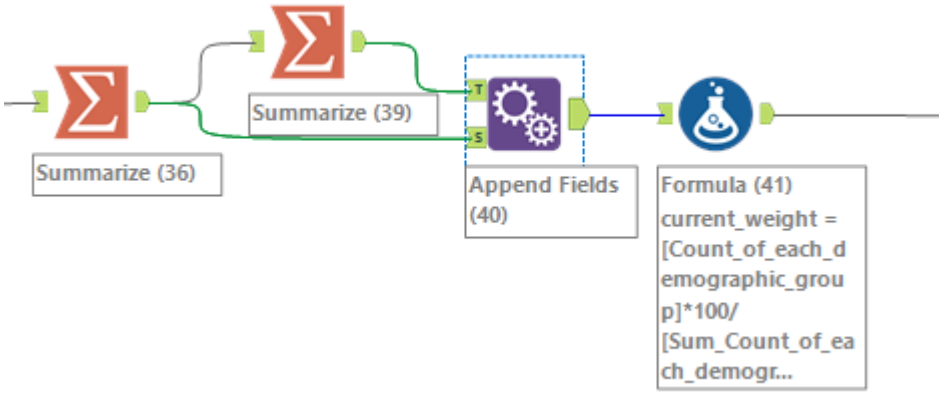
This is the screenshot of the initial survey data which has 786 respondents from 53 different cities of Turkey.

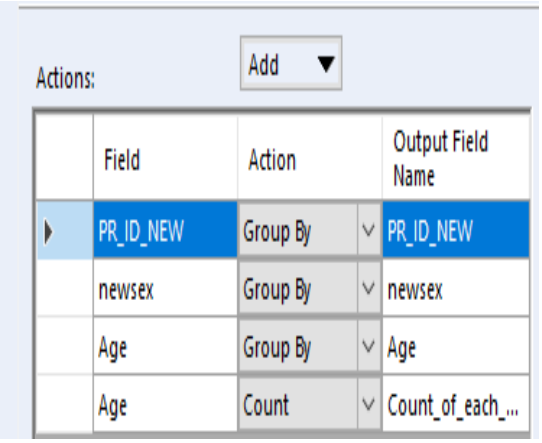
21 of 21 Fields · Cell Viewer · 786 records displayed · Search · Data · Metadata

