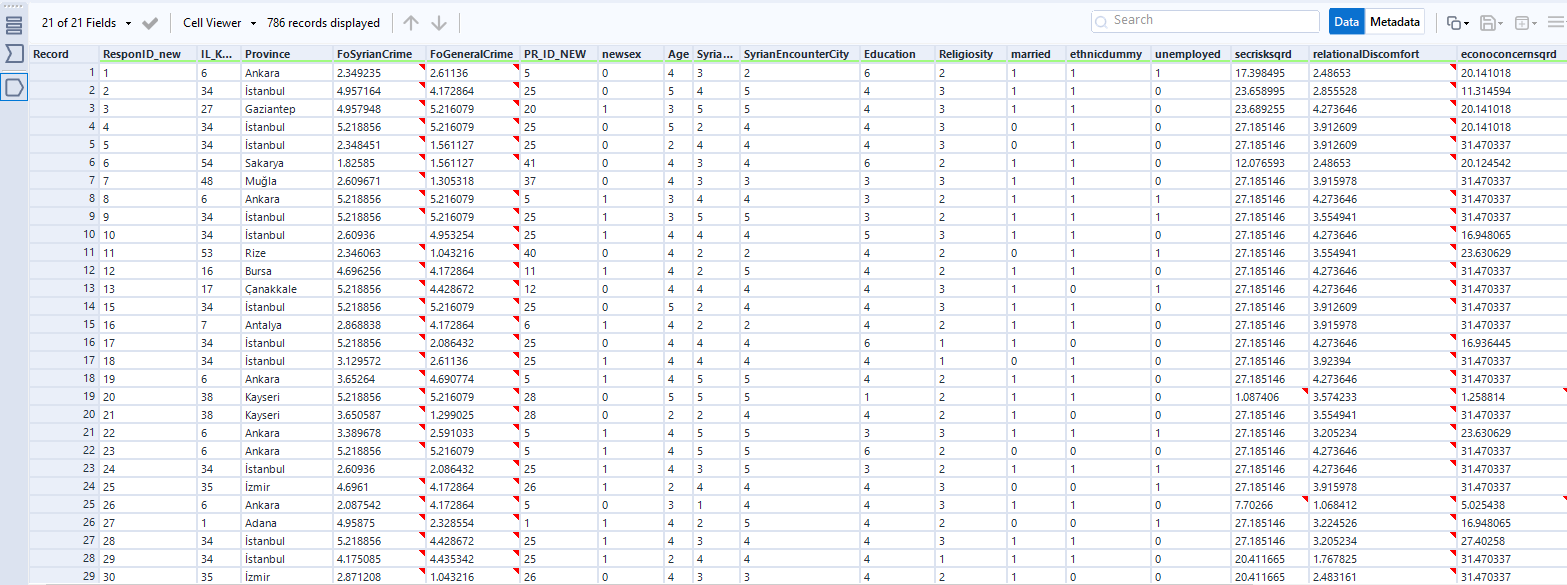
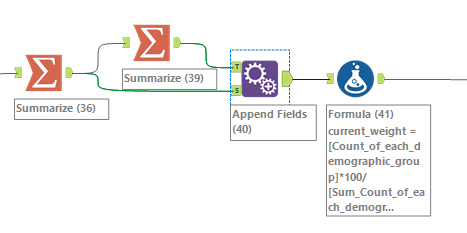
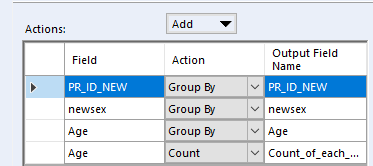
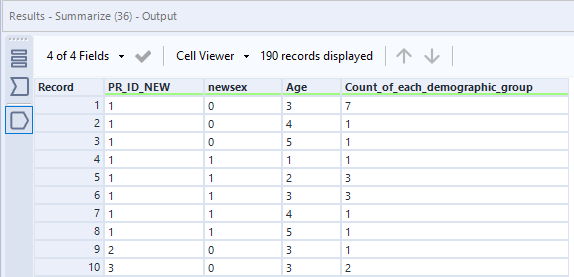
This is the screenshot of the initial survey data which has 786 respondents from 53 different cities of Turkey.



