

Vishal Grover

Sr. Data Analyst

(972) 948 5194

Professional Summary:

- More than 11 years of experience in **Hive, Python, Hadoop Big Data, MS SQL Server, MSBI (SSIS, SSRS, SSAS), ASP.Net with C#**.
- Implemented Proofs of Concept on Hadoop stack and different big data analytic tools, migration from different databases (i.e Teradata, Oracle,MySQL) to **Hadoop**.
- Strong Knowledge of **Hadoop and Hive** and Hive's analytical functions.
- Capturing data from existing databases that provide SQL interfaces using **Sqoop**.
- Efficient in building **hive and pig scripts**.
- Having good experience in Hadoop Big Data processing. Expertise in developing the queries in Hive, Pig.
- Mastering/Leading in the development of applications/tools using Python.
- Worked on several python packages like **numpy, scipy, pytables** etc..
- Having Good Experience in Object Oriented Concepts with Python.
- Integrated different data sources, data wrangling: cleaning, transforming, merging and reshaping data data sets by writing **Python scripts**.
- Successfully loaded files to Hive and HDFS from Mysql.
- Loaded the dataset into Hive for ETL Operation.
- Expert in **SSIS Package development, deployment, Job Scheduling and Error Trapping, SSRS Report development, deployment, subscription, SSAS Tabular and Multidimensional Cube development**.
- Created Transformation and Modelling in **Power BI Visualizations**.
- Generated **Tableau Dashboard with quick/context/global filters, parameters and calculated fields** on Tableau (9.x) reports.
- Created Tableau Dashboards with interactive views, trends and drill downs along with user level security.
- Strong experience as Business Intelligence Developer and Data Analyst in Production, Development and Staging Environments.
- Defining **data warehouse** (star and snow flake schema), fact table, cubes, dimensions, measures using **SQL Server Analysis Services with DAX**.
- Designing and Deployment of Reports for the end-user Requests using **SharePoint & SSRS**.
- Expert in creating **Parameterized Reports**.
- Expert in **T-SQL** to create Table, view, user Defined Function (UDF), Stored Procedures, Triggers etc.

Technical Skills:

Programming language **T-SQL, C#, Python**

RDBMS **SQL Server 2005/2008/R2/2012/2014, PGAdmin 1.4**

Big Data Ecosystem	Hadoop, Hive, Pig, Sqoop, Talend
Internet Technologies	ASP.Net with C# and VB , JavaScript HTML
GUI Tools	Visual Studio 2005,2008,2010,2012
Operating Systems	Window 2000 Server, Windows Server 2003, XP , Windows 7,
Version Control System	VSS, TFS. SVN
Internet	IIS, HTML, HTTP
ETL Tool	SSIS, Talend
Visualization Tools	Tableau, Power BI
Graphical Database	Neo4j

Professional Experience:

Otsuka, Los Angeles, CA

Role : Sr. Data Analyst

May'17 to Sept'17

Project Description: The project goal is to analyze pharma business by combining data from different sources (i.e. csv, excel, mysql, hive, json, postgres, oracle) to one centralized place and then create nodes and relationships in neo4j graphical database.

Responsibilities:

- Explored different implementations in hadoop environment for data extraction and summarization by using packages like Hive, Pig.
- Importing and exporting data into HDFS using Sqoop.
- Worked on python files to load the data from csv, json, mysql, hive files to Neo4j Graphical database.
- Created Neo4j Graphical database nodes and relationship.
- Implement cypher queries to manipulate data on Neo4j database.
- Involved in Design, analysis, Implementation, Testing and support of ETL processes for Stage, ODS and Mart.
- Used Talend as a Data Cleansing tool to correct the data before loading into the staging area.
- Collect and link metadata from diverse sources, including relational databases and flat files.
- Extracted data from different Flat files, MS Excel, Hive and transformed the data based on user requirement using Talend and loaded data into target, by scheduling the sessions.
- Supporting daily loads and work with business users to handle rejected data.
- Implemented data cleansing for files using Talend.
- Created data model in Postgres using dimensional model.
- Load data into Postgres using snow flake schema.
- Performed Unit Testing and tuned for better performance.
- Created Reusable Transformations and multiple Mappings.

Environment: Python, Hive, Pig, Sqoop, Neo4j, Talend 9.6, PGAdmin 1.4

Florida Power and Light, Miami, FL
Role : Sr. Data Analyst

Jul'16 to Apr'17

Project Description: The System will streamline the verification of the sections like revenues billed, cash postings, refunds issued, meters uploaded and cash deposited to banks. The System has ability to create dashboard, custom reporting and analytics. The Revenue balance data can be drilled down by measures and dimensions, enabling users to gain access to accurate, up-to-date Information for better decision making.

