

# Mahesh K. Yadav Ph.D.

e-mail: mahesh.yadav@tcs.com,  
mahesh\_y@hotmail.com  
phone/whatsapp + 62-82311332512

## ◆ OBJECTIVE

To work in a role primarily as Data scientist in an environment that is friendly, fast paced and give opportunities to solve challenging problems. The tasks should involve running various ML algorithms for developing optimized models, understanding data, dealing with missing data and outliers, generating graphics for visualization, learning and experimenting with cutting edge developments. The preferred language is R/Rstudio and Python/Jupyter with open source libraries.

I could also partially supplement my experience with data modeling, SQL query, data base tuning, and documentation as the situation demands in the project.

I am currently looking for position as data scientist in Singapore within 4-6 months.

## ◆ DATA SCIENCE RELATED EXPERIENCE

- I. **Education/Self Executed Project:** (2016-2017): I have self-taught and self-executed Data Science project with real-life data that is show cased in: <https://mahesh-y.github.io/> . This is result of learning data science for the past two years holding my current job. I have good command on the data mining methods in line with book: Element of Statistical Learning T. Hastie, *et. al.*

I have participated in Kaggle competitions (without making significant ranking but got great deal of experience). I learnt to deal with large data sets, making use of multiple CPU for parallel processing and executing 10x faster algorithms such as XGBOOST and RF/GBM from H2O package. I used personal computer with 32 GB RAM.

- II. **Professional Experience:** (2017-2018): In the current data warehouse project, I have created set of testing reports for cross-checking important dimension and fact entities of new data warehouse to the legacy data warehouse to check the validity of data. The testing reports were generated in html from Rstudio/R. The floating-menus showed difference in the data for each column. This method helped compare large dataset in a quick and precise manner. This was vital to check the data accuracy after each day load and used during the Data Warehouse SIT and cooling-off period
- III. **Complex Requirements: (2003-2017)** Involved and played critical role in various risk-management projects that involved complex business understanding – Common Purchase Point in Visa Card Fraud detection, Actuarial related metrics in Insurance domain, performance monitoring of scoring models in Banking.

I have good understanding of Black-Scholes equation and use this knowledge in my personal investments which is usually as an Option seller and managing risk. (Created programs in VBA Excel/Mathematica to download Options data and process using Black-Scholes equation)

- IV. **Ph.D. with Data Analysis: (1997)** My Ph.D thesis in Experimental physics involved data analysis of large data taken from the particle accelerator in Mainz. The work involved analysis of large data sets that had to be cleaned and modeled based on software provided by CERN. The thesis is available at following link:  
[https://mahesh-y.github.io/thesis\\_mahesh\\_yadav\\_1997.pdf](https://mahesh-y.github.io/thesis_mahesh_yadav_1997.pdf)

#### ◆ Technical Skills

- R/Rstudio (intermediate level), Python/Jupyter (beginner level),
- Unix, Cygwin, Awk
- Excel with VBA (Advanced), Notepad++
- Data Modeling (generic 3NF, 3NF, OLAP), SQL (Advanced)
- Graph Database (Neo4j) (beginner level)
- RDBMS – Teradata, Oracle and SQL Server

#### ◆ IT BACKGROUND

I have been in IT industry for 20 years and worked for various high profile companies – Aviva, General Electric Capital Solutions, Deutsche Bank, Standard Chartered Bank, Visa International Asia Pacific. I have played multiple roles as project manager, consultant, data/solution architect, and business/system analyst. Most of the work was focused in Data Warehousing projects.

I also very good communication and documentation skills that helps in fewer and productive interactions with users.

