

Marketing Analytics Crosstabs

PI2 Team: Goldie Zhu, Feiyang Fan, Ashley Yuan, Montgomery Gole, Thomas Cooney | Client: Polaris Intelligence | Supervisor: Dr. Michael Gruninger

Problem Definition

Polaris Intelligence's current application already has many functionalities but it lacks a simple and straightforward way for companies to be able to narrow down the scope of their analytics by selecting multiple variables.

Our solution will allow Polaris' customers to select and view data for two variables in a cross table for an aggregated selection of postal codes. This module will efficiently query and display searches in a few seconds while also storing previous queries for easy retrieval and recurrent review.

Written using a combination of PHP, React, Javascript, HTML, and CSS, it will be integratable into Polaris' existing tech stack. Our scalable solution does not include changing Polaris' existing services, softwares, and data.

Stakeholders

Stakeholders in Polaris include Consumer Packaged Goods (CPG) companies, municipalities, educational institutions, and other small businesses.

Product

Stakeholders	Interests	Impact on Functional Requirement
Existing customers	The data table is easy to use. Labels are specific with details. Necessary colour code is used. The format of the solution is consistent with other Polaris data tables. Price is reasonable.	Keep the cost of the solution below a certain amount. The solution does not need lots of money to build or maintain. The format needs to be consistent with existing tables.
Prospective customers	Polaris provides unique data tables that other companies cannot provide. Price is reasonable.	Inexpensive solution that provides new information for all industries.
Polaris software data engineering team	Easy to ingest the data. Easy to update data. Code is clearly written. Instruction and wiki include all the aspects of the solution.	Use the same coding format Polaris is using now. Write a detailed wiki page for future reference. Use the consistent coding language. Write tests that cover all the functions so future updates can run them as reference. Use reasonable names for variables.
Polaris accounts team	Need to understand the solution so they can show clients.	The tables are easy to explain. Maybe provide an adaptor/mock version for the accounts team to learn and show customers.
Polaris data science team	Data is very clean and easy to read.	An accessible and reliable design.

Example Result

postalcode	populationage014	westasian	Total Of Variables	Selected Total (population15total)
L6V0A2	16.14	0	16.14	3
L6Z0A1	12.6	0.0009	12.6009	3
L6Z0A2	12.6	0	12.6	7
L6Z0A3	11.98	0.2055000000000004	12.18550000000001	15
L6Z0A7	15.54	0	15.54	3
L6Z0A9	16.12	0	16.12	3
L6Z0B1	15.35	0	15.35	11
L6Z0B2	13.95	0.0504	14.00039999999999	3
L6Z0B4	19.5	0.0222	19.5222	3
L6Z0B8	13.2	0.748	13.94799999999999	85
L6Z0B9	11.57	0	11.57	9
L6Z0C1	13.95	0.4008	14.3508	24

Example of a Crosstab

Performance

Using Apachebench, we tested the performance of our module with various amounts of concurrent requests to a maximum level of 40.

Concurrency Level:	40
Time taken for tests:	10.501 seconds
Complete requests:	500
Failed requests:	0
Total transferred:	201010500 bytes
HTML transferred:	200929500 bytes
Requests per second:	47.61 [#/sec] (mean)
Time per request:	840.078 [ms] (mean)
Time per request:	21.002 [ms] (mean, across all concurrent requests)
Transfer rate:	18693.43 [Kbytes/sec] received

Based on the "Time per request" results, the worst case performance for each request from 40 concurrent users is handled in 840ms. The best performance is for one concurrent user request handled in 21ms. As such, our design is proven to meet the concurrent load of forty users and the performance requirements of less than two seconds.



UNIVERSITY OF
TORONTO

