

```
library("RMariaDB")
library(tidyverse)
library(ggplot2)
library(eeptools)
library(scales)

con <- dbConnect(RMariaDB::MariaDB(),
                  dbname = "team07",
                  username = "team07",
                  password = "te@m24ot",
                  host = "giniewicz.it")

query <- "SELECT * FROM klienci"
```

## 1 Odsetek naprawianych marek pojazdów

Przeprowadzono analizę, aby sprawdzić jakich marek pojazdy najczęściej pojawiają się w warsztacie do naprawy. Przyjęto, że każdy samochód lub motor jest liczony tylko raz, nawet jeżeli był kilkakrotnie naprawiany.

```
query <- "SELECT DISTINCT marka,
ROUND(100*count(marka)/SUM(COUNT(DISTINCT id_samochodu)) OVER (), 2) as procent
FROM usługi JOIN samochody
ON usługi.id_samochodu = samochody.id
GROUP BY marka ORDER BY count(marka) DESC"

df <- dbGetQuery(con, query)
df[,2] <- as.numeric(df[,2])

minimalne <- df %>% filter(procent == min(procent))
maksymalne <- df %>% filter(procent == max(procent))

minimalne_tekst <- if_else(length(minimalne[,1])>1, paste(paste0(minimalne[,1][2:length(minimalne[,1])]),
maksymalne_tekst <- if_else(length(maksymalne[,1])>1, paste(paste0(maksymalne[,1][2:length(maksymalne[,1])]),
ranking <- paste0(seq_len(nrow(df)), ". ", df$marka, ": ", paste(df$procent,"%", sep=""), collapse = "\n")
```

Można przedstawić procentowy udział marek pojazdów klientów salonu, zaczynając od najczęściej się pojawiającej:

1. Volkswagen: 14.22%
2. Opel: 11.61%
3. Ford: 8.77%
4. Renault: 6.16%
5. Audi: 6.16%
6. Hyundai: 4.98%
7. Skoda: 4.98%
8. Fiat: 4.74%
9. Bajaj: 4.27%
10. Mercedes-Benz: 3.79%
11. SEAT: 3.08%
12. BMW: 3.08%

13. Citroen: 3.08%
14. Honda: 2.61%
15. Suzuki: 2.37%
16. Kia: 2.37%
17. Nissan: 2.13%
18. Peugeot: 2.13%
19. Mazda: 1.9%
20. Yamaha: 1.66%
21. TVS: 1.42%
22. Hero: 1.42%
23. Toyota: 1.42%
24. Volvo: 1.18%
25. Mitsubishi: 0.95%
26. Dacia: 0.71%
27. Porsche: 0.71%
28. Royal: 0.47%
29. Chevrolet: 0.24%
30. Mahindra: 0.24%
31. Jaguar: 0.24%
32. SsangYong: 0.24%
33. MINI: 0.24%
34. Dodge: 0.24%
35. Vespa: 0.24%
36. Microcar: 0.24%
37. Jeep: 0.24%

### PROCENTY ZNAK

Jak widać, najpopularniejsze są pojazdy marki Volkswagen. Pojazdy tej marki stanowią 14.22 wszystkich.

Najmniej popularne są marki Mahindra, Jaguar, SsangYong, MINI, Dodge, Vespa, Microcar, Jeep i Chevrolet. Każdą z nich reprezentowało 0.24 pojazdów, które się pojawiły w warsztacie.

