

# Sanjay Gilki

Contact: +44 07760513983 | e-mail: [sanjaygilki@gmail.com](mailto:sanjaygilki@gmail.com)

---

Johns Hopkins University Certified Data Analyst, IBM & Open Group Certified Data Scientist, Microsoft Azure & Azure Data Certified, AWS Partner and IBM Certified Architect. Seeking to Increase Data Efficiency by Recommendations & Fore-sights using Descriptive, Diagnostic, Predictive and Prescriptive Analytics. Key Achievements include 100% RU Billing Accuracy, \$ 5 Million Savings YOY by proliferating FDMS to more than 25+ countries, Optimized SQL script's compute time from 23 min to only 4 minutes , 60% Reduction of Infra tickets to avoided server failures and downtimes and Automated Several manual reports with Dashboards with Key ML Insights. Innovative, Highly skilled in Machine Learning, Data Visualization Expert and Creative Thinking.

---

## IT AND TECHNICAL SKILLS



- **Programming Languages:** R, SQL, Python, VB6
  - **Data Visualisation:** R shiny, Tableau, Power BI, Cognos Analytics, AWS QuickSight Q, Kibana.
  - **Databases/ETL/BI:** SQL Stored Procedures, Custom SQL, Abinitio , PostgreSQL, Teradata, SAP HANA, MS SQL Server , SOQL, IBMDB2, Oracle.
  - **Cloud Platforms:** AWS (EC2, S3, Amazon Forecast, AWS Athena, QuickSight, QuickSight Q, Spark cluster setup using AWS EMR. AWS EMR Setup using Dockers and Containers. GCP (Google Storage, Google Data Studio, BigQuery ML) Built-In ML ARIMA and DeepAR+. Azure (Azure DevOps (ADO) Wiki Repositories, DAX, Power App, ADO Boards for Sprints Salesforce), Customer Analytics using Snowflake & Tableau Dashboards. PowerShell, CICD, PySpark, Schema Builder, SQL, Mural
  - **Machine & Deep Learning :** Prediction and Forecasting, ARIMA, Linear Regression, NLP, Text Mining, TensorFlow, Tesrat, RapidMiner
  - **Architect :** Data Dictionary, Data Modelling, Mapping Document, Gap Analysis, Data Flow Diagrams, Designing and Proposing As-IS and To-Be Architectures, Enterprise Architect, Visio, Ervin, LDM, PDM
  - **Version Control:** Github, IBM Box and Jira.
  - **FinTech:** P&L Analysis, Quantitative Finance Research, FTE and Revenue Forecast, Prediction, RU Billing, Accounts Payable, Account Receivable and FDMS.
  - **Quantum Computing:** Qiskit, IBM Box, GitHub.
  - **Designs Tools:** Mind Mapping, Mural, Bubble Diagrams, Silver Run, LDM, PDM, Indexing, Explain Plan, SOD, and Component Diagram.
  - **Influencer/Digital Marketing:** Product Research, B2B, B2C, Product Listing, Video and Webpage Ad Creation, Vidnami, Google Ad, Youtube, SEO, email marketing, Affiliated Marketing, Managing Ad Campaigns, Optimizing Shipment cost.
- 

## POCs (Individually Designed and Developed):

- Exploratory Data Analysis (EDA) using PySpark – Spark Cluster Setup, RDD & SQL
  - Customer Analytics Cloud using Tableau Dashboards from Snowflake Data.
  - Predicting and Forecasting using GCP Bigquery & Amazon Forecast (Deep AR+ & ARIMA), AWS S3, Athena, QuickSight and SQL
  - Read data from a Json file and streaming data using PySpark.
  - Create quantum computing program using Qiskit and Python
  - Spark cluster setup using AWS EMR and Dockers and Containers.
  - EDA using Amazon QuickSight Q
  - CICD for Dev Ops , ICFF
- 

## EXPERIENCE AND ACHIEVEMENTS

I have had the opportunity to work across multiple industry verticals which including Telecom, Auto, Health Care, HR Oil, Gas & Energy services.

**IBM India Pvt Ltd**  **2012 – 2021**  **Bangalore, India**

---

**Project: Global Supply Chain Analytics, Intelligence and Report | Client: Illumina****Role: Data Engineer**

**Description:** Illumina wanted to optimize their Supply Chain and reduce the time to deliver, improve their self-service BI by increasing the customer satisfaction and reduce the backlogs.

Worked as Data Engineer with Illumina's all the Business and Support functions, Report Owners, SCM Analytics Delivery Teams, SCM Director, SAP Technical Teams and all the stakeholders for this Project.

