

# **INDIAN JANSSEN RDE**

Project report submitted in partial fulfillment of the requirement for the degree of

Bachelor of Technology

In

**Computer Science and Engineering/Information Technology**

By

Vikal Choudhary (181236)

Under the supervision of

**Mr. Subhash Nadkarni**

( Consultant, ZS Associates )

and

**Dr. Rajni Mohana**



Department of Computer Science & Engineering and

Information Technology

**Jaypee University of Information Technology**

**Waknaghat, Solan-173234, Himachal Pradesh**

## Candidate's Declaration

I hereby declare that the work presented and done in this document entitled "Integrated Account Management" in partial fulfilment of the requirements for the award of the degree of Bachelor of Technology in Computer Science and Engineering/Information Technology submitted in the department of Computer Science & Engineering and Information Technology, Jaypee University of Information Technology, Waknaghat is an authentic record of our own work carried out over a period from Feb 2022 to May 2022 under the supervision of Mr. Subhash Nadkarni and Dr. Rajni Mohana (Faculty, Computer Science and Engineering And Information Technology)

The matter embodied in the report has not been submitted for the award of any other degree or diploma.

Vikal Choudhary, 181236

This is to certify that the above statement made by the candidate is true to the best of my knowledge.

Mr. Subhash Nadkarni

Consultant (ZS)

Dated: 28/05/2022



## Project Report Undertaking

I Mr. Vikal Choudhary Roll No.\_181236\_Branch  
Information Technology is doing my internship with\_\_\_\_\_ ZS  
Associates from \_\_\_\_1<sup>st</sup> Feb 2022\_\_\_\_\_ to \_\_\_\_\_31<sup>st</sup> July  
\_\_\_\_\_  
2022\_\_\_\_\_

As per procedure I have to submit my project report to the university related to my work that I have done during this internship.

I have compiled my project report. But due to COVID-19 situation my project mentor in the company is not able to sign my project report.

So I hereby declare that the project report is fully designed/developed by me and no part of the work is borrowed or purchased from any agency. And I'll produce a certificate/document of my internship completion with the company to TnP Cell whenever COVID-19 situation gets normal.

Signature\_\_\_\_\_

Name\_Vikal Choudhary\_

Roll No.\_181236\_\_\_\_\_

Date\_\_29 May 2022 \_\_\_\_\_

## **Acknowledgement**

I would like to express my special thanks of gratitude to my supervisor in College Dr. Rajni Mohana and Mr. Subhash Nadkarni (Consultant, ZS Associates) who gave me this golden opportunity to do this project on the Data Analyzing and DBMS tools which also helped me in doing all of the Research and we came to know about so many new things. We are really thankful to him.

Secondly, I would also like to thank my mentor and team mates who helped me a lot in finalizing this project within the limited time frame.

Vikal Choudhary, 181236

-  
-  
-  
-  
-  
-  
-  
-  
-

# Table of Contents

1) Introduction.....	(6)
1 Pharmaceuticals Industry.....	(6)
2 ZS Associates.....	(10)
3 Contracting Capability Evolution.....	(11)
2) Tools and Tech used.....	(12)
3) Literature Survey.....	(19)
1 Survey 1.....	(19)
2 Survey 2.....	(20)
3 Survey 3.....	(20)
4) Proposed Work.....	(23)
4.1 Intelligent Cubes.....	(28)
5) Result and Discussion.....	(39)
6) Conclusion.....	(40)
7) References.....	(41)

## **Introduction**

### Pharmaceuticals Industry

The pharmaceuticals industry acquires, develop, manufactures, and sells drugs or drugs for use as drugs that will be given (or self-administered) to patients, for the purpose of treating, vaccinating, or reducing symptoms. Pharmaceutical companies can work with the generic or branded medicines and medical device. They are subject to various laws and regulations governing patent, testing, safety, efficacy and marketing of drugs.

### Research and Development

Drug discovery is that the method by that potential medicine area unit discovered or designed. within the past, several medicine were obtained by separating the active ingredient from ancient remedies or by the discretion of the drug. trendy biotechnology usually focuses on understanding the metabolic pathways related to a malady or infectious agent, and applying these techniques exploitation biological science or organic chemistry. vital early drug discovery has been created through universities and analysis institutes.

Medication improvement alludes to the exercises performed after the compound has been distinguished as an expected medication to decide its adequacy as a medication. The destinations of medication advancement are to decide the suitable plan and measurement, just as to set up wellbeing. Exploration there regularly includes a blend of in vitro considers, in vivo examines, and clinical preliminaries. The expense of late-stage improvement implies that it is generally done by huge drug organizations.

## Marketing

Advertising is common in health care newspapers and in the mainstream media. In some countries, especially the US, they are allowed to advertise directly to the general public. Pharmaceutical companies often use retailers (often referred to as 'drug reps' or, more commonly, 'men's details') to market themselves directly and personally to doctors and other health care providers. Since the 1980's new ways to market prescription drugs to consumers have become increasingly important. Consumer-directed media advertising is officially approved by the FDA Guidance for Industry on Advertising Directed Advertising.

## Health Insurance In USA

Health insurance within the u. s. is any program that helps obtain medical bills, be it in camera purchased, social welfare, or a state-sponsored welfare program. [1] Similar terms for this use embrace "health coverage", "health care", and "health benefits". in a very additional technical sense, the term "health insurance" is employed to explain any sort of insurance that gives protection against the price of medical services. These uses embrace each non-public insurance and social welfare programs like health care, which includes resources and spreads the monetary risks related to high medical prices for all to safeguard everybody, yet as social programs like Medicaid and also the kid

insurance Program. , that provides help to people that cannot afford health care.

## Medicaid

Medicaid was developed for the poor in 1965. As registrants must pass the means test, Medicaid is a social or social security program and not a social insurance scheme. Despite its inception, the percentage of US residents without any form of health insurance has increased since 1994. The Affordable Care Act significantly increased Medicaid. The plan now covers everyone with less than 133% of the level of poverty that does not qualify for Medicare, as long as this increase in access is accepted by the country in which the person lives. In the meantime, Medicaid's benefits should be matched by a significant benefit to the newly created state exchange. The provincial government will fully fund Medicaid expansion initially, with a specific financial obligation (10% of medical expenses) gradually returning to the provinces by 2020.

## Group Purchasing Organization

In the United States, a gathering buying bunch (GPO) is an association made to utilize the buying force of a business gathering to get limits from retailers dependent on the joint buying power [1] of GPO individuals.

Numerous GPOs are subsidized by authoritative charges paid by GPO-oversaw retailers. A few GPOs are financed by buying individuals. A few GPOs are upheld by a blend of the two methodologies. These sums can be set as a buy rate or set as a base yearly rate. A few GPOs set obligatory investment rates for their individuals, while others are completely dedicated. Taking an interest individuals as indicated by their buy needs and their degree of trust in what ought to be the cutthroat costs haggled by their GPOs.

Gathering buys are utilized in numerous enterprises to purchase crude materials and crude materials, yet it is basic in the retail, wellbeing, innovation, producing and farming businesses. Lately, bunch buys have started to zero in on a charitable society. Gathering buys between not-for-profits are still moderately new, however are before long getting ordinary as philanthropies intend to discover approaches to lessen further expenses. In the medical services area, GPOs are frequently found in concentrated consideration associations, however charitable public facilities and wellbeing focuses all through the U.S. They were associated with bunch buys.



## Health Care GPO

Healthcare Group (GPO) assists in promoting quality health care and assists various providers in managing costs effectively. The GPO combines the purchasing power of its members with a variety of goods and services and has contracts with suppliers where members can purchase at group rates and terms of choice. GPOs often offer contract discounts on medical, nutrition, pharmacy and laboratory services. Some large GPOs have expanded their portfolio of contracts to offer discounts on office supplies and non-medical services. As the healthcare industry is full of GPOs, pricing is another way for GPOs to bring in new members or convince members of another GPO to change.

## CoPay

Copay is a cash based sum paid by the safety net provider for incorporated administrations. It is a typical piece of numerous medical coverage plans. Safety net providers frequently charge co-installments for administrations, for example, specialist visits or doctor prescribed medications.

Copays is a predefined dollar esteem instead of a level of the bill, and they were normally paid during activity. Not all clinical benefits demand copay. For instance, some insurance agencies don't need a copay of yearly items.

### How Co-Pay functions

Copay expenses change between guarantors however are normally \$ 25 or less. For instance, a protected protection plan may require the safety net provider to pay \$ 25 for a specialist's visit or \$ 10 for each specialist's remedy. Survey the details of your protection intend to decide your installment choice.

### How Does Copays Affect Insurance Premiums?

A premium is the amount paid for an insurance policy. In most cases, programs with very high premiums may pay less, while programs with lower premiums may receive higher co-pay.

## KEY TAKEWAYS

Not all medical visit required payment for patients.

Offline visits may have higher payouts or copayers than in-service medical

providers. Withdrawals are larger than copays.

Copays and insurance aren't identical things. insurance could be a share of the bill, copay a set quantity

Similarly there is a large market of Pharmaceuticals and Insurance plan companies (Payers) In USA. And they want to increase their business and sales with the growing market. Here comes the role of ZS Associates.

# ZS Associates

ZS Associates is a consulting and technology consulting company focused on consulting, software, and technology, headquartered in Evanston, Illinois, providing customer equity, private health care and technology. ZS was founded in 1983 by Prabhakant Sinha and Andris Zoltners, who work together as professors of marketing at Kellogg School of Management at Northwestern University.

We are passionately committed for helping our clients and their customer thrive, working together to drive customer values and results.

## Strategy and Advice

From portfolio management and growth strategy to customer experience and digital transformation, we help clients improve performance and deliver lasting value.

Sales Enhance the efficiency and effectiveness of your organization with smart marketing strategies, team formation and performance management solutions.

Technology and performance Tap into our comprehensive collection of technology solutions — design strategy, awesome implementation and automation - to help your company thrive.

With in-depth data sources, analytics and trading strategies, we can assist you with market-going strategies, sales and structure team size, collaborative strategies, incentive compensation programs, performance management and more.

ZS is an expert administrations firm that works one next to the other with organizations to help create and convey items that drive client worth and friends results. We influence our profound industry ability, driving edge investigation, innovation and methodology to make arrangements that work in reality. With over 35 years of involvement and in excess of 8,000 ZSers in 28+ workplaces around the world, we are energetically dedicated to aiding organizations and their clients flourish. We have three core values rooted into our culture, Get it Right, Do the Right Thing and Treat People Right.

# Contracting Capability Evolution

Pharma companies in the USA sign contracts with Insurance companies in order for their product to be placed in higher tiers of the insurance plans, so that the customer opts for their product. ZS analyzes the data, both current and historical regarding the products and the scenarios and provides insights on multiple key levels in this system. We provide guidance through data analysis on what contract should be created for maximum efficiency. We provide data visualizations used to understand the impact of the contract on the sales, or other important metrics.