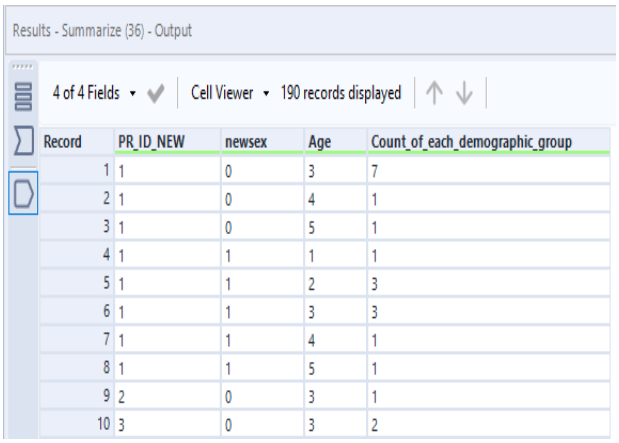
Record	ResponID_new	IL_K...	Province	FoSyrianCrime	FoGeneralCrime	PR_ID_NEW	newsex	Age	Syria...	SyrianEncounterCity	Education	Religiosity	married	ethnicdummy	unemployed	secrisksqrd	relationalDiscomfort	econoconcernsqrd
1	1	6	Ankara	2.349235	2.61136	5	0	4	3	2	6	2	1	1	1	17.398495	2.48653	20.141018
2	2	34	İstanbul	4.957164	4.172864	25	0	5	4	5	4	3	1	1	0	23.658995	2.855528	11.314594
3	3	27	Gaziantep	4.957948	5.216079	20	1	3	5	5	4	3	1	1	0	23.689255	4.273646	20.141018
4	4	34	İstanbul	5.218856	5.216079	25	0	5	2	4	4	3	0	1	0	27.185146	3.912609	20.141018
5	5	34	İstanbul	2.348451	1.561127	25	0	2	4	4	4	3	0	1	0	27.185146	3.912609	31.470337
6	6	54	Sakarya	1.82585	1.561127	41	0	4	3	4	6	2	1	1	0	12.076593	2.48653	20.124542
7	7	48	Muğla	2.609671	1.305318	37	0	4	3	3	3	3	1	1	0	27.185146	3.915978	31.470337
8	8	6	Ankara	5.218856	5.216079	5	1	3	4	4	3	2	1	1	1	27.185146	4.273646	31.470337
9	9	34	İstanbul	5.218856	5.216079	25	1	3	5	5	3	2	1	1	1	27.185146	3.554941	31.470337
10	10	34	İstanbul	2.60936	4.953254	25	1	4	4	4	5	3	1	1	0	27.185146	4.273646	16.948065
11	11	53	Rize	2.346063	1.043216	40	0	4	2	2	4	2	0	1	1	27.185146	3.554941	23.630629
12	12	16	Bursa	4.696256	4.172864	11	1	4	2	5	4	2	1	1	0	27.185146	4.273646	31.470337
13	13	17	Çanakkale	5.218856	4.428672	12	0	4	4	4	4	3	1	0	1	27.185146	4.273646	31.470337
14	15	34	İstanbul	5.218856	5.216079	25	0	5	2	4	4	3	1	1	0	27.185146	3.912609	31.470337
15	16	7	Antalya	2.868838	4.172864	6	1	4	2	2	4	2	1	1	0	27.185146	3.915978	31.470337
16	17	34	İstanbul	5.218856	2.086432	25	0	4	4	4	6	1	1	0	0	27.185146	4.273646	16.936445
17	18	34	İstanbul	3.129572	2.61136	25	1	4	4	4	4	1	0	1	0	27.185146	3.92394	31.470337
18	19	6	Ankara	3.65264	4.690774	5	1	4	5	5	4	2	1	1	0	27.185146	4.273646	31.470337
19	20	38	Kayseri	5.218856	5.216079	28	0	5	5	5	1	2	1	1	0	1.087406	3.574233	1.258814
20	21	38	Kayseri	3.650587	1.299025	28	0	2	2	4	4	2	1	0	0	27.185146	3.554941	31.470337
21	22	6	Ankara	3.389678	2.591033	5	1	4	5	5	3	3	1	1	1	27.185146	3.205234	23.630629
22	23	6	Ankara	5.218856	5.216079	5	1	4	5	5	6	2	0	0	0	27.185146	4.273646	31.470337
23	24	34	İstanbul	2.60936	2.086432	25	1	4	3	5	3	2	1	1	1	27.185146	4.273646	31.470337
24	25	35	İzmir	4.6961	4.172864	26	1	2	4	4	4	3	0	0	1	27.185146	3.915978	31.470337
25	26	6	Ankara	2.087542	4.172864	5	0	3	1	4	4	3	1	1	0	7.70266	1.068412	5.025438
26	27	1	Adana	4.95875	2.328554	1	1	4	2	5	4	2	0	0	1	27.185146	3.224526	16.948065
27	28	34	İstanbul	5.218856	4.428672	25	1	4	3	4	4	3	1	1	0	27.185146	3.205234	27.40258
28	29	34	İstanbul	4.175085	4.435342	25	1	2	4	4	4	1	1	1	0	20.411665	1.767825	31.470337
29	30	35	İzmir	2.871208	1.043216	26	0	4	3	3	4	2	1	0	0	20.411665	2.483161	31.470337

First, we created our combination groups with the demographic variables we would like to weight. They are respectively; 1) population of the city 2) gender and 3) five different age groups. We need to get the number of the respondents in each combination of the three demographic groups. The first summarize tool below in the Alteryx workflow returns the number of respondents in each combination of the demographic groups.



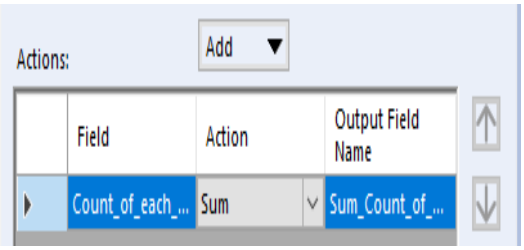
Below is the configuration of the first summarize tool and the output.

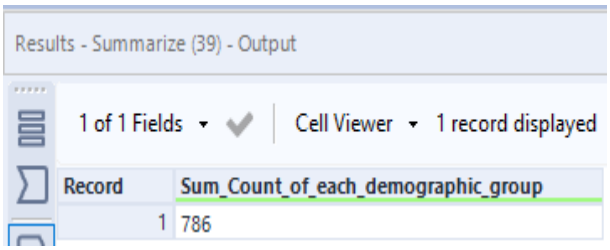




Record	PR_ID_NEW	newsex	Age	Count_of_each_demographic_group
1	1	0	3	7
2	1	0	4	1
3	1	0	5	1
4	1	1	1	1
5	1	1	2	3
6	1	1	3	3
7	1	1	4	1
8	1	1	5	1
9	2	0	3	1
10	3	0	3	2

Then we need to get the total number of the respondents in our sample (sample size=786). Second Summarize tool above in the Alteryx workflow does that and below is the configuration of the workflow and the output.

