

## PROFESSIONAL OVERVIEW

- **5.2 years of total experience in IT industry as Data Engineer, Big Data Technology like Hadoop, Hive, Sqoop.**
- **Around 4.9 Years of experience in Python.**
- **Experience of writing UNIX Script.**
- **Extensive Knowledge of Regex.**
- **Job responsibilities include data analysis and reporting.**
- **Knowledge of Excel(Conditional Formatting,v lookup and Pivots etc)**
- Created packages and implemented custom logging in Python for easy debugging.
- Created script for transfer of our data backup to ensure Disaster recovery in Python.
- **Around 5.2 Years of experience in MySQL.**
- Around 4.7 years of experience in SAP BO reporting tool.
- Adaptable and proficient in all the phases of Data warehouse life cycle involving Requirement Analysis, Design, Coding, Testing, and Deployment.
- Extensive knowledge on Reporting using complex SQL queries, stored procedures, functions triggers, Indexing, Join, Views, and transactions in MySQL.
- Working experience of Big Data, Hive and Sqoop.
- Strong knowledge of systems and software development life cycles, application integration techniques and relational database design.
- **Interested in learning and working on Development in Machine Learning.**

## PROFESSIONAL EXPERIENCE

- Working as Python Developer with Paytm Noida from Feb, 2016 to till date.

## TECHNICAL SKILLS

Technical Expertise	:	Python,Pandas,Numpy,SQL Procedure, Function, Trigger, Indexes, Query Optimization, Hadoop
Operating System	:	UNIX,Windows 7, 8, 10, XP
Language	:	Python,Shell Scripting,SQL and SQL Server
Tools	:	BIRT, Putty, Kettle,SAP BO (Business Object)
Databases	:	SQL,MySQL, MariaDB, Hive

## PROJECT WORK:

### Project #1

Organization	:	Paytm (One97)
Team Size	:	25
Role	:	Python Developer
Database	:	SQL SERVER , MySQL and Hadoop
Technology	:	UNIX, Hive, Sqoop
Tools	:	SAP Business Object, BIRT

### Project Description:

The project aimed to provide comprehensive business intelligence solution for major telecom operator. The flow was designed to read, transform and integrate data from multiple databases and finally present per-defined intelligent dashboards and reports to business for constant intelligence support.

### Roles & Responsibilities:

- Create custom ETL process (In-House, written in python) using Python and SQL procedure and functions for data load into Paytm databases.
- Develop MSISDN SEARCH PORTAL for data retrieval, ETL and Data Reconciliation Scripts using Python,Django and Pandas.
- Scheduling the scripts in Crontab for auto run of the procedures and shell scripts.
- Created script for transfer of our data backup to ensure Disaster recovery in Python.
- Created packages and implemented custom logging in Python for easy debugging.
- Store Source data into SQL SERVER database and MySQL.
- Fire queries on daily basis to check the data on circle and central servers.
- Performed analysis and presented results using SQL queries and MS Excel.
- Involved in query optimization and troubleshooting issues in development environment.
- Handle heavy data volume on the daily basis and optimize query and ETL process.
- Create BIRT report and SAP BO reports and dashboard on the basis of key processing indicator.
- Create design SAP BO universe and map business requirements.
- Archive daily data into Hadoop cluster using HIVE integrated with Sqoop.
- Data analysis of historical data using HIVE queries.
- Involved in analysis of end user requirements and business rules.

### Project Details :

- MSISDN Search Portal is a Panel which is designed to serve the purpose of fetching data and KPIs from the files which are generated at given path in the server by using some shell script or Python script. Panel can fetch the data from the generated files by using **Python and Pandas** according to the client requirement. I have started learning and development in Python,Pandas and Django with this project as a single member of team.

Role : **Full Stack Developer**, Language : **Python**,Framework : **Django**

- Knowledge Kart is a Panel which is designed to serve the purpose of fetching data and KPIs at any granular level for different product of a domain, So that user don't have to ask and wait for developer for fetching data, now he/she can direct fetch data directly from panel without a single minute wait. Panel can connect with multiple databases according to specification : **MySQL, SQL SERVER, Hive**. It is designed in a way that different user can have different /customized privileges according to their authentication.

Role : **Full Stack Developer**, Language : **Python**,Framework : **Django**

- Have developed Data Reconciliation panel which is useful with the data reconciliation between operational DB and DWH (Hadoop) on the granularity of packname.

Role : **Full Stack Developer**, Language : **Python**

- E1 utilization panel which is designed to update the Numbers of E1 of the servers by client or sales team in order to calculate the utilization of the server.

Role : **Full Stack Developer**, Language : **Python**,Framework : **Django**

**CURRENT ORGANIZATION DETAILS:**

Company : One97 Communication Limited (Paytm)  
Role : Developer  
Notice Period : 2 month  
Job Type : Permanent

**Declaration:**

I hereby declare that the information furnished above is true to the best of my information and belief.

Mohit Khatri