# **Creating Cross Tables in Vivintel**

## Introduction

Welcome to The Library's video on creating cross tables in Vivintel, formerly known as Vividata. Vivintel is a data analysis platform that gives you access to survey data across a range of consumer products and service sectors.

By the end of this video, you will be able to:

- Navigate the Vivintel interface
- Create a cross table
- Analyze demographic distributions
- And interpret key metrics such as Column %, Row % and Index

#### Navigating the Interface

Once logged in, the Demo Project loads automatically. This interface consists of categories on the left and the cross table builder in the center of the screen. The interface resets each time you sign-in to Vivintel, so be sure to export your data before leaving.

A cross table, or cross tab, is how Vivintel displays the relationship between two or more variables from a dataset. The default dataset includes all individuals in Canada aged 14 and older.

At the time of recording, the most recent dataset is Fall 2024. This study represents approximately 35 million Canadians. This figure is calculated by adding three zeroes to the population numbers displayed in the totals column.

The sample size refers to the number of survey participants for a given category. The Fall 2024 study surveyed over 54,000 people.

#### Selecting a Study

To change the dataset, use the Study or "Waves" drop-down menu in the upper right corner of the page. In most cases, selecting the most recent study is recommended. For this tutorial, we will be using the Fall 2024 study.

## **Building a Cross Table**

Now we'll explore how to build a cross table. For our example, let's say we want to analyze fitness club membership by age group, or generation.

We will begin by selecting and organizing variables in the table builder.

Expand the Demographics category in the Variables Tree, then expand Age and Generation.

You can select individual generations by checking them one at a time or selecting all generations under their variable heading.

Next, drag the selected generations into the Row section in the upper left of the cross table builder and choose the input method.

For our example, selecting the "And" input method would not be appropriate, as it would attempt to find individuals who belong to multiple generations simultaneously, which is not possible. Instead, select the "Separate" option.

Now that the generational groups have been added as separate rows, we can view their population breakdowns.

For example, we can see that Gen Z in Canada has a total population of 6,794,000.

The "Column %" metric indicates the share of a category within the total population. For instance, the column percentage for Gen Z shows that they make up roughly 19% of the Canadian population over the age of 14.

If we navigate to metrics options, indicated by a gear icon, and hide all other data in the cell, except for column percentage. We can see that the column percentage from all generations totals to 100%.

#### Adding a Behavioural Variable

Let's continue with our example of fitness club memberships across generations, but now, let's add another variable to our table.

Instead of expanding categories manually, we can use the Search bar in the Variables Tree to find a specific category. Searching for "fitness club" will generate a refined list of variables.

Expand the most relevant category and check the variable corresponding to individuals who answered "Yes" to personally belonging to a fitness club.

Drag this variable into the Column section to display its relationship to the generational cohorts.

## Understanding the Results

With this new data added, we can now compare fitness club membership rates across different generations.

The Column % for each generation among those who answered "Yes" to belonging to a fitness club differs from the Column % representing each generation's share of the total population. This suggests that some generations are overrepresented or underrepresented in fitness club memberships relative to their total population size. Marketers may also use the terms "over indexed" or "under indexed" when referring to these relationships.

For example, although Gen Z comprises 19.36% of the total population, they account for 31.71% of fitness club members. This means they are overrepresented, that is, they over index, as fitness club members compared to their population share.

While Column % shows the total share of fitness club members from each generation, it does not indicate the proportion of each generation that belongs to a fitness club.

To see this, navigate to the metrics options, indicated by a gear icon, to unhide the "Row %" metric.

The Row % metric shows what proportion of each generational subgroup belongs to a fitness club. This is useful for understanding behaviors within specific demographics.

For example, when Row % is enabled, we can see that 31.2% of Gen Z belongs to a fitness club.

## Using the Index Metric

Another useful metric is Index, which helps determine whether a group is more or less likely than the general population to engage in a behavior.

The index is a measure of likelihood compared to baseline. If you take away one skill from this section of the video, it should be that a high index signals a high likelihood of variables appearing together.

A number between 90 and 110 means that the likelihood of a person having the qualities in both the row and column is about what you would expect based on their prevalence in the population. In our example, the index is 164, which means that the pairing of Gen Z and being a fitness club member is 64% more likely than baseline. What is baseline? well, in this case it is members of Gen Z. Reading the index in this way is a great method for identifying your target market.

You can calculate the index yourself using the expected frequency and the observed frequency. In this case, our baseline is the percent column for Gen Z, which is 19.36%, and the observed frequency of fitness club membership for Gen Z is 31.71%. If there was no association between age and fitness club membership, we would expect these percentages to match. But they don't. Dividing 31.71% by 19.36% and multiplying by 100 calculates the index of 164. We might not know why being younger is linked with fitness club membership but based on this index we know that it is.

To recap, the larger the index, the larger the likelihood. So, if you are trying to identify your target market, you are looking for an index above 110. Conversely, a low index means that the pairing of the variables is less likely. For example, the index for Gen X is 76 so they are 24% less likely than baseline to belong to a fitness club. Another way that people phrase these relationships is with the use of under index and over index. You can say that Gen Z over indexes on fitness club membership and Gen X under indexes.

It might seem confusing that the column percentage is higher in Gen Y than in Gen Z. What this means is that Gen Y makes up a larger proportion of fitness club members, and they also over index. Even so, they don't over index as much as Gen Z. In this case, both would be part of the target market for anyone marketing to fitness club members.

In summary, a high index signals a high likelihood that these variables appear together more often than you would expect, helping you to identify your target market.

# Conclusion

This concludes our video on getting started with Vivintel, formerly known as Vividata.

This video demonstrated how to navigate the Vivintel interface, create a cross table, analyze demographic distributions, and interpret key metrics such as Column %, Row %, and Index.

If you have any questions, visit library.humber.ca to connect with a Library team member.