

Unikanie pułapek w budowaniu zespołu Data Science

Omówimy dwie główne pułapki, które wpływają na zespoły analityków danych. Po pierwsze, jeśli zespół zbyt szybko osiąga konsensus, hamuje to odkrywanie i jest znakiem, że zespół ma ślepe punkty i jest skłonny do myślenia grupowego. Wreszcie, skąd wiesz, że Twój zespół spędza zbyt dużo czasu na niewłaściwym pytaniu lub zadaje pytania dotyczące niewłaściwych rzeczy? Nazywa się to wędrówką i podajemy kilka wskazówek, jak tego uniknąć.

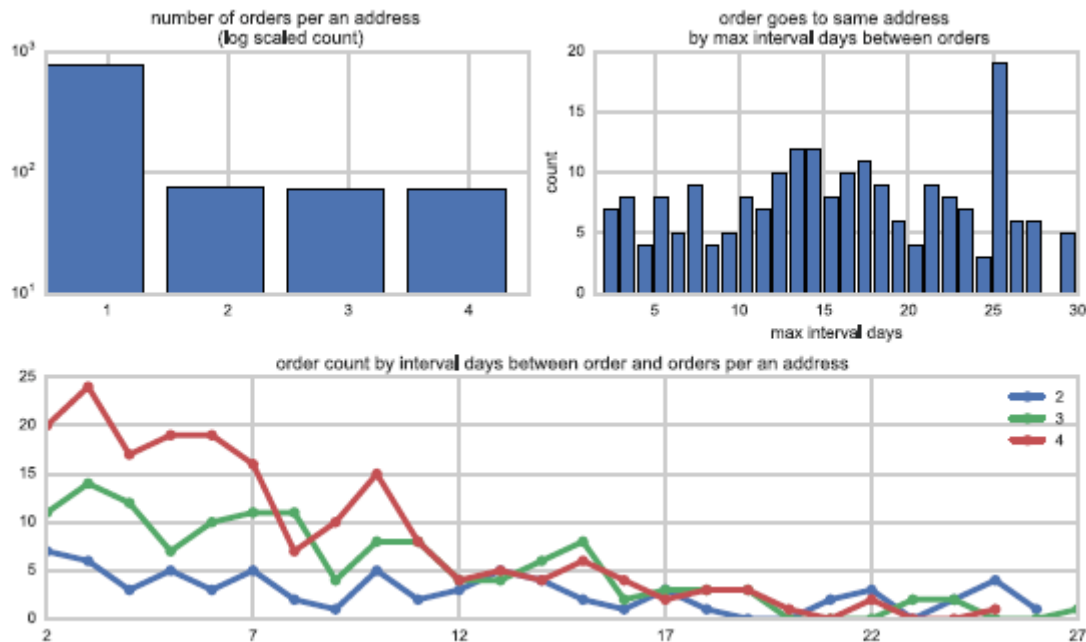
Unikanie konsensusu

W większości organizacji ludzie w naturalny sposób starają się osiągnąć konsensus. Różne organizacje nazywają to różnymi rzeczami. Niektórzy zachęcają wszystkich do „dogadywania się”. Inni używają słów takich jak „socjalizacja”. Nauka o danych jest zupełnie inna. Konsensus może być dużym problemem. Chcesz, aby Twój zespół odkrywał nowe pomysły. Jeśli wszyscy dojdą do konsensusu zbyt szybko, może to oznaczać, że wszyscy podzielają powszechne nieporozumienie. Pamiętaj, że nauka o danych polega na eksploracji. Szukasz wiedzy i spostrzeżeń. Nie ma potrzeby, aby wszyscy byli na tej samej stronie. W rzeczywistości chcesz, aby wszyscy byli wystarczająco spokojni, aby dyskutować o tym, jak interpretować dane. Zespół zajmujący się analizą danych powinien bardziej przypominać niezręczny rodzinny obiad niż spokojną przejażdżkę autobusem. Chcesz, aby zespół rozmawiał, badał, a nawet drażnił się nawzajem. Ten rodzaj wymiany jest znacznie bardziej prawdopodobny, aby odkryć nowe pomysły. Jest kilka rzeczy, które możesz zrobić, aby Twój zespół nie osiągnął szybkiego konsensusu:

- Bądź świadomy niebezpieczeństwa konsensusu. Pamiętaj, że szybkie porozumienie w złożonym temacie jest często oznaką myślenia grupowego.
- Upewnij się, że Twój zespół jest wystarczająco mały, aby wszyscy czuli się komfortowo, jeśli się nie zgadzają. Twój zespół powinien mieć kierownika badań, kilku analityków danych i kierownika projektu. Staraj się, aby Twój zespół liczył mniej niż sześć osób. Większe grupy mają tendencję do wypierania cichszych głosów, które często mają jedno z najlepszych spostrzeżeń.
- Zadbaj o to, aby Twój kierownik badań sprowadzał osoby spoza zespołu.