#### ◆ EXPERIENCE SUMMARY

Client	Domain	Location	Role	Technology	Period	Year
Allianz	Insurance	Jakarta, Indonesia	Data Modeler/ Data Analyst	Oracle, SQL, Erwin, Sparx	1.5 year	16-17
Deutsche Bank	Banking	Singapore	DBA Support	Oracle, PL/SQL	2 year	13-16
Best Buy & Supervalu	Retail	Minneapolis, MN, USA	Data Architect/ System Analyst	Teradata/ Informatica/ Oracle/ Trillium	3 years	09-13
Aviva	Insurance	Norwich, UK & India	Functional Analyst/ Data Architect	Teradata, Informatica	8 months	09
Avery-Dennison	Manufacturing	Brea, CA, USA	Consultant	Cognos	1.5 months	08
Cadbury Schweppes	Manufacturing	Melbourne, Australia	Project & Data Manager	SAP 3.0/ECC6	5 months	07
Cummins	Manufacturing	Columbus, IN, USA	DW Consultant	N/A	2 months	06
GE – Capital Solutions	Financial Services	Danbury, CT, USA	IT Analyst	Oracle, Business Objects	6 months	05-06
Canadian Bank - CIBC	Banking	Toronto, Canada	Business Analyst	SQL Server, Actuate, Matlab	1.5 months	04
Visa Int. Asia Pacific (Visa)	Financial Services	Singapore	PM/Data/Solution Architect	SQL Server 2000 MicroStrategy 7i	1.5 years	03-04
Dell	Retail	India	Solution Architect	SQL Server, ASP	7 months	02-03
AT&T	Telecom	USA	Data Architect/ Developer	Sybase, Oracle Visual Basic, PB	3 years	98-01

## ◆ PROFESSIONAL EXPERIENCE

**Tata Consultancy Services, Indonesia**                      **Role: Data Modeler/Analyst**                      **2016-2018**  
Allianz Indonesia – From Aug 15, 2016 to present, I have worked as Data Modeler and Data Analyst for building the data warehouse for Individual Life, Individual Health, Group Life, Group Health and other LoB.

There were four data modelers in the project. My role was to create data mapping for the ETL team in Individual Life. Once the data modeling was done, my role was to check the data integrity of data warehouse and match with the source query and the legacy data warehouse. The data model was generic 3NF. I found this generic 3NF quite fascinating as it had architecture pattern of graph database (Neo4j).

Once the modeling was finished, I switched my role as Data Analyst and had the role to check the integrity of Data Warehouse comparing to Source query and legacy Data Warehouse. I created various reports in R that would compare major dimensional entities like Agent, Policy and Facts. The RStudio/R helped in creating nifty reports that could help making showing difference in data by columns. For large set of data, we used random chunks of data from heterogeneous sources for comparison that could be generated in minutes on demand. Many such random chunk comparisons give quantifiable statistical measure of confidence for data integrity.

**Tata Consultancy Services, Singapore**                      **Role: DBA Support**                      **2013-2016**  
Deutsche Bank, Singapore – From Sep 15, 2013 to June 30, 2016, I was in the DBA Support role to improve the performance and tuning of DBDI – a global payment application for corporate customers.

The DBDI application has five instances of application managing different geography and one common reporting server which were all used in customer transactions. Since it was a more than 15 year old application, it had many performance problems due to growth of data and increasing customer base. As a result the extensive queries were tuned, and unnecessary data were purged from the table and some processes were optimized.

In this role, I developed the plan for purging and tuning, developed the code, managed testing and managed the implementation. Following task were done in detail.

- Analyze and identify the expensive queries through AWR report. These expensive queries resided in Java code of the application, Informatica mappings, Stored procedures, Alert queries managed by L2 team.
- Identify the root cause and develop a solution. The execution of queries were either taking long time, having large Buffer Gets or being called unreasonably large number of times.
- Solution consisted of developing code for one Time ad hoc purge followed by code changes for regular purge of tables, creating new indexes, rewriting queries, removing snapshots of Materialized views, change of Informatica workflows, removing Data Cache in Informatica mappings.
- Plan and review for testing.
- Managing the deployment through Global Change Management application.