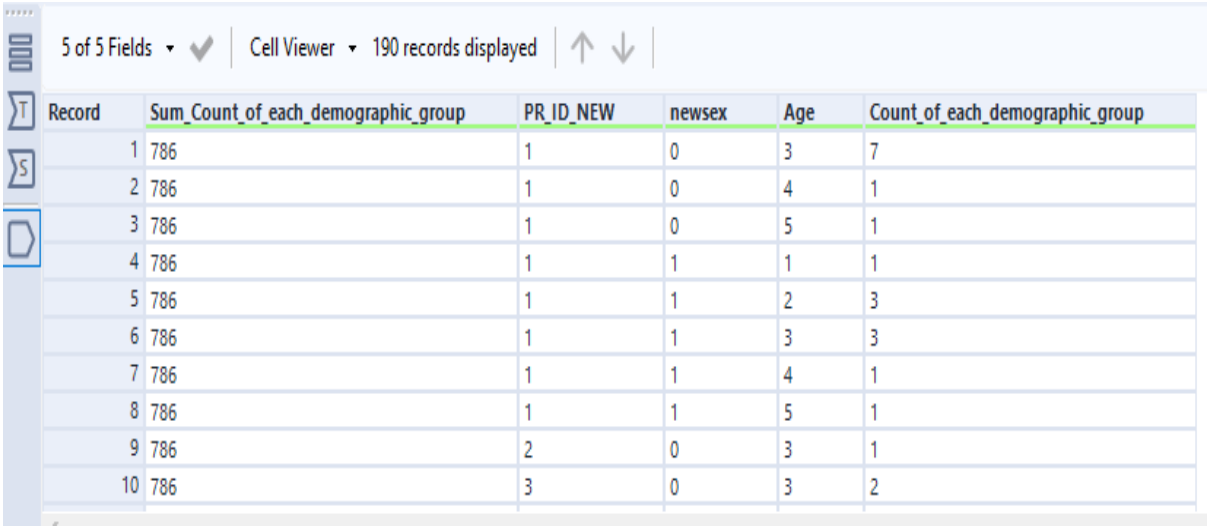




Record	Sum_Count_of_each_demographic_group
1	786

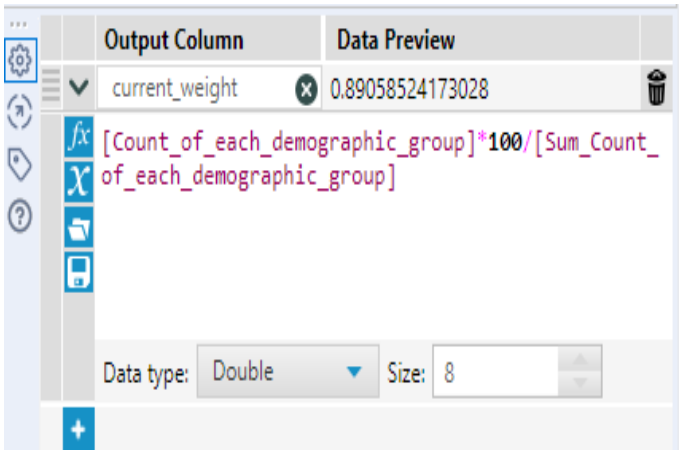
Then

With the Append tool in the Alteryx workflow above, we joined the two tables. Below is the output.

