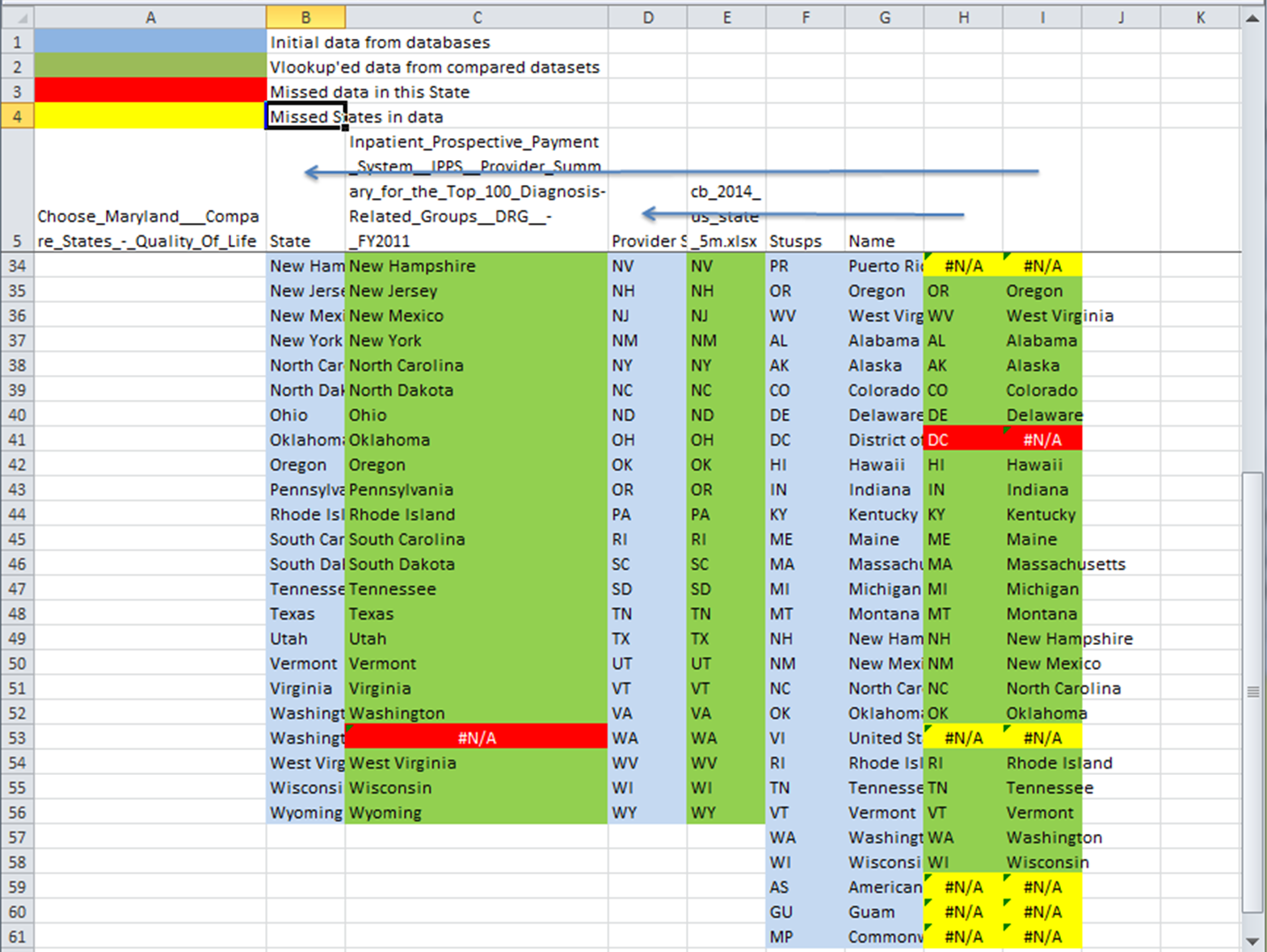
**Step 1. Validating data.** We will start by checking to see if we have the accurate and complete list of States in each file and ensure compatibility by ensuring that we have equal number of States in each file and identify missing States.



