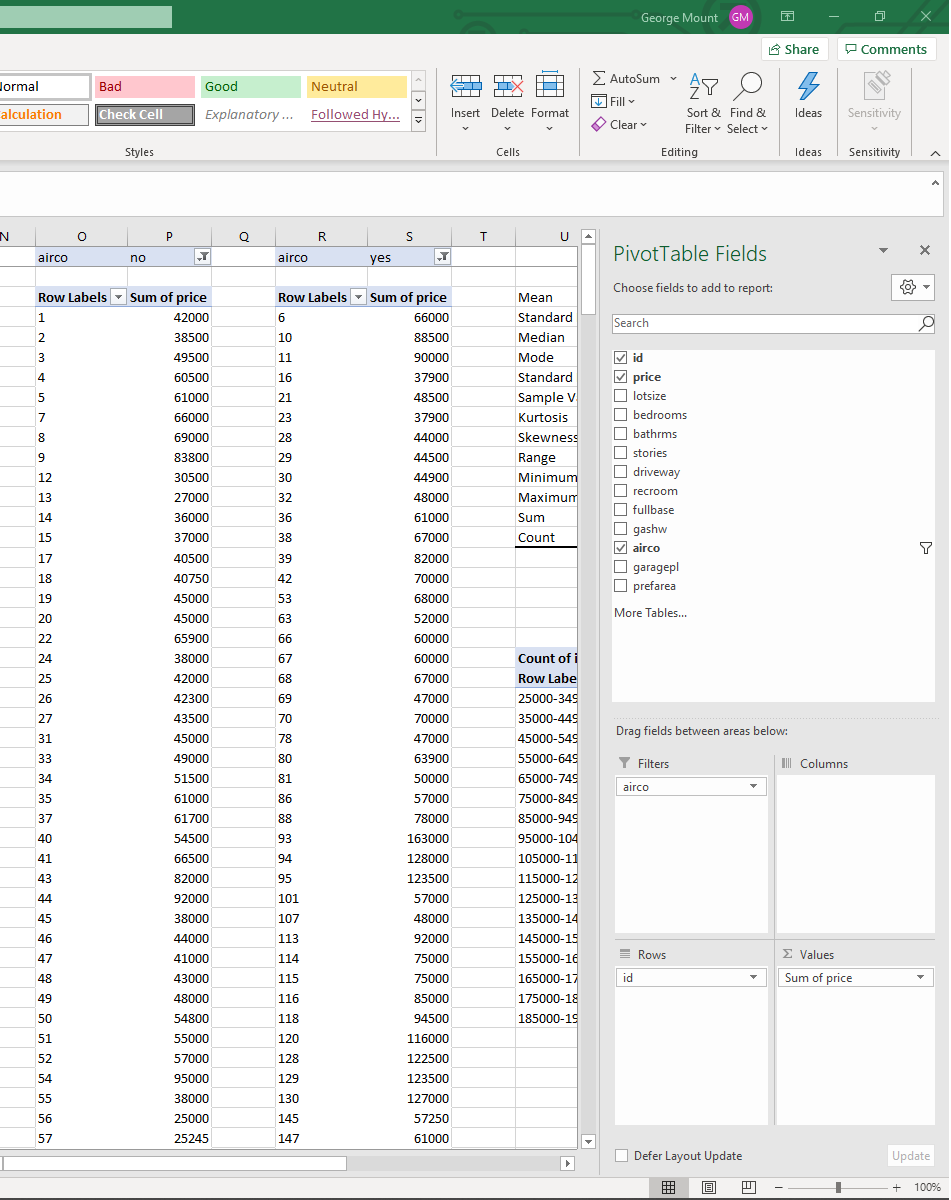
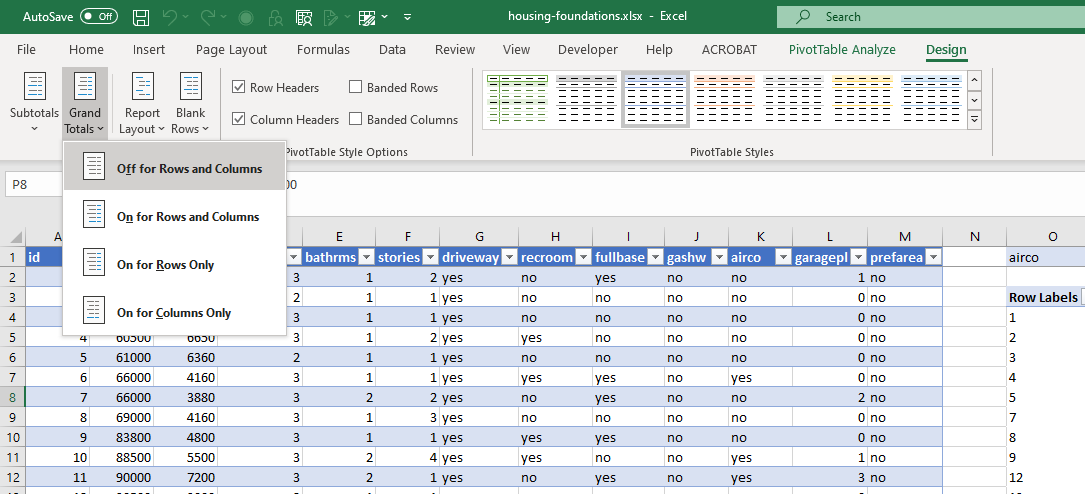
**FOUNDATIONS OF INFERENTIAL STATISTICS – DEMO NOTES**