Sprawdzono także, jak prezentowałyby się rozkład marek, gdyby brano pod uwagę jedynie samochody

```
query <- "SELECT DISTINCT marka,
ROUND(100*count(marka)/SUM(COUNT(DISTINCT id_samochodu)) OVER (), 2) as procent
FROM usługi JOIN (SELECT id, marka FROM samochody WHERE liczba_kół='4') as samochód
ON usługi.id_samochodu = samochód.id
GROUP BY marka ORDER BY count(marka) DESC"

df <- dbGetQuery(con, query)
df[,2] <- as.numeric(df[,2])

minimalne <- df %>% filter(procent == min(procent))
maksymalne <- df %>% filter(procent == max(procent))

minimalne_tekst <- if_else(length(minimalne[,1])>1, paste(paste0(minimalne[,1][2:length(minimalne[,1])])

maksymalne_tekst <- if_else(length(maksymalne[,1])>1, paste(paste0(maksymalne[,1][2:length(maksymalne[,1])])

df[,2] <- paste(df[,2], "%", sep="")
ranking <- paste0(seq_len(nrow(df)), ". ", df$marka, ": ", df$procent, collapse = "\n")
```

Można przedstawić ranking marek, zaczynając od najczęściej się pojawiającej:

1. Volkswagen: 16.13%

2. Opel: 13.17%
3. Ford: 9.95%
4. Audi: 6.99%
5. Renault: 6.99%
6. Skoda: 5.65%
7. Hyundai: 5.65%
8. Fiat: 5.38%
9. Mercedes-Benz: 4.3%
10. Citroen: 3.49%
11. BMW: 3.49%
12. SEAT: 3.49%
13. Kia: 2.69%
14. Peugeot: 2.42%
15. Nissan: 2.42%
16. Mazda: 2.15%
17. Suzuki: 1.88%
18. Toyota: 1.61%
19. Volvo: 1.34%
20. Mitsubishi: 1.08%
21. Dacia: 0.81%
22. Porsche: 0.81%
23. Honda: 0.81%
24. SsangYong: 0.27%
25. Microcar: 0.27%
26. Jaguar: 0.27%
27. Dodge: 0.27%
28. Chevrolet: 0.27%
29. MINI: 0.27%
30. Jeep: 0.27%

Wśród aut prym wiedzie Volkswagen, reprezentując 16.13 naprawianych samochodów.

Najmniej było samochodów marek Microcar, Jaguar, Dodge, Chevrolet, MINI, Jeep i SsangYong. Było ich 0.27 dla każdej.

Pozostały do sprawdzenia motory.

```
query <- "SELECT DISTINCT marka,
ROUND(100*count(marka)/SUM(COUNT(DISTINCT id_samochodu)) OVER (), 2) as procent
FROM usługi JOIN (SELECT id, marka FROM samochody WHERE liczba_kół='2') as motor
ON usługi.id_samochodu = motor.id
GROUP BY marka ORDER BY count(marka) DESC"

df <- dbGetQuery(con, query)
df[,2] <- as.numeric(df[,2])

minimalne <- df %>% filter(procent == min(procent))
maksymalne <- df %>% filter(procent == max(procent))

minimalne_tekst <- if_else(length(minimalne[,1])>1, paste(paste0(minimalne[,1][2:length(minimalne[,1])])

maksymalne_tekst <- if_else(length(maksymalne[,1])>1, paste(paste0(maksymalne[,1][2:length(maksymalne[,1])])

df[,2] <- paste(df[,2], "%", sep="")
```

```
ranking <- paste0(seq_len(nrow(df)), ". ", df$marka, ": ", df$procent, collapse = "\n")
```

Udział procentowy poszczególnych marek wśród nich prezentuje się następująco:

1. Bajaj: 36%
2. Honda: 16%
3. Yamaha: 14%
4. TVS: 12%
5. Hero: 12%
6. Suzuki: 6%
7. Royal: 4%
8. Vespa: 2%
9. Mahindra: 2%

Wśród nich najczęściej naprawiano motory marki Bajaj. Pojazdy tej marki stanowią 36.00 wszystkich.

Najmniej popularnymi motorami są z kolei modele marek Mahindra i Vespa. W warsztacie było zaledwie 2.00 motorów, tej marki.