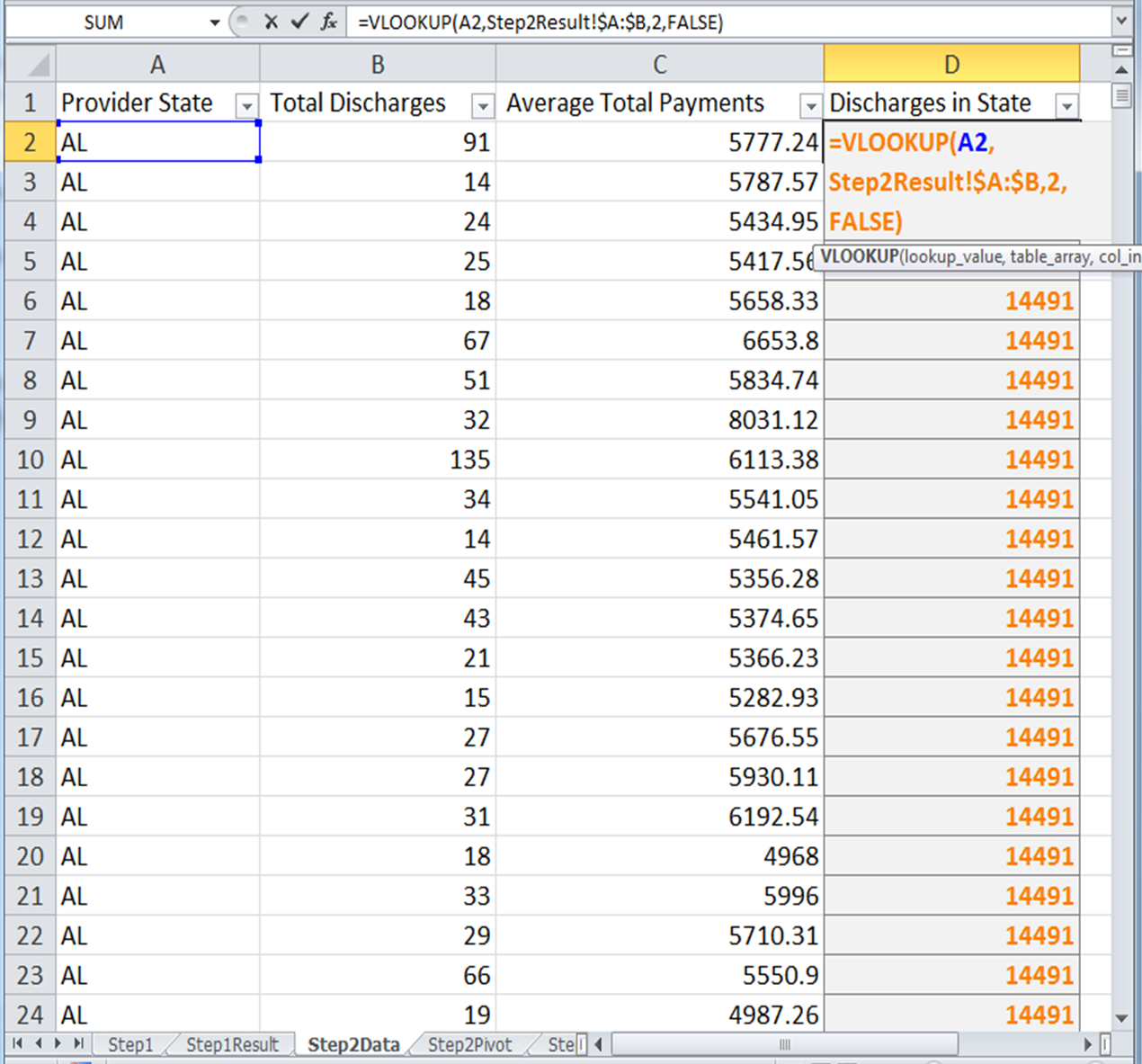
**Figure 1.** Data validation results.

We had to copy “State” (column B) from [2] to our Excel file (“Step1” sheet), “Provider State” (column D) from [1], and F and G columns from [3]. We then remove duplicates in “Provider State” using “Data->Remove Duplicates” (only for column D (to get to “Step1” sheet in current view). Then using Excel’s “Vlookup” function in green cells we can see the differences between values (columns C, E, H and I on figure 1) . We then get to know we have lost Puerto Rico, United States Virgin Islands, American Samoa, Guam, and Commonwealth of the Northern Mariana Islands – marked yellow. These territories are missing in both files with data, So we can safely get to exclude them.

