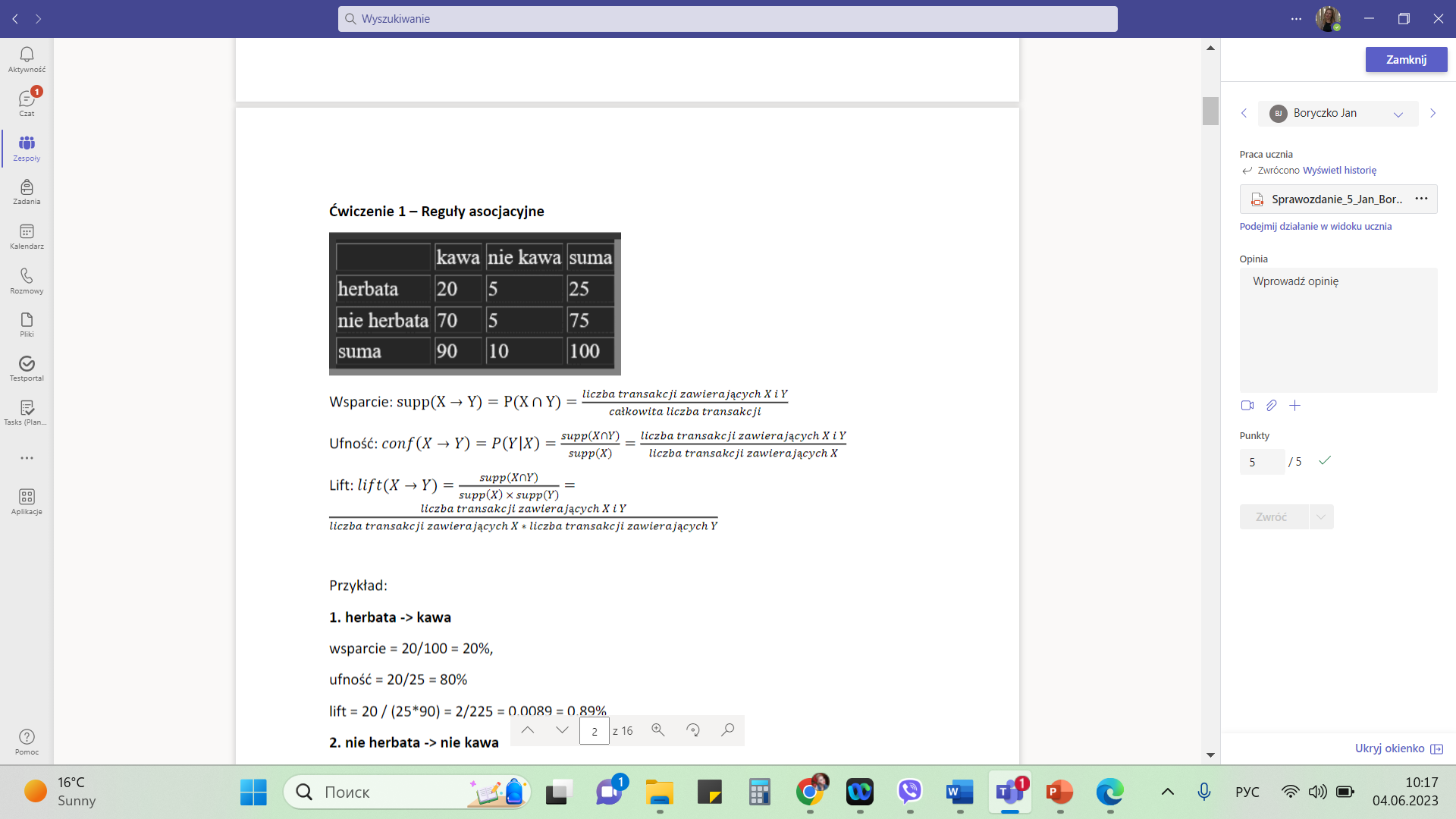
**Reguły asocjacyjne**

**Reguły asocjacyjne**

Rozważmy tabelę przedstawiającą wyniki ankiety dotyczącej preferencji klientów w zakresie herbaty i kawy:

|  |  |  |  |
| --- | --- | --- | --- |
|  | kawa | nie kawa | suma |
| herbata | 20 | 5 | 25 |
| nie herbata | 70 | 5 | 75 |
| suma | 90 | 10 | 100 |



