Objective: Analyze bike shop transactions. Explore data and look for trends. Determine (TBD after analysis). Create Pivot tables and a dashboard for presenting the data. Practice/showcase use of vlookups, pivot tables, and charts in Excel.

Data Source:

- https://www.kaggle.com/datasets/rohitsahoo/bicycle-store-dataset) dataset.

- I took inspiration and learned a few new skills from https://www.youtube.com/watch?v=bjLIA1vSqGs) by Lean Excel Solutions

Observations/To Do

* Date format. Some have dashes, other slashes. In dd-mm-yy format Ones starting with 1-12 were converted to Excel serial. Had trouble converting them because Excel got month and day backwards. Ultimately addressed the 2 styles separately. Used =DATE(YEAR(D6),DAY(D6),(MONTH(D6))) for the ones in Excel Serial format (ones starting with 1-12) and =DATE(RIGHT(D2,4),MID(D2,4,2),LEFT(D2,2))
* product\_first\_sold\_date is in Excel format. (or not, may not need)
* Demographics tab has a “Default” column with what appears to be computer garbage. Source has very little data so no idea what this means
* Address/Property Valuation – no idea what this means
* Demographic/Tenure – unsure what this means. Time at Job?
* 197 Rows missing most data – deleted from working sheet
* Clean Male/Female columns. Sometimes spelled out, sometimes just M/F
* In Transactions, the same 197 lines are missing several fields: brand product\_line, product\_class, product\_size, list\_price, standard\_cost, product\_first\_sold\_date. Deleted from working sheet, but left in demographic tab
* There is no actual sale price or cost data. Used list price and standard cost.
* Customer list has no customer IDs, but customer names are in Demographic tab. May just not use customer tab at all
* Data is very unform from month to month, 50/50 online vs offline, very evenly split between brands. 50% profit seems high. This could be legit for Australia, but user has not responded to requests for more information. Perhaps the data was artificially created?