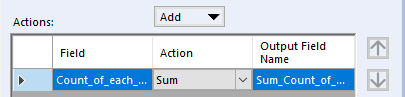
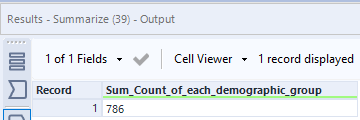
First, we created our combination groups with the demographic variables we would like to weight. They are respectively; 1) population of the city 2) gender and 3) five different age groups. We need to get the number of the respondents in each combination of the three demographic groups. The first summarize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

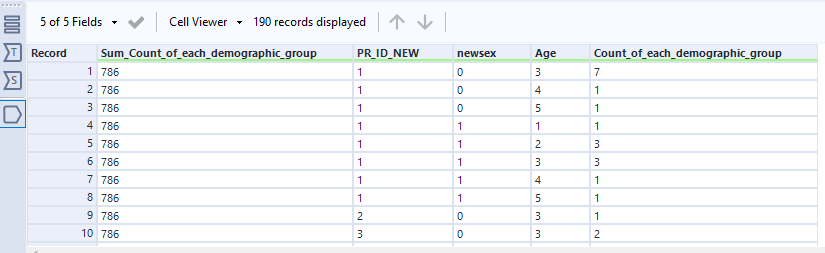
 

Then we need to get the total number of the respondents in our sample (sample size=786). Second Summarize tool above in the Alteryx workflow does that and below is the configuration of the workflow and the output.

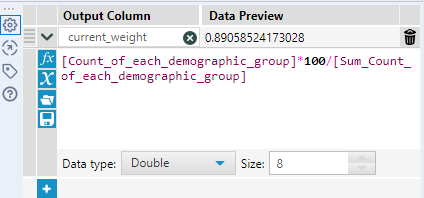
 

Then

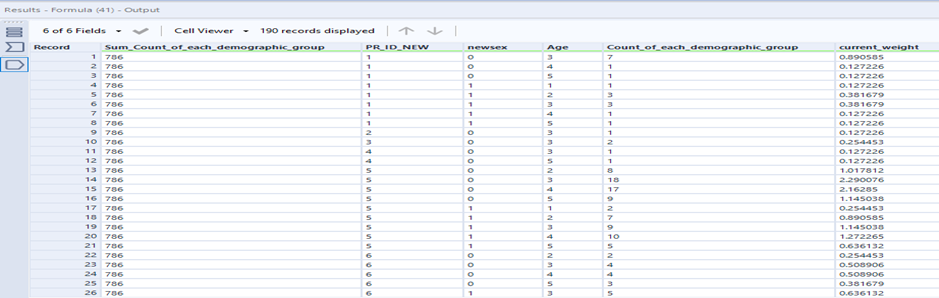
With the Append tool in the Alteryx workflow above, we joined the two tables. Below is the output.



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.

