

Перепись как основа для бизнес-аналитики

ЕЭК ООН – Неделя переписи – сентябрь 2022 г.

Джон Неллиган

Технический директор
Энвироникс Аналитика
Торонто, Онтарио
Канада

Перепись как основа для бизнес-аналитики

Вы не можете улучшить то, что не можете измерить

Вы не можете измерить что-либо без мерила

Вы не можете ссылаться на что-либо или сравнивать что-то с чем-то без стандартизации

Нельзя делить без знаменателя

Существует огромное количество данных, которые можно использовать с большей пользой и получать от этого ценные результаты, если спроецировать их в рамки, установленные переписью.

Расширение охвата переписи на новые продукты

В этой презентации я расскажу о следующих продуктах компании «Энвироникс Аналитика» (Environics Analytics (EA)), которые используют рамки переписи (Census Framework) для создания новой ценности:

- Актуализация оценок численности населения и размера доходов
- Расширение потребительской корзины новыми продуктами
- Готовая коммерческая система сегментации
- Расширение использования данных опросов, собранных для измерения СМИ
- Применение атрибуции и аналоговых методов к другим данным обследования
- Пассивно собираемые данные сетевого отслеживания и большие данные
- Расширение экосистемы технологий рекламы (Adtech)

- Социально-экономические концепции: население, домохозяйства, жилища, рабочая сила.
- Структура жилища и типовая модель расселения: тип жилья, миграция, плотность городского населения.
- Географическая структура отчетности: географическая единица отчетности.

Создание новых данных о малых территориях с использованием рамок переписи

Геодемографический подход основан на 3 идеях:

1. Атрибуты/характеристики людей/домохозяйств, которые находятся близко друг к другу, более похожи, чем те, которые находятся далеко друг от друга (*положительная пространственная автокорреляция*).
2. Знание того, кто проживает по соседству с человеком (в небольшом пространственном масштабе) помогает сделать правильные выводы об этом человеке.
3. Когда соблюдение конфиденциальности является критически важным аспектом, можно использовать атрибуты небольших территорий в качестве обоснованных оценок атрибутов жителей этих небольших территорий.

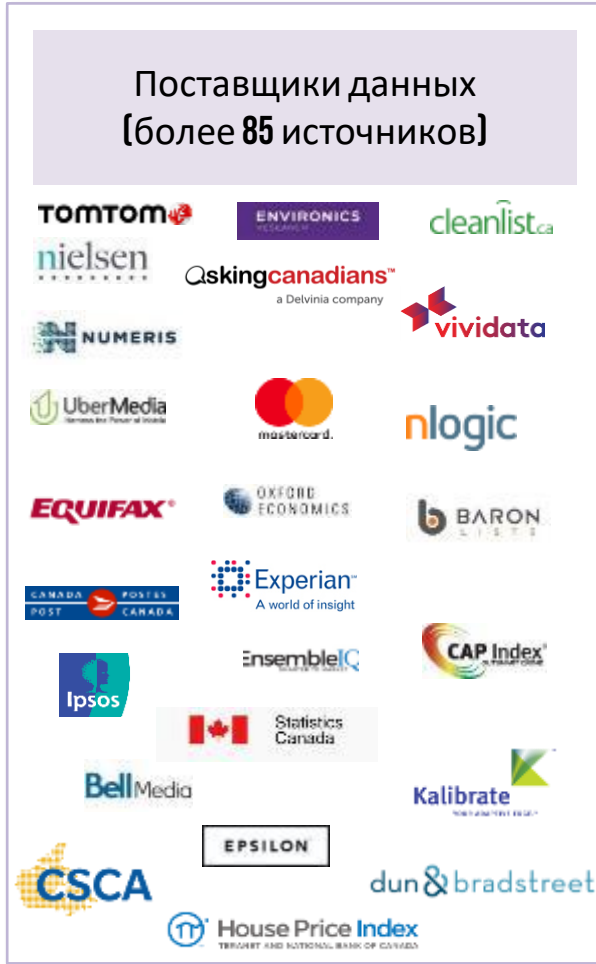
Построено на базе геодемографии

- Сочетание географии и демографии
- Используются географические коды (например, почтовые коды или индекс +4 цифры) в качестве уникального индикатора для объединения и обобщения больших объемов данных.









Используя многие источники, применяя лучшие методы

Поставщики данных
(более 85 источников)



Геодемография
Метод ближайших соседей (KNN)
Микро-симуляция
Предиктивная аналитика
Типологическая проекция
Машинное обучение
Бенчмаркинг (сравнение)
Общая контрольная калибровка
Кластеризация
Нормализация
Бутстрэпинг

-  Demographic
-  Segmentation
-  Psychographic
-  Behavioural
-  Financial
-  Mobile Analytics

50 000 переменных данных, доступных на уровне 6-значного почтового индекса

На базе географических рамок переписи

Поставщики данных
(более 85 источников)




Геодемография

Метод ближайших соседей (KNN)
Микро-симуляция
Предиктивная аналитика
Типологическая
Машинное обучение
Общая контрольная
калибровка
Кластеризация
Нормализация
Бутстрэпнинг

 Demographic

 Segmentation

 Psychographic

 Behavioural

 Financial

 Mobile Analytics

50 000 переменных данных, доступных на уровне 6-значного почтового индекса © 2022 Environics Analytics

Делая обширные данные обследования полезными для жителей района

На базе коэффициента

- Присвоить респондентам обследования код сегмента (PRIZM) на базе их почтового индекса.
- Составить таблицу численности людей или домохозяйств, которые покупают продукт на уровне национального сегмента. Так, если в 5-м сегменте продукт приобретут 672 931 человек из 2 301 764, то можно сделать прогноз, что покупку этого продукта совершат 29,2% людей.
- Следовательно, все почтовые индексы 5-го сегмента имеют одинаковую склонность к совершению покупки независимо от географического положения.

