Sub Assignment 3: Transforming CSV Data into a MySQL Relational Database [Level Hard]

Steps that I follow:

Step 1: After cleaning and processing data update this changes in dataset file.

import csv

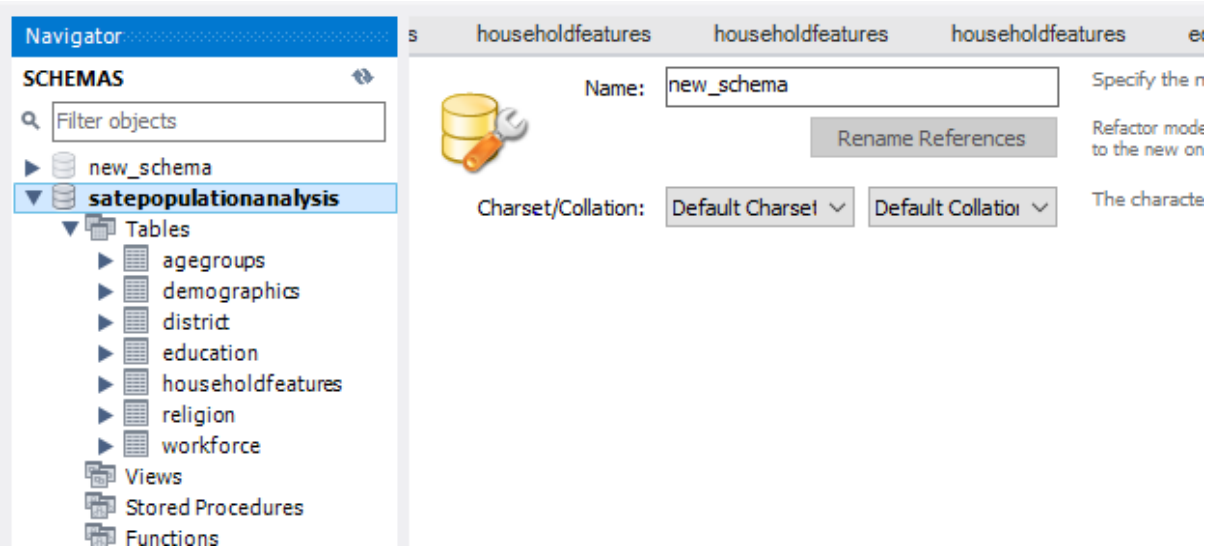
csv_file_path = 'india-state-wise-data-analysis.csv'

df.to_csv(csv_file_path, index=False)

I use this to update changes in file

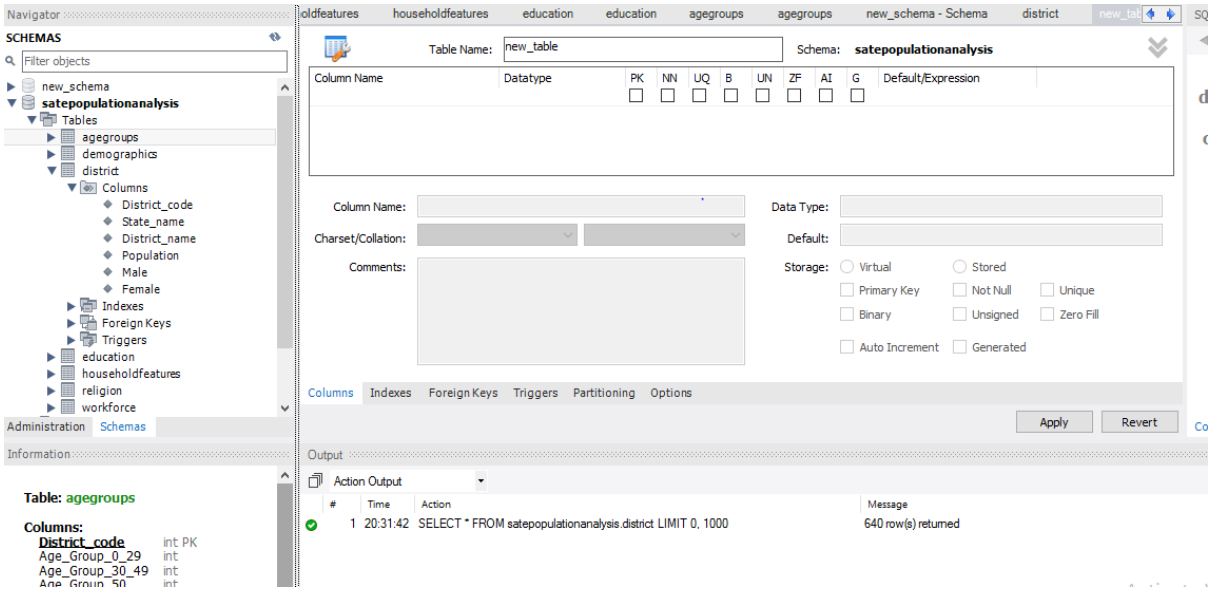
Step 2: I used MySQL Workbench for create the schema.