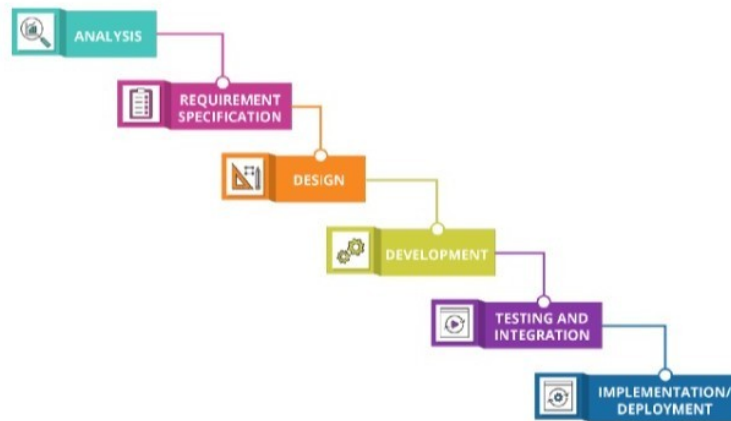
This is one big team and department at ZS working for our clients and helping them in all contracting and growth related models and operations. Also providing them the best solutions for future and current contracts which in turn will be helping the client as a whole.

Contractors' contracts define and define a provider's reimbursement program for delivering health care services to patients covered by a particular health system. Contracts cover everything from refund prices and network providers depending on the medical need and provider verification.

Understanding complex taxpayer contracts is important in ensuring timely refunds, and keeping the doors of practice open. Being able to enter and exit each contract is also important in avoiding denial of claims, dragging patients into the practice, and providing complete (and non-refundable) services to patients.

IAM is one such project under it i.e. Integrated Account Management.

## APPROACH:



# Tools and Technologies Used

## Business Intelligence Software

Business intelligence software is a type of software application designed to detect, analyze, modify and report business intelligence. Applications typically read previously stored data, usually - or otherwise - in the database or in the data center. A Business Intelligence (BII) tool is software that collects, modifies, and presents information to help decision-makers drive business growth. BI tools incorporate a large amount of structured and unstructured data from a variety of sources, modify it and assist in gaining effective business insights from the data.

- SAS Business Intelligence
- Clear Analytics
- SAP BusinessObject
- Domo
- METRIC INSIGHTS
- Zoho Analytics
- Microstrategy

We have integrated Amazon Web Services into the core of our project and all the components of the project are cloud based starting from the data stores on amazon S3 to the tables stored in amazon Redshift clusters as well as High level data analytic performed using Sagemaker. We use Microstrategy as the

data visualization tool of choice. It acts as a platform to develop reporting applications which are intuitive and informative.

## Tableau

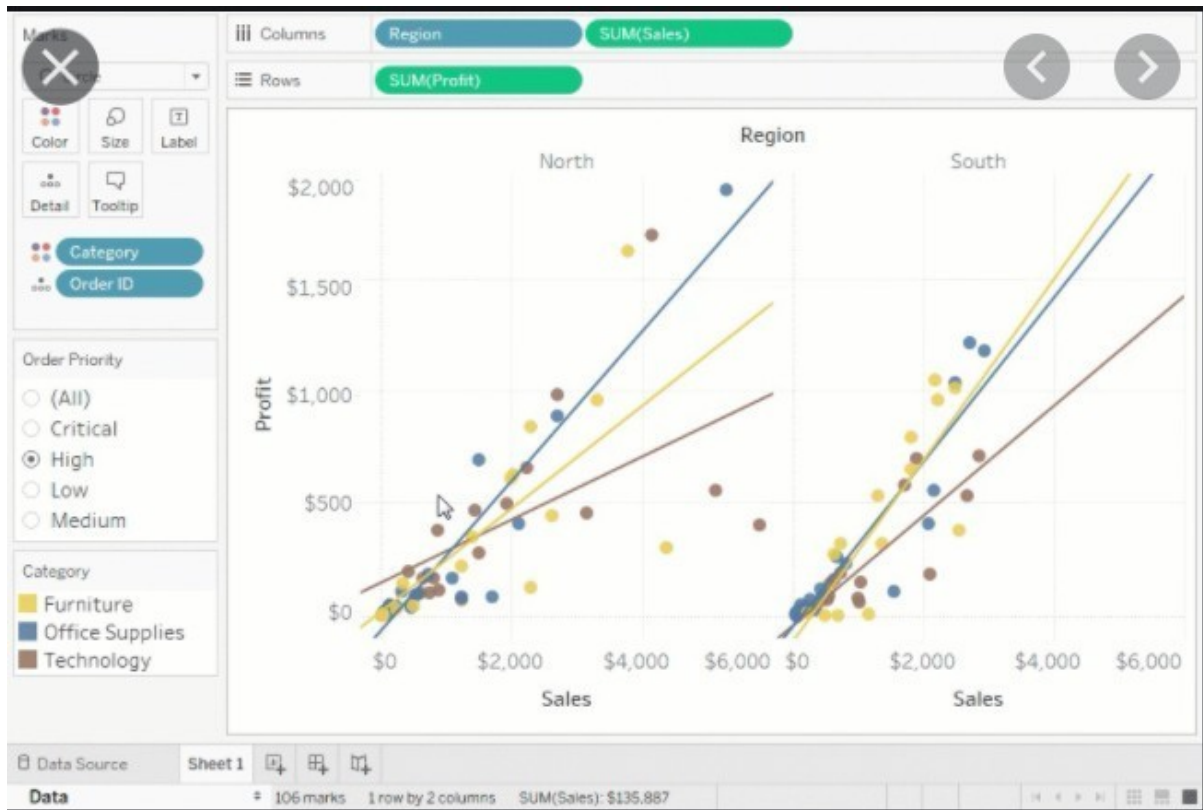
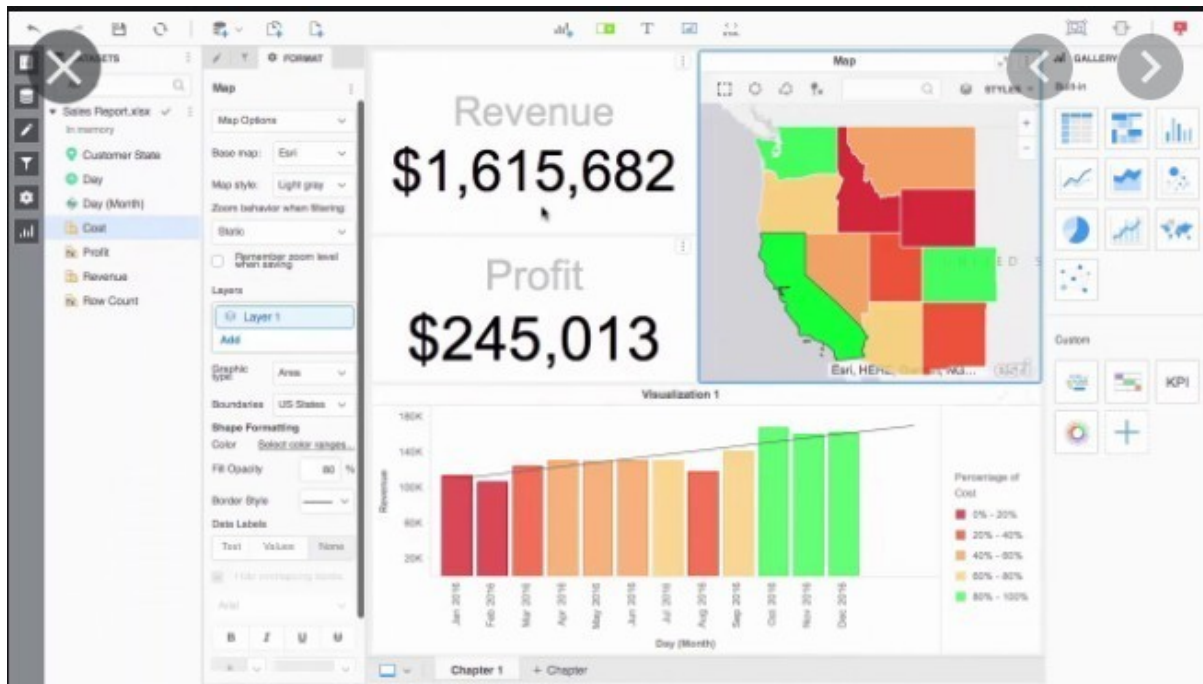


Tableau code is associate yank knowledge consulting code company centered on business intelligence. it had been supported in 2003 in Mountain read, California, and is presently headquartered in metropolis, Washington. In 2019 the corporate was nonheritable by Salesforce for \$ fifteen.7 billion. CNBC rumored that the acquisition was the most important acquisition by Salesforce, thought-about to be the strongest within the CRM sector, since its origin. The company's founders, Christian Chabot, Pat Hanrahan and Chris Stolte, were researchers within the Department of Technical Science at Stanford. They became aware of the visual techniques of testing and analyzing relationships associated with knowledge and knowledge cubes, and commenced the corporate as a look center in Stanford from 1999 to 2002. Table product enkindle relationship-related data, on-line process cubes, cloud data, and spreadsheets to get graph-type knowledge visibility. code also can extract, store, and restore knowledge from a memory engine. Tableau has map practicality, and is in a position to edit latitude and meridian links and link to location files like Esri Shapefiles, KML, and GeoJSON to show custom geographics. integral geo-coding permits body areas (state, state / province, region / region), communicating codes, America general assembly regions, America CBSA / MSA, Area Codes, Airports and European Economic Community applied math locations ( kookie codes) ar

mechanically mapped. Geographies may be collected to make custom locations or custom geocoding wont to expand existing native roles in an exceedingly product.

## MicroStrategy



MicroStr Strategy Incorporate is one company that provide business intelligence (BI), mobile software and cloud services.

MicroStr Strategy business analysis and mobility platforms help businesses build and deploy analytics and mobility applications to transform their business.

Founded in 1989 by Michael Saylor, Sanju Bansal, and Thomas Spahrr, the firm develops internal and the external data analysis software to make business decisions and improve mobile applications. It is a public company headquartered in Tysons Corner, Virginia, Washington. The main business competitors include SAP AG Business Objects, IBM Cognos, and the BI Platform or Platform Corporation. Saylor is the CEO and chairman of the board. MicroStr Strategy 2020, the company's new speaker release for its intelligence software, includes the development of HyperIntelligence vendor skills, an embedded analytics system using advanced intelligence and machine learning technology.

MicroStr Strategy 2019, a previous speaker released, focuses on three areas: integrated analytics, allowing extended connections to data sources and applications; flexible mobility, easy mobile performance improvements; and HyperIntelligence, which includes Bluetooth and voice recognition. The

previous software version, MicroStr Strategy 10, contained MicroStr Strategy Analytics, MicroStr Strategy Mobile, and Usher. MicroStr Strategy 10.10, released in December 2017, added the MicroStr Strategy Workstation. It uses business intelligence and forecasting predictions to search and analyze analytics for large data from a variety of sources, including data repository, excel files and Apache Hadoop distribution.