Bezpośrednim rachunkiem oblicz wsparcie, ufność, lift dla reguł, które można utworzyć. Zbuduj odpowiedni plik z danymi i wykonaj to samo w Rattle

**Reguły asocjacyjne w Rattle**

Do zilustrowania analizy koszyka w Rattle, możemy użyć bardzo prostego zestawu danych składający się z filmów DVD zakupionych przez klientów. Dane są dostępne w postaci pliku CSV (dvdtrans.csv) z pakietu Rattle.  
Najprostszym sposobem, aby załadować ten zestaw danych do Rattle jest załadowanie najpierw domyślnego zbioru weather z pliku weather.csv plik do Rattle.

1. Klikając na przycisk Execute w zakładce Data (zaraz po starcie pakietu).
2. Następnie klikamy na przycisk Filename (który będzie pokazywać weather.csv), aby wyświetlić zawartość folderu CSV pakietu Rattle.
3. Wybieramy dvdtrans.csv i klikamy przycisk Open , a następnie Execute.
4. Zmienna ID automatycznie zostanie wybrana jako Ident, ale musimy zmienić rolę zmiennej Item na Target (rysunek).
5. Na karcie Associate, upewniamy się, że przycisk Baskets jest zaznaczony. Kliknijmy przycisk Execute, aby zbudować model, który będzie składał się z kolekcji reguł asocjacyjnych.

**Przeanalizować reguły asocjacyjne dla sklepu warzywnego.**

Wczytać do Rattle dane z pliku warzywniak.csv. I dokonać analizy reguł asocjacyjnych. Dane z pliku można wcześniej podejrzeć w edytorze lub arkuszu kalkulacyjnym.

**Reguły asocjacyjne w R**

Wykonać następujący kod tworzący reguły asocjacyjne w R:

library(arules)

library(rattle)

dvdtrans <- read.csv(system.file("csv","dvdtrans.csv", package="rattle"))

dvdDS <- new.env()

dvdDS$data<- as(split(dvdtrans$Item, dvdtrans$ID),"transactions")

dvdDS$data

dvdAPRIORI <- new.env(parent=dvdDS)

evalq({

model <- apriori(data, parameter=list(

support=0.2, confidence=0.1))

}, dvdAPRIORI)

Reguły mogą być wyświetlone i posortowane wg ufności za pomocą funkcji

inspect(sort(dvdAPRIORI$model, by="confidence")[1:5])

**Reguły asocjacyjne w R**

W R zbiór transakcji otwieamy przy pomocy funkcji:

read.transactions()

tak więc chcąc wczytać plik z produktami spożywczymi należy

g<- read.transactions("groceries.csv", sep = ",")

Wykonujemy funkcję apriori

reguly.zakupy <- apriori(g, parameter = list(supp = 0.001, conf = 0.8))

Następnie możemy wyświetlić reguły, bądź dokonać sumaryzacji

inspect(reguly.zakupy[1:7])

summary(reguly.zakupy)

Reguły można utworzyć również według wybranej wartości (tutaj mamy pytanie:Co kupują klienci, którzy kupili mleko?)

r<-apriori(data=g, parameter=list(supp=0.001,conf = 0.15,minlen=2),

appearance = list(default="rhs",lhs="whole milk"),

control = list(verbose=F))

r<-sort(r, decreasing=TRUE,by="confidence")

inspect(r[1:5])

Utworzone reguły można zapisać do pliku

write(reguly.zakupy, file = "reguly.csv", quote=TRUE, sep = ",", col.names = NA)

write(r, file = "regulymleko.csv", quote=TRUE, sep = ",", col.names = NA)

**Praca domowa**

Dokonaj analizy reguł asocjacji zbioru groceries, według poleceń ze strony: [Analiza koszyka](http://www.salemmarafi.com/code/market-basket-analysis-with-r/) Przyslij dwa pliki. Jeden z regułami wskazującymi co kupują klienci, którzy kupili masło a drugi z listą reguł, wskazujących, które towary kupują klienci wraz z masłem.