## 2 Analiza bilansu

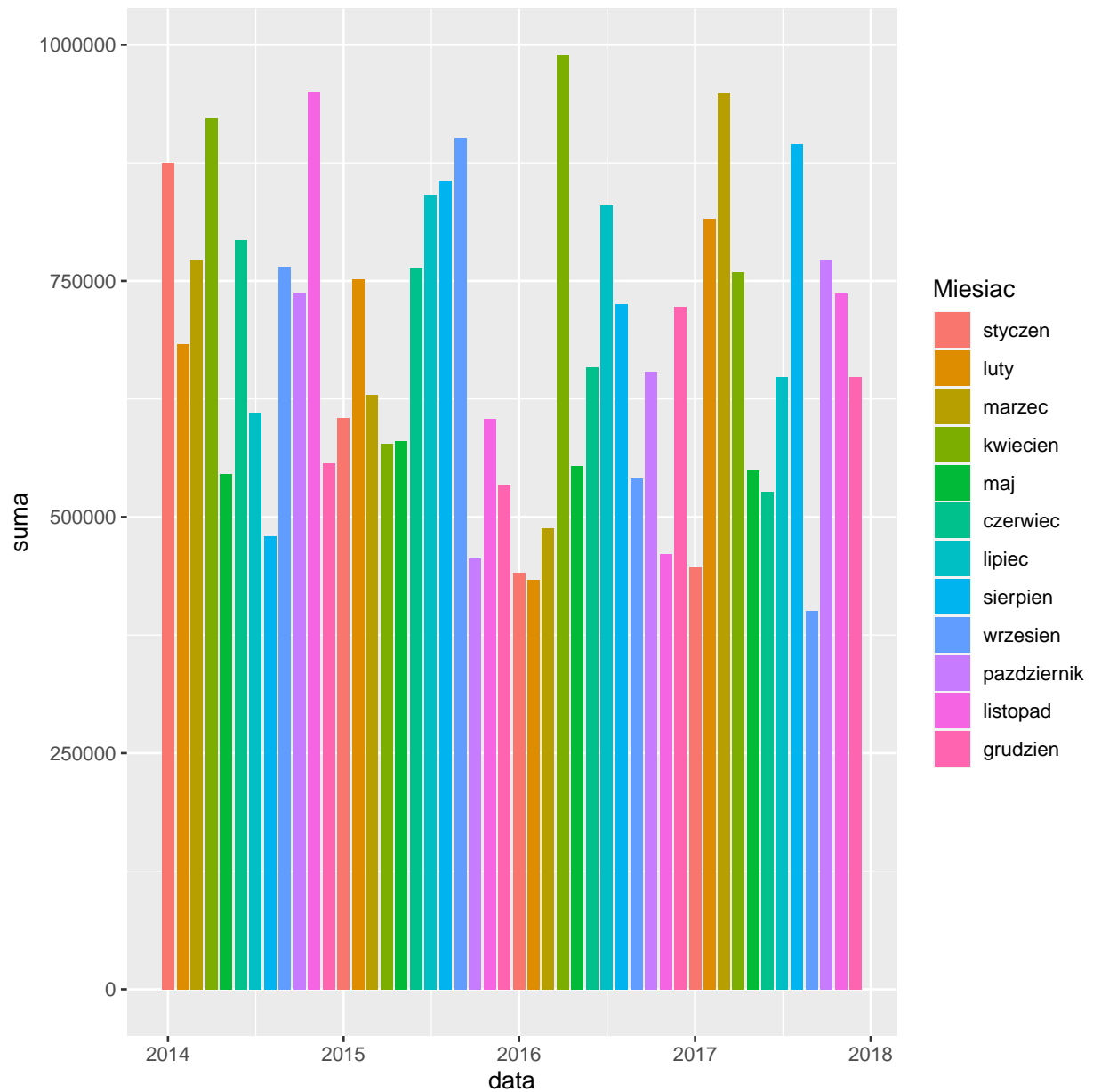
### 2.1 Analiza wydatków na zakup pojazdów

```
query <- "SELECT CONCAT(DATE_FORMAT(data_zaksięgowania, '%Y-%m'), '-01')
AS data, SUM(cena) as suma
FROM zakup_samochodu GROUP BY data ORDER BY data"

df <- dbGetQuery(con, query)
df[,1] <- as.Date(df[,1])
df[,2] <- as.numeric(df[,2])

wszystkie_daty <- expand.grid(1:12, 2014:2017) |> reframe(data = make_date(Var2, Var1, 1))
df <- merge(x=df, y=wszystkie_daty, by="data", all.y=T)
df$suma[is.na(df$suma)] <- 0

df |> ggplot(aes(x=data, y=suma, , fill=format(data, "%m"))) +
  geom_bar(stat = "identity") +
  scale_fill_manual(name="Miesiąc", labels=unique(format(df$data, "%B")), values=hue_pal()(12))
```



