

Data Engineer Data Engineer Data Engineer San Jose, CA Strong expertise on Hadoop big data, SQL and HQL. Worked on Kafka p.o.c for end-to-end implementation to receive real time data flow from AWS server. Experience in data visualization, data analytics, data integration, data quality using Python. Used NLTK libraries to detect the customer satisfaction through their verbatim comments. (Tokenization and standardization are part of this process) Querying on large data sets to process and refine the data for down streams. Automated the audit system to recognize any errors on fly (Null percentage and validity) Expertise in RDBMS, Agile, Scrum and waterfall methodologies. Creating, Reading and Updating Data in MongoDB. (Inclusive optimization and monitoring) CSS, HTML and JavaScript along with good knowledge of backend technologies including Python. Tableau expertise with 12 months of experience. Strong understanding of project life cycle and Software Development Life Cycle (SDLC) . WordPress, MS office tools, CMS and multiple platforms Windows and Mac. Strong command on MS Excel with VLOOKUP, MATCH and INDEX functions. Work Experience Data Engineer Telemetry - San Jose, CA June 2018 to March 2019 As a Data engineer worked on huge raw data sets (Telemetry data) Ingesting the data by creating tables in Hive and achieving high-level optimization of the ingested data by altering and refining through HQL. Data classification and text analysis using NLTK libraries in Python. Collecting and compiling the various datasets from various data generating sources (AWS, MYSQL servers, Oracle Database) Resolved compatibility issues by working on different areas of data architecture including data ingestion and pipeline design. Worked on Machine learning and advanced data processing to see the customer trends (scrutinize) in cumulative. Developed and implemented scripts using machine-learning algorithms like Na ve Bayes Classifier Algorithm, k means clustering, Random Forest classifier for regression, clustering etc. Created many plots, bar charts, graphs, histograms of the complex data by visualizing the data using matplotlib in python. Maintain source code repository in subversion and handled branching, tagging & merging process. Research, design and develop computer software systems, applications which require use of advanced computational and quantitative methodologies Python Developer Ravin Solutions May 2015 to December 2015 Worked as a Python developer at Ravin Solutions using Django for a

content management system project that allowed collaborative workflows across users. Designed ER diagrams. Aided in managing performance indexes and built/maintained ETL pipelines to feed the data. Worked in Tableau environment to create dashboards like weekly, monthly, daily reports using Tableau. Various analytical reports were designed and developed from multiple data sources by combining data on a worksheet in Tableau Desktop. Web developer, Intern Ravin Solutions January 2013 to December 2015 Worked as a web developer at Ravin Solutions using HTML, CSS and JavaScript. ? Projects included various extensions to the company site. Was also responsible for monitoring and fixing website crashes and any web log errors. Assisted in managing and enhancing the company's CRM system. Also performed graphic designing using cs6 and cc illustrator. ? Black box testing is done manually to check the websites performance and stability. Education Database Management Systems Enterprise System University Alliance May 2016 to November 2016 Master of Science in Management Fairleigh Dickinson University - Teaneck, NJ August 2014 to December 2014

Name: Tom Gilbert

Email: lgarcia@example.net

Phone: 520-259-6213x5328