Stats NZ
Tatauranga Aotearoa

Миграция на основе наблюдаемого поведения путешествующих

Подход Новой Зеландии к оценке миграции на основе административных данных

Условия, способствующие измерению результатов

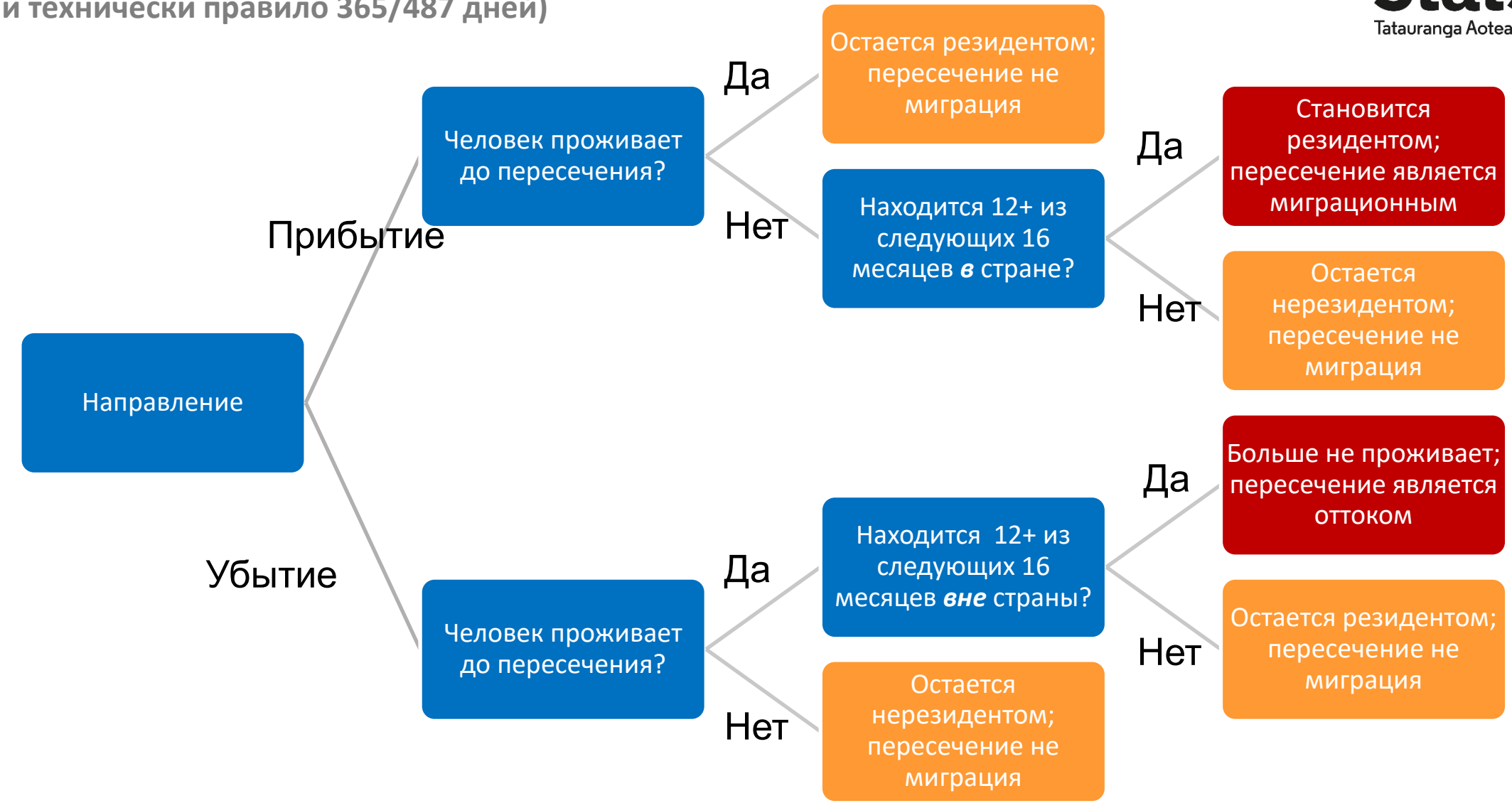
- Осведомленность об ограничениях намерений
- Меняющийся характер миграции
  - Временные визы
  - Мигранты, пересекающие границу несколько раз
- Возможность связать записи прибытия и убытия
  - Наличие связывающих переменных
  - Улучшена методология/методы связывания
  - Рядовые сотрудники ИТ
- Растущее количество пересечений границы
- Предлагаемое упразднение карты убытия