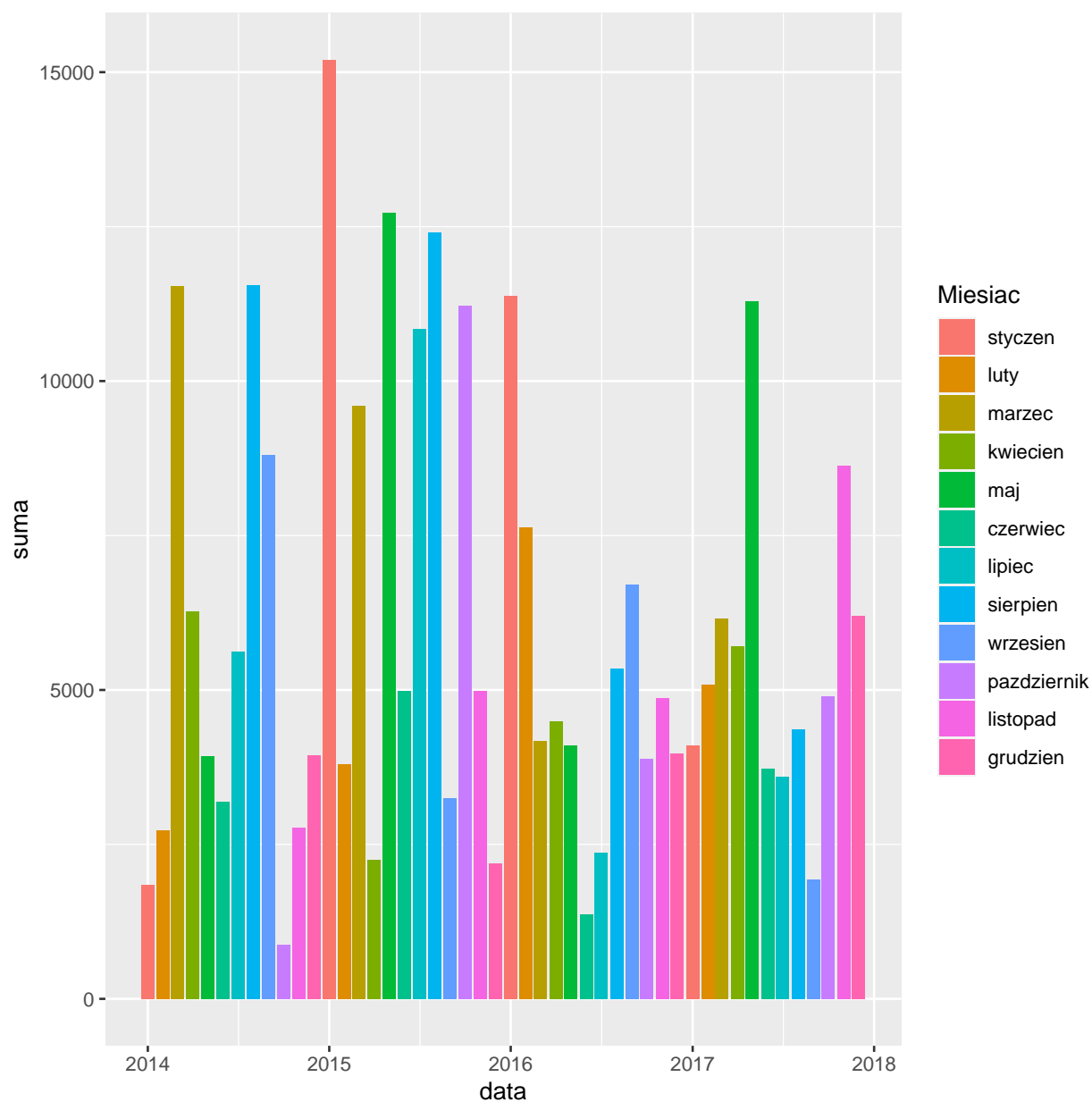
## 2.2 Analiza wydatków na zakup części

```
query <- "SELECT CONCAT(DATE_FORMAT(data_zaksięgowania, '%Y-%m'), '-01') as data,
SUM(użyte_części.cena*użyte_części.ilość) as suma
FROM usługi JOIN użyte_części
ON usługi.id = użyte_części.id_usługi
GROUP BY data ORDER BY data"

df <- dbGetQuery(con, query)
df[,1] <- as.Date(df[,1])
df[,2] <- as.numeric(df[,2])
```

```
wszystkie_daty <- expand.grid(1:12, 2014:2017) |> reframe(data = make_date(Var2, Var1, 1))
df <- merge(x=df, y=wszystkie_daty, by="data", all.y=T)
df$suma[is.na(df$suma)] <- 0

df |> ggplot(aes(x=data, y=suma, , fill=format(data, "%m")))+
  geom_bar(stat = "identity") +
  scale_fill_manual(name="Miesiac", labels=unique(format(df$data, "%B")), values=hue_pal()(12))
```