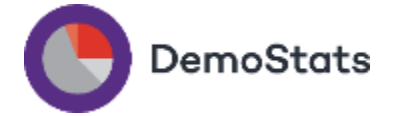
На базе модели

- Данные моделируются непосредственно на уровне почтовых индексов с использованием демографических данных респондента и демографических характеристик района.
- Модели могут быть ограничены по региону их применения.
- Прямого отношения к PRIZM нет.
- Теоретически каждый почтовый индекс получает уникальный набор свойств.

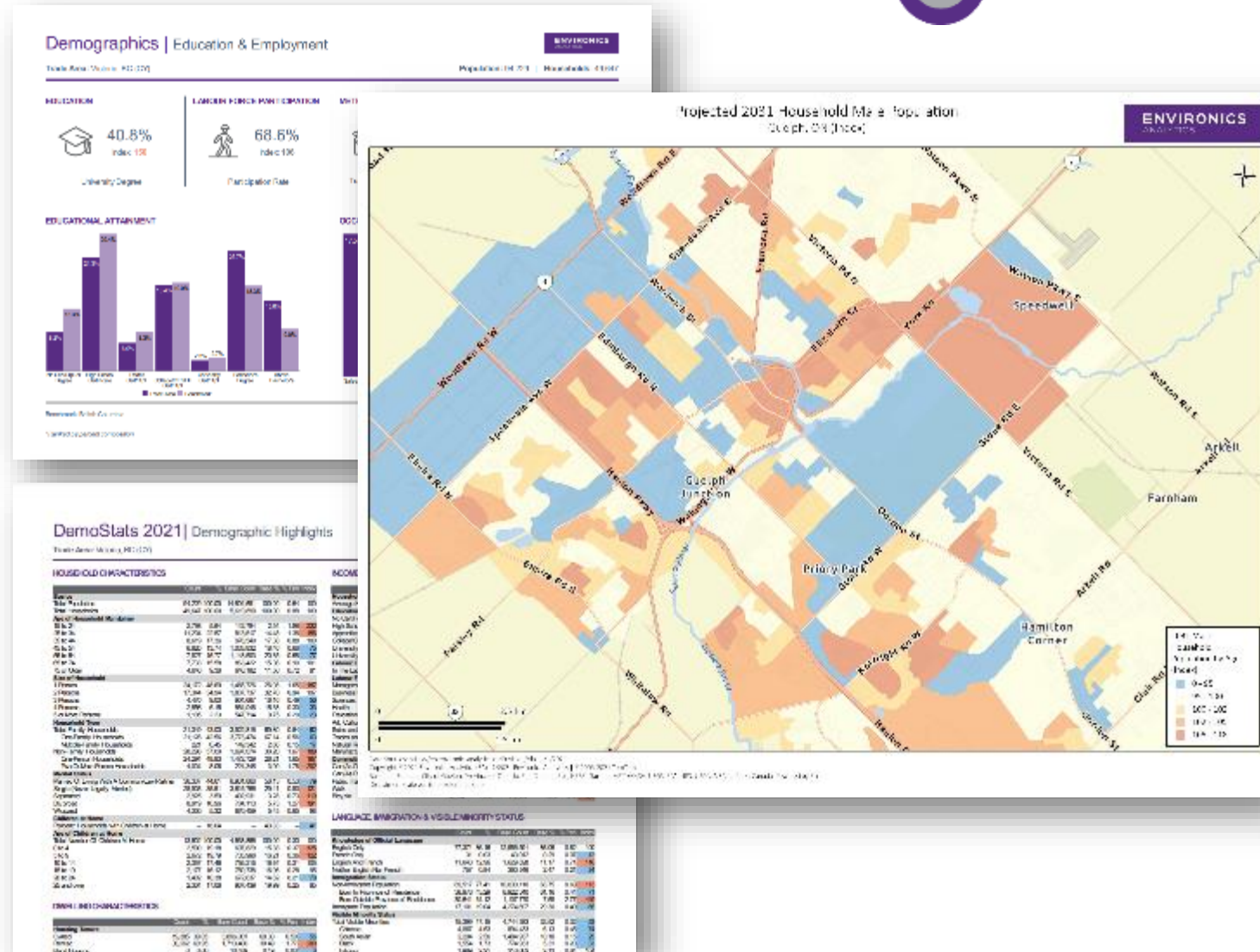
Геодемография – почему она все еще работает

- Использует данные переписи, поэтому их можно использовать во многих юрисдикциях (странах).
- Привязка к земле, а не к человеку.
- Соблюдение конфиденциальности личных данных.
- Легко назвать, описать и использовать.
- Сегменты имеют смысл для большинства людей как способ описания реальности.
- Разработано для детализации, поэтому возможно объединение в более крупные пользовательские группы.
- Возможность объединения разрозненных баз данных – большое преимущество.
- Определение размера рынка на небольшом географическом уровне – большое преимущество.
- Используются новые типы данных.

