

Paweł Hanczar, Jacek Zasada, Paweł Sobolik

Uniwersytet Ekonomiczny we Wrocławiu

Identyfikacja obserwacji odstających w szeregach czasowych na przykładzie branży FMCG

Identification of outliers in the time series on an example of the FMCG industry

Synopsis. Niniejsza praca ma na celu przeprowadzenie weryfikacji i analizy wybranych metod identyfikacji obserwacji odstających w branży FMCG (ang. *fast – moving consumer goods*). Do przeprowadzenia badań użyto rzeczywistych danych sprzedaży amerykańskiej sieci sklepów Walmart, udostępnionych przez firmę w Internecie. Analiza została wykonana za pomocą klasycznych metod identyfikacji obserwacji odstających: metody trzech sigm oraz metody kwartylowej, a także zaawansowanej metody opartej na modelu autoregresji zaproponowanej przez Chena i Liu w 1993 roku. Pracę rozpoczyna krótkie wprowadzenie i charakterystyka branży FMCG. Następnie przedstawiono definicje i rodzaje obserwacji odstających oraz używanych przy ich badaniu metod. Pracę kończy prezentacja wyników badań identyfikacji obserwacji odstających dla dwóch jednostek z użyciem wszystkich trzech metod oraz porównanie wyników badań.

Słowa kluczowe: obserwacja odstająca, FMCG, autoregresja, identyfikacja

Abstract. The aim of the article is to verify and analyse selected methods for the identification of outliers in the FMCG industry (fast-moving consumer goods). Real tests were used to carry out the sales data of the American chain of Walmart stores, made available by the company on the Internet¹. The analysis has been made using classic methods for identifying outliers: three sigma method and quartile method, as well as an advanced method based on the auto regression model proposed in the 1993 work of Chen and Liu. Article begins with a short introduction and characteristics of the FMCG industry. Next, the definitions, types of outliers and methods for identification are presented. The paper ends with the presentation of research results for two units using all three methods.

Key words: outliers, FMCG, auto regression, identification

Wstęp

W dzisiejszych czasach odpowiednie planowanie zapotrzebowania na produkty jest niezmiernie ważne. Zwłaszcza jest to istotne w branży FMCG (ang. *fast-moving consumer goods*), w której drobne odchylenia mogą spowodować braki magazynowe lub

konieczność utylizacji produktów, których nie udało się sprzedać ze względu na krótki czas przydatności do spożycia.

Planiści dokonując prognoz popytu, najczęściej wykorzystują dane historyczne sprzedaży. Należy jednak pamiętać, że w ciągu roku zdarzają się okresy, które znacząco odbiegają swoją wartością od reszty obserwacji. Takie zdarzenia nazywa się obserwacjami odstającymi. Nie powinny być brane pod uwagę w sposób bezpośredni podczas planowania zapotrzebowania bądź sprzedaży. W przeciwnym razie mogą one znacząco wpłynąć na wysokość prognoz.

Niniejsza praca ma na celu przeprowadzenie weryfikacji i analizy metod identyfikacji obserwacji odstających w branży FMCG. Do przeprowadzenia badań użyto rzeczywistych danych amerykańskiej sieci sklepów Walmart, udostępnionych przez firmę w Internecie¹. Analiza została wykonana za pomocą klasycznych metod identyfikacji obserwacji odstających, a także tych bardziej skomplikowanych, dostępnych z użyciem narzędzi informatycznych.

Przedstawiona praca ma również wskazać, jakie metody powinny być używane przy identyfikacji obserwacji odstających, oraz porównać klasyczne metody z tą zaproponowaną przez Chunga Chena i Lona Mu Liu.

Charakterystyka branży FMCG

Branża FMCG to część rynku zajmująca się produkcją i dystrybucją tak zwanych dóbr szybkozwywalnych – często nabywanych przez klientów. Jest to bardzo rozległa gama produktów obejmująca takie artykuły, jak: dobra spożywcze, napoje, alkohol i tytoń, chemia gospodarcza, kosmetyki.

Artykuły z branży FMCG to produkty charakteryzujące się przede wszystkim: krótkim terminem ważności produktu, szybką zbywalnością, małymi kosztami produkcji, małym zyskiem jednostkowym, choć przy sprzedaży dużych ilości zysk może być znaczny.

Powyższe cechy produktów stawiają wyzwania prognozującym zamówienie. Ich szybka zbywalność oraz wymóg dostępności produktu przez klienta powoduje, że odpowiednie planowanie zakupu tych artykułów jest niezwykle istotne. Nieodpowiednio przygotowana prognoza zapotrzebowania na dobra branży FMCG może spowodować nadwyżki magazynowe bądź też zupełnie odwrotnie – braki towarów w sklepach.

Obserwacje odstające

Analiza obserwacji odstających (także anomalia, obserwacja nietypowa, obserwacja oddalona, *outliers*) może sprawić wiele problemów osobom odpowiedzialnym za planowanie zaopatrzenia. To zjawisko rozumiane jest jako obserwację (lub ich cały zbiór), która jest niespójna z pozostałymi obserwacjami [Majewska 2012]. Przykładem ilustrującym obserwacje odstające może być wzrost drużyny koszykarzy. Większość z nich mierzy 195 cm. Jednakże wśród nich znajduje się niski gracz ma jedynie 165 cm wzrostu. Ten zawodnik mierzy zdecydowanie mniej od reszty swoich kolegów. Wzrost tego

¹ <https://www.kaggle.com/c/walmart-recruiting-store-sales-forecasting> [dostęp: 18.04.2018].

