which subsequently led to faster data refresh of **Power BI** dashboards.
- Found underlying reasons for having ethnic disparity in the teacher student ratio by analyzing the data for the past 10 years using MsSQL and Power BI and **effectively communicated** the analysis and trends to the stakeholders.

Data Assistant – Northeastern University – Boston, USA

Aug 2020 - Dec 2020

Co-Op MGEN

- Designed a data warehouse and established connection with Power BI to **decrease the overall processing time to generate reports by more than 50%**.
- Collected data form programming survey Qualtrics and **performed QA** to maintain the quality of data.
- Reduced the issue rate resulting by **ambiguous data by 90%** by analysing the data and setting a set of pre-defined rules.
- Analysed the previous 5-year job posting data and created visualization in Power BI which directly impacted the outreach program and **help our department Increase the job postings by 13%**.
- Managed the bi-weekly data focus meeting** with the internal stakeholders to answer any data related question and provide ad-hoc data reports in excel.

Database Developer and Analyst Intern – Granite Telecommunications LLC – Quincy, USA

Aug 2019 - Dec 2019

Information Technology Department

- Performed Feature Engineering** in Data Pipelining project that decreased communication time by 25% and increased the interaction rate 21%, resulting in revenue generation of ~\$3 million.
- Automated the migration process** from MsSQL to Hadoop by creating Python script that helped eliminate the manual process of migrating each table individually and save time.
- Architected metadata and wrote Stored Procedures** in MsSQL to populate the website content dynamically by collaborating with cross-functional teams.

Data Analyst – TeknoPoint – Mumbai, India

May 2017 - Jun 2018

Information Technology Department

- Automated a report in Tableau for the sales team using filters, parameters and calculated sets that automatically qualified sales lead resulting in a lift in **conversion rate of 20%**.
- Responsible for maintaining the integrity of the Database and reporting any issues to the manager.
- Monitored expenses, sales and revenue; conducted weekly reports in Tableau and presented findings to executives, helping **reduce 18% expenses and increase 11% sales** over 12-month period.

RELATED PROJECTS

IMDB Data and Analysis | (MsSQL, Talend, Alteryx, Tableau, Looker)

Mar 2020 - Apr 2020

- Created a one-click execution process that ingested the Database with more than 250 Million rows in less than 60 mins.
- Learning outcome** – process of ingesting using data integration tool, ETL process before loading the data in BI tool and visualizing, Connection process from Database to Talend to BI tool and the issues that can surface during the connection.

IBM Data Science Capstone, Battle of neighborhoods | (Python, IBM Watson, Foursquare API)

May 2020 - Aug 2020

- Recognized neighborhoods based on features where a Fast-food restaurant has potential of being more than 90% (CI) successful.
- Learning outcome** - This project was a culmination of a 9-course part wherein I learnt working in IBM Watson environment, using SQL in Python, how to use API call in Python, process of identifying data points, Clustering techniques, Visualizing clusters using Folium and working with real life data and issues.