## AWS S3



Amazon S3 or Amazon Simple Storage Service is a help given by Amazon Web Services (AWS) that gives object stockpiling through a web administration interface. Amazon S3 utilizes a similar stockpiling foundation utilized by Amazon.com utilizing its worldwide online business organization. Amazon S3 can be recruited to store any sort of article, taking into account uses like Internet stockpiling, reinforcement and reestablish, calamity recuperation, information vaults, investigation information pools, and coordinated distributed storage.

AWS dispatched the Amazon S3 in the United States on March 14, 2006.

While Amazon Web Services (AWS) doesn't freely unveil S3 innovation assemble subtleties, Amazon S3 oversees information on design which means to give vigor, high accessibility, and low dormancy with 99.99999999% sturdiness and - 99.95% up to 99.99% (with no firm assistance level understanding).

The fundamental units of the Amazon S3 are pail measured things. Every thing is recognized by a remarkable key, allocated to the client. Pails can be overseen utilizing either the reassurance given by Amazon S3, utilizing the AWS SDK framework, or the Amazon S3 REST RSS application programming interface (API) Items can be overseen utilizing the AWS SDK or through the Amazon S3 REST API and it very well may be up to five terabytes of information for each two kilobytes of metadata. Moreover, things can be downloaded utilizing the HTTP GET interface and the BitTorrent convention.



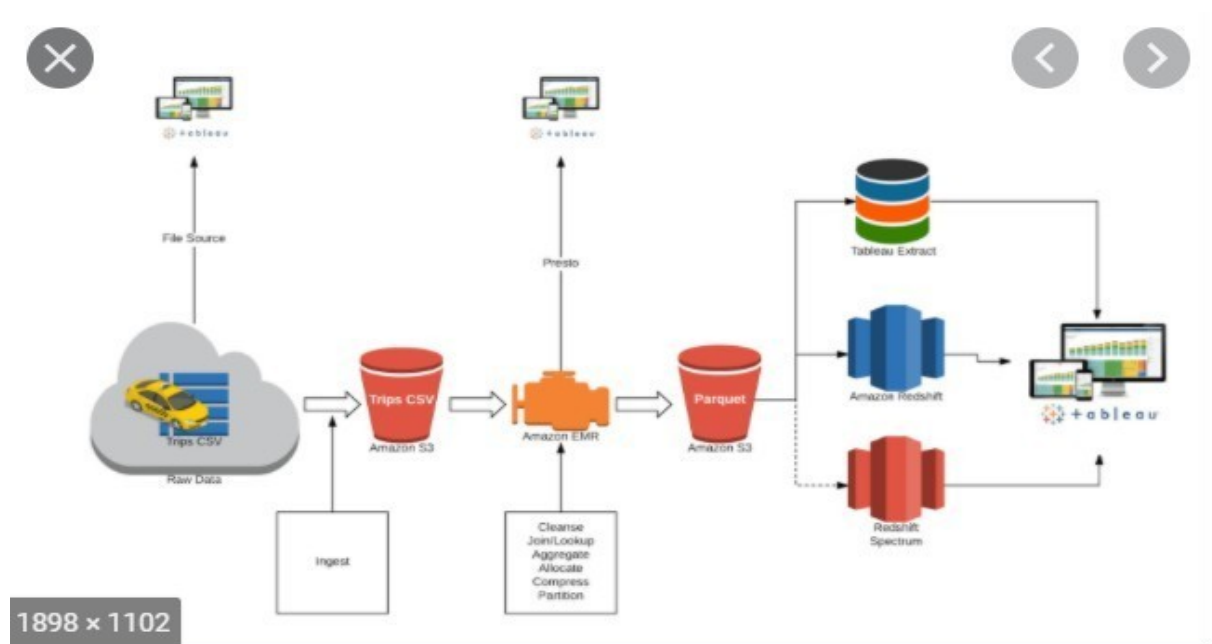
Amazon S3 offers four classes that offer various degrees of strength, accessibility, and execution prerequisites.

Amazon S3 Standard is the default classification. Amazon S3 Standard Infrequent Access (IA) is intended for rare information. Basic instances of utilization of catastrophe hazard the executives arrangements.

The Amazon S3 One Zone-Infrequent Access is intended for regularly undesirable information however if necessary, it needs quick access. Information is put away in one spot and if that area is erased, all information is lost.

Amazon Glacier is intended for long haul stockpiling of bizarrely accessible information and when recovery of minutes or hours is worthy. The "Ice sheet Deep Archive" is another approach to make some recuperation memories of in any event 12 hours, however 1/4 cost. It is planned as an option in contrast to attractive tape libraries, and is intended for long haul information stockpiling for 7 to 10 years.

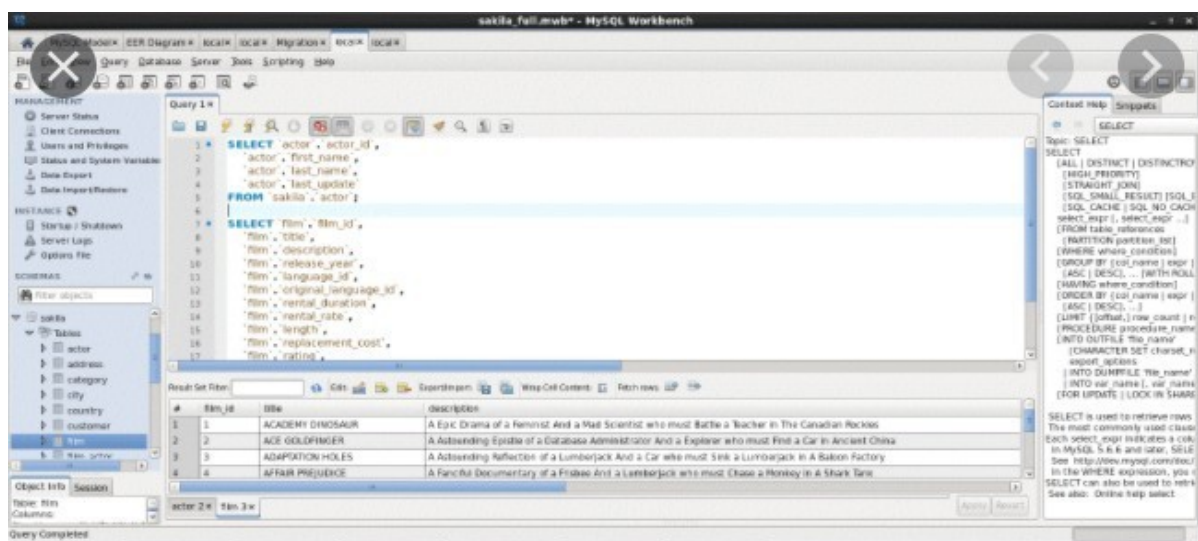
## Redshift



Amazon's Redshift is a data storing product that is part of Amazon Web Services' largest cloud platform. The name means from Oracle, red being the object of Oracle, its red color and randomly called "Big Red." Built on technology from state-of-the-art parallel (MPP) storage company ParAccel

(further acquired by Actian), large data management and data migration. Redshift is different from other Amazon's hosted database offerings, Amazon RDS in its ability to handle analytical workloads on large datasets maintained by the DBMS-based DBMS policy. Redshift allows up to 16 petabytes of data in the collection compared to Amazon RDS 16TB database size. Amazon Redshift is predicated on the older version of PostgreSQL eight.0.2, and Redshift has created changes thereto version. The preview beta was free in Nov 2012 and therefore the full unleash was created on the market on February fifteen, 2013. The service will manage connections from several different programs victimisation ODBC and JDBC communications. in keeping with a Cloud knowledge Warehouse report printed by Forrester in this autumn 2018, Amazon Redshift has the biggest readying of cloud knowledge, with over half-dozen,500 usage. Redshift uses an equivalent process and pressure to scale back command time. this permits Redshift to perform operations on ample lines at the same time. This conjointly makes Redshift helpful for storing and analyzing giant knowledge from logs or live feeds via a supply like Amazon response knowledge Firehose. Amazon has listed variety of homeowners of intelligence software package as partners and tools tested in their "APN Partner" program, together with Actian, Actuate Corporation, Alteryx, Dundas knowledge visual image, IBM Cognos, InetSoft, Infor, Logi Analytics, Looker (company), MicroStr Strategy, Pentaho, Qlik, Sense, Tableau software package, and tunny. Partner firms that give knowledge assortment tools embrace Informatica and SnapLogic. System integration with consultation partners includes Accenture, Deloitte, Capgemini and DXC Technology.

## MySQL Workbench



SQL Workbench is a graphic design tool that integrates SQL development, management, data creation, creation and storage in a single MySQL database system development. You are a fan of DBDesigner 4 from faceFORCE.net, and have replaced the previous software package, the MySQL GUI Tools Bundle, creating and creating data within MySQL data. Developed by MySQL AB and later by Sun Microsystems and released under the GPL. Bulk GUI tool development has stopped, and is now only stored under the download Archive of the MySQL site. On April 5, 2018, the MySQL work table Team declared the primary public unharness of version eight.0.11 as unharness Candidate (RC) and MySQL Community Server eight.0.11. the primary unharness of General accessibility (GA) appeared on Gregorian calendar month twenty seven, 2018 and at the side of the server follows a brand new policy to align translation numbers across all MySQL merchandise. works with object editors (causes, views, saved processes, and functions) within the visual SQL editor and models. The new varieties add support for brand spanking new language options in MySQL eight.0, like descriptions and customary table and roles. there's additionally support for intangible indicators and persistence within the diversity of the worldwide system. The new default caching\_sha2\_password word for MySQL eight.0 is currently supported by work table, thus resetting user accounts to different varieties of authentication isn't any longer necessary once connecting to the newest servers. Management tabs are updated with the newest configuration choices and also the interface is customised between tabs.

The main features of the MySQL Workbench are:

#### General

- Database Connection & Instance Management
- Items of action run by Wizard
- Completely written in Python and Lua
- Support for the Custom plugin
- MSAA (Window Accessibility API) is compliant in it
- Supports the MySQL enterprise features (Audit log, Firewall , and the Enterprise Backup)

#### SQL editor

- Browse for browsing, exploring and for searching
- SQL syntax highlighting and editorial statement
- SQL code termination and critical context help
- A set of multiple and systematic effects
- VISIBLE DEFINITION
- SQL captions collections
- SSH connection tunnel
- Unicode support

## Data modeling

- ER design
- Drag visual modeling
- Restore engineering from SQL scripts and live database
- Transfer engineering to SQL documents and live database
- Schema Sync
- Printing models
- Import from FATFORCE.net DBDesigner4

## Data domain management

- Start and configure database conditions
- Status suspension
- Data account management
- Browse flexibility
- Browsing login file
- Dispose of data for export and import
- Performance monitoring
- Performance Schema metrics
- MySQL example dashboards
- Question statistics

# **Literature Survey**

## **1) Micro Strategy and Strategizing : Towards an Activity-Based View**

Gerry Johnson, Leif Melin, Richard Whittington

This starting paper gives foundation, topics and papers for this exceptional issue of sub-systems and key arranging. Our principle contention is that, while the essential area has generally centered around an enormous number of associations, it now needs to address low level occasions. We propose a work-based viewpoint of a system that spotlights on definite cycles and practices that shape the everyday activities of the association and related key results. This paper fosters this viewpoint by analyzing two spaces of vision, specifically asset and institutional point of view; two constructions of underlying work, those of various organizations and structures; lastly the way of life of the essential exploration measure. This paper recognizes the advantages of a work-based point of view and presents a portion of the difficulties of additional examination. It finishes up by presenting the papers in this unique issue.