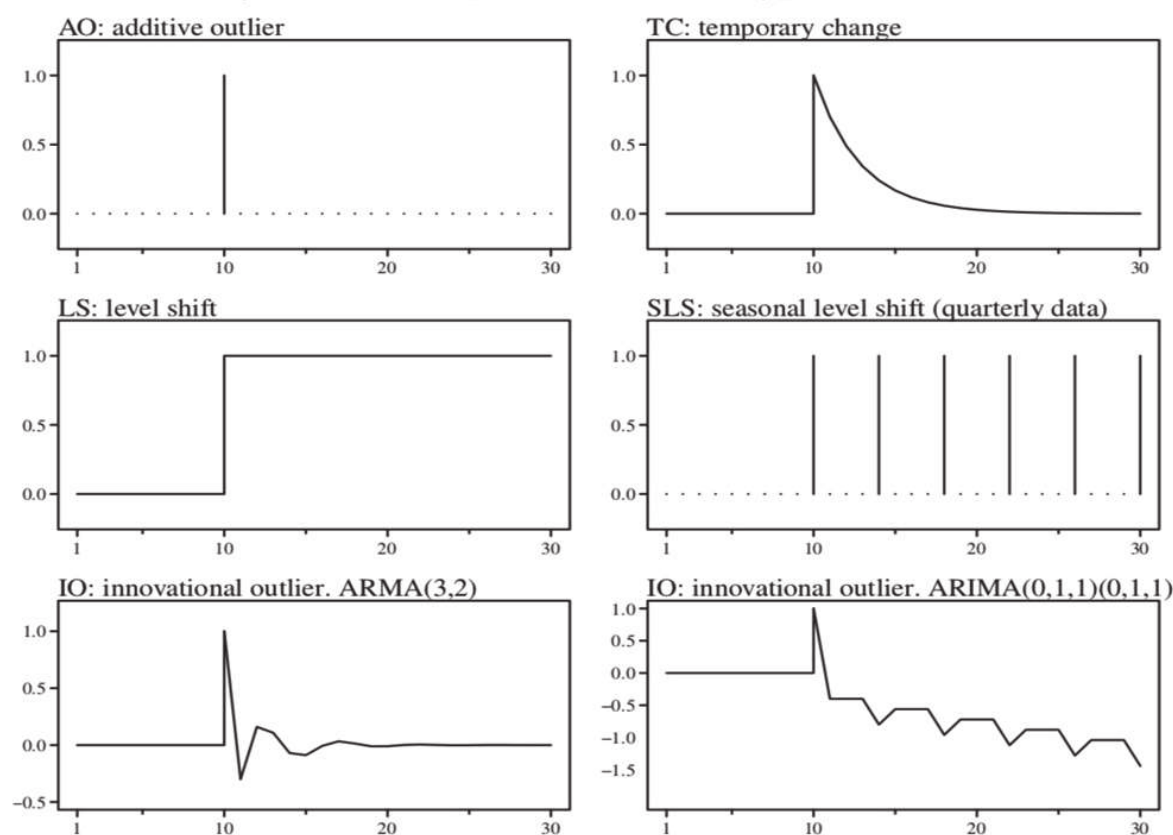
zawodnika możemy określić jako obserwację odstającą – jednostkę, która odbiega swoją wartością od większości obserwacji.

Charakterystyka branży FMCG sprawia, że powodów powstawania obserwacji odstających może być kilka. W szczególności do powodów anomalii można zaliczyć: promocje, wahania sezonowe, rozwój technologiczny, powodujący krótki cykl życia produktów, braki magazynowe, zamówienia oczekujące (ang. *back orders*).

Pierwszym z czynników mogących spowodować obserwacje odstające są promocje. Zalicza się je do powodów powstawania anomalii, ponieważ wynika to bezpośrednio ze specyfiki rynkowej FMCG. Codziennie w sklepach pojawiają się krótkoterminowe promocje na pojedyncze produkty.

Obserwacje odstające mogą powstawać także w dwóch przypadkach. Pierwszym z nich są braki magazynowe. W przypadku, kiedy przedsiębiorstwo nie ma dóbr na magazynie, w danym okresie straci możliwość do uzyskania wartość sprzedaży. Rozpatrując to w dłuższym horyzoncie, ten okres będzie obserwacją odstającą od reszty, ponieważ wartość sprzedaży będzie zdecydowanie mniejsza niż w okresach bez braków magazynowych.

Konsekwencją braków magazynowych, a jednocześnie kolejną przyczyną powstawania *outliers* są tak zwane zamówienia oczekujące. Jest to zamówienie przyjęte do realizacji z dostępnego zapasu, którego w danym momencie nie można zrealizować, ponieważ



Rysunek 1. Typy obserwacji odstających

Figure 1. Types of outliers

Źródło: López-de-Lacalle [2016].

zapas ten jest mniejszy niż wielkość zamówienia klienta². Takie zamówienie jest realizowane w następnym okresie. Z racji tego, że zamówienie przechodzi z poprzedniego okresu, powstaje obserwacja odstająca – wielkość zapasu normalnie przeznaczonego na dany okres jest powiększona o wielkość zamówienia przerastającego.

Outliers mają duży wpływ na prognozowanie zapotrzebowania na produkty. Przede wszystkim mają one wpływ na dane historyczne, co utrudnia prognozowanie na ich podstawie. Konsekwencją tego zjawiska są odchylenia od realnego popytu na produkty. Planisci rozpatrując dane zaburzone przez obserwacje odstające, mogą nie doszacować lub przeszacować poziom zapotrzebowania na dany produkt w rozpatrywanym okresie.

W przypadku rozpatrywania szeregów czasowych możemy wyróżnić sześć rodzajów obserwacji odstających [rys. 1, Chen i Liu 1993, López-de-Lacalle 2016]: zmiana jednorazowa (AO), zmiana czasowa (TC), zmiana poziomu (LS), zmiana sezonowa (SLS), mała zmiana innowacyjna (IO), duża zmiana innowacyjna (IO).