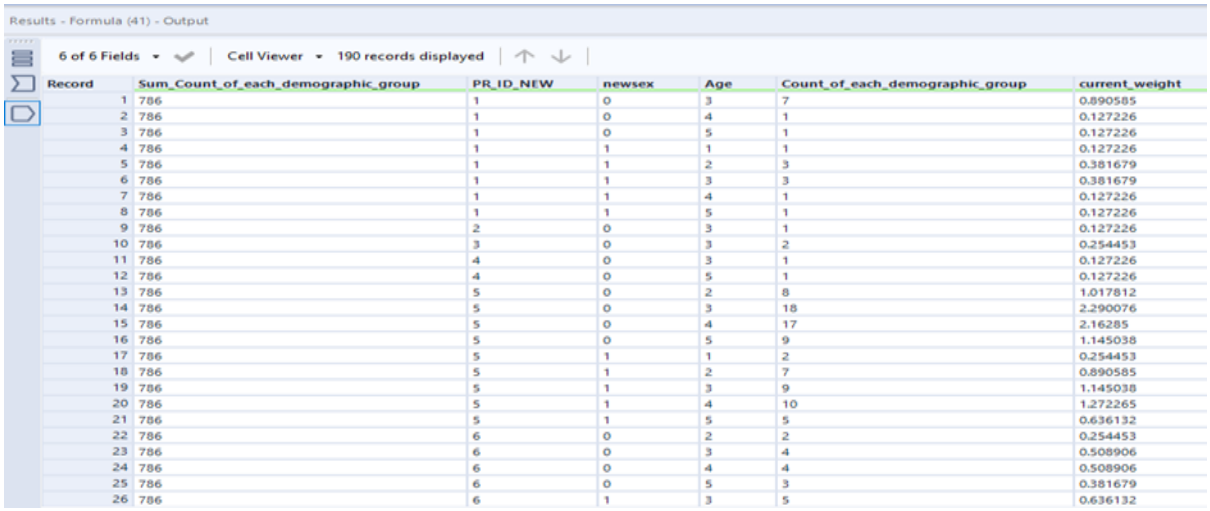


Record	Sum_Count_of_each_demographic_group	PR_ID_NEW	newsex	Age	Count_of_each_demographic_group
1	786	1	0	3	7
2	786	1	0	4	1
3	786	1	0	5	1
4	786	1	1	1	1
5	786	1	1	2	3
6	786	1	1	3	3
7	786	1	1	4	1
8	786	1	1	5	1
9	786	2	0	3	1
10	786	3	0	3	2

Now we can calculate the proportion of each combination in our sample by dividing the count number of respondents in each row (combination) by the total number of the respondents (786) and multiply it with 100. E. g. 7/786*100 = 0.8905



Now we got the current weight measure which corresponds to the proportion of each combination of the demographic groups in our sample.



Record	Sum_Count_of_each_demographic_group	PR_ID_NEW	newsex	Age	Count_of_each_demographic_group	current_weight
1	786	1	0	3	7	0.890585
2	786	1	0	4	1	0.127226
3	786	1	0	5	1	0.127226
4	786	1	1	1	1	0.127226
5	786	1	1	2	3	0.381679
6	786	1	1	3	3	0.381679
7	786	1	1	4	1	0.127226
8	786	1	1	5	1	0.127226
9	786	2	0	3	1	0.127226
10	786	3	0	3	2	0.254453
11	786	4	0	3	1	0.127226
12	786	4	0	5	1	0.127226
13	786	5	0	2	8	1.017812
14	786	5	0	3	18	2.290076
15	786	5	0	4	17	2.16285
16	786	5	0	5	9	1.145038
17	786	5	1	1	2	0.254453
18	786	5	1	2	7	0.890585
19	786	5	1	3	9	1.145038
20	786	5	1	4	10	1.272265
21	786	5	1	5	5	0.636132
22	786	6	0	2	2	0.254453
23	786	6	0	3	4	0.508906
24	786	6	0	4	4	0.508906
25	786	6	0	5	3	0.381679
26	786	6	1	3	5	0.636132

Note that, in Adana province (PR_ID_NEW = 1), the number of the males (newsex = 0) in the age group 3 (Age =3) is 7 and the representation of this category in our sample is % 0.89. In Adana (PR_ID_NEW = 1), the number of the females (newsex = 1) in age group 3 (Age =3) is 3 and the representation for this group in the sample is % 0.38.

As you may notice in our sample, we do not have respondents for all combinations of the demographic groups. We have respondents from 53 cities of Turkey out of 81. That means if we had respondents from all the demographic groups, we would have 530 different combinations of demographic groups. (population of the cities-53 * categories of gender -2 * age groups-5 =530). However, as you can see, we do not have respondents from all but 190 groups.

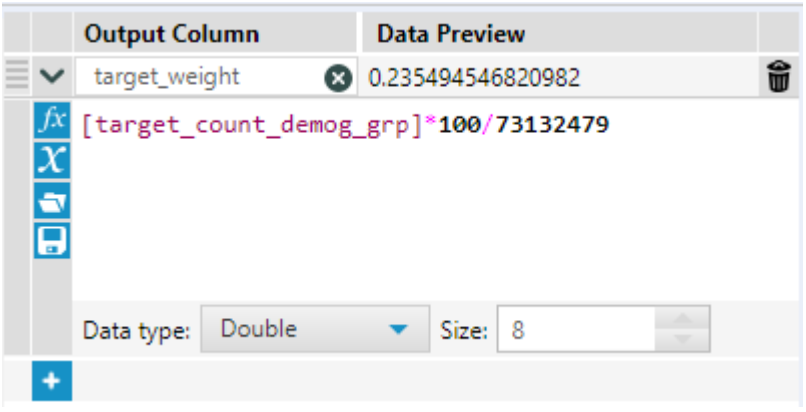
For example, in Adana (PR_ID_NEW = 1) there is no male respondent in the age group 1. That is what makes the calculation of the survey weight more complicated in this case because before calculating the survey_weight, we need to create the same combination of the demographic groups for our target population and obtain the relevant statistics for each, calculate the proportion of each combination in the overall population which we will call target_weight.