First I created new database-



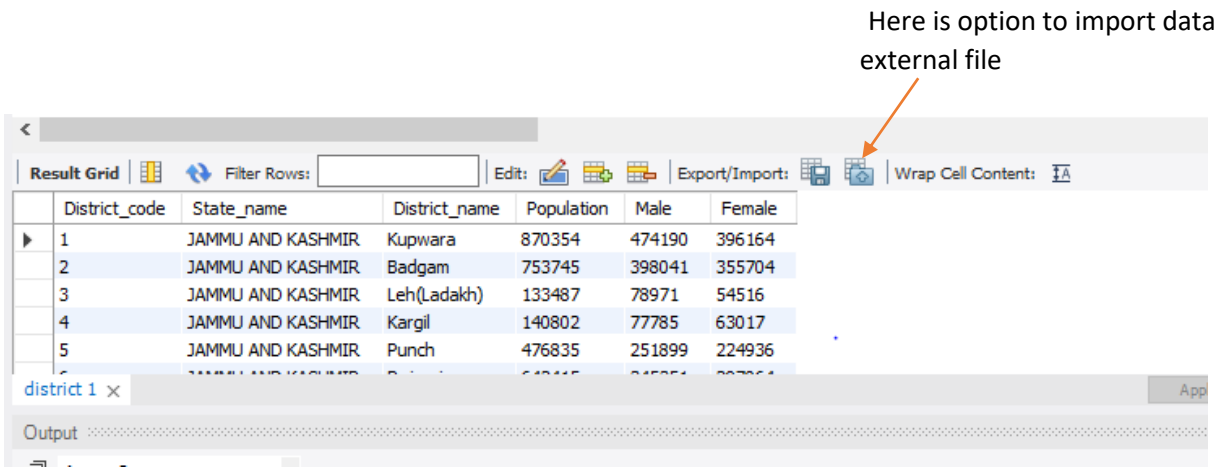
Here, **statepopulationanalysis** is my databse name

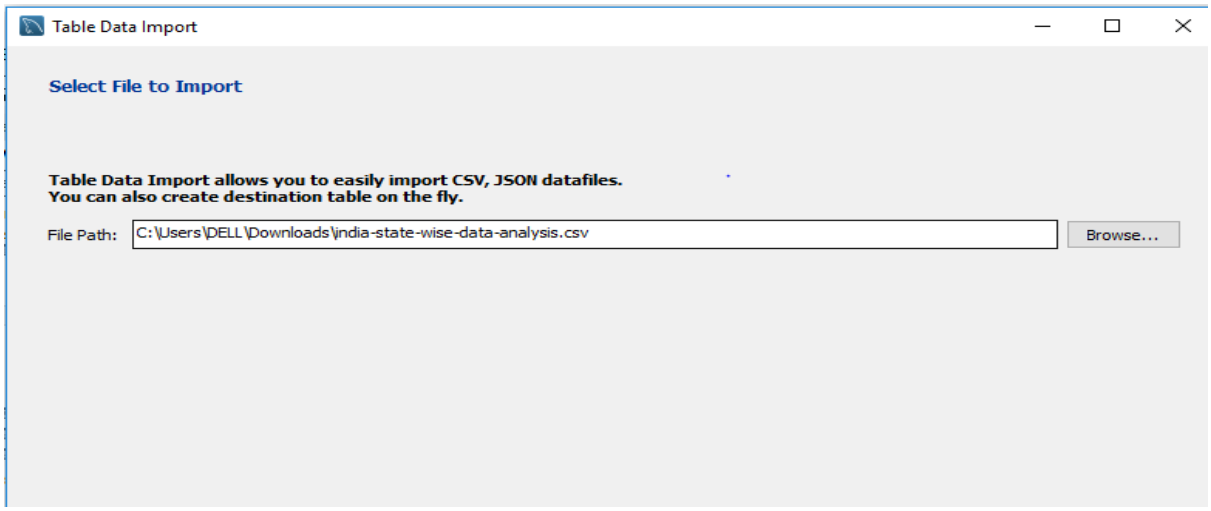
Step 3: Then I created new table ‘District’