Klasyczne metody identyfikacji obserwacji odstających

Metoda trzech sigm

Pierwszą metodą klasyczną używaną do wykrycia obserwacji odstających w szeregu czasowym jest metoda trzech sigm. Zakłada ona, że dla danego rozkładu normalnego $N(\mu, \sigma)$ 99,7% wszystkich rozpatrywanych wartości zawiera się w przedziale:

$$(\mu - 3\sigma, \mu + 3\sigma)$$

gdzie:

μ – wartość oczekiwana rozkładu normalnego,
 σ – odchylenie standardowe rozkładu normalnego.

Ta metoda jest łatwa w użyciu. Dla danego szeregu czasowego należy policzyć średnią oraz odchylenie standardowe. Następnie należy weryfikować, czy dana wartość empiryczna należy do przedziału 3σ . Obserwacje nienależące do tego przedziału są obserwacjami odstającymi. Analizę metodą trzech sigm wykonuje się dla szeregów czasowych symetrycznych.

Wadą tej metody jest założenie, że wartość sprzedaży ma rozkład normalny, ale prostota tej metody powoduje jej częste użycie w prognozowaniu.

Metoda kwartyłowa

Kolejna metoda klasyczna używana do identyfikacji obserwacji odstających w szeregu czasowym wykorzystuje do analizy kwartyle. Przedział wyraża się wzorem:

$$[-1,5(Q_3 - Q_1) + Q_1 ; Q_3 + 1,5(Q_3 - Q_1)]$$

gdzie:

Q_1 – kwartył pierwszy,
 Q_3 – kwartył trzeci.

² https://www.logistyka.net.pl/slownik-logistyczny/szczegoly/1640,zamowienie_przerastajace [dostęp: 18.04.2018].

Kwartylem pierwszym oznacza się wartość w szeregu, poniżej której znajduje się 25% obserwacji. Kwartyl trzeci jest 75. percentylem, czyli poniżej jego wartości znajduje się 75% obserwacji [Aczel Amir 2005].

W tym sposobie należy wyznaczyć wartości kwartyli pierwszego oraz trzeciego. Następnie oblicza się rozstęp międzykwartyłowy oraz przemnaża się go przez 1,5. W zależności od obliczania granicy przedziału do wyniku dodaje się:

- w przypadku dolnej granicy przedziału – wartość kwartyla pierwszego,
- do górnej granicy – wartość kwartyla trzeciego.

Za obserwacje odstające uważa się wszystkie te, które nie zawierają się w ustalonym przedziale. Analizę metodą kwartyłową wykonuje się przede wszystkim dla szeregów niesymetrycznych.

Zaawansowana metoda zaproponowana przez Chena i Liu

Metoda zaproponowana przez Chunga Chena i Lona Mu Liu w 1993 roku jest bardziej zaawansowanym sposobem identyfikacji obserwacji odstających. Jest ona metodą dedykowaną i jest ściśle powiązana z modelami ARMA oraz ARIMA [RoqueWave 2016]. Model ARMA składa się z powiązania modelu autoregresji oraz średniej ruchomej. Z kolei ARIMA jest modelem autoregresji zintegrowanym ze średnią ruchomą.

Model Chena i Liu [1993] jest iteracyjnym procesem, który można opisać w trzech [Trzęsiok 2016] powtarzających się krokach:

- krok 1 – oszacowanie parametrów modelu ARIMA dla szeregu czasowego;
- krok 2 – mając wyestymowane parametry z krok 1, należy założyć, że we wszystkich momentach, które są zaburzeniem szeregu, wystąpiła obserwacja odstająca. Należy oszacować wielkość odchylenia tej obserwacji. Jeżeli przekraczała ona ustaloną wcześniej wartość progową (czyli błędu założonego przy układania modelu), to należy przyjąć, że jest to obserwacja odstająca;
- krok 3 – należy usunąć efekt wystąpienia obserwacji przez odjęcie od wielkości obserwacji odstającej wartości odchylenia, a następnie należy skorygować wszystkie następne obserwacje zgodnie z modelem ARIMA, zbudowanym w kroku 1.

Model ten jest skomplikowany i złożony³.

Identyfikacja obserwacji odstających – studium przypadku

Jak zostało wspomniane, analiza i interpretacja *outliers* w branży FMCG jest kluczowa przy planowaniu zamówień artykułów. W tej części zostanie przedstawiona analiza i interpretacja obserwacji odstających występujących w sieci sklepów Walmart z użyciem wspomnianych wcześniej metod.

Dane udostępnione przez Walmart dotyczą przedziału czasowego od 05.02.2010 r. do 26.10.2012 r. Są one dostępne dla 45 sklepów sieci, z podziałem na około 98 departamentów (w zależności od sklepu). Do badań zostały użyte dane dla dwóch wybranych sklepów i jednego departamentu występującego w każdym z nich. Dane, które zostały użyte do identyfikacji obserwacji odstających, dotyczą wielkości sprzedaży w danym departamencie.

³ Artykuł twórców koncepcji (Chen i Liu) dokładnie przedstawia wszystkie założenia wraz z ich wyjaśnieniem.

Walmart w swoich danych nie ujawnił konkretnie, które sklepy i departamenty są w nich zawarte, dlatego nie jest możliwe dokładne stwierdzenie, jaki departament i sklep są rozpatrywane. Jedynie po analizie wielkości sprzedaży oraz typach *outliers* w danych tygodniach można domyślać się, czym dany departament się zajmuje.

