

Brief na prace zaliczeniową

Przedmiot: Marketing Mix Modelling w praktyce

Uczelnia: Wydział Nauk Ekonomicznych UW

Koordinator przedmiotu: Natalia Nehrebecka

Semestr: zimowy 2022/2023

1. Informacje ogólne:

- praca zaliczeniowa polega na przeprowadzeniu projektu MMM dla wskazanej marki
- głównym celem projektu jest pokazanie przez studenta, że:
 - rozumie ideę projektu MMM
 - rozumie charakter danych zawartych w modelowaniu i potrafi je przygotować
 - potrafi stworzyć model i przygotować na jego podstawie analizy
 - identyfikuje problemy w projekcie MMM i próbuje je rozwiązać
 - potrafi zaprezentować przygotowane przez siebie analizy i rekomendacje z perspektywy działań biznesowych klienta (a nie z perspektywy technicznej)
- dostarczony przez studenta model i wyniki analiz nie będą porównywane z żadnym „szablonem”:
 - student zatem nie powinien się przejmować, czy uzyskany przez niego wynik jest spójny z jakimś innym modelem lub projektem

2. Output z projektu zaliczeniowego:

- skrypt R z kodem:
 - jeden plik z całym kodem (nie kilka plików)
 - skrypt powinien być podzielony na części (np. przygotowanie danych o temperaturze, adstockowanie zmiennych, analiza graficzna, model, walidacja modelu, dekompozycja modelu, ...)
 - istotna jest czytelność kodu (dodatkowe punkty)
 - student powinien zamieszczać w kodzie najważniejsze komentarze
 - kompletny kod projektu w R wraz z komentarzami należy przestać na adres bartosz.kowalski@mediacom.com **do północy przed dniem prezentacji**
- prezentacja w formacie pptx powinna:
 - zawierać wyniki analiz i kluczowe rekomendacje
 - być zaprezentowana w sposób atrakcyjny, zarówno z perspektywy wniosków, jak i akcji biznesowych klienta/marki
 - prezentacja powinna trwać 10 minut. Znaczne przedłużenie lub prezentacja niepełna wpłyną negatywnie na ocenę projektu
 - estetyka slajdów i sposób opowiadania wpływają w znacznym stopniu na jej odbiór, a tym samym na ocenę
 - przygotowane slajdy w pptx również należy przestać na adres bartosz.kowalski@mediacom.com, ale **do północy w dniu samej prezentacji** (czyli można po spotkaniu)
 - zapisy na termin prezentacji odbywają się pod poniższym linkiem: https://bit.ly/MMM_zapisy

3. Dane potrzebne do przygotowania zadań:

- lokalizacja:
 - <https://github.com/KarolinaDrowanowska/Marketing-Mix-Modeling/Zaliczenie>
- data.csv:
 - plik zawiera wszystkie dane niezbędne do przygotowania projektu

- *data_slownik.csv* :
 - plik zawiera opis wszystkich zmiennych zawartych w pliku *data.csv*
- *temperatura.xlsx*:
 - plik zawiera dane potrzebne do przygotowania temperatury

4. Przydział marki i jednostki panelu do modelowania:

- każdemu studentowi została przydzielona konkretna marka i jednostka panelu, na podstawie której ma przygotować model
- lista z przydziałem marek i zmiennych znajduje się w pliku *przydział_modeli.xlsx*, umieszczonym w linku w niniejszym briefie
- student może poprosić o zmianę marki i jednostki panelu na inną, znajdującą się w puli marek i jednostek panelu do wolnego wyboru:
 - w tym celu student musi wybrać inną markę i jednostkę panelu z puli dostępnych marek
 - następnie powinien przestać wiadomość email ze preferencją (preferencjami) na adres leszek.lauks@mediacom.com
 - jeśli do tego momentu żaden inny student nie wybrał wskazanej marki w danej jednostce panelu, to student uzyska informację zwrotną, że dana marka została mu przydzielona (w przeciwnym wypadku dostanie informację zwrotną, że dana opcja jest już zajęta)
 - plik *przydział_modeli.xlsx* zamieszczony na sieci będzie uaktualniany raz w tygodniu
 - o kolejności przydziału marek i jednostek panelu z puli decyduje czas nadstania zgłoszenia na adres email wskazany powyżej
 - liczba marek i jednostek panelu w puli do wolnego wyboru jest ograniczona, zatem student nie ma 100% gwarancji, że jest prośba o zmianę zostanie zaakceptowana

