

Lab 2: Shaping Tables

Part 1: Breaking a file into multiple tables

Open the Power BI Desktop file named `2_1_split_table.pbix` in the `Exercises` folder on the Desktop and go to Power Query.

Duplicate the `Manufacturing Data` table and rename it to `NAICS`.

- In this new `NAICS` table, keep only `2017 NAICS code` and `Meaning of NAICS code`.
- Remove the duplicate codes from `2017 NAICS code`.

Which "Meaning of NAICS code" is shown on the 4th row of the NAICS table?

- ☐ Grain and oilseed milling
- ☐ Manufacturing
- ☐ Food manufacturing

Part 2: Appending files

Within Power Query, load each of the following files as new datasets: `Timeseries_1979.csv`, `Timeseries_1980.csv`, `Timeseries_1981.csv`. They are located in the `Datasets` folder on the Desktop.

For each of the tables you just loaded, remove the first row and use the new first row as headers.

Create a new table, called `Establishments by Age`, with all of the `Timeseries_19XX` tables appended.

Apply the Power Query changes and return to the *Report* view in Power BI. Hide all separate `Timeseries_19XX` tables.

Create a *Clustered bar chart* visual. From the `Establishments by Age` table use the `Meaning of Establishment age code` as *Axis* and `Number of employees` for the *Values*.

How many employees were hired at firms which had existed for two years?

Part 3: Column extraction

Go to Power Query and select the `Manufacturing Data` table. Navigate to the last column, which should be `Range indicating percent of total employees imputed`.

Duplicate the `Range indicating percent of total employees imputed` column **twice**.

- From the first duplicated column, only keep the numbers before the first percent sign (%). You can use "%" as a delimiter.
- Change the resulting column's data type to "Whole Number" and rename the column to `Low Range Total Employees`.
- From the second duplicated column, only keep the numbers before the second percent sign (%). First extract the last four characters, then extract again using "%" as a delimiter.
- Change the resulting column's data type to "Whole Number" and rename the column to `High Range Total Employees`.

Close and apply your steps and create new page with a clustered column chart visual. From `Manufacturing Data`, use the `High Range Total Employees` column as the *Axis* and `First-quarter payroll ($1,000)` for the *Values*.

Which is the second-largest "High Range Total Employees" category in terms of payroll?