**Step 2. Data preparation.**Getting “Provider State”, “Total Discharges”, “Average Total Payments” columns from [1] to our Excel file (sheet Step2Data columns A-C) we can make pivot table from these data to find sum of Total Discharges in each state (Step2Pivot), because each provider performance depend not only on“Average total payments”, but also on “Total discharges” of the provider too. The result of calculation was copied to “Step2Result” sheet (figure 2). Then using Vlookup function in column D of “Step2Data” sheet we can populate these results for each provider (figure 3).

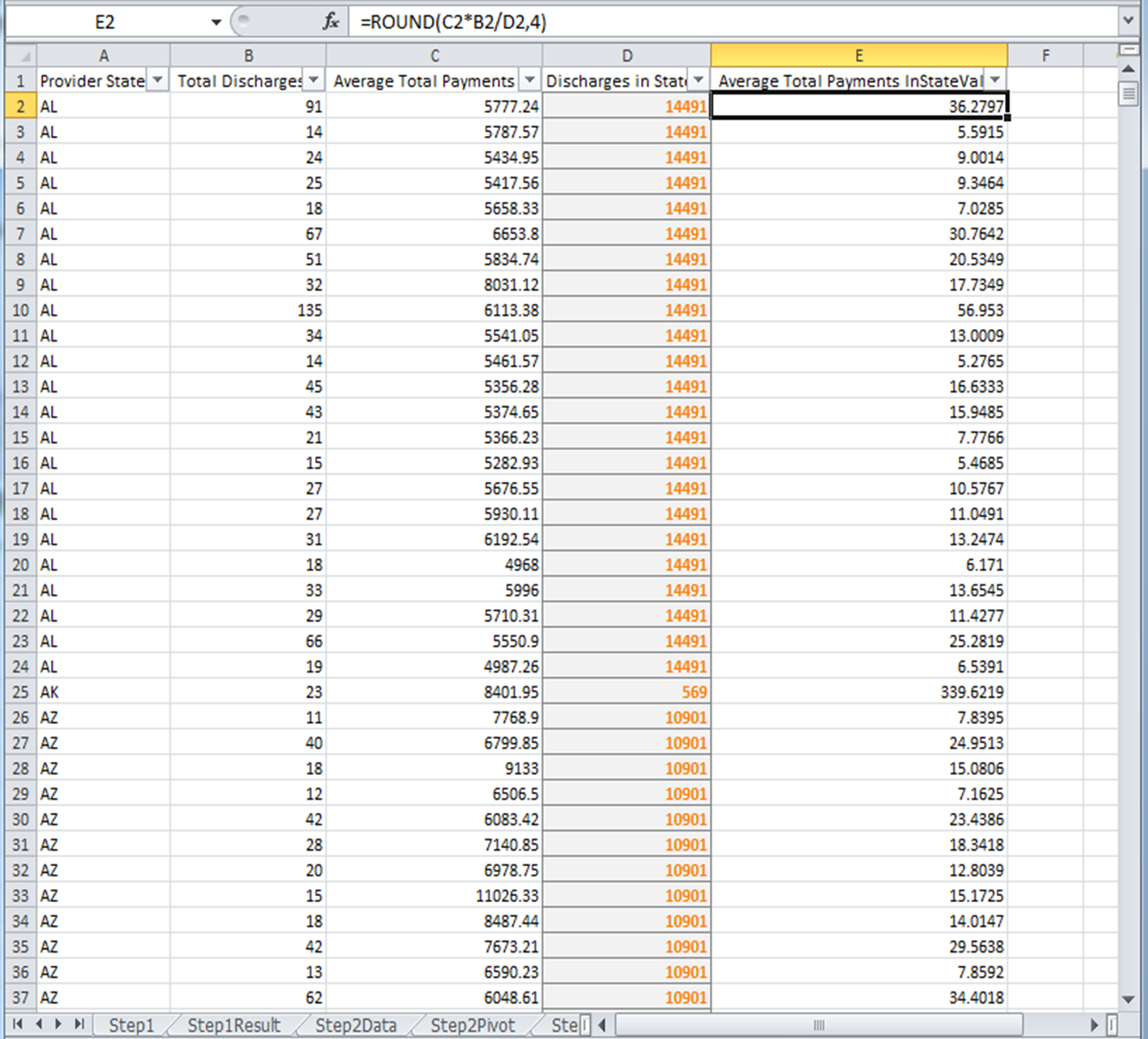
**Figure 2.** Summary of total discharges on state level, exported from pivot table **Three**