5. Przygotowanie danych:

- w pliku *data.csv* znajdują się wszystkie potrzebne zmienne do stworzenia modelu
- chcemy mieć pewność, że student rozumie **pojęcia adstocku oraz metodę przygotowania zmiennych dotyczących temperatury**, stąd pomimo kompletności bazy danych *data.csv*, prosimy o przygotowanie następujących zmiennych:
 - adstockowanie zmiennych mediowych:
 - student powinien zatem dostarczyć kod w skrypcie R, który będzie adstockował zmienne mediowe dla przydzielonej mu marki modelowanej i jednostce panelu
 - dane do adstockowania zawarte są w zmiennej z adstockiem 0%
 - zaadstockowane przez studenta dane mogą nieznacznie różnić się od danych zawartych w bazie, ze względu na zastosowany przez *Business Science* algorytm „redukcji” bardzo małych wartości
 - przygotowanie zmiennych dla temperatury:
 - student powinien przygotować zmienną dotyczącą rzeczywistej temperatury w całej Polsce i jej normy (cyklu)
 - student powinien dostarczyć kod w skrypcie R, który na podstawie pliku inputowego przygotuje dane dla temperatury
- student nie musi wykorzystywać w modelu przygotowanych przez siebie zaadstockowanych zmiennych mediowych oraz zmiennych dotyczących temperatury (gdyż zmienne te zawarte są już w bazie)

6. Przygotowanie pozostałych zmiennych:

- jeśli student uważa, że dostarczona baza (*data.csv*) nie zawiera jakiegś istotnej zmiennej, i jest w stanie przygotować taką zmienną (na podstawie publicznie

dostępnego źródła), to może przygotować taką zmienną i wykorzystać w trakcie modelowania (nie jest to jednak obowiązkowe)

- jeśli student zdecyduje się takie zmienne wykorzystać, proszony jest o podanie źródła danych
- w tym punkcie należy umieścić również wszystkie zmienne tworzone na podstawie zmiennych z bazy *data.csv*, które student finalnie wykorzystuje w projekcie

7. Analiza graficzna danych:

- student powinien dostarczyć kod w skrypcie do stworzenia wykresów, przedstawiających najważniejsze informacje o przydzielonej marce modelowanej (w danej jednostce panelu)
- rekomendowane jest dostarczenie komentarzy do wykresów z najważniejszymi wnioskami

8. Modelowanie:

- student powinien przygotować model ekonometryczny
- student proszony jest o umieszczenie tylko jednego, finalnego modelu
- kluczowe informacje o modelu:
 - forma modelu: model addytywny
 - struktura danych: szereg czasowy na całym dostępnym okresie
 - zmienna modelowana: wolumen sprzedaży marki w konkretnej jednostce panelu (rekomendowane jest, aby zmienną modelowaną podzielić przez jej średnią wartość, w celu łatwiejszej interpretacji parametrów)
 - metoda estymacji: MNK
 - jeśli student uważa, że jakaś zmienna w modelu ma sens marketingowy, to akceptujemy *p.value* nawet do poziomu 0,30

9. Walidacja marketingowa i statystyczna modelu modelu:

- student powinien zwalidować model pod kątem marketingowym
- student powinien zwalidować model pod kątem statystycznym
- z perspektywy *Business Science* głównym celem zadania jest dowód, że student rozumie ideę walidacji marketingowej modelu, rozumie dany problem, zdaje sobie sprawę z jego istnienia, i próbuje go rozwiązać. W związku z tym:
 - jeśli jakiś czynnik jest np. „nieco przeszacowany” w modelu, i pomimo prób nie udało się tego zmienić, to taki fakt **nie dyskwalifikuje** modelu
 - jeśli model np. „nie przechodzi” jakiegoś testu statystycznego, i pomimo prób nie udało się tego zmienić, to taki fakt **nie dyskwalifikuje** modelu