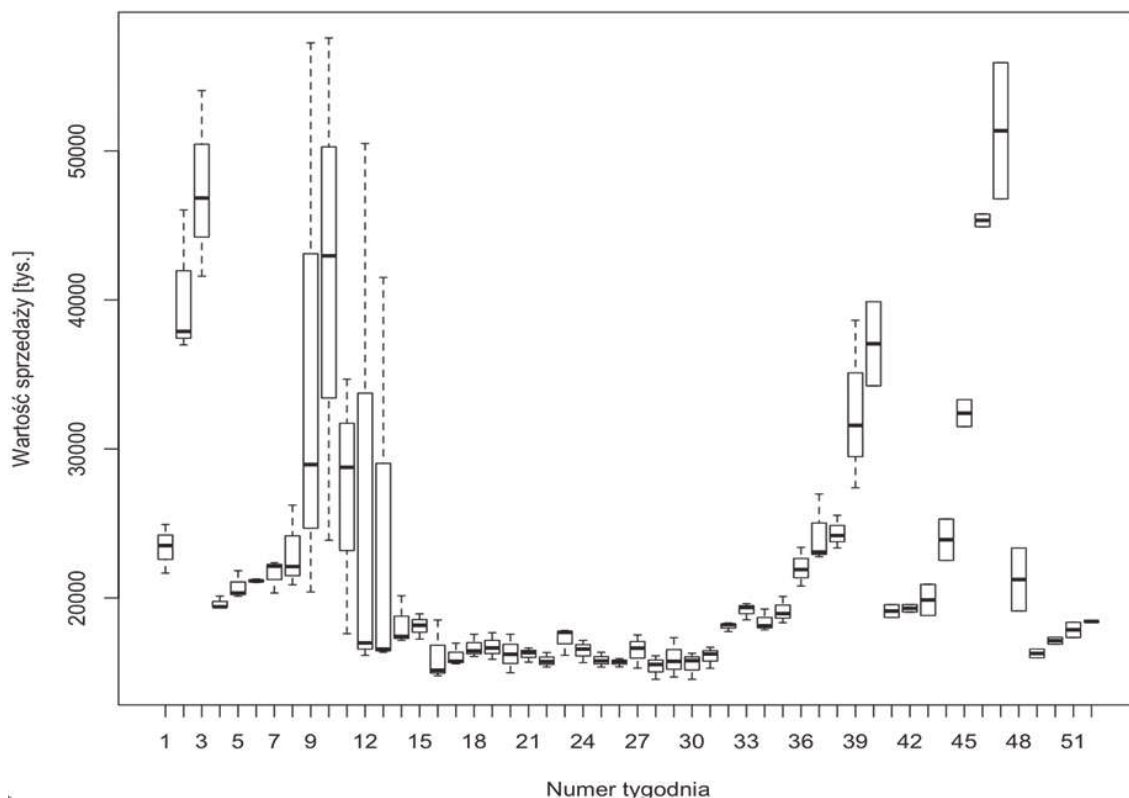
Analiza podstawowych danych statystycznych dla sklepów Walmart

W pierwszym kroku analizy wykorzystujemy podstawowe dane statystyczne. Dla każdej badanej jednostki przedstawione zostaną dane wielkości sprzedaży w rozpatrywanym okresie. Zaprezentowane one zostaną w dwóch graficznych formach: wykresach pudełkowych oraz wykresach liniowych szeregów czasowych.

Wykresy pudełkowe

Wykresy te to graficzne przedstawienie podstawowych wartości statystycznych. Z takich wykresów łatwo odczytać takie wielkości, jak: pierwszy i trzeci kwartył, medianę, wartości maksymalne i minimalne.

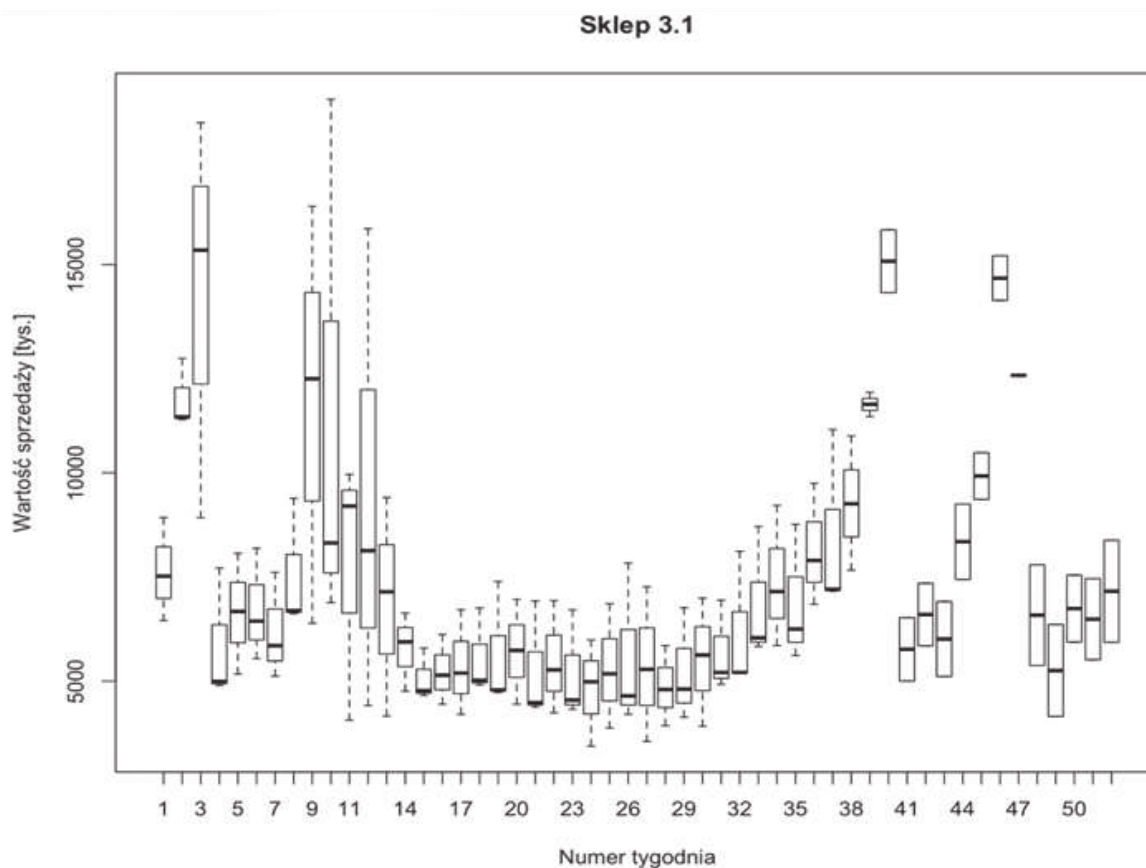
Na rysunkach 2 i 3 zaprezentowano wielkość sprzedaży danego departamentu oraz numer tygodnia, w którym wystąpiła dana sprzedaż. Z racji zróżnicowania w całkowitym obrocie sklepu wykresy mają różną skalę. Oznaczenia na wykresach dotyczą: liczba pierwsza – numer sklepu, liczba druga – numer departamentu.