**Figure 3.** ThePopulating of Total Discharges at State level in [1] dataset.

****

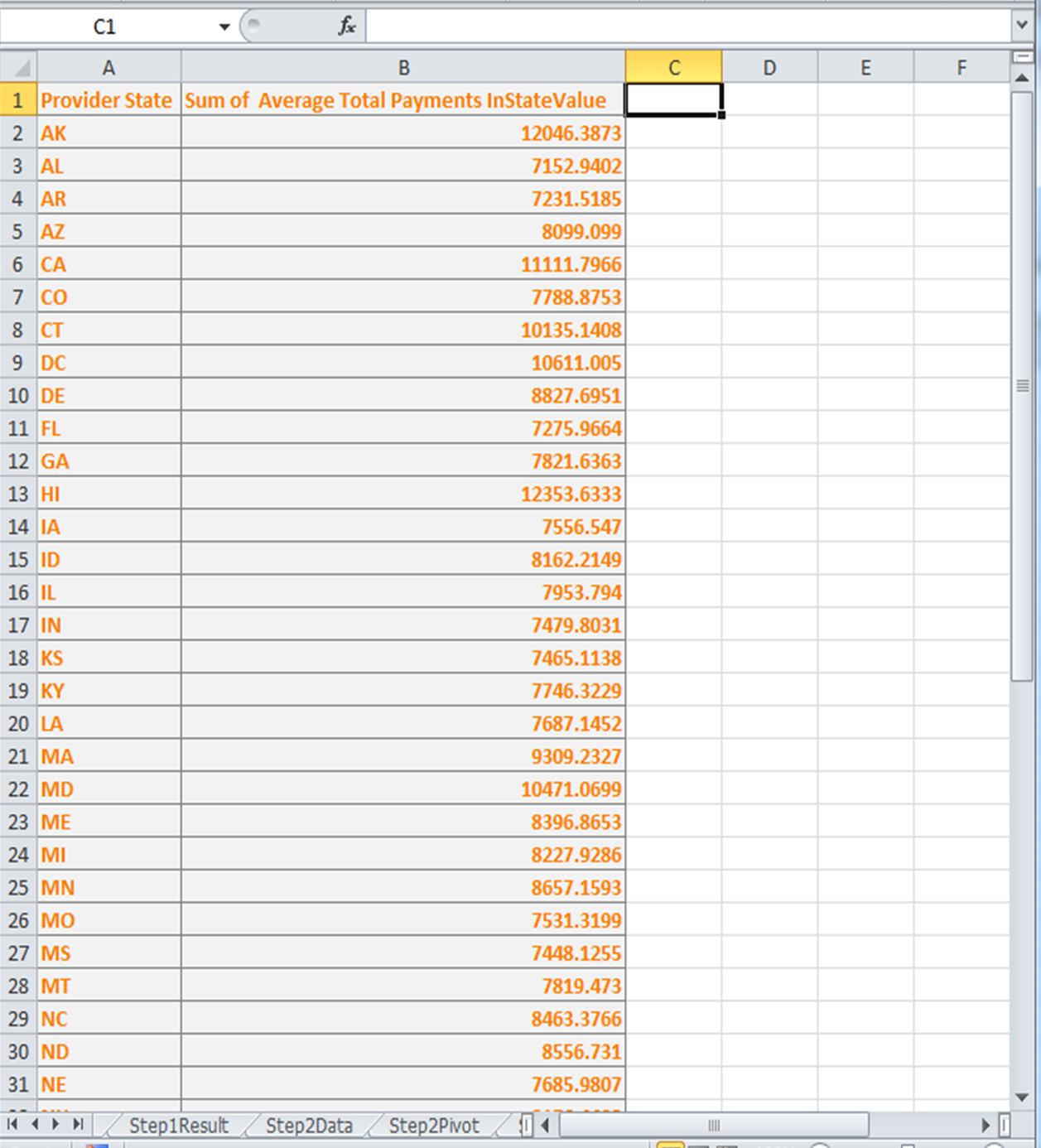
**Step 3.** Now we can populate and calculate Average Total Payments in StateValue. Every provider has different count of Discharges. The “Average total payments” was calculated for one Discharge. To obtain “Average total payment” in State level I have to multiply “Average Total Payments” of provider on Count of discharges of this provider and divide on count of Discharges of all providers of State. (figure 4).

