



STUDENT DATA MANAGEMENT

Project

DATA GYAN
Batch – 68 (Freshers)

INTRODUCTION:

- **PROJECT TITLE** - Student Data Management for Data Gyan
- **PROJECT PURPOSE** – To highlight the steps to acquire the student data and maintain a student Database which would help in targeting potential students.
- **SOURCE OF DATA** – Internal Data Source such as Registration data from the Organization and External Data Source such as Facebook, LinkedIn, & Naukri.
- **TOOLS USED** – Excel, MS SQL Server, Tableau, & PowerPoint Presentation.

Aim of the Project –

- To acquire student data from across various source such as LinkedIn, Facebook, Naukri and Internal Data is collected into single Database.
- This can help in identifying potential students for the company for its future business.
- Thus, different marketing campaign and advertising strategy can be directed towards these student segments
- To have a specified Student Database, where each Information along with several other can be maintained for future purpose.

Sequential steps for carrying out the Project –

THE DATA ANALYSIS PROCESS

Step 1:

Define the question

Step 2:

Collect the data

Step 3:

Clean the data

Step 4:

Analyze the data

Step 5:

Visualize and share your findings

PHASE 1 : DATA COLLECTION

Data Collection Steps –

- Data pertaining to various attributes such as name, city, states, course pursued, skills known were collected in Excel Sheets from external sources as LinkedIn, Facebook, & Naukri.com as well as from Internal Source such as Registered Student Information at Data Gyan.
- Initially, a template in MS-Excel for the Master Dataset Sheet was prepared and a few important attributes which need to be captured in it and subsequently into the Database were finalised after thorough discussion with the team.

DATA COLLECTION METHODS

- | | |
|-------------------|--------------------------------------|
| ✓ Surveys | ✓ Marketing analysis |
| ✓ Interviews | ✓ Monitoring social media |
| ✓ Focus groups | ✓ Subscription and registration data |
| ✓ Online tracking | ✓ Monitoring in-store traffic |



Data Preparation Steps –

- **Facebook Dataset –** Attributes such ID, FIRST NAME, LAST NAME, PROFESSION, DOB, AGE, EMAIL ID, GENDER, PHONE NUMBER, BLOOD GROUP, LANGUAGE KNOWN, HOBBIES, CITY, COUNTRY, PIN-CODE, BRANCH, SCHOOL NAME, LAST ACTIVE, and PROFILE LINK as were collected from Facebook for 200 students into another work sheet of the same work sheet shown below.

A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	Last_Ac
1	ID	First Name	Last Name	Profession	Date of Birth	Age	Email	Gender	Phone_Number	Blood Group	Language Known	Hobbies	City	Country	Pincode	Branch	School_Name
2	1	Charlotte	Jobson	Unemployed	2/16/2002	25	Charlotte_Jobson7251@sveldo.biz	Female	9195234578	A+	Filipino	Shopping	Boston	Nauru	456941	Power Engg	High school
3	2	Grace	Miller	Unemployed	02-02-2010	28	Grace.Miller5085@supunk.biz	Female	7854213698	A-	Sotho	Cricket	New Orleans	Romania	635943	Aerospace Engg	Gymnasium
4	3	Tania	Giles	Unemployed	11-03-2008	28	Tania_Giles874@famism.biz	Female	9625478541	B+	Bosnian	PUBG	Huntsville	Burundi	848707	Civil Engg	University college
5	4	Louise	Bullock	Employed	12/31/1995	24	Louise_Bullock4253@nickia.com	Female	8932578914	B-	Telugu	Travelling	Jacksonville	Egypt	952585	Telecommunication	Graduate school
6	5	Aileen	Cattell	Unemployed	9/26/1998	22	Aileen_Cattell6600@jiman.org	Female	8896541237	O+	Somali	Singing	Orlando	Pakistan	696785	EEE	University-preparatory school
7	6	Angela	Young	Student	12/21/1994	25	Angela_Young7157@twipet.com	Female	6023697412	O-	Estonian	Shopping	Berna	Congo, Republic of the	490164	Biotech Engg	Public university
8	7	Chloe	Amstead	Student	7/23/1999	21	Chloe_Amstead5097@qater.org	Female	7412369850	AB+	Somali	Singing	Columbus	Singapore	691840	Industrial Engg	Technical college
9	8	Estrella	Fall	Student	7/24/2007	25	Estrella_Fall8124@tonsy.org	Female	8745123690	AB-	Japanese	Cricket	London	Congo, Republic of the	486037	EE	Gymnasium
10	9	William	Jobson	Employed	01-01-2002	24	William_Jobson9244@vetan.org	Male	8030142057	A+	German	Football	Saint Paul	Guinea	746630	Power Engg	University-preparatory school
11	10	Sylvia	Nielson	Unemployed	10-02-2010	27	Sylvia_Nielson289@twace.org	Female	9350061247	A-	French	Painting	Berna	Mozambique	547430	Biotech Engg	Technical college
12	11	Logan	Stanton	Student	02-10-1994	24	Logan_Stanton6191@vetan.org	Male	8271465390	O+	Fijian	Singing	Stockton	Nauru	530686	Power Engg	Public university
13	12	Logan	Fisher	Unemployed	6/23/2010	28	Logan_Fisher1707@famism.biz	Male	7896541230	O-	Mongolian	Dacing	Atlanta	Moldova	588164	EEE	Corporate university
14	13	Violet	Hunt	Unemployed	9/18/2010	20	Violet_Hunt4089@irrepsy.com	Female	8056301478	B+	Tswana	Cricket	Kansas City	Brunei	957927	Petroleum Engg	Corporate university
15	14	Chester	Lowe	Unemployed	9/23/1999	20	Chester_Lowe6789@bulufly.com	Male	9632105874	B-	Haitian Creole	Dacing	Cincinnati	Venezuela	900622	ECE	Public university
16	15	Chloe	Poole	Unemployed	03-11-2009	22	Chloe_Poole860@grannar.com	Female	9325064781	AB+	Finnish	Football	Amarillo	Montenegro	720426	Telecommunication	Graduate school
17	16	Leroy	Bowen	Student	01-10-2010	21	Leroy_Bowen597@joiniaa.com	Male	8542103697	AB-	Hebrew	Shopping	Huntsville	Antigua and Barbuda	775083	Mech Engg	Junior college
18	17	Chris	Levin	Student	11/30/2008	25	Chris_Levin8225@acrlt.org	Male	9632014587	A+	Punjabi	Travelling	Philadelphia	Senegal	851267	Aeronautical Engg	University college
19	18	Harriet	Larkin	Employed	8/19/2003	25	Harriet_Larkin7410@twipet.com	Female	7412509630	A-	Maori	Painting	Sacramento	Japan	492109	Aerospace Engg	Junior college

- LinkedIn Dataset –** Attributes such ID, FIRST NAME, LAST NAME, AGE, GENDER, QUALIFICATION, DEPARTMENT, DESIGNATION, EXPERIENCE, COURSE INTERESTED, SKILLS, CITY, CONTACT, EMAIL ID, STATE, COUNTRY, PASS OUT, MONTH and PROOF TYPE as were collected from LinkedIn for 200 students into another work sheet of the same work sheet shown below.

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P
1	ID	FNAME	LNAME	AGE	GENDER	QUALIFICATION	DEPARTMENT	DESIGNATION	EXPERIENCE	COURSE INTERESTED	SKILLS	CITY	CONTACT	EMAIL	STATE	COUN
2	1	rohit	patnaik	24	m	B.TECH	CSE	FRESHER	0	software	java,c++,android,HTML	coimbatore	+91 788 774 8554	patnakrohit21@gmail.com	Tamil-Nadu	India
3	2	sumit	ghosh	21	m	B.TECH	ME	FRESHER	0	Database	fusion360,Ansys,CNC	kochi	+91 788 774 8554	sumit@yahoo.com	Kerala	India
4	3	amit	rampal	23	m	B.TECH	ECE	PROFESSIONAL	5	Business Analytics	Embedded design,java,html	guntur	+91 61279 39726	amit102@gmail.com	AP	India
5	4	ram	gupta	20	m	M.TECH	CIVIL	FRESHER	0	Data analytics	beam designing,manufacturing	balasore	+91 96864 45492	ram0201@gmail.com	Orissa	India
6	5	shyam	lal	25	m	B.SC	MATHS	FRESHER	0	Data Analytics	phy,chem	howrah		shayam887@gmail.com	WB	India
7	6	raju	srivastav	21	m	B.TECH	ECE	FRESHER	0	Programming	java,c++,SQL,android,HTML	allahabad	+91 65797 03594	rajusrusv@yahoo.com	Uttar Pradesh	India
8	7	bablu	roy	21	m	MCA	CSE	FRESHER	0	Programming	java,.net,css,	indore	+91 86117 47725	bablu7@hotmail.com	Madhya Pradesh	India
9	8	denver	silva	23	m	B.TECH	CSE	FRESHER	0	Data Analytics	java,.net,spring,php	New York	+1 6127975653	denversilva221@yahoo.com	Michigan	USA
10	9	ola	gunju	26	m	BCA	CSE	PROFESSIONAL	2	Programming	java,.net,css,php,html,	Lagos	+27 6127920098	ola@gmail.com	Cape Town	South Africa
11	10	pooja	sharma	25	f	BCA	CSE	FRESHER	0	Data Analytics	java,.net,css,php	tinsukia	+91 61279 58492	pooja23@gmail.com	Assam	India
12	11	rakesh	singh	28	m	B.TECH	CIVIL	FRESHER	0	software	beam designing,manufacturing	mumbai	+91 61279 75346	rakesh98@yahoo.com	Maharashtra	India
13	12	jitesh	rai	22	m	B.TECH	CSE	FRESHER	0	Database	salesforce,python	ahmedabad	+91 61279 38842	j	Gujarat	India
14	13	rohit	pawar	20	m	B.TECH	ECE	FRESHER	0	Business Analytics		jaipur	+91 61279 93913	rohitpr@yahoo.com	Rajasthan	India
15	14	pranav	sahu	28	m	M.TECH	CSE	FRESHER	0	Data analytics	python,R,C	chandigarh	+91 81603 91766	pranav89@hotmail.com	Punjab	India
16	15	swagat	bhati	21	m	B.SC		FRESHER	0	Data Analytics	Data structure, c++,java,jvm	palwal	+91 61279 81077	swaat44@gmail.com	Haryana	India
17	16	saroj	patra	22	m	B.TECH	ME	FRESHER	0	Programming	angular java,python,data structure	ranchi	+91 68894 07184	saroj5@gmail.com	Jharkhand	India
18	17	raghavendra	canthus	23	m	MCA	CSE	PROFESSIONAL	1	Data Analytics	java script,python,perl,kotlin	patna	+91 79798 29606	raghu32@yahoo.com	Bihar	India
19	18	abhishhek	awasthi	21	m	B.TECH	CIVIL	FRESHER	0	Programming	matlab,sql,perl	mangalore	+91 61279 29161	awasthi32@yahoo.com	KA	India
20	19	madhur	sharma	25	m	BCA	CSE	FRESHER	0	Data Analytics	scala,perl,kotlin,javascript,html,cs	chennai	+91 80794 35958	sharmamadhur21@yahoo.com	TN	India
21	20	shalini	bansal	24	f	BCA	ECE	FRESHER	0	Software		kochi	+91 80794 30885	shalini@gmail.com	Kerala	India
22	21	ritu	kumari	21	f	B.TECH	CSE	PROFESSIONAL	3	Software		eluru	+91 80798 30885	rtu42@yahoo.com	AP	India

- Naukri Dataset –** Attributes such ID, FIRST NAME, LAST NAME, AGE, DOB, GENDER, CONTACT, EMAIL ID, QUALIFICATION, SPECIALIZATION, PASS OUT, ENROLLED COURSE, WORK EXPRIENCE, SKILL KNOWN, CITY, STATE, and COUNTRY as were collected from LinkedIn for 200 students into another work sheet of the same work sheet shown below.

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q
1	ID	First Name	Middle Nar	Last Name	Age	DOB	Gender	Mobile Number	Email ID	Qualification	Specialization	Passout Year	Enrollment Courses	Work Experience	Skills Known	City	State
2	1	Ram	Devi	Kumar	25	10-02-1995	Male	7974923455	ramkumardev1@gmail.com	B.Sc	Zoology	2018	Data Science	3 Years	C++	Bengaluru	Karnataka
3	2	Mohammae	Amir	Hussian	21	28/02/2000	Male	6269945120	mohamadir007@gmail.com	B.Tech	Computer Science	2022	Business Analyst	NA	Python	Puri	Odisha
4	3	Sunita		Singh	27	24/11/1994	Female	9963215456	singhsunita24@gmail.com	B.Ed	Computer Application	2016	Database	4 Years	C	Kannur	Madhya Prades
5	4	Sri		Das	30	04-12-1992	Male	8283848586	sridas1992@gmail.com	M.Sc	Physics	2013	Data Analytic	6 Years	Java	Indore	Bhubaneswar
6	5	Anita		Kaur	23	15/08/1998	Female	9596949391	anitakaur1508@gmail.com	BBA	Fincance	2021	Machine Learning	NA	MS SQL Server	Odisha	West Bengal
7	6	Santosh		Deshmukh	31	12-10-1990	Male	7578747971	shrmukh10santosh@gmail.cc	M.Tech	Data Science	2012	Business Analyst	8 Years	Django	Durg	Chhattisgarh
8	7	Sanjay		Jain	22	29/01/1999	Male	6268656065	sanjayjain01@gmail.com	BCA		2020	Machine Learning	NA	Udupi	Karnataka	Karnataka
9	8	Sunil		Lal	28	17/10/1993	Male	7878981542	sunila171093@gmail.com	MBA	Marketing	2015	Data Science	2 Years	Machine Learning	Nagpur	Maharashtra
10	9	Rajesh		Yadav	27	05-05-1994	Male	7985462132	yadvrajaesh1994@gmail.com	B.Com	Business Analytics	2016	Data Science	NA	PHP	Palakkad	Kerala
11	10	Gita	Ram	Singh	29	11-03-1992	Female	8589874163	gitaramsingh311@gmail.com	B.Ed	Computer Application	2014	Database	3 Years	CSS	Chennai	Tamil Nahu
12	11	Ramesh		Reddy	22	20/07/1999	Male	7389824565	reddy07ramesh@gmail.com	B.E.	Mechanical	2020		NA	Python	Jangaon	Telangana
13	12	Ashok		Kumari	24	07-06-1997	Male	7898456513	ashok797singh@gmail.com	M.Sc	Business Studies	2019	Machine Learning	NA	Java	Panaji	Goa
14	13	Rekha		Jain	21	13/02/2000	Female	7189886396	umarirekha2000@gmail.com	B.Tech	Civil	2021	Business Analyst	NA	C++	Bhopal	Madhya Prades
15	14	Manoj	K	Jain	25	26/06/1996	Male	9694921385	kalimanol0@gmail.com	BBA	Human Resources	2018	Database	10 Months	Tableau	Buxar	Bihar
16	15	Anil		Pathak	28	14/03/1993	Male	9173826455	anilpathak14@gmail.com	B.E.	Petro-Chemical	2015	Data Science	NA	Advanced Excel	Kolkata	Tamil Nahu
17	16	Suresh	Raja	Nair	30	27/07/1991	Male	8246937185	sureshrajana91@gmail.com	B.Ed	Banking Management	2013	Data Science	NA	Machine Learning	West Bengal	West Bengal
18	17	Vijay	Devi	Prasad	21	21/09/2000	Male	7463128533	vijayprasad109@gmail.com	Civil	2021	Data Analytic	NA	My SQL	Raipur	Chhattisgarh	
19	18	Lakshmi	Ben	Sharma	23	31/12/1998	Female	9897959663	armabenlakshmi98@gmail.cc	M.Sc	Statistics	2020	Machine Learning	NA	Machine Learning	Kozhikode	Kerala
20	19	Abdul	Shah	Ali	26	10-01-1995	Male	9969636115	abdulalishah95@gmail.com	B.Com	Fincance	2018	Data Analytic	3 Months	My SQL	Bhubaneswar	Odisha
21	20	Raju		Sahu	22	09-09-1999	Male	9595654545	raju99sahu@gmail.com	BCA		2020	Data Science	NA	MS SQL Server	Lucknow	Uttar Prades
22	21	Manju		Gupta	29	05-08-1992	Female	8589851122	manjugupta0805@gmail.com	B.Com	Business Analytics	2013	Data Analytic	NA	Java	Satavahanana University	Uttar Prades

- Internal Source Dataset –** Attributes such ID, FIRST NAME, LAST NAME, FULL NAME, GENDER, EMAIL ID, CONTACT, COURSE, CITY, STATE, COUNTRY, EDUCATION, SPECIALIZATION, MONTH, YEAR OF PASSING, UNIVERSITY, and EXPERIENCE as were collected from LinkedIn for 200 students into another work sheet of the same work sheet shown below.

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q		
1	ID	FIRST NAME	LAST NAME	FULL NAME	GENDER	EMAIL	CONTACT NO.	COURSE	CITY	STATE	COUNTRY	EDUCATION	SPECIALIZATION	MONTH	ID	PROOF	EXPERIENCE	YEAR OF PASSING	UNIVERSITY
2	1	logon	currie	logon currie	M	logon32@gmail.com	758695204	java	NEW YORK	NEW YORK	US	MSC	physics	6	pan card	6		2012	Jawaharlal Nehru Technological University
3	2	michel	paul	michel paul	M	paul98@hotmail.com	8524876992	python	CHICAGO	ILLINOIS	US	BSC	chemistry	5	aadhar card	3		2013	Jawaharlal Nehru Technological University
4	3	raj	sharma	rajsharma	M	rsharma9@yahoo.com	8763859421	sql	PATNA	Bihar	IN	BE	math	5	pan card	0		2017	Kakatiya University
5	4	ankit	khandelwal	ankitkhandelwal	M	rkhandelwal9@gmail.com	8764859421	excel	VILAG	AP	IN	BTECH	civil	6	voter id	2		2018	Krishna University
6	5	abhishik	sathpathy	abhishik sathpathy	M	sathpathy9@yahoo.com	8765859421	java	PUNE	MH	IN	BA	english	9	pan card	0		2015	Mahatma Gandhi University, Noida
7	6	anshika	rout	anshikarout	F	route9786@gmail.com	5626896632	c	HYDERABAD	TN	IN	MCA	computer science	5	aadhar card	1		2016	National Academy of Legal Studies and Research
8	7	minakshi	dey	minakshidey	F	medey9@gmail.com	6366996636	tableau	BANGLORE	KA	IN	MSC	physics	6	pan card	1		2019	Osmania University
9	8	surbhi	das	surbhidas	F	sdsas@yahoo.com	6965998858	c	BHUBNESWAR	OD	IN	BSC	chemistry	9	aadhar card	2		2018	Palamuru University
10	9	swati	ghose	swatiaghose	F	rgtose9@gmail.com	266588999	c	CUTTACK	OD	IN	BE	math	9	pan card	4		2017	Rajeev Ramaiah University of Technology
11	10	shrabani	sethi	shrabani sethi	F	ssethi9@yahoo.com	87610859421	pythont	VIJAG	TN	IN	BTECH	IT	5	aadhar card	1		2018	Rajiv Gandhi University of Knowledge Technology
12	11	anima	kar	animakar	F	animas5@rediffmail.com	6558885653	tableau	PUNE	MH	IN	BA	english	6	pan card	2		2016	Rajaleeswara University
13	12	saurav	joshi	sauravjoshi	M	sjoshi95@gmail.com	6962326568	sql	HYDERABAD	TN	IN	MCA	computer science	6	voter id	0		2014	Satavahanana University
14	13	ashish	chandlani	ashishchandlani	M	chandlani45@yahoo.com	15451212454	excel	KOLKATA	WB	IN	MSC	physics	5	pan card	3		2018	Sri Krishnadevaraya University
15	14	abhisek	malan	abhisek malan	M	abhisek45@gmail.com	4554212	java	PUNE	MH	IN	BSC	chemistry	5	aadhar card	0		2019	Sri Padmavati Mahila Viswavidyalaya
16	15	hitesh	das	hiteshdas	M	hiteshd45@gmail.com	5478731444	python	MUMBAI	MH	IN	BE	math	4	pan card	4		2020	Sri Venkateswara University
17	16	ankit	bhat	ankitbhat	M	medey9@gmail.com	255326545	c	KANPUR	UP	IN	BTECH	etc	5	pan card	1		2017	Sri Venkateswara University
18	17	ankita	lakhwande	ankitalakhwande	F	ankita@y.com	2978745144	python	PURI	OD	IN	BA	english	5	pan card	0		2016	Sri Venkateswara Vedic University
19	18	souvik	dey	souvikdey	M	medey9@gmail.com	2485454554	sql	KANPUR	UP	IN	MCA	computer science	6	aadhar card	4		2014	Sri Venkateswara Veterinary University
20	19	shanti	das	shantidas	F	fdey@yahoo.com	5268894477	tableau	BHUBNESWAR	OD	IN	MSC	physics	6	voter id	2		2014	Telangana University
21	20	dolly	monanta	dollymonanta	F	dolly88@rediffmail.com	247982647	python	KOLKATA	WB	IN	BSC	chemistry	8	pan card	5			

PHASE 2 : DATA CLEANING

Data forms the backbone of any data analytics you do. Regarding data, there are many things to go wrong – be it the construction, arrangement, formatting, spellings, duplication, extra spaces, and so on. To perform the data analytics properly we need various data cleaning techniques so that our data is ready for analysis. It's commonly said that,

“Data scientists spend 80% of their time cleaning and manipulating data and only 20% of their time actually analyzing it.”

What is Data Cleaning –

Data cleansing or data cleaning is the process of identifying and removing (or correcting) inaccurate records from a dataset, table, or database and refers to recognizing unfinished, unreliable, inaccurate, or non-relevant parts of the data and then restoring, remodelling, or removing the dirty or crude data.

Data cleaning techniques may be performed as batch processing through scripting or interactively with data cleansing tools.

After cleaning, a dataset should be uniform with other related datasets in the operation. The discrepancies identified or eliminated may have been basically caused by user entry mistakes, by corruption in storage or transmission, or by various data dictionary descriptions of similar items in various stores.

Data Cleaning Steps –

1. Get Rid of Extra Spaces.
2. Select and Treat All Blank Cells.
3. Convert Numbers Stored as Text into Numbers.
4. Remove Duplicates.
5. Highlight Errors.
6. Change Text to Lower/Upper/Proper Case.
7. Spell Check.
8. Delete all Formatting.
9. Removing the Short Names using VLOOKUP method.
10. Using Concatenation method for combining some split data.
11. Change the Number format into Date format.

