# Mantej Singh Dhanjal

### **Business Intelligence Analyst**

Four years of analytical experience with data extracting applications, practical knowledge of data reporting/analysis tools, Business Intelligence tools in Manufacturing & Retail domains. Exposure to various tools: Informatica 9.1, SQL, Tableau, Spreadsheet, R, Python(PANDAS), HIVE and Apache Spark.





### **Work Experience**

### **Insight Global**

ADP: Data Engineer

02/2018 - Present

Parsippany, NJ

- Data reporting, data migration, and creating dictionaries to catalog the data for ADP's DataCloud & GSS Team
- Environment: Sqoop, AWS, Alation, Python, HIVE SQL, VBA, Tableau, Confluence and Rally

#### Citi

### AML Advisory- Machine Learning Engineer

06/2017 - 02/2018

Warren,NJ

- Solely handled a POC for the Citi's AML Discovery Analytics Program
- Analyzed Customer, Account and Transaction data to identify & predict unusual behaviors using Apache Spark MLlib,H2O Flow and Alpine Data
- Environment: HIVE SQL, Apache Spark ML, H2O Flow, DataRobot, Rstudio, Python and Alpine Data

### Alpha IT Partners

#### Data Analyst

12/2016 - 05/2017

NYC

- Responsible for performing and checking regularly scheduled tasks such as usage processing, generating customer bills, and troubleshooting any issues that arise
- Executed manual operations and monitored the performance of automated tasks in the billing systems

#### Bluefly

### Business Intelligence Analyst Intern

06/2016 - 08/2016

NYC

- Wrote complex SQL queries to get relevant data for analysis from customer's feedback, Sale, and Stock datasets
- Implemented Joins, Data Blending, Custom SQL in Tableau, created dashboards & pivots in Tableau and Excel
- Environment: Tableau Desktop, Excel, SQL Developer, R studio, Python, Google Analytics & Google Trends

#### Accenture

### Associate Software Engineer/BI Developer

11/2013 - 07/2015

Mumbai, India

- Build dashboards, automated reports, and ad hoc reports based on agile business requirements
- Developed and overhauled ETL mappings and mapplets to extract data from source systems into the Staging area
- Automated SQL queries, fetched data directly from the Informatica repository and listed output as succeeded, failed, long running jobs saved overall production time by 2+ hours
- Environment: Informatica Power Center 9.1, SQL, QlikView, Crontabs, OBIEE, BO, Toad 10.5, Control M

### **Education**

### **Master's Degree- Computer Science**

New Jersey Institute of Technology, US. 09/2015 - 08/2016

#### **Bachelor's Degree-Information Technology**

University of Mumbai, India. 05/2010 - 06/2013

### **Skills & Competences**



### **Projects**

#### Word Count | Apache Spark 07/2017 - 08/2017

 Ran Apache Spark with YARN in client mode, compared and observed performance for groupByKey & reduceByKey in Spark UI.

### Web Scraping Project | Python 12/2016 - 01/2017

 Crawled the data to parse person's name, email address, and net worth, generated report from it using BeautifulSoup, Pandas, RE, NLTK and D3 API.

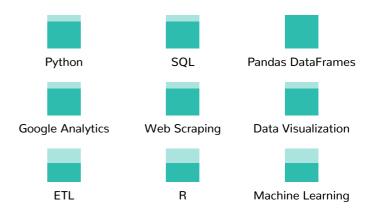
#### ETL Pipeline | Python 07/2016 - 08/2016

 Python code to extract, clean & dump the data with email notification using Pandas.

#### Social Media Analyzer | R 06/2016 - 07/2016

Sentiment analysis over Twitter API using R.

## **Programming**



### **Interests**

Big Data | HIVE | Data Analysis | Data Visualization | SQL