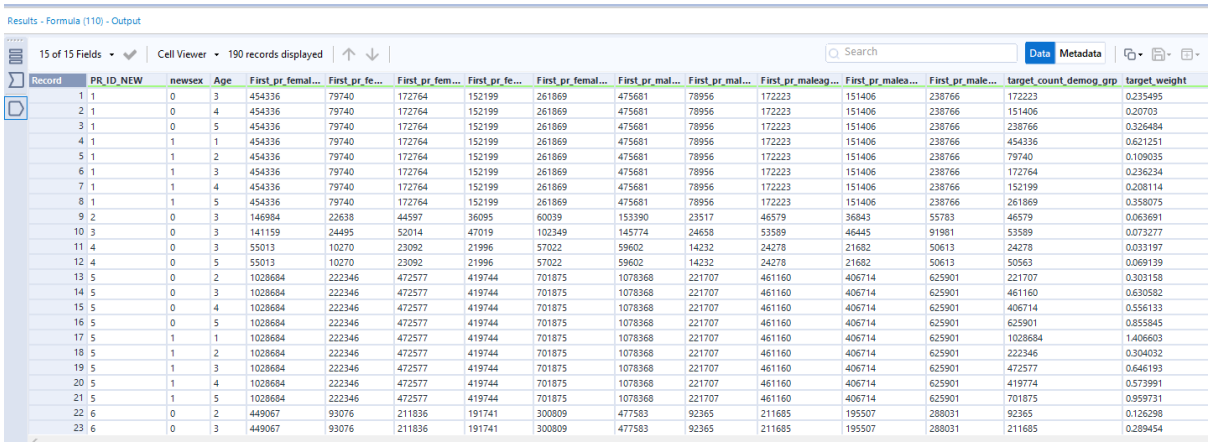
Here is how we calculated the target_weight. At this stage, we need the data for the same combinations of groups in our target population. Below is the data we obtained from the Turkish Statistical Institute, TUIK. Based on this; we can find out the male and female population in each age group in each city of Turkey. Now all we need to do is to create a new dataset from the overall population matching the same combinations of demographic groups in our sample.

13 of 13 Fields | Cell Viewer | 190 records displayed | Search | Data | Metadata

Rec...	PR_ID_NEW	newsex	Age	First_pr_femaleage1	First_pr_femaleage2	First_pr_femaleage3	First_pr_femaleage4	First_pr_femaleage5	First_pr_maleage1	First_pr_maleage2	First_pr_maleage3	First_pr_maleage4	First_pr_maleage5
1	1	0	3	454336	79740	172764	152199	261869	475681	78956	172223	151406	238766
2	1	0	4	454336	79740	172764	152199	261869	475681	78956	172223	151406	238766
3	1	0	5	454336	79740	172764	152199	261869	475681	78956	172223	151406	238766
4	1	1	1	454336	79740	172764	152199	261869	475681	78956	172223	151406	238766
5	1	1	2	454336	79740	172764	152199	261869	475681	78956	172223	151406	238766
6	1	1	3	454336	79740	172764	152199	261869	475681	78956	172223	151406	238766
7	1	1	4	454336	79740	172764	152199	261869	475681	78956	172223	151406	238766
8	1	1	5	454336	79740	172764	152199	261869	475681	78956	172223	151406	238766
9	2	0	3	146984	22638	44597	36095	60039	153390	23517	46579	36843	55783
10	3	0	3	141159	24495	52014	47019	102349	145774	24658	53589	46445	91981
11	4	0	3	55013	10270	23092	21996	57022	59602	14232	24278	21682	50613
12	4	0	5	55013	10270	23092	21996	57022	59602	14232	24278	21682	50613
13	5	0	2	1028684	222346	472577	419744	701875	1078368	221707	461160	406714	625901
14	5	0	3	1028684	222346	472577	419744	701875	1078368	221707	461160	406714	625901
15	5	0	4	1028684	222346	472577	419744	701875	1078368	221707	461160	406714	625901
16	5	0	5	1028684	222346	472577	419744	701875	1078368	221707	461160	406714	625901
17	5	1	1	1028684	222346	472577	419744	701875	1078368	221707	461160	406714	625901
18	5	1	2	1028684	222346	472577	419744	701875	1078368	221707	461160	406714	625901
19	5	1	3	1028684	222346	472577	419744	701875	1078368	221707	461160	406714	625901
20	5	1	4	1028684	222346	472577	419744	701875	1078368	221707	461160	406714	625901
21	5	1	5	1028684	222346	472577	419744	701875	1078368	221707	461160	406714	625901
22	6	0	2	449067	93076	211836	191741	300809	477583	92365	211685	195507	288031
23	6	0	3	449067	93076	211836	191741	300809	477583	92365	211685	195507	288031
24	6	0	4	449067	93076	211836	191741	300809	477583	92365	211685	195507	288031
25	6	0	5	449067	93076	211836	191741	300809	477583	92365	211685	195507	288031
26	6	1	3	449067	93076	211836	191741	300809	477583	92365	211685	195507	288031
27	6	1	4	449067	93076	211836	191741	300809	477583	92365	211685	195507	288031
28	6	1	5	449067	93076	211836	191741	300809	477583	92365	211685	195507	288031
29	7	0	3	185135	36186	79191	77017	180817	192945	37512	80462	76748	164959
30	7	0	4	185135	36186	79191	77017	180817	192945	37512	80462	76748	164959
31	7	1	3	185135	36186	79191	77017	180817	192945	37512	80462	76748	164959
32	7	1	4	185135	36186	79191	77017	180817	192945	37512	80462	76748	164959
33	8	0	2	182515	36508	84492	88851	222780	198010	39734	87020	88579	200131
34	8	0	3	182515	36508	84492	88851	222780	198010	39734	87020	88579	200131