**Figure 4.** Calculation of subject’s values for summary in [1] dataset

****

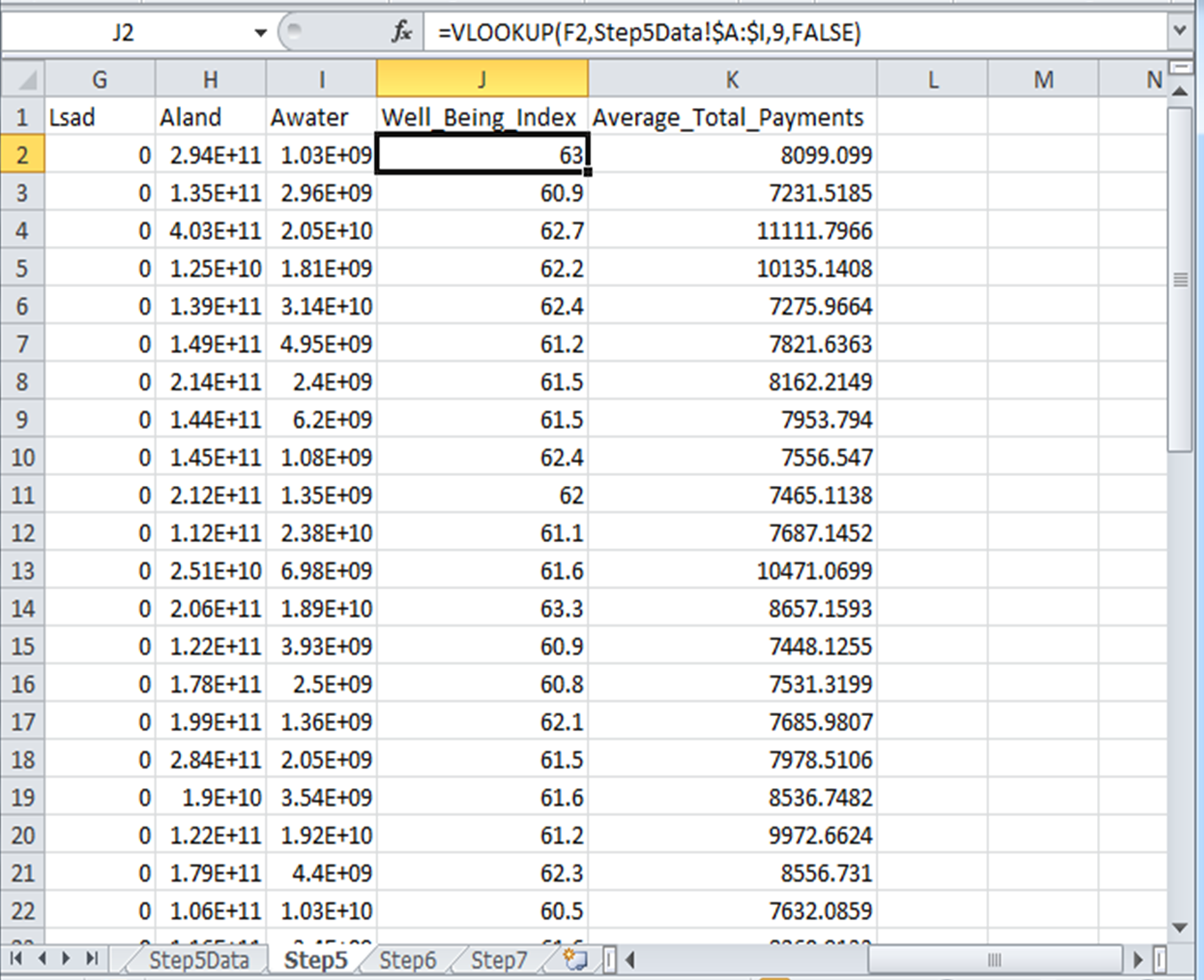
**Step 4**. The sum of “Average Total Payments in State Value” was calculated in each State. (figure 5). To achieve this we created a pivot table from “Step3” data, then we displayed  “Provider State” in Row lables, and “Average Total Payments in State Value” in “sum Values”.  The result was then copied from Pivot to “Step4” sheet.