**Descriptive statistics for two categories**

1. Create a PivotTable from the raw data.
2. Group create two PivotTables, each displaying the variable you want to measure in the Values, the ID variable in the Rows, and the category in the Filter:



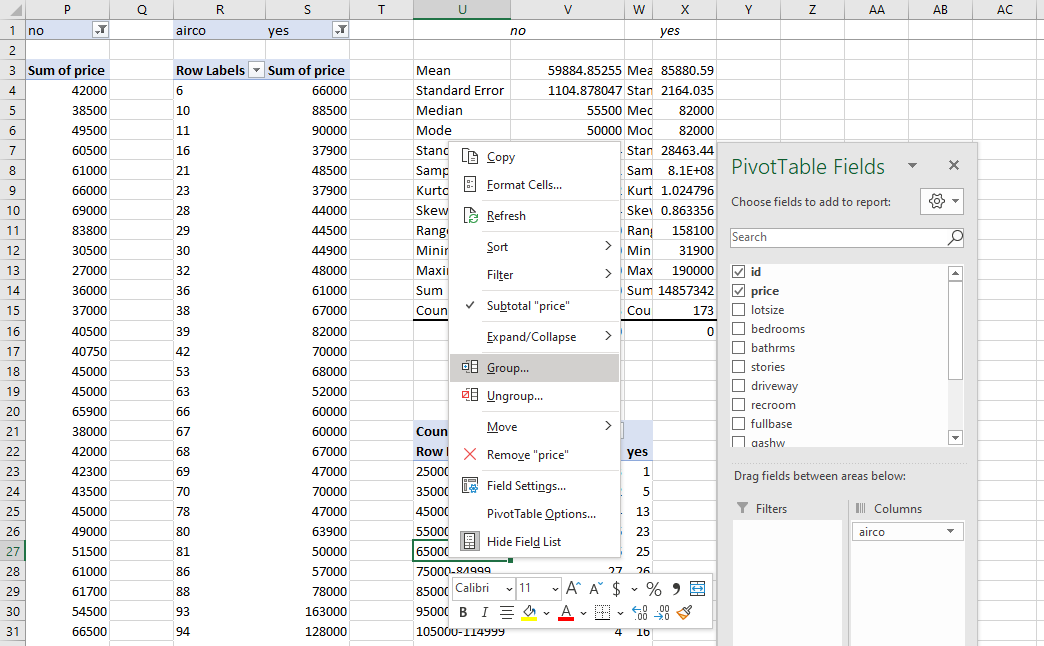
1. Remove the Grand Totals for the PivotTables by clicking inside the PivotTable, selecting Design from the home ribbon, then in the Layout group, under Grand Totals, select Off for Rows and Columns.