Responsibilities:

- Developed and maintained Python ETL scripts to scrape data from external sources and load cleansed data into a Sql Server. The data was used for daily electrical power virtual trading activities in several markets.
- Implemented discretization and binning, data wrangling: cleaning, transforming, merging and reshaping data frames using Python.
- Integrated applications with designing database architecture and server scripting, studying & establishing
- Designed data visualization to present current impact and growth
- Loaded the dataset into Hive for ETL Operation.
- Written Hive scripts to extract data from staging tables
- Building, publishing customized interactive reports and dashboards, report scheduling using Power BI Desktop.
- Used Power BI Power Pivot to develop data analysis prototype, and used Power View and Power Map to visualize reports.
- Used DAX (Data Analysis Expressions) functions for the creation of calculations and measures in the Power BI Models.
- Designing and apply Microsoft Project Plan.
- Requirement, Design, Development and Testing.
- Data Modeling and Cube design in SSAS using Tabular Model.
- Monitoring overall progress and use of resources, initiating corrective action where necessary.
- Planning and monitoring the project.
- Preparing and maintaining project, stage and exception plans as required.

Environment: Hadoop, Hive, Python, SSAS, Power BI

HMS, Irvine, CA
(NASDAQ: HMSY)
Project: AmeriHealth
Role : Data Analyst

Mar'14 to Jul'16

Project Description: Client is an innovative healthcare services company with a Comprehensive, patient – centered approach to diabetes management. SSIS has been used as an ETL tool for building of data warehouse and data consolidation. Client provides the claim, provider and membership files on daily, weekly and monthly basis on FTP or EDI Server.

Package load the claim data into staging database. After loading Medicare and Medicaid data in staging edits (Rules provided by State and Federal government for claims) applied on staging database and then after finally loaded into the Data Warehouse and then mail delivery report to the client at every step as Sanity Check, Return file and Cleanup process. Every step maintains in Sql server log table. Error handling and monitoring the package is done by the BlxPress tools.

Responsibilities:

- Designing, developing, deploying , job scheduling reports and success/failure notifications in MS SQL Server environment using SSIS in Sql Server data Tools (SSDT).
- Used ETL (SSIS) tasks to develop jobs for extracting, cleaning, transforming and loading data into data warehouse.
- Extracted data from various sources like SQL Server 2008, Oracle, CSV, Excel and Text file from Client servers and through FTP.
- Experience working in a large data environment doing SQL development.
- SQL query optimization (Knowledge of Transactions, Isolation Levels, Indexes, Blocks, Locks, and Database Partitioning etc.)
- Implemented SQL constraints (Primary Key, Foreign Key, Index, Unique, Not Null, Default)
- Experience managing large projects with a close attention to detail and high-quality performance.
- Data modelling and mapping using UML in visio.
- Experience in handling Enterprise datasets.
- Identify issues and fix those issues while troubleshooting.
- Experience in writing MDX queries to access cubes data in SSAS.
- Monitor scheduled SSIS Package through the BlxPress tools
- Error Trapping through BlxPress tool, Script Writing in .net, Transaction, Email Alert and some complex Transformation in SSIS
- Deployment and Scheduling SSIS Package through SQL Server and file system Report development in SSRS
- Create SSIS packages, Email alert for SSIS packages and schedule it through SQL Server and file system.
- Manage the Dev, SQA and Production Environments.

Environment: SQL Server 2012 with Integration, Reporting Services and Analysis Services, C#.net code for SSIS scripting, BlxPress, Microsoft Business intelligence development studio, SQL Server 2012.

Sanare, Miami, FL
Role : Data Analyst

Sep'12 to Mar'14

Project Description: The system is a database driven online Reporting model and SSIS package for data upload and send reports by the packages, which is for Health care domain in USA, for this report we have create our own data warehouse and fetch all report from this data warehouse. Around 100+ business reports have been created using functionality of SSRS.

SSIS has been used as an ETL tool for building of data warehouse and data consolidation. ETL process has been designed on following lines

- Download daily extracts from SFTP server and upload them into data warehouse after performing various data integrity checks
- Preparation of data comparison sheet to compare record counts in tables against the counts which are supposed to be their
- Execution of data maintenance tasks e.g. Rebuild Indices
- After successful completion of data upload this package kicks off second package to generate various reports and distribute them in encrypted form to intended recipients
- Preparation of logs has been maintained during execution of package
- Notifications to users of failure of package, if occurred
- Notifications to users of success of important tasks to keep them aware during execution
- At last summary has been sent to show the statistics of package
- Real time data has been picked from Other production systems for integrated reporting and Dashboards
- Package is configurable and data driven so it can be changed easily as per the changing business environment.