- Quickly Understood the Business reporting requirements from various business functions.
- Optimized the custom extract SQLs for Tableau Dashboard Development.
- Design, Develop, Test, Validate and Deploy the Dashboards.
- Granted Dashboard access to end users to validate and promote it to production.
- Worked with SAP Technical Team accommodate new Self-service BI requirements to download the reports.
- Wrote ARIMA ML model for prediction and forecasting any Time Series Data for optimizing the SCM operations. e.g., multi plant orders.
- Designed and Developed Real Time and Near Real Time Dashboards, Reports for Middle Level Management.
- Create SQL scripts, performed the unit testing and capture the test results in the document and shared.

**Key Accomplishments:** Helped in Optimizing the Illumina's Global Supply Chain Teams to Plan, Predict and Forecast Multi Plant Orders, also helped in Reducing their Shipment Delivery Time and in gaining huge Customer Satisfaction. As a senior BI Consultant I, proposed CRISP DM methodology. After implementing this process, the execution of the projects became easy and smooth and thus reduced the reporting backlogs and which quickly helped to focus on key reporting deliverables.

**Technologies used:** R, Python, Tableau, SAP HANA, Custom SQL, SSRS, Amazon Workspaces and JIRA.

---

**Project: CPaaS - IBM Talent Platform | Client: UHG | Role: Senior Data Lead**

**Description:** IBM wanted Talent Platform on Cloud with SaaS offering that helps companies accelerate time to hire and wanted to optimize the Hiring process of Recruiting, Sourcing & Care Centre for Middle Level Management

- Worked as Sr Data Analyst with Data Team, Service Team, Data Science Team, DevOps Team, UI Team, Testing Team and all the stakeholders.
- My responsibility was to optimize the complex SQL of Material views which were taking more compute time in data processing, Data Governance & Deployment Process Improvement, POC: Spark Cluster Setup, Data Model, Data Mapping for Prediction and Forecasting for ML insights, Model Diagnostic, BOT Simulation, CI&DC implementation, Restore and Backup, Recommendations of To-Be Architectural diagrams with Tools and Technologies.
- Understood the Business requirements, Strategize Data Science and Analytics Insights and quickly performed the EDA and shared the key insights in Tableau Dashboard to all the stakeholders.
- Designed Data Model, Data Mapping Documents, Creation Architectural Flow Diagrams.
- Research on new open sources and cost-effective solutions for huge data processing in less compute time and visualization tools for insights.
- As a Sr. Data Lead Reviewed the Implementation Plan of DDL, DML and SQL objects.
- Defined the procedure for CICD implementation of Data Team by creating the folder structures and defining the process for other team members to follow approval workflow for smooth migration.
- Conducting technical interviews and assess the candidates to expand the team, Giving the KT and ramping then quickly onboard and deliver, setup to reduce the com Review of code and various Documents. Etc.
- Creating Entities, ER Diagrams, created data models, mapping, and data dictionary and configuration documents.
- Create SQL scripts, perform the unit testing and capture the test results in the document and shared through JIRA. On available change window promoted the objects to Test and other environments.
- Reviewing the open source visualization tools which are strategically aligned to the projects.

**Key Accomplishments:** Clients are happy with new capabilities of IBM Talent Platform, able to track activities by alter systems, and they were able to plan their billing and FTEs by using Predicting and forecasting, Compute time of material views brought down from 30+ min to 4 minutes and Vey Quick Deployment time taken due to CICD Implementation.

---

**Technologies used:** Tableau, R, SQL, PostgreSQL, CI & CD, DevOps, Jupyter Notebook, OpenShift, Docker, Containers, PowerShell, Maven, Jenkins.

**Program: IT Operations Analytics (ITOA) | Client: Kaiser Permanente |**

**Role: Data Engineer, Data Scientist @ Analytics and Insight Services (AIS) - GTS Labs | Duration: Oct 2018 to Apr 2020.**

**Project: Performance Capacity Management (PCM) Analytics:** PCM team was using XL based application to compute the plan the capacity, which had limitation of data processing and it was manual process to generate the reports and share.