## **2) Business Intelligence Softwares for the Classrooms: Microstrategy Resources on the Teradata University Network**

Barbara H. Wixom, University of Virginia

Faculty members area unit challenged by staying au courant business acumen and teaching the topic in acceptable ways in which. Recent developments of Teradata University Network (www.TeradataUniversityNetwork.com) to feature technology code the web site currently offers MicroStr Strategy 7i, associate degree integrated business reportage and analysis platform and plenty of MicroStr Strategy analytic modules centered on the analysis of specific business processes. New code is offered for manual use by college and students. This tutorial describes these business resources and provides many ways within which resources will be wont to produce effective schoolroom info. Resources area unit obtainable to any or all college and students freed from charge by registering with Teradata University Network.

### **3) MicroStrategy from MacroLeadership: Distributed Intelligence via New Science**

A. Y. Lewin and H. Volberda

While the advocates of the new Science still see leadership as very important for nonlinear rapidly changing world, many writers also see the break between the position of command-and-control hierarchy and type of broadcast intelligence coming from that sustainable economic employment. I use complexity theory's evolving tensions to show how CEOs can speed it up simultaneous DI information measurement suppressing the emergence of authority. Being complex science recognizes that complex species are not unrepentant; they are the result of changing tensions. Knowing this, when leaders change changing tensions set in the system, its type of difficulty changes. Specifically, the correction of variable variables is between 1 and the second critical values that

produce an emerging network properties. The science of stress, therefore, does not only provide complete methods of social expression events but also provides clear ways in which they can CEOs can make a fundamental change in intrafirm the programs they are responsible for. I say leadership with heroic ideas unemployment due to the "pay down" problem with several levels of intervention — it is more apt to suppress a company's brain rather than increase its IQ. In In contrast, macro leadership produces emerging DI without regulatory frameworks. The Macro leadership vision is not just one of the highest standards the path to leadership, many of which come from Dansereau and Jammarino (1998a, b). Build on file for the ideas of leadership discussed by Bryman were circulated (1996) and ideas for high echelon leadership by Waldman and Yammarino (1999). Identifies the roles of CEOs to use who should lead all firms, that is, "lead down" through the many levels of intervention of the organization. It avoids leading leads. Although my use of microcoevolution, DI, and complexity the theory of identifying the functions of major leadership strategies that novel, works themselves reflect OD texts. Given this, my article sets out the theory of CEO-level leadership in a compact micro / macro theory text and attach it hiring production as the same purpose for both CEOs and researchers

#### **4) A Survey on Business Intelligence Tools for Marketing, Financial, and Transportation Services**

Chavva Subba Reddy, B. Srinivasa Rao

---

The big fire at the back of this paper is to present a detailed study of Business Intelligence (BI) tools for Marketing, Transport and Finance. As research, our aim should be to introduce more BI tools such as Tableau, Power BI, Pentaho, QlikView, and Microsoft Strategy Analytical Express. These tools are easily identified in organizational use. Comparisons are made with these tools to test their effectiveness and try to find the best ones. Based on the available data collected from various sources, one can clearly conclude that Pentaho is the leading data processing provider in the other BI tools available in the specified services.

The concept of business intelligence has been widely used over the years, and now, this organization of words has been used in a variety of fields from modeling business to data technology.

Business intelligence identifies areas that apply for diversity and technology for the collection, storage, analysis and provision of information to improve business processes quality.

Business intelligence statement, Getting the right information from the right people at the right time focuses on the fact that business intelligence works with Data and not data due to the inclusion of data processing skills in smart data, which means that it works and is universally accepted and can be used consistently in the modeling process. The era of business intelligence was introduced by the Gartner Group in the mid-90s. The idea, on the other hand, existed long before it was used in high-quality reporting systems.



# Proposed Work

## INDIAN JANSSEN RDE

It is a model and project proposed for the Research and Development team of the client to help them better analyse their datasets and demand data leading to growth in the performance of AEs and hence the company(client).

### METHODOLOGY:

We follow AGILE methodology and use AWS provided by the client to manage the project. We have offshore meetings on daily basis to assign/take updates about tasks. Our work is divided into sprints. One sprint is of two weeks where we define all kind of task that need to be completed in given timeframe. Set of all the task is defined as accumulation. An accumulation is a focused on rundown of assignments that groups need to chip away at inside the extent of an undertaking the board procedure. It is gotten from the necessities and guide of the undertaking. Significant assignments in the accumulation are normally appeared on first spot on the list to tell the group which undertakings they need to convey first.

We interact with client on multiple stages. Employees at higher position keep in check with them by providing them with consistent updates on the project status and taking feedback from the client. Although we follow a waterfall approach wherein the phases of the project are well defined and the process is clearly laid out, We accommodate the changes in requirements and based on the user feedback at every stage.

Things we used upon on this Project.

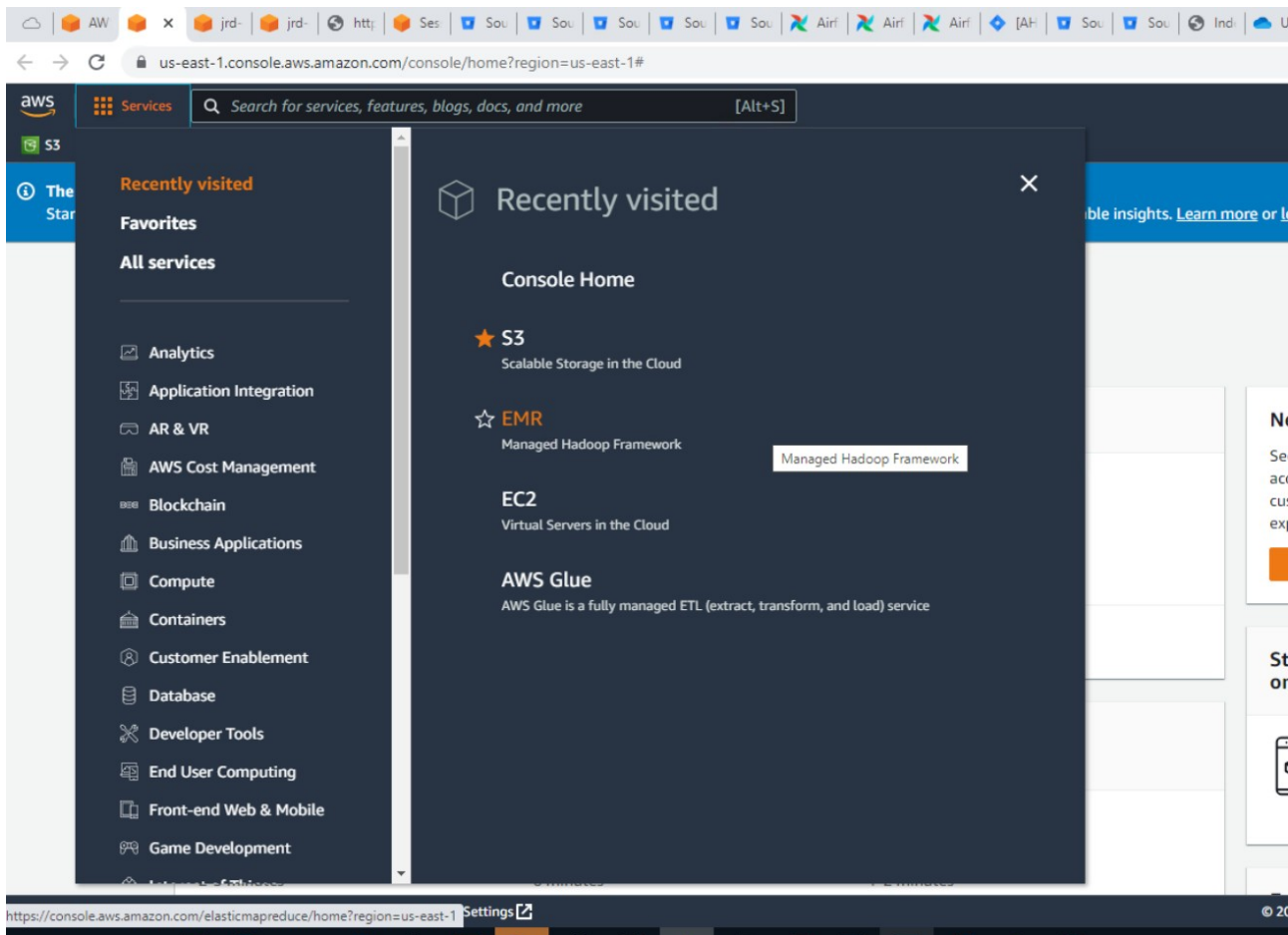
### AWS

Amazon Web Services (AWS) is the world's most comprehensive broadly adopted cloud platform, offering services from data centers globally.

AWS provide secure, reliable and globally distributed architecture to build your application.

Provides fine grained control over your account and services you use. Also, flexible 'Pay as Use' policies for better cost management.

Used in Multiple domains like Compute, Storage, Machine Learning, IOT, Automation, Game Development, Analytics and many more.



## Advantages of AWS



### Pros

- Has plenty of service, enough for application development
- Provides unlimited storage
- Excellent IT support
- Help develop scalable Infrastructure
- Has separate compute, storage and notification services.

### Cons

- Billing can get confusing, if not taken care right from beginning.
- Costly services
- Backup re-store is time consuming and costly.
- Requires prior programming knowledge

## **Features of AWS**

### Global Architecture

A combination of best-of-breed cloud services, and easy access to those.

- Deployment

Allows easy application deployment and even integration with non-AWS tools

- Resilience

Ensures recovery of systems from infrastructure or service failures to mitigate disruptions

- Performance

Efficiently utilizes computing resources to meet system requirements as demands evolve

- Security

Protects information using best security measures like having IAM roles being defined for every access required.

- Compliance

Enforces change control management, regional/country compliance needs and vulnerability remediation

- Optimization

Helps avoid or eliminate unneeded cost or suboptimal system resources

- Support

Provides prompt incident management and system maintenance assistance

## **AWS EC2**

Amazon Elastic Compute Cloud (Amazon EC2) is a web service that provides secure, resizable compute capacity in the cloud.

It is designed to make web-scale cloud computing easier for developers.

Amazon EC2's simple web service interface allows you to obtain and configure capacity with minimal friction.

It provides you with complete control of your computing resources and lets you run on Amazon's proven computing environment

AWS provides wide selection of instance types optimized to fit different use cases. Instance type comprise varying combination of CPU, memory, storage & networking capacity and give you the flexibility to choose the appropriate mix of resources for your application.

Developer can choose between General Purpose, Compute Optimized, Memory Optimized and Storage optimized instances.

