

**OPENSTAT PSA: SUMMARY STATISTICS FOR MANUFACTURING
ESTABLISHMENTS BY INDUSTRY DESCRIPTION, SIZE AND DATA ITEM**

A capstone project
For PROJECT SPARTA Y5
Data Engineer Pathway

KARL JOSEPH C. CACAO, RCE

November 2024

ACKNOWLEDGEMENT

This section presents the scholar's gratitude for their contributions to the project's success:

The **Development Academy of the Philippines** (DAP), for granting him a scholarship for the PROJECT SPARTA Y5 in order to be enhanced in data literacy.

The **Department of Science and Technology - Philippine Council for Industry, Energy, and Emerging Technology Research and Development** (DOST-PCIEERD), being one of the agencies for science and technology innovation and research play a national role for STEM developments in the country.

The **Philippine Statistics Authority** (PSA) for providing the Filipinos handful of open data for the benefit of public good. To its website OpenSTAT, it has been an interesting platform accessible to all to gather information about statuses of the Philippine sectors.

Pleasantries and gratitude also to the **Data Engineering Pilipinas**, a PyData community in the Philippines for inspiring young students, professionals and out-of-school people, to be data-driven with the help of their advocacy. The journey of the scholar started from Data Analytics, but keep on moving on Data Engineering to understand more of the big data.

To **Engr. Rammel F. Mistica, Engr. Crispino A. Aranilla**, the scholar's former academic instructors for believing in him for the opportunity of learning the Data Engineer Pathway.

To the scholar's *family* who morally support his decision and potential in developing his skills in data and integrate these to various domain knowledge.

Most especially to **Almighty God** who showed tenacity and determination in completing this research.

K.J.C.C

INTRODUCTION

Philippine economics seemed to be abstract in the point of view of ordinary Filipinos. Most, if not all are not aware of the ongoing status of the economy for various sectors. One of the challenges lie in the information dissemination.

The Philippine Statistics Authority holds all the necessary information for several aspects. They are responsible in the national censuses and surveys as the legal way of collecting information for the record and may be used for academic and socio-political developments. For social statistics, information such as demographics, population and labor could be accessed to the agency. For information on resources, environmental, agricultural—fisheries, mining and construction could be present also. PSA has a website for open data in the Philippine context. OpenSTAT (<https://openstat.psa.gov.ph>) is the website made accessible to the public for viewing statistics on the different sectors of the economy.

As cited on United Nations Department of Economic and Social Affairs, there are 17 goals under the Sustainable Development Goals (SDGs). In line with the statistics for economy, SDG8 or the goal number 8 states “Promote sustained, inclusive and sustainable economic growth, full and productive employment and decent work for all. As of 2024 sustainable development report dashboard, (<https://dashboards.sdgindex.org/map/goals/SDG8>) it is indicated that the Philippines remain on facing major challenges. Indicators supporting the goal are similarly marked as facing major challenges also for the Philippines.

STATEMENT OF THE PROBLEM

Specifically, the scholar aims to conduct a data engineering workflow to deliver clean and understandable data for further analytics as the data has to be presented to the general public. The said workflow will be in a form of a relational database

For economic statistics, the data will focus on manufacturing industry. Available records are for years 2015, 2016 and 2017 only. The summary statistics gathered for the relational database will be based on the **Annual Survey of Philippine Business and Industry** available in the PSA OpenSTAT website.

Furthermore, the statistics are classified into following:

- Industry Description
- Year of observation
- Employment size
- Data Item (Statistical Measure)

METHODOLOGY

Data Integration Approach

Due to accessibility presented in the OpenSTAT website, the **ELT (Extract, Load, Transform)** was used to be the data processing method.

For extracting data, the database is in PXWeb, a web-based tool used for statistical data. The tool provides the statistical data in granulated columns so that the user could easily select the information based on the desired variable and corresponding statistics.