**Figure 5.** Summarize “Average Total Payments in State Value” for each “Provider State”.

****

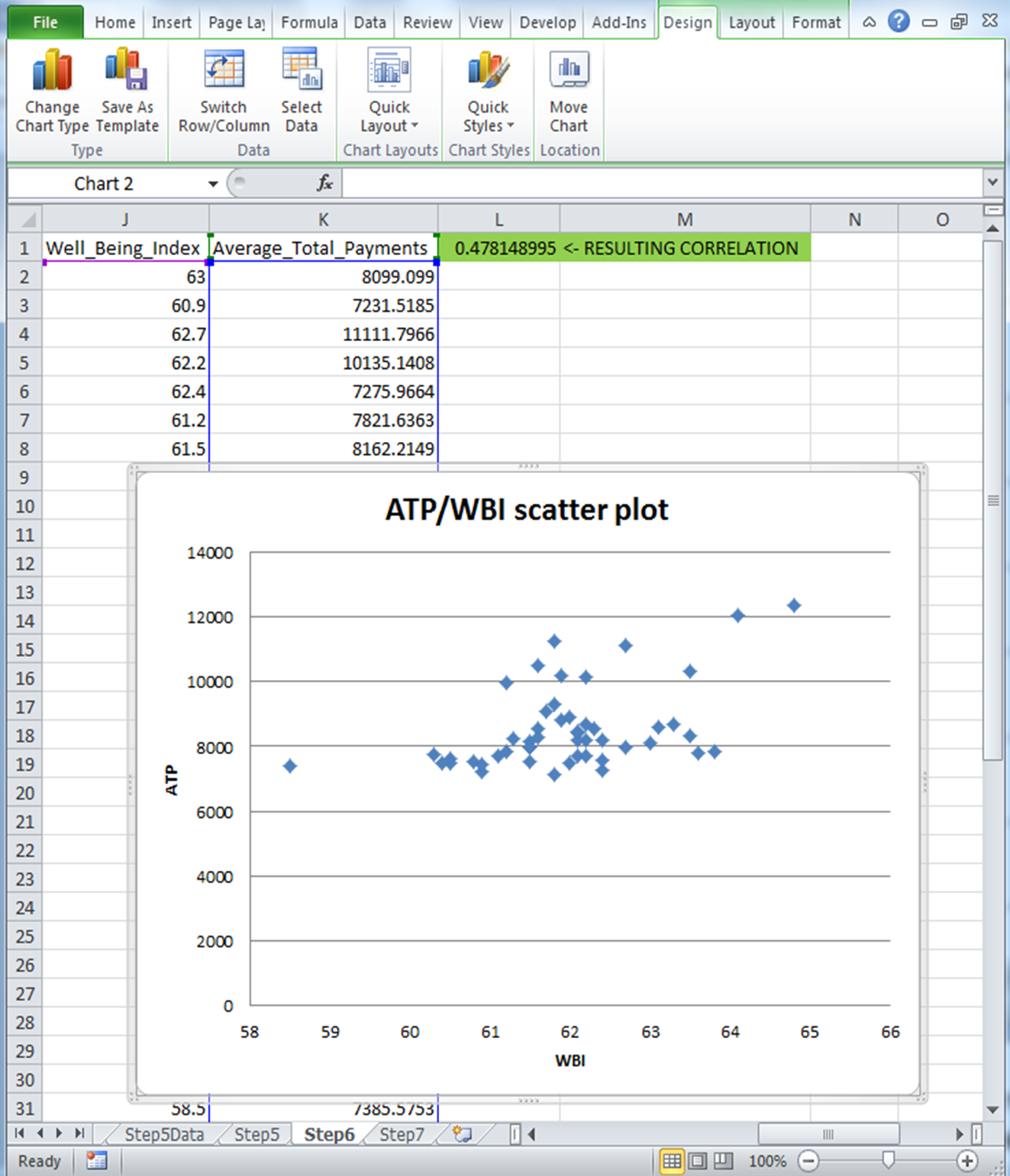
**Step 5.** We can copy data from [3] to our Excel fiile (Columns A-I of “Step5” sheet) and copy [2] to “Step5Data” sheet, then using “VLOOKUP” we can populate “Average\_Total\_Payments” (for State level) from “Step4” sheet and “Well\_Being\_Index” from “Step5Data” to“Step5” columns  J and K. (figure 6).

**Figure 6. Populate of WBI(Well Being Index) and ATP(Average Total Payments) data in the resulting dataset.**

****

**Step 6. Analysis process.** To identify whether or not Well Being index has anything to do with Average total payments in hospitals at State level we start from correlation analysis between these variables. The correlation coefficient is 0.478148995. (Figure 7.)