We can login to EC2 using putty and can also see it on the console

## **AWS EMR**

Amazon EMR is the industry leading cloud-native big data platform for processing

vast amounts of data quickly and cost-effectively at scale

Amazon EMR can be used to analyze click stream data in order to segment users and understand user preferences. Advertisers can also analyze click streams and advertising impression logs to deliver more effective ads.

EMR gives teams the flexibility to run use cases on single-purpose, short lived clusters that automatically scale to meet demand, or on long running highly available clusters using the new multi-master deployment mode

## EMR Features

- Easy to use
- Low cost
- Elastic
- Reliable
- Secure
- Flexible

Clone Terminate AWS CLI export

⚠ User: arn:aws:sts:885736535748:federated-user/vchoudh8 is not authorized to perform: elasticmapreduce:GetAutoTerminationPolicy on resource: arn:aws:elasticmapreduce:us-east-1:885736535748:cluster/j-3LRJICO46DIRQ because no identity-based policy allows the elasticmapreduce:GetAutoTerminationPolicy action

Cluster: JRD-ORBIT-DE-L2-DataWarehousing-Weekly-Full-Dev **Terminating** Terminated by user request

Summary Application user interfaces Monitoring Hardware Configurations Events Steps Bootstrap actions

### Summary

ID: j-3LRJICO46DIRQ  
Creation date: 2022-05-27 15:01 (UTC+5:30)  
Elapsed time: 15 hours  
After last step completes: Cluster waits  
Termination protection: Off  
Tags: Application = APP000010041301, Cost Center = 0000001270, Creation Date = 20210611, Environment = Dev ... [View All](#)  
Master public DNS: ec2-3-231-19-70.compute-1.amazonaws.com [Connect to the Master Node Using SSH](#)

### Configuration details

Release label: emr-5.25.0  
Hadoop distribution: Amazon 2.8.5  
Applications: Hive 2.3.5, Pig 0.17.0, Ganglia 3.7.2, Hue 4.4.0, Spark 2.4.3, Sqoop 1.4.7, Zeppelin 0.8.1, Livy 0.6.0, Tez 0.9.2, Presto 0.220  
Log URI: s3://jrd-orbit-logs-dev/emrlogs/ [View](#)  
EMRFS consistent view: Disabled  
Custom AMI ID: --

### Application user interfaces

Persistent user interfaces [View](#): Spark history server  
On-cluster user interfaces [View](#):

### Network and hardware

Availability zone: us-east-1c  
Subnet ID: [subnet-095c2171c668f1039](#) [View](#)  
Master: **Terminating** 1 r5.2xlarge  
Core: **Terminating** 10 r5.2xlarge  
Spot (max \$0.441/hr)

## AWS S3

Amazon Simple Storage Service (Amazon S3) is an object storage service that offers industry-leading scalability, data availability, security, and performance.

Amazon S3 provides easy-to-use management features so you can organize your data and configure finely-tuned access controls to meet your specific business,

organizational, and compliance requirements.

Amazon S3 is designed for 99.999999999% (11 9's) of durability, and stores data for millions of applications for companies all around the world.

Bucket : A bucket is a container for objects stored in S3. Every object is contained in a bucket.

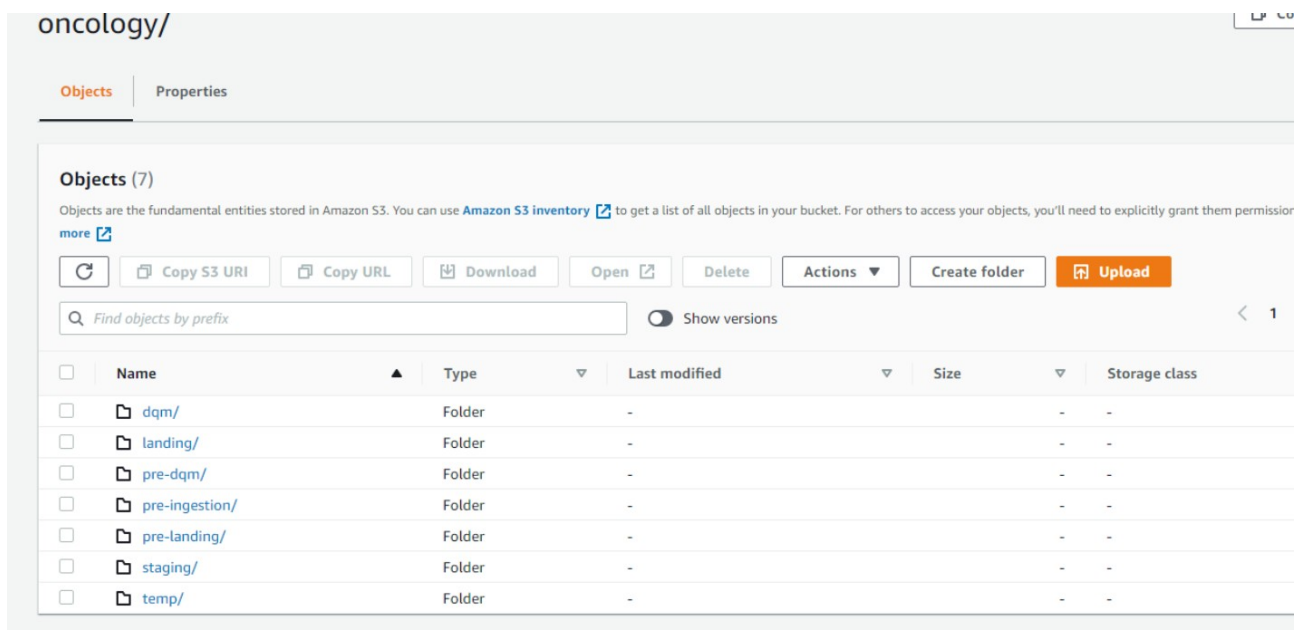
Object : Objects are fundamental entities stored in S3. Objects consist of object data and metadata. Metadata is set of key-value pairs that describe the object.

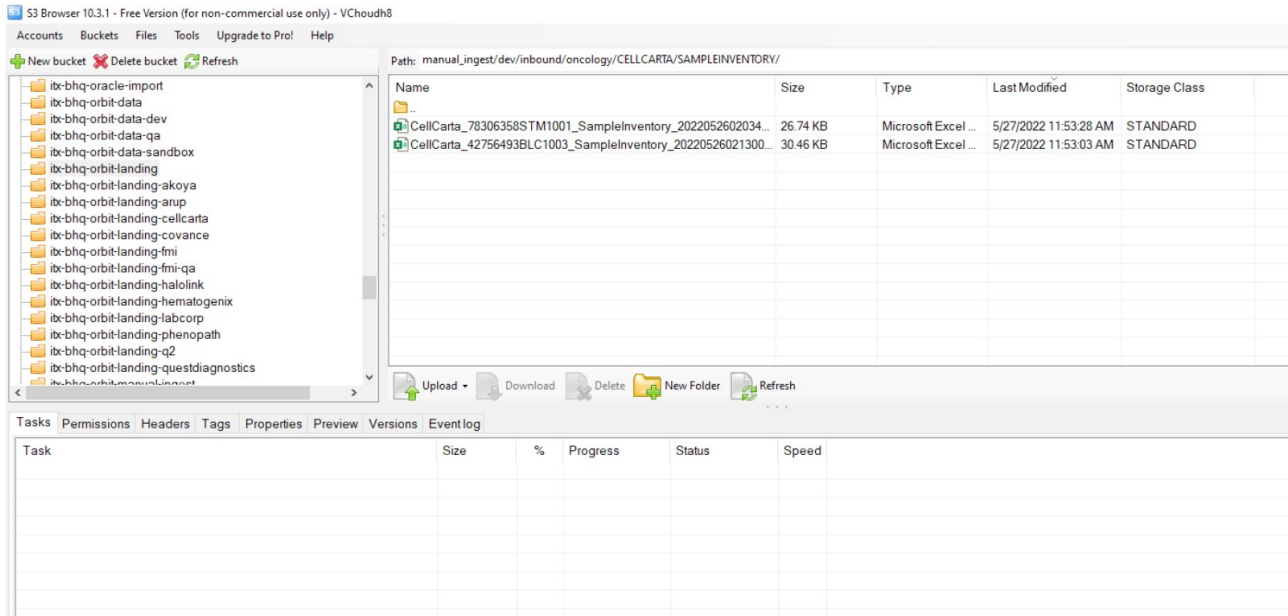
Key : Key is unique identifier for an object within a bucket. Every object in a bucket has exactly one key.

Metadata : Metadata is a set of name-value pairs that describe the object. These include some default metadata such as last modified date, Content-Type etc

Bucket policy : Bucket policy are JSON format document that define access and permissions on bucket level

Encryption : Uses 256 bit Advanced Encryption Standard (AES 256)





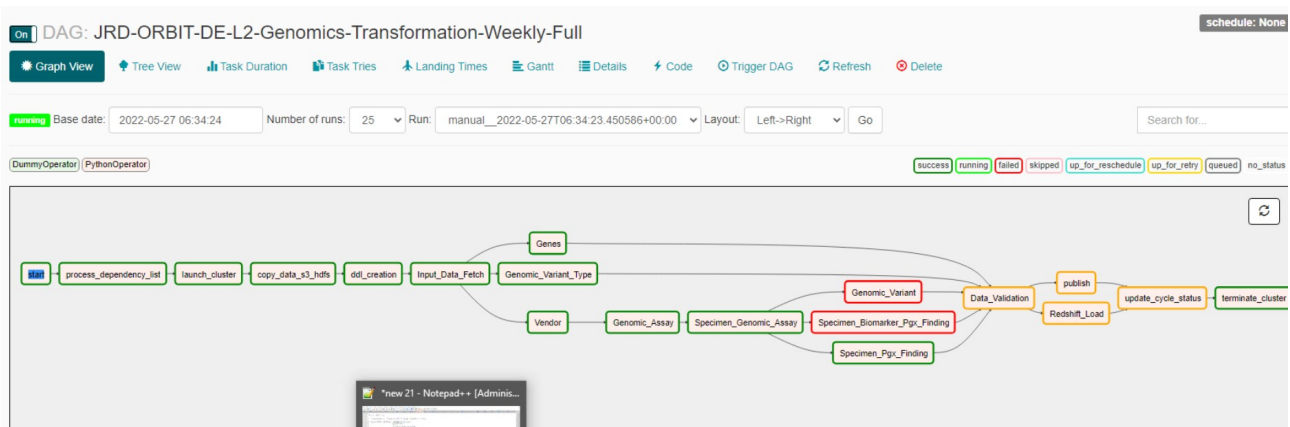
## AIRFLOW

Airflow is a platform to programmatically author, schedule and monitor workflows. Use Airflow to author workflows as Directed Acyclic Graphs (DAGs) of tasks. The Airflow scheduler executes your tasks on an array of workers while following the specified dependencies. Rich command line utilities make performing complex surgeries on DAGs a snap. The rich user interface makes it easy to visualize pipelines running in production, monitor progress, and troubleshoot issues when needed.

When workflows are defined as code, they become more maintainable, versionable, testable, and collaborative.