Datasets After Cleaned –

• Facebook Dataset –

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P
1	ID	First Name	Last Name	Full Name	Profession	Date of Birth	Age	Email	Gender	Phone Number	Blood Group	Language Known	Hobbies	City	State	Country
2	1	Charlotte	Jobson	Charlotte Jobson	Unemployed	2/16/2002	25	Charlotte.Jobson7251@sveldo.biz	Female	9195234578	A+	Odia	Shopping	Bhubaneswar	Odisha	India
3	2	Grace	Miller	Grace Miller	Unemployed	02-02-2010	28	Grace_Miller5085@supunk.biz	Female	7854213698	A-	Odia	Cricket	Rourkela	Odisha	India
4	3	Tania	Giles	Tania Giles	Unemployed	11-03-2008	28	Tania_Giles8774@famism.biz	Female	9625478541	B+	Hindi	PUBG	Ranchi	Jharkhand	India
5	4	Louise	Bullock	Louise Bullock	Employed	12/31/1995	24	Louise_Bullock4253@nickia.com	Female	8932578914	B-	Bengali	Travelling	Kolkata	West Bengal	India
6	5	Aileen	Cattell	Aileen Cattell	Unemployed	9/26/1998	22	Aileen_Cattell6600@iman.org	Female	8896541237	O+	Odia	Singing	Berhampur	Odisha	India
7	6	Angela	Young	Angela Young	Student	12/22/1994	25	Angela_Young7157@twipet.com	Female	6023697412	O-	Odia	Shopping	Rayagada	Odisha	India
8	7	Chloe	Amstead	Chloe Amstead	Student	7/23/1999	21	Chloe_Amstead5097@qater.org	Female	7412369850	AB+	Bengali	Singing	Kolkata	West Bengal	India
9	8	Estrella	Fall	Estrella Fall	Student	7/24/2007	25	Estrella_Fall8124@tonys.org	Female	8745123690	AB-	Odia	Cricket	Bhubaneswar	Odisha	India
10	9	William	Jobson	William Jobson	Employed	11-01-2002	24	William_Jobson9244@vetan.org	Male	8030142057	A+	Odia	Football	Rourkela	Odisha	India
11	10	Sylvia	Nielson	Sylvia Nielson	Unemployed	10-02-2010	27	Sylvia_Nielson2893@twace.org	Female	9350061247	A-	Odia	Painting	Koraput	Odisha	India
12	11	Logan	Stanton	Logan Stanton	Student	02-10-1994	24	Logan_Stanton6191@vetan.org	Male	8271465390	O+	Rajasthan	Singing	Kota	Rajasthan	India
13	12	Logan	Fisher	Logan Fisher	Unemployed	6/23/2010	28	Logan_Fisher1707@famism.biz	Male	7896541230	O-	Odia	Dancing	Jharsuguda	Odisha	India
14	13	Violet	Hunt	Violet Hunt	Unemployed	9/18/2010	20	Violet_Hunt4089@irrepsy.com	Female	8056301478	B+	Odia	Cricket	Rayagada	Odisha	India
15	14	Chester	Lowe	Chester Lowe	Unemployed	9/23/1999	20	Chester_Lowe8789@bulaffy.com	Male	9632105874	B-	Marathi	Dancing	Mumbai	Maharashtra	India
16	15	Chloe	Poole	Chloe Poole	Unemployed	03-11-2009	22	Chloe_Poole6860@grannar.com	Female	9325050781	AB+	Odia	Football	Bhubaneswar	Odisha	India
17	16	Leroy	Bowen	Leroy Bowen	Student	01-10-2010	21	Leroy_Bowen5976@jolniaa.com	Male	8542103697	AB-	Hindi	Shopping	Surat	Gujarat	India
18	17	Chris	Lewin	Chris Lewin	Student	11/30/2008	25	Chris_Lewin8225@acrit.org	Male	9632014587	A+	Kannada	Travelling	Bangalore	Karnataka	India
19	18	Harriet	Larkin	Harriet Larkin	Employed	8/19/2003	25	Harriet_Larkin7410@twipet.com	Female	7412589630	A-	Odia	Painting	Rourkela	Odisha	India
20	19	Logan	Mcneill	Logan Mcneill	Employed	9/14/2008	25	Logan_Mcneill3031@deons.tech	Female	9777994229	B+	Hindi	Travelling	Surat	Gujarat	India
21	20	Sofie	Vane	Sofie Vane	Employed	5/18/1994	25	Sofie_Vane5779@gmail.com	Female	8249200575	B-	Odia	PUBG	Bhubaneswar	Odisha	India
22	21	Doris	Holmes	Doris Holmes	Employed	2/23/2000	29	Doris_Holmes4761@mualltech	Female	755540781	A+	Telugu	Cricket	Visakhapatnam	Andhra Pradesh	India
23	22	Bart	Furnell	Bart Furnell	Employed	12-06-2003	20	Bart_Furnell8372@gembat.biz	Male	9655510984	A-	Odia	Singing	Bhubaneswar	Odisha	India
24	23	Angela	Bennett	Angela Bennett	Employed	8/17/1995	27	Angela_Bennett8941@sveldo.biz	Female	985555692	O+	Kannada	Travelling	Kochi	Kerala	India
25	24	Mason	Denton	Mason Denton	Unemployed	02-03-2002	23	Mason_Denton2477@jiman.org	Male	9906960566	O-	Odia	Dancing	Berhampur	Odisha	India
26	25	Rachael	Lloyd	Rachael Lloyd	Student	10-10-2010	20	Rachael_Lloyd2381@iret.org	Female	5900874374	AB+	Odia	Painting	Rayagada	Odisha	India
27	26	Martin	Purvis	Martin Purvis	Student	07-09-2010	23	Martin_Purvis2449@deavo.com	Male	4536011688	AB-	Hindi	Singing	Delhi	Delhi	India
28	27	Analise	Vallins	Analise Vallins	Employed	7/23/1997	28	Analise_Vallins9156@eirey.tech	Female	1032274597	B+	Hindi	Painting	Ranchi	Jharkhand	India
29	28	Candace	Blackwall	Candace Blackwall	Unemployed	11/16/2006	20	Candace_Blackwall117@zorer.org	Female	1664438923	B-	Odia	Athletic	Bhadrapur	Odisha	India
30	29	Skylar	Vincent	Skylar Vincent	Employed	4/13/2004	29	Skylar_Vincent3535@nimogy.biz	Female	9563201052	AB+	Odia	Football	Rayagada	Odisha	India
31	30	Alan	Ogilvy	Alan Ogilvy	Employed	11-05-2008	21	Alan_Ogilvy968@jiman.org	Male	9645214785	A+	Odia	Painting	Berhampur	Odisha	India
32	31	Tyson	Gonzales	Tyson Gonzales	Employed	5/17/1994	21	Tyson_Gonzales5777@nimogy.biz	Male	6737600518	A-	Kannada	Volleyball	Bangalore	Karnataka	India
33	32	Mya	Simpson	Mya Simpson	Student	11-04-1997	23	Mya_Simpson8141@grannar.com	Female	9671602033	O+	Odia	Painting	Bhubaneswar	Odisha	India
34	33	Rocco	Richards	Rocco Richards	Unemployed	3/22/2010	23	Rocco_Richards2241@atink.com	Male	7770639356	O-	Odia	Singing	Nalanda	Bihar	India
35	34	Crystal	Gates	Crystal Gates	Employed	06-02-1994	20	Crystal_Gates7586@supunk.biz	Female	9446055414	B+	Odia	PUBG	Bhubaneswar	Odisha	India
36	35	Mike	Hood	Mike Hood	Student	08-08-2008	22	Mike_Hood8233@dionrab.com	Male	6579979894	B-	Hindi	Athletic	Junagarh	Chhattisgarh	India
37	36	Harvey	Truscott	Harvey Truscott	Employed	1/21/2006	24	Harvey_Truscott5251@mattyth.com	Male	9654785210	A+	Odia	Travelling	Bhubaneswar	Odisha	India
38	37	Tom	Powell	Tom Powell	Student	8/30/2005	23	Tom_Powell5899@iret.org	Male	7381285565	A-	Telugu	Volleyball	Mysuru	Karnataka	India

• Internal Source Dataset –

	A	B	C	D	E	G	H	I	K	L	M	N	P	Q	R
1	ID	FIRST NAME	LAST NAME	FULL NAME	GENDER	EMAIL	CONTACT NO	COURSE	CITY	State	STATE	Country	EDUCATION	SPECIALIZATION	MONTH
2	1	Ara	Jim	Ara Jim	Female	Jim@gmail.com	9876585214	Programming	Salem	Tamil-Nadu	TN	India	B.Tech	EEE	May
3	2	Deepika	Mondal	Deepika Mondal	Female	Deepi@rediffmail.com	987654	Business Analytics	Jaipur	Rajasthan	RAJ	India	BA	ECE	January
4	3	Bernard	Rice	Bernard Rice	Male	riceben@gmail.com	987654	Data analytics	CHENNAI	Karnataka	KA	India	CA	ECE	February
5	4	abhisek	malan	abhisek malan	Male	abhisek45@hotmail.com	4554212	Programming	PUNE	Maharashtra	MH	India	BSC	chemistry	September
6	5	Kevin	Nash	Kevin Nash	Male	Kevin@yahoo.com	6778899	Data analytics	Jamshedpur	Jharkhand	JH	India	Diploma	Mech	october
7	6	Jose	Robbins	Jose Robbins	Male	Jose@yahoo.com	9012345	Database	Amritsar	Punjab	PU	India	CA	Economics	August
8	7	Kevin	Nashi	Kevin Nashi	Male	Kevin@yahoo.com	2346789	Data analytics	Amritsar	Punjab	PU	India	Diploma	Mech	october
9	8	Sony	Peterson	Sony Peterson	Female	Sony@rediffmail.com	7654321	Data Analytics	Gaya	Bihar	BI	India	BSC	Civil	November
10	9	Rose	Parks	Rose Parks	Female	Rose@yahoo.com	2345678	Programming	marathali	Karnataka	KA	India	MSC	Finance	December
11	10	minakshi	dey	minakshi dey	Female	mdey9@gmail.com	63669696	Data Analytics	BANGLORE	Karnataka	KA	India	MSC	physics	June
12	11	Shilpa	Das	Shilpa Das	Female	Shilpa@rediffmail.com	98765432	software	Mumbai	Maharashtra	MH	India	M.Tech	Civil	November
13	12	Ajit	Mohanan	Ajit Mohanan	Male	Ajit@rediffmail.com	56784321	Programming	Amritsar	Punjab	PU	India	Diploma	Sociology	April
14	13	Heemani	Pat	Heemani Pat	Female	Pat@rediffmail.com	35678998	software	JP Nagar	Andhra Prades	AP	India	MBA	HR	September
15	14	Abhijit	Kumar	Abhijit Kumar	Male	Abhi@gmail.com	45678907	Business Analytics	Ansalsol	West Bengal	WB	India	CA	Civil	November
16	15	Berry	Wilkins	Berry Wilkins	Female	BW@rediffmail.com	45678901	Programming	Patna	Bihar	BI	India	MSC	Marketing	March
17	16	Raj	Moran	Raj Moran	Male	moronraj@yahoo.com	98765432	Database	patan	Bihar	BI	India	MBA	Finance	December
18	17	Aditya	Parks	Aditya Parks	Male	adityaparks@yahoo.com	56784321	Programming	Chennai	Karnataka	KA	India	BSC	Maths	May
19	18	Shilpa	Pandey	Shilpa Pandey	Female	shilpap@rediffmail.com	35678998	Database	Chennai	Tamil-Nadu	TN	India	M.Tech	ECE	February
20	19	Deepika	Pattanik	Deepika Pattanik	Female	deepikapatt@redifmail.co	34567789	Data Analytics	Kormangala	Andhra Prades	AP	India	BA	Sociology	April
21	20	Abhishek	Mahapatra	Abhishek Mahapatra	Male	abhishek45M@hotmail.co	45678901	Data Analytics	Bhopal	Madhya Prade	MP	India	BSC	Economics	August
22	21	swati	ghose	swati ghose	Female	rghose9@gmail.com	266588999	Programming	CUTTACK	Orissa	OD	India	BE	math	September
23	22	ankit	bhat	ankit bhat	Male	mdey9@gmail.com	25532645	Programming	KANPUR	Uttar Pradesh	UP	India	B.TECH	etc	May
24	23	saurav	kar	saurav kar	Male	karsa58@gmail.com	245545451	Programming	HYDERABAD	Tamil-Nadu	TN	India	B.TECH	etc	May
25	24	Danish	D'Souza	Danish D'Souza	Male	Danish@gmail.com	453601168	Business Analytics	Kormangala	Andhra Prades	AP	India	MBA	Marketing	March
26	25	Ganesh	Rahu	Ganesh Rahu	Male	Ganesh@gmail.com	637360051	Data Analytics	Not Available	Himachal Prad	HP	India	BSC	Economics	August
27	26	Alisha	Pattanik	Alisha Pattanik	Female	alisha@gmail.com	898756745	Database	Kochi	Kerala	KE	India	M.Tech	ECE	February
28	27	Raj	Paul	Raj Paul	Male	Raj@yahoo.com	387685360	Data analytics	Bhubaneswar	Orissa	OD	India	BA	Sociology	April
29	28	Neha	Patro	Neha Patro	Female	Patro@gmail.com	227887357	Data analytics	Amritsar	Punjab	PU	India	CA	ECE	February
30	29	Trupti	Sharma	Trupti Sharma	Male	Ts@rediffmail.com	678907479	Data Analytics	KOCHI	Kerala	KE	India	M.Tech	CS	August
31	30	Domingo	Ross	Domingo Ross	Female	Ross9@gmail.com	980945562	Data analytics	Lucknow	Uttar Pradesh	UP	India	MS	Finance	December
32	31	Marvin	Ballard	Marvin Ballard	Male	Ballard@gmail.com	345678909	Programming	Shimla	Himachal Prad	HP	India	BSC	ECE	February
33	32	David	William	David William	Male	David@gmail.com	234567789	Data Analytics	Kormangala	Assam	AS	India	B.Tech	CS	April
34	33	Muhammad	Khan	Muhammad Khan	Male	Khan@rediffmail.com	456789022	Programming	Pune	Maharashtra	MH	India	M.Tech	Maths	May
35	34	Ivan	P	Ivan P	Male	Ivan@rediffmail.com	789012345	Data Analytics	Rajkot	Gujar	GUJ	India	MBA	Bio Tech	June
36	35	Hunter	Moran	Hunter Moran	Male	Hunter@yahoo.com	889901245	software	Kota	Rajasthan	RAJ	India	BA	Maths	July
37	36	Raj	Nashi	Raj Nashi	Male	rainashi@yahoo.com	227887357	Data Analytics	PUNE	Maharashtra	MH	India	MS	Marketing	March
38	37	Divya	Sharma	Divya Sharma	Male	divyasma@gmail.com	678907479	software	BANARAS	Uttar Pradesh	UP	India	MBA	HR	September
39	38	Jitendra	Panda	Jitendra Panda	Male	jltpanda@rediffmail.com	909087655	Business Analytics	Kochi	Kerala	KE	India	MBA	Marketing	March

• LinkedIn Dataset –

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
1	ID	FNAME	LNAME	FULLNAME	AGE	Gender	QUALIFICATION	DEPARTMENT	DESIGNATION	EXPERIENCE	COURSE_INTERESTED	EMAIL	CONTACT	CITY	STA
2	1	Rohit	Patnaik	Rohit Patnaik	24	M	B.tech	EE	Fresher	0	Business Analytics	patnaikrohit21@gmail.com	+91 788 774 8554	Vzag	Tamil-Nadu
3	2	Sumit	Ghosh	Sumit Ghosh	21	M	B.tech	EEE	Professional	2	Programming	sumit@yahoo.com	+91 788 774 8554	Vzag	Kerala
4	3	Amit	Rampal	Amit Rampal	23	M	B.tech	EE	Professional	2	Programming	amit0102@gmail.com	+91 61279 39726	Guntur	AP
5	4	Ram	Gupta	Ram Gupta	20	M	B.TECH	ECE	PROFESSIONAL	5	Business Analytics	ram0201@gmail.com	+91 96864 45492	Guntur	Orissa
6	5	Shyam	Lal	Shyam Lal	25	M	B.TECH	CSE	PROFESSIONAL	3	software	shayam887@gmail.com	+91 80790 58492	Eluru	WB
7	6	Raju	Srivastav	Raju Srivastav	21	M	B.TECH	CIVIL	FRESHER	0	Business Analytics	raju9rsrvt@yahoo.com	+91 65797 03594	Kadapa	Uttar Pradesh
8	7	Bablu	Roy	Bablu Roy	21	M	B.tech	electrical	Professional	6	Programming	bablu8@hotmail.com	+91 86117 47725	Bhimavram	Madhya Pradesh
9	8	Denver	Silva	Denver Silva	23	M	BCA	CSE	FRESHER	0	Data Analytics	denversilva221@yahoo.com	+1 6127975653	Tinsukia	Michigan
10	9	Ola	Gunju	Ola Gunju	26	M	B.SC	AGRI	PROFESSIONAL	2	Data Analytics	ola@gmail.com	+27 6127920000	Kazirana	Cape Town
11	10	Pooja	Sharma	Pooja Sharma	25	F	BCA	CSE	PROFESSIONAL	5	Data Analytics	pooja23@gmail.com	+91 80790 58492	Kamakhya	Assam
12	11	Rakesh	Singh	Rakesh Singh	28	M	MCA	CSE	PROFESSIONAL	1	Programming	rakesh98@yahoo.com	+91 61279 75346	Patna	Maharashtra
13	12	Jitender	Rai	Jitender Rai	22	M	M.TECH	ECE	FRESHER	0	Programming	jitender45@gmail.com	+91 61279 38842	Hajipur	Gujarat
14	13	Rohit	Pawnar	Rohit Pawnar	20	M	B.TECH	CSE	FRESHER	0	Data Analytics	rohitpr@yahoo.com	+91 61279 93913	Pali	Rajasthan
15	14	Pranav	Sahu	Pranav Sahu	28	M	B.TECH	ECE	FRESHER	0	Data Analytics	pranav8@hotmail.com	+91 81603 91766	Gaya	Punjab
16	15	Swagat	Bhati	Swagat Bhati	21	M	BCA	CSE	PROFESSIONAL	2	Programming	swaat44@gmail.com	+91 61279 81077	Lagos	Haryana
17	16	Saroj	Patra	Saroj Patra	22	M	B.TECH	CSE	0	0	Database	saroj55@gmail.com	+91 68094 07184	Ahmedabad	Jharkhand
18	17	Raghuvendra	Anshu	Raghuvendra Anshu	23	M	MCA	CSE	FRESHER	0	Data Analytics	raghuan32@yahoo.com	+91 79798 29606	Bhuj	Bihar
19	18	Abhishek	Awasthi	Abhishek Awasthi	21	M	B.TECH	ECE	PROFESSIONAL	2	Data Analytics	awasthi32@yahoo.com	+91 61279 21916	Vadodra	KA
20	19	Madhur	Sharma	Madhur Sharma	25	M	Biology	zoology	Professional	4	database	sharmamadhur21@yahoo.com	+91 81677 48312	Ahmedabad	TN
21	20	Shalini	Bansal	Shalini Bansal	24	F	MBA	digital marketing	FRESHER	0	Data analytics	shalini@gmail.com	+91 80794 35958	Ahmedabad	Kerala
22	21	Ritu	Kumari	Ritu Kumari	21	F	MBA	digital marketing	FRESHER	0	Data analytics	ritu42@yahoo.com	+91 80790 30885	Ahmedabad	AP
23	22	Subhangi	Joshi	Subhangi Joshi	23	F	MBA	digital marketing	FRESHER	0	Business Analytics	subhangis22@gmail.com	+91 66192 21999	Ahmedabad	Orissa
24	23	Preeti	Dhariwal	Preeti Dhariwal	21	F	MBA	marketing	FRESHER	0	Data analytics	preetidhariwal@yahoo.com	+91 81672 06300	Ahmedabad	WB
25	24	Ahmed	Sheikh	Ahmed Sheikh	21	M	MBA	HR	FRESHER	0	Data analytics	ahmeds@hotmail.com	+91 61283 22619	Ahmedabad	UP
26	25	Subhash	Kumar	Subhash Kumar	26	M	MBA	Accounting	PROFESSIONAL	4	Data Science	tamannmg21@yahoo.com	+91 788 775 4706	Gandhinagar	MP
27	26	Sarabjit	Singh	Sarabjit Singh	25	M	MBA	HR	PROFESSIONAL	4	Business Analytics	singhsarabjit@yahoo.com	+91 79799 52096	Gandhinagar	HP
28	27	Jaspreet	Kaur	Jaspreet Kaur	21	F	MBA	HR	PROFESSIONAL	4	Data Science	jaspreetkaur78@yahoo.com	+91 6413 99810	Gandhinagar	Bihar
29	28	Rahul	Sharma	Rahul Sharma	25	M	MBA	HR	PROFESSIONAL	4	Data Science	rahulsr@gmail.com	+91 61279 86958	Gandhinagar	Assam
30	29	Subham	Saran	Subham Saran	23	M	M.SC	Metallurgy	FRESHER	0	Data analytics	subhamkr@yahoo.com	+91 79313 67954	Jamnagar	Maharashtra
31	30	Amar	Awasthi	Amar Awasthi	21	M	B.SC	Metallurgy	0	0	Data Analytics	amaraws@yahoo.com	+91 61279 43954	Palwal	Gujarat
32	31	Amol	Singh	Amol Singh	25	M	BCA	CSE	FRESHER	0	Business Analytics	singhanmol99@gmail.com	+91 788 767 5689	Rohitak	Rajasthan
33	32	Kanishka	Anand	Kanishka Anand	35	M	B.TECH	MECH	PROFESSIONAL	5	Data Analytics	kanishkaan44@gmail.com	+91 63006 24603	Faridabad	Punjab
34	33	Ranjan	Prasad	Ranjan Prasad	36	M	B.TECH	ECE	FRESHER	0	Programming	ranjanpr47@yahoo.com	+91 61279 17078	Manali	Haryana
35	34	Ravi	Kumar	Ravi Kumar	34	M	MCA	ME	FRESHER	0	Data analytics	ravirk32332@gmail.com	+91 74770 95320	Kasaul	Jharkhand
36	35	Roy	Banerjee	Roy Banerjee	30	M	M.SC	zoology	FRESHER	0	Programming	banerieer23@yahoo.com	+91 61279 99383	Hazirbagh	Bihar
37	36	Namrita	Rao	Namrita Rao	25	F	M.Sc	mathematics	FRESHER	0	Data analytics	namritarao323@yahoo.com	+91 73100 42786	Latdak	KA
38	37	Shallesh	Kumar	Shallesh Kumar	31	M	B.TECH	Metallurgy	FRESHER	0	Programming	shalleshr3@yahoo.com	+91 61279 88521	Ranchi	TN