Rozwińmy ten ostatni punkt, ponieważ trzymanie się z dala od szybkiego konsensusu ma kluczowe znaczenie. Załóżmy, że przeglądasz interesujące pytanie dotyczące witryny internetowej poświęconej butom do biegania. Próbujesz sprawdzić, czy ludzie, którzy są w romantycznym związku, prawdopodobnie namówią swojego partnera na bieganie. Zespół próbuje dowiedzieć się, jak dotrzeć do tego pytania za pomocą danych ze strony internetowej. Kierownik badania może chcieć zaprosić kogoś ze sprzedaży do podzielenia się anegdotycznymi historiami lub zaprosić kilka znanych im osób, które biegają ze swoją dziewczyną, chłopakiem, mężem lub żoną, aby opowiedzieć o tym, jak zaczęli biegać ze swoim partnerem. Osoby te mogą być w stanie dodać spostrzeżenia, których zespół mógłby nie mieć w innym przypadku. Pamiętaj tylko, aby zespół był mały. Więc może po prostu dodaj jedną lub dwie osoby na spotkanie. Coś innego, co Twój zespół może spróbować, to zakończyć każdą dyskusję zakładając, że się myli. Może to być tak proste, jak pytanie: „A co, jeśli się mylimy?” Twój zespół powinien być w stanie odpowiedzieć na to pytanie. Może zdadzą sobie sprawę, że mylą się w tym lub mylą się w wielu innych sprawach. Jeśli to prawda, przejrzyj te wszystkie inne rzeczy. Zwróć uwagę na czasy, kiedy zespół nie ma odpowiedzi na to pytanie. To jeden z sygnałów ostrzegawczych, że zespół może paść ofiarą myślenia grupowego. Powiedzmy, że zespół postanawia podejść do kwestii romantycznych partnerów do biegania. Decydują, że najlepszym sposobem na przejrzenie tych danych jest sprawdzenie, czy istnieją klienci, którzy mają ten sam adres, a następnie porównanie ich

zamówień, aby sprawdzić, czy jeden klient zaczął składać zamówienia przed drugim, jak pokazano na Rysunku.



Ale co, jeśli zespół się myli? Jak będą wyglądać dane? Może współlokatorzy będą biegać razem. Mogą być studentami lub próbować podzielić czynsz. Mogą to być słabe argumenty, ale należy je rozważyć. Najcenniejsza jest dyskusja o tym, co to znaczy być w błędzie.

Większość adresów ma tylko jedno zamówienie, a maksymalnie cztery zamówienia na adres w oknie jednego miesiąca. Z wyłączeniem tych, które mają tylko jedno zamówienie, maksymalny odstęp dni między zamówieniami na jeden adres wynosi średnio 16,5 dnia. Jeśli porównamy przedział czasowy między zamówieniami, najprawdopodobniej ludzie złożą kolejne zamówienie w ciągu 15 dni, a im więcej zamówią w ciągu jednego miesiąca, tym więcej zamówień złożą w ciągu pierwszych 7 dni.

Inną rzeczą, którą należy spróbować, jeśli zespół zbyt szybko osiąga konsensus, jest znalezienie kogoś, kto prawdopodobnie się nie zgodzi i będzie adwokatem diabła. Na przykład na stronie internetowej poświęconej butom do biegania możesz chcieć zaprosić kogoś, kto nigdy nie biegałby z romantycznym partnerem. Zapytaj ją, dlaczego tak się czuje. Może o wiele częściej ludzie wykorzystują ten czas na samotność. Możliwe, że biegacze lubią biegać, ponieważ jest to sport dla samotników. Są to spostrzeżenia, które naprawdę mogą wnieść wiele do dyskusji zespołu.