- **Dynamic:** Airflow pipelines are configuration as code (Python), allowing for dynamic pipeline generation. This allows for writing code that instantiates pipelines dynamically.
- **Extensible:** Easily define your own operators, executors and extend the library so that it fits the level of abstraction that suits your environment.
- **Elegant:** Airflow pipelines are lean and explicit. Parameterizing your scripts is built into the core of Airflow using the powerful **Jinja** templating engine.
- **Scalable:** Airflow has a modular architecture and uses a message queue to orchestrate an arbitrary number of workers. Airflow is ready to scale to infinity.

## L2



## Log options

## Components of Apache Airflow

- **DAG:** It is the Directed Acyclic Graph – a collection of all the tasks that you want to run which is organized and shows the relationship between different tasks. It is defined in a python script.
- **Web Server:** It is the user interface built on the Flask. It allows us to monitor the status of the DAGs and trigger them.
- **Metadata Database:** Airflow stores the status of all the tasks in a database and do all read/write operations of a workflow from here.
- **Scheduler:** As the name suggests, this component is responsible for scheduling the execution of DAGs. It retrieves and updates the status of the task in the database.

## User Interface

Now that you have installed the Airflow, let's have a quick overview of some of the components of the user interface.

### DAGS VIEW

It is the default view of the user interface. This will list down all the DAGS present in your system. It will give you a summarized view of the DAGS like how many times a particular DAG was run successfully, how many times it failed, the last execution time, and some other useful links.

L1



## SQL

SQL is a PC information language intended for information recovery and the board on a connected site. SQL represents Structured Query Language. This instructional exercise will give you a fast beginning to SQL. It covers the greater part of the themes required for an essential comprehension of SQL and how it functions.

### Why Learn SQL?

SQL is a Formal Query Language, which is a definitive coding, to deceive and recover information put away seeing someone.

SQL is the standard language of the Relational Database System. All Relational Database Management Systems (RDMS, for example, MySQL, MS Access, Oracle, Sybase, Informix, Postgres and SQL Server use SQL as their standard information language.

Likewise, they utilize an assortment of dialects, for example, -

MS SQL Server utilizes T-SQL,

Prophet utilizes PL/SQL,



The MS Access SQL type is called JET SQL (customary configuration) and so on

## SQL Requests

As referenced before, SQL is quite possibly the most broadly utilized inquiry dialects as far as information. I will list a couple of them here:

Permits clients to get to information in data set administration frameworks.

Permits clients to decipher information.

Permits clients to characterize information in a data set and utilize that information.

Permits implanting in different dialects utilizing SQL modules, libraries and pre-compilers.

Permits clients to make and drop information and tables.

Permits clients to make sees, measure cycles, and information base capacities.

Permits clients to set authorizations on tables, cycles and perspectives.

## **Audience**

SQL is intended for beginners to assist them perceive the essential ideas and advanced ideas associated with SQL languages. this can offer you enough insight into the assorted parts of SQL and relevant examples. Prerequisite Before you begin obtaining aware of the assorted forms of examples provided during this tutorial, i believe you already apprehend what a info is, particularly RDBMS and what a artificial language is.

RDBMS is the short form for Relational Database Management system.

## SQL Connection Page

Manage Server Connections

MySQL Connections  
Orbit\_dev

Connection Name: Orbit\_dev

Connection Remote Management System Profile

Connection Method: Standard (TCP/IP) Method to use to connect to the RDBMS

Parameters SSL Advanced

Hostname: [Redacted] Port: [Redacted] Name or IP address of the server host - and TCP/IP port.

Username: [Redacted] Name of the user to connect with.

Password: [Redacted] Store in Vault ... Clear The user's password. Will be requested later if it's not set.

Default Schema: [Redacted] The schema to use as default schema. Leave blank to select it later.

New Delete Duplicate Move Up Move Down Test Connection Close

## SQL Sheets

MySQL Workbench

Orbit\_dev (orchestration\_db) x Orbit\_dev x

File Edit View Query Database Server Tools Scripting Help

orbit\_dev SQL File 3\* SQL File 4\*

Limit to 1000 rows

```

1 • select COUNT(*), STUDY_ID from ctl_src_tgt_mapping_edc where tbl_name = '[Redacted]' and form_name = '[Redacted]' group by study_id
2
3 UPDATE ctl_src_tgt_mapping_edc SET SRC_COL = 'cast(NULL AS string)' where tgt_col = 'consent_value'
4
5 select * from ctl_src_tgt_mapping_edc where tbl_name = '[Redacted]' and study_id = '61186372NSC3001' and form_name = 'BE_GL_900'
6
7 insert into ctl_src_tgt_mapping_edc values('2', '[Redacted]', '[Redacted]', 'bestdat', 'bestdat', 'edc_rave_be_gl', '1', 'Y', now(), now());
8 • insert into ctl_src_tgt_mapping_edc values('2', '[Redacted]', '[Redacted]', 'subject', 'subject', 'edc_rave_be_gl', '2', 'Y', now(), now());

```

PER	application_id	study_id	form_name	src_col	tgt_col	tbl_name	column_sequence_number	active_flg	created_date	updated_date
1	61186372NSC3001	61186372NSC3001	BE_GL_900	bestdat	bestdat	edc_rave_be_gl	1	Y	2022-05-27 09:58:29	2022-05-27 09:58:29
2	61186372NSC3001	61186372NSC3001	BE_GL_900	subject	subject	edc_rave_be_gl	2	Y	2022-05-27 09:58:29	2022-05-27 09:58:29
3	61186372NSC3001	61186372NSC3001	BE_GL_900	sitenumber	sitenumber	edc_rave_be_gl	3	Y	2022-05-27 09:58:29	2022-05-27 09:58:29
4	61186372NSC3001	61186372NSC3001	BE_GL_900	project	project	edc_rave_be_gl	4	Y	2022-05-27 09:58:29	2022-05-27 09:58:29
5	61186372NSC3001	61186372NSC3001	BE_GL_900	datapage_name	datapage_name	edc_rave_be_gl	5	Y	2022-05-27 09:58:29	2022-05-27 09:58:29
6	61186372NSC3001	61186372NSC3001	BE_GL_900	form_name	form_name	edc_rave_be_gl	6	Y	2022-05-27 09:58:29	2022-05-27 09:58:29
7	61186372NSC3001	61186372NSC3001	BE_GL_900	CAST(null as string)	consent_value	edc_rave_be_gl	7	Y	2022-05-27 09:58:29	2022-05-27 09:58:29

Result Grid Filter Rows: Export: Wrap Cell Content: I

ct1\_src\_tgt\_mapping\_edc30 x

Output

Action Output

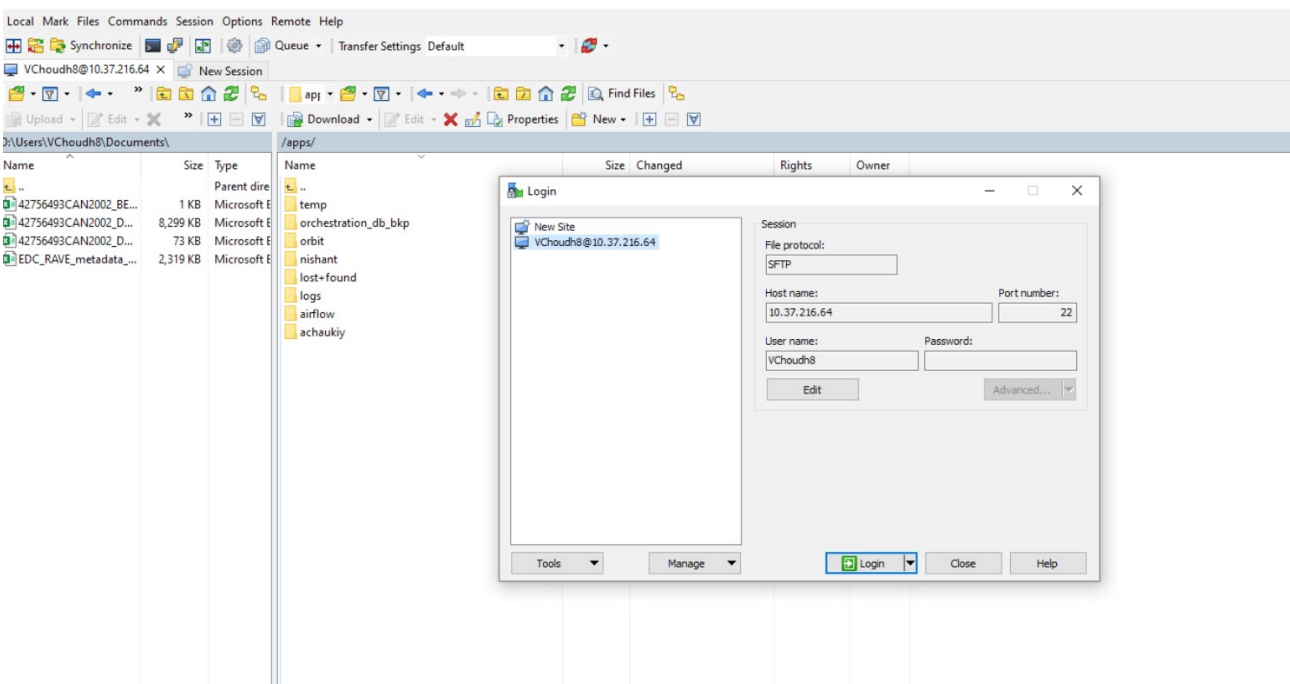
# WINSCP

WinSCP is a popular free SFTP and FTP client for Windows, a powerful file manager that will improve your productivity. It offers an easy to use GUI to copy files between a local and remote computer using multiple protocols: Amazon S3, FTP, FTPS, SCP, SFTP or WebDAV. Power users can automate WinSCP using .NET assembly. WinSCP is available in English and many other languages.