Как работают подходы, основанные на результатах

- Выявление отдельных лиц в данных о пересечении границы
- Составление «истории путешествий» для каждого человека
  - Связывание воедино все прибытия и убытия для одного и того же человека
- Применение правила классификации
  - Классифицируйте каждое прибытие/убытие как «миграция» и «не миграция».
- 12/16 - правило месяца - один подход, основанный на результатах

Правило 12/16 месяцев

(Или технически правило 365/487 дней)



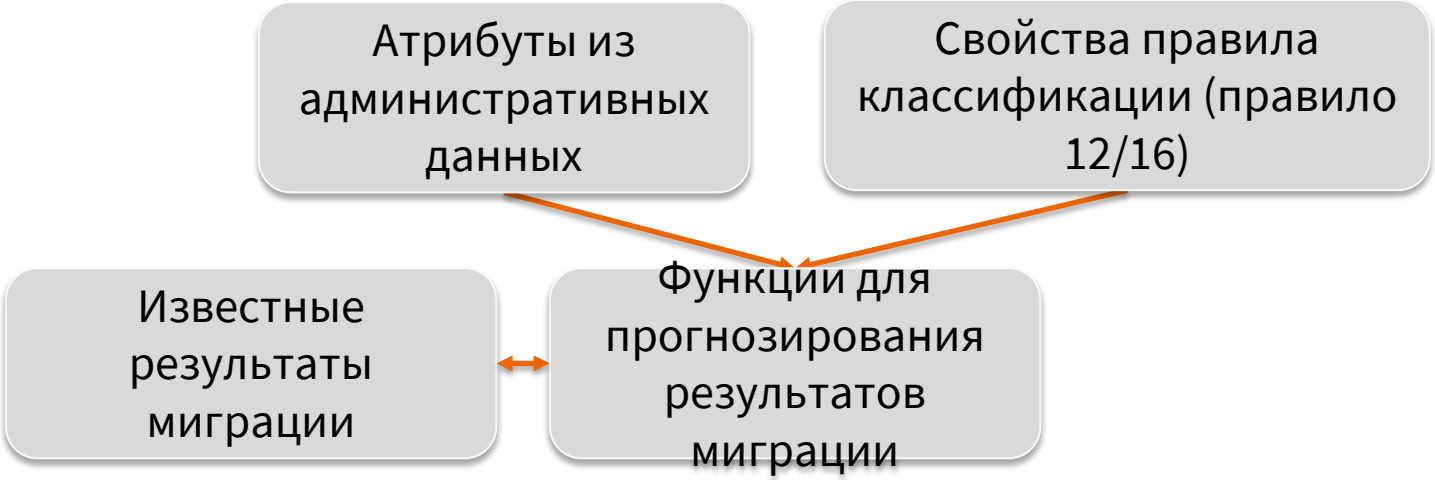
Преодоление потери своевременности с помощью прогнозирования

- Клиенты из Новой Зеландии не могут ждать миграционных мер более 16 месяцев.
  - Они ожидают результатов примерно через 1 месяц после истечения отчетного периода.
- Чтобы применить правило 12/16 к пересечению границы, потенциально придется подождать 16 месяцев после того, как произошло пересечение границы.
- Из всех пунктов пересечения границы (до COVID) только 1-2% приходились на пересечение мигрантами
- Можем ли мы предсказать классификацию этих переходов?