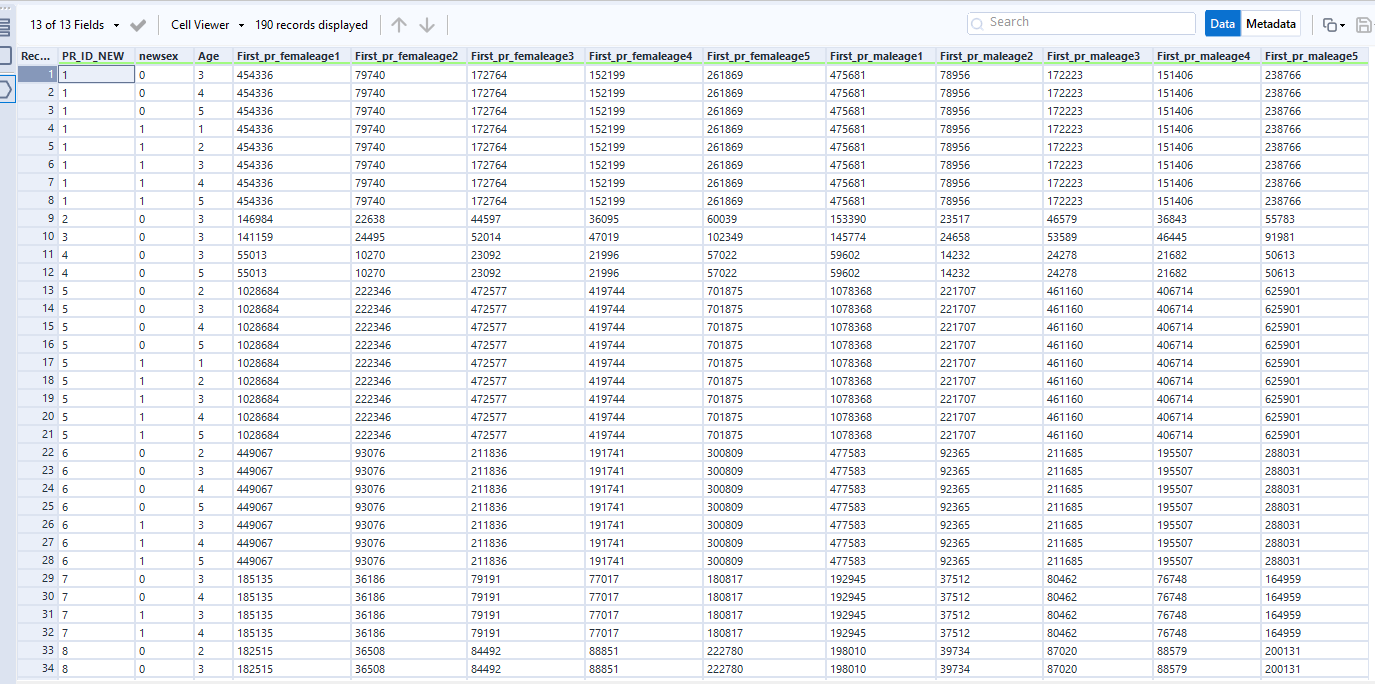


Note that, in Adana province (PR\_ID\_NEW = 1), the number of the males (newsex = 0) in the age group 3 (Age =3) is 7 and the representation of this category in our sample is % 0.89. In Adana (PR\_ID\_NEW = 1), the number of the females (newsex = 1) in age group 3 (Age =3) is 3 and the representation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

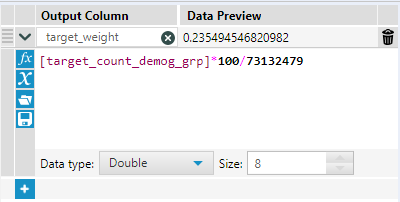
As you may notice in our sample, we do not have respondents for all combinations of the demographic groups. We have respondents from 53 cities of Turkey out of 81. That means if we had respondents from all the demographic groups, we would have 530 different combinations of demographic groups. (population 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.

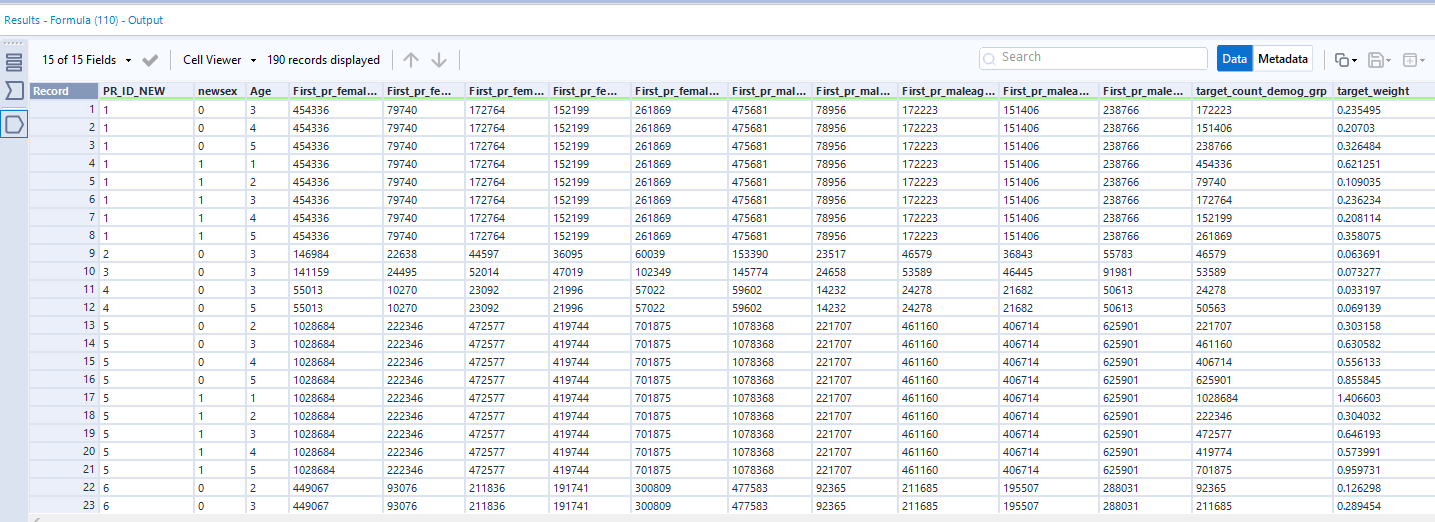
For example, in Adana (PR\_ID\_NEW = 1) there is no male respondent in the age group 1. That is what makes the calculation of the survey weight more complicated in this case because before calculating the survey\_weight, we need to create the same combination of the demographic groups for our target 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.