Jeśli pracujesz dla dużej organizacji, możesz być bardziej skłonny do osiągnięcia szybkiego konsensusu, niż myślisz. Typowy projekt w dużej organizacji wymaga dużej koordynacji. W przypadku tych projektów potrzebujesz konsensusu, aby wszystko załatwić. Zespół analityki danych ma inny zestaw priorytetów. Kiedy szukasz nowej wiedzy i spostrzeżeń, zbyt szybkie osiągnięcie konsensusu może być prawdziwym niebezpieczeństwem. Jeśli użyjesz technik omówionych w tej sekcji, z większym prawdopodobieństwem sprawisz, że Twój zespół będzie odkrywał i otwierał się na nowe pomysły.

Utrzymanie zespołu od wędrówki

Jak widzieliśmy, istnieje niebezpieczeństwo, że zespół zbyt szybko osiągnie konsensus. Z drugiej strony, skąd wiesz, że Twój zespół spędza zbyt dużo czasu na niewłaściwym pytaniu? Kiedy powinieneś porzucić pytanie i zacząć kolejne? To jest przeciwne wyzwanie. Zamiast zgody wszystkich, nadal zadają

Odkryli, że ludzie, którzy kupili farbę, częściej kupują rano, osoby, które dokonały dużych zakupów, chętniej kupowały urządzenie, a klienci, którzy kupili dywan, chętniej kupowali go w piątek.

■ **Uwaga:** Im większy kwadrat, tym więcej pieniędzy wydano na dany dzień tygodnia. Ciemniejszy kolor reprezentuje sprzedaż dokonaną później w ciągu dnia; jaśniejszy kolor reprezentuje sprzedaż dokonaną wcześniej w ciągu dnia. W przypadku farb i dekoracji całkowite wartości są dość małe w porównaniu do dywanów, naczyń kuchennych i wentylatorów. Farby kupuje się najczęściej rano. Dywan ma nieco wyższą wartość sprzedaży w piątek, a naczynia kuchenne mają nieco wyższą sprzedaż w poniedziałek.

Każdy z tych małych punktów danych był w porządku, ale analitycy danych nie podjęli próby stworzenia historii. Jakie są motywacje i wyzwania profesjonalnego klienta? To jest problem z zagłębieniem przez dziurkę od klucza w szklanych drzwiach. Zdarza się to czasami, gdy zespół ma zbyt wielu analityków danych. Koncentrują się na dziesiątkach małych punktów danych, zamiast zmagać się z większą prawdą. Jeśli nie zostanie to rozwiązane, zespół może dalej wędrować po danych i stracić zainteresowanie zadawaniem ważnych pytań. Kluczowym obowiązkiem kierownika badań jest upewnienie się, że tak się nie stanie, zmuszając zespół do zadawania interesujących pytań i dostarczania cennych informacji. Kierownik badań reprezentuje interesy biznesowe w zespole data science. Chcesz, aby zespół czuł się swobodnie w eksploracji, ale jednocześnie chcesz, aby nie wędrował. Jeśli zespół nie dostarcza spostrzeżeń, nie będą one bardzo cenne dla firmy. To jeden z powodów, dla których chcesz, aby zespół był mały i równo reprezentowany — jeden lub dwóch analityków danych, kierownik badań i kierownik projektu. Jeśli Twój zespół jest zbyt obciążony analitykami danych, istnieje ryzyko gromadzenia danych bez wglądu. Małe zespoły mogą zadawać ważne pytania. Pamiętaj też, że jesteś właścicielem wiedzy i spostrzeżeń zespołu. Jeśli uważasz, że zespół nie myśli wystarczająco dobrze, musisz rzucić mu wyzwanie, aby podązał za większymi historiami.

PODSUMOWANIE

Poznałeś kilka wskazówek, jak uniknąć dwóch typowych pułapek w zespołach zajmujących się analizą danych. Jeden zbyt szybko dochodził do konsensusu, a drugi błąkał się. Jeśli zespół zbyt szybko osiąga konsensus, hamuje to odkrywanie i jest znakiem, że zespół ma martwe punkty. Jeśli Twój zespół wędruje, spędza zbyt dużo czasu na jednym pytaniu. Mogą również zadawać pytania dotyczące niewłaściwej rzeczy. W Sekcji IV dowiesz się, w jaki sposób zespół ds. analityki danych powinien koncentrować się na dostarczaniu wartości