- Worked as Data Engineer with Client, Customers, SMEs, PCM Capacity Analysts, Platform Team, SDMs, BAs, and Delivery Project Executive (DPE) Operation Head and even with 3rd part monitoring tool Vendors and all the stakeholders.
- My responsibility was to understand the business and convert the XL based application to PCM Analytics Dashboard with Analytics Insight for Middle Level Management and avoid manual process of creating and sharing the reports to key platform teams and management teams
- Gather all the raw data, XL application, sample reports and email formats.
- Understood the Business and Data (Data Science and Analytics) requirements and quickly performed the EDA and shared the key insights in Tableau Dashboard to all the stakeholders.
- Strategically designed the data model and wrote separate python programming for CPU and Memory (AIX, Physical & Virtual Servers). Computed the business logic (CPU & Memory  $\geq 80\%$  &  $\geq 4$  Days) to identify the breached Servers and generated the final pre-processed data set from sub data sets.
- Designed and Developed Dashboard with Insights using that pre-processed de-normalized data set.
- Validated and Tested the Dashboards, captured the test results in the document and share with all the stake holders by attaching to the JIRA tasks. And update the accordingly update the data model documents, other user guide documents and updated the details in assigned JIRA task as well.
- Presented and gave PCM Analytics Dashboard demo to client, management, and end users, educated them on the adoption of the dashboard.
- Collected end users list for dashboard access and planned for the production migration. Documented all the project related documents (like user guide documents).
- Worked closely with Tableau Admin, depends on the available change window, planed, and promoted dashboards to Test and other environments.

---

**Key Accomplishments:** Saved weekly 12 hours manual work of report generation and avoided legacy XL application. Made centralized view for PCM insights with single version of the truth (SVOT). Which helped in CPU and Memory Upsizing/Degradation for Performance Management, Trend Analysis and Capacity Modelling too. Capacity Analysts can quickly do the Performance Monitoring, Performance Analysis, Performance Tuning, and Performance Reporting to all the stakeholders and, they can download the reports Monthly with Drill-through/Download and email options to share with the management thus ease consolidation of various source data and compute the logic. And in the second phase developed PCM Analytics Dashboard for historic data.

**Technologies used:** Python, Tableau, Cognos Analytics, SQL, TIBCO JasperSoft, IBM Box and GitHub, CRISP-DM, JIRA.

---

**Project: Dynamic Automation (DA) Analytics:** DA team was unable to deep dive and figure out impacted CI with High SWAP, High CPU, High Memory & High Page details from DA logs which was logged by the automata jobs.

- Worked as Data Engineer with Client, Customers, SMEs, DA Team, Platform Team, SDMs, BAs, and DPE Operation Head and even with 3rd part monitoring tool Vendors and all the stakeholders.
- My responsibility was to engage DA team and SMEs, understand and gather the business and data requirements to identify the process and metrics causing these High SWAP, High CPU, High Memory & High Page issue tickets.
- Extracted and collected the data from these High SWAP, High CPU, High Memory & High Page issues from ServiceNow ticketing tool.
- Strategically designed the data model and wrote separate NLP python programming to extract the required information using regular expressions for each High SWAP, High CPU, and High Memory

& High Page issues to extract the metrics from the long text and computed the values and captured them in pre-processed data set in de-normalized fashion.

- Designed and Developed Dynamic Automation Analytics Dashboard with Insights (e.g., Number of java processes were impacting the server) using that pre-processed de-normalized data set.
- Validated and Tested the Dashboards, captured the test results in the document and share with all the stake holders by attaching to the JIRA tasks. And update the accordingly update the data model documents, other user guide documents and updated the details in assigned JIRA task as well.
- Worked closely with Tableau Admin, depends on the available change window, planed, and promoted dashboards to Test and other environments.
- Reviewed the all the dashboard with SMEs and DPE and presented and given a demo to client too.
- Documented and maintained all the process and files in secured IBM Box cloud folder.

---

**Key Accomplishments:** High CPU, High Memory, High SWAP & High SPACE Dashboards helped DA Team to easily identify which process and how many numbers of count which were casing these High SWAP, High CPU, High Memory & High Page issues and impacting the respective servers. These insights quickly helped the respective platform team to take immediate proactive actions which helped to reduce the 60% generation of Infra tickets and Reduction of server downtimes and prevention of server failures, improved accuracy on alerts and SOPs (Set of Procedures), Will recommend information to update the SOPs for optimization, adjusting the thresholds.

**Technologies used:** Python, Tableau, Cognos Analytics, SQL, IBM Box, JIRA, CRISP-DM and GitHub

**Project: RU Billing Analytics:** Resource Units (RU) Billing Team was facing difficulties with counting of RUs used to create variable portion of invoices, Significant manual activity required to produce & verify data Customer dissatisfaction with invoice accuracy, Key billing information missing or incomplete when variable invoice due, Methodology for gathering and calculating RU data is not always tested prior to 1st invoice. Loss of IBM billing revenue from missing or disputed charges, no single version of the truth (SVOT), Reports are not centralized. Because of incorrect data from several monitoring systems, it was complex and challenge to consolidate and identify the used Resource Unit for that month and bill to the client.