Моделирование единичных записей с помощью машинного обучения

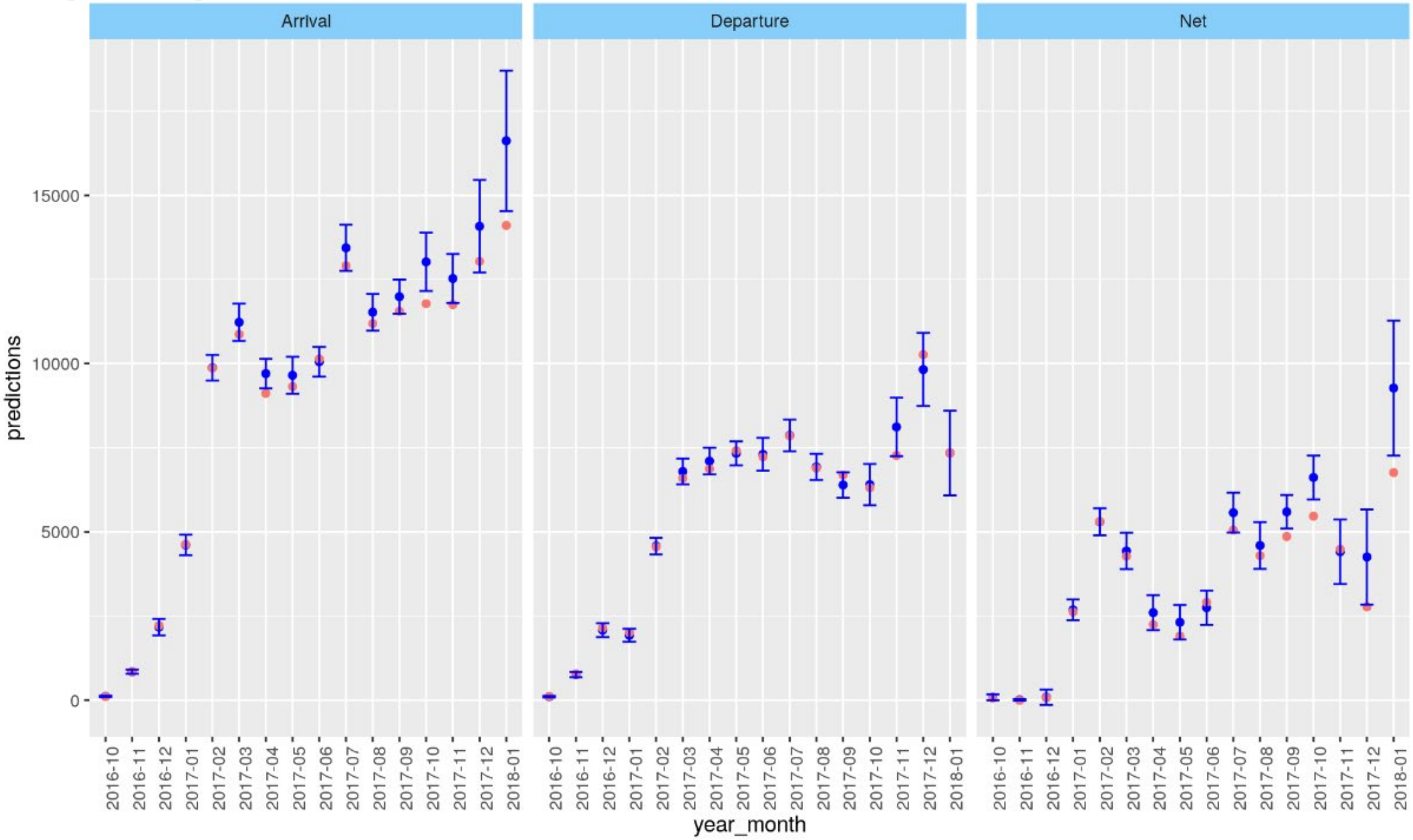
- Подход машинного обучения к классификации границ - **на уровне единичных записей**
- Используется классификатор с градиентным усилением, специально реализованный через XGBoost.
  - внедрение усиленных древовидных схем
  - Известно, что алгоритм хорошо работает в задачах классификации.
- Для каждого непроясненного пересечения классификатор оценивает вероятность того, что это пересечение мигрантами.
- Интегрирует некоторую неопределенность модели - путем запуска модели на нескольких подвыборках обучающих данных