**Tata Consultancy Services, USA**                      **Role: Project Manager/Data Architect**                      **2012-2013**

Best Buy, MN, USA – From Aug 15, 2012 to July 31, 2013, I was managing and playing the role of Data Architect for two large projects. This was the first project for TCS as a vendor, I managed the entire life cycle of the project that gave client the confidence that TCS can handle large scale EDW projects. I played a pivotal role in managing, architecting, and identifying near and long time risks. It was a zero defect deliverable. It was one on my career best performance on the deliverable.

The project involved replacement of manual monthly budget feeds of various channels like PAC, MDC,



**Tata Consultancy Services, UK                      Role: Functional Analyst, Data Architect                      2009**

Aviva (Norwich Union Insurance), Norwich, UK – From Jan 12, 09 to July, 09 (7 months), I was involved in migration of functionality/reports from Main Frame/DB2 platform to newly built enterprise data warehouse in Teradata.

My responsibilities include:

- Gathering requirement from SME relating to complex reports – Earning, Actuarial, IBNR derivation.
- Create functional requirement for offshore development
- Data Modeling for Teradata specific architecture
- Create solution design for the offshore development
- Monitor and help offshore to resolve the requirement complexities
- Review and develop some of the key component to help jumpstart relating to architecture, database, Informatica Mapping, and Business Reports,

The role is quite challenging in coordinating with offshore team as the project is based on Agile methodology where the requirements are not finalized.

**Tata Consultancy Services, South Africa                      Role: BI Consultant                      2008**

CELL-C, Sandton, South Africa - From March 14 to May 14, 08 (2 months), I was involved in laying down the strategy to establish the Business Intelligence Competency Center (BICC) that is the first step in larger effort to implement the enterprise wide BI solution.

Main deliverable was proposing the organization structure, having roles and responsibilities, providing assets and addressing the problem points related to process and governance

**Tata Consultancy Services, USA                      Role: BI Consultant                      2008**

Avery-Dennison, Brea, CA, USA - From Jan 2, 08 to Feb 13, 08 (1.5 months), I was involved as a consultant for the setting up of Cognos Competency Center at Avery-Dennison. Avery-Dennison's strategy was to align from highly de-centralized to the centralized approach for information delivery and have a centralized competency center established.

The main deliverables from this study were charter & framework document, Service level document to Business units and presentation for the director to be given to business units. The deliverables contained the governance structure, roles and responsibilities, services provided to business units, interaction with other parties, sample illustrations of various process for establishing competency center. In addition, various documents related to methodology, process and services were given to build a matured competency center in relatively short time frame.

**Tata Consultancy Services Asia Pacific Pty Ltd, Australia                      Role: Data/Project Manager                      2007**

Cadbury Schweppes, Melbourne, Australia - From April 2 to August 31, 2007 (5 months), I was involved as Project Manager/Data Manager for migration of Confectionary data from SAP 3.0 to another existing Food & Beverage SAP ECC6 system. The data migration was related to FI/CO, SD, MM (Finance, Costing, Sales & Distribution, and Material Management). Most of the data migrations were either manually done or automated with LSMW (Legacy System Migration Workbench).

My role was the following:

- Overall co-ordination of manual and automated LSMW development between functional groups and

- onsite/offshore developers
- Creating detailed project timelines
- Resource estimation
- Updating weekly status report
- Detailed plan on data load schedule
- Highlighting the risks to Process Board members and plan on mitigation.
- Creation of process documents for data quality and data load

These measures were important to the success of the project as the development and migration were done with an aggressive time line of five months.

**Tata Consultancy Services, US                      Role: BI Consultant                      2006**

Cummins Inc, Columbus, IN, USA - From May 1, 2006 to June 30, 2006 (2 months), I was involved as Business Lead/Analyst for Business Intelligence Health Check of current data warehouses and prepare the Blue Print that will produce the road map for future enterprise data warehouse. The work required gathering information related to business needs, governance, and existing systems. This was followed by analysis phase and finally giving recommendation for the architecture, governance and road map for next three years. Cummins manufactures and markets the diesel and natural gas powered engines and parts. The subject area of focus was supply chain management.