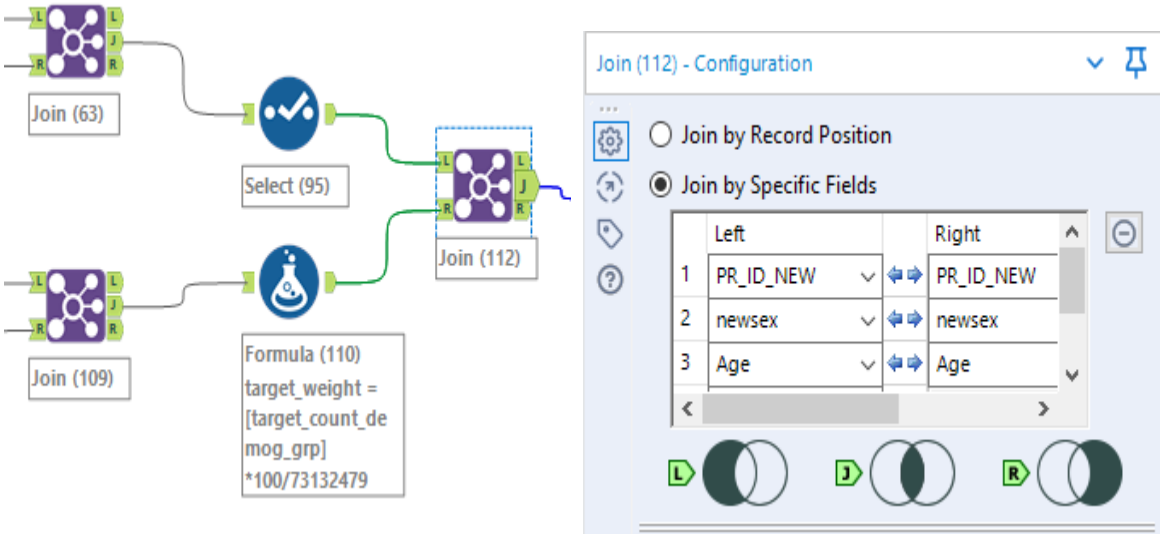
It required some tedious work to obtain each statistics but afterwards applying the same procedures we easily calculated the target weight. We divided the number of the individuals in each demographic group by the overall population of the 53 cities of Turkey. If we had respondents from all the cities of Turkey this figure would be 81, but we could not. Below is the configuration of the formula tool in Alteryx and the output.