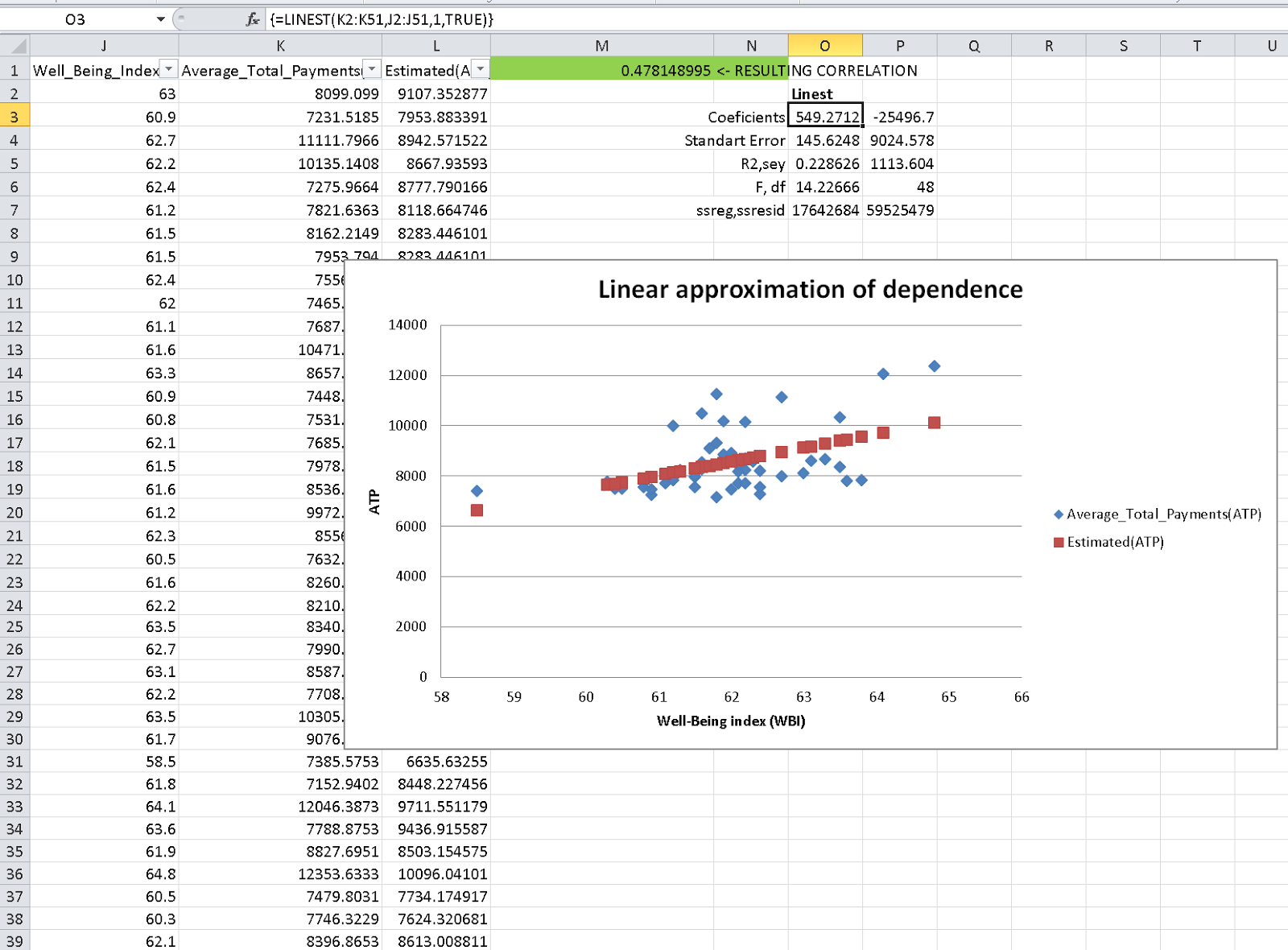
**Figure 7. Obtaining of correlation coefficient ( =CORREL(K2:K51,J2:J51) )**

****

**Step 7. Analysis process part 2.**

There seems to be a good correlation between WBI and ATP. We can study this further using a linear regression approximation to identify possible linear dependence between WBI and ATP.  Using LINEST function in excel we made a linear approximation. Shown on figure 8 the data and calculated estimation (from approximation) values in column L we can see in the chart how close are our approximation (red points)  to real data (blue). (Figure 8.)

**Fiqure 8. Linear approximation of dependence.**

****