• Naukri Dataset –

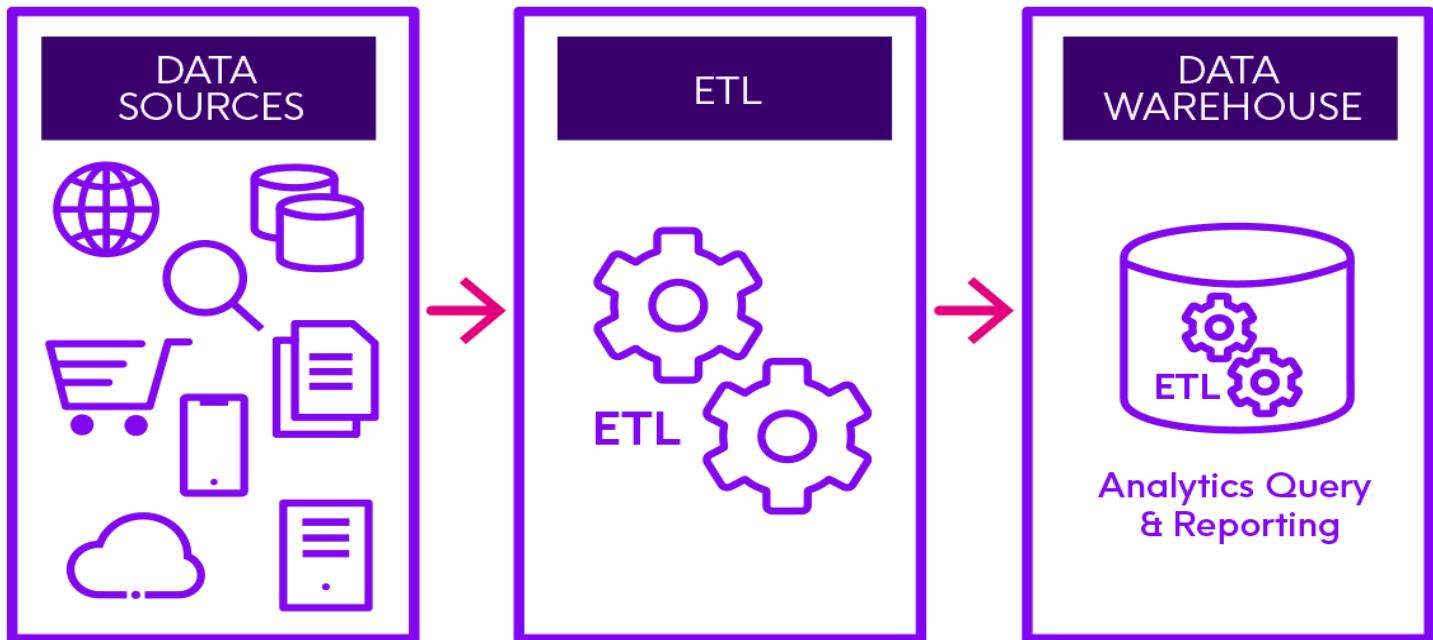
	A	B	D	E	F	G	H	I	J	K	L	M	N	O	P
1	ID	First Name	Last Name	Full Name	Age	DOB	Gender	Mobile Number	Email ID	Qualification	Specialization	Passout Year	Enrollment Courses	Work experien	Skills Known
2	1	Ram	Kumar	Ram Devi Kumar	25	10-02-1995	M	7974923455	ram@gmail.com	B.Sc	Zoology	2018	Data Science	3 Years	C++
3	2	Mohammaed	Hussian	Mohammaed Amir Hussian	21	28/02/2000	M	6269945120	mohammaed@gmail.com	B.Tech	Computer Science	2022	Business Analyst	NA	Python
4	3	Sunita	Singh	Sunita Singh	27	24/11/1994	F	9963215456	sunita@gmail.com	B.Ed	Computer Application	2016	Database	4 Years	C
5	4	Sri	Das	Sri Das	30	04-12-1992	M	8283848586	sri@gmail.com	M.Sc	Physics	2013	Data Analytic	6 Years	Java
6	5	Anita	Kaur	Anita Kaur	23	15/08/1998	F	9596949391	anita@gmail.com	BBA	Finance	2021	Machine Learning	NA	MS SQL Server
7	6	Santosh	Deshmukh	Santosh Deshmukh	31	12-10-1990	M	7578747971	santosh@gmail.com	M.Tech	Data Science	2012	Business Analyst	8 Years	Django
8	7	Sanjay	Jain	Sanjay Jain	22	29/01/1999	M	6268656965	sanjay@gmail.com	BCA	Data Science	2020	Machine Learning	NA	Django
9	8	Sunil	Lal	Sunil Lal	28	17/10/1993	M	7878981542	sunil@gmail.com	MBA	Marketing	2015	Data Science	2 Years	Machine Learnin
10	9	Rajesh	Yadav	Rajesh Yadav	27	05-05-1994	M	7985462132	rajesh@gmail.com	B.Com	Business Analytics	2016	Data Science	NA	PHP
11	10	Gita	Singh	Gita Ram Singh	29	11-03-1992	F	8589874163	gita@gmail.com	B.Ed	Computer Application	2014	Database	3 Years	CSS
12	11	Ramesh	Reddy	Ramesh Reddy	22	20/07/1999	M	7389824565	ramesh@gmail.com	B.E.	Mechanical	2020	Database	NA	Python
13	12	Ashok		Ashok	24	07-06-1997	M	7898456513	ashok@gmail.com	M.Sc	Business Studies	2019	Machine Learning	NA	Java
14	13	Rekha	Kumari	Rekha Kumari	21	13/02/2000	F	7189886396	rekha@gmail.com	B.Tech	Civil	2021	Business Analyst	NA	C++
15	14	Manoj	Jain	Manoj K Jain	25	26/06/1996	M	9694921385	manoj@gmail.com	BBA	Human Resources	2018	Database	10 Months	CSS
16	15	Anil	Pathak	Anil Pathak	28	14/03/1993	M	9173826455	anil@gmail.com	B.E.	Petro-Chemical	2015	Database	3 Years	Django
17	16	Suresh	Nair	Suresh Raja Nair	30	27/07/1991	M	8246937185	suresh@gmail.com	B.Ed	Banking Management	2013	Data Science	NA	Tableau
18	17	Vijay	Prasad	Vijay Devi Prasad	21	21/09/2000	M	7463128533	vijay@gmail.com	B.Ed	Civil	2021	Data Analytic	NA	Advanced Excel
19	18	Lakshmi	Sharma	Lakshmi Ben Sharma	23	31/12/1998	F	98798759663	lakshmi@gmail.com	M.Sc	Statistics	2020	Machine Learning	NA	Machine Learnin
20	19	Abdul	Ali	Abdul Shah Ali	26	10-01-1995	M	9996363115	abdul@gmail.com	B.Com	Finance	2018	Data Analytic	3 Months	My SQL
21	20	Raju	Sahu	Raju Sahu	22	09-09-1999	M	9595654545	raju@gmail.com	BCA	Finance	2020	Data Science	NA	MS SQL Server
22	21	Manju	Gupta	Manju Gupta	29	05-08-1992	F	8589581122	manju@gmail.com	B.Com	Business Analytics	2013	Data Analytic	NA	Java
23	22	Shanti	Paramar	Shanti Paramar	24	12-12-1998	F	9413526877	shanti@gmail.com	M.Sc	Business Studies	2020	Data Analytic	NA	Java
24	23	Vinod	Jadhav	Vinod Jadhav	20	31/07/2001	M	9632145879	vinod@gmail.com	BBA	Marketing	2022	Machine Learning	NA	Oracle
25	24	Suman	Rani	Suman Rani	27	14/12/1994	F	9797985956	suman@gmail.com	B.Tech	Electrical	2016	Data Analytic	2 Years	CSS
26	25	Rajendra	Mishra	Rajendra Mishra	23	19/01/1998	M	8660048981	rajendra@gmail.com	BBA	Hospitality	2021	Data Analytic	NA	HTML
27	26	Raj	Chand	Raj Roy Chand	29	30/10/1992	M	7086352419	raj@gmail.com	M.Tech	Data Science	2015	Data Analytic	4 Years	JQuery
28	27	Usha	Saha	Usha Saha	28	31/10/1998	F	9091325687	usha@gmail.com	B.Com	Information Systems	2015	Machine Learning	NA	PHP
29	28	Asha	Thakur	Asha Thakur	21	12-10-2000	F	9865320147	asha@gmail.com	B.E.	Information Technology	2022	Data Analytic	NA	PHP
30	29	Dinesh		Dinesh	20	22/02/2001	M	9780002564	dinesh@gmail.com	BBA	Human Resources	2021	Data Analytic	NA	Python
31	30	Krishna	Pradhan	Krishna Pradhan	26	16/08/1995	M	9691784532	krishna@gmail.com	BBA	Human Resources	2017	Data Analytic	NA	Machine Learnin
32	31	Mina	Malik	Mina Malik	24	03-07-1997	F	9504682007	mina@gmail.com	B.Sc	Forestry	2019	Business Analyst	5 Months	HTML
33	32	Kiran	Patra	Kiran Patra	21	25/01/2000	F	6265646317	kiran@gmail.com	BCA	Finance	2021	Database	NA	CSS
34	33	Dilip	Jha	Dilip Jha	25	12-06-1996	M	6360698745	dilip@gmail.com	B.Com	Zoology	2018	Database	10 Months	Oracle
35	34	Pushpa	Pandey	Pushpa Pandey	30	23/07/1991	F	7877745651	pushpa@gmail.com	M.Tech	Security & Cyber	2015	Machine Learning	3 Years	Oracle
36	35	Ajay	Singh	Ajay Raj Singh	22	12-05-1999	M	9001232781	ajay@gmail.com	M.Sc	Biotechnology	2020	Machine Learning	NA	Django
37	36	Urmila	Pandit	Urmila Pandit	23	18/05/1998	F	6561630201	urmila@gmail.com	BCA	Biotechnology	2021	Business Analyst	NA	JQuery
38	37	Mukesh	Kamble	Mukesh Kamble	20	04-01-2001	M	8786005251	mukesh@gmail.com	B.Tech	Information Technology	2022	Data Analytic	NA	C++
39	38	Amit	Rana	Amit Rana	29	24/02/1992	M	7890123654	amit@gmail.com	MBA	General Management	2017	Data Analytic	2 Years	Machine Learnin
40	39	Shankar	Malla	Shankar Malla	27	22/01/2004	xx	9007112220	shankar.malla@gmail.com	B.Com	Accountancy	2016	Business Analyst	2 Years	Python

PHASE 3 : ETL

These days, companies have access to more data sources and formats than ever before: databases, websites, SaaS (software as a service) applications, and analytics tools, to name a few. Unfortunately, the ways businesses often store this data make it challenging to extract the valuable insights hidden within — especially when you need it for smarter data-driven business decision-making.

Companies need a way to analyze their data for critical business decisions. **ETL** provide a method of moving the data from various source into a data warehouse. A Company data may be scattered in different locations and in different formats. **ETL** allows you to: Migrate the data into a data warehouse. Convert the various formats and types to adhere to one consistent system. **ETL** is a predefined process for access and manipulate source data and loading it into a target database.

What is ETL?



ETL stands for **E**xtract, **T**ransform and **L**oad, which are the three steps of the ETL process. ETL collects and processes data from various sources into a single data store (e.g. a data warehouse or data lake), making it much easier to analyze.

Extract

Extracting data is the act of pulling data from one or more data sources. During the extraction phase of ETL, you may handle a variety of sources with data, such as:

- Relational and non-relational databases
- Flat files (e.g. XML, JSON, CSV, Microsoft Excel spreadsheets, etc.)
- SaaS applications, such as CRM (customer relationship management) and ERP (enterprise resource planning) systems
- APIs (application programming interfaces)
- Websites
- Analytics and monitoring tools
- System logs and metadata

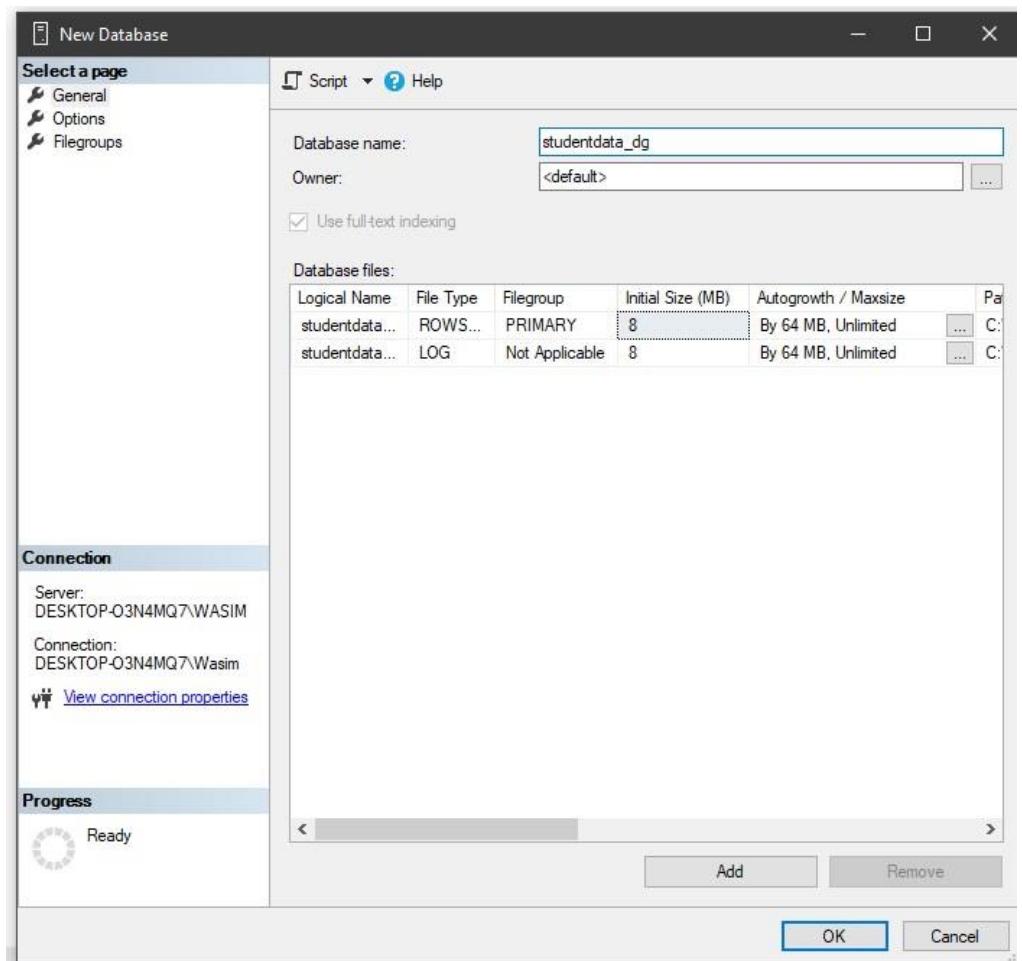
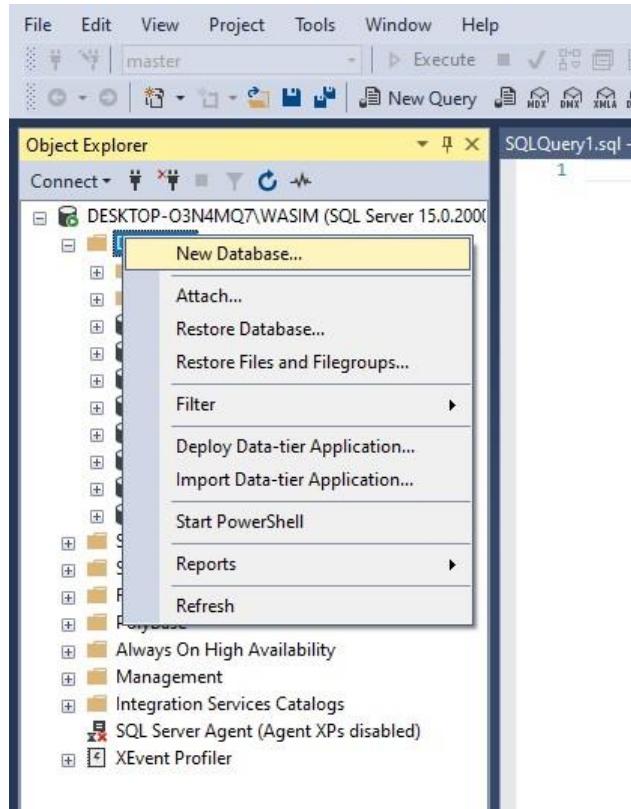
We've a dataset in Excel format, which we need to import first.

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q
1	ID	FULL NAME	GENDE	AGE	QUALIFICATI	COURSE	SKILLS	COURSE DURATI	PASS OUT YE	YEAR OF EXPERIEN	PROFESSION	MONTH JOINED	MAIL ID	CONTACT NO	CITY	STATE	COUN
2	1	Ram Devi Kumar	Male	25	B.Sc	Data Science	AI	4	2018	1	PROFESSIONAL	April	ram@gmail.com	7974923455	Bengaluru	Karnataka	India
3	2	Mohammaed Amir Hus	Male	24	B.Tech	Business Analys	Power Bi	5	2022	2	PROFESSIONAL	December	nmaed@gmail.com	626945120	Puri	Odisha	India
4	3	Sunita Singh	Female	27	B.Ed	Database	Oracle	4	2016	1	PROFESSIONAL	July	sunita@gmail.com	9963215456	Kannur	Kerala	India
5	4	Sri Das	Male	30	M.Sc	Data Analyst	Tableau	3	2013	0	FRESHER	May	sri@gmail.com	8283848586	Indore	Idhya Prade	India
6	5	Anita Kaur	Female	23	BBA	Machine Learnin	AI	3	2021	0	FRESHER	November	anita@gmail.com	959649399	Bhubaneswar	Odisha	India
7	6	Santosh Deshmukh	Male	31	M.Tech	Business Analys	Power Bi	3	2012	0	FRESHER	December	antosh@gmail.com	7578747971	Durg	Jhattisgarl	India
8	7	Sanjay Jain	Male	22	BCA	Machine Learnin	AI	3	2020	0	FRESHER	July	sanjay@gmail.com	6268656965	Udupi	Karnataka	India
9	8	Sunil Lal	Male	28	MBA	Data Science	AI	3	2015	2	PROFESSIONAL	April	sunil@gmail.com	7878981542	Nagpur	Aharashtr	India
10	9	Rajesh Yadav	Male	27	B.Com	Data Science	AI	3	2016	0	FRESHER	February	ralesh@gmail.com	7985462132	Palakkad	Kerala	India
11	10	Gita Ram Singh	Female	29	B.Ed	Database	Oracle	3	2014	1	PROFESSIONAL	March	gita@gmail.com	8589874163	Chennai	Tamil-Nadi	India
12	11	Ramesh Reddy	Male	22	B.E	Database	Oracle	3	2020	5	PROFESSIONAL	November	ramesh@gmail.com	7389824565	Jangaon	Telangana	India
13	12	Ashok	Male	24	M.Sc	Machine Learnin	AI	3	2019	1	PROFESSIONAL	May	ashok@gmail.com	7898456513	Panaji	Goa	India
14	13	Rekha Kumari	Female	21	B.Tech	Business Analys	Power Bi	3	2021	0	FRESHER	July	rekha@gmail.com	7189886396	Bhopal	Idhya Prade	India
15	14	Manoj K Jain	Male	25	BBA	Database	Oracle	5	2018	0	FRESHER	June	manoj@gmail.com	9694921385	Buxar	Bihar	India
16	15	Anil Pathak	Male	28	B.E	Database	Oracle	5	2015	0	FRESHER	August	anil@gmail.com	917326455	Agra	Ittar Prades	India
17	16	Suresh Raja Nair	Male	30	B.Ed	Data Science	AI	4	2013	2	PROFESSIONAL	August	suresh@gmail.com	8246937185	Agra	Ittar Prades	India
18	17	Vijay Devi Prasad	Male	21	B.Ed	Data Analyst	Tableau	5	2021	0	FRESHER	January	vijay@gmail.com	7463128533	Kolkata	West Benga	India
19	18	Lakshmi Ben Sharma	Female	23	M.Sc	Machine Learnin	AI	5	2020	1	PROFESSIONAL	February	lakshmi@gmail.com	9879759663	Raipur	Jhattisgarl	India
20	19	Abdul Shah Ali	Male	26	B.Com	Data Analyst	Tableau	3	2018	3	PROFESSIONAL	December	abdul@gmail.com	996365115	Kozhikode	Kerala	India
21	20	Raju Sahu	Male	22	BCA	Data Science	AI	3	2020	0	FRESHER	June	raju@gmail.com	9595654545	Bhubaneswar	Odisha	India
22	21	Manju Gupta	Male	29	B.Com	Data Analyst	Tableau	4	2013	4	PROFESSIONAL	August	manju@gmail.com	858985112	Lucknow	Ittar Prades	India
23	22	Shanti Paramar	Female	24	M.Sc	Data Analyst	Tableau	5	2020	1	PROFESSIONAL	November	shanti@gmail.com	9413526877	Mumbai	Aharashtr	India
24	23	Vinod Jadhav	Female	20	BBA	Machine Learnin	AI	5	2022	3	PROFESSIONAL	January	vinod@gmail.com	9632145879	Kozhikode	Kerala	India
25	24	Suman Rani	Female	23	B.Tech	Data Analyst	Excel	5	2016	2	PROFESSIONAL	December	suman@gmail.com	9797985956	Hyderabad	Telangana	India
26	25	Rajendra Mishra	Male	23	BBA	Data Analyst	Tableau	3	2021	0	FRESHER	February	jendra@gmail.com	866049891	Visakhapatnamidhr	Prade	India
27	26	Raj Roy Chand	Male	29	M.Tech	Data Analyst	Tableau	3	2015	1	PROFESSIONAL	December	raj@gmail.com	7086352419	Chennai	Tamil-Nadi	India
28	27	Usha Saha	Female	28	B.Com	Machine Learnin	AI	5	2015	3	PROFESSIONAL	May	usha@gmail.com	9091325687	Patna	Bihar	India
29	28	Asha Thakur	Female	21	B.E	Data Analyst	Tableau	4	2022	1	PROFESSIONAL	August	asha@gmail.com	9865320147	Kolkata	West Benga	India
30	29	Dinesh	Male	20	BBA	Data Analyst	Tableau	3	2021	1	PROFESSIONAL	September	dinesh@gmail.com	9780002564	Chennai	TamilNadu	India
31	30	Krishna Pradhan	Female	26	BBA	Data Analyst	Tableau	5	2017	3	PROFESSIONAL	January	rishna@gmail.com	9691784532	Panaji	Goa	India
32	31	Mina Malik	Female	24	B.Sc	Business Analys	Power Bi	5	2019	2	PROFESSIONAL	November	mina@gmail.com	9504682007	Ahmedabad	Gujarat	India
33	32	Kiran Patra	Female	21	BCA	Database	Oracle	3	2021	0	FRESHER	October	kiran@gmail.com	6265646317	Bilaspur	Jhattisgarl	India
34	33	Dilip Jha	Male	25	B.Com	Database	Oracle	5	2018	1	PROFESSIONAL	October	dilip@gmail.com	6360698745	Agra	Ittar Prades	India
35	34	Pushpa Pandey	Male	30	M.Tech	Machine Learnin	AI	4	2015	0	FRESHER	May	ushpa@gmail.com	7877475651	New Delhi	New Delhi	India