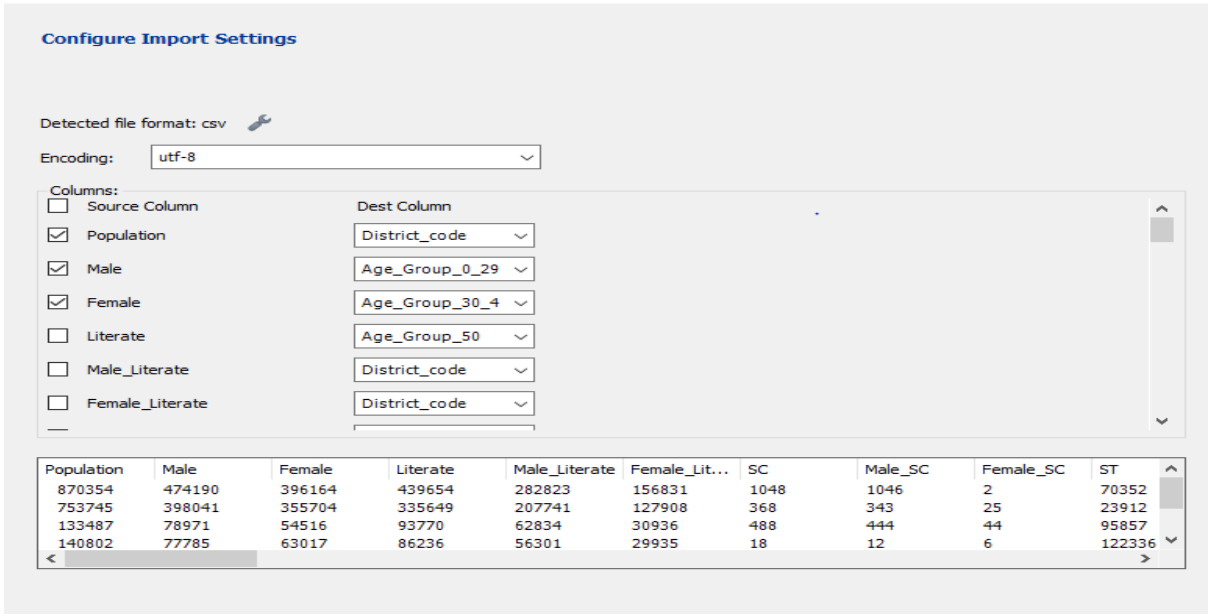
When creating table I kept name of column same as name of column in dataset ''india-state-wise-data-analysis.csv'' for ex .for district code column name is District _code in dataset so, in table it is kept as same 'District _code'

Step 4: After creating table ,I imported data from CSV file(dataset file) into table .s