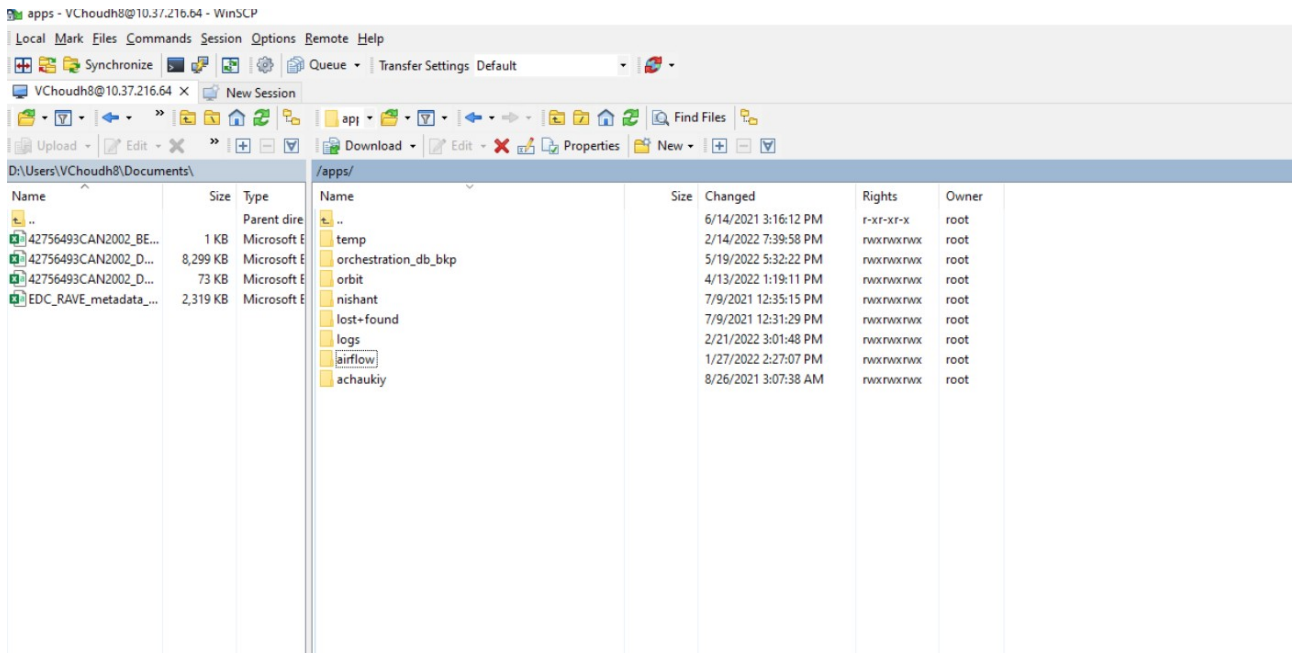
## Features

- All common operations with files
- Support for Amazon S3, FTP, FTPS, SCP, SFTP or WebDAV
- Directory synchronization in several semi or fully automatic ways
- Binary and text transfer modes, transfer queue and background transfers
- Graphical user interface with choice of Windows Explorer-like or Norton Commander-like interfaces
- Integrated text editor and support for external text editors
- Integration with Windows, PuTTY and other applications
- Scripting and task automation with batch file scripting, command-line interface and .NET assembly
- Optionally supports portable operation using a configuration file instead of registry entries, suitable for operation from removable media
- A comprehensive WinSCP documentation and an active user forum are freely accessible at WinSCP site.

### WinSCP Connection



## WinSCP file



## PUTTY

PuTTY is a free implementation of SSH (and telnet) for PCs running Microsoft Windows (it also includes an xterm terminal emulator). You will find PuTTY useful if you want to access an account on a Unix or other multi-user system from a PC (for example your own or one in an internet cafe). In the case of the University of Sussex, this would include users of the central ITS Unix system and of the Unix systems provided by some other schools and units. The central ITS Unix system supports the use of SSH clients. Users of other systems should ask their system administrator if SSH is supported.

PuTTY is an alternative to telnet clients. Its primary advantage is that SSH provides a secure, encrypted connection to the remote system. It's also small and self-contained and can be carried around on a floppy disk. This makes it ideal for accessing Sussex systems securely from other locations on the public Internet.

Although still only a beta (pre-production) release, PuTTY is full-featured, stable and its terminal emulation is good. Since version 0.53, it also supports the passthrough printing feature of VT100 terminals.



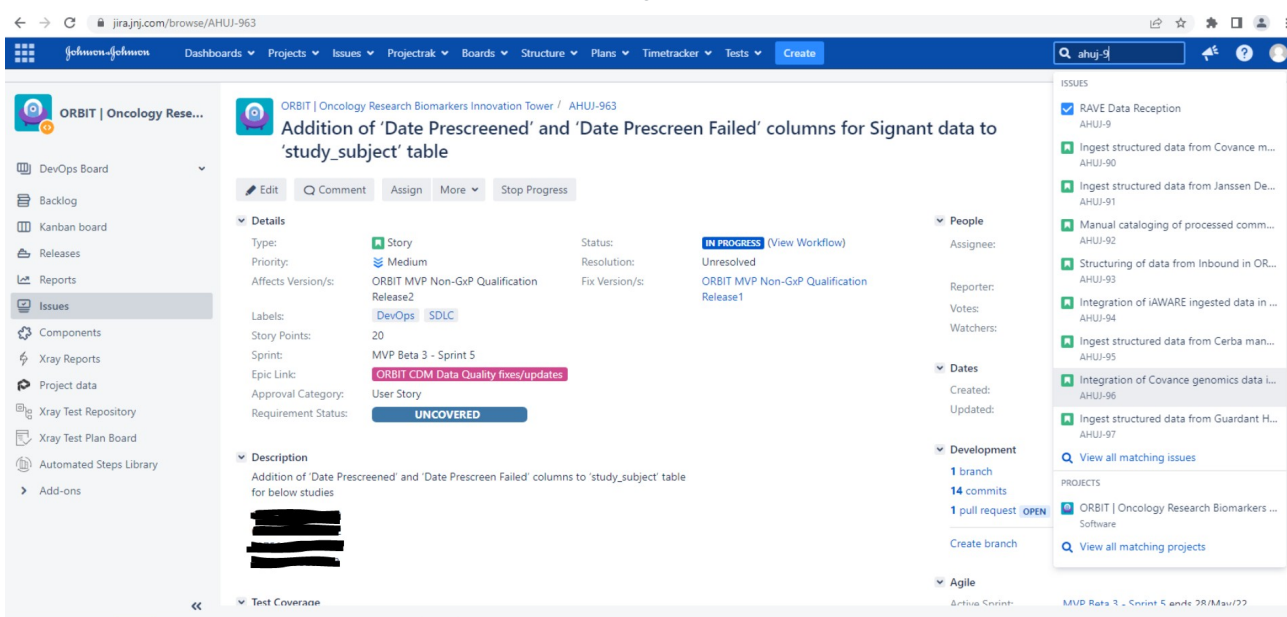
## BITBUCKET

BitBucket is a cloud-based service that helps developers store and manage their code, as well as track and control the changes to their code. BitBucket provides a cloud-based Git repository hosting service. Its interface is user-friendly enough so even novice coders can take advantage of Git. We generally require a bit more technical knowledge and use of the command line to use Git alone. Additionally, BitBucket provides a variety of services like it gives teams to collaborate and create projects, test and deploy the code.

BitBucket Advantages:

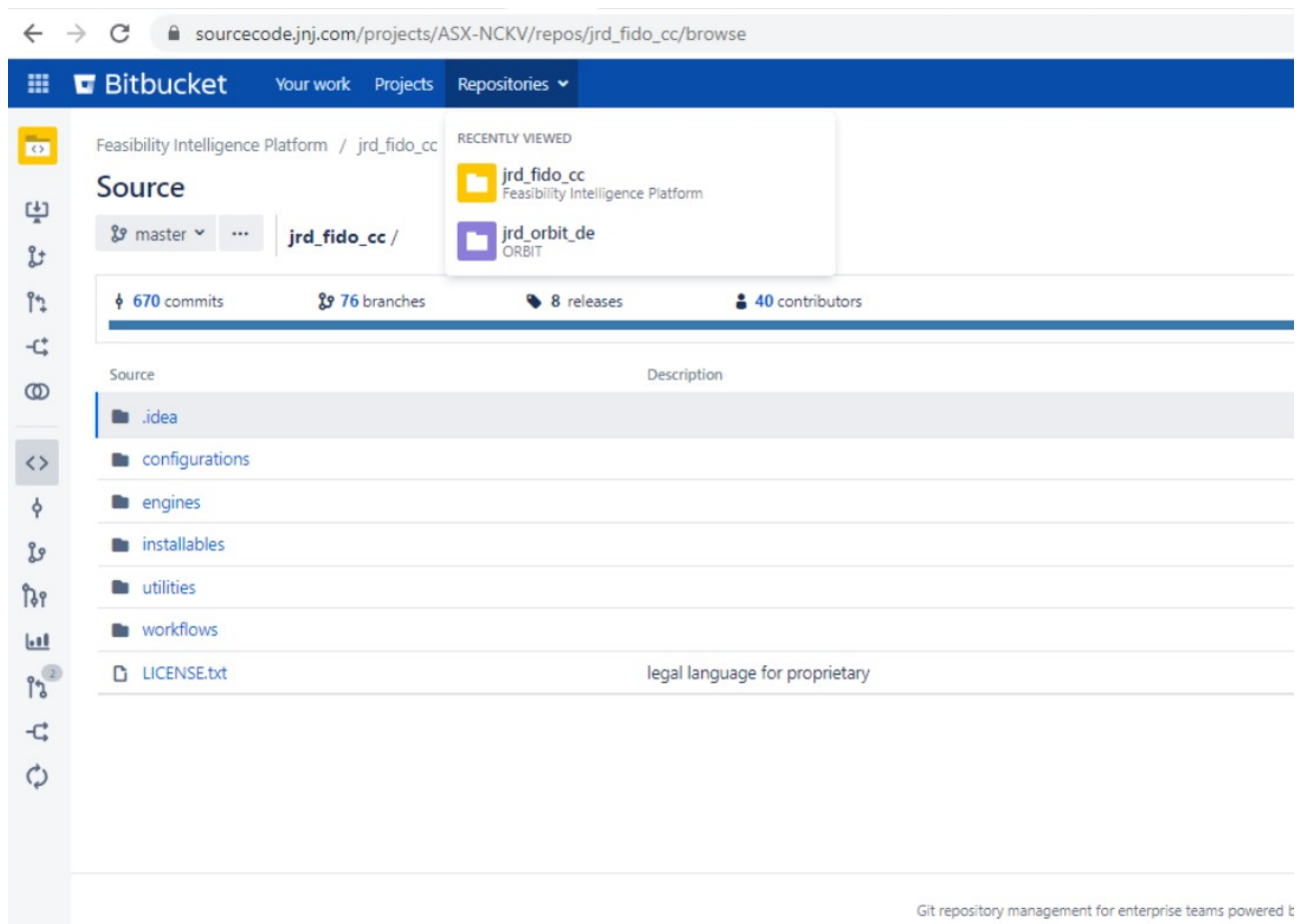
- Query Management System – The query management system provided in BitBucket makes the work easy by providing an integrated system.
- Authentication via Github – BitBucket allows the user to authenticate and add social media via Github.
- Integrated JIRA Tool – BitBucket also supports Bug tracking tool JIRA. It automatically detects bugs in the code and updates the information when a new commit is performed.
- Import Repositories – We can also import old repositories from third-party services.
- Pricing – BitBucket is free for small teams (0-5). You can use an unlimited number of free private repositories. The students and teachers can also get the free benefits of BitBucket.

## JIRA



The screenshot displays the JIRA web interface for a project named 'ORBIT | Oncology Research Biomarkers Innovation Tower / AHUJ-963'. The main issue title is 'Addition of 'Date Prescreened' and 'Date Prescreen Failed' columns for Signant data to 'study\_subject' table'. The issue is currently in the 'IN PROGRESS' state. The 'Details' section shows the issue is a 'Story' with a 'Medium' priority, affecting 'ORBIT MVP Non-GxP Qualification Release2'. It is labeled with 'DevOps' and 'SDLC', and is assigned to 'User Story'. The 'Description' section contains the text: 'Addition of 'Date Prescreened' and 'Date Prescreen Failed' columns to 'study\_subject' table for below studies'. The 'Issues' sidebar on the right lists several other issues, including 'RAVE Data Reception AHUJ-9', 'Ingest structured data from Covance m... AHUJ-90', and 'Integration of iAWARE ingested data in ... AHUJ-94'. The 'Development' sidebar shows '1 branch', '14 commits', and '1 pull request OPEN'. The 'Agile' sidebar shows 'MVP Beta 3 - Sprint 5 and 6 28/Mau/22'.