FIG. SOURCE DATA

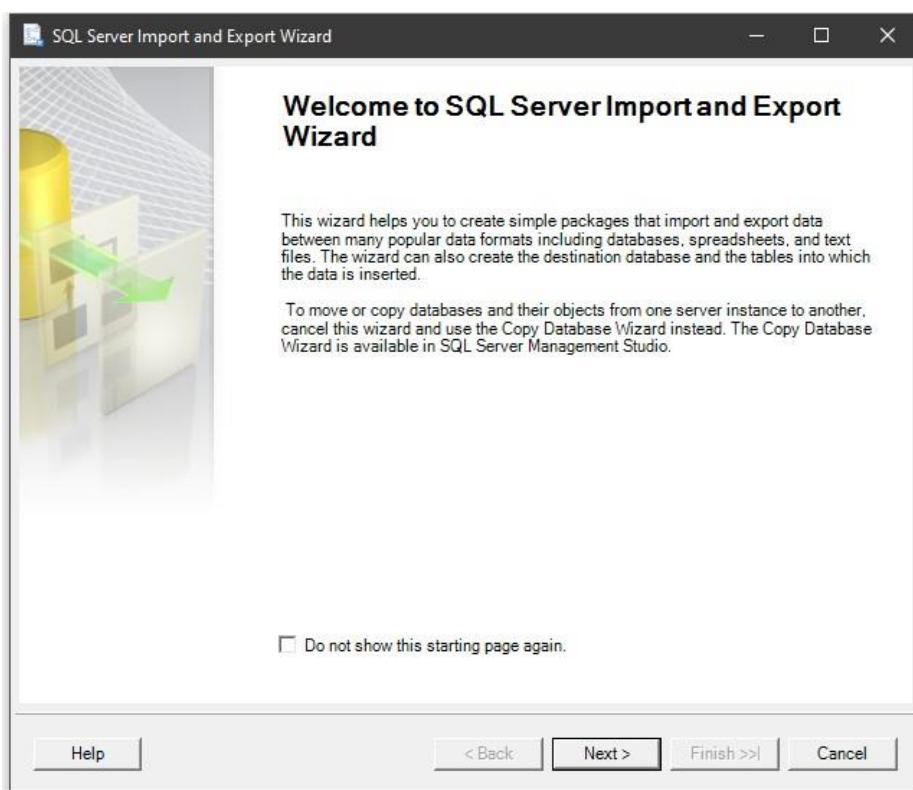
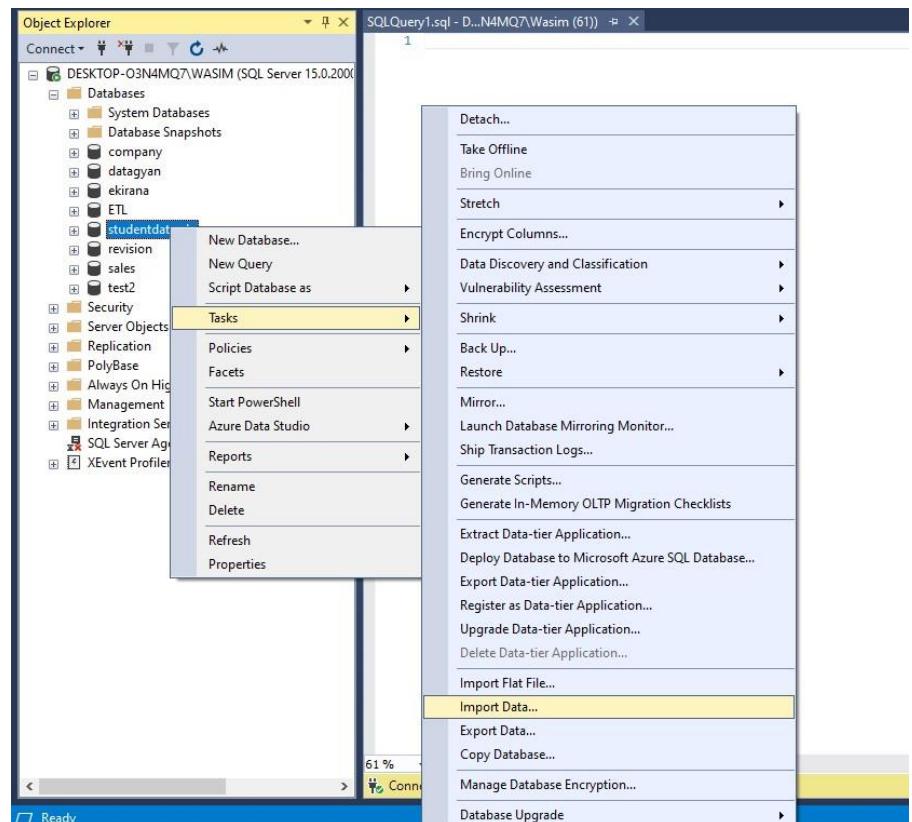
Step 1 – Create Database in SQL Server:

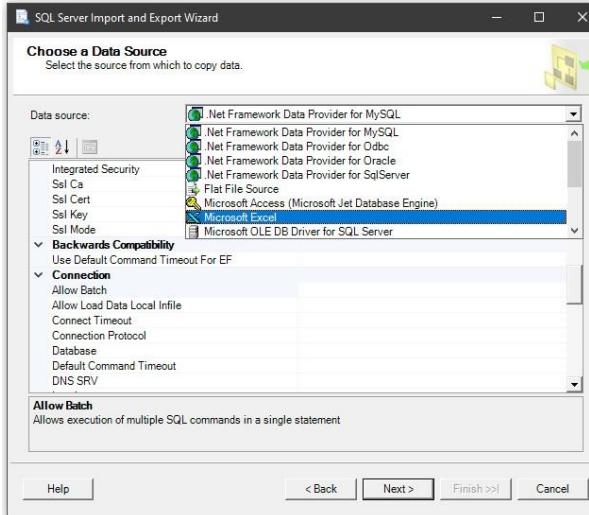
Before going next, you need to create a SQL server database connection. Right click on database and click CREATE NEW DATABASE.



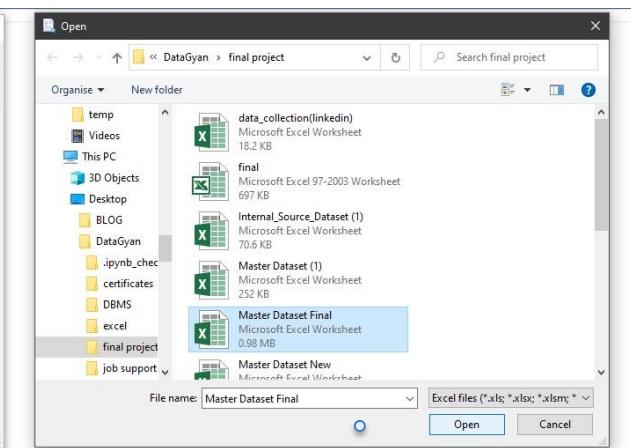
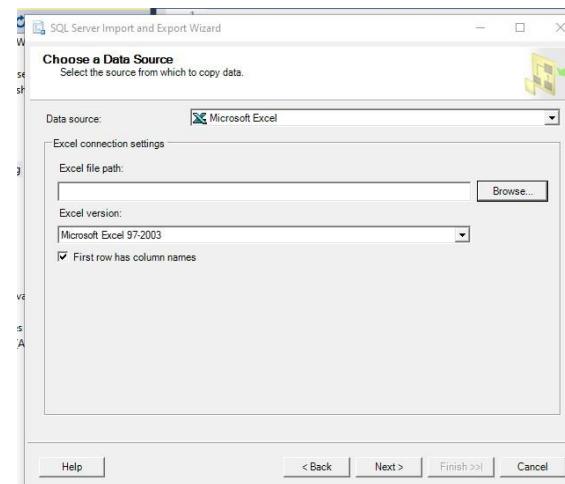
Step 2 – Import Dataset in Database:

Now, we need to load the dataset in our Database which we created now, Right Click on Database Select Task Option, Click Import Data

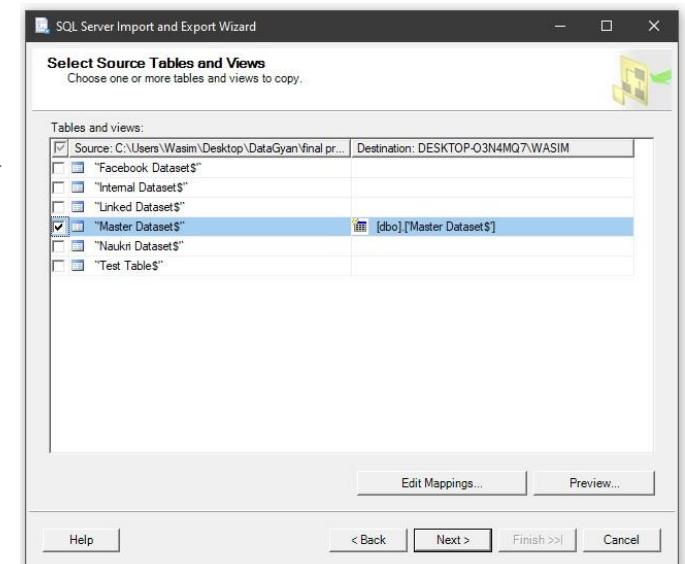
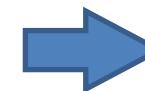
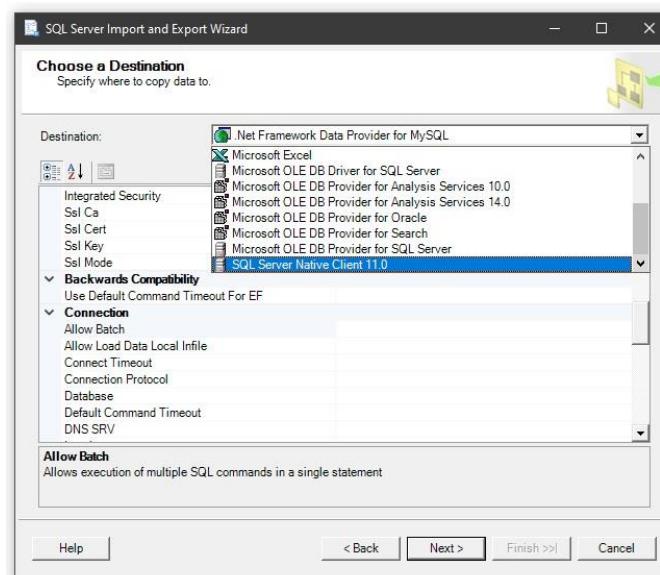




Select the source file type: Microsoft Excel



Select file

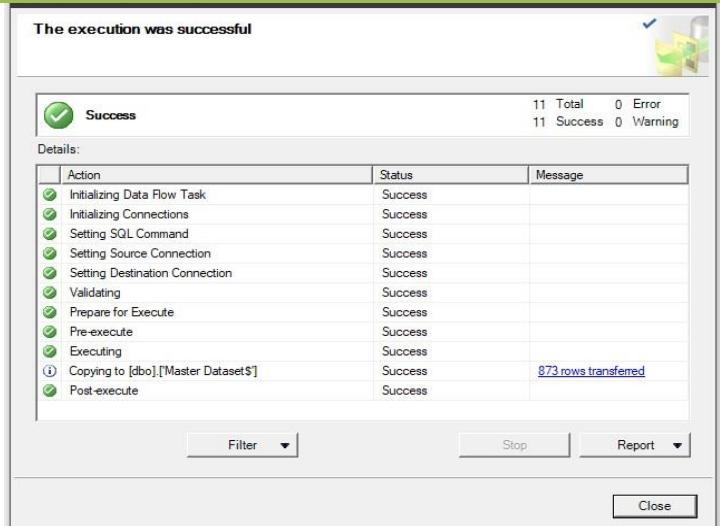
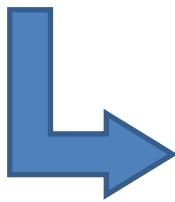


Choose destination to SQL Server Native Client 11.0 then Click Next:

Tick the checkbox to that only which file you want to proceed for ETL

It's the final step of data extraction into the server.

The data which we extracted will be stored in the database what we created and now we can go with our next phase of ETL, i.e. Transformation



The execution was successful

Success

11 Total 0 Error
11 Success 0 Warning

Action	Status	Message
Initializing Data Flow Task	Success	
Initializing Connections	Success	
Setting SQL Command	Success	
Setting Source Connection	Success	
Setting Destination Connection	Success	
Validating	Success	
Prepare for Execute	Success	
Pre-execute	Success	
Executing	Success	
Copying to [dbo].[Master Dataset\$]	Success	873 rows transferred
Post-execute	Success	

Filter Stop Report Close

Transform

It's rarely the case that your extracted data is already in the exact format that you need it to be. For example, you may want to:

Rearrange unstructured data into a structured format.

Limit the data you've extracted to just a few fields.

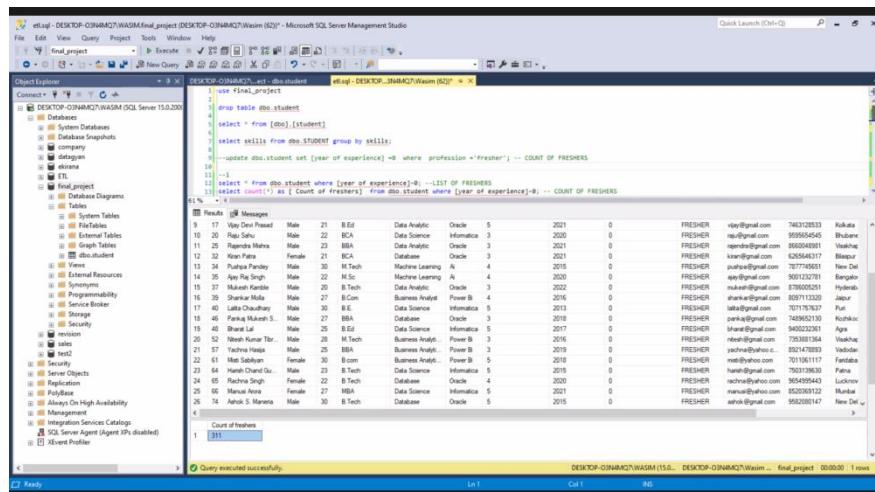
Sort the data so that all the columns are in a certain order.

Join multiple tables together.

Clean the data to eliminate duplicate and out-of-date records.

All these changes and more take place during the **transformation** phase of ETL. There are many types of data transformations that you can execute, from data cleansing and aggregation to filtering and validation.

Step 3: Implementing Queries in SQL server for Insights:



```
use final_project
go
drop table gpa_student
go
select * from [dbo].[student]
go
select skills from dbo.STUDENT group by skills
go
--update dbo.student set [year of experience]=0 where profession='Fresher' -- COUNT OF FRESHERS
go
select * from dbo.student where [year of experience]=0 --LIST OF FRESHERS
go
select count(*) from dbo.student where [year of experience]=0 -- COUNT OF FRESHERS
```

Results

Count of freshers
311

Query executed successfully.

After creating a new database, we used the query to see that the whole dataset had been imported or not. So, for this we executed the **Select * from [dbo] . [student]** query. This query will help to run the whole dataset for our reference.

Query 1

After the execution of query 1 it shows the list of freshers from the table. And then by using **COUNT () and WHERE clause** we got the details of freshers having an experience of 0 years.

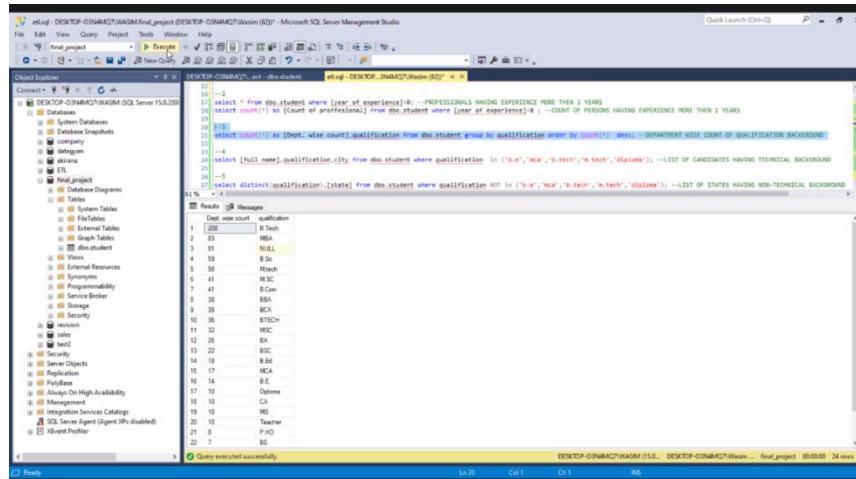
Query 2

This query is giving a brief visual of the professional candidates those who are having an experience of more than 1 year
And by using **COUNT ()** and **WHERE clause**, it shows the no.of candidates having an experience of more than 1 year

Query 3

By using COUNT (), GROUP BY () and ORDER BY () it executed the qualification of candidates in descending order.

So here it basically shows the no. of candidates from every department.

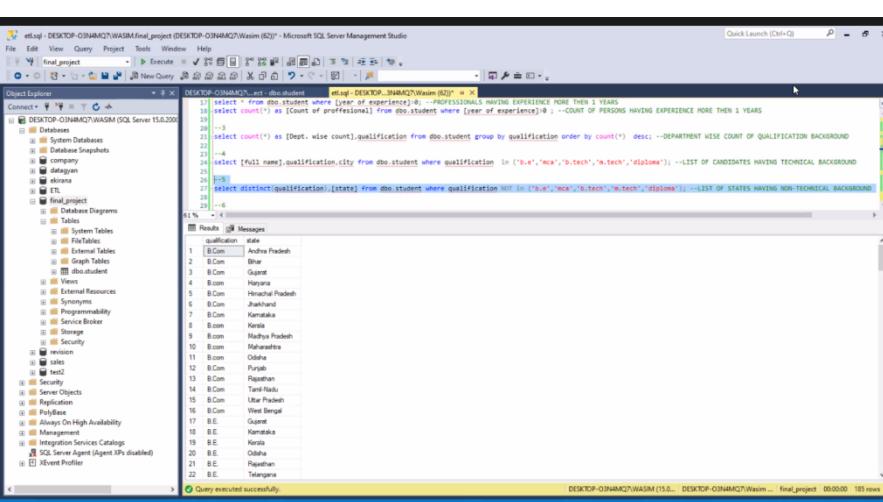


Query 4

To find the candidates with technical background by using **WHERE** clause **IN** operator

Query 5

By using **DISTINCT()** and **WHERE** clause and **NOT IN()** operator we executed state wise candidates having non-technical background.



```
22
23 --4
24 select [full name].qualification,city from dbo.student where qualification in ('b.e','mca','b.tech','m.tech','diploma'); --LIST OF C
25
26 --5
27 select distinct(qualification),[state] from dbo.student where qualification NOT IN ('b.e','mca','b.tech','m.tech','diploma'); --LIST (
28
29 --6
30 select gender,[count(*)] as [Gender wise count] from dbo.student group by gender; -- COUNT OF STUDENTS BASED ON THEIR GENDER
31
32 --7
33 select COURSE,[count(*)] as [Course wise count] from dbo.student group by COURSE; -- COUNT OF STUDENTS BASED ON THEIR COURSE
```

Results	Messages
gender	Gender wise count
NULL	81
Male	474
Female	318

Query 6

By using **COUNT(*)** and **GROUP BY** clause we found the count distribution by gender.