Rysunek 2. Wykres pudełkowy sprzedaży w sklepie 1.1

Figure 2. A box chart of sales in the store 1.1

Źródło: opracowanie własne w pakiecie R.



Rysunek 3. Wykres pudełkowy sprzedaży w sklepie 3.1

Figure 3. A box chart of sales in the store 3.

Źródło: opracowanie własne w pakiecie R.

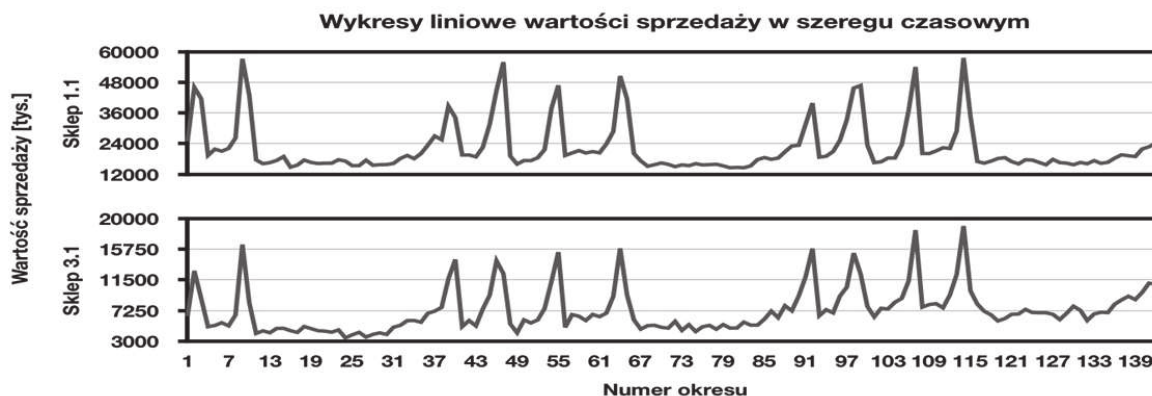
Po analizie wykresów danych wnioskuje się, że:

- departament pierwszy zajmuje się sprzedażą artykułów, które lepiej sprzedają się w okresie zimowym; w początkowych i ostatnich tygodniach roku kalendarzowego wahania poziomu sprzedaży są największe;
- sklep numer trzy jest sklepem zdecydowanie mniejszym od sklepu numer jeden; poziom przychodów (nawet maksymalnych) tej jednostki jest zdecydowanie niższy niż w pierwszym sklepie.

Wykresy pudełkowe to proste narzędzie, dzięki którym możliwe jest uzyskanie podstawowych informacji statystycznych o badanej próbie. Ich graficzna forma pozwala na łatwe i przejrzyste interpretacje danych empirycznych.

Wykres szeregu czasowego dla wielkości sprzedaży w Walmart

Kolejnym podstawowym narzędziem obrazującym dane jest wykres szeregów czasowych. Umożliwia on zobrazowanie dynamiki danych – pokazuje, jakie wartości przyjmuje badana zmienna co okres. Dzięki temu wykresowi łatwo można zauważyć powstawanie odchyleń lub inne nieprzewidziane zjawiska.



Rysunek 4. Wykresy liniowe wartości sprzedaży

Figure 4. Line charts of sales value

Źródło: opracowanie własne.

Rysunek 4 przedstawia, jak kształtowała się wielkość sprzedaży w przeciągu badanego okresu w rozpatrywanych sklepach. Łatwo można z nich zauważyć, że operacje odstające występują. W dalszej części pracy z użyciem wcześniej opisanych metod podjęto próbę zidentyfikowania typów *outliers* oraz interpretacji ich możliwego źródła pochodzenia.

Identyfikacja obserwacji odstających

Metoda trzech sigm oraz metoda kwartyłowa

Tabela 1 przedstawia wyniki badań przeprowadzonych na danych sieci Walmart z użyciem metody trzech sigm oraz metody kwartyłowej. W tabeli 1 pokazane zostały oznaczenia analizowanych jednostek, wartość przedziału 3σ i przedziału kwartyłowego oraz liczbę tygodni, które są według tych metod obserwacjami odstającymi.

