Business Description

PRIVATE PRE-SCHOOL EDUCATION INSTITUTIONS IN NEW ZEALAND

Business: Private pre-school education institutions in New Zealand.

Stage 1. Data search.

The main difficulty in finding data was that almost no countries keep statistics of their preschool education institutions (or at least do not put it in the public domain).

The only dataset I could find was data on New Zealand kindergartens.

Fortunately, the customer was satisfied with them, as he was interested in the possibility of building a private preschool education institution on a paid basis with competitors, the number of potential customers and their solvency.

Next, it was necessary to find data on New Zealand cities in order to have an idea of the location of existing institutions and to evaluate the most promising cities.

The found dataset contains the main cities and localities of most countries of the world and also has data on the population of cities as of the fourth quarter of 2020.

It was most difficult to find information on birth rates in New Zealand cities, as access to some data has been restricted since publication, while other data has been archived and previous links have been broken.

A similar situation was with the data on salaries, which are needed to assess the standard of living of citizens and, accordingly, their ability to pay. Since the institution is private and paid, it does not make sense to open it in an area where people will not be physically able to pay for it.

Stage 2. Processing the data, uploading it to the database.

After creating a new database and a new schema, I created tables with new columns that were already contained in CSV files, so that I could then import data into them.

But two of the four tables (“Births” and “Earnings”) in the CSV files contained columns with the same names and therefore did not open in the database. Even if you create a separate connection for the CSV and set the path to the file folder, the files still won't open.

Therefore, I had to manually change the table structure using Notepad++.

I was going to split the table with cities into two separate ones, taking out the population, but when loading the data, errors occurred, because the CSV file had commas as a separator and they were also contained in some cell values, so the row could have more attributes than it should. Again, replacing characters with Notepad++, as well as renaming a column containing a character that does not work with the cp-1521 encoding.

To fill in the timetable, a function was written that calculates the values for a given time interval, which adds them to the table.

Since at this stage of database development I didn't know that it was possible to export tables from a CSV connection without first creating tables manually, a function was written to create a Births table that contains a significant number of columns, reading the first row of the CSV file into an array and creating columns with unique names corresponding to the array elements.

But further operation with the data in this form would be inconvenient, especially in the geographical reference, so a function similar to UNPIVOT was applied to the inserted data, changing its structure.

The table with pre-school education institutions directly was created without first declaring the table structure, via export from the CSV connection. All fields were created as VARCHAR (32767). They will not take up extra space in memory (unlike CHAR), but may cause inconvenience when counting or inserting new data. But the DBMS does not allow you to change the data type of cells whose value is set to "", so a function was written that replaces the values of "" with NULL in the specified table. Columns with questionable unconfirmed conflicting information were also removed.

Initial dataset.

1.

FILE NAME
BIRTHS.CSV

DATA PROVIDED BY
Stats NZ

DATASET NAME
Live births by area, city district councils (Maori and total population) (Qrtly-MarJunSepDec)

WEBPAGE:
http://archive.stats.govt.nz/infoshare/

HOW TO FIND THE DATA
At URL provided, select 'Population > Births - VSB > Live births by area, city district councils (Maori and total population) (Qrtly-MarJunSepDec)'.

2.

FILE NAME
WORLDCITIES.CSV

DATA PROVIDED BY
SimpleMaps.com

DATASET NAME
World Cities Database (Basic)

WEBPAGE:
https://simplemaps.com/data/world-cities

HOW TO FIND THE DATA
At URL provided, select 'Download'.

3.

FILE NAME
EARNINGS.CSV

DATA PROVIDED BY
Stats NZ

DATASET NAME
QEX - Average Weekly Earnings (Employees) by Regional Council and Sex (Qrtly-MarJunSepDec)

WEBPAGE:
http://archive.stats.govt.nz/infoshare/

HOW TO FIND THE DATA
At URL provided, select 'Work income and spending > Earnings and Employment Survey (QES) - QEX > QEX - Average Weekly Earnings (Employees) by Regional Council and Sex (Qrtly-MarJunSepDec)'.

4.