Query 7

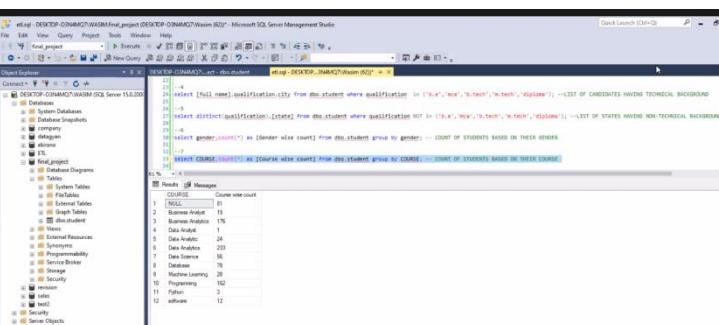
By using **COUNT(*)** and **GROUP BY** clause we found the count of candidates state wise

The screenshot shows the Microsoft SQL Server Management Studio interface. The Object Explorer on the left lists the database structure, including tables like 'student' and 'course'. The main pane displays a query result grid for a query that retrieves student details based on their qualification. The status bar at the bottom indicates the query was executed successfully.

```
etlog@DESKTOP-0JN8M4Q\WASIM:final_project (DESKTOP-0JN8M4Q\WASIM (SQL)) - Microsoft SQL Server Management Studio
File Edit View Query Project Tools Window Help
... Execute ... Find Replace ... Refresh ... Stop ...
Quick Launch [Ctrl+Q] P
Object Explorer
...
DESKTOP-0JN8M4Q\WASIM (DESKTOP-0JN8M4Q\WASIM (SQL)) - *
1 select distinct(qualification),state from dbo.student where qualification NOT IN ('b.e','mca','b.tech','m.tech','diploma'); --LIST OF STATES HAVING NON-TECHNICAL BACKGROUND
2
3 select gender,cnt(*) as [Gender wise count] from dbo.student group by gender; -- COUNT OF STUDENTS BASED ON THEIR GENDER
4
5 select course,cnt(*) as [Course wise count] from dbo.student group by course; -- COUNT OF STUDENTS BASED ON THEIR COURSE
6
7
8 SELECT * FROM (SELECT city ,count(students) AS [City wise count of students] FROM dbo.student GROUP BY city ORDER BY [Count] DESC) AS CITY_WISE_COUNT_OF_STUDENTS;
9
10
11 % Results
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
279
280
281
282
283
284
285
286
287
288
289
289
290
291
292
293
294
295
296
297
298
299
299
300
301
302
303
304
305
306
307
308
309
309
310
311
312
313
314
315
316
317
318
319
319
320
321
322
323
324
325
326
327
328
329
329
330
331
332
333
334
335
336
337
338
339
339
340
341
342
343
344
345
346
347
348
349
349
350
351
352
353
354
355
356
357
358
359
359
360
361
362
363
364
365
366
367
368
369
369
370
371
372
373
374
375
376
377
378
379
379
380
381
382
383
384
385
386
387
388
389
389
390
391
392
393
394
395
396
397
398
399
399
400
401
402
403
404
405
406
407
408
409
409
410
411
412
413
414
415
416
417
418
419
419
420
421
422
423
424
425
426
427
428
429
429
430
431
432
433
434
435
436
437
438
439
439
440
441
442
443
444
445
446
447
448
449
449
450
451
452
453
454
455
456
457
458
459
459
460
461
462
463
464
465
466
467
468
469
469
470
471
472
473
474
475
476
477
478
479
479
480
481
482
483
484
485
486
487
488
489
489
490
491
492
493
494
495
496
497
498
499
499
500
501
502
503
504
505
506
507
508
509
509
510
511
512
513
514
515
516
517
518
519
519
520
521
522
523
524
525
526
527
528
529
529
530
531
532
533
534
535
536
537
538
539
539
540
541
542
543
544
545
546
547
548
549
549
550
551
552
553
554
555
556
557
558
559
559
560
561
562
563
564
565
566
567
568
569
569
570
571
572
573
574
575
576
577
578
579
579
580
581
582
583
584
585
586
587
588
589
589
590
591
592
593
594
595
596
597
598
599
599
600
601
602
603
604
605
606
607
608
609
609
610
611
612
613
614
615
616
617
618
619
619
620
621
622
623
624
625
626
627
628
629
629
630
631
632
633
634
635
636
637
638
639
639
640
641
642
643
644
645
646
647
648
649
649
650
651
652
653
654
655
656
657
658
659
659
660
661
662
663
664
665
666
667
668
669
669
670
671
672
673
674
675
676
677
678
679
679
680
681
682
683
684
685
686
687
688
689
689
690
691
692
693
694
695
696
697
698
698
699
700
701
702
703
704
705
706
707
708
709
709
710
711
712
713
714
715
716
717
718
719
719
720
721
722
723
724
725
726
727
728
729
729
730
731
732
733
734
735
736
737
738
739
739
740
741
742
743
744
745
746
747
748
749
749
750
751
752
753
754
755
756
757
758
759
759
760
761
762
763
764
765
766
767
768
769
769
770
771
772
773
774
775
776
777
778
779
779
780
781
782
783
784
785
786
787
788
788
789
789
790
791
792
793
794
795
796
797
798
798
799
799
800
801
802
803
804
805
806
807
808
809
809
810
811
812
813
814
815
816
817
818
819
819
820
821
822
823
824
825
826
827
828
829
829
830
831
832
833
834
835
836
837
838
839
839
840
841
842
843
844
845
846
847
848
849
849
850
851
852
853
854
855
856
857
858
859
859
860
861
862
863
864
865
866
867
868
869
869
870
871
872
873
874
875
876
877
878
879
879
880
881
882
883
884
885
886
887
888
888
889
889
890
891
892
893
894
895
896
897
898
898
899
899
900
901
902
903
904
905
906
907
908
909
909
910
911
912
913
914
915
916
917
918
919
919
920
921
922
923
924
925
926
927
928
929
929
930
931
932
933
934
935
936
937
938
939
939
940
941
942
943
944
945
946
947
948
949
949
950
951
952
953
954
955
956
957
958
959
959
960
961
962
963
964
965
966
967
968
969
969
970
971
972
973
974
975
976
977
978
979
979
980
981
982
983
984
985
986
987
988
988
989
989
990
991
992
993
994
995
996
997
997
998
998
999
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1019
1020
1021
1022
1023
1024
1025
1026
1027
1028
1029
1029
1030
1031
1032
1033
1034
1035
1036
1037
1038
1039
1039
1040
1041
1042
1043
1044
1045
1046
1047
1048
1049
1049
1050
1051
1052
1053
1054
1055
1056
1057
1058
1059
1059
1060
1061
1062
1063
1064
1065
1066
1067
1068
1069
1069
1070
1071
1072
1073
1074
1075
1076
1077
1078
1079
1079
1080
1081
1082
1083
1084
1085
1086
1087
1088
1088
1089
1089
1090
1091
1092
1093
1094
1095
1096
1097
1097
1098
1098
1099
1099
1100
1101
1102
1103
1104
1105
1106
1107
1108
1109
1109
1110
1111
1112
1113
1114
1115
1116
1117
1118
1119
1119
1120
1121
1122
1123
1124
1125
1126
1127
1128
1129
1129
1130
1131
1132
1133
1134
1135
1136
1137
1138
1139
1139
1140
1141
1142
1143
1144
1145
1146
1147
1148
1149
1149
1150
1151
1152
1153
1154
1155
1156
1157
1158
1159
1159
1160
1161
1162
1163
1164
1165
1166
1167
1168
1169
1169
1170
1171
1172
1173
1174
1175
1176
1177
1178
1179
1179
1180
1181
1182
1183
1184
1185
1186
1187
1188
1188
1189
1189
1190
1191
1192
1193
1194
1195
1196
1197
1197
1198
1198
1199
1199
1200
1201
1202
1203
1204
1205
1206
1207
1208
1209
1209
1210
1211
1212
1213
1214
1215
1216
1217
1218
1219
1219
1220
1221
1222
1223
1224
1225
1226
1227
1228
1229
1229
1230
1231
1232
1233
1234
1235
1236
1237
1238
1239
1239
1240
1241
1242
1243
1244
1245
1246
1247
1248
1249
1249
1250
1251
1252
1253
1254
1255
1256
1257
1258
1259
1259
1260
1261
1262
1263
1264
1265
1266
1267
1268
1269
1269
1270
1271
1272
1273
1274
1275
1276
1277
1278
1279
1279
1280
1281
1282
1283
1284
1285
1286
1287
1288
1288
1289
1289
1290
1291
1292
1293
1294
1295
1296
1297
1297
1298
1298
1299
1299
1300
1301
1302
1303
1304
1305
1306
1307
1308
1309
1309
1310
1311
1312
1313
1314
1315
1316
1317
1318
1319
1319
1320
1321
1322
1323
1324
1325
1326
1327
1328
1329
1329
1330
1331
1332
1333
1334
1335
1336
1337
1338
1339
1339
1340
1341
1342
1343
1344
1345
1346
1347
1348
1349
1349
1350
1351
1352
1353
1354
1355
1356
1357
1358
1359
1359
1360
1361
1362
1363
1364
1365
1366
1367
1368
1369
1369
1370
1371
1372
1373
1374
1375
1376
1377
1378
1379
1379
1380
1381
1382
1383
1384
1385
1386
1387
1388
1388
1389
1389
1390
1391
1392
1393
1394
1395
1396
1397
1397
1398
1398
1399
1399
1400
1401
1402
1403
1404
1405
1406
1407
1408
1409
1409
1410
1411
1412
1413
1414
1415
1416
1417
1418
1419
1419
1420
1421
1422
1423
1424
1425
1426
1427
1428
1429
1429
1430
1431
1432
1433
1434
1435
1436
1437
1438
1439
1439
1440
1441
1442
1443
1444
1445
1446
1447
1448
1449
1449
1450
1451
1452
1453
1454
1455
1456
1457
1458
1459
1459
1460
1461
1462
1463
1464
1465
1466
1467
1468
1469
1469
1470
1471
1472
1473
1474
1475
1476
1477
1478
1479
1479
1480
1481
1482
1483
1484
1485
1486
1487
1488
1488
1489
1489
1490
1491
1492
1493
1494
1495
1496
1497
1497
1498
1498
1499
1499
1500
1501
1502
1503
1504
1505
1506
1507
1508
1509
1509
1510
1511
1512
1513
1514
1515
1516
1517
1518
1519
1519
1520
1521
1522
1523
1524
1525
1526
1527
1528
1529
1529
1530
1531
1532
1533
1534
1535
1536
1537
1538
1539
1539
1540
1541
1542
1543
1544
1545
1546
1547
1548
1549
1549
1550
1551
1552
1553
1554
1555
1556
1557
1558
1559
1559
1560
1561
1562
1563
1564
1565
1566
1567
1568
1569
1569
1570
1571
1572
1573
1574
1575
1576
1577
1578
1579
1579
1580
1581
1582
1583
1584
1585
1586
1587
1588
1588
1589
1589
1590
1591
1592
1593
1594
1595
1596
1597
1597
1598
1598
1599
1599
1600
1601
1602
1603
1604
1605
1606
1607
1608
1609
1609
1610
1611
1612
1613
1614
1615
1616
1617
1618
1619
1619
1620
1621
1622
1623
1624
1625
1626
1627
1628
1629
1629
1630
1631
1632
1633
1634
1635
1636
1637
1638
1639
1639
1640
1641
1642
1643
1644
1645
1646
1647
1648
1649
1649
1650
1651
1652
1653
1654
1655
1656
1657
1658
1659
1659
1660
1661
1662
1663
1664
1665
1666
1667
1668
1669
1669
1670
1671
1672
1673
1674
1675
1676
1677
1678
1679
1679
1680
1681
1682
1683
1684
1685
1686
1687
1688
1688
1689
1689
1690
1691
1692
1693
1694
1695
1696
1697
1697
1698
1698
1699
1699
1700
1701
1702
1703
1704
1705
1706
1707
1708
1709
1709
1710
1711
1712
1713
1714
1715
1716
1717
1718
1719
1719
1720
1721
1722
1723
1724
1725
1726
1727
1728
1729
1729
1730
1731
1732
1733
1734
1735
1736
1737
1738
1739
1739
1740
1741
1742
1743
1744
1745
1746
1747
1748
1749
1749
1750
1751
1752
1753
1754
1755
1756
1757
1758
1759
1759
1760
1761
1762
1763
1764
1765
1766
1767
1768
1769
1769
1770
1771
1772
1773
1774
1775
1776
1777
1778
1779
1779
1780
1781
1782
1783
1784
1785
1786
1787
1788
1788
1789
1789
1790
1791
1792
1793
1794
1795
1796
1797
1797
1798
1798
1799
1799
1800
1801
1802
1803
1804
1805
1806
1807
1808
1809
1809
1810
1811
1812
1813
1814
1815
1816
1817
1818
1819
1819
1820
1821
1822
1823
1824
1825
1826
1827
1828
1829
1829
1830
1831
1832
1833
1834
1835
1836
1837
1838
1839
1839
1840
1841
1842
1843
1844
1845
1846
1847
1848
1849
1849
1850
1851
1852
1853
1854
1855
1856
1857
1858
1859
1859
1860
1861
1862
1863
1864
1865
1866
1867
1868
1869
1869
1870
1871
1872
1873
1874
1875
1876
1877
1878
1879
1879
1880
1881
1882
1883
1884
1885
1886
1887
1888
1888
1889
1889
1890
1891
1892
1893
1894
1895
1896
1897
1897
1898
1898
1899
1899
1900
1901
1902
1903
1904
1905
1906
1907
1908
1909
1909
1910
1911
1912
1913
1914
1915
1916
1917
1918
1919
1919
1920
1921
1922
1923
1924
1925
1926
1927
1928
1929
1929
1930
1931
1932
1933
1934
1935
1936
1937
1938
1939
1939
1940
1941
1942
1943
1944
1945
1946
1947
1948
1949
1949
1950
1951
1952
1953
1954
1955
1956
1957
1958
1959
1959
1960
1961
1962
1963
1964
1965
1966
1967
1968
1969
1969
1970
1971
1972
1973
1974
1975
1976
1977
1978
1979
1979
1980
1981
1982
1983
1984
1985
1986
1987
1988
1988
1989
1989
1990
1991
1992
1993
1994
1995
1996
1997
1997
1998
1998
1999
1999
2000
2001
2002
2003
2004
2005
2006
2007
2008
2009
2009
2010
2011
2012
2013
2014
2015
2016
2017
2018
2019
2020
2021
2022
2023
2024
2025
2026
2027
2028
2029
2030
2031
2032
2033
2034
2035
2036
2037
2038
2039
2039
2040
2041
2042
2043
2044
2045
2046
2047
2048
2049
2049
2050
2051
2052
2053
2054
2055
2056
2057
2058
2059
2059
2060
2061
2062
2063
2064
2065
2066
2067
2068
2069
2069
2070
2071
2072
2073
2074
2075
2076
2077
2078
2079
2079
2080
2081
2082
2083
2084
2085
2086
2087
2088
2088
2089
2089
2090
2091
2092
2093
2094
2095
2096
2097
2097
2098
2098
2099
2099
2100
2101
2102
2103
2104
2105
2106
2107
2108
2109
2109
2110
2111
2112
2113
2114
2115
2116
2117
2118
2119
2119
2120
2121
2122
2123
2124
2125
2126
2127
2128
2129
2129
2130
2131
2132
2133
2134
2135
2136
2137
2138
2139
2139
2140
2141
2142
2143
2144
2145
2146
2147
2148
2149
2149
2150
2151
2152
2153
2154
2155
2156
2157
2158
2159
2159
2160
2161
2162
2163
2164
2165
2166
2167
2168
2169
2169
2170
2171
2172
2173
2174
2175
2176
2177
2178
2179
2179
2180
2181
2182
2183
2184
2185
2186
2187
2188
2188
2189
2189
2190
2191
2192
2193
2194
2195
2196
2197
2197
2198
2198
2199
2199
2200
2201
2202
2203
2204
2205
2206
2207
2208
2209
2209
2210
2211
2212
2213
2214
2215
2216
2217
2218
2219
2219
2220
2221
2222
2223
2224
2225
2226
2227
2228
2229
2229
2230
2231
2232
2233
2234
2235
2236
2237
2238
2239
2239
2240
2241
2242
2243
2244
2245
2246
2247
2248
2249
2249
2250
2251
2252
2253
2254
2255
2256
2257
2258
2259
2259
2260
2261
2262
2263
2264
2265
2266
2267
2268
2269
2269
2270
2271
2272
2273
2274
2275
2276
2277
2278
2279
2279
2280
2281
2282
2283
2284
2285
2286
2287
2288
2288
2289
2289
2290
2291
2292
2293
2294
2295
2296
2297
2297
2298
2298
2299
2299
2300
2301
2302
2303
2304
2305
2306
2307
2308
2309
2309
2310
2311
2312
2313
2314
2315
2316
2317
2318
2319
2319
2320
2321
2322
2323
2324
2325
2326
2327
2328
2329
2329
2330
2331
2332
2333
2334
2335
2336
2337
2338
2339
2339
2340
2341
2342
2343
2344
2345
2346
2347
2348
2349
2349
2350
2351
2352
2353
2354
2355
2356
2357
2358
2359
2359
2360
2361
2362
2363
2364
2365
2366
2367
2368
2369
2369
2370
2371
2372
2373
2374
2375
2376
2377
2378
2379
2379
2380
2381
2382
2383
2384
2
```

Query 8

By using **COUNT(*)** and **GROUP BY** clause we've found the count of peoples interested in distinct courses.



The screenshot shows a Microsoft SQL Server Management Studio (SSMS) interface. The top bar displays the title 'DESKTOP-ONHARCT-WASM\sqlasm (System) - Microsoft SQL Server Management Studio'. Below the title bar is a toolbar with various icons for file operations, search, and database management. The main window is divided into several panes:

- Object Explorer**: Shows the database structure, including 'DESKTOP-ONHARCT-WASM (SQL Server 15.0.2000)' and its objects like 'Tables', 'Views', 'Procedures', and 'Triggers'.
- Script pane**: Contains the T-SQL query being executed:

```
USE [DESKTOP-ONHARCT-WASM]
GO
SELECT [dbo].[name],qualification,city FROM dbo.student WHERE qualification IN ('B.A','MCA','B.TECH','M.TECH','DIPLOMA') --LIST OF CANDIDATES HAVING TECHNICAL BACKGROUND
SELECT DISTINCT qualification,[State] FROM dbo.student WHERE qualification NOT IN ('B.A','MCA','B.TECH','M.TECH','DIPLOMA') --LIST OF STATES HAVING NON-TECHNICAL BACKGROUND
SELECT gender,[count(*)] AS [Gender wise count] FROM dbo.student GROUP BY gender -- COUNT OF STUDENTS BASED ON THEIR GENDER
SELECT course,[count(*)] AS [Course wise count] FROM dbo.student GROUP BY course -- COUNT OF STUDENTS BASED ON THEIR COURSE
```
- Status bar**: Shows 'Query executed successfully.' and the execution time '00:00:00'.

The central area of the interface is a large grid representing the query results. The columns are labeled 'Course' and 'Count'. The data is as follows:

Course	Count
NULL	0
Business Analyst	2
Data Science	75
Data Analyst	1
Data Mining	24
Data Analytics	233
Data Science	16
Data Mining	16
Machine Learning	28
Programming	162
Finance	1
Software	12

Query 9

In this query we've found there're numerous experienced peoples who're looking for job change.

With the help of **SELECT, WHERE, LOGICAL OPERATORS, NOT IN** keywords we made this insight

```
Desktop-CINAMCY-WIN3M (SQL Server 15.0.2000) - [master] - Microsoft SQL Server Management Studio [Quick Launch (Ctrl+Q)]
```