## 2.3 Analiza przychodów z usług warsztatu

```

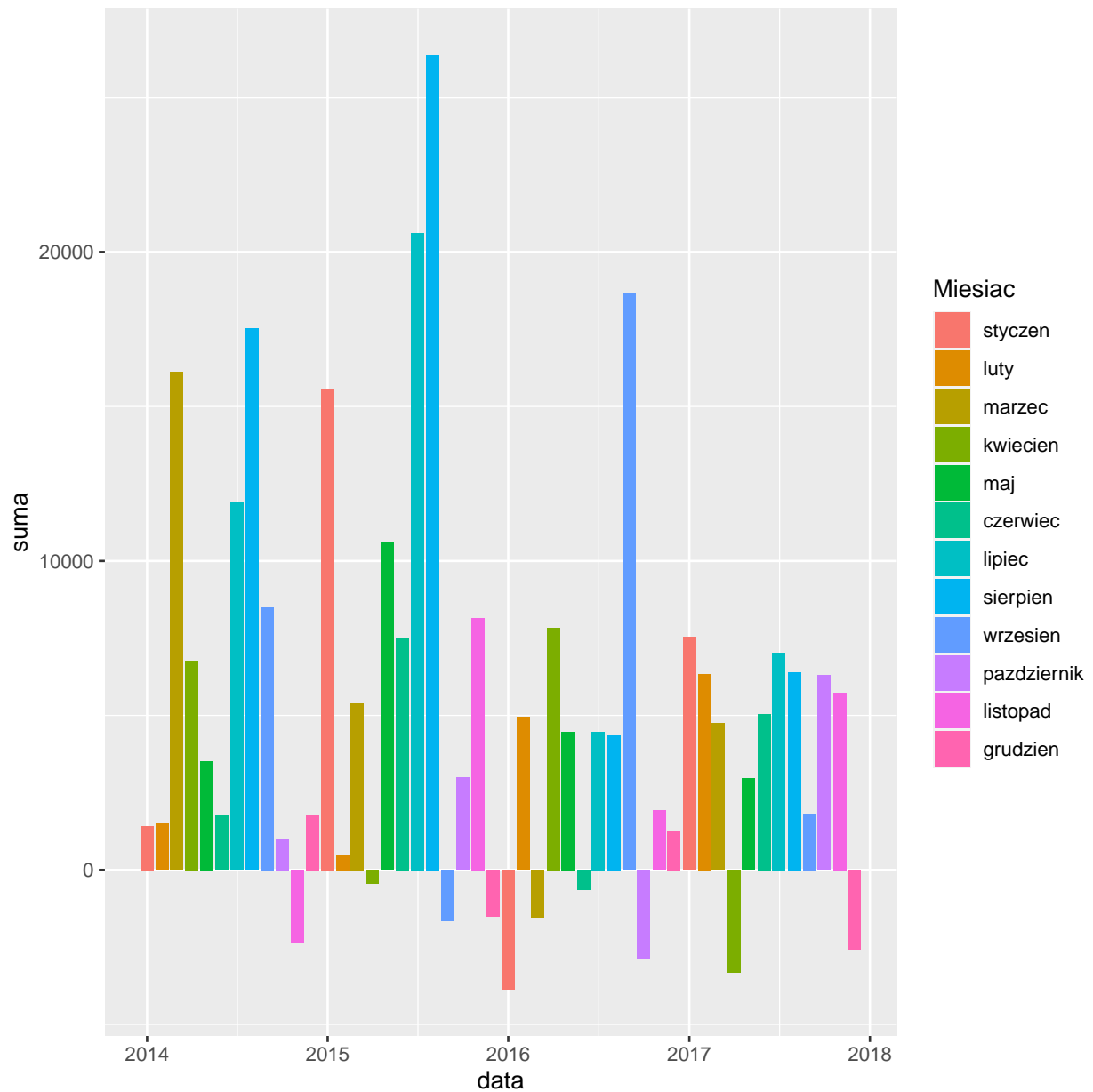
query <- "SELECT CONCAT(DATE_FORMAT(data_zaksięgowania, '%Y-%m'), '-01') as data,
SUM(usługi.cena) - SUM(użyte_części.cena*użyte_części.ilosc) as suma
FROM usługi JOIN użyte_części
ON usługi.id = użyte_części.id_usługi
GROUP BY data ORDER BY data"

df <- dbGetQuery(con, query)
df[,1] <- as.Date(df[,1])
df[,2] <- as.numeric(df[,2])

wszystkie_daty <- expand.grid(1:12, 2014:2017) |> reframe(data = make_date(Var2, Var1, 1))
df <- merge(x=df, y=wszystkie_daty, by="data", all.y=T)
df$suma[is.na(df$suma)] <- 0

df |> ggplot(aes(x=data, y=suma, , fill=format(data, "%m"))) +
  geom_bar(stat = "identity") +
  scale_fill_manual(name="Miesiąc", labels=unique(format(df$data, "%B")), values=hue_pal()(12))