Начало работы с демографическими данными малых территорий



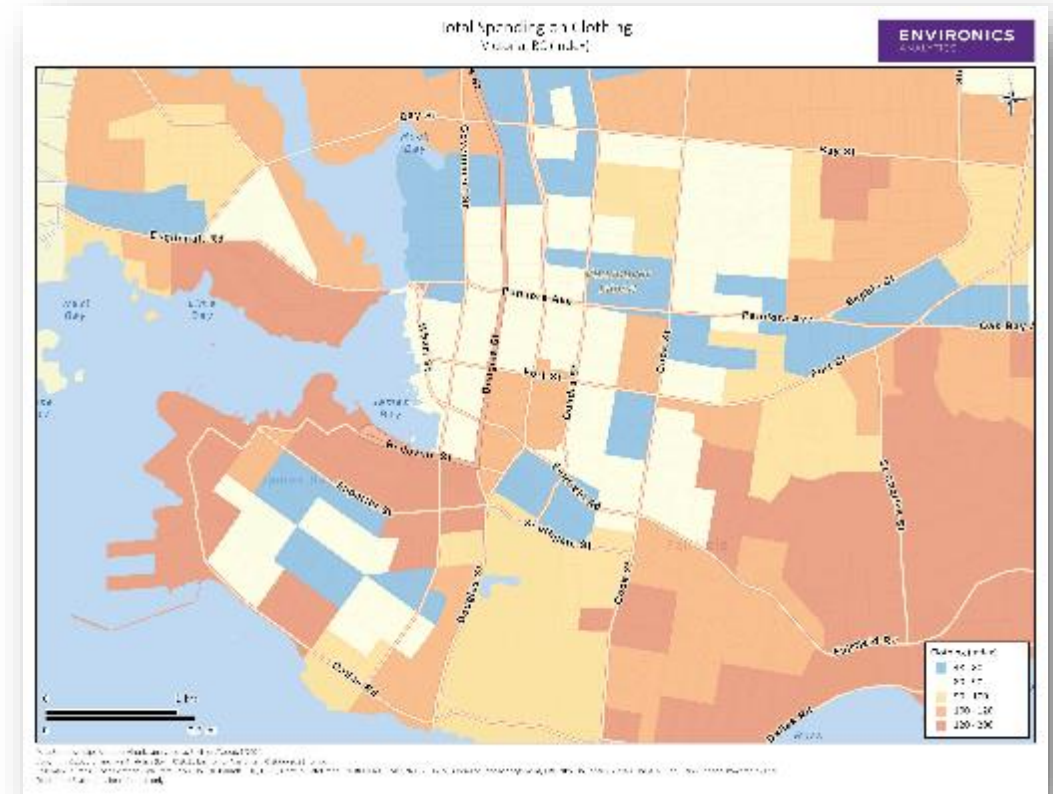
- Оценки на текущий год в отношении более чем 750 человек: образование, культурное разнообразие и уровень доходов
- Создано на основе данных Environics Analytics, Статистического управления Канады, Oxford Economics, Equifax, СМНС и Canada Post (Почта Канады), а также путем сочетания эконометрических, демографических и географических моделей.
- В наличии на уровне почтового индекса.



Добавление потребительской корзины товаров



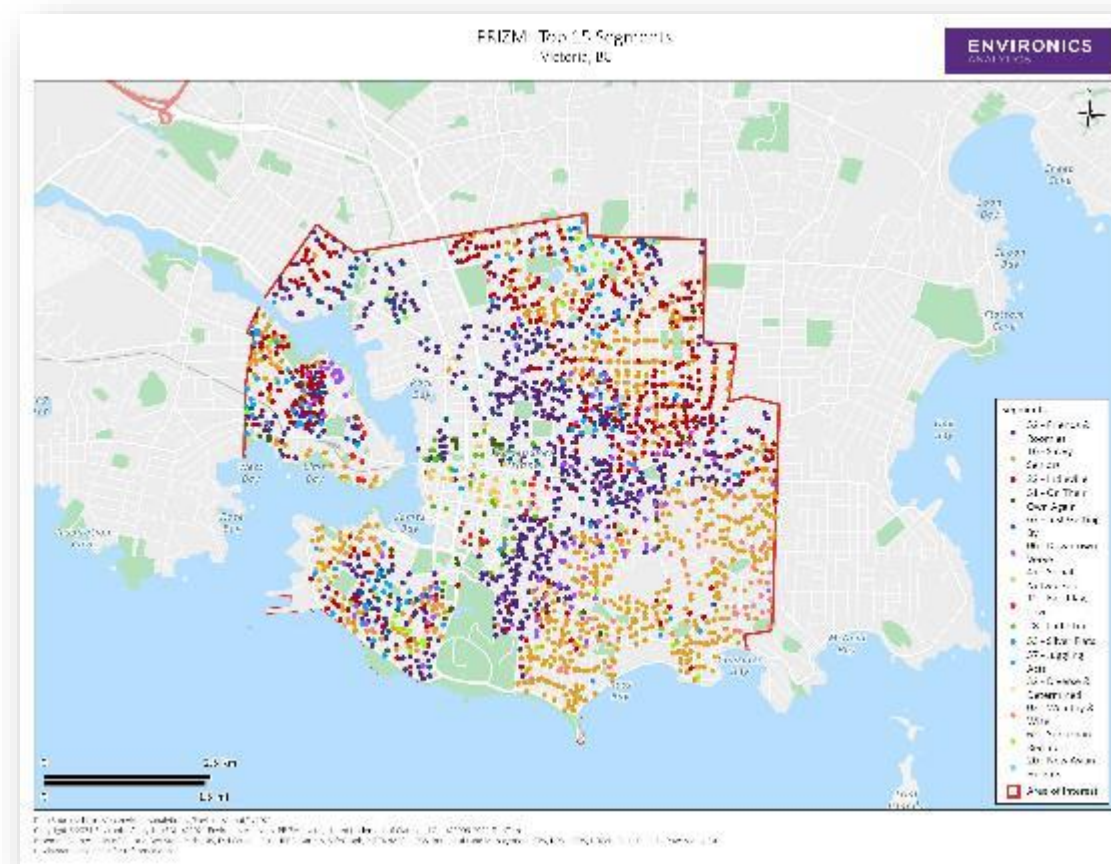
- Модуль «Расходы домохозяйств» (HouseholdSpend) предоставляет оценки ежегодных расходов почти для 500 переменных по 18 категориям товаров и услуг, потребляемых канадскими домохозяйствами.
- Обычно это используется для определения тенденций совершения покупок в торговых зонах, для выявления возможностей перекрестного продвижения и определения потенциала расходов в разбивке по категориям.
- Это позволяет оценивать расходы как в среднем в долларах на домохозяйство, так и в долларах, потраченных на любом географическом уровне.
- В процессе подготовки используются данные DemoStats, PRIZM™ и Обследования расходов домохозяйств Статистического управления Канады.
- Приведено в соответствие с контрольными показателями, предоставленными Канадской системой национальных счетов и Статистическим управлением Канады.