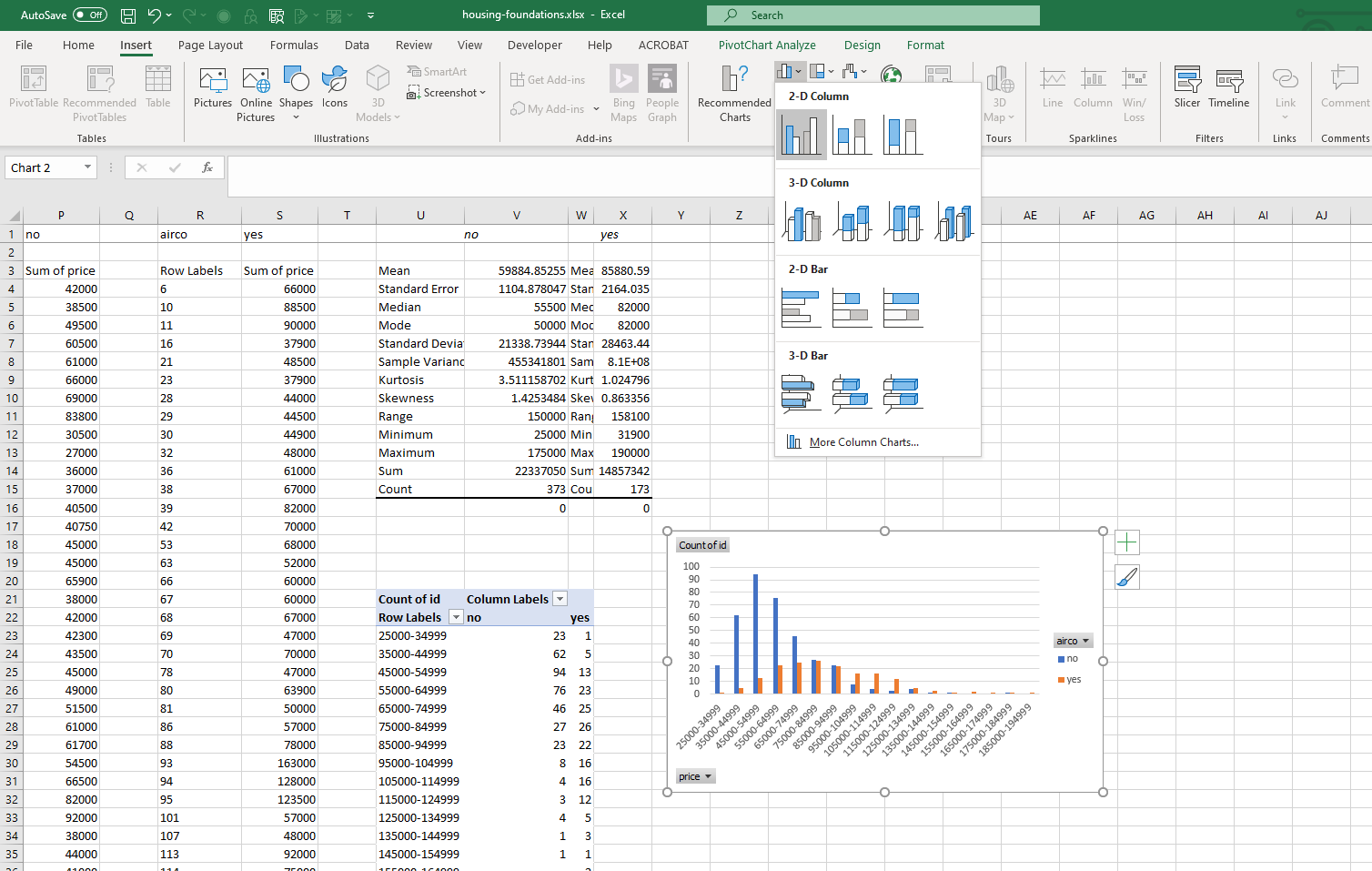
1. Run the descriptives for each category using the Analysis ToolPak. See demo notes from Section 1 for a refresher.