Summary Statistics for Manufacturing Establishments by Industry Description, Size and Data Item: 2015

Select variable

About table

Mark your selections and choose between table on screen and file format. Marking tips
For variables marked * you need to select at least one value

| | |
|--|---|
| <div>Industry Description *</div> <div><div><div></div><div></div><div></div><div></div></div></div> <div>Total 415 Selected: 0</div> <div>Manufacturing Slaughtering and meat packing Production processing and preserving of meat and meat products Canning/packing of fish and other marine products Drying of fish and other marine products Smoking of fish and other marine products</div> <div>Search <input type="text"/></div> <div><input type="checkbox"/> Beginning of row</div> | <div>Year *</div> <div><div><div></div><div></div><div></div><div></div></div></div> <div>Total 1 Selected: 0</div> <div>2015</div> <div>Search <input type="text"/></div> <div><input type="checkbox"/> Beginning of row</div> |
| <div>Size *</div> <div><div><div></div><div></div><div></div><div></div></div></div> <div>Total 3 Selected: 0</div> <div>All Employment Sizes Total Employment of 20 and Over Total Employment of Less than 20</div> <div>Search <input type="text"/></div> <div><input type="checkbox"/> Beginning of row</div> | <div>Data Item *</div> <div><div><div></div><div></div><div></div><div></div></div></div> <div>Total 14 Selected: 0</div> <div>Number of establishments Total employment Paid employees Total income Total expense Compensation</div> <div>Search <input type="text"/></div> <div><input type="checkbox"/> Beginning of row</div> |

Number of selected data cells are: 0 (maximum number allowed is 100,000)

Presentation on screen is limited to 1,000 rows and 30 columns

Table - Layout 1

Sample database using the PXWEB tool for data collection

The said platform can extract the necessary data and can be presented into tabulated datasets, graphs and charts. Additionally, the exported data are also available in various file format such as comma-separated values, excel spreadsheet file, etc.

Metadata

Attached to the database are the definitions regarding information about the statistical measure or data item. The data items are the fields of the dataset gathered.

Metadata

Metadata on Mining, Manufacturing, Construction

Summary Statistics for Manufacturing Establishments by Size, Industry Description, Year and Data Items

Table ID: 2G4CASC0

| | |
|--------------------------------|--|
| Indicator/Variable | Number of Establishments |
| Definition | An establishment is defined as an economic unit, which engages, under a single ownership or control, i.e. under a single legal entity, in one or predominantly one kind of economic activity at a single fixed physical location. |
| Method of Computation | 2015 ASPBI Survey Design (PDF) |
| Source of Basic Data | 2015 Annual Survey of Philippine Business and Industry (ASPBI) |
| Source | Philippine Statistics Authority |
| Document/Agency | Philippine Statistics Authority |
| Frequency of Release | Annually |
| Latest Available | 2015 |
| Level of Disaggregation | Employment Sizes, National/Regional |
| Contacts | |
| Last Updated | |
| Indicator/Variable | Total employment |
| Definition | Total Employment refers to the total number of persons who work in or for the establishment as of 15 November of the reference year. This includes paid employees, working owners, unpaid workers and all employees who work full-time or part-time including seasonal workers. Included also are persons on short term leave such as those on sick, vacation or annual leaves and on strike. Excluded from the count of total employment are consultants, home workers, workers receiving pure commission only, workers on indefinite leave, silent or inactive partners and members of cooperative who are not involved in the operation of the cooperative. Paid employees are all full-time and part-time employees working in or for the establishment and receiving regular pay, as well as those working away from the establishment and paid by and under the control of the establishment. |

General Information

About the OpenSTAT Metadata

Statistical Laws and Other Legal Provisions

Philippine Statistical System (PSS)

Philippine Statistical Development Program (PSDP)

PSA Structure and Organization

System of Designated Statistics

Inventory of Statistical Standards in the Philippines

Classification Systems

Specific Metadata

Open Data Product Agreement

Release Calendar and Updates

Contact Information

For data inquiries, contact:

Knowledge Management and Communications Division
Philippine Statistics Authority
5/F C/VEA Bldg., East Avenue,
Quezon City

<https://openstat.psa.gov.ph/Metadata/2G4CASC0>

The values were in thousand pesos; except for the industry description and employment data. In addition, footnotes for the database stated the following:

- “s” – suppressed data
- “c” – combined data
- “..” – unavailable data

Note that the rows containing these data are to be omitted as these does not help with the summary statistics.