Record	PR_ID_NEW	newsex	Age	First_pr_femal...	First_pr_fe...	First_pr_fem...	First_pr_fe...	First_pr_femal...	First_pr_mal...	First_pr_mal...	First_pr_maleag...	First_pr_malea...	First_pr_male...	target_count_demog_grp	target_weight
1	1	0	3	454336	79740	172764	152199	261869	475681	78956	172223	151406	238766	172223	0.235495
2	1	0	4	454336	79740	172764	152199	261869	475681	78956	172223	151406	238766	151406	0.20703
3	1	0	5	454336	79740	172764	152199	261869	475681	78956	172223	151406	238766	238766	0.326484
4	1	1	1	454336	79740	172764	152199	261869	475681	78956	172223	151406	238766	454336	0.621251
5	1	1	2	454336	79740	172764	152199	261869	475681	78956	172223	151406	238766	79740	0.109035
6	1	1	3	454336	79740	172764	152199	261869	475681	78956	172223	151406	238766	172764	0.236234
7	1	1	4	454336	79740	172764	152199	261869	475681	78956	172223	151406	238766	152199	0.208114
8	1	1	5	454336	79740	172764	152199	261869	475681	78956	172223	151406	238766	261869	0.358075
9	2	0	3	146984	22638	44597	36095	60039	153390	23517	46579	36843	55783	46579	0.063691
10	3	0	3	141159	24495	52014	47019	102349	145774	24658	53589	46445	91981	53589	0.073277
11	4	0	3	55013	10270	23092	21996	57022	59602	14232	24278	21682	50613	24278	0.033197
12	4	0	5	55013	10270	23092	21996	57022	59602	14232	24278	21682	50613	50563	0.069139
13	5	0	2	1028684	222346	472577	419744	701875	1078368	221707	461160	406714	625901	221707	0.303158
14	5	0	3	1028684	222346	472577	419744	701875	1078368	221707	461160	406714	625901	461160	0.630582
15	5	0	4	1028684	222346	472577	419744	701875	1078368	221707	461160	406714	625901	406714	0.556133
16	5	0	5	1028684	222346	472577	419744	701875	1078368	221707	461160	406714	625901	625901	0.855845
17	5	1	1	1028684	222346	472577	419744	701875	1078368	221707	461160	406714	625901	1028684	1.406603
18	5	1	2	1028684	222346	472577	419744	701875	1078368	221707	461160	406714	625901	222346	0.304032
19	5	1	3	1028684	222346	472577	419744	701875	1078368	221707	461160	406714	625901	472577	0.646193
20	5	1	4	1028684	222346	472577	419744	701875	1078368	221707	461160	406714	625901	419774	0.573991
21	5	1	5	1028684	222346	472577	419744	701875	1078368	221707	461160	406714	625901	701875	0.959731
22	6	0	2	449067	93076	211836	191741	300809	477583	92365	211685	195507	288031	92365	0.126298
23	6	0	3	449067	93076	211836	191741	300809	477583	92365	211685	195507	288031	211685	0.289454

Now we can calculate the survey weight. To this end, we need to join the two datasets which are our sample and the other overall population, based on the same demographic weight variables. Below is the Alteryx workflow configuration to do that.



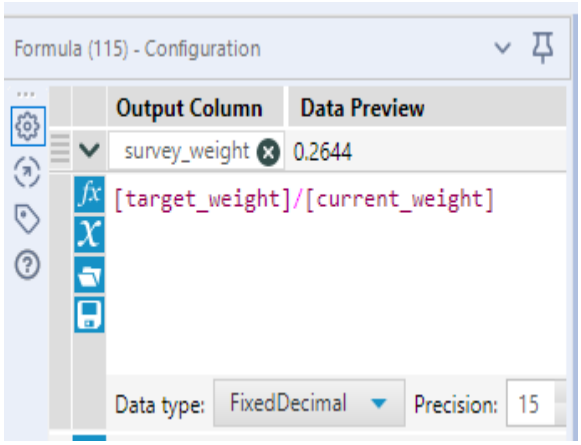
We have both the target_weight and current_weight now and all wee need to do is to calculate the survey_weight dividing the target_weight by the current_weight.