**Tata Consultancy Services Asia Pacific Pte Ltd, Singapore    Role:Data Analyst                      2006**

Deutsche Bank, Singapore - From July 15, 2006 to Dec 15, 2006 (5 months), I was involved as Data Manager/Analyst for Financial System Renewal project. My role required consolidation of the reference tables, evaluate environments for hosting reference tables, propose various architecture options for optimizing data flow and error handling. The reference table pertained to data that was required by cross Streams or at interfaces. The project organization is structured in various streams – Portal, SAP R3, SAP BW, and dbFeeds (ETL). The deliverable was Data Dictionary, hosting decision paper, architecture option paper, and Technical Design Document. My role also required to mediate between various stream owners to reach a decision on architecture and table design.

**Tata Consultancy Services, US                      Role: BI Analyst                      2005-2006**

General Electric Capital Solutions, Danbury, CT, USA - From Sep 24, 2005 to April 30, 2006 (6 months), I was involved as Project Manager/IT Analyst for a highly utilized data warehouse. I was involved in maintenance and enhancement of the system that had Oracle database and BO reports. The data was loaded into the data warehouse with Informatica. The business of General Electric Capital Solutions focused on financial services – business lending and business equipment leasing. The data warehouse had the integrated data for the entire cycle of the deals as described below:

- The sales system (Siebel) where the opportunities were created, negotiated and approved.
- The risk system where the deals would be analyzed and then awarded or rejected
- The booking system that would book the deal in installment as given by the risk system

My role was the following:

- Single point of contact with Business for Data Warehouse issues.
- Finding root cause, assigning required development group for issue resolution.
- Keeping log of all the pending and resolved issues.
- Requirement gathering for all the enhancement requirements
- Project plan, resource estimation, and impact analysis for enhancements.

- Helping in UAT and production rollout

Having the integrated data created a high demand for the data warehouse from the entire spectrum of the business.

**Tata Consultancy Services Asia Pacific Pte Ltd, Singapore    Role: Business Analyst                    2005**

Companhia de Telecomunicações de Macau S.A.R.L (CTM) - From May 11, 2005 to June 22, 2005 (2 months), I was in a role of Business Analyst for gathering user requirements for developing the data warehouse. The user requirement was to be translated to put in a Request of Proposal (RFP) document for inviting the vendors for developing data warehouse. The requirement resulted in over 30 interviews with users from all business units – mobile, internet, Fixed line, accounting, finance, Fraud Management and Risk Assurance. I prepared Meeting minutes for each of the interviews and then wrote the user requirement document. Additionally, since users were not very familiar with analytical report, I gave a demonstration of the power of analytical report from pivot table in MS Excel (based on one of their requirement).

**Tata Consultancy Services Asia Pacific Pte Ltd, Singapore    Role: PM/Data Architect                    2005**

Standard Chartered Bank, Singapore - From Jan 24, 2005 to Apr 15, 2005 (3 months), I was involved as a project manager/Data architect leader for six member team catering to scenario calculations for Basel II compliance project. This was end-to-end project that involved enhancing the existing data mart and existing Expected Loss Calculation Engine to run for various Scenarios. Each Scenario calculates vital figures like EAD, LGD depending on the input data set by the business. The Scenarios were to be managed by developing web based application accessible to business users.

There were many interfaces on the project that was required to be considered: (a) controlling the flow of job-execution by iLog rule and Unix based polling mechanism, (b) DB2 database changes (c) Cognos reports changes (d) changes to custom developed JSP screens which controlled the meta data and reference table.

My role involved the following:

- Create project plan
- Understanding the requirement
- Creating the Business Requirement Specification and System Requirement Document
- Understanding various interfaces such as
  - web application for Business inputs
  - control for job execution
  - Impact on downstream application such as Cognos reports
- Evaluating risks and risk mitigation

The Calculation Engine was implemented in DataStage modules and new screen were made using JSP. The project was delivered in time.