FILE NAME
KINDERGARTENS.CSV

DATA PROVIDED BY
Ministry of Education of New Zealand

DATASET NAME
Early Childhood Services (ECE) Directory

WEBPAGE:
https://catalogue.data.govt.nz/dataset/directory-of-educational-institutions

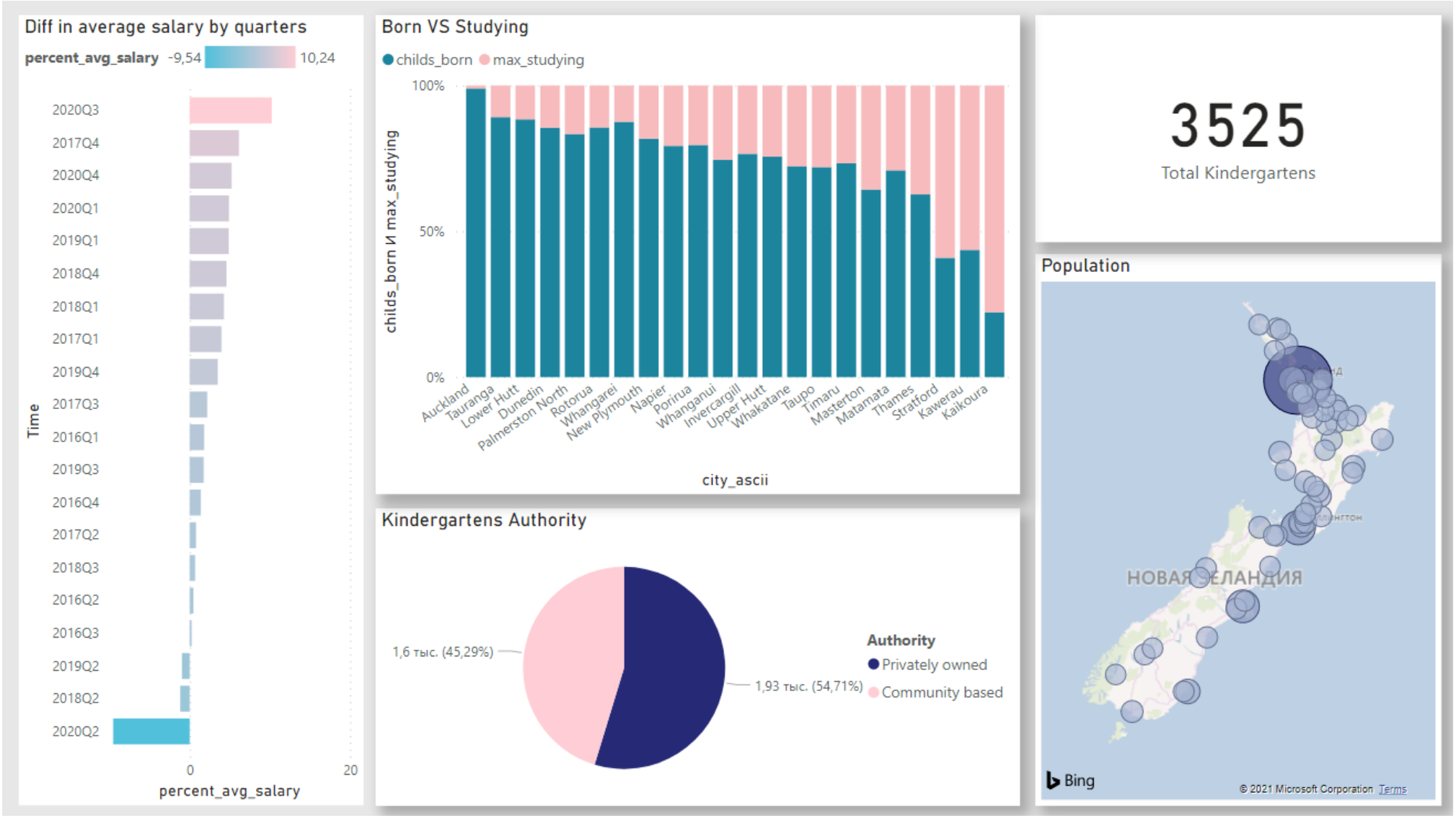
HOW TO FIND THE DATA
At URL provided, select ' Early Childhood Services (ECE) Directory '.

The purpose.

The main purpose of the report was to identify the most promising places for opening a private preschool education institution.

To do this, we need to:

- ✓ analyze the correspondence of the number of preschool institutions in different cities of New Zealand to the number of children aged 0 to 5 years living there; the priority is given to those cities where there is a shortage of educational places (and not 3 places in kindergarten per child) (*VIEW fp.analysis_of_free_places*);
- ✓ analyze the dynamics of the birth rate of children in cities; give preference to cities where the percentage of newborns is growing every year (*VIEW fp.births_analyse*);
- ✓ estimate the standard of living of the local urban population based on the average quarterly salary, because it is not advisable to open an institution in poor regions, even if there is a shortage of educational places (*VIEW fp.salary_analyse*);
- ✓ evaluate the overall ratio of preschool education institutions to the resident population (*VIEW fp.population_kindergarten_analyse*).



Pic.1. Visualization Result