Object Explorer Task List Results Explorer Object Explorer (New Query) New Query

REDSHIFT-DW[DW] - [master] - [Log: DESKTOP-CINAMCY-WIN3M (SQL Server 15.0.2000)]

```
1 USE [Log]
2 GO
3 SELECT * FROM [full_name].[app].[qualification_profession]([year_of_experience] FROM
4 [dbo].[student] WHERE age >= 25 AND qualification NOT IN ('C', 'Mech', 'Tech', 'M.Tech', 'B.Tech')
5 AND profession='professional' ) EXCEPTED PERSON LOOKING FOR CAREER TRANSACTION
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
559
560
561
562
563
564
565
566
567
568
569
569
570
571
572
573
574
575
576
577
578
579
579
580
581
582
583
584
585
586
587
588
589
589
590
591
592
593
594
595
596
597
598
599
599
600
601
602
603
604
605
606
607
608
609
609
610
611
612
613
614
615
616
617
617
618
619
619
620
621
622
623
624
625
625
626
627
627
628
628
629
629
630
630
631
631
632
632
633
633
634
634
635
635
636
636
637
637
638
638
639
639
640
640
641
641
642
642
643
643
644
644
645
645
646
646
647
647
648
648
649
649
650
650
651
651
652
652
653
653
654
654
655
655
656
656
657
657
658
658
659
659
660
660
661
661
662
662
663
663
664
664
665
665
666
666
667
667
668
668
669
669
670
670
671
671
672
672
673
673
674
674
675
675
676
676
677
677
678
678
679
679
680
680
681
681
682
682
683
683
684
684
685
685
686
686
687
687
688
688
689
689
690
690
691
691
692
692
693
693
694
694
695
695
696
696
697
697
698
698
699
699
700
700
701
701
702
702
703
703
704
704
705
705
706
706
707
707
708
708
709
709
710
710
711
711
712
712
713
713
714
714
715
715
716
716
717
717
718
718
719
719
720
720
721
721
722
722
723
723
724
724
725
725
726
726
727
727
728
728
729
729
730
730
731
731
732
732
733
733
734
734
735
735
736
736
737
737
738
738
739
739
740
740
741
741
742
742
743
743
744
744
745
745
746
746
747
747
748
748
749
749
750
750
751
751
752
752
753
753
754
754
755
755
756
756
757
757
758
758
759
759
760
760
761
761
762
762
763
763
764
764
765
765
766
766
767
767
768
768
769
769
770
770
771
771
772
772
773
773
774
774
775
775
776
776
777
777
778
778
779
779
780
780
781
781
782
782
783
783
784
784
785
785
786
786
787
787
788
788
789
789
790
790
791
791
792
792
793
793
794
794
795
795
796
796
797
797
798
798
799
799
800
800
801
801
802
802
803
803
804
804
805
805
806
806
807
807
808
808
809
809
810
810
811
811
812
812
813
813
814
814
815
815
816
816
817
817
818
818
819
819
820
820
821
821
822
822
823
823
824
824
825
825
826
826
827
827
828
828
829
829
830
830
831
831
832
832
833
833
834
834
835
835
836
836
837
837
838
838
839
839
840
840
841
841
842
842
843
843
844
844
845
845
846
846
847
847
848
848
849
849
850
850
851
851
852
852
853
853
854
854
855
855
856
856
857
857
858
858
859
859
860
860
861
861
862
862
863
863
864
864
865
865
866
866
867
867
868
868
869
869
870
870
871
871
872
872
873
873
874
874
875
875
876
876
877
877
878
878
879
879
880
880
881
881
882
882
883
883
884
884
885
885
886
886
887
887
888
888
889
889
890
890
891
891
892
892
893
893
894
894
895
895
896
896
897
897
898
898
899
899
900
900
901
901
902
902
903
903
904
904
905
905
906
906
907
907
908
908
909
909
910
910
911
911
912
912
913
913
914
914
915
915
916
916
917
917
918
918
919
919
920
920
921
921
922
922
923
923
924
924
925
925
926
926
927
927
928
928
929
929
930
930
931
931
932
932
933
933
934
934
935
935
936
936
937
937
938
938
939
939
940
940
941
941
942
942
943
943
944
944
945
945
946
946
947
947
948
948
949
949
950
950
951
951
952
952
953
953
954
954
955
955
956
956
957
957
958
958
959
959
960
960
961
961
962
962
963
963
964
964
965
965
966
966
967
967
968
968
969
969
970
970
971
971
972
972
973
973
974
974
975
975
976
976
977
977
978
978
979
979
980
980
981
981
982
982
983
983
984
984
985
985
986
986
987
987
988
988
989
989
990
990
991
991
992
992
993
993
994
994
995
995
996
996
997
997
998
998
999
999
1000
1000
1001
1001
1002
1002
1003
1003
1004
1004
1005
1005
1006
1006
1007
1007
1008
1008
1009
1009
1010
1010
1011
1011
1012
1012
1013
1013
1014
1014
1015
1015
1016
1016
1017
1017
1018
1018
1019
1019
1020
1020
1021
1021
1022
1022
1023
1023
1024
1024
1025
1025
1026
1026
1027
1027
1028
1028
1029
1029
1030
1030
1031
1031
1032
1032
1033
1033
1034
1034
1035
1035
1036
1036
1037
1037
1038
1038
1039
1039
1040
1040
1041
1041
1042
1042
1043
1043
1044
1044
1045
1045
1046
1046
1047
1047
1048
1048
1049
1049
1050
1050
1051
1051
1052
1052
1053
1053
1054
1054
1055
1055
1056
1056
1057
1057
1058
1058
1059
1059
1060
1060
1061
1061
1062
1062
1063
1063
1064
1064
1065
1065
1066
1066
1067
1067
1068
1068
1069
1069
1070
1070
1071
1071
1072
1072
1073
1073
1074
1074
1075
1075
1076
1076
1077
1077
1078
1078
1079
1079
1080
1080
1081
1081
1082
1082
1083
1083
1084
1084
1085
1085
1086
1086
1087
1087
1088
1088
1089
1089
1090
1090
1091
1091
1092
1092
1093
1093
1094
1094
1095
1095
1096
1096
1097
1097
1098
1098
1099
1099
1100
1100
1101
1101
1102
1102
1103
1103
1104
1104
1105
1105
1106
1106
1107
1107
1108
1108
1109
1109
1110
1110
1111
1111
1112
1112
1113
1113
1114
1114
1115
1115
1116
1116
1117
1117
1118
1118
1119
1119
1120
1120
1121
1121
1122
1122
1123
1123
1124
1124
1125
1125
1126
1126
1127
1127
1128
1128
1129
1129
1130
1130
1131
1131
1132
1132
1133
1133
1134
1134
1135
1135
1136
1136
1137
1137
1138
1138
1139
1139
1140
1140
1141
1141
1142
1142
1143
1143
1144
1144
1145
1145
1146
1146
1147
1147
1148
1148
1149
1149
1150
1150
1151
1151
1152
1152
1153
1153
1154
1154
1155
1155
1156
1156
1157
1157
1158
1158
1159
1159
1160
1160
1161
1161
1162
1162
1163
1163
1164
1164
1165
1165
1166
1166
1167
1167
1168
1168
1169
1169
1170
1170
1171
1171
1172
1172
1173
1173
1174
1174
1175
1175
1176
1176
1177
1177
1178
1178
1179
1179
1180
1180
1181
1181
1182
1182
1183
1183
1184
1184
1185
1185
1186
1186
1187
1187
1188
1188
1189
1189
1190
1190
1191
1191
1192
1192
1193
1193
1194
1194
1195
1195
1196
1196
1197
1197
1198
1198
1199
1199
1200
1200
1201
1201
1202
1202
1203
1203
1204
1204
1205
1205
1206
1206
1207
1207
1208
1208
1209
1209
1210
1210
1211
1211
1212
1212
1213
1213
1214
1214
1215
1215
1216
1216
1217
1217
1218
1218
1219
1219
1220
1220
1221
1221
1222
1222
1223
1223
1224
1224
1225
1225
1226
1226
1227
1227
1228
1228
1229
1229
1230
1230
1231
1231
1232
1232
1233
1233
1234
1234
1235
1235
1236
1236
1237
1237
1238
1238
1239
1239
1240
1240
1241
1241
1242
1242
1243
1243
1244
1244
1245
1245
1246
1246
1247
1247
1248
1248
1249
1249
1250
1250
1251
1251
1252
1252
1253
1253
1254
1254
1255
1255
1256
1256
1257
1257
1258
1258
1259
1259
1260
1260
1261
1261
1262
1262
1263
1263
1264
1264
1265
1265
1266
1266
1267
1267
1268
1268
1269
1269
1270
1270
1271
1271
1272
1272
1273
1273
1274
1274
1275
1275
1276
1276
1277
1277
1278
1278
1279
1279
1280
1280
1281
1281
1282
1282
1283
1283
1284
1284
1285
1285
1286
1286
1287
1287
1288
1288
1289
1289
1290
1290
1291
1291
1292
1292
1293
1293
1294
1294
1295
1295
1296
1296
1297
1297
1298
1298
1299
1299
1300
1300
1301
1301
1302
1302
1303
1303
1304
1304
1305
1305
1306
1306
1307
1307
1308
1308
1309
1309
1310
1310
1311
1311
1312
1312
1313
1313
1314
1314
1315
1315
1316
1316
1317
1317
1318
1318
1319
1319
1320
1320
1321
1321
1322
1322
1323
1323
1324
1324
1325
1325
1326
1326
1327
1327
1328
1328
1329
1329
1330
1330
1331
1331
1332
1332
1333
1333
1334
1334
1335
1335
1336
1336
1337
1337
1338
1338
1339
1339
1340
1340
1341
1341
1342
1342
1343
1343
1344
1344
1345
1345
1346
1346
1347
1347
1348
1348
1349
1349
1350
1350
1351
1351
1352
1352
1353
1353
1354
1354
1355
1355
1356
1356
1357
1357
1358
1358
1359
1359
1360
1360
1361
1361
1362
1362
1363
1363
1364
1364
1365
1365
1366
1366
1367
1367
1368
1368
1369
1369
1370
1370
1371
1371
1372
1372
1373
1373
1374
1374
1375
1375
1376
1376
1377
1377
1378
1378
1379
1379
1380
1380
1381
1381
1382
1382
1383
1383
1384
1384
1385
1385
1386
1386
1387
1387
1388
1388
1389
1389
1390
1390
1391
1391
1392
1392
1393
1393
1394
1394
1395
1395
1396
1396
1397
1397
1398
1398
1399
1399
1400
1400
1401
1401
1402
1402
1403
1403
1404
1404
1405
1405
1406
1406
1407
1407
1408
1408
1409
1409
1410
1410
1411
1411
1412
1412
1413
1413
1414
1414
1415
1415
1416
1416
1417
1417
1418
1418
1419
1419
1420
1420
1421
1421
1422
1422
1423
1423
1424
1424
1425
1425
1426
1426
1427
1427
1428
1428
1429
1429
1430
1430
1431
1431
1432
1432
1433
1433
1434
1434
1435
1435
1436
1436
1437
1437
1438
1438
1439
1439
1440
1440
1441
1441
1442
1442
1443
1443
1444
1444
1445
1445
1446
1446
1447
1447
1448
1448
1449
1449
1450
1450
1451
1451
1452
1452
1453
1453
1454
1454
1455
1455
1456
1456
1457
1457
1458
1458
1459
1459
1460
1460
1461
1461
1462
1462
1463
1463
1464
1464
1465
1465
1466
1466
1467
1467
1468
1468
1469
1469
1470
1470
1471
1471
1472
1472
1473
1473
1474
1474
1475
1475
1476
1476
1477
1477
1478
1478
1479
1479
1480
1480
1481
1481
1482
1482
1483
1483
1484
1484
1485
1485
1486
1486
1487
1487
1488
1488
1489
1489
1490
1490
1491
1491
1492
1492
1493
1493
1494
1494
1495
1495
1496
1496
1497
1497
1498
1498
1499
1499
1500
1500
1501
1501
1502
1502
1503
1503
1504
1504
1505
1505
1506
1506
1507
1507
1508
1508
1509
1509
1510
1510
1511
1511
1512
1512
1513
1513
1514
1514
1515
1515
1516
1516
1517
1517
1518
1518
1519
1519
1520
1520
1521
1521
1522
1522
1523
1523
1524
1524
1525
1525
1526
1526
1527
1527
1528
1528
1529
1529
1530
1530
1531
1531
1532
1532
1533
1533
1534
1534
1535
1535
1536
1536
1537
1537
1538
1538
1539
1539
1540
1540
1541
1541
1542
1542
1543
1543
1544
1544
1545
1545
1546
1546
1547
1547
1548
1548
1549
1549
1550
1550
1551
1551
1552
1552
1553
1553
1554
1554
1555
1555
1556
1556
1557
1557
1558
1558
1559
1559
1560
1560
1561
1561
1562
1562
1563
1563
1564
1564
1565
1565
1566
1566
1567
1567
1568
1568
1569
1569
1570
1570
1571
1571
1572
1572
1573
1573
1574
1574
1575
1575
1576
1576
1577
1577
1578
1578
1579
1579
1580
1580
1581
1581
1582
1582
1583
1583
1584
1584
1585
1585
1586
1586
1587
1587
1588
1588
1589
1589
1590
1590
1591
1591
1592
1592
1593
1593
1594
1594
1595
1595
1596
1596
1597
1597
1598
1598
1599
1599
1600
1600
1601
1601
1602
1602
1603
1603
1604
1604
1605
1605
1606
1606
1607
1607
1608
1608
1609
1609
1610
1610
1611
1611
1612
1612
1613
1613
1614
1614
1615
1615
1616
1616
1617
1617
1618
1618
1619
1619
1620
1620
1621
1621
1622
1622
1623
1623
1624
1624
1625
1625
1626
1626
1627
1627
1628
1628
1629
1629
1630
1630
1631
1631
1632
1632
1633
1633
1634
1634
1635
1635
1636
1636
1637
1637
1638
1638
1639
1639
1640
1640
1641
1641
1642
1642
1643
1643
1644
1644
1645
1645
1646
1646
1647
1647
1648
1648
1649
1649
1650
1650
1651
1651
1652
1652
1653
1653
1654
1654
1655
1655
1656
1656
1657
1657
1658
1658
1659
1659
1660
1660
1661
1661
1662
1662
1663
1663
1664
1664
1665
1665
1666
1666
1667
1667
1668
1668
1669
1669
1670
1670
1671
1671
1672
1672
1673
1673
16
```

Load

Finally, once the process has transformed, sorted, cleaned, validated, and prepared the data, you need to load it into data storage somewhere. The most common target database is a data warehouse, a centralized repository designed to work with BI and analytics systems.

In our case, we've used the above insights as a reference for our visualization.

PHASE 4 : EXCEL VISUALIZATION

Data visualization refers to the techniques used to communicate data or information by encoding it as visual objects (e.g., points, lines, or bars) contained in graphics. The goal is to communicate information clearly and efficiently to users. It is one of the steps in D According to Vitaly Friedman (2008) the "main goal of data visualization is to communicate information clearly and effectively through graphical means. It doesn't mean that data visualization needs to look boring to be functional or extremely sophisticated to look beautiful. To convey ideas effectively, both aesthetic form and functionality need to go hand in hand, providing insights into a rather sparse and complex data set by communicating its key aspects in a more intuitive way.

To communicate information clearly and efficiently, data visualization uses statistical graphics, plots, Information graphics and other tools. Numerical data may be encoded using dots, lines, or bars, to visually communicate a quantitative message. Effective visualization helps users analyze and reason about data and evidence. It makes complex data more accessible, understandable, and usable. Users may have particular analytical tasks, such as making comparisons or understanding causality, and the design principle of the graphic (i.e., showing comparisons or showing causality) follows the task. Tables are generally used where users will look up a specific measurement, while charts of various types are used to show patterns or relationships in the data for one or more variables.

1. Data Visualization in EXCEL:

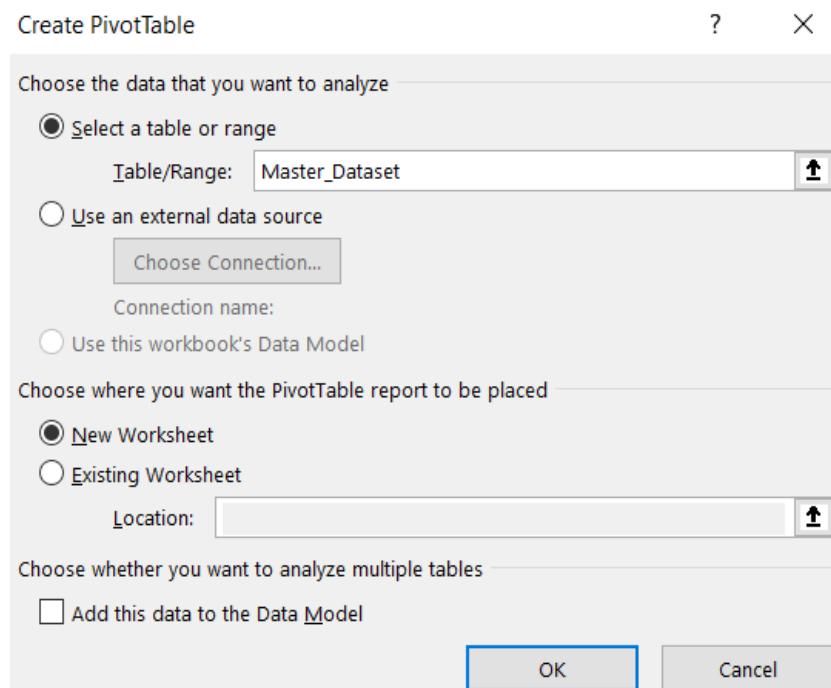


Step 1 – Create Table from the available dataset i.e., Master Table:

	A	B	C	D	E	F	G	H	I	J	K	L	M
1	ID	FULL NAME	GENDER	AGE	QUALIFICATION	COURSE	SKILLS	COURSE DURATION	PASS OUT YEAR	YEAR OF EXPERIENCE	PROFESSION	MONTH JOINED	MAIL ID
2	1	Ram Devi Kumar	Male	25	B.Sc	Data Science	AI	4	2018	1	PROFESSIONAL	April	ram@gmail.com
3	2	Mohammaed Amir Hussian	Male	24	B.Tech	Business Analyst	Power Bi	5	2022	2	PROFESSIONAL	December	mohammaed@gmail.com
4	3	Sunita Singh	Female	27	B.Ed	Database	Oracle	4	2016	1	PROFESSIONAL	July	sunita@gmail.com
5	4	Sri Das	Male	30	M.Sc	Data Analyst	Tableau	3	2013	0	FRESHER	May	sri@gmail.com
6	5	Anita Kaur	Female	23	BBA	Machine Learning	AI	3	2021	0	FRESHER	November	anita@gmail.com
7	6	Santosh Deshmukh	Male	31	M.Tech	Business Analyst	Power Bi	3	2012	0	FRESHER	December	santosh@gmail.com
8	7	Sanjay Jain	Male	22	BCA	Machine Learning	AI	3	2020	0	FRESHER	July	sanjay@gmail.com
9	8	Sunil Lal	Male	28	MBA	Data Science	AI	3	2015	2	PROFESSIONAL	April	sunil@gmail.com
10	9	Rajesh Yadav	Male	27	B.Com	Data Science	AI	3	2016	0	FRESHER	February	rajesh@gmail.com
11	10	Gita Ram Singh	Female	29	B.Ed	Database	Oracle	3	2014	1	PROFESSIONAL	March	gita@gmail.com
12	11	Ramesh Reddy	Male	22	B.E	Database	Oracle	3	2020	5	PROFESSIONAL	November	ramesh@gmail.com
13	12	Ashok	Male	24	M.Sc	Machine Learning	AI	3	2019	1	PROFESSIONAL	May	ashok@gmail.com
14	13	Rekha Kumari	Female	21	B.Tech	Business Analyst	Power Bi	3	2021	0	FRESHER	July	rekha@gmail.com
15	14	Manoj K Jain	Male	25	BBA	Database	Oracle	5	2018	0	FRESHER	June	manoj@gmail.com
16	15	Anil Pathak	Male	28	B.E	Database	Oracle	5	2015	0	FRESHER	August	anil@gmail.com
17	16	Suresh Raja Nair	Male	30	B.Ed	Data Science	AI	4	2013	2	PROFESSIONAL	August	suresh@gmail.com
18	17	Vijay Devi Prasad	Male	21	B.Ed	Data Analyst	Tableau	5	2021	0	FRESHER	January	vijay@gmail.com
19	18	Lakshmi Ben Sharma	Female	23	M.Sc	Machine Learning	AI	5	2020	1	PROFESSIONAL	February	lakshmi@gmail.com
20	19	Abdul Shah Ali	Male	26	B.Com	Data Analyst	Tableau	3	2018	3	PROFESSIONAL	December	abdul@gmail.com
21	20	Raju Sahu	Male	22	BCA	Data Science	AI	3	2020	0	FRESHER	June	raju@gmail.com
22	21	Manju Gupta	Male	29	B.Com	Data Analyst	Tableau	4	2013	4	PROFESSIONAL	August	manju@gmail.com
23	22	Shanti Paramar	Female	24	M.Sc	Data Analyst	Tableau	5	2020	1	PROFESSIONAL	November	shanti@gmail.com
24	23	Vinod Jadhav	Female	20	BBA	Machine Learning	AI	5	2022	3	PROFESSIONAL	January	vinod@gmail.com
25	24	Suman Rani	Female	23	B.Tech	Data Analyst	Excel	5	2016	2	PROFESSIONAL	December	suman@gmail.com
26	25	Rajendra Mishra	Male	23	BBA	Data Analyst	Tableau	3	2021	0	FRESHER	February	rajendra@gmail.com
27	26	Raj Roy Chand	Male	29	M.Tech	Data Analyst	Tableau	3	2015	1	PROFESSIONAL	December	raj@gmail.com
28	27	Usha Saha	Female	28	B.Com	Machine Learning	AI	5	2015	3	PROFESSIONAL	May	usha@gmail.com
29	28	Asha Thakur	Female	21	B.E	Data Analyst	Tableau	4	2022	1	PROFESSIONAL	August	asha@gmail.com
30	29	Dinesh	Male	20	BBA	Data Analyst	Tableau	3	2021	1	PROFESSIONAL	September	dinesh@gmail.com
31	30	Krishna Pradhan	Female	26	BBA	Data Analyst	Tableau	5	2017	3	PROFESSIONAL	January	krishna@gmail.com
32	31	Mina Malik	Female	24	B.Sc	Business Analyst	Power Bi	5	2019	2	PROFESSIONAL	November	mina@gmail.com
33	32	Kiran Patra	Female	21	BCA	Database	Oracle	3	2021	0	FRESHER	October	kiran@gmail.com
34	33	Dilip Jha	Male	25	B.Com	Database	Oracle	5	2018	1	PROFESSIONAL	October	dilip@gmail.com
35	34	Pushpa Pandey	Male	30	M.Tech	Machine Learning	AI	4	2015	0	FRESHER	May	pushpa@gmail.com
36	35	Ajay Raj Singh	Male	22	M.Sc	Machine Learning	AI	4	2020	0	FRESHER	May	ajay@gmail.com
37	36	Urmila Pandit	Female	23	BCA	Business Analyst	Power Bi	3	2021	1	PROFESSIONAL	May	urmila@gmail.com
38	37	Mukesh Kamble	Male	21	B.Tech	Data Analyst	Excel	3	2022	0	FRESHER	July	mukesh@gmail.com

FIG. MASTER DATASET

Step 2 – Create Pivot Table from the Master Dataset:

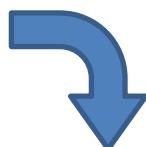


Step 3 – Age Wise distribution [Column Chart]:

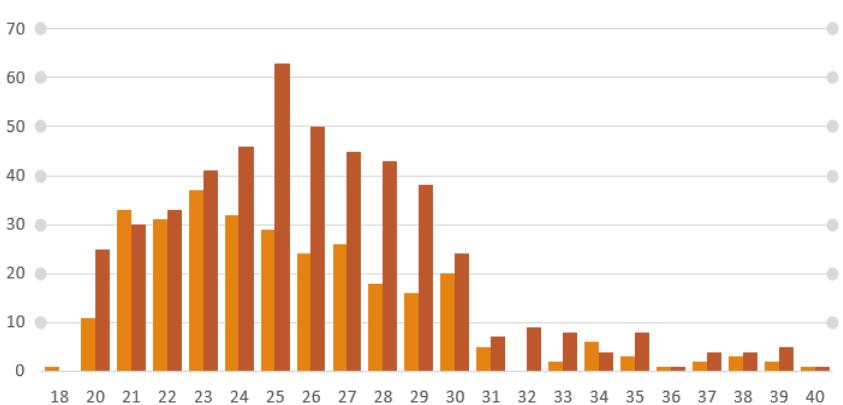
Count of PROFESSION	Column Labels	FRESHER	PROFESSIONAL	Grand Total
Row Labels				
18		1		1
20		11	25	36
21		33	30	63
22		31	33	64
23		37	41	78
24		32	46	78
25		29	63	92
26		24	50	74
27		26	45	71
28		18	43	61
29		16	38	54
30		20	24	44
31		5	7	12
32			9	9
33		2	8	10
34		6	4	10
35		3	8	11
36		1	1	2
37		2	4	6
38		3	4	7
39		2	5	7
40		1	1	2
Grand Total		303	489	792

FIG. AGE WISE

The AGE value is dragged into the Row Section and PROFESSION is dragged into the Value & Columns Section of the Pivot, to form the Pivot Table. (FIG. AGE WISE)



AGE-WISE DISTRIBUTIONS



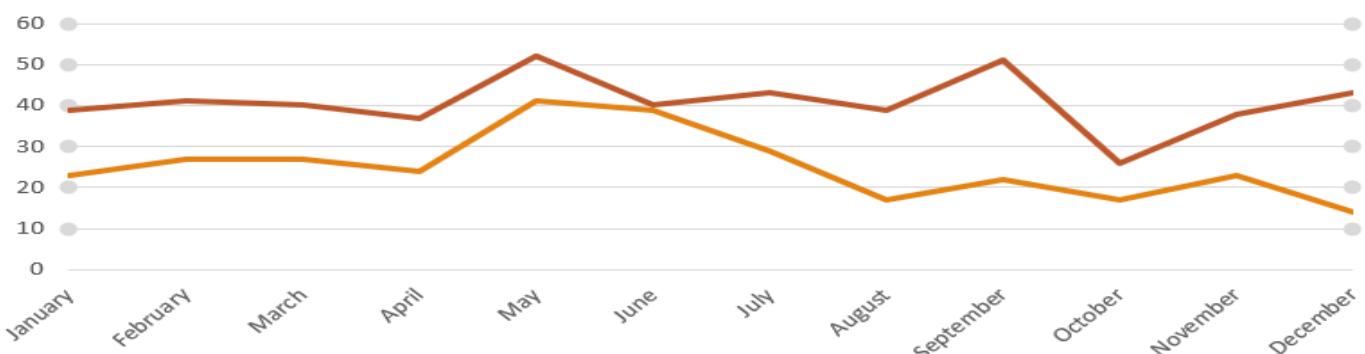
Step 4 – Month Wise distribution [Line Chart]:

Count of PROFESSION	Column Labels	FRESHER	PROFESSIONAL	Grand Total
Row Labels				
January		23	39	62
February		27	41	68
March		27	40	67
April		24	37	61
May		41	52	93
June		39	40	79
July		29	43	72
August		17	39	56
September		22	51	73
October		17	26	43
November		23	38	61
December		14	43	57
Grand Total		303	489	792

FIG. MONTH WISE

The MONTH of Joining value is dragged into the Row Section and PROFESSION Is dragged into the Value & Columns Section of the Pivot, to form the Pivot Table. (FIG. MONTH WISE)

MONTH WISE DISTRIBUTION



Step 4 – State Wise distribution [Bar Chart]:

Row Labels	Column Labels		
	FRESHER	PROFESSIONAL	Grand Total
Leh	1	1	1
Goa		3	3
Himachal Pradesh	2	4	6
New Delhi	3	3	6
Assam	2	6	8
Chattisgarh	5	6	11
Haryana	7	8	15
Delhi	6	9	15
Telangana	8	10	18
Madhya Pradesh	10	9	19
Jharkhand	4	16	20
Punjab	7	13	20
Rajasthan	12	15	27
Andhra Pradesh	9	21	30
Uttar Pradesh	12	21	33
Bihar	16	21	37
Kerala	17	23	40
Gujarat	17	23	40
West Bengal	15	31	46
Tamil-Nadu	18	29	47
Maharashtra	14	50	64
Karnataka	37	45	82
Odisha	81	123	204
Grand Total	303	489	792