Data Extraction Proper

Summary Statistics for Manufacturing Establishments by Industry Description, Size and Data Item: 2015

[Select variable](#) [About table](#)

Mark your selections and choose between table on screen and file format. Marking tips
For variables marked * you need to select at least one value

| | | |
|---|--|---|
| Industry Description * ✓ - + - + Total 415 Selected 0 Manufacturing Slaughtering and meat packing Production processing and preserving of meat and meat products Canning/packing of fish and other marine products Drying of fish and other marine products Smoking of fish and other marine products Search <input type="text"/> <input type="button" value="▶"/> <input type="checkbox"/> Beginning of row | | Year * ✓ - + - + Total 1 Selected 0 2015 Search <input type="text"/> <input type="button" value="▶"/> <input type="checkbox"/> Beginning of row |
| Size * ✓ - + - + Total 3 Selected 0 All Employment Sizes Total Employment of 20 and Over Total Employment of Less than 20 Search <input type="text"/> <input type="button" value="▶"/> <input type="checkbox"/> Beginning of row | Data Item * ✓ - + - + Total 14 Selected 0 Number of establishments Total employment Paid employees Total income Total expense Compensation Search <input type="text"/> <input type="button" value="▶"/> <input type="checkbox"/> Beginning of row | |

Number of selected data cells are: 0 (maximum number allowed is 100,000)

Presentation on screen is limited to 1,000 rows and 30 columns

Table - Layout 1

Continue

The simplest extraction for the database here, is aggregating into two “Size(s)”, the **Total Employment of 20 and over**, and the **Total employment of less than 20**. In this situation, if the dataset hasn’t separated data for various years, my extraction will be based on the annual reports before aggregating into employment sizes.

Hence the extraction process was done by six times to six CSV files. Two extractions were done per year as per employment size, and the years **2015, 2016 and 2017** were involved.

Loading Data

Using PostgreSQL, the initial loading of data takes place there. The column names were renamed due to complexity or length. The following are the newly assigned column names and the actual metadata:

- **industry_desc** : Industry Description
- **est_number** : Number of establishments
- **total_employment** : Total employment
- **paid_employees**: Paid employees
- **total_income**: Total income
- **total_exp**: Total Expense
- **compensation**: Compensation
- **other_exp**; Other expense
- **value_of_output**: Value of Output
- **intermediate_exp**: Intermediate Expense
- **value_added**: Value added
- **tf_assets**: Gross addition to tangible fixed assets
- **change_inv** : Change in inventories
- **subsidy**
- **commerce_transaction** : Sales from e-commerce transactions

```
Query  Query History
1
2  -----
3  --CHECK IF TABLE ALREADY EXISTS--
4  -----
5  DROP TABLE IF EXISTS large_manufacturer_2015;
6  DROP TABLE IF EXISTS small_manufacturer_2015;
7  DROP TABLE IF EXISTS large_manufacturer_2016;
8  DROP TABLE IF EXISTS small_manufacturer_2016;
9  DROP TABLE IF EXISTS large_manufacturer_2017;
10 DROP TABLE IF EXISTS small_manufacturer_2017;
11 -----
12 --CREATE TABLE large_manufacturer_2015--
13 -----
14 CREATE TABLE large_manufacturer_2015 (
15 industry_desc VARCHAR(512),
16 est_number VARCHAR(64),
17 total_employment VARCHAR(64),
18 paid_employees VARCHAR(64),
19 total_income VARCHAR(64),
20 total_exp VARCHAR(64),
21 compensation VARCHAR(64),
22 other_exp VARCHAR(64),
23 value_of_output VARCHAR(64),
24 intermediate_exp VARCHAR(64),
25 value_added VARCHAR(64),
26 tf_assets VARCHAR(64),
27 change_inv VARCHAR(64),
28 subsidy VARCHAR(64),
29 commerce_transaction VARCHAR(64)
30 );
31
32 --importing data from OpenSTAT
33 COPY large_manufacturer_2015
34 FROM 'C:\jcc\DA\Project SPARTA V5\Capstone\OpenstatPSA\EconomicStat\Manufacturing\2015\large_manufacturer_2015.csv'
35 DELIMITER ','
36 CSV HEADER;
```

Part of the query in importing a dataset.