10. Analizy na podstawie wyestymowanego modelu

- Dekompozycja modelu – przygotowanie dekompozycji w R z opisanymi poziomami debazowania oraz zaprezentowanie wyników analizy w prezentacji z interpretacją najważniejszych wniosków dla klienta
- Analizy czynników niemediowych – wybranie najciekawszych i najistotniejszych czynników oraz wizualizacja wniosków z analiz (np. konkurencji lub elastyczności cenowej) – zdecyduj na czym chciałbyś się skupić opowiadając swoje wnioski klientowi i zwizualizuj odpowiednie wnioski
- Analizy mediowe, których podstawą jest wskaźnik ROI. Założenia:
 - rzeczywisty koszt tygodniowy danego medium (dotyczy mediów: Kino, Outdoor, Radio) określa zmienna mediowa dla danego medium z adstockiem zerowym
 - medium TV wyrażone jest w GRP, zatem do policzenia kosztu TV przyjmij koszt: 1 GRP = 1200 PLN
 - załóż, że suma sprzedaży marki modelowanej w danej jednostka panelu stanowi 5% całkowitej sprzedaży w całej Polsce. W związku z tym inkrementalną sprzedaż

wygenerowaną przez dany kanał mediowy w modelu należy „wyekstrapolować” na całą Polskę

- marża na produkcie = 1 PLN za litr
- uzyskane ROI-e mogą wyjść bardzo wysokie lub bardzo niskie (np. 50 lub 0,1), wynika to z faktu, że do policzenia ROI przyjęliśmy wiele założeń (np. marża, ekstrapolacja, itp.), które w rzeczywistości mogły wyglądać inaczej

11. Optymalizacja media mixu

- obejmuje stworzenie krzywych dla mediów w wybranym przez studenta roku kalendarzowym (2010 lub 2011):
 - warunek konieczny: krzywe dla minimum dwóch kanałów mediowych
- podanie optymalnej realokacji historycznego budżetu oraz korzyści, jakie klient dzięki temu osiągnie
- Założenia:
 - załóż, że liczba tygodni, w których dane medium będzie w wybranym roku, jest taka jak historycznie w wybranym roku
 - dla medium TV przyjmij koszt 1 GRP = 1200 PLN
 - dla pozostałych mediów koszt medium opisuje zmienna dla tego medium z adstockiem zerowym
 - jeśli jakieś medium nie wyszło w modelu (nawet po próbie rozdzielenia go na lata lub półrocza), to przyjmij, że krzywa Response curve ma kształt linii prostej o wartościach zerowych
 - wielkość rocznego budżetu do optymalizacji jest taka, jaka była historycznie w wybranym przez studenta roku
 - załóż, że na dany kanał mediowy nie można wydać mniej niż 50% historycznego budżetu (w wybranym roku) – nawet jeśli dane medium nie wyszło w modelu w tym roku
 - załóż, że na dany kanał mediowy można wydać maksymalnie o 100% więcej niż wynosił historyczny budżet (w wybranym roku)

12. Kryteria oceny pracy zaliczeniowej

Zadanie	Element	Oczekiwania	Maks. punkty
Przygotowanie danych	kod	Adstockowanie danych mediowych i przygotowaniem danych dla temperatury	5
Analiza graficzna danych	kod	Kod z wykresami odpowiednich zmiennych i komentarzami	5
Stworzenie modelu	kod	Kod z finalnym modelem + komentarze do zmiennych i modelu	10
Weryfikacja marketingowa i statystyczna modelu	kod	Kod z testami statystycznymi i obliczeniami + komentarze	5
Analizy, symulacje, rekomendacje	kod	Kod do: kontrybucji, policzenia ROI, stworzenia Response curves, optymalizacji mediów, ...	5
Czytelność kodu w R	kod	Czytelność, przejrzystość i komentarze w kodzie	5
Prezentacja (dokument) i prezentacja (ustna)	prezentacja		35