```



## 2.4 Koszty wypłat dla pracowników

## 3 Kim są najlepszy mechanik i sprzedawca w warsztacie?

Przez czas działania warsztatu „Pimp My Wheels” osoba zarządzająca warsztatem nie była skłonna do dawania podwyżek, jednak postanowiła zlecić informatykowi by przeanalizował bazę danych i znalazł pracowników, którzy zasługują na większe wynagrodzenie.

```
query <- "SELECT CONCAT(imię, ' ',nazwisko) as sprzedawca, SUM(sprzedaż_samochodu.cena) as suma,
COUNT(sprzedaż_samochodu.cena) as ile_sprzedanych, płaça, id_pracownika
FROM sprzedaż_samochodu
JOIN pracownicy ON sprzedaż_samochodu.id_pracownika=pracownicy.id"
```



```

GROUP BY sprzedaż_samochodu.id_pracownika ORDER BY suma"

df <- dbGetQuery(con, query)

najwiecej_zarobił <- df %>% filter(suma == max(suma))
najwiecej_sprzedaży <- df %>% filter(ile_sprzedanych == max(ile_sprzedanych))

średnia_płaca <- mean(df$płaca)
średnia_suma <- mean(df$suma)
średnia_ilość <- mean(df$ile_sprzedanych)

procent_suma_zarobił <- (najwiecej_zarobił$suma / średnia_suma) * 100
procent_placa_zarobił <- (najwiecej_zarobił$płaca / średnia_płaca) * 100

różnica_zarobił <- round(procent_suma_zarobił - procent_placa_zarobił, 2)

procent_suma_sprzedaży <- (najwiecej_sprzedaży$ile_sprzedanych / średnia_ilość) * 100
procent_placa_sprzedaży <- (najwiecej_sprzedaży$płaca / średnia_płaca) * 100

różnica_sprzedaż <- round(procent_suma_sprzedaży - procent_placa_sprzedaży, 2)

if (najwiecej_sprzedaży[,5] == najwiecej_zarobił[,5]){ten_sam=c("ponownie ")} else {ten_sam=rep("", 1)}

if (length(najwiecej_zarobił[,1])>1){
  tekst_sprzedaż1=paste("Osoby, które sprzedały w firmie najwięcej pojazdów to ", najwiecej_zarobił[,1])
} else { tekst_sprzedaż1= paste("Osoba, która sprzedała w firmie najbardziej wartościowe pojazdy to "
". Ta osoba zarobiła dla firmy ", format(round(najwiecej_zarobił$suma,2), nsmall = 2), " zł na sprzedaż