As for the footnotes, there are strings ('c', 's', '..') involved in the extracted data sets and has to be excluded. Therefore the dataset to import here are initially all in VARCHAR() datatype.


```

Query Query History
185 --Check out the industry description
186 DROP TABLE IF EXISTS industry_description CASCADE;
187 CREATE TABLE industry_description(
188   id SERIAL NOT NULL,
189   industry_desc VARCHAR(512) NOT NULL,
190   ind_code INTEGER PRIMARY KEY
191 );
192
193 INSERT INTO industry_description (industry_desc, ind_code)
194 (SELECT industry_desc, ('15' || ROW_NUMBER() OVER()):: INTEGER AS ind_code
195 FROM large_manufacturer_2015)
196 ;
197
198 INSERT INTO industry_description (industry_desc, ind_code)
199 (SELECT industry_desc, ('16' || ROW_NUMBER() OVER()):: INTEGER AS ind_code
200 FROM large_manufacturer_2016)
201 ;
202
203 INSERT INTO industry_description (industry_desc, ind_code)
204 (SELECT industry_desc, ('17' || ROW_NUMBER() OVER()):: INTEGER AS ind_code
205 FROM large_manufacturer_2017)
206 ;
207
208 SELECT industry_desc, COUNT(*) FROM industry_description
209 GROUP BY ind_code
210 ORDER BY COUNT(*) DESC;
211
212 --Create YEAR Table
213 DROP TABLE IF EXISTS year;
214 CREATE TABLE year(
215   year integer PRIMARY KEY);
216
217 INSERT INTO year VALUES (2015);
218 INSERT INTO year VALUES (2016);
219 INSERT INTO year VALUES (2017);

```

Creation of keys for future referencing

The query shown is necessary for creating primary and foreign keys of the respective datasets. NOT NULL statements are used to prevent the workflow from having NULL values. The consideration of concatenating the last two digits of year of observation to the ROW_NUMBER statements so that redundancy will be avoided. The *ind_code* serves as a key has to be typecasted into INTEGER.

Cleaning Data

```

Query Query History
1 -----
2 --Deleting unnecessary data for manufacturer data--
3 -----
4
5 DELETE
6 FROM large_manufacturer_2015
7 WHERE
8   est_number IN ('c','s','.') OR
9   total_exp IN ('c','s','.') OR
10  total_employment IN ('c','s','.') OR
11  paid_employees IN ('c','s','.') OR
12  total_income IN ('c','s','.') OR
13  compensation IN ('c','s','.') OR
14  other_exp IN ('c','s','.') OR
15  value_of_output IN ('c','s','.') OR
16  intermediate_exp IN ('c','s','.') OR
17  value_added IN ('c','s','.') OR
18  r_assets IN ('c','s','.') OR
19  change_inv IN ('c','s','.') OR
20  subsidy IN ('c','s','.') OR
21  commerce_transaction IN ('c','s','.');
22
23 DELETE
24 FROM small_manufacturer_2015
25 WHERE
26   est_number IN ('c','s','.') OR
27   total_exp IN ('c','s','.') OR
28   total_employment IN ('c','s','.') OR
29   paid_employees IN ('c','s','.') OR
30   total_income IN ('c','s','.') OR
31   compensation IN ('c','s','.') OR
32   other_exp IN ('c','s','.') OR
33   value_of_output IN ('c','s','.') OR
34   intermediate_exp IN ('c','s','.') OR
35   value_added IN ('c','s','.') OR

```

Excluding the non-numerical data

As mentioned earlier about the footnotes, these data are not significant to the whole data workflow, if not excluded, summary statistics will not work as expected due to the non-numerical values in the dataset.

Transforming Data

After the cleaning proper, the fields must be able to be converted into the respective and/or necessary datatype. Recall that the values were in thousand pesos; except for the industry description and employment data. Hence we are using the INTEGER and DECIMAL as datatype. In addition the *year*, *emp_code* and *ind_code* are to be the new columns of the datasets for key-connections.

```
Query  Query History
1
2  -----
3  --Transforming Data--
4  -----
5
6  --Adding column for foreign keys
7  ALTER TABLE large_manufacturer_2015 ADD COLUMN ind_code INTEGER;
8  ALTER TABLE large_manufacturer_2016 ADD COLUMN ind_code INTEGER;
9  ALTER TABLE large_manufacturer_2017 ADD COLUMN ind_code INTEGER;
10 ALTER TABLE small_manufacturer_2015 ADD COLUMN ind_code INTEGER;
11 ALTER TABLE small_manufacturer_2016 ADD COLUMN ind_code INTEGER;
12 ALTER TABLE small_manufacturer_2017 ADD COLUMN ind_code INTEGER;
13
14 ALTER TABLE large_manufacturer_2015 ADD COLUMN year INTEGER;
15 ALTER TABLE large_manufacturer_2016 ADD COLUMN year INTEGER;
16 ALTER TABLE large_manufacturer_2017 ADD COLUMN year INTEGER;
17 ALTER TABLE small_manufacturer_2015 ADD COLUMN year INTEGER;
18 ALTER TABLE small_manufacturer_2016 ADD COLUMN year INTEGER;
19 ALTER TABLE small_manufacturer_2017 ADD COLUMN year INTEGER;
20
21 ALTER TABLE large_manufacturer_2015 ADD COLUMN emp_code VARCHAR(5);
22 ALTER TABLE large_manufacturer_2016 ADD COLUMN emp_code VARCHAR(5);
23 ALTER TABLE large_manufacturer_2017 ADD COLUMN emp_code VARCHAR(5);
24 ALTER TABLE small_manufacturer_2015 ADD COLUMN emp_code VARCHAR(5);
25 ALTER TABLE small_manufacturer_2016 ADD COLUMN emp_code VARCHAR(5);
26 ALTER TABLE small_manufacturer_2017 ADD COLUMN emp_code VARCHAR(5);
27
```