Разделение территории на однородные сегменты



- PRIZM® — это наша новаторская система сегментации, которая классифицирует районы Канады по 67 уникальным типам образа жизни.
- Объединяет данные почти из дюжины географических, демографических, медийных и психографических источников.
- Является основой для создания пользовательских сегментов для различных потребностей клиентов.
- Возможность делать привязку к более чем 30000 поведенческих переменных, чтобы помочь вам лучше анализировать, понимать и находить клиентов и рынки.
- Уникальное присвоение на уровне 6-значного почтового индекса или на уровне территории распространения результатов переписи.



PRIZM Segments



Индикатор социально-экономического положения (SESI)
 Название сегмента
 Изображение сегмента

Социальная группа
 Группа этапа жизни

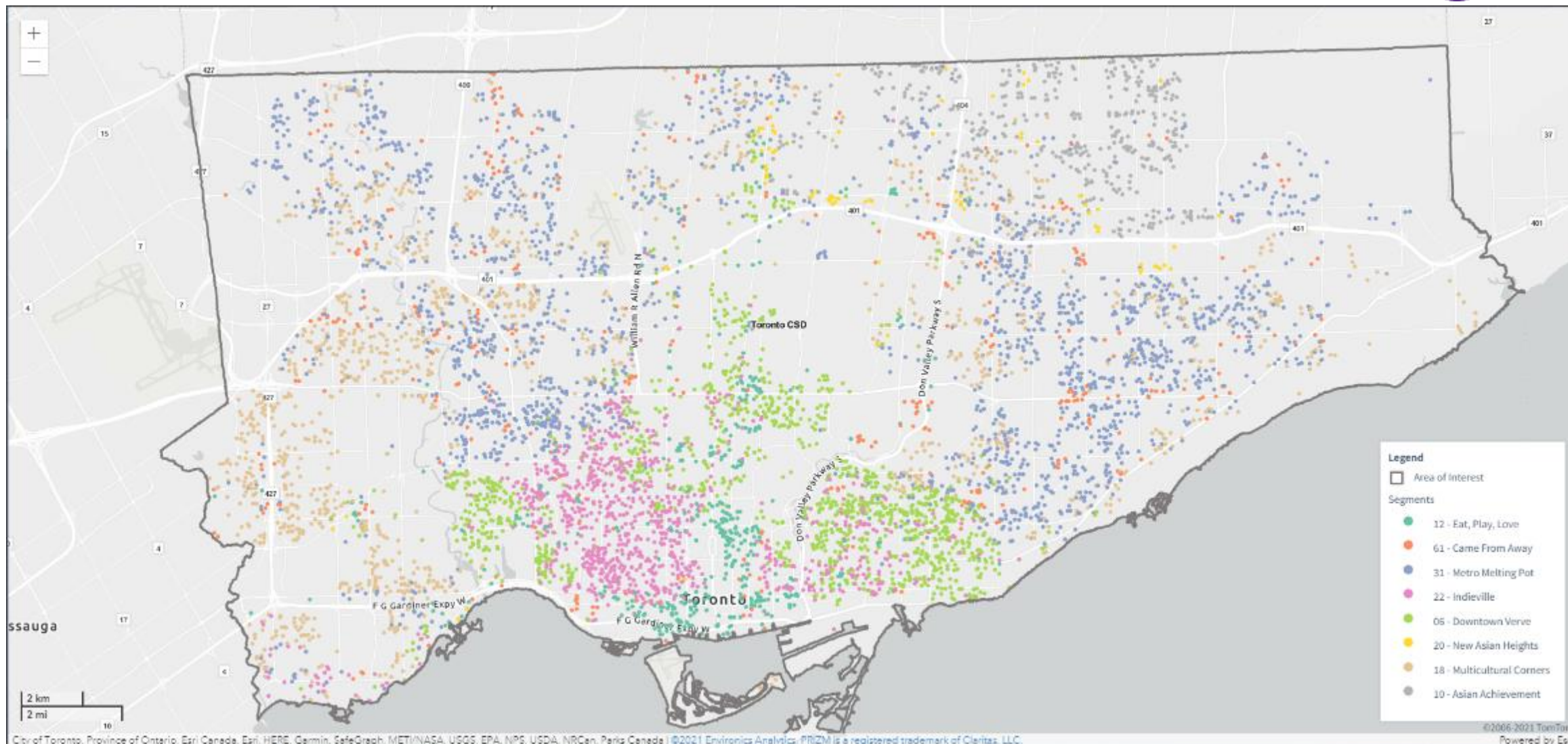
Вводные данные

- Демографический
- Психологический
- Финансовый
- Географический
- Покупка
- Досуг
- СМИ
- Мотиваторы

Как это сделано?