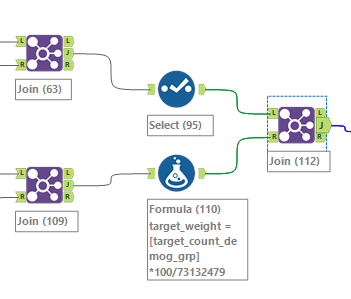
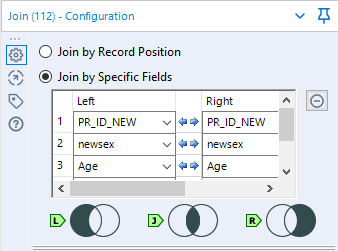


It required some tedious work to obtain each statistics but afterwards applying the same procedures we easily calculated the target weight. We divided the number of the individuals in each demographic group by the overall population of the 53 cities of Turkey. If we had respondents from all the cities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

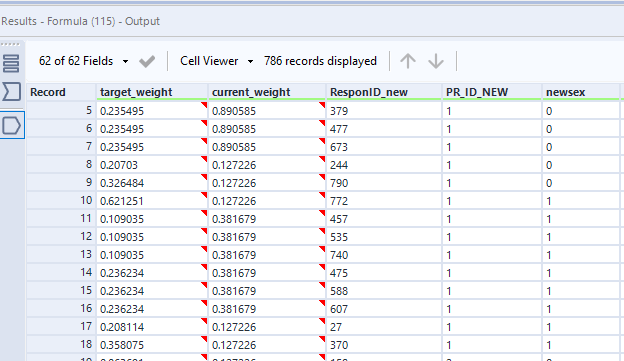




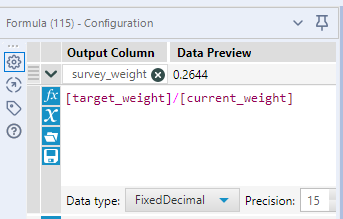
Now we can calculate the survey weight. To this end, we need to join the two datasets which are our sample and the other overall population, based on the same demographic weight variables. Below is the Alteryx workflow configuration to do that.

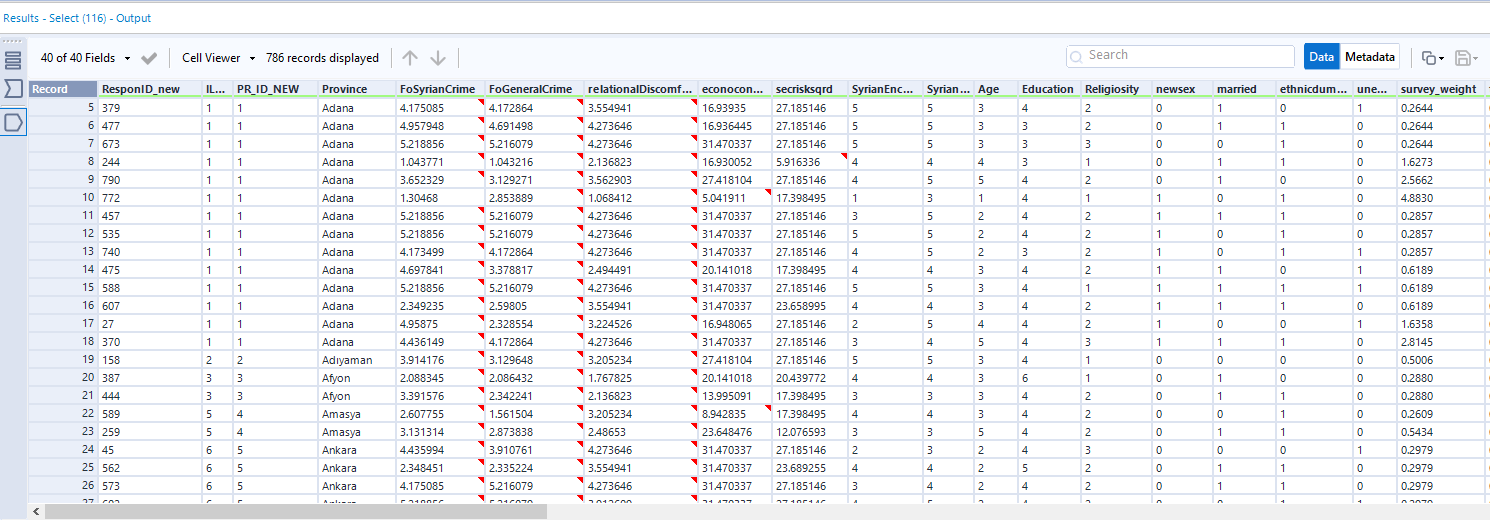
 

We have both the target\_weight and current\_weight now and all wee need to do is to calculate the survey\_weight dividing the target\_weight by the current\_weight.



Survey\_weight= target\_weight/current\_weight





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.