Additional columns for every dataset

```
Query  Query History
237
238 -----
239 --Key Relationships--
240 -----
241
242 ALTER TABLE large_manufacturer_2015
243   ADD CONSTRAINT fk_ind_code FOREIGN KEY (ind_code) REFERENCES industry_description(ind_code);
244 ALTER TABLE large_manufacturer_2016
245   ADD CONSTRAINT fk_ind_code FOREIGN KEY (ind_code) REFERENCES industry_description(ind_code);
246 ALTER TABLE large_manufacturer_2017
247   ADD CONSTRAINT fk_ind_code FOREIGN KEY (ind_code) REFERENCES industry_description(ind_code);
248
249 ALTER TABLE small_manufacturer_2015
250   ADD CONSTRAINT fk_ind_code FOREIGN KEY (ind_code) REFERENCES industry_description(ind_code);
251 ALTER TABLE small_manufacturer_2016
252   ADD CONSTRAINT fk_ind_code FOREIGN KEY (ind_code) REFERENCES industry_description(ind_code);
253 ALTER TABLE small_manufacturer_2017
254   ADD CONSTRAINT fk_ind_code FOREIGN KEY (ind_code) REFERENCES industry_description(ind_code);
255
256 ALTER TABLE large_manufacturer_2015
257   ADD CONSTRAINT fk_year FOREIGN KEY (year) REFERENCES year(year);
258 ALTER TABLE large_manufacturer_2016
259   ADD CONSTRAINT fk_year FOREIGN KEY (year) REFERENCES year(year);
260 ALTER TABLE large_manufacturer_2017
261   ADD CONSTRAINT fk_year FOREIGN KEY (year) REFERENCES year(year);
262
263 ALTER TABLE small_manufacturer_2015
264   ADD CONSTRAINT fk_year FOREIGN KEY (year) REFERENCES year(year);
265 ALTER TABLE small_manufacturer_2016
266   ADD CONSTRAINT fk_year FOREIGN KEY (year) REFERENCES year(year);
267 ALTER TABLE small_manufacturer_2017
268   ADD CONSTRAINT fk_year FOREIGN KEY (year) REFERENCES year(year);
269
270 ALTER TABLE large_manufacturer_2015
271   ADD CONSTRAINT fk_emp_size FOREIGN KEY (emp_code) REFERENCES employment_size(emp_code);
272 ALTER TABLE large_manufacturer_2016
273   ADD CONSTRAINT fk_emp_size FOREIGN KEY (emp_code) REFERENCES employment_size(emp_code);
274 ALTER TABLE large_manufacturer_2017
275   ADD CONSTRAINT fk_emp_size FOREIGN KEY (emp_code) REFERENCES employment_size(emp_code);
276
```