**Tata Consultancy Services, Toronto, Canada                    Role: Business Analyst                    2004**

Canadian Imperial Bank of Commerce - From Nov 12, 2004 to Dec 29, 2004 (1.5 months), I was involved as a consultant in a five-member onsite team for Risk Quantification group, CIBC. The Risk Quantification Project is a part of the Retail Work stream created under the auspice of CIBC Basel Program to satisfy the requirements of the Advanced Internal Ratings (AIRB) approach under Basel II.

Risk Quantification group had developed many scoring models for evaluation of various risk parameters. The group has also developed methodology for performance monitoring of scoring models. The performance monitoring involving various statistics Komlogrov-Smirnov test, Receiver Operating Characteristic,

population stability index, and bootstrap method for calculating standard error. The risk parameters included Probability of Default (PD), Loss Given Default (LGD) and Exposure at Default (EAD).

Our team's objective was to understand the business requirement for putting the prototype processes for evaluating risk parameters and performance monitoring of score cards in production environment. This involved in developing data model, integration with existing data warehouse, optimization of evaluation of risk parameters and statistical calculations, selecting tools for loading, transformation of data and creating reports which meets the business requirement. The deliverables were business requirement specification and system requirement specification documents.

**Tata Consultancy Services, Hyderabad, India**

**Role: Business Analyst**

**2004**

Tata TeleServices Limited – From Oct 4, 2004 to Nov 1, 2004 (1 month), I was involved as a consultant for developing the strategy for data warehouse solution. TTSL volume has grown at a fast pace in a relatively short time. This increased the burden on existing data warehouse which was not very scalable. The objective of the project was to consider integrating information from various source systems in a central repository, choosing the database that could handle the current and future volume requirement, selection of reporting tools, and developing a road map for developing the data warehouse.

**Tata Consultancy Services Asia Pacific Pte Ltd, Singapore**

**Role: Solution Architect**

**2003 to 2004**

Visa International Asia Pacific - From January 2003 to August 2004 (1.5 years), I was involved as project manager and solution architect for CPP project in Risk Management, Visa International - Asia Pacific. The aim of CPP Project was to identify a specific type of fraud called Skimming. The CPPs (Common Purchase Points) is identified by analyzing the patterns (rules) and various metrics on six month of legitimate authorized transactions.

I interacted with users in Risk Management and developed the logical data model (3NF) with about 80 tables to support their requirements. This exercise helped in streamlining the thoughts of clients into concrete and consistent requirements. Also written most of the process related Functional Requirement Specification (FRS) document.

My role was the following

- Requirement gathering and preparation of Functional Requirement Specification (FRS) document
- Developing Logical Data Model
- Preparation of project plan
- Single point of contact with business and development team
- Evaluation of risk and risk mitigation plan
- Coordinating with offshore for development
- Helping develop detail user acceptance testing plan
- Coordinating the UAT with business users and development team
- Coordinating with release engineer and development team to rollout in production
- Managing cool down phase

The logical data model and FRS served as a foundation for software development at offshore in India. The project was completed on time.

**eClerx Services Pvt. Ltd, Mumbai, India**

**2002 to 2002**

DELL (Software and Peripherals) - From April 2002 to October 2002 (7 months), I was project manager and solution architect for DELL-USA. The primary goal was to assure consistency of pricing information in various databases and DELL web sites.



I played a pivotal role in the startup company to develop data mart, design reports and design Ticket Tracking System (TTS) to assure consistency of pricing information. All the issues were logged in TTS and resolved based on SLA. I was involved in the entire project lifecycle of the project.

**Zyga/FutureNext Consulting, Inc., New Jersey USA**

**1997 to 2001**

AT&T Solutions (Network Solution Provider) - From June 2000 to March 2001 (one year), I was in the role of solution architect for enhancement and maintenance of the client/server application. The in-house application in Power Builder and Oracle application managed the invoices of third party telecom operators to payments from corporate/retail customers.