STATE WISE

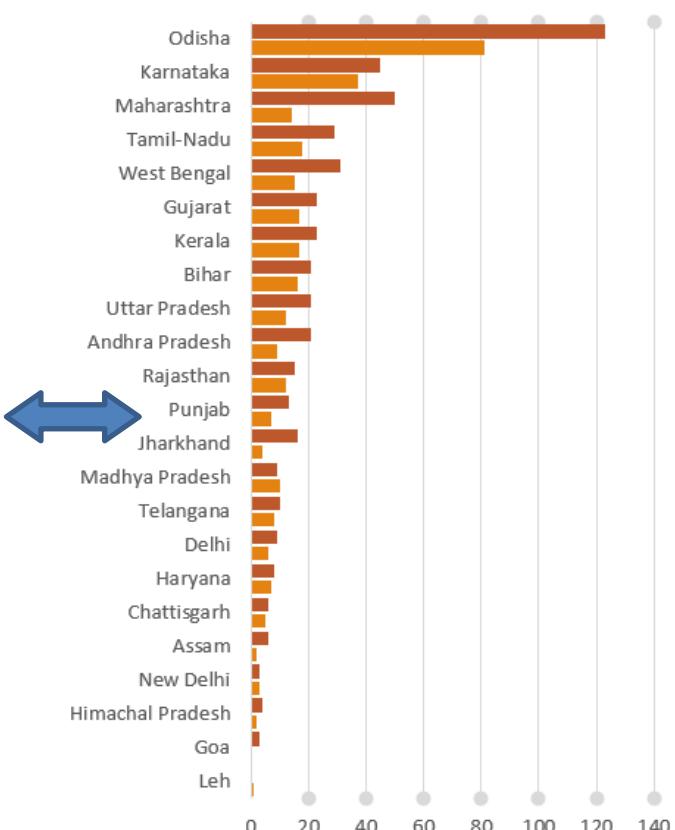


FIG. STATE WISE

The STATE is dragged into the Row Section and PROFESSION Is dragged into the Value & Columns Section of the Pivot, to form the Pivot Table. (FIG. STATE WISE)

Step 5 – Most Used Skill's distribution [Tree Map]:

SKILLS	TOTAL
AI	49
C++	23
Excel	105
IT	12
Java	43
Oracle	67
Power Bi	190
Python	94
R	17
SQL	105
Tableau	87

As Tree Map is not possible within Pivot Table, so is formed separately using the values and formed the table in order to apply the Tree Maps, by which we get to know which is most used skills through it. (FIG. USED SKILLS)

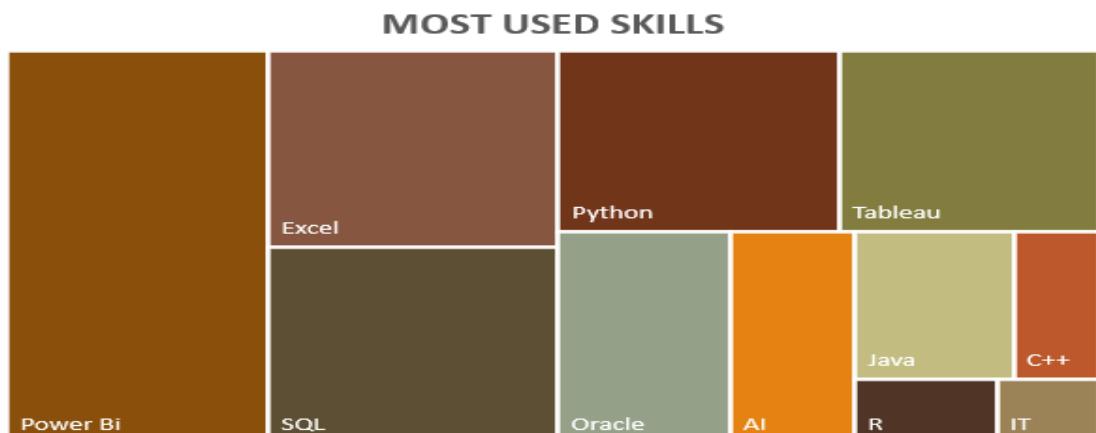


FIG. USED SKILLS

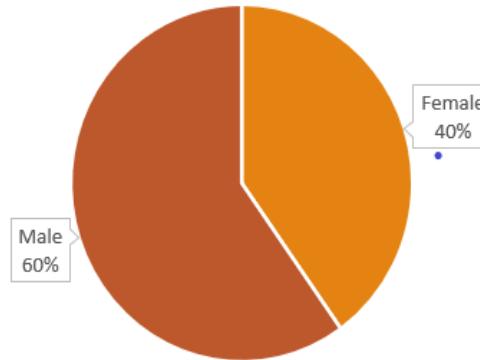
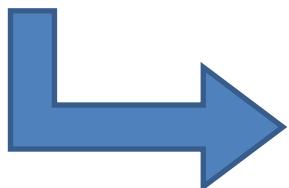
Step 6 –Gender Wise distribution [Pie Chart]:

Row Labels	Count of GENDER
Female	318
Male	474
Grand Total	792

The GENDER is dragged into the Row Section and also the GENDER Is dragged into the Value Section of the Pivot, to form the Pivot Table. (FIG. GENDER WISE)

FIG. GENDER WISE

GENDER WISE



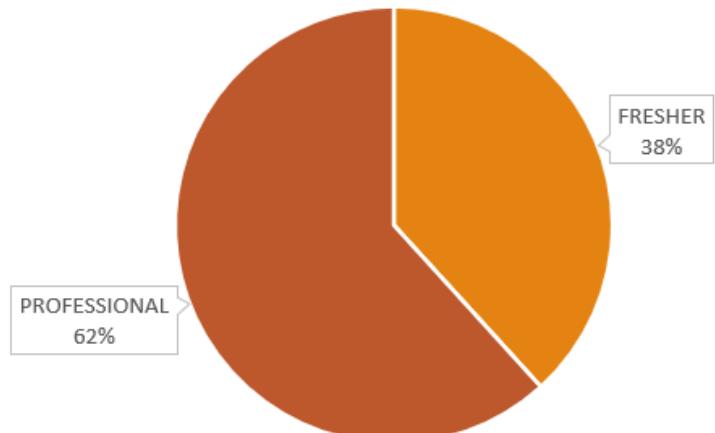
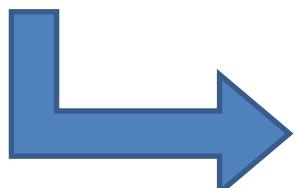
Step 6 –Professional Wise distribution [Pie Chart]:

Row Labels	Count of PROFESSION
FRESHER	303
PROFESSIONAL	489
Grand Total	792

The PROFESSIONAL is dragged into the Row Section and also the PROFESSIONAL Is dragged into the Value Section of the Pivot, to form the Pivot Table. (FIG. PROFESSIONAL WISE)

FIG. PROFESSIONAL WISE

EXPERIENCE WISE



Step 6 – Dashboard [Student Data Management for Data Gyan]:

A dashboard is a tool used for **information management** and **business intelligence**. Much like the dashboard of a car, data dashboards organize, store, and display important information from multiple data sources into one, easy-to-access place.

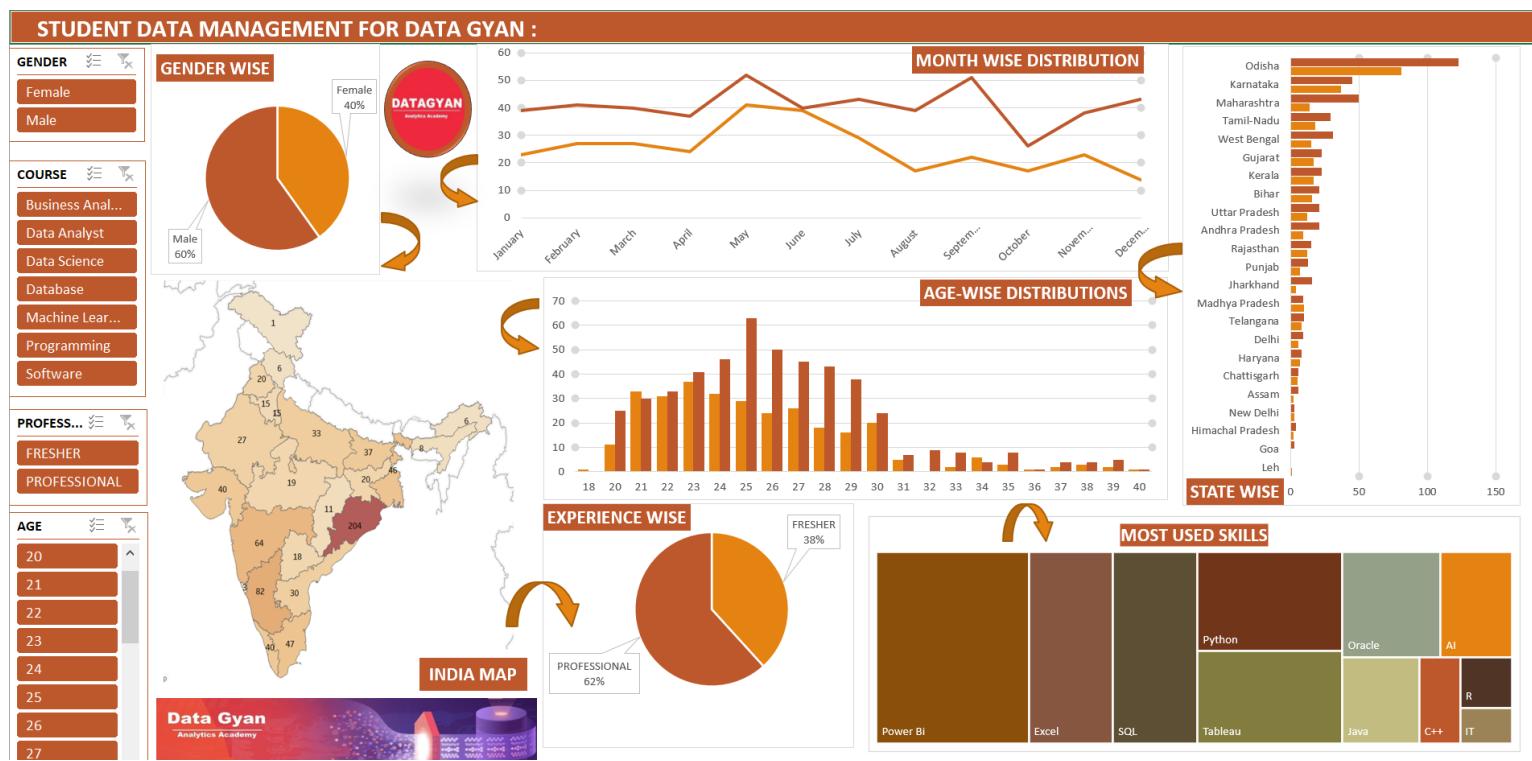


FIG. DASHBOARD FOR THE DATA VISUALIZATION

PHASE 5 : TABLEAU VISUALIZATION

There are various tools available for data visualization which helps in presenting the visualizations in a very simple manner. Tableau is one of the fastest-growing data visualization tools which is presently being used a lot in the BI industry. This software provides the most efficient way to change or transform the raw data into an easily understandable format.

Why Tableau Is the Most Popular Data Visualization Tool?

Tableau is used more frequently as the tool allows to analysis the data more quickly and visualizations are generated as dashboards and worksheets. Tableau enables us to make dashboards that give actionable insights and spreads the business faster. Tableau products are always operated in virtualized environments if they are configured with proper hardware and operating system. There are plenty of reasons listed below that will tell you why businesses use Tableau as a Data Visualization tool

- Great visualizations
- Detailed Insights
- User-friendly Approach
- Working with the different Data Sources
- Data set adding is easy

Process of Import Master Dataset into Tableau from MS-Excel?

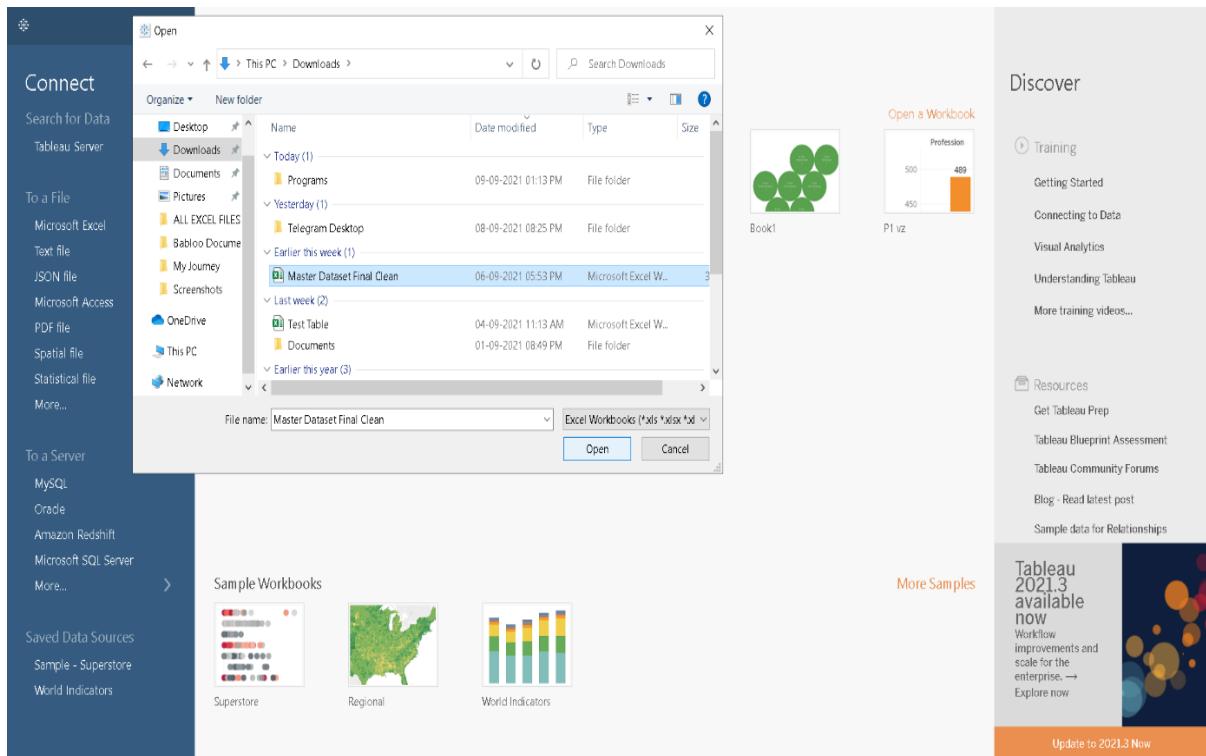
Step 1- Open the Tableau Terminal:

- Firstly, we have to open the Tableau software.



Step 2 – Import Master Dataset From MS-Excel:

- Click on Microsoft Excel option in Tableau software and select the master dataset from preferred location in your desktop.



Step 3 – Drag and Drop Method Using on Data Source:

- After successfully imported the dataset into tableau then drag the required dataset to Data Source and then we have to change the data type based on object type.

This screenshot shows the Tableau Data Source editor. The 'Connections' pane on the left lists a single connection named 'Master Dataset Final Clean' (Microsoft Excel). The 'Sheets' pane below it lists several sheets: Facebook Dataset, Internal Dataset, Linked Dataset, Master Dataset, Nauki Dataset, Text Table, and New Union. The main workspace shows a grid icon with the text 'Drag tables here'. At the bottom, there are buttons for 'Sort fields', 'Data source order', 'Show aliases', 'Show hidden fields', and a 'rows' dropdown. The status bar at the bottom indicates '0 Data Source' and 'Sheet 1'.



Screenshot of the Tableau Data Source page:

The interface shows the "Master Dataset (Master Dataset Final Clean)" connection selected in the Connections pane. The Sheets pane lists various datasets: Facebook Dataset, Internal Dataset, Linked Dataset, Master Dataset, Naukri Dataset, Test Table, and New Union. The "New Union" sheet is currently active, displaying a data preview with 792 rows. The columns include ID, Full Name, Gender, AGE, Qualification, Course, Skills, Course Duration, Pass Out Year, Year Of Experi..., Profession, Month Joined, and Mail.

ID	Full Name	Gender	AGE	Qualification	Course	Skills	Course Duration	Pass Out Year	Year Of Experi...	Profession	Month Joined	Mail
1	Ram Devi Kumar	Male	25	B.Sc	Data Science	AI		4	2,018	1	PROFESSIONAL	April
2	Mohammaed Ami...	Male	24	B.Tech	Business Analyst	Power Bi		5	2,022	2	PROFESSIONAL	December
3	Sunita Singh	Female	27	B.Ed	Database	Oracle		4	2,016	1	PROFESSIONAL	July
4	Sri Das	Male	30	M.Sc	Data Analyst	Tableau		3	2,013	0	FRESHER	May
5	Anita Kaur	Female	23	BBA	Machine Learning	AI		3	2,021	0	FRESHER	November
6	Santosh Deshmu...	Male	31	M.Tech	Business Analyst	Power Bi		3	2,012	0	FRESHER	December
7	Sanjay Jain	Male	22	BCA	Machine Learning	AI		3	2,020	0	FRESHER	July
8	Sunil Lal	Male	28	MBA	Data Science	AI		3	2,015	2	PROFESSIONAL	April
9	Rajesh Yadav	Male	27	B.Com	Data Science	AI		3	2,016	0	FRESHER	February
10	Gita Ram Singh	Female	29	B.Ed	Database	Oracle		3	2,014	1	PROFESSIONAL	March

At the bottom, there are tabs for "Data Source" and "Sheet1".

FIG. DATA SOURCE PAGE

Working Mechanism of Tableau

The below steps we were followed to do before visualization on Tableau:

- Change the datatype of required columns.
- Push the master data to our new work sheet by clicking the +(plus) icon.
- Select the required fields for visualization and then apply drag and drop method.
- Placing the Dimensions and Measures in the Columns and Rows shelves to get the required Charts and Graphs according to the requirements.
- By following the above steps, we did some visualization.

VISUALIZATION

1. Count of Profession:

[Year of Experience Vs Profession]

Here it displays the number of freshers and professional available in master dataset.

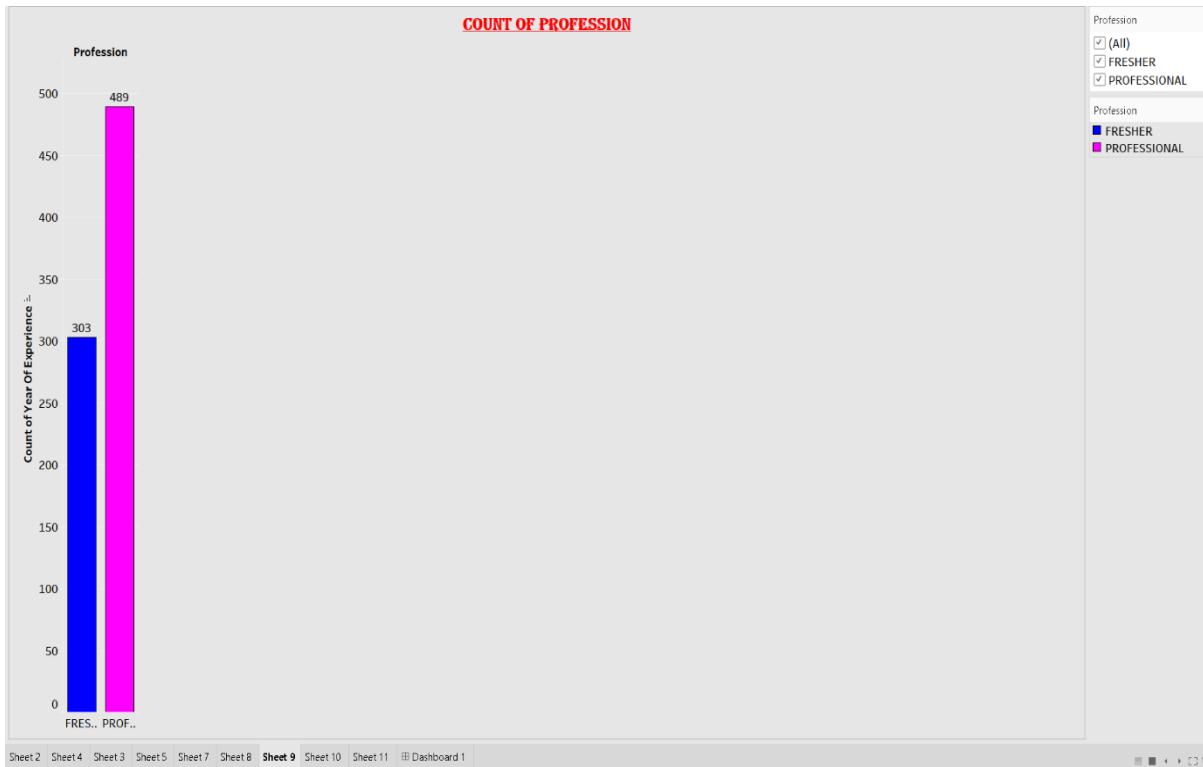
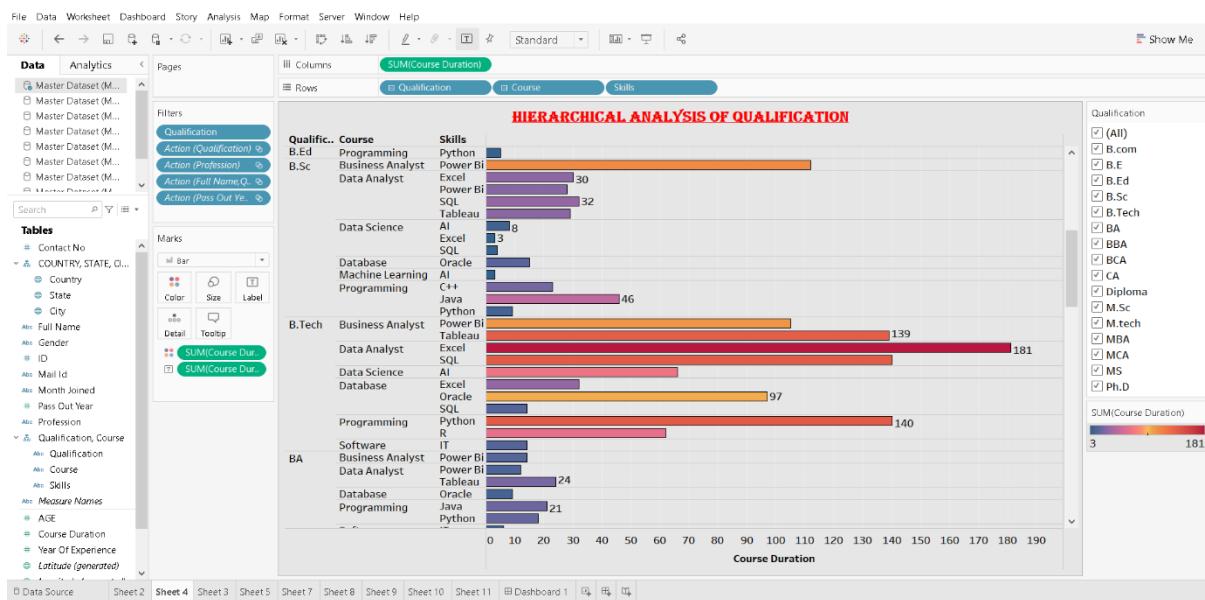


FIG. COUNT OF PROFESSION

2. Hierarchical Analysis:

Here It displays hierarchical analysis (using Drill down method) of qualification, course and skills



3. Higher Experience wise Candidate:

Here it displays that experience wise candidate name with their qualification.