Характеристики модели машинного обучения



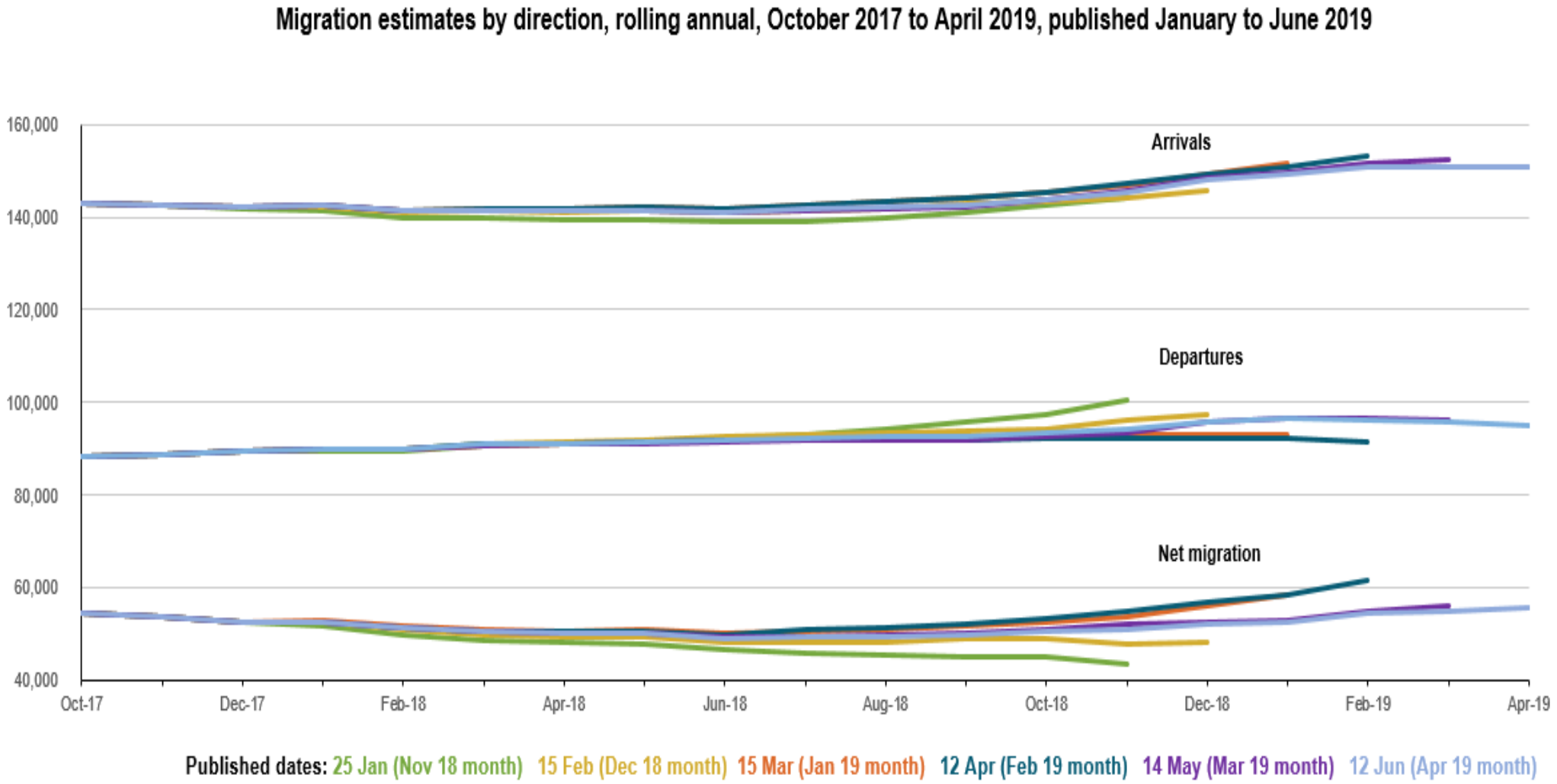
- Наблюдаемые дни: дни, прошедшие с момента пересечения границы
- Дни в стране: дни в Новой Зеландии с момента пересечения границы.
- Подсчитать будущую сумму: пересечения границы после того, которое будет классифицировано
- Направление, возраст, пол, тип визы, регион гражданства, месяц года

Пример выполнения общей оценки



Предварительные оценки миграции

Stats NZ
Tatauranga Aotearoa



Несколько заключительных замечаний и точка зрения клиента / пользователя

- Новая Зеландия использует ежедневные данные о пересечении границ для измерения миграции
- Используется метод, основанный на результатах, который основан на времени, проведенном в/вне страны - чисто демографическая концепция.
  - Независимо от правового статуса, типа визы или других подобных факторов
  - Независимо от намерений путешествующего - это просто основано на их поведении.
- Заказчики в Новой Зеландии требуют своевременную (около 1 месяца давности) оценку миграции - мы используем модель машинного обучения, чтобы предоставить эту информацию.

Stats NZ
Tatauranga Aotearoa

Спасибо, вопросы?