Tabela 1. Wyniki analizy danych metodami klasycznymi

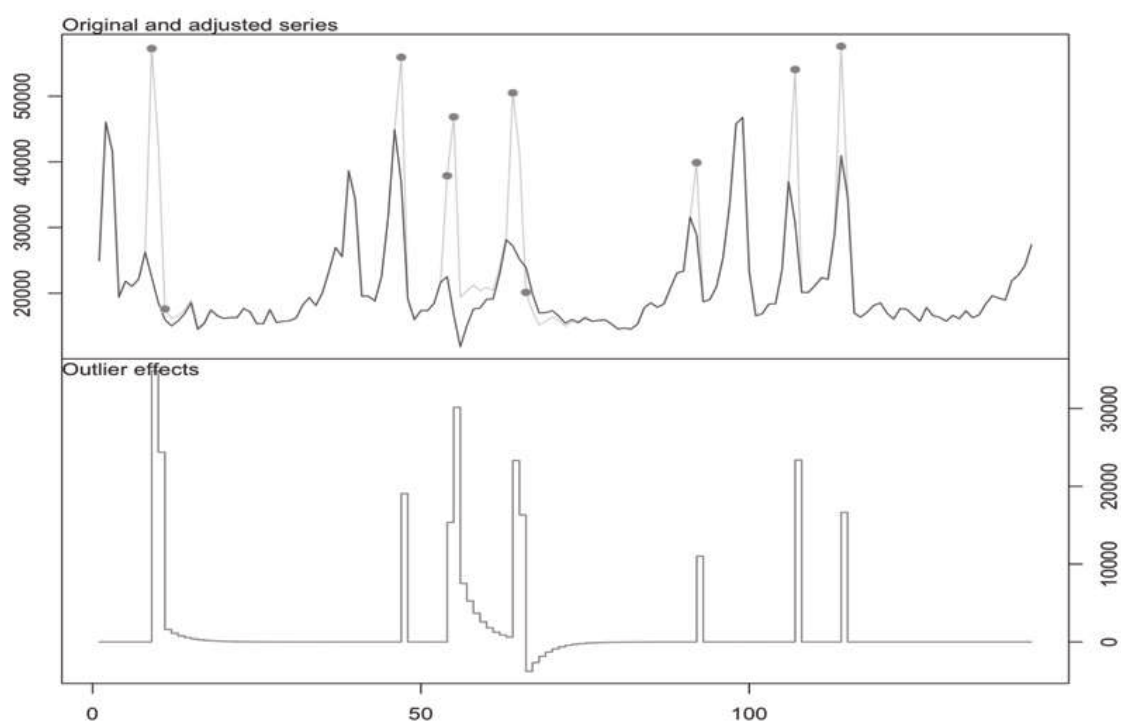
Table 1. The results of data analysis by classical methods

Numer sklepu	Przedział sigmowy	Przedział kwartyłowy	Liczba <i>outliers</i> dla 3σ	Liczba <i>outliers</i> dla kwartyli
1.1	(-6 946,2; 51 972,8)	(6 415,3; 33 293,6)	4	20
3.1	(-1 989,58; 16 646,83)	(443,81; 12 940,77)	2	9

Źródło: opracowanie własne.

Metoda zaprezentowana przez Chena i Liu

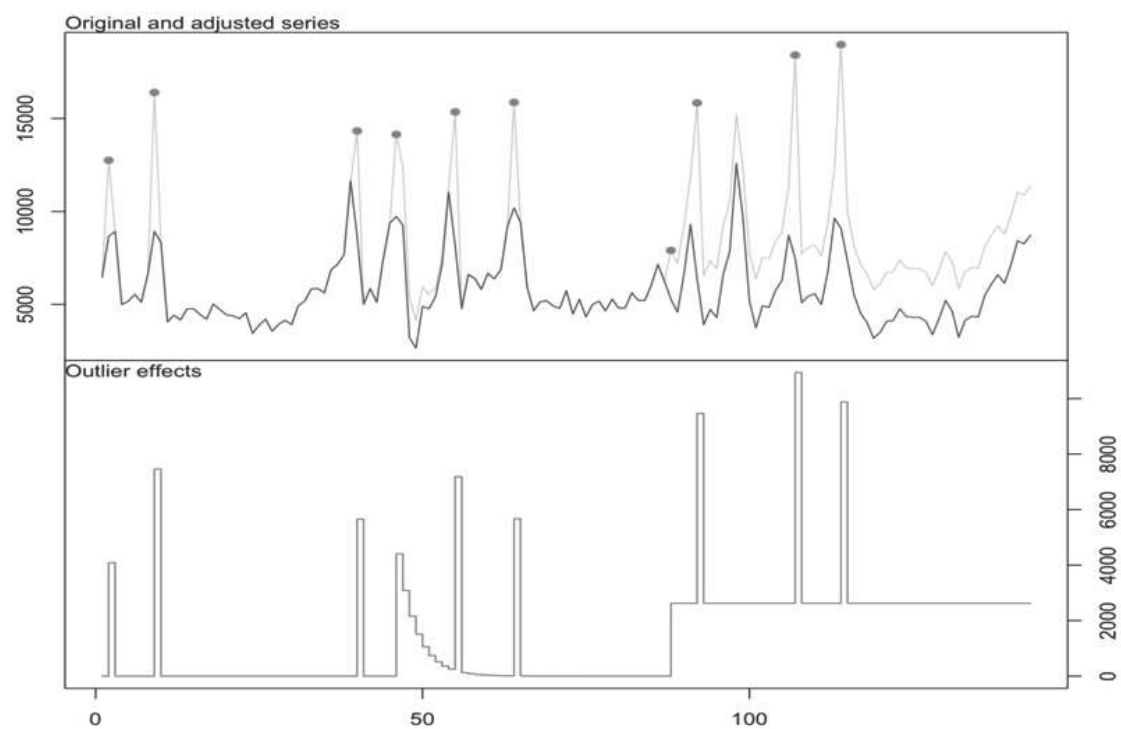
Na rysunkach 5 i 6 przedstawiono wykaz tygodni oraz typów obserwacji odstających, które wystąpiły w badanym okresie w sklepach Walmart. W pracy zostaną przedstawione wykresy *outliers* tylko dla dwóch wybranych jednostek wraz z tabelami przedstawiającymi wyniki badania obserwacji odstających (tab. 2, 3).