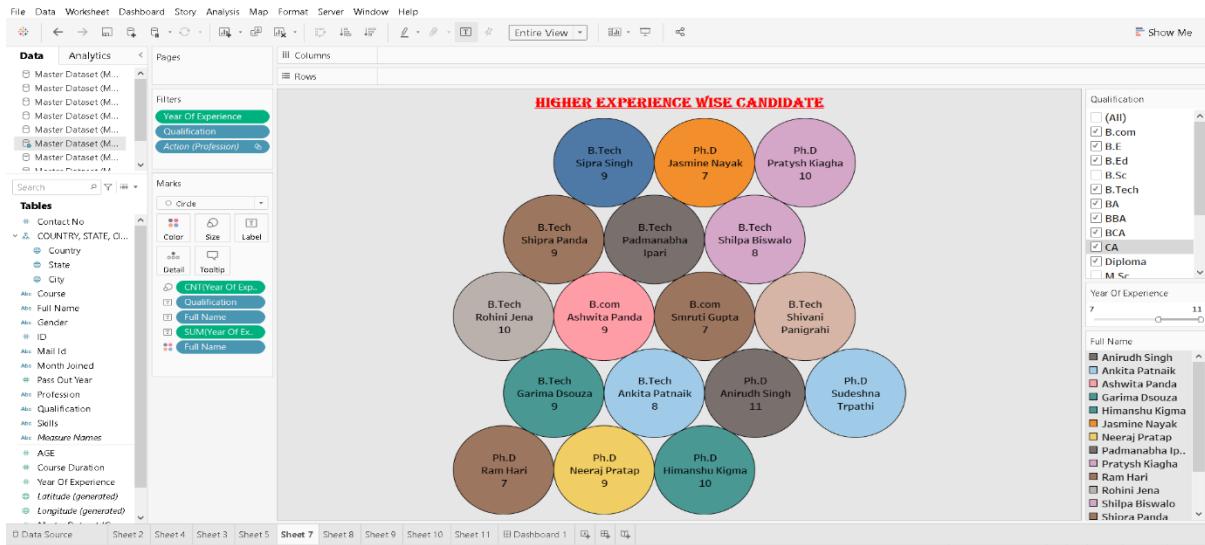


FIG.HIGHER EXPERIENCE WISE CANDIDATES

4. Gender wise profession and qualification:

It states that profession wise male and female present in qualification background.

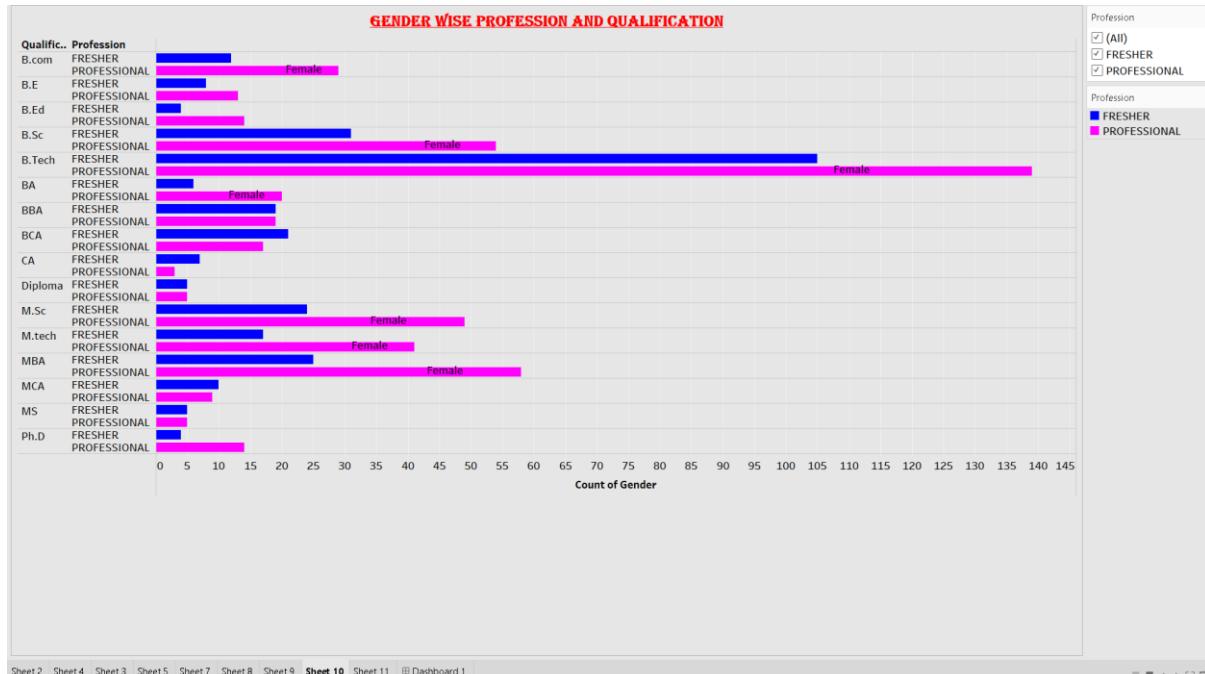


FIG. GENDER WISE PROFESSION AND QUALIFICATION

5. Pass out Year wise Qualification Background:

Now we can see that how many freshers and professional are pass out from individual qualification in respective years as shown in the figure.

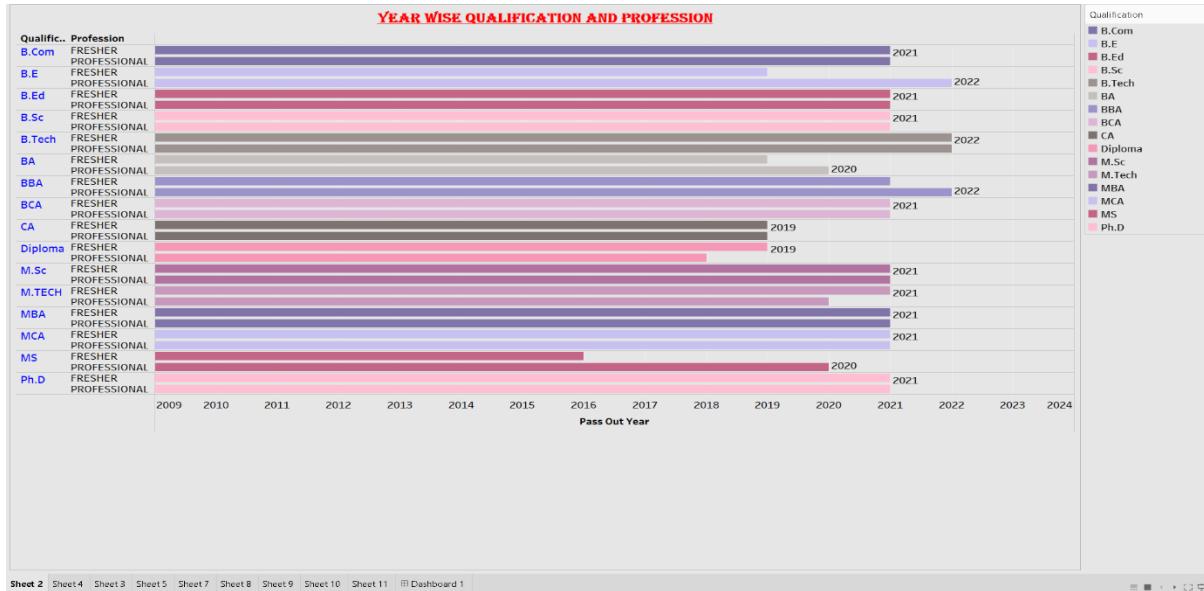


FIG.YEAR WISE QUALIFICATION BACKGROUND

6. India map on professional and skills:

This visualization sites that state wise professional joined with respective skills.

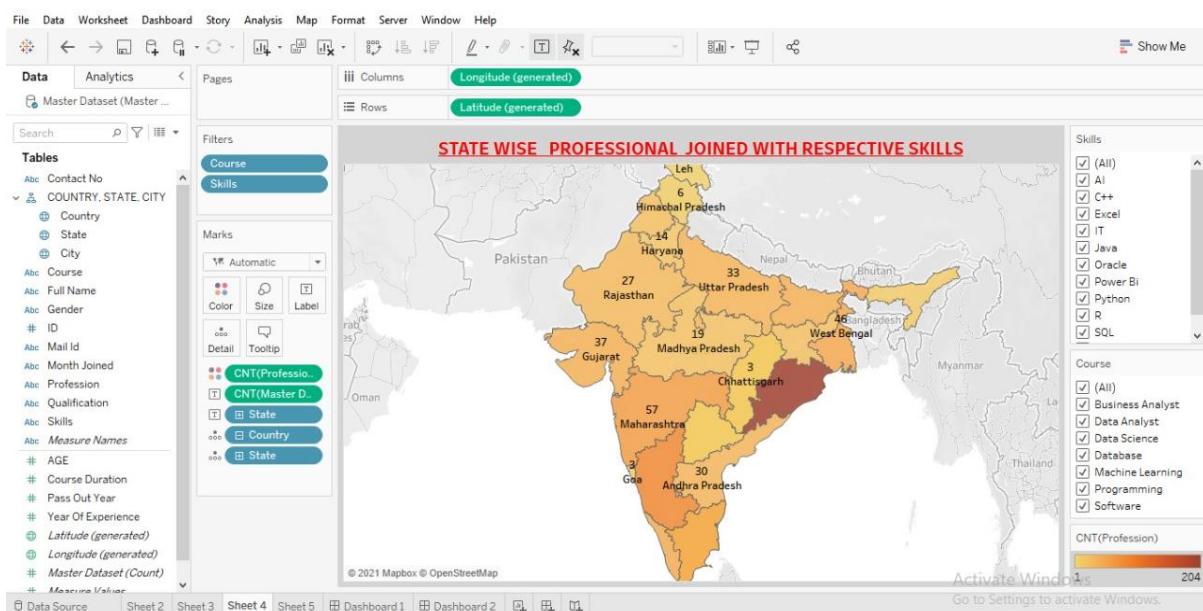


FIG. INDIA MAP ON PROFESSIONAL AND SKILLS

7. Course and Skills with Gender:

It displays that count of gender with skills and course.

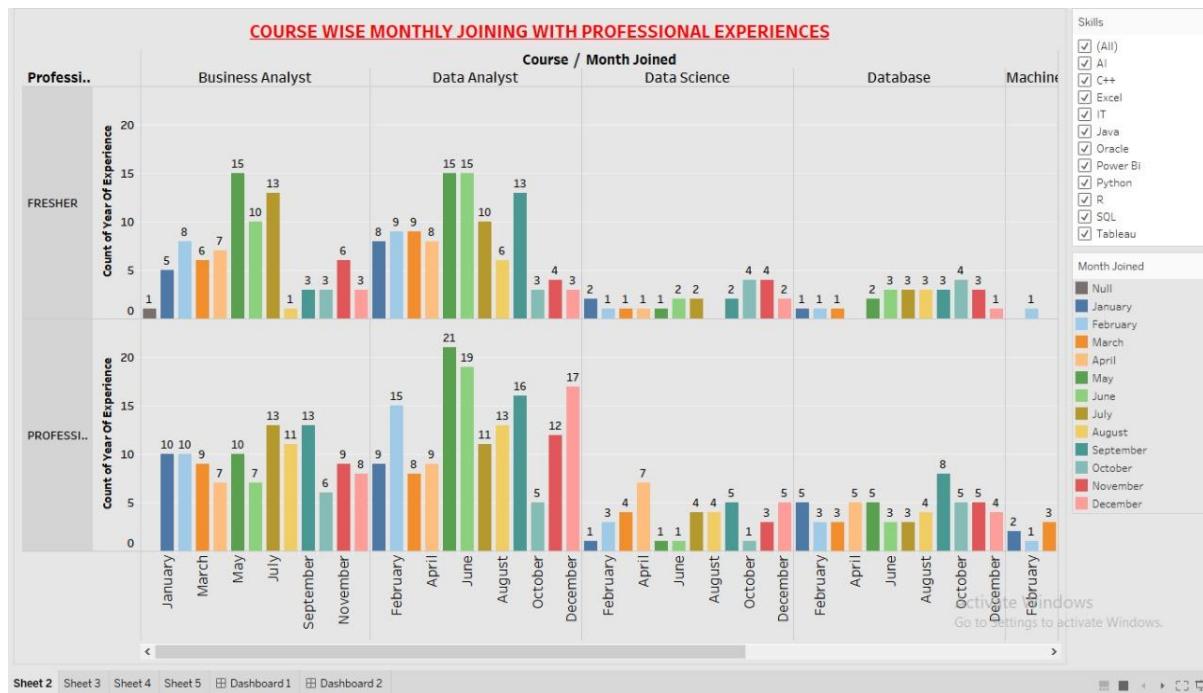


FIG. COURSE AND SKILLS GENDER WISE

8. Skill and Course wise count of Professional:

It shows how many freshers and professionals are joined with individual courses.

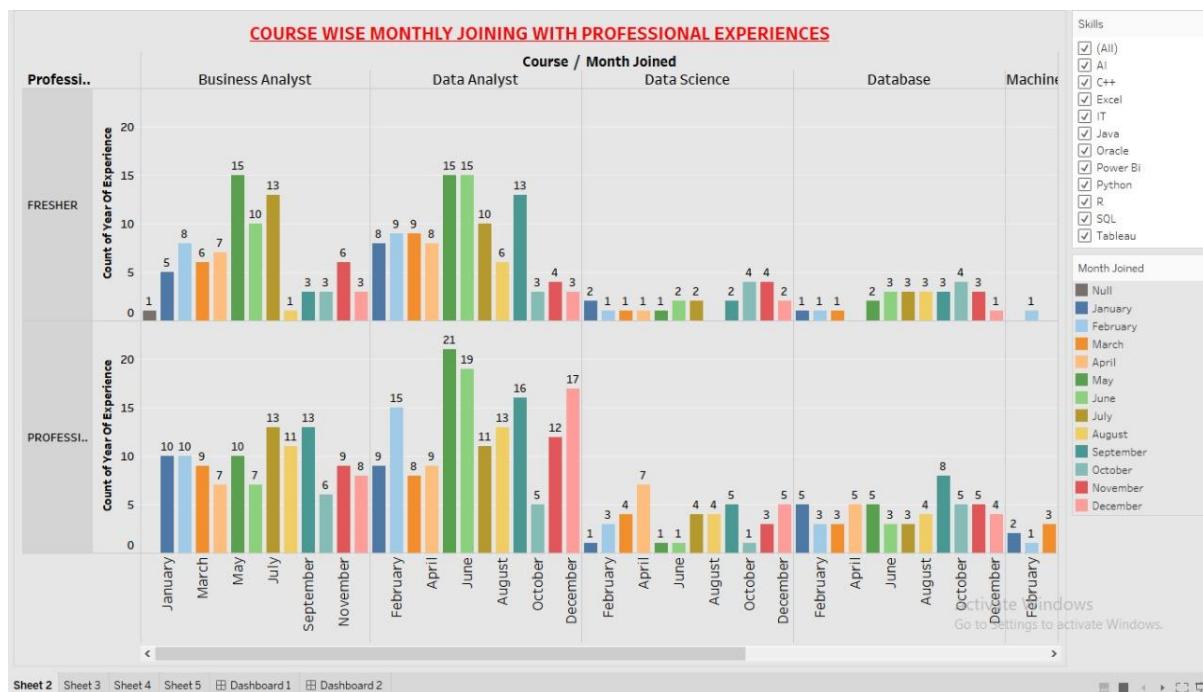


FIG. COURSE AND SKILLS PROFESSIONAL WISE

9. Skill and Course wise count of Professional:

This Tree map implies that depending on individual qualification how many skills and courses are present.



FIG. SKILL AND COURSE WISE COUNT OF PROFESSIONAL

10. Dashboard [Student Data Management for Data Gyan]:

A dashboard is a consolidated display of many worksheets and related information in a single place. It is used to compare and monitor a variety of data simultaneously. The different data views are displayed all at once. Dashboards are shown as tabs at the bottom of the workbook and they usually get updated with the most recent data from the data source. While creating a dashboard, you can add views from any worksheet in the workbook along with many supporting objects such as text areas, web pages, and images.

Each view you add to the dashboard is connected to its corresponding worksheet. So, when you modify the worksheet, the dashboard is updated and when you modify the view in the dashboard, the worksheet is updated.

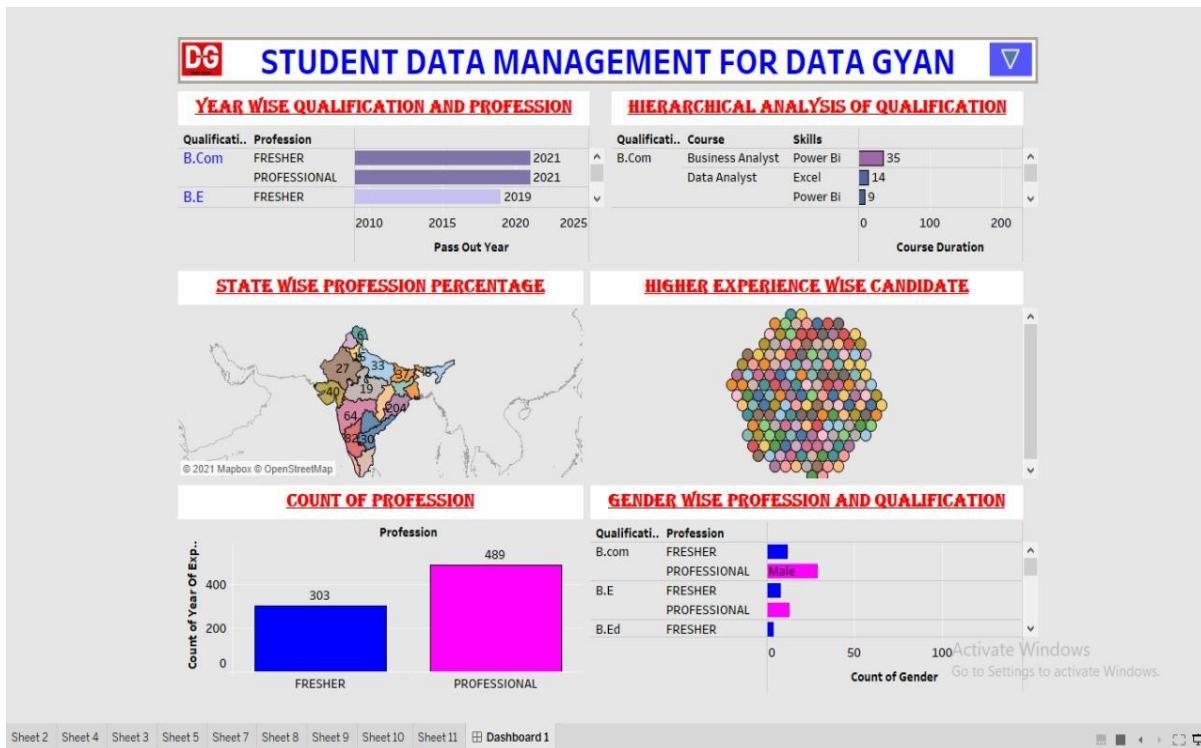


FIG. DASHBOARD 1 FOR DATA MANAGEMENT FOR DATA GYAN

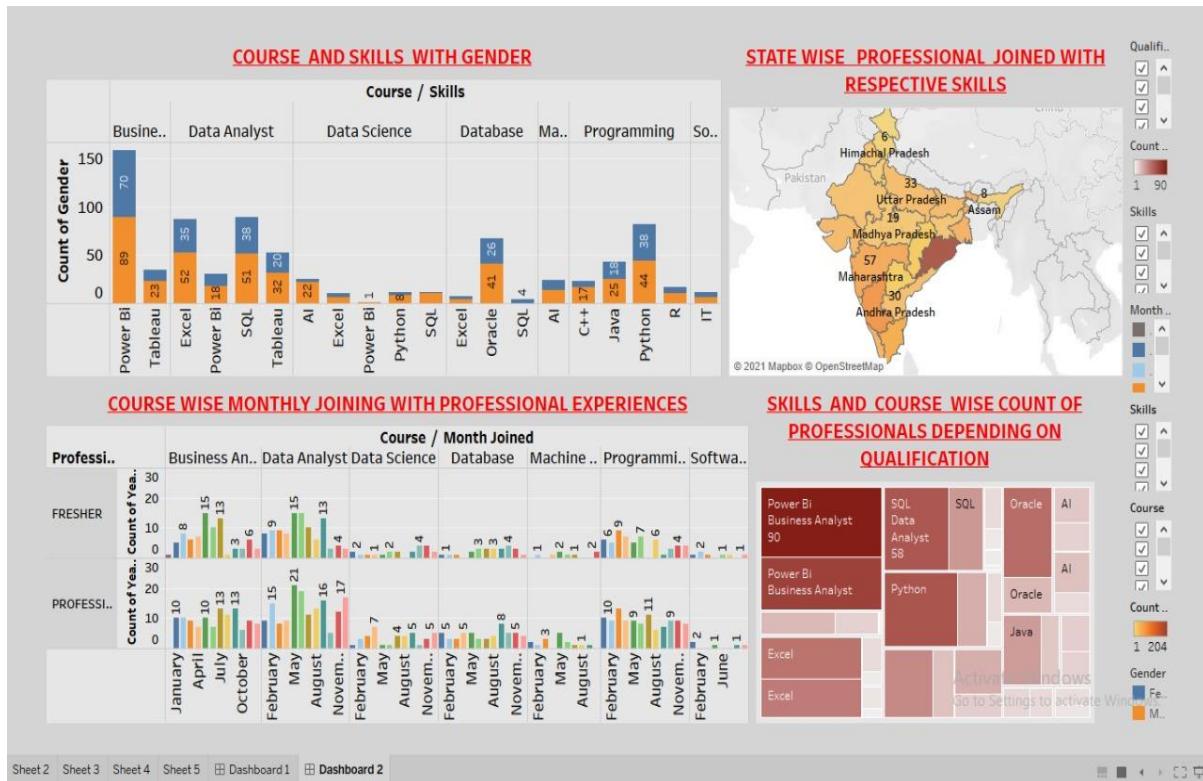


FIG. DASHBOARD 2 FOR DATA MANAGEMENT FOR DATA GYAN