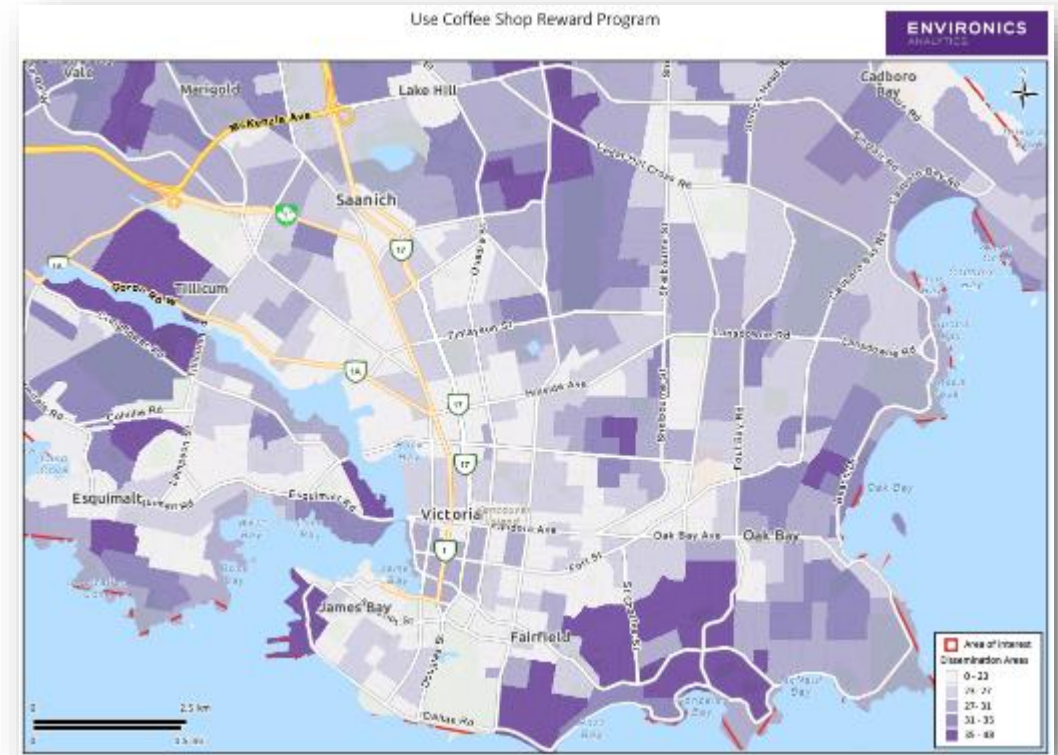


PRIZM: карта точек сегмента



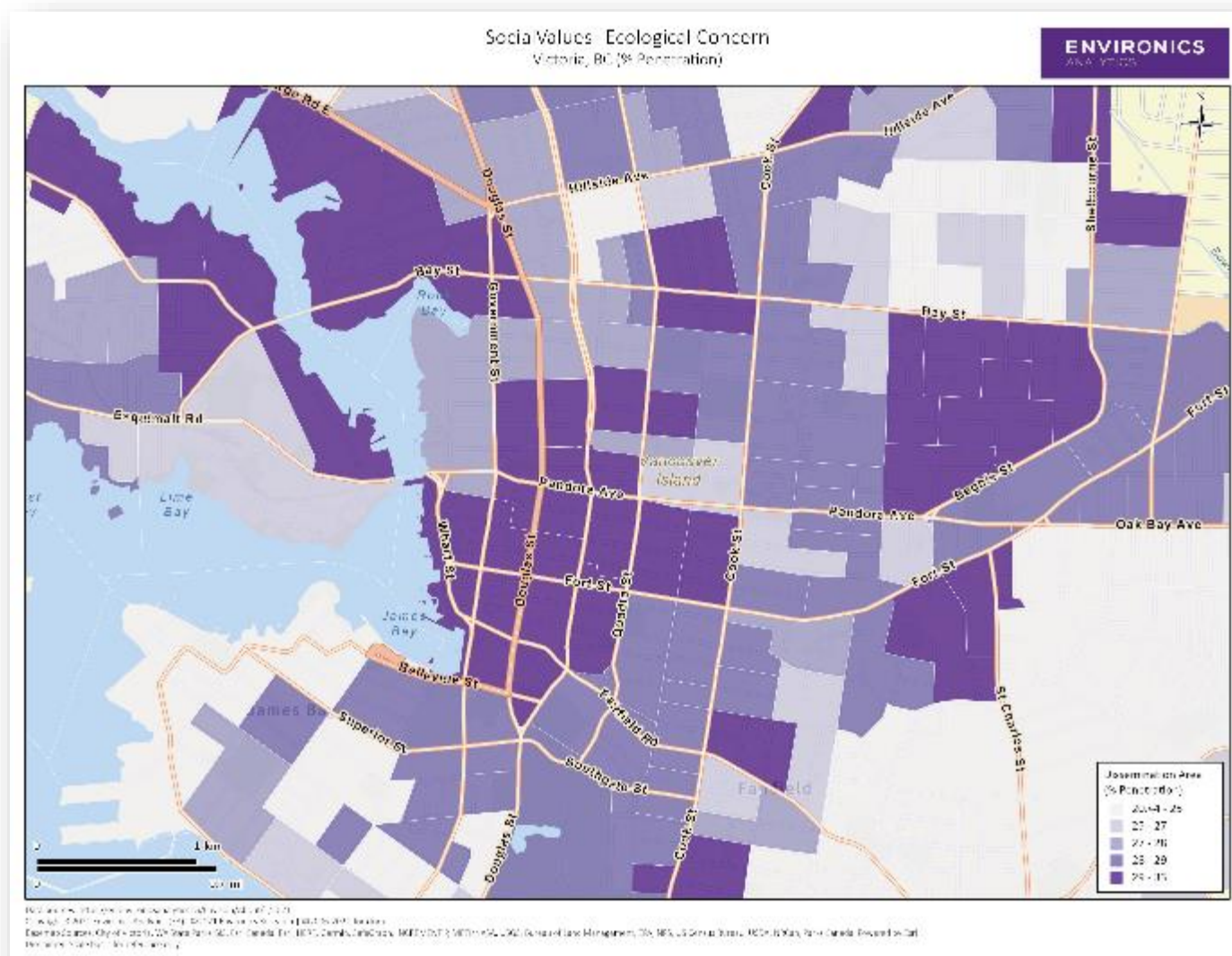
Расширение сферы применения на широко используемые обследования по измерению СМИ

- Включает более 5 000 поведенческих профилей, применимых к широкому спектру продуктов, видов деятельности, услуг и СМИ.
- Опирается на данные ведущего синдицированного канадского исследования кросс-медийной аудитории, потребительского поведения, использования продуктов и образа жизни.
- За основу берутся результаты выборки скользящего 52-недельного онлайн-опроса, в котором принимает участие более 34 000 респондентов.
- Детализация до уровня почтового индекса.