"\\% średniej kwoty zarobionej dla warsztatu przez cały okres pracy. W porównaniu, zarabiana przez tego
format(round(średnia_płaca,2), nsmall = 2) , " zł). Różnica procentowa wynosi ",
format(abs(różnica_zarobił), nsmall = 2), "\\%",
if_else(różnica_zarobił < 0 , " na korzyść pracownika, czyli warsztat nie musi koniecznzie zwiększać st
różnica_zarobił < 5, ", czyli warsztat powinien rozważyć podwyżkę dla tego pracownika, by zachęcić go
}

if (length(najwiecej_sprzedaży[,1])>1){
  tekst_sprzedaż2=paste("Osoby, które sprzedały w firmie najwięcej pojazdów to ", najwiecej_sprzedaży[,1])
} else {
tekst_sprzedaż2 = paste("Pracownik, który sprzedał najwięcej pojazdów to ", ten_sam[1], najwiecej_sprzedaży[,1])
". Ta osoba sprzedała aż ", format(round(najwiecej_sprzedaży$ile_sprzedanych,2), nsmall = 2), " pojazdów
"\\% średnich sprzedaży przypadających na pracownika. Patrząc na zarobki pracownika można zauważyć, że z
format(round(średnia_płaca,2), nsmall = 2) , " zł). Między tymi dwoma wartościami wynosi ",
format(abs(różnica_sprzedaż), nsmall = 2), "\\%",
if_else(różnica_sprzedaż < 0 , " na korzyść pracownika, czyli zwiększenie stawki tego sprzedawcy nie j
różnica_sprzedaż < 5, ", wobec czego podwyżka dla tego sprzedawcy powinna zostać rozważona, ponieważ m
}

if (najwiecej_sprzedaży[,5] == najwiecej_zarobił[,5]){tekst_sprzedaż3=c("ponownie ")} else {tekst_sprzedaż3=c("nie ponownie ")}

```

Osoba, która sprzedała w firmie najbardziej wartościowe pojazdy to Janusz Kończal. Ta osoba zarobiła dla firmy 25277132.54 zł na sprzedaży pojazdów, co stanowi 112.19% średniej kwoty zarobionej dla warsztatu przez cały okres pracy. W porównaniu, zarabiana przez tego pracownika kwota za godzinę pracy (59.50 zł) stanowi 105.03 % średniej płacy sprzedawcy (56.65 zł). Różnica procentowa wynosi 7.16%, czyli warsztat zdecydowanie powinien rozważyć podwyżkę dla tego pracownika, by lepiej wyróżnić jego wpływ na

dochody firmy.