- 
- Worked as Data Engineer with Client, Customers, SDM, Finance Team, SMEs, Das, and DPE Operation Head and even with 3rd part monitoring tool Vendors and all the stakeholders.
  - My responsibility was to make RU Billing information as single version of the truth (SVOT), Improve transparency between IBM & KP and gain customer satisfaction with invoice accuracy, Ease of consolidation of various RU tools data and eliminate manual effort, create a centralized view of master RU list with analytics insights for immediate actions, make it easy for operation and financial management to produce accurate invoices to the clients.
  - My responsibility was to engage SDM and Finance Team and understand and gather the business and data requirements.
  - I found the RU (Resources Units) which are supported by IBM to client from all 5 monitoring source data, gathered the finance data which kept at very secure folder with restricted access, Designed and Developed Resource Units (RU) Billing Analytics Dashboard by consolidating various monitoring tools data, pre-process the data to find gain the targeted profit how many virtual servers to be sold and setup in the client's network using Linear Regression and also enabled drill-through/download option.
  - Additional security access setup at Table Servers for both Data files and Dashboard Access too with the help of Tableau Admin due to financial sensitive data to compute profit and loss.
  - Validated and Tested the Dashboards, captured the test results in the document and share with all the stake holders by attaching to the JIRA tasks. And update the accordingly update the data model documents, other user guide documents and updated the details in assigned JIRA task as well.
  - Worked closely with Tableau Admin, depends on the available change window, planed, and promoted dashboard to Test and other environments.
  - Reviewed the all the dashboard with SDM, SMEs and DPE and eventually presented and given a demo to client.

---

**Key Accomplishments:** RU Billing Analytics Dashboard helped to gain faith and confidence from the client with invoice accuracy, eased consolidation of various RU monitoring tools data, and eliminated manual effort of 40 hours per month, created a centralized view of master RU list with analytics insights for immediate actions and Easy for operation and financial management. And helped RU Management with critical insights like Profit & Loss from the RUs, Billing Trend and RU Type churn, Through Cross Mapper insight the



management was able to find-out which monitoring tool is accurate, and Deviation tracker. Making RU Billing information as single version of the truth (SVOT) improved transparency between IBM & KP, which. RU Billing Dashboard helps in optimize/accurate the RU Billing process and upsurge profits and minimize the losses to IBM, it also helped trending platform team's servers/RUs to be focused to minimize the losses to IBM. Repository Cross mapper will help analyze the RU counts analysis from various monitoring tools.

**Technologies used:** Python, Tableau, Cognos Analytics, SQL, IBM Box and GitHub, JIRA

**Project: Problem & Change Management Analytics:** Middle Level Management of KP wanted to optimize the Problem Management and Change Management so my responsibilities were to understand business requirements, process, and Data from Middle Level Management (Service Delivery Managers [SDMs] and Data Analyst [DAs])

---

- Worked as Data Engineer with Client, Customers, SDM, Finance Team, SMEs, Das, and DPE Operation Head and even with 3rd part monitoring tool Vendors and all the stakeholders.
  - I quickly understood the Business, Process, Data and workflow of both Problem and Change Managements. For proactive prevention of problems.
  - Insight 1) I wrote R programming for text mining / extracting the root cause analysis test from the free form text filed from the ticketing data to help platform team to get solution to quickly fix the tickets and close basically to reduce resolution time of the tickets. Insight
  - Insight 2) I wrote ARIMA forecasting machine learning algorithm in R for predicting the Critical and high priority tickets for future weeks to optimize and plan their operations.
  - I identified the Root Cause Analysis (RCAs) of existing problem tickets by text mining using R and captured the pre-processed dataset. Also computed the insight for problem tickets and how many incidents were contributing the same so that the problem management dashboard was having Descriptive and Diagnostic Analytics actionable Insights.
  - I also designed and developed the Change Management Dashboard with Descriptive and Diagnostic Analytics actionable Insights using Python and built ARIMA model using R to forecast and predict the problem tickets for future months and extracted root cause long text fields. Later migrated the dashboard from Tableau to Cognos Analytics and created key analytics actionable insights using ARIMA, Linear Regression, NLP, and Text mining statistics models.
  - Validated and Tested the Dashboards, captured the test results in the document and share with all the stake holders by attaching to the JIRA tasks. And update the accordingly update the data model documents, other user guide documents and updated the details in assigned JIRA task as well.
  - Prepared the user access list by engaging with all the stakeholders, granted the access, performed the dashboard usage analysis, and conducted the user adoption of dashboards sessions to end users.
  - Worked closely with Tableau Admin, depends on the available change window, planed, and promoted dashboard to Test and other environments.
- 

**Key Accomplishments:** eventually helped in 20% reduction of incidents and saved weekly 8 hours of manual effort in creation and sharing the reports hence both the middle level management teams were able to proactively take the decisions and optimize their operations.