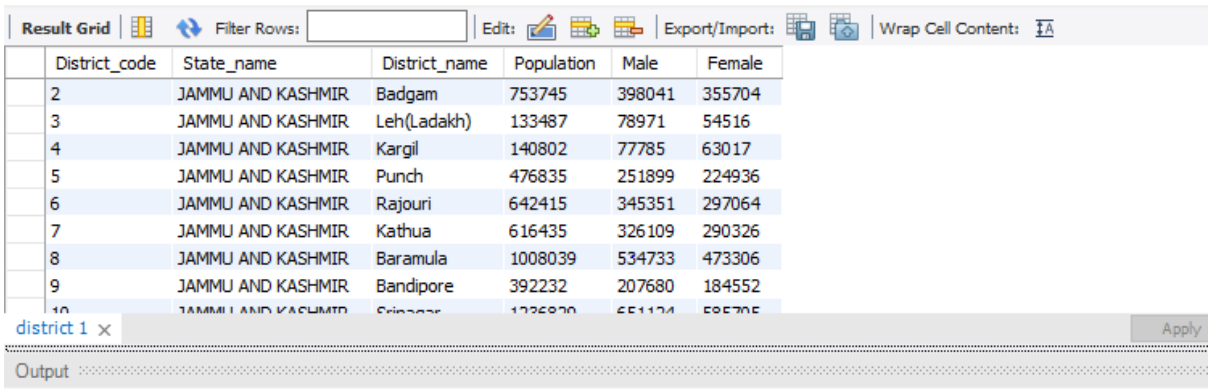




Select CSV file



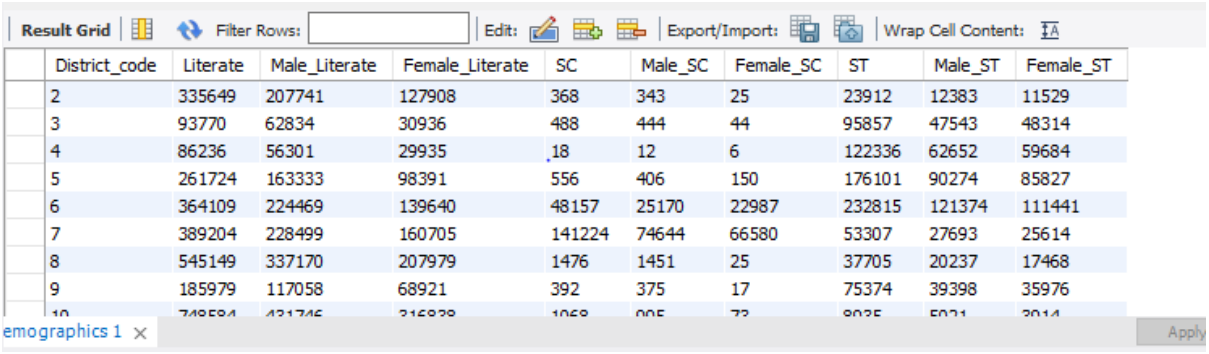
Choose columns which data we want to be inserted.



All data imported in selected fields

Stpe 5: By follow above steps I created all other tables.

Demographics table:

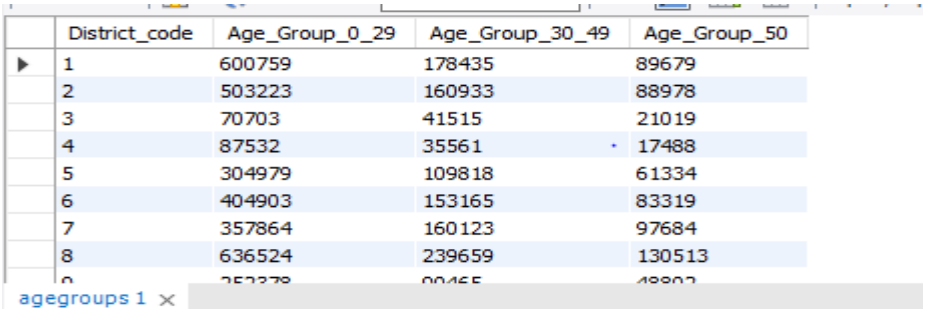


Result Grid | Filter Rows: | Edit: | Export/Import: | Wrap Cell Content:

District_code	Literate	Male_Literate	Female_Literate	SC	Male_SC	Female_SC	ST	Male_ST	Female_ST
2	335649	207741	127908	368	343	25	23912	12383	11529
3	93770	62834	30936	488	444	44	95857	47543	48314
4	86236	56301	29935	18	12	6	122336	62652	59684
5	261724	163333	98391	556	406	150	176101	90274	85827
6	364109	224469	139640	48157	25170	22987	232815	121374	111441
7	389204	228499	160705	141224	74644	66580	53307	27693	25614
8	545149	337170	207979	1476	1451	25	37705	20237	17468
9	185979	117058	68921	392	375	17	75374	39398	35976

emographics 1 | Apply

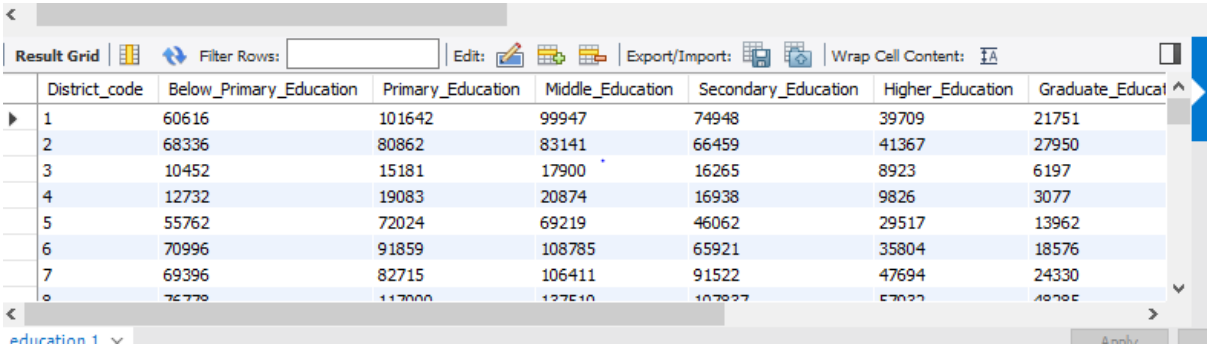
Agegroups table:



District_code	Age_Group_0_29	Age_Group_30_49	Age_Group_50
1	600759	178435	89679
2	503223	160933	88978
3	70703	41515	21019
4	87532	35561	17488
5	304979	109818	61334
6	404903	153165	83319
7	357864	160123	97684
8	636524	239659	130513

agegroups 1

Education table:

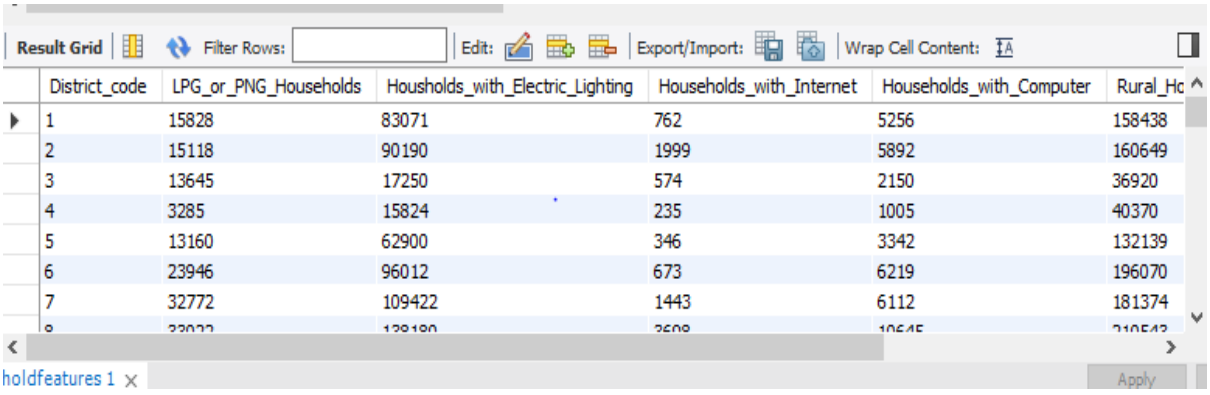


Result Grid | Filter Rows: | Edit: | Export/Import: | Wrap Cell Content:

District_code	Below_Primary_Education	Primary_Education	Middle_Education	Secondary_Education	Higher_Education	Graduate_Educat
1	60616	101642	99947	74948	39709	21751
2	68336	80862	83141	66459	41367	27950
3	10452	15181	17900	16265	8923	6197
4	12732	19083	20874	16938	9826	3077
5	55762	72024	69219	46062	29517	13962
6	70996	91859	108785	65921	35804	18576
7	69396	82715	106411	91522	47694	24330

education 1 | Apply

Householdsfeature table:

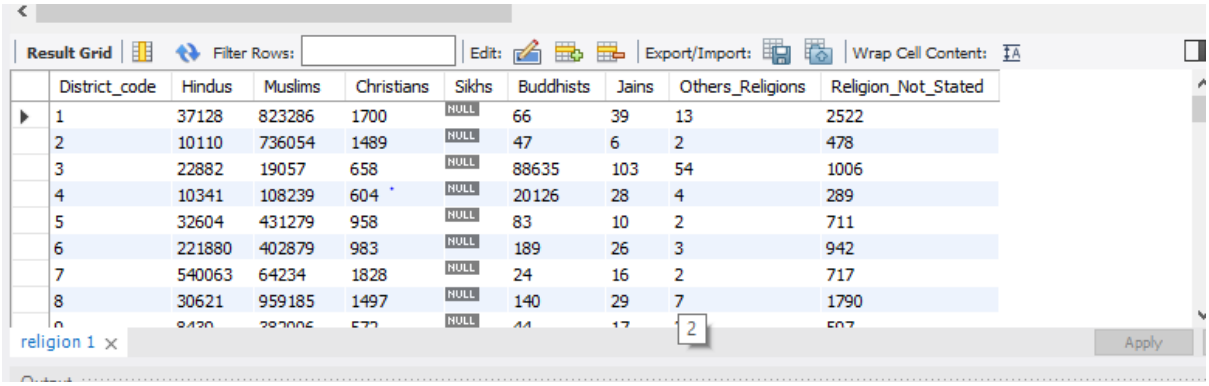


Result Grid | Filter Rows: | Edit: | Export/Import: | Wrap Cell Content:

District_code	LPG_or_PNG_Households	Housholds_with_Electric_Lighting	Households_with_Internet	Households_with_Computer	Rural_Ho
1	15828	83071	762	5256	158438
2	15118	90190	1999	5892	160649
3	13645	17250	574	2150	36920
4	3285	15824	235	1005	40370
5	13160	62900	346	3342	132139
6	23946	96012	673	6219	196070
7	32772	109422	1443	6112	181374

holdfeatures 1 | Apply

Religion table:

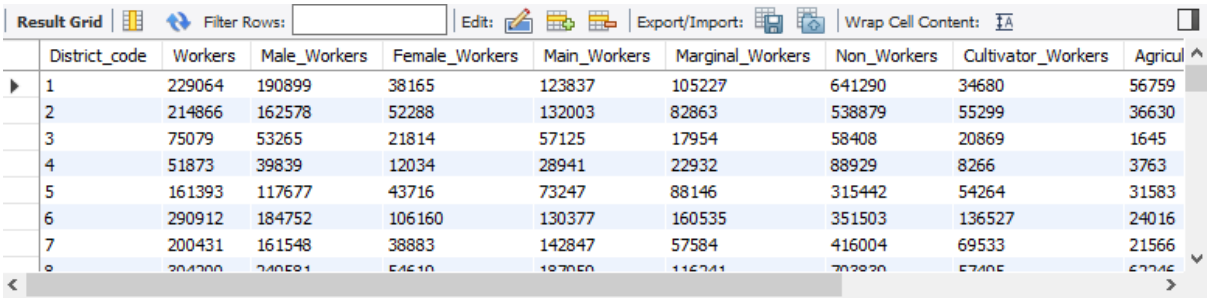


Result Grid | Filter Rows: | Edit: | Export/Import: | Wrap Cell Content:

District_code	Hindus	Muslims	Christians	Sikhs	Buddhists	Jains	Others_Religions	Religion_Not_Stated
1	37128	823286	1700	NULL	66	39	13	2522
2	10110	736054	1489	NULL	47	6	2	478
3	22882	19057	658	NULL	88635	103	54	1006
4	10341	108239	604	NULL	20126	28	4	289
5	32604	431279	958	NULL	83	10	2	711
6	221880	402879	983	NULL	189	26	3	942
7	540063	64234	1828	NULL	24	16	2	717
8	30621	959185	1497	NULL	140	29	7	1790

religion 1 | Apply

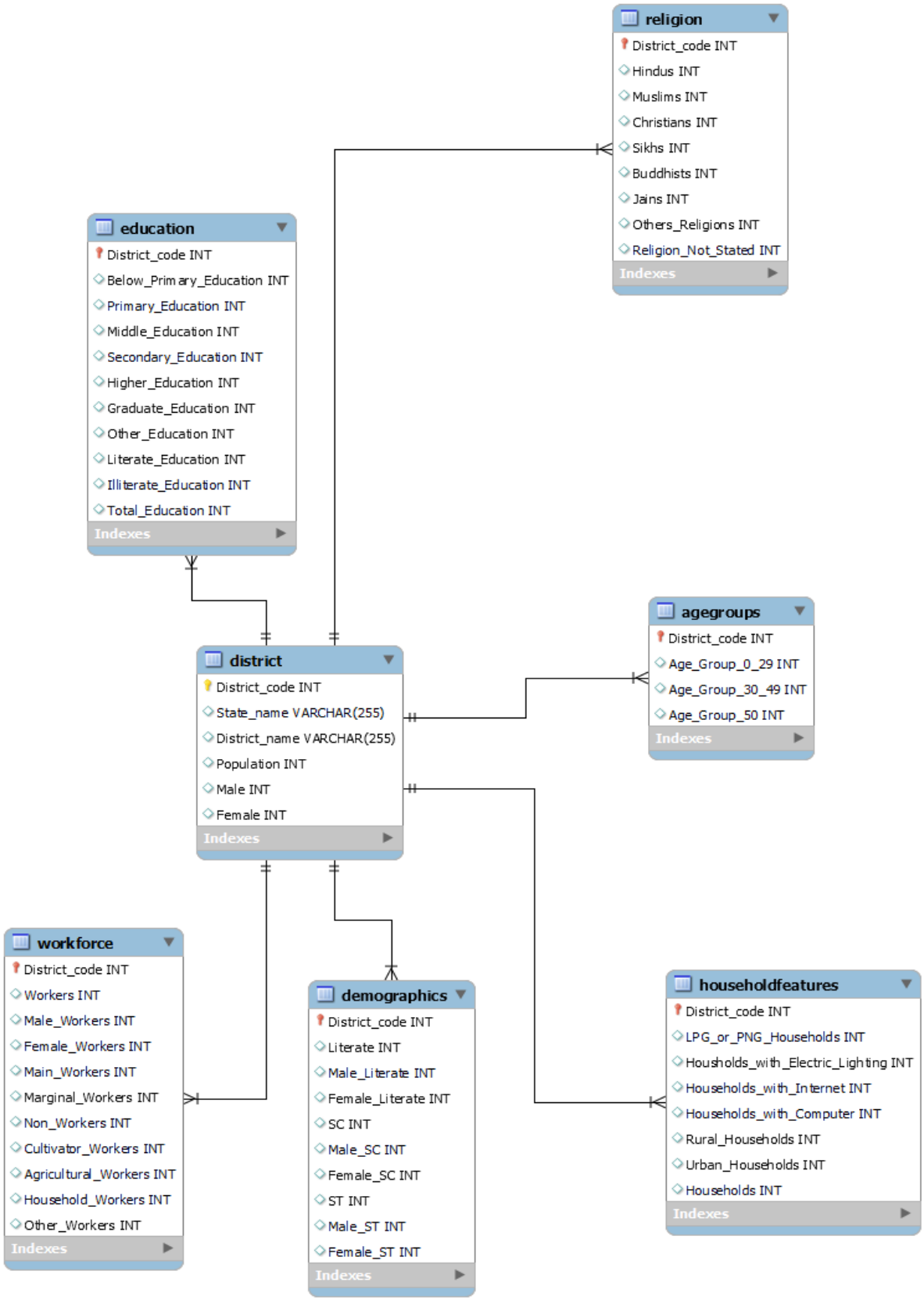
Workforce table:



Result Grid | Filter Rows: | Edit: | Export/Import: | Wrap Cell Content:

District_code	Workers	Male_Workers	Female_Workers	Main_Workers	Marginal_Workers	Non_Workers	Cultivator_Workers	Agricul
1	229064	190899	38165	123837	105227	641290	34680	56759
2	214866	162578	52288	132003	82863	538879	55299	36630
3	75079	53265	21814	57125	17954	58408	20869	1645
4	51873	39839	12034	28941	22932	88929	8266	3763
5	161393	117677	43716	73247	88146	315442	54264	31583
6	290912	184752	106160	130377	160535	351503	136527	24016
7	200431	161548	38883	142847	57584	416004	69533	21566

Step 6: After creating all table get final schema


religion
District_code INT
Hindus INT
Muslims INT
Christians INT
Sikhs INT
Buddhists INT
Jains INT
Others_Religions INT
Religion_Not_Stated INT
Indexes
education
District_code INT
Below_Primary_Education INT
Primary_Education INT
Middle_Education INT
Secondary_Education INT
Higher_Education INT
Graduate_Education INT
Other_Education INT
Literate_Education INT
Illiterate_Education INT
Total_Education INT
Indexes
agegroups
District_code INT
Age_Group_0_29 INT
Age_Group_30_49 INT
Age_Group_50 INT
Indexes
district
District_code INT
State_name VARCHAR(255)
District_name VARCHAR(255)
Population INT
Male INT
Female INT
Indexes
workforce
District_code INT
Workers INT
Male_Workers INT
Female_Workers INT
Main_Workers INT
Marginal_Workers INT
Non_Workers INT
Cultivator_Workers INT
Agricultural_Workers INT
Household_Workers INT
Other_Workers INT
Indexes
demographics
District_code INT
Literate INT
Male_Literate INT
Female_Literate INT
SC INT
Male_SC INT
Female_SC INT
ST INT
Male_ST INT
Female_ST INT
Indexes
householdfeatures
District_code INT
LPG_or_PNG_Households INT
Housholds_with_Electric_Lighting INT
Households_with_Internet INT
Households_with_Computer INT
Rural_Households INT
Urban_Households INT
Households INT
Indexes

Entity and its Attributes:

District Entity:

Atrributes: DistrictCode (Primary Key)**,** StateName **,**DistrictName **,**Population**,**Male**,**Female

Demographics Entity:

Attributes: DistrictCode (Foreign Key), Literate,MaleLiterate, FemaleLiterate, SC ,MaleSC ,FemaleSC,ST, MaleST,FemaleST

Workforce Entity:

Attributes:

DistrictCode (Foreign Key)**,** Workers **,**Male Workers **,**Female Workers**,**Main Workers**,**Marginal Workers**,**Non Workers**,**Cultivator Workers**,** Agricultural Workers **,**Household Workers**,** Other Workers

Religion Entity:

Attributes: DistrictCode (Foreign Key)**,** Hindus **,**Muslims**,** Christians**,** Sikhs**,**Buddhists**,**Jains**,** Others Religions**,**Religion NotStated

Household Features Entity:

Attributes: DistrictCode (Foreign Key) **,**LPGorPNGHouseholds **,**HouseholdsElectricLighting**,**HouseholdsInternet**,**HouseholdsComputer**,**RuralHouseholds**,**UrbanHouseholds HouseholdsTotal

Education Entity

Attributes: DistrictCode (Foreign Key)BelowPrimaryEducation, PrimaryEducation, MiddleEducation, SecondaryEducation**,** HigherEducation **,**GraduateEducation**,** OtherEducation**,**LiterateEducation**,**IlliterateEducation**,**TotalEducation

Age Groups Entity

Attributes: DistrictCode (Foreign Key), AgeGroup0_29, AgeGroup30_49, AgeGroup50