Responsibilities:

- Data warehouse architecture designing.
- ETL package Development in SSIS.
- Deployment and Scheduling of SSIS Package in SSIS.
- Populate data marts with the right data at the right time.
- Performance Tuning/ SQL Query Optimization, Complex Query Design.
- Stored Procedure Writing, Function Writing.
- Creating Reports in SSRS.
- Project monitoring, Status reporting and client interaction.
- Monitoring the critical components of the System.
- Implementation and support of systems in production environment.

Environment: SQL Server 2008 R2 with **Integration and Reporting Services**, SSIS Scripting, Microsoft Business intelligence development studio, SQL Server 2008 R2

Aspen, UK

Apr'12 to Sep'12

Project: Data Repository Manager

Role : ETL Developer

Project Description: Data Repository collects all clinical, operational, and financial extracts across the organization and uploads them in data warehouse after performing various data integrity checks. It also integrate real time cloud data with the ware house and provides an open environment for powerful reporting and decision support to healthcare managerial authorities. SSIS has been used as an ETL tool with following features:

- Created Complex ETL Packages using SSIS to extract data from staging tables to partitioned tables with incremental load.

- Created SSIS Reusable Packages to extract data from Multi formatted Flat files, Excel, XML files into Database Billing Systems.
- Created **SSIS** packages for File Transfer from one location to the other using FTP task.
- Developed, deployed, and monitored SSIS Packages
- Transform extract data after performing integrity checks.
- Implement Email notification in case of package failure.
- Integration with other package for clean-up task.
- Develop various script tasks for validating and transform data.
- Designed SSIS Packages to transfer data from flat files to SQL Server using Business Intelligence Development Studio.
- Extensively used SSIS transformations such as Lookup, Derived column, Data conversion, Aggregate, Conditional split, SQL task, Script task and Send Mail task etc.
- Implement configuration files to deploy the SSIS packages across all environments.

Responsibilities:

- ETL package Development in SSIS.
- Integration of real-time cloud data (NetSuite) with data Warehouse for critical business requirements.
- Report development in SSRS.
- Project monitoring, Status reporting and client interaction.
- Operate as the liaison between Digital Analytics team and IT Development team for all projects and releases including testing pre/post deployment of changes to ensure timely and accurate reporting and insights

Environment: SQL Server 2008 R2 with **Integration and Reporting Services**, Microsoft Business intelligence development studio, SSMS, SQL Server 2008 R2

Aspen, UK

Sep'11 to Mar'12

Project: Patient Reporting System

Role : ETL and Report Developer

Project Description: PMIS is an application which stores information about various types of Patients disease and taking input from different chain management solutions around the globe. PMIS is used to manage patient historical information. Various business reports have been developed for dynamic data analysis. Comparison of actual health data with industry standard in graphical analysis has been shown. Following features have been implemented in this project:

- Designing, developing and deploying reports in MS SQL Server environment using SSRS-2008 and SSIS in Business Intelligence Development Studio (BIDS).
- Used ETL (SSIS) to develop jobs for extracting, cleaning, transforming and loading data into data warehouse.
- Designed SSIS Packages to transfer data from flat files to SQL Server using Business Intelligence Development Studio.
- Creating multiple parameterized stored procedures which were used by the reports to get the data.

- Worked on formatting SSRS reports using the Global variables and expressions.
- Created parameterized reports, Drill down and Drill through reports using SSRS.
- Used Execution Plan, SQL Profiler and Database Engine Tuning Advisor to optimize queries and enhance the performance of databases.
- Optimized query performance by creating indexes.
- Write T-SQL statements for retrieval of data and performance tuning of TSQL.

Responsibilities:

- ETL package Development in SSIS.
- Report development in SSRS.
- Database Design.
- Understanding the business problem, identifying relevant data, gathering and summarizing data meaningfully.
- Performance Tuning/Query Optimization in SPs.
- Complex Query Design, Stored Procedure Writing.
- Function Writing, Trigger Writing.