Primary Key to Foreign Key links

```

Query  Query History
1
2 -----
3 --Converting the proper data type--
4 -----
5
6 ALTER TABLE large_manufacturer_2015 ALTER COLUMN est_number TYPE INTEGER USING est_number::INTEGER;
7 ALTER TABLE large_manufacturer_2015 ALTER COLUMN total_employment TYPE INTEGER USING total_employment::INTEGER;
8 ALTER TABLE large_manufacturer_2015 ALTER COLUMN paid_employees TYPE INTEGER USING paid_employees::INTEGER ;
9 ALTER TABLE large_manufacturer_2015 ALTER COLUMN total_income TYPE DECIMAL (15,2) USING total_income::DECIMAL(15,2);
10 ALTER TABLE large_manufacturer_2015 ALTER COLUMN total_exp TYPE DECIMAL (15,2) USING total_exp::DECIMAL(15,2);
11 ALTER TABLE large_manufacturer_2015 ALTER COLUMN compensation TYPE DECIMAL (15,2) USING compensation::DECIMAL (15,2);
12 ALTER TABLE large_manufacturer_2015 ALTER COLUMN other_exp TYPE DECIMAL (15,2) USING other_exp::DECIMAL (15,2);
13 ALTER TABLE large_manufacturer_2015 ALTER COLUMN value_of_output TYPE DECIMAL (15,2) USING value_of_output::DECIMAL (15,2);
14 ALTER TABLE large_manufacturer_2015 ALTER COLUMN intermediate_exp TYPE DECIMAL (15,2) USING intermediate_exp::DECIMAL (15,2);
15 ALTER TABLE large_manufacturer_2015 ALTER COLUMN value_added TYPE DECIMAL (15,2) USING value_added::DECIMAL (15,2);
16 ALTER TABLE large_manufacturer_2015 ALTER COLUMN tf_assets TYPE DECIMAL (15,2) USING tf_assets::DECIMAL (15,2);
17 ALTER TABLE large_manufacturer_2015 ALTER COLUMN change_inv TYPE DECIMAL (15,2) USING change_inv::DECIMAL (15,2);
18 ALTER TABLE large_manufacturer_2015 ALTER COLUMN subsidy TYPE DECIMAL (15,2) USING subsidy::DECIMAL (15,2);
19 ALTER TABLE large_manufacturer_2015 ALTER COLUMN commerce_transaction TYPE DECIMAL (15,2) USING commerce_transaction::DECIMAL (15,2);
20
21 ALTER TABLE large_manufacturer_2016 ALTER COLUMN est_number TYPE INTEGER USING est_number::INTEGER;
22 ALTER TABLE large_manufacturer_2016 ALTER COLUMN total_employment TYPE INTEGER USING total_employment::INTEGER;
23 ALTER TABLE large_manufacturer_2016 ALTER COLUMN paid_employees TYPE INTEGER USING paid_employees::INTEGER ;
24 ALTER TABLE large_manufacturer_2016 ALTER COLUMN total_income TYPE DECIMAL (15,2) USING total_income::DECIMAL(15,2);
25 ALTER TABLE large_manufacturer_2016 ALTER COLUMN total_exp TYPE DECIMAL (15,2) USING total_exp::DECIMAL(15,2);
26 ALTER TABLE large_manufacturer_2016 ALTER COLUMN compensation TYPE DECIMAL (15,2) USING compensation::DECIMAL (15,2);
27 ALTER TABLE large_manufacturer_2016 ALTER COLUMN other_exp TYPE DECIMAL (15,2) USING other_exp::DECIMAL (15,2);
28 ALTER TABLE large_manufacturer_2016 ALTER COLUMN value_of_output TYPE DECIMAL (15,2) USING value_of_output::DECIMAL (15,2);
29 ALTER TABLE large_manufacturer_2016 ALTER COLUMN intermediate_exp TYPE DECIMAL (15,2) USING intermediate_exp::DECIMAL (15,2);
30 ALTER TABLE large_manufacturer_2016 ALTER COLUMN value_added TYPE DECIMAL (15,2) USING value_added::DECIMAL (15,2);
31 ALTER TABLE large_manufacturer_2016 ALTER COLUMN tf_assets TYPE DECIMAL (15,2) USING tf_assets::DECIMAL (15,2);
32 ALTER TABLE large_manufacturer_2016 ALTER COLUMN change_inv TYPE DECIMAL (15,2) USING change_inv::DECIMAL (15,2);
33 ALTER TABLE large_manufacturer_2016 ALTER COLUMN subsidy TYPE DECIMAL (15,2) USING subsidy::DECIMAL (15,2);
34 ALTER TABLE large_manufacturer_2016 ALTER COLUMN commerce_transaction TYPE DECIMAL (15,2) USING commerce_transaction::DECIMAL (15,2);

```

Notice that the ALTER queries have additional statements such as TYPE and USING.

These are significant statements to ensure the typecasting will be secured and none of the records will be completely converted into. The reason is all of these data are actually in numbers. It only differs that some of the records are in Philippine Pesos, that is why it has to be in DECIMAL data type.