От геодемографии к психографике?

- Описывает образ мышления канадцев, чтобы найти подход к ним и обеспечить взаимодействие с ними.
- Опирается на данные обследования социальных ценностей, проводимого исследовательской группой Environics Research.
- 10 000 респондентов, что позволяет измерить 250 показателей мотивации и социальных отношений.
- Поведение часто используется в качестве заменителя, но образ мышления является главным фактором и особенно в цифровых технологиях.



Обращение к источникам больших данных

- Объем, разнообразие, скорость... Обычно с пространственными и временными атрибутами (признаками).
- Обычно эти данные собираются ежедневно или в режиме реального времени; проводятся миллиарды или триллионы наблюдений в год с той или иной формой ассоциации местоположения.
- Источники данных:
 - Общественная инфраструктура;
 - Мобильные устройства;
 - Сетевая инфраструктура;
 - Подключенные к Интернету устройства (автомобили, Интернет вещей, кассовые терминалы, датчики и т.д.).

MobileScapes — данные о перемещениях мобильных устройств, обновляемые ежедневно

- Объединяет в себе три источника разрешенных и анонимных данных, собранных с мобильных устройств с функцией определения местоположения, и использует лучшие методы обработки и анализа пространственных данных.
- Мобильная аналитика помогает предприятиям и организациям расширить свои знания о моделях передвижения, используя для этого данные о перемещениях мобильных устройств (соответствующие требованиям соблюдения конфиденциальности), собранные с мобильных устройств с поддержкой определения их местоположения.
- Самая точная, полная и актуальная база данных перемещений мобильных устройств, доступная для маркетинговых и бизнес-приложений.
- Может описать, кто посещает ваши офисы или офисы конкурентов, как часто, где они живут и работают, чтобы давать информацию для принятия решений, касающихся продуктов, для разработки программ, маркетинга, обмена сообщениями, планирования восстановления и подбора персонала.
- Действует как заменитель данных клиентов для привязки к системе Environics Analytics PRIZM.

Появление данных о передвижениях мобильных устройств

Преимущества

- Своевременность
- Низкая стоимость сбора данных
- Низкая нагрузка на респондентов

Недостатки необработанных данных

- Предвзятость во вселенной «пингов» (передачи данных)
- Уменьшение популярности приложений, привязанных к местоположению
- Опасность неправильного использования
- Риск неверных выводов

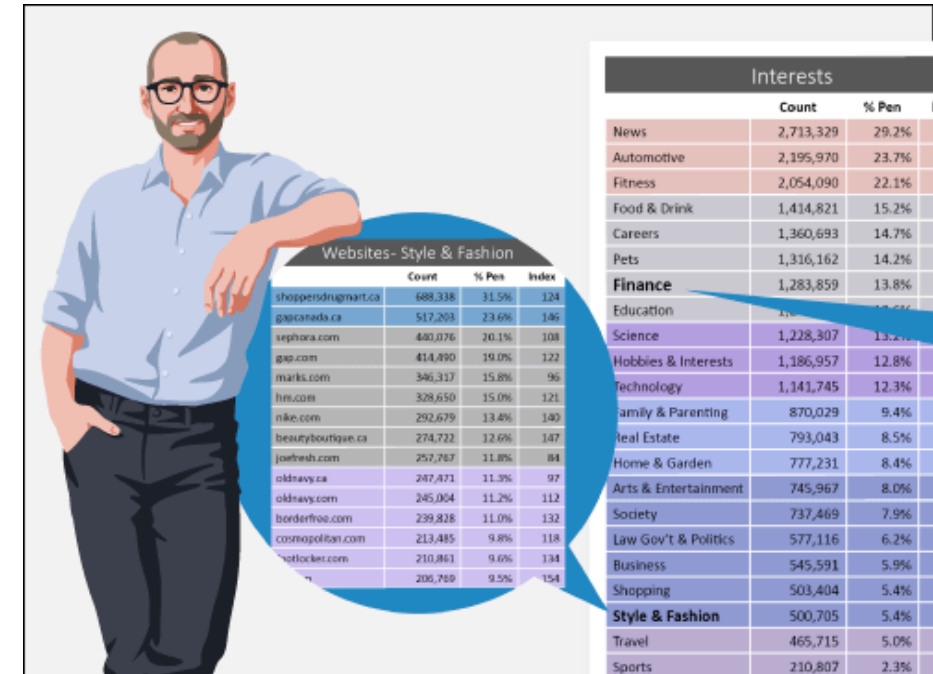
Преобразование БОЛЬШИХ данных в идеи

- Анонимизированные мобильные устройства можно рассматривать как пространственно-временную выборку населения.
- Не путайте «большие данные» с «аналитикой».
- Пространственный анализ может использовать большие данные, чтобы превратить их в полезную информацию.
- Методы сглаживания, интерполяции, агрегации и синтеза
 - защита конфиденциальности
 - проект для населения в целом
- Взвешено по эталонной совокупности для исправления поведенческих и отчетных предубеждений.
- Результаты представляются как расчетное население территории.
- Затем для описания демографии, поведения и психографических характеристик посетителей можно использовать традиционные геодемографические методы.