Environment: SQL Server 2008 R2/SSIS/SSRS, Microsoft Business intelligence development studio, **SQL Server 2008 R2**

Oil Field Services, Germany

Sep'10 to Sep'11

Role : Asp.Net Developer

Project Description: OFS is web based application that keeps track of products between stores/rigs/well. OFS application offers clients a secure and efficient management tool in processing their business. OFS application is built as an n-tier Web based application. The Application User interface is built as Web application using ASP.NET Master Pages, AJAX controls and JavaScript. To Support SOA based architecture, Business Web services in middle tier is used and it acts as core business engine. The web services consume/ provide the Business Entities. Business Entities are used for passing the data across the tiers. Database tier is the oracle database. To maintain application portability across databases, data transactions are handled in business services tier.

Responsibilities:

- Coordination and handling tasks among team members.
- Designing of HTML pages, JavaScript validations.
- Assist Project Manager in Feasibility study and requirement Analysis phase.
- Implement web service as business layer.
- Deploy application on Client Server.
- Designing Database, Creation and implementation of View, Stored Procedures.
- Documentation for coding.
- Configuring, deploying Database and web application for testing environment.
- Code review during different stage of development life cycle.

Environment: Web, Framework 2.0, ASP.Net 2.0, C#, Web Service, TFS, Visual Studio 2005, JavaScript, **Oracle 10G**

Insurity, Hartford, CT
Project: Insurance Policy Decision
Role : Asp.Net Developer

Jan'10 to Aug'10

Project Description: NGA (Next Generation Application) is a web-based insurance-product of one of the Famous American product-based company in commercial lines. It allow their clients to do policy-administration as,

- Automated rating and issuance support for all major commercial lines of business.
- Full policy lifecycle transaction support
- Full maintenance of ISO and customer-specific rates, forms, and rules changes.

Responsibilities:

- Requirement Gathering of defects.
- Onsite Coordination.
- Estimation of defects.
- Defect Fixing
- Code review

Environment: Web, Framework 2.0, VB.Net, VB6, **SQL Server 2005**

Insurity, Hartford, CT
Role : Asp.Net Developer

Mar'09 to Jan'10

Project: CHealth

Project Description: CHealth is a Collaborative Consumer Engagement, which is based on Microsoft connected Healthcare framework. It empowers consumers through its Wellness module and at the same time provides an integrated platform wherein Payers, TPAs, Brokers, ASOs & Providers collaborate seamlessly. With an innovative business model cHealth is supported by SOA based architecture. To Support SOA based architecture, Business Web services in middle tier is used and it acts as core business engine.

Responsibilities:

- Coordination and handling tasks among team members.
- Designing of HTML pages.
- Assist Project Manager in Feasibility study and requirement Analysis phase.
- Deploy application on Client Server.
- Database Design, Creation and implementation of View, Stored Procedures.
- Taking Backup and Restore Database.
- Implement Web services as business tier.
- Documentation for coding.
- Configuring, deploying Database and web application for testing environment.
- Code review during different stage of development life cycle.
- Developed HL7 translator.
- Write Store Procedure to implement HL7 translator.

Environment: Web, Framework 2.0, Asp.Net with C#, **SQL Server 2005**

Munich Re, Princeton, NJ

Aug'07 to Feb'09

Project: AutoFac

Role : Asp.Net Developer

Project Description: AutoFac is a web-based application that allows Re-Insurance clients to conduct Property and Commercial Facultative re-insurance business operations with MRAM. AutoFac application offers clients a secure and efficient management tool in processing their business. AutoFac application is built as an n-tier Web based application. The Application User interface is built as Web application using ASP.NET Master Pages, AJAX controls and JavaScript. To Support SOA based architecture, Business Web services in middle tier is used and it acts as core business engine. The web services consume/ provide the Business Entities. Business Entities are used for passing the data across the tiers. Database tier is the oracle database. To maintain application portability across databases, data transactions are handled in business services tier.

Responsibilities:

- Designing of HTML pages, JavaScript validations.
- Programming for Referral, People, and System admin, Organization, Policy, Quotes Location Upload and Reporting.
- Identifying the Functionality of the Application.
- Designing of framework.
- Designing Database Model, Creating Views and Stored Procedure.
- Requirement gathering phase and UAT.
- Extensive work on web services as business tier.
- Deploying application for testing environment.
- Documentation for specific modules
- Integration and System Testing
- Programming in XSLT for location upload

Environment: Web, Framework 2.0, Microsoft Framework 2.0(ASP.Net) with C#, **Oracle 10G**

Education:

- Masters in Computer Application, 2004, India.

Certification:

- Microsoft .Net Framework 2.0 Application Development (70-536).