Rysunek 5. Wykres outliers dla jednostki 1.1

Figure 5. Outliers chart for the unit 1.1

Źródło: opracowanie własne w pakiecie R.



Rysunek 6. Wykres outliers dla jednostki 3.1

Figure 6. Outliers chart for the unit 3.1

Źródło: opracowanie własne w pakiecie R.

Rysunki 5 i 6 przedstawiają w dokładny sposób obserwacje odstające – są one zaznaczone kropkami. Wykresy górne rysunków 5 i 6 pokazują oryginalny szereg czasowy z wartościami sprzedaży oraz szereg z danymi oczyszczonymi z *outliers*. Szereg oczyszczony jest zaznaczony ciemniejszą linią. Z kolei wykresy dolne rysunków 5 i 6 przedstawiają typy obserwacji odstających.

Tabela 2. Interpretacja obserwacji odstających dla jednostki 1.1

Table 2. Interpretation of outliers for the entity 1.1

Lp.	Typ obserwacji	Tydzień	Coefhat	Tstat
1	TC	9	34 833	8,755
2	TC	11	-15 469	-3,925
3	AO	47	19 056	9,014
4	TC	54	15 360	3,547
5	AO	55	19 402	6,946
6	TC	64	22 884	5,864
7	TC	66	-15 208	-3,905
8	AO	92	11 019	5,048
9	AO	107	23 369	11,058
10	AO	114	16 638	7,141

Źródło: opracowanie własne w pakiecie R.

Tabela 3. Interpretacja obserwacji odstających dla jednostki 3.1

Table 3. Interpretation of outliers for the entity 3.1

Lp.	Typ obserwacji	Tydzień	Coefhat	Tstat
1	AO	2	4087	4,064
2	AO	9	7461	7,810
3	AO	40	5656	6,312
4	TC	46	4406	3,491
5	AO	55	7008	7,800
6	AO	64	5665	6,219
7	LS	88	2624	4,854
8	AO	92	6843	7,321
9	AO	107	8313	9,070
10	AO	114	7253	8,056

Źródło: opracowanie własne w pakiecie R.

Porównanie metod

Analizując wyniki badań, wyciągnięto kilka wniosków dotyczących metod identyfikacji *outliers* w branży FMCG:

- metoda trzech sigm wskazuje najmniej tygodni, w których powstają operacje odstające;

- metoda kwartylowa wskazuje najwięcej tygodni, w których występują obserwacje odstające; jest to spowodowane tym, że dane wielkości sprzedaży z założenia są asymetryczne;
- model Chena i Liu jest najbardziej dokładną metodą spośród przedstawionych.

Dzięki analizie danych sposobem zaproponowanym przez Chena i Liu przedsiębiorstwa mogą dowiedzieć się, jakie dokładnie typy *outliers* pojawiły się w szeregu czasowym, a także mogą poznać wielkość odchylenia i okres, w którym ono powstało. Umożliwia to firmom lepszą interpretację zjawisk obserwacji odstających, co w konsekwencji pozwala się zabezpieczyć przedsiębiorstwom przed podobnymi wypadkami w przyszłości. Jednakże wadą tego modelu jest jego wysoki poziom skomplikowania.

Tabela 4 przedstawiono podsumowanie wyników badań wszystkimi rozpatrywanymi sposobami identyfikacji obserwacji odstających. Przedstawia ona porównanie liczby wskazanych obserwacji przez każdą z trzech metod dla wszystkich badanych jednostek.

Tabela 4. Porównanie liczby obserwacji odstających

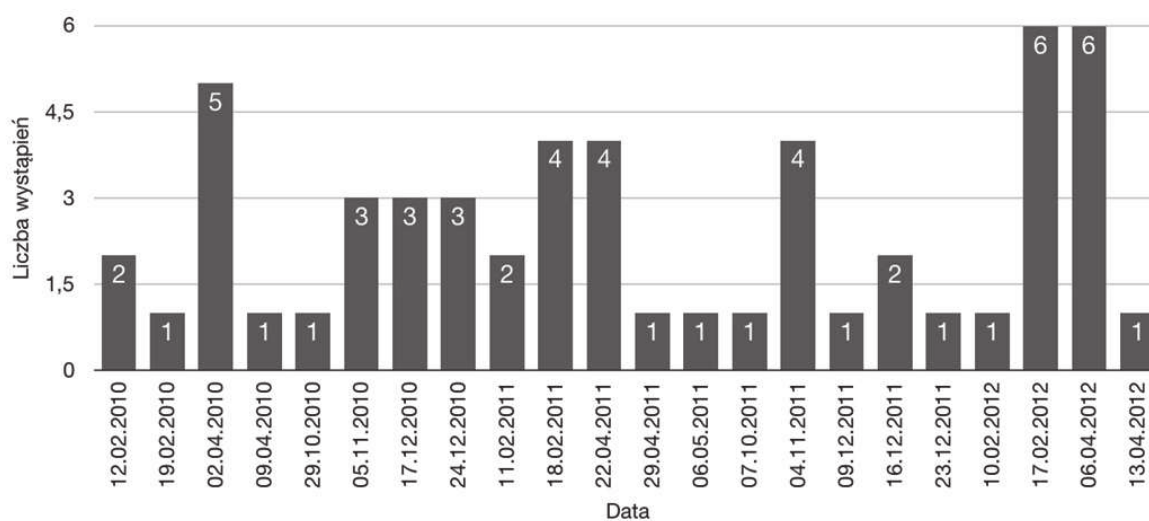
Table 4. Comparison of the number of outliers

Numer sklepu	Metoda trzech sigm	Metoda kwartylowa	Model Chena i Liu
1.1	4	20	10
3.1	2	9	10

Źródło: opracowanie własne.

