

Matlab Table Summarize and Aggregate by Groups

back to [Fan's Intro Math for Econ](#), [Matlab Examples](#), or [MEconTools Repositories](#)

Group Table Rows and Sum within Group

There is a table where subsets of rows belong to different simulations, with exogenous fixed ρ parameters. For each ρ parameter combination, there are, stored as different rows, a number of model predictions and data moments, and corresponding difference. Find the total difference between model and data for subsets of rows based for each ρ parameter set.

First, create a table where each ρ group is identified jointly by ρ_a and ρ_b , stored in the 3rd and 4th rows.

```
% Make N by 2 matrix of fieldname + value type
mt_st_variable_names_types = [...
    ["year", "double"];["category", "string"];...
    ["rhoa", "double"];["rhob", "double"]; ...
    ["numberWorkersSimu", "logical"]; ["numberWorkersData", "double"]];
% Make table using fieldnames & value types from above
tb_agg_exa = table('Size',[0,size(mt_st_variable_names_types,1)],...
    'VariableNames', mt_st_variable_names_types(:,1),...
    'VariableTypes', mt_st_variable_names_types(:,2));
% Table with data inputs
tb_agg_exa = [tb_agg_exa;...
    {1, 'C001', 0.50, 0.50, 5.5, 6.05}; {2, 'C002', 0.50, 0.50, 3.7, 4.4}; ...
    {1, 'C001', 0.25, 0.30, 2.5, 3.65}; {2, 'C002', 0.25, 0.30, 0.1, 1.6}; ...
    {3, 'C001', 0.25, 0.50, 0.01, 1.66}];
% Generate model and data difference
tb_agg_exa(:, "diff") = tb_agg_exa(:, "numberWorkersSimu") - tb_agg_exa(:, "numberWorkersData");
% Display
disp(tb_agg_exa);
```

| year | category | rhoa | rhob | numberWorkersSimu | numberWorkersData | diff |
|------|----------|------|------|-------------------|-------------------|-------|
| 1 | "C001" | 0.5 | 0.5 | 5.5 | 6.05 | -0.55 |
| 2 | "C002" | 0.5 | 0.5 | 3.7 | 4.4 | -0.7 |
| 1 | "C001" | 0.25 | 0.3 | 2.5 | 3.65 | -1.15 |
| 2 | "C002" | 0.25 | 0.3 | 0.1 | 1.6 | -1.5 |
| 3 | "C001" | 0.25 | 0.5 | 0.01 | 1.66 | -1.65 |

Second, select the subset of columns that are relevant for aggregation.

```
% Select
tb_agg_exa = tb_agg_exa(:, ["rhoa", "rhob", "diff"]);
% Display
disp(tb_agg_exa);
```

| rhoa | rhob | diff |
|------|------|-------|
| 0.5 | 0.5 | -0.55 |
| 0.5 | 0.5 | -0.7 |
| 0.25 | 0.3 | -1.15 |
| 0.25 | 0.3 | -1.5 |

0.25 0.5 -1.65

Third, group by unique combinations of rhoa, rhob, and aggregate. Then generate group ID.

```
% Sum within groupo
tb_groupby_agg_sum = groupsummary(tb_agg_exa, ["rhoa", "rhob"], "sum");
% Generate grouping ID
tb_groupby_agg_sum{:, "ID"} = (1:1:size(tb_groupby_agg_sum, 1))';
tb_groupby_agg_sum = movevars(tb_groupby_agg_sum, "ID", "Before", 1);
disp(tb_groupby_agg_sum);
```

| ID | rhoa | rhob | GroupCount | sum_diff |
|----|------|------|------------|----------|
| 1 | 0.25 | 0.3 | 2 | -2.65 |
| 2 | 0.25 | 0.5 | 1 | -1.65 |
| 3 | 0.5 | 0.5 | 2 | -1.25 |