AT&T (Telecommunications firm) - From May 1999 to February 2000 (10 months), I was involved in client/server startup project for managing orders. The application was designed in Visual Basic with Oracle database. I was responsible in designing Control Desk part of the application, which gave the user access to make changes to about 40 reference and non-reference tables. I created many stored procedures in a PL/SQL package in server side for optimizing the client/server architecture. The Visual Basic application had class modules enclosing Data Services and User Interfaces to make the code manageable and easily adaptable to changes. The relational database was created and maintained in Erwin. The Microsoft Visual Source Safe was used for managing Visual Basic project files between developers. I also designed the reports using Crystal Reports.

AT&T Solutions (Network Solution Provider) - From July 1998 to May 1999 (1 year), I was involved in CIS (Client Invoice Service) project for loading the vendor invoice data (hierarchal database) to 3NF relational database. When I joined the CIS team, it was a fully functional client/server application serving AT&T for over two years using Power Builder and Sybase. My work was focused in enhancement of the application and to automate the loading of invoice data from files provided by vendor to CIS database. I designed the module reading the source data into few ASCII files, which were uploaded in Sybase by BCP process. The data was then validated at various checkpoints. Once the data passed at all the checkpoints, it was loaded to CIS database. I was responsible in designing the data model, loading procedures and validation procedures. The automated process could load most of data successfully.

Hoffmann-La Roche (Manufacture of pharmaceuticals and diagnostic systems – Tox. and Path. Dept) - From May 1997 to April 1998 (1 year), I was involved in Data Warehouse project in which I was responsible for getting files from various source systems to be loaded into the Data Warehouse. The extraction of source required the knowledge of FORTRAN and C. The ToxPath project was a complete Data Warehouse project, which involved understanding the legacy source files, having Oracle database for storing the data, and using OLAP tool from Microstrategy. I designed Visual Basic application to automate the extraction of files from legacy systems (VAX/VMS) by RFTP, to execute the transformation process (store procedures in Oracle), and to execute the load process to Oracle database. The two databases - Star Schema for decision support and third normal form for OLTP- were created and maintained in Erwin.

**◆ EDUCATION**

**Ph.D in Physics, Rutgers, The State University of New Jersey, USA**

**1986 to 1997**

From September 1986 to May 1997, I was a graduate student of Physics in Rutgers University. During the graduate studies and thesis work in Experimental Nuclear Physics, I gained experience in data acquisition system, data analysis, numerical methods, and statistical methods for understanding data. I analyzed the data in FORTRAN, C, Unix shell scripts in Unix and VAX/VMS platform. I used AutoCAD in MS DOS for creating the detector (which had the shape of truncated icosahedron). I also used various CERN library packages like MINIUT, HBOOK, PAW, and GEANT for statistical analysis, plotting histograms, n-tuple

analysis, and Monte Carlo simulations respectively. My thesis is available at [https://mahesh-y.github.io/thesis\\_mahesh\\_yadav\\_1997.pdf](https://mahesh-y.github.io/thesis_mahesh_yadav_1997.pdf)

**M.Sc. in Physics, University of Delhi, India**

**1983 to 1986**

During September 1983 to May 1986, I was a student of Physics in Delhi University. I got an average of 67% and placed in first division.

**B.Sc. in Physics, Hans Raj College, University of Delhi, India**

**1980 to 1983**

During September 1980 to May 1983, I was a student of Physics in Hans Raj College, Delhi University. I got an average of 68% and placed in first division.

**◆HOBBIES & INTEREST**

- Take great interest in maintaining good health. I jog about 15-20 miles/week that act as great stress reducer and keep me active and fit.
- I am very interested in cooking vegetarian food that is good for health and not necessarily on taste.
- Based on opportunity I love to hike and travel. I have hiked many national parks in United States and done lot of day hikes in New Jersey/New York, United States.