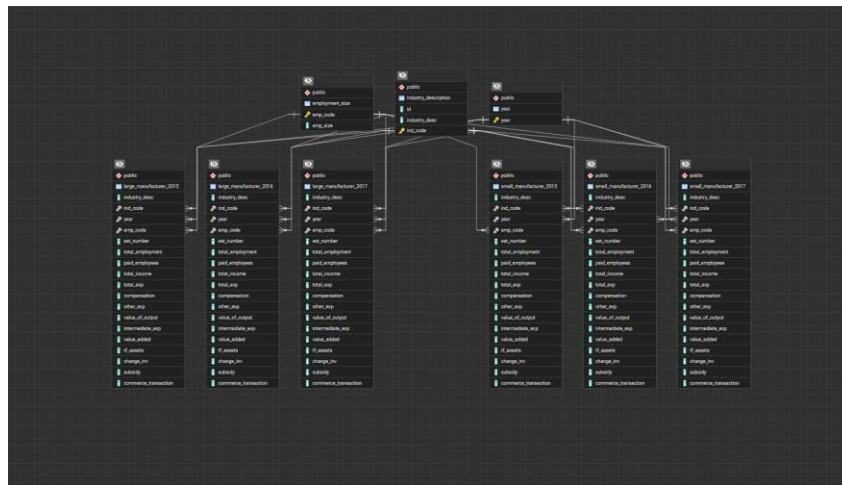
Exporting Data

```

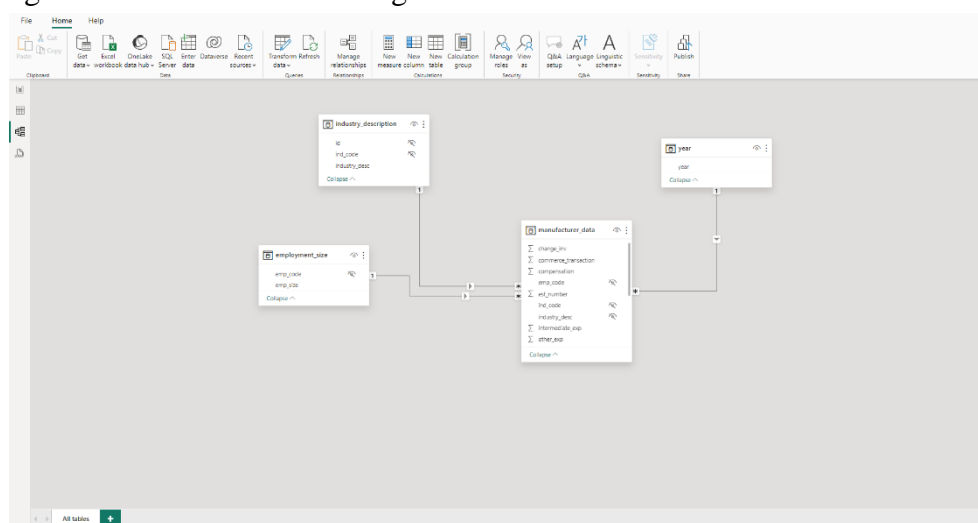
Query  Query History
1
2 -----
3 --PRODUCT ETL FOR VISUALIZATION--
4 -----
5
6 --Manufacturer Data for 2015 to 2017
7 COPY (SELECT * FROM large_manufacturer_2015
8 UNION ALL
9 SELECT * FROM small_manufacturer_2015
10 UNION ALL
11 SELECT * FROM large_manufacturer_2016
12 UNION ALL
13 SELECT * FROM small_manufacturer_2016
14 UNION ALL
15 SELECT * FROM large_manufacturer_2017
16 UNION ALL
17 SELECT * FROM small_manufacturer_2017) TO 'C:\kjc\DA\Project SPARTA Y5\Capstone\OpenstatPSA\EconomicStat\Manufacturing\manufacturer_data.csv' WITH (FORMAT CSV, HEADER);
18
19
20 --Export the primary/foreign key tables
21 COPY (SELECT * FROM year) TO 'C:\kjc\DA\Project SPARTA Y5\Capstone\OpenstatPSA\EconomicStat\Manufacturing\year.csv' WITH (FORMAT CSV, HEADER);
22 COPY (SELECT * FROM employment_size) TO 'C:\kjc\DA\Project SPARTA Y5\Capstone\OpenstatPSA\EconomicStat\Manufacturing\employment_size.csv' WITH (FORMAT CSV, HEADER);
23 COPY (SELECT * FROM industry_description) TO 'C:\kjc\DA\Project SPARTA Y5\Capstone\OpenstatPSA\EconomicStat\Manufacturing\industry_description.csv' WITH (FORMAT CSV, HEADER);
24
25