Большие данные – Интернет вещей – клики



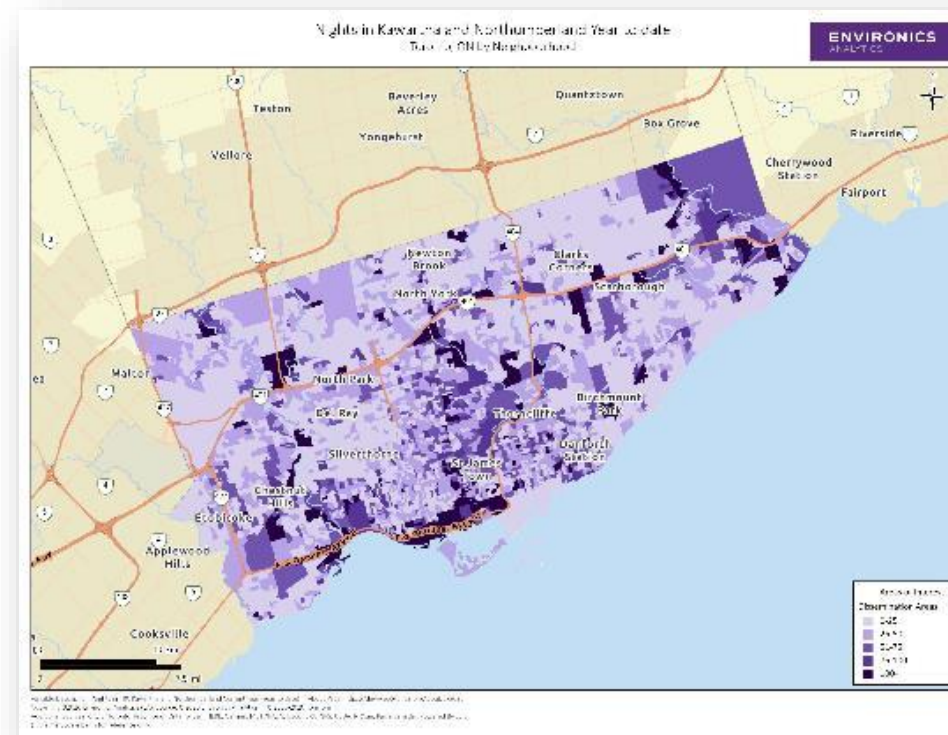
- Показатели посещения веб-ресурсов для тысяч программных приложений и веб-сайтов с агрегированием данных по компаниям и интересам.
- Данные для мобильных или домашних интернет-сетей.
- Доступ через проверенную, безопасную и отвечающую требованиям конфиденциальности экосистему Environics Analytics (EA).
- Ежегодные или ежемесячные обновления с еженедельной детализацией.
- Привязка модуля ClickScapes (Клики) с данными о клиентах или другими данными из Environics Analytics (EA), чтобы получать ранее недоступную информацию о посетителях и географические данные.