**Plot a histogram for two categories**

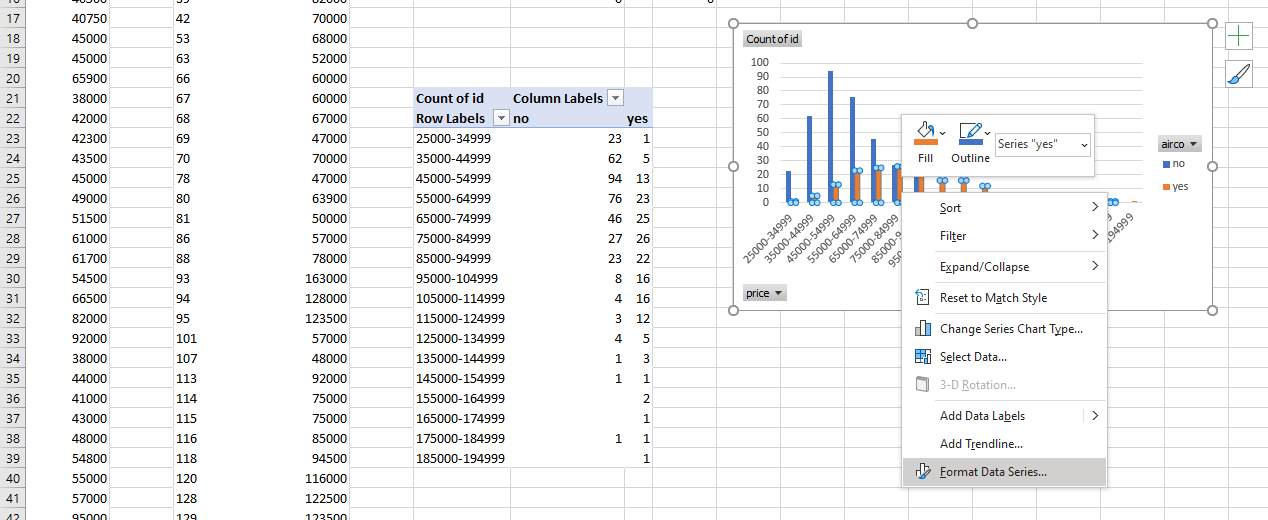
1. Create another PivotTable.
2. Place the two categories along the Columns, the continuous variable of interest down the Rows, and the Count of the ID variable in the Values.
3. Right-click the Rows area and select Group.



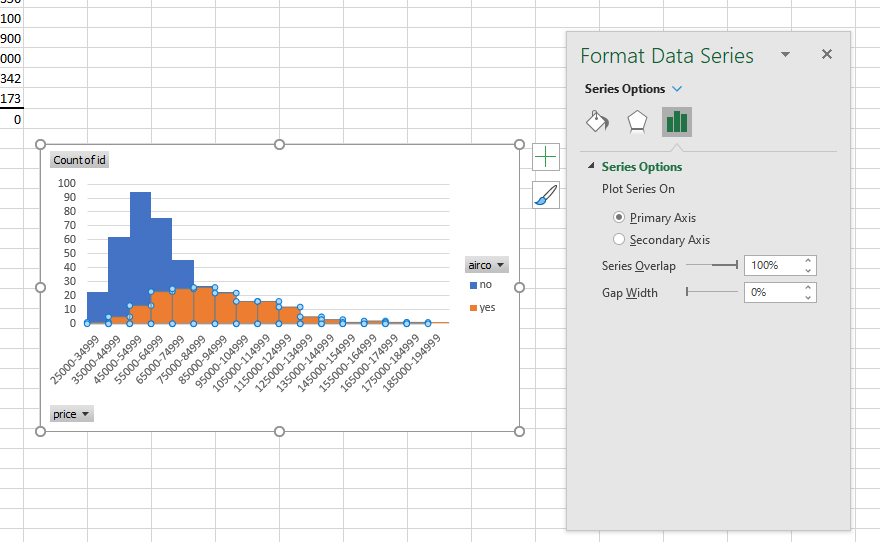
1. Go to Insert on the home ribbon and select a 2-D Column chart.



1. Right-click on any of the bars in the resulting bar chart. Select Format Data Series.



1. Adjust the Series Overlap to 100% and the Gap Width to 0%.



Demo: margin-of-error.xlsx

Here is our housing dataset. We want to examine how likely we expect our true mean to fall within our margin of error. To approximate this, we will calculate the percent of the margin of error to the mean. The lower the margin of error, the more likely we expect our population mean to fall within the bounds.

1. In Column C we calculate the sample mean of the entire sample thus far using mixed referencing: =AVERAGE($B$2:INDEX($B$2:$B$547,$A2))
2. In Column D we calculate the sample variance of the entire sample thus far: =VAR.S($B$2:INDEX($B$2:$B$547,$A3))
3. In Column E we calculate the standard error of our sample thus far: =SQRT(D5)/SQRT(A5)
4. In Column F we calculate our critical value. Technically the 1.96 critical value is for larger sample sizes (at least greater than 200), so we will use a lookup table to give a more generous critical value for smaller sample sizes: =VLOOKUP($A4,'critical-value'!$A$1:$B$34,2)
5. In Column G we can now calculate our margin of error, which is our standard error times our critical value: =E2\*F2
6. Last but not least, in Column H calculate the margin of error as % of the mean. =G2/C2
7. Now plot Column H as a line chart.
8. We can see that over time, we get increasingly closer to 3%. If our sample were even larger, this number would dip lower.
   1. Remember, this is only going to help us if our sample is indeed representative of the population!