```

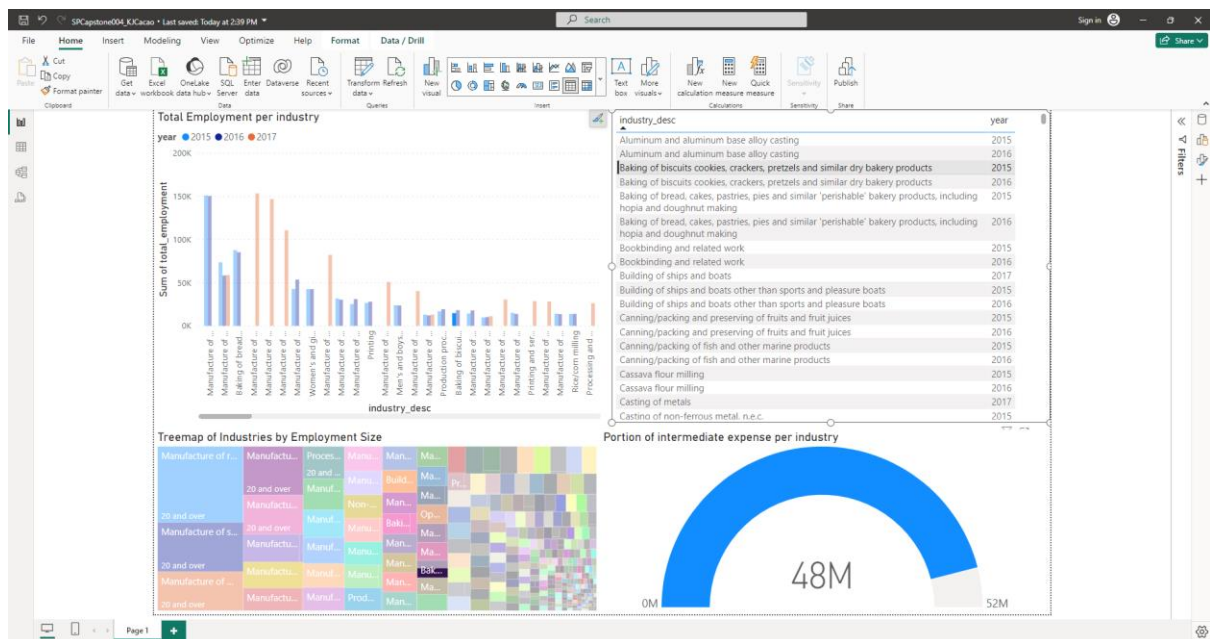
CONCLUSION



The ERD for the manufacturing data schema showcases a well-structured database design, capturing comprehensive information about large and small manufacturers across three years (2015, 2016, and 2017). The central `industry_description` table, identified by the `ind_code`, anchors the relationships with various attributes of the summary statistics. This normalization ensures data consistency and integrity, facilitating robust data analysis, reporting, and decision-making. Through this ERD, audience, or the general public can efficiently navigate and interpret the interconnected datasets, leading to more informed insights and strategies within the manufacturing sector.



The figure above shows the simplified schema made for visualization. In Power BI, a schema is essential for linking various visualizations used.



The visualization above effectively highlights the summary statistics across various industries. These figures can be filtered based on the year of observation, industry type, or employment sizes.

The following visualizations used are bar chart, table, treemap and gauge/circular chart. For summary statistics for manufacturing, these are the nearest visualization appropriate for understanding the economy for these relative sectors. The visualization is interactive, and audience may acquire various insights based on the selection of variables on the visualization.

REFERENCES:

**Summary Statistics for Manufacturing Establishments by Industry Description, Size
and Data Item: 2015-2017;**

https://openstat.psa.gov.ph/PXWeb/pxweb/en/DB/DB_2G_MNFG_ASPBI/?tablelist=true&rxid=bdf9d8da-96f1-4100-ae09-18cb3eae313; 2024

Metadata on Mining, Manufacturing, Construction;

<https://openstat.psa.gov.ph/Metadata/2G4CASC0>

Dashboards on Sustainable Development Goal 8;

<https://dashboards.sdgindex.org/map/goals/SDG8>