Pracownik, który sprzedał najwięcej pojazdów to ponownie Janusz Kończal. Ta osoba sprzedała aż 413 pojazdów, co stanowi 106.99% średnich sprzedaży przypadających na pracownika. Patrząc na zarobki pracownika można zauważyć, że zarabia on za godzinę pracy (59.50 zł), czyli 105.03 % średniej płacy sprzedawcy (56.65 zł). Między tymi dwoma wartościami wynosi 1.96%, czyli zwiększenie stawki tego sprzedawcy nie jest niezbędne, z powodu niewielkiej różnicy. Drobną podwyżka mogłaby za to pozytywnie wpłynąć na ilość sprzedaży i motywację pracownika,

```
query <- "SELECT CONCAT(imię, ' ',nazwisko) as mechanik,
SUM(usługi.cena) - SUM(użyte_części.cena*użyte_części.ilość) as suma,
COUNT(usługi.id) as ile_napraw, płaca
FROM usługi
JOIN pracownicy ON usługi.id_pracownika=pracownicy.id
JOIN użyte_części ON usługi.id = użyte_części.id_usługi
GROUP BY usługi.id_pracownika ORDER BY suma"

df <- dbGetQuery(con, query)

najwiecej_zarobil <- (df %>% filter(suma == max_zarobki))

## Error in 'filter()':
## i In argument: 'suma == max_zarobki'.
## Caused by error:
## ! nie znaleziono obiektu 'max_zarobki'

najwiecej_napraw <- (df %>% filter(ile_napraw == max(ile_napraw)))

srednia_placa <- mean(df$płaca)
srednia_suma <- mean(df$suma)

procent_suma_zarobil <- ( najwiecej_zarobil$suma/ srednia_suma) * 100 - 100

## Error in eval(expr, envir, enclos): nie znaleziono obiektu 'najwiecej_zarobil'

procent_placa_zarobil <- ( najwiecej_zarobil$płaca/ srednia_placa) * 100 - 100

## Error in eval(expr, envir, enclos): nie znaleziono obiektu 'najwiecej_zarobil'

procent_suma_napraw <- (najwiecej_napraw$suma / srednia_suma) * 100
procent_placa_napraw <- (najwiecej_napraw$płaca / srednia_placa) * 100

# Wyświetlenie wyników
cat("Pracownik, który zarobił najwięcej dla firmy:\n")

## Pracownik, który zarobił najwięcej dla firmy:

cat("Nazwa: ", najwiecej_zarobil$mechanik, "\n")

## Error in eval(expr, envir, enclos): nie znaleziono obiektu 'najwiecej_zarobil'

cat("Procent średniej sumy: ", round(procent_suma_zarobil, 2), "%\n")

## Error in eval(expr, envir, enclos): nie znaleziono obiektu 'procent_suma_zarobil'

cat("Procent średniej płacy: ", round(procent_placa_zarobil, 2), "%\n\n")

## Error in eval(expr, envir, enclos): nie znaleziono obiektu 'procent_placa_zarobil'
```

```

cat("Pracownik, który dokonał największej ilości napraw:\n")
## Pracownik, który dokonał największej ilości napraw:
cat("Nazwa: ", najwiecej_napraw$mechanik, "\n")
## Nazwa: Dariusz Kuraś
cat("Procent średniej sumy: ", round(procent_suma_napraw, 2), "%\n")
## Procent średniej sumy: 100.71 %
cat("Procent średniej płacy: ", round(procent_placa_napraw, 2), "%\n")
## Procent średniej płacy: 86.48 %
if (length(najwiecej_zarobil)>1){
  liczba_osób=c("Osoby, które sprzedały w firmie najwięcej pojazdów to ")
} else {
  liczba_osób = c("Osoba, która sprzedała w firmie najwięcej pojazdów to ")
}
## Error in eval(expr, envir, enclos): nie znaleziono obiektu 'najwiecej_zarobil'

dbDisconnect(con)

```