## BITBUCKET



## EXCEL

Excel is a spreadsheet program from Microsoft and a component of its Office product group for business applications. Microsoft Excel enables users to format, organize and calculate data in a spreadsheet.

By organizing data using software like Excel, data analysts and other users can make information easier to view as data is added or changed. Excel contains a large number of boxes called cells that are ordered in rows and columns. Data is placed in these cells.

Excel is a part of the Microsoft Office and is compatible with other applications in the Office suite. The spreadsheet software is available for Windows, macOS, Android and iOS platforms.

We use Microsoft Excel for the following:

- collection and verification of business data;
- business analysis;
- data entry and storage;
- data analysis;
- performance reporting;

- strategic analysis;
- accounting and budgeting;
- administrative and managerial management;
- account management;
- office administration.

## Mapping Excel Sheets

The screenshot shows an Excel spreadsheet with the following columns: Target Table, Source Table, Source Column, Target Column, Transformation Logic, and Comments. The rows contain detailed mappings for various data fields, including study and subject identifiers, dates, and specific data points. Some rows are highlighted in yellow, and there are blue circular markers in some cells. The spreadsheet is titled 'EDC\_RAVE\_FORM' and is part of a larger workbook named 'L2\_Study\_Subject'.

## L1 config in Exel

The screenshot shows an Excel spreadsheet with the following columns: Item ID, Item Name, Item Description, Item Type, Item Status, Item Location, and Item Category. The rows contain detailed configuration items, including study and subject identifiers, dates, and specific data points. The spreadsheet is titled 'L1 Configuration v1.0.xlsm' and is part of a larger workbook named 'S3 folder structure by files'. The formula bar shows a VLOOKUP function: '=VLOOKUP(E20,Info!\$M\$8:\$N\$17,2,0)'. The spreadsheet is titled 'L1 Configuration v1.0.xlsm' and is part of a larger workbook named 'S3 folder structure by files'.



## **Datasets**

A data set (or database) may be a assortment of knowledge. within the case of table knowledge, a collection of knowledge corresponds to at least one or additional keep tables, during which all columns of a table represent an explicit variable, and every row corresponds to a given recording of the desired knowledge. an information set calculates the values of every variable object, like the peak associate degreed weight of an object, in every member of the info set. every worth is understood as a data point. knowledge sets might also contain collections of texts or files.

## **Unit Testing**

UNIT TESTING could be a form of software package take a look at within which individual units or software package parts square measure tested. the aim is to confirm that every unit of software package code works needless to say. Unit testing is finished throughout the event (coding phase) of the program by developers. The Unit take a look at separates the code section and verifies its accuracy. A unit are often a private perform, method, process, module, or object. Unit testing, the testing method exploitation that individual modules square measure tested to see if there square measure any engineers themselves. it's to try and do with the accuracy of the performance of the freelance modules. The main purpose of separating every system unit is to spot, analyze and proper errors. Unit testing is vital as a result of software package developers generally attempt to save time doing tiny unit tests and this can be a story as a result of improper unit testing results in price overruns throughout system testing, take a look at integration and Beta testing once the app is made. If correct unit testing is finished on initial growth, then you save time and cash within the finish.

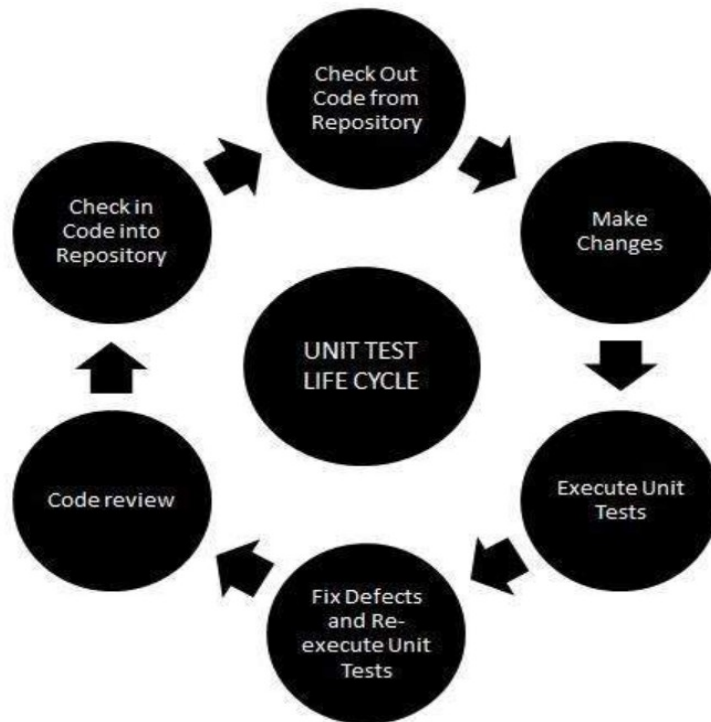
Nit Test - Benefits:

Reduce Errors in newly developed features or reduce bugs when changing existing functionality.

Reduce Test Costs as Disability is considered early.

It improves structure and allows you to reuse code better.

Unit Testing, combined with construction also provides construction quality as well.



To perform Unit testing, developer writes a code section to test the specific functions in a software applications. Developers could also splits this function into a rigorous test that produces unnecessary dependencies between the test function and other units so that the dependency is not eliminated. Engineers often use the UnitTest framework to develop automatic test case testing cases.

Unit testing has two types

By hand

Default

Unit testing is usually automatic but could still be done manually. Software Engineering does not like each other but automation is very popular. The unit standard method used can use the instruction manual for each step.

Under default mode-

The builder composes the code segment of the application to test the capacity. They will remark later lastly erase the check code once the application has been introduced.

The designer may isolate the work to test it all the more proficiently. This is a finished testing technique for units that incorporate reordering code into its test site instead of its current circumstance. Isolating code assists with uncovering superfluous reliance between the tried code and different units or information spaces in the item. This reliance can be eliminated.

The coder normally utilizes the UnitTest System to make programmed experiments.

Utilizing the default system, engineer code techniques in the test to confirm the legitimacy of the code. During experiments, the logs of the system fall flat in experiments. Numerous organizations will naturally stamp and report, to put it plainly, these bombed experiments. Contingent upon the size of the disappointment, the system may suspend the following test.

The work process of the Testing Unit is 1) Make Experiments 2) Audit/Adjust 3) Standard 4) Perform Testing Cases.

#### Good Practice Testing Unit

Unit check cases ought to be freelance. If there area unit any enhancements or changes in needs, unit check cases shouldn't be affected. Check just one code at a time. Follow clear and consistent style conferences for your unit's testing In the event that a code is modified in any module, certify there's a compatible legal action section of the module, and therefore the module passes the tests before switch use Disruptions known throughout unit testing ought to be corrected before moving on to future stage in SDLC.

Accept the 'test as yours code" method. When writing the code beyond testing, many methods should check for errors.

## RESULTS AND DISCUSSION

Internship study:

- Acquired expertise in SQL operations. As is the case with data analysis tasks mainly using SQL queries, acquired expertise in writing different types of sql questions based on the use case.
- Gained Skills in Business intelligence software and data analysis models. Learnt about creating Dossiers and standard reports on Microstrategy platform.
- Communication: The project had a large team of more than 20 people, working with them a group of this size needs effective and efficient communication. So, an internship help to acquire the ability to communicate in both written and oral forms.
- Understanding US Pharma Industry Performance: working on client projects has helped understand the functioning of the industry. The type of data used in this industry as well the importance of certain KPIs to certain types of users while another set of KPIs may be useful to other user sets depending on the management category.

### **Project results:**

- A project is an ongoing project that helps the client make informed decisions based on the data.
- The project contains many modules, which can work in one or different areas depending on the operating case. The client can access these programs from the system the choice could be a Web Browser or Ipad application. Applications perform on both platforms.

## CONCLUSION

1. Reports are the best way to understand data, and it also helps to improve the system.
2. Automation is best for reducing a work , when a specific operation needs to be done based on some predefined condition.
3. Contracting is an essential component of the pharma industry.
4. Understanding the data is essential to performing data analysis and taking decisions.
5. A work or task tracker is always a good practice while moving on with the project.

## References

<https://onlinelibrary.wiley.com/doi/abs/10.1111/1467-6486.t01-2-00002>

[https://en.wikipedia.org/wiki/MySQL\\_Workbench](https://en.wikipedia.org/wiki/MySQL_Workbench)

[https://onlinelibrary.wiley.com/doi/abs/10.1111/1467-6486.t01-2-00002\](https://onlinelibrary.wiley.com/doi/abs/10.1111/1467-6486.t01-2-00002)

<https://aisel.aisnet.org/cais/vol14/iss1/46/>

[https://www.researchgate.net/profile/Bill-Mckelvey/publication/254397115\\_MicroStrategy\\_from\\_MacroLeadership\\_Distributed\\_Intelligence\\_via\\_New\\_Science/links/544be5380cf2bcc9b1d6c1f1/MicroStrategy-from-MacroLeadership-Distributed-Intelligence-via-New-Science.pdf](https://www.researchgate.net/profile/Bill-Mckelvey/publication/254397115_MicroStrategy_from_MacroLeadership_Distributed_Intelligence_via_New_Science/links/544be5380cf2bcc9b1d6c1f1/MicroStrategy-from-MacroLeadership-Distributed-Intelligence-via-New-Science.pdf)

[https://link.springer.com/chapter/10.1007/978-981-13-1927-3\\_53](https://link.springer.com/chapter/10.1007/978-981-13-1927-3_53)

<https://profisee.com/master-data-management-what-why-how-who/>

[https://www2.microstrategy.com/producthelp/Current/InMemoryAnalytics/WebHelp/Lang\\_1033/Content/InMemoryAnalysis/Creating\\_Intelligent\\_Cubes.htm](https://www2.microstrategy.com/producthelp/Current/InMemoryAnalytics/WebHelp/Lang_1033/Content/InMemoryAnalysis/Creating_Intelligent_Cubes.htm)

[https://www.tutorialspoint.com/microstrategy/microstrategy\\_creating\\_metrics.htm#:~:text=Metrics%20in%20MicroStrategy%20are%20the,a%20column%20in%20source%20data.&text=Creation%20of%20a%20metric%20involves,functions%20already%20available%20in%20MicroStrategy.](https://www.tutorialspoint.com/microstrategy/microstrategy_creating_metrics.htm#:~:text=Metrics%20in%20MicroStrategy%20are%20the,a%20column%20in%20source%20data.&text=Creation%20of%20a%20metric%20involves,functions%20already%20available%20in%20MicroStrategy.)