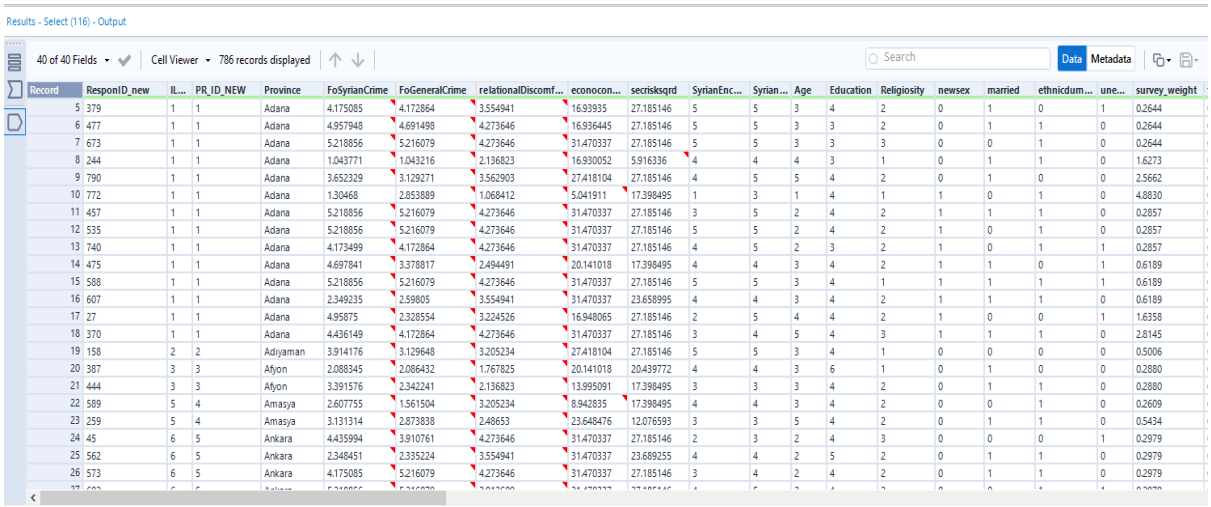
Results - Formula (115) - Output

62 of 62 Fields | Cell Viewer | 786 records displayed

Record	target_weight	current_weight	ResponID_new	PR_ID_NEW	newsex
5	0.235495	0.890585	379	1	0
6	0.235495	0.890585	477	1	0
7	0.235495	0.890585	673	1	0
8	0.20703	0.127226	244	1	0
9	0.326484	0.127226	790	1	0
10	0.621251	0.127226	772	1	1
11	0.109035	0.381679	457	1	1
12	0.109035	0.381679	535	1	1
13	0.109035	0.381679	740	1	1
14	0.236234	0.381679	475	1	1
15	0.236234	0.381679	588	1	1
16	0.236234	0.381679	607	1	1
17	0.208114	0.127226	27	1	1
18	0.358075	0.127226	370	1	1

Survey_weight= target_weight/current_weight





Record	ResponID_new	IL...	PR_ID_NEW	Province	FoSyrianCrime	FoGeneralCrime	relationalDiscomf...	econocon...	secrisksqrd	SyrianEnc...	Syrian ...	Age	Education	Religiosity	newsex	married	ethnicdum...	une...	survey_weight
5	379	1	1	Adana	4.175085	4.172864	3.554941	16.93935	27.185146	5	5	3	4	2	0	1	0	1	0.2644
6	477	1	1	Adana	4.957948	4.691498	4.273646	16.936445	27.185146	5	5	3	3	2	0	1	1	0	0.2644
7	673	1	1	Adana	5.218856	5.216079	4.273646	31.470337	27.185146	5	5	3	3	3	0	0	1	0	0.2644
8	244	1	1	Adana	1.043771	1.043216	2.136823	16.930052	5.916336	4	4	4	3	1	0	1	1	0	1.6273
9	790	1	1	Adana	3.652329	3.129271	3.562903	27.418104	27.185146	4	5	5	4	2	0	1	0	0	2.5662
10	772	1	1	Adana	1.30468	2.853889	1.068412	5.041911	17.398495	1	3	1	4	1	1	0	1	0	4.8830
11	457	1	1	Adana	5.218856	5.216079	4.273646	31.470337	27.185146	3	5	2	4	2	1	1	1	0	0.2857
12	535	1	1	Adana	5.218856	5.216079	4.273646	31.470337	27.185146	5	5	2	4	2	1	0	1	0	0.2857
13	740	1	1	Adana	4.173499	4.172864	4.273646	31.470337	27.185146	4	5	2	3	2	1	0	1	1	0.2857
14	475	1	1	Adana	4.697841	3.378817	2.494491	20.141018	17.398495	4	4	3	4	2	1	1	0	1	0.6189
15	588	1	1	Adana	5.218856	5.216079	4.273646	31.470337	27.185146	5	5	3	4	1	1	1	1	1	0.6189
16	607	1	1	Adana	2.349235	2.59805	3.554941	31.470337	23.658995	4	4	3	4	2	1	1	1	0	0.6189
17	27	1	1	Adana	4.95875	2.328554	3.224526	16.948065	27.185146	2	5	4	4	2	1	0	0	1	1.6358
18	370	1	1	Adana	4.436149	4.172864	4.273646	31.470337	27.185146	3	4	5	4	3	1	1	1	0	2.8145
19	158	2	2	Adiyaman	3.914176	3.129648	3.205234	27.418104	27.185146	5	5	3	4	1	0	0	0	0	0.5006
20	387	3	3	Afyon	2.088345	2.086432	1.767825	20.141018	20.439772	4	4	3	6	1	0	1	0	0	0.2880
21	444	3	3	Afyon	3.391576	2.342241	2.136823	13.995091	17.398495	3	3	3	4	2	0	1	1	0	0.2880
22	589	5	4	Amasya	2.607755	1.561504	3.205234	8.942835	17.398495	4	4	3	4	2	0	0	1	0	0.2609
23	259	5	4	Amasya	3.131314	2.873838	2.48653	23.648476	12.076593	3	3	5	4	2	0	1	1	0	0.5434
24	45	6	5	Ankara	4.435994	3.910761	4.273646	31.470337	27.185146	2	3	2	4	3	0	0	0	1	0.2979
25	562	6	5	Ankara	2.348451	2.335224	3.554941	31.470337	23.689255	4	4	2	5	2	0	1	1	0	0.2979
26	573	6	5	Ankara	4.175085	5.216079	4.273646	31.470337	27.185146	3	4	2	4	2	0	1	1	0	0.2979

After calculating the survey weight, all we need to do is including it in the regression analysis together with the other predictor variables. The outcome will give the weighted regression results.

Since we have two dependent variables and we run the regression with and without the survey_weight, you see four different regression models in the Alteryx workflow below.