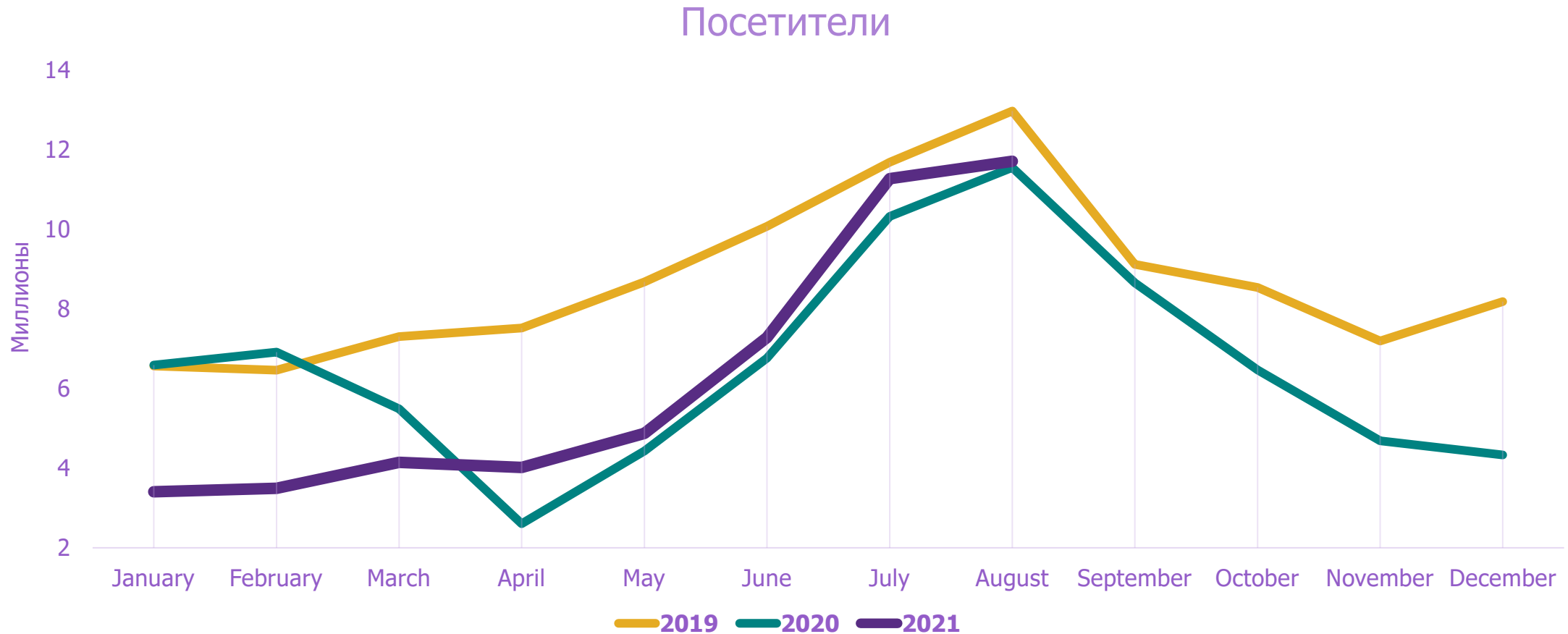
Большие данные – Интернет вещей – пинги



- Функция VisitorView представляет собой оценки на уровне районов о ночевках канадских путешественников в Канаде, ее провинциях и территориях, или в одном из 85 туристических регионов Канады согласно определению Destination Canada (портал, посвященный туризму в Канаде).
- Даются расчетные показатели общего числа посетителей, количество поездок и ночей, проведенных в каждой провинции или в туристическом регионе.
- Ежемесячное обновление, чтобы вы могли быстро и эффективно разрабатывать и проводить маркетинговые кампании.
- Есть привязка к нашей системе классификации районов PRIZM, включающей 67 сегментов, которые охватывают аспекты складывающейся демографической ситуации, образа жизни и ценностей канадцев.



Тенденции посещаемости с 2019 по 2021 год



Некоторые вопросы, на которые можно получить ответ



- Какие районы Торонто привлекают наибольшее количество посетителей в туристический регион Ниагарского водопада и Виноградной страны?
- Какие сегменты и рынки Онтарио привлекают наибольшее количество посетителей в канадские Скалистые горы в лыжный сезон?
- Каково общее количество ночей, проведенных посетителями в туристическом регионе Скалистых гор Кутеней?
- Где больше всего путешественников в межсезонье?
- На какие почтовые индексы следует настроить таргетированные (адресные) кампании прямого маркетинга?
- ... все ответы даны в рамках переписи населения.

60 информационных продуктов для канадского рынка



Demographic

DemoStats (Демографическая статистика)
CensusPlus (Перепись+)
DaytimePop (Население в течение дня)
CrimeStats (Криминальная статистика)
AccultuRates (Частота посещения объектов культуры)



Segmentation

PRIZM
PRIZM QC
DELTA



Financial

HouseholdSpend (Бытовые расходы)
FoodSpend (Расходы на питание)
WealthScapes (Профиль богатства)
WealthScapes Lite (Упрощенный профиль богатства)
WealthScapes Daytime (Профиль изменения богатства)
WealthScapes Fundraiser (Сбор средств)
AgeByIncome (Разбивка по возрасту и доходу)
LiquidAssets (Ликвидные активы)
MoneyMatters (Финансы) Данные от Canadian Financial Monitor
Neighbourhood View™ (Вид на окрестности)
ClickSpend™ (клики) Данные от J.C. Williams Group
WealthCare (Обеспеченность богатством)
WealthTransfer (Передача богатства)
Financial Vulnerability Index (Индекс финансовой уязвимости)



Behavioural

Opticks Данные от Vividata
Opticks Данные от Numeris
Opticks Social (социальные аспекты) Данные от AskingCanadians™
Opticks Mobile (мобильные аспекты) Данные от AskingCanadians™
Opticks eShopper (онлайн-покупки) Данные от AskingCanadians™
Opticks Automotive (автомобили) Данные от IHS Markit™
CannabisInsights (потребление марихуаны) Данные от Vividata
GreenLiving («Зеленый» образ жизни)
CommunityLife (Общественная жизнь)
GivingBack (Отдам даром)
Профили Homescan® (Дома)
AutoRank™ (Авторейтинг)
Рейтинг донора™ (DonorRank™)
FireScapes (Пожары)
ShopperChoice (Выбор покупателей)



Health

Social Vulnerability Index (Индекс социальной уязвимости)
Frailty Index (Индекс хрупкости)
CommunityHealth (Общественное здоровье)
VaccineInsights (Вакцинирование)



Psychographic

SocialValues (Социальные ценности)



Mobile Movement

MobileScapes Out & About (Мобильные технологии)
MobileScapes ENVISION (Мобильность)
VisitorView (Обзоры от посетителей)
ClickScapes (Клики)



Geographic

Enhanced PCCF (Леса)
Streets & Boundaries (Улицы и границы районов)
Postal Code Boundaries (Границы почтового индекса)



Location

Businesses (Компании)
Financial Institutions (Финансовые учреждения)
Spectra Trade Areas (Торговые зоны Spectra)
ShoppingCentres (Торговые центры)
TrafficCounts (Счетчик трафика)
Points of Interest (Интересные



Business

BusinessProfiles (Бизнес-профили)
BusinessProfiles Lite (Упрощенные бизнес-профили)
ResponseCanada™ Business (Реагирование канадского бизнеса)



Contact

ResponseCanada™ Consumer (Потребители)
ResponseCanada™ Movers (Мигранты)
ResponseCanada™ Pre-Movers (Планирующие мигрировать)

Использование оценок малых территорий в качестве показателя данных о клиентах

Данные малых территорий можно использовать для следующего:

- Описать
 - Понять
 - Сфокусироваться и нацелиться
-
- Использование заданных рамок в качестве показателя / знаменателя

Следующий шаг: Исполнение и атрибуция

Применение фреймворка (рамок) и данных к технологиям рекламы (Adtech)



Адаптация к отраслям

Мы работаем во многих отраслях

- Автомобилестроение
- Банковское дело
- Потребительские товары
- Кредитные союзы
- Образование
- Энергетика
- Правительство
- Здравоохранение
- Страхование
- Некоммерческий сектор
- Розничная торговля
- Недвижимость
- Спорт и развлечения
- Путешествия и отдых
- Телекоммуникации
- СМИ

Перепись как основа для бизнес-аналитики

ЕЭК ООН – Неделя переписи – сентябрь 2022 г.

Джон Неллиган

Технический директор
Энвироникс Аналитика
Торонто, Онтарио
Канада