**Technologies used:** Tableau, Python, R, Cognos Analytics, SQL, IBM Box, JIRA

**Project: Aging Ticket Analytics:** ACE Insurance/CHUBB wanted to optimize it's operation of aging tickets.

---

- As a Data Engineer, collected raw data, understood the requirements
  - Prepared the TS Data Sets, Designed and Developed R and R Shiny code for iterative visualizations
  - Wrote ARIMA Model to predict and forecast the high priority ticket counts
  - Trained and Tested the model by model diagnostics
  - Combined the TS Raw Data and Predicted Data.
  - Published into the dashboard and validated
- 

**Key Accomplishments:** IT Operation Managers to efficiently manage the high priority tickets interactively by slider, with cut off line which helped to plan the resources and avoided the missing SLA and Penalties too. ARIMA forecasting machine learning algorithm in R for predicting the high priority tickets for future weeks to optimize and plan their operations.

**Technologies used:** R, R Shiny, Tableau, Python, Cognos Analytics, SQL, IBM Box, JIRA, XL

Used R Markdown for sharing my analysis details with stakeholders, for sentiment analysis on the survey data I used R programming and recently taken course on R for Quantitative Finance and learnt Advanced Mathematical and Statistical Finance for Products Evaluation, Optimize the Portfolio and Performance & Risk Return with few tickers.

**Previous project list:**

| Projects  | Client   | Role   | Duration                    |
|---|--|--|-----------------------------|
| Container Ship-Schedule Optimization  | Walmart  | Data Engineer  | 2019 to 2019                |
| Business Operations (BO) Dashboard  | Westpac  | Data Engineer  | 2018 to 2019                |
| Call Data Analytics   | Suncorp Bank Australia                           | Data Engineer  | 2018 to 2019                |
| Aging Ticket Analytics  | ACE Insurance/CHUBB                              | Data Engineer  | 2017 to 2018                |
| ECM Reporting & CIRATS Reporting -GTS Compliance Analytics  | IBM GTS Labs (Identity & Access Management)      | Data Engineer  | 2016 to 2017                |
| Global Systems Management Reporting Technology (GSMRT) Analytics  | Thomson Reuters, Time Inc                        | Requirement Architect/DA                                       | 2013 to 2016                |
| Retail-MIS, 1DQ (Data Quality) and P&P (Permissions and Preferences)  | Vodafone UK                                      | BI Architect   | 2013 to 2016                |
| DOC (Daily Operation Control), My Honda, PLCC (Prospect to Loyal Customer Care):  | Honda Europe                                     | BI Architect   | 2013 to 2016                |
| AD and AMO - Vodafone Portugal - Managed VF Portugal's BI Projects and AMO operation with 40 resources 35 IBM Employees and 5 contractors too.  | Vodafone Portugal (Telecom Industry)             | Delivery Manager   | 2012 to 2013 (~1 Year only) |
| Intel Premier Support IPS Report, E-Customs, Shipment Visibility, SMCCS, SAP Re-platform (Program), WFDS (Workflow Data Staging), E-Pipe Assembly/Test (A/T), E-Pipe Global Platform Solutions (GPS), ART (Account Reconciliation Tool), SAP P2P (Purchase to Pay) (Program), IION EDW Data Conversion Phase-3, POC: ICFF | Intel  | ETL Developer, ETL Architect, ETL Lead                         | 2008 to 2012                |
| Global FDMS (Financial Document Management System) program, eDDAR (electronic Data and Document Archive and Retrieval) program, MA MARK (Marketing Automated Rational Knowledgebase), RK, KTBR/KTBE, IT Alignment,  | Intel  | System Integrator, Application Architect, Enterprise Architect | 2004 to 2008                |
| CyKnit Knowledge Management Portal  | Nanyang Technological University (NTU) Singapore | Developer, Lead  | 2002 to 2004                |
| CyDOC Manager [Product]   | Nanyang Technological University (NTU) Singapore | Developer, Product Owner                                       | 2002 to 2004                |

**ACADEMIC HIGHLIGHTS**

MBA in Business Analytics | Sunstone Business School | 2015 – 2016

B.E in Computer Science | PDA Engineering College | 1993 – 1998

Diploma in Electronics | Govt Polytechnic | 1990 – 1993