Interpretacja wyników badań

Rysunek 7 przedstawia, ile razy dany tydzień został zidentyfikowany jako obserwacja odstająca. Wykres uwzględnia identyfikację tygodnia przez wszystkie trzy rozpatrywane metody łącznie. Każdy z tygodni mógł zostać wskazany maksymalnie 6 razy.



Rysunek 7. Liczba wystąpień danego tygodnia jako odstający

Figure 7. The number of instances of the week as outliers

Źródło: opracowanie własne.

Po porównaniu rysunku 7 z kalendarzem, można zauważyć, że:

- najczęściej jako *outliers* rozpoznane zostały tygodnie z datami 17.02.2012 r. oraz 06.04.2012 r. – są to tygodnie, na które przypadają odpowiednio Super Bowl (finał amerykańskiej ligi futbolu) oraz święta wielkanocne; daty te zostały wskazane jako odstające przez wszystkie metody w obu badanych sklepach;
- tydzień 02.04.2010 r. to również święta wielkanocne; tydzień ten osiągnął 5 wskazań.

Tak jak można było się spodziewać, obserwacje odstające w branży FMCG powstają najczęściej przy okazji świąt i dni wolnych od pracy. Jednakże zauważalne są także inne przyczyny powstawania *outliers*.

Na rysunku 5 widać, że w 88. analizowanym tygodniu, który odpowiada 07.10.2011 r., nastąpiła zmiana poziomu sprzedaży. Średnia wielkość sprzedaży wzrosła względem poprzednich tygodni. Takie zjawisko może przykładowo nastąpić w wyniku podpisania długoterminowego kontraktu na zakup dóbr od dostawcy po niższej niż wcześniej cenie.

Podsumowanie

Praca ta miała na celu przybliżenie statystycznych metod identyfikacji obserwacji odstających w szeregach czasowych na podstawie danych z branży FMCG. Każda z przedstawionych metod wskazywała na różną liczbę *outliers* w badanym okresie. Wynika to przede wszystkim z założeń tych metod. Najbardziej dokładnym z przedstawionych sposobów jest metoda zaprezentowana przez Chena i Liu, która nie tylko wskazuje na istnienie obserwacji odstającej, ale także na typ *outliers*, co ułatwia interpretację zaistniałej sytuacji.

Model zaproponowany przez Chena i Liu jest ciekawym i rzetelnym narzędziem używanym do prognozowania poziomu sprzedaży. Mimo jego wysokiego poziomu skomplikowania warto zapoznać się z tą metodą, ponieważ dostarcza ona wielu informacji przydatnych przy planowaniu zapotrzebowania i prognozowaniu poziomu sprzedaży.

Literatura

- Aczel Amir D., 2005: Statystyka w zarządzaniu, Wydawnictwo Naukowe PWN, Warszawa.
- Chen Ch., Liu L-M., 1993: Joint estimation of model parameters and outlier effects in time series, *Journal of the American Statistical Association* 88 (421), 284–297.
- Logistyka.net.pl. Słownik logistyczny, hasło: zamówienie przerastające, [źródło elektroniczne] https://www.logistyka.net.pl/slownik-logistyczny/szczegoly/1640,zamowienie_przerastajace [dostęp: 18.04.2018].
- López-de-Lacalle J., 2016: tsoutliers R Package for Detection of Outliers in Time Series, [źródło elektroniczne] <https://jalobe.com/doc/tsoutliers.pdf> [dostęp: 18.04.2018].
- Majewska J., 2012: Obserwacje odstające w (dużych) zbiorach danych, *Statystyka a Data Science*, [w:] G. Trzpiot (red.), Wydawnictwo Uniwersytetu Ekonomicznego w Katowicach, Katowice.

- RogueWave.com, 2016: Tech tutorial: embedding analytics into a database using SourcePro and JMSL, [źródło elektroniczne] <https://www.roguewave.com/sites/rw/files/attachments/IMSL-AutoARIMA-WP.pdf> [dostęp: 18.04.2018].
- Trzęsiok M., 2016: Identyfikacja obserwacji oddalonych w szeregach czasowych, *Studia Ekonomiczne, Zeszyty Naukowe Uniwersytetu Ekonomicznego w Katowicach* 265, 95–105.
- Walmart Recruiting – Store Sales Forecasting, [źródło elektroniczne] <https://www.kaggle.com/c/walmart-recruiting-store-sales-forecasting> [dostęp: 18.04.2018].

Adres do korespondencji:

dr hab. Paweł Hanczar, prof. UE
(<https://orcid.org/0000-0002-5144-7975>)
Uniwersytet Ekonomiczny we Wrocławiu
Wydział Zarządzania, Informatyki i Finansów
Katedra Logistyki
ul. Komandorska 118/120, 53-345 Wrocław
tel. (+48) 71 36 80 345
e-mail: pawel.hanczar@ue.wroc.pl

Jacek Zasada

(<https://orcid.org/0000-0002-8993-9414>)

Paweł Sobolik

(